Institute of Hydrology

Report No. 123

Allowance for discretization in hydrological and environmental risk estimation

Natural Environment Research Council

Report No. 123

Allowance for discretization in hydrological and environmental risk estimation

I J Dwyer & D W Reed

January 1995

Institute of Hydrology
Crowmarsh Gifford
Wallingford
Oxfordshire OX10 8BB
United Kingdom

ISBN 0 948540 67 2

IH Report No. 126

published by the Institute of Hydrology

January 1995

Cover picture: Wind damage, October 1989, Hove Park, Sussex.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Executive summary

Disasters, natural or otherwise, are often precipitated by extremes of an environmental variable such as rainfall or wind speed. In order to make provision for alleviating the effects of such events, reliable risk estimation is required. This is usually obtained by measuring the relevant variable and analysing the data in an extreme value context. The data, however, inevitably contain some form of discretization (such as averaging over discrete time steps) which can degrade the risk assessment.

The report examines the effect of data discretization upon the estimation of period maxima. A correction model is proposed for converting *fixed* maxima, derived from discretized data, to *true* maxima as would be derived from continuous data. The model can be applied over a range of event durations.

The report takes a particular interest in rainfall extremes, for which discretization effects can be marked. The results suggest that previously reported correction factors are too low. The analysis of hourly rainfall data from various locations has enabled correction factors for daily rainfall to be discerned according to climate regime. Wind speed and air temperature extremes are also examined. A relationship between correction factors and *effective fractal dimension* is demonstrated, providing a means for deriving correction factors for other environmental variables.

The report concludes by consolidating the results into practical recommendations.

Contents

1 Introduction

Extremes of natural processes may present risk to the public. For example, heavy rainfall can cause flooding, strong winds may topple trees or buildings, and extreme air temperatures can result in the death of the weak and elderly.

Thousands, sometimes millions, of pounds are invested in preventing or alleviating such effects. To achieve this effectively and efficiently, a good quantitative assessment of the likelihood of occurrence is imperative: this is normally obtained by systematically measuring the environmental variable of interest and statistically analysing the resulting data in an extreme value context.

However, the measurement and recording process inevitably involves some form of discretization since continuous measurement, if practicable, is costly to perform and difficult to record. Instead, it is common to take instantaneous measurements at regularly spaced intervals or infer averages over discrete time steps.

In the case of assessing extreme rainfall, it is common to measure daily rainfall accumulations and extract the annual maxima. Daily accumulations are conventionally collected between the *fixed* hours of 0900 each day; if they were collected between the hours of, say, 2100 each day, a different value for the annual maximum would result in many years. With respect to the flooding of a river, the clock timing of the rainfall is irrelevant — it is the maximum accumulation that is important. In the absence of continuous data, however, the *true* 24-hour maximum is unavailable, and so the annual maximum based upon the fixed time intervals is used; this is, in general, lower and thus the extremes are under-estimated.

There are two distinct mechanisms which cause the true and fixed maxima to differ: firstly, where the fixed maximum results from the same event as the true maximum but fails to capture it fully (Figure 1.1a); secondly, where the fixed maximum results from a separate event which is better synchronized with the timing of the fixed intervals (Figure 1.1b). The same problem arises for any variable measured by averages or accumulations over discrete time intervals.

To correct for this under-estimation of the extremes, a relationship between the mean

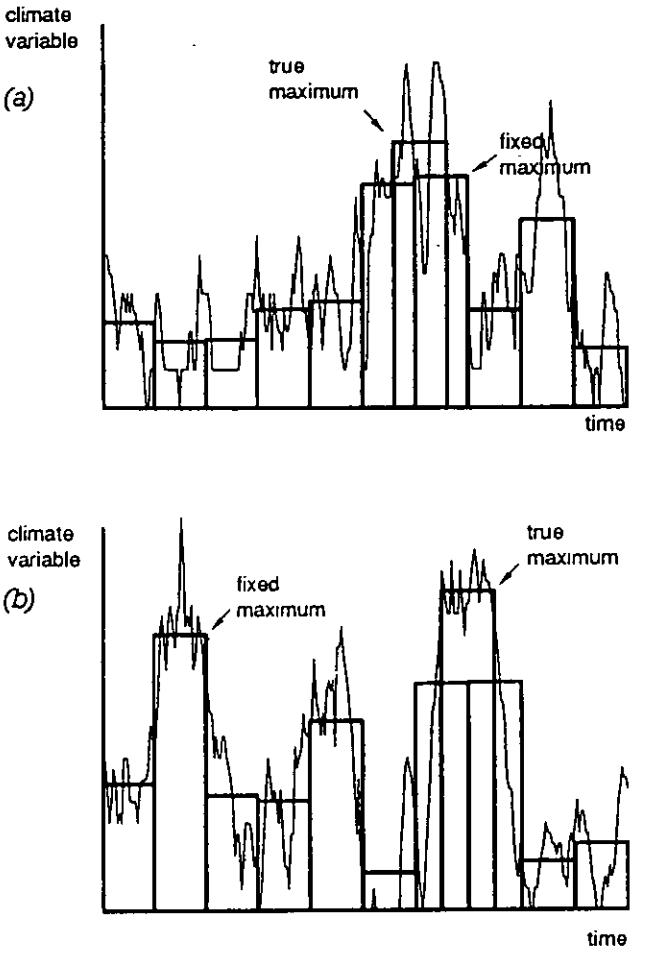


Figure 1.1 *True and fixed maxima as (a) part of the same event and (b) arising from different events*

fixed maximum and the mean true maximum is sought. Given hourly data, the fixed daily accumulations are easily constructed and the maxima derived. The true 24-hour maxima still escape calculation, however, since the data are still discretized: rather than sliding continuously, a 24-hour window 'shunts' along in increments of one hour. If 1-minute accumulations were available then the shunting window would be more nearly continuous and a better approximation of the true maximum would be obtained. The maxima derived from these shunting windows are referred to as *sliding* maxima (as opposed to the *fixed* and *true* maxima). The data resolution pertaining to a sliding maximum is implied in the notation: for example, 2-day sliding maxima are of duration 2 days and obtained from daily data, whereas 48-hour sliding maxima are also of duration 2

days but obtained from hourly data. The problem is generalized to maxima of various durations.

The ratios of the mean sliding to the mean fixed maxima are calculated and modelled to infer *multipliers* or *correction factors* which convert one type of maximum to another for a specified duration. As well as the mean of the maxima, discretization could affect the higher order moments which would also require correction. This is therefore investigated as part of the study.

A historical overview is presented in Chapter 2. Chapter 3 details the theoretical aspects of the problem, the methodology used in the study and the proposed correction model. A detailed examination of hourly rainfall data is presented in Chapter 4 to derive correction factors for rainfall maxima of durations one to 32 hours. A comparison between these and similar correction factors for wind speed and air temperature is given in Chapter 5 where a link to fractals is suggested. Chapter 6 compares the results from hourly and non-hourly rainfall data and Chapter 7 details some incidental investigations which proved to be of interest during the course of the study. Chapter 8 consolidates the results of the previous chapters into recommendations for use. The Appendices contain some technical expansions of the main text. In particular, Appendix C contains details of all the data records used in the study, including summary statistics, principal results, climate details and other particulars such as missing data if applicable.

2 Historical overview

One of the earliest and most commonly cited references to the discretization problem is that of Hershfield and Wilson (1958). This wide-ranging paper on extreme rainfall estimation is rather concise, devoting just a single paragraph to discretization which neglects to give clear details of the data and methods used. Nevertheless, they report a multiplier of 1.13 for converting "...observation-day rainfall for a particular frequency to the maximum 1440-minute rainfall for the same frequency". They find the same multiplier for converting clock-hour to 60-minute extremes, which they consider to be coincidental.

Another frequently quoted paper on the topic is that of Weiss (1964). He approached the problem theoretically using a simple probabilistic model to argue that the expected ratio of true to fixed maxima is 8/7 (1.143). Whilst this is reassuringly close to Hershfield's result, the analysis is mathematically flawed: improper statistical formulation has led to incorrect heuristic arguments. Correction of these errors results in an expected multiplier of 4/3 (1.33), which is disconcertingly higher than Hershfield's result. Close examination of the model assumptions, however, reveals that this result is inevitably an over-estimate. A full discussion of Weiss's paper can be found in Appendix A.

The study of Kerr *et al.* (1970) used extensive data to examine rainfall-frequency-duration relationships in the State of Pennsylvania, USA. Some 45 stations were used to investigate the discretization problem, each with 17 or more years of hourly and daily data (collected separately). For each station, the sliding 24-hour and the fixed 1-day annual maxima series were obtained and Gumbel distributions fitted to each; the respective mean annual rainfalls (2.33-year return period) were calculated thereof and their ratio obtained. The average of the 45 ratios pertaining to each site came to 1.12, close to Hershfield's figure. Sliding 60-minute annual maxima were available at six sites, enabling a similar calculation for the average ratio of 60-minute to clock-hour maxima: this resulted in 1.16 which does not agree so well with Hershfield's findings. Whilst the extent of the data used to obtain these results is made clear, the quality is not.

A similar analysis was done by Harishara and Tripathi (1973) for 67 sites across India. Ratios for the T-year 24-hour to 1-day rainfalls were obtained at each site for T=2, 5, 10, 25 and 50. Averaging across T resulted in a mean value at each site; averaging across sites then resulted in an overall mean correction factor of 1.15. However, only five stations had 25 or more years of record, which means that the ratios for T=25 and T=50 are somewhat unreliable. Indeed, with some stations having less than ten years of record, some kind of pooling of the data would have been beneficial. Nevertheless, comprehensive reporting of the data and results at each station enabled the 2-year return period ratios from stations having 15 or more years of record to be isolated. Averaging the resulting 23 ratios (weighted according to the number of years of record at each site), gives a mean correction factor of 1.144.

The Flood Studies Report (Natural Environment Research Council, 1975, Vol. II, Chapter 3), concentrates on the 5-year return period (M5) rainfall. Using data from around the UK, M5 rainfall was estimated for observational-day, 24-hour, clock-hour and 60-minute durations. A multiplier of 1.11 was found for converting 1-day to 24-hour M5 rainfall, though the quantity and quality of the data used to obtain this result are not clearly specified. A higher multiplier of 1.15 is reported for converting clock-hour to 60-minute M5 rainfall; this figure is based upon 50 stations with both hourly annual maxima data for estimating the M5 clock-hour values, and annual frequency data (frequencies of 60-minute exceedances of 5, 10, 15, and 25 mm) for estimating the M5 60-minute values.

More recently, van Montfort (1991) examined the problem of estimating extreme value distribution parameters for sliding maxima, given that only fixed maxima are available. A common method of fitting distributions to data is the maximum likelihood procedure which renders those parameter values which maximize the chance of the observed data being reproduced by simulation. Denoting, for year i, the sliding 24-hour maximum by A_{1i} and the fixed 1-day and 2-day maxima by F_{1i} and F_{2i} respectively, van Montfort utilizes the inequality $F_{1i} \leq A_{1i} \leq F_{2i}$ to estimate GEV parameters for the sliding maxima using only the fixed maxima. This is done by using the maximum likelihood procedure to maximize the chance of the

differences F_{2i}-F_{1i} being reproduced in simulation. The method is demonstrated on 58 years of data from Kelburn, New Zealand. It is claimed that the resulting parameter values are as reliable as those obtained directly from the observed sliding maxima also available at the site. In terms of correcting maxima, the mean annual rainfalls (2.33 return period) extracted from the EV1 distributions fitted to the observed fixed and sliding maxima, result in a correction factor of 1.15. Van Montfort quotes 1.14 for correcting the 50-year rainfall, although estimating this from a single record of 58 years is somewhat sample-dependent.

The Allowance for Discretization in Hydrological and Environmental Risk Estimation (ADHERE) project, funded by the Natural Environment Research Council, was set up as a comprehensive investigation into the discretization effect. The work of Coyle *et al.* (1991) is a precursor to that reported here. Correction factors are calculated directly from maxima using good quality, high resolution data, the details of which are fully reported. A range of environmental variables, event durations, data resolutions and climate regimes are examined. A generalized correction model is proposed.

3 Method and theory

3.1 Terminology and conventions

The discussion so far has concentrated upon 24-hour versus 1-day annual maxima for rainfall. The study investigates discretization in a more general context, however, as described below.

Data resolution and event duration
Speaking of a 24-hour rainfall means that a rainfall time series of hourly resolution has been used to calculate an accumulation of duration 24 times that resolution. In general, durations of D times the data resolution are considered. Also, the data resolution, denoted by τ, may be something other than 1-hour. Thus, τ=1 day and D=4 refer to 4-day accumulations calculated from daily data. Note that τ has dimensions of time but D is dimensionless.

Length of record
The length of a record is expressed in terms of the number of data values, n. To ensure equitable comparison of results, n is fixed at 16384 (2^{14}) unless stated otherwise. This pertains to nearly two years of hourly data or 45 years of daily data. The figure has been chosen to reflect the lengths of record generally available and constrained to a power of two for methodological reasons.

Period length
Whilst 45 years of data enable a reasonable calculation of mean annual maximum values, two years of (hourly) data do not. Hence, each record is divided not into years but into m=32 (2^5) periods of equal length (about 512 values). A maximum is extracted for each period, the average of which is termed the *mean period maximum*. Maxima are therefore extracted from periods of approximately 21 days for hourly data and 1.4 years for daily data.

Climate variables
Climate variables other than rainfall are also examined, such as wind speed and air temperature. For variables where the term 'accumulation' over some duration is inappropriate, it is taken to mean the 'average value' over the duration.

Thus, the extremes of a climate variable are examined by reference to a time series consisting of n data values measured at a resolution τ. The record is divided into m periods of equal length and the maximum accumulation of duration D (fixed or sliding) is extracted for each. The values for n and m are fixed at 16384 and 32 respectively; sensitivity of the results to these choices is investigated.

3.2 Calculating the fixed and sliding period maxima

Fixed maxima
Let the first period consist of data values x(1), x(2), x(3), ... etc. The first fixed accumulation of duration D is obtained by summing (or averaging) the data values x(1), x(2), ... , x(D). The second fixed accumulation is obtained from the values x(D+1), x(D+2), ... , x(2D), and so on. This is illustrated for D=4 in Fig. 3.1a. The maximum of all the resulting fixed accumulations is the fixed maximum for that period. Fixed maxima for other periods are calculated similarly.

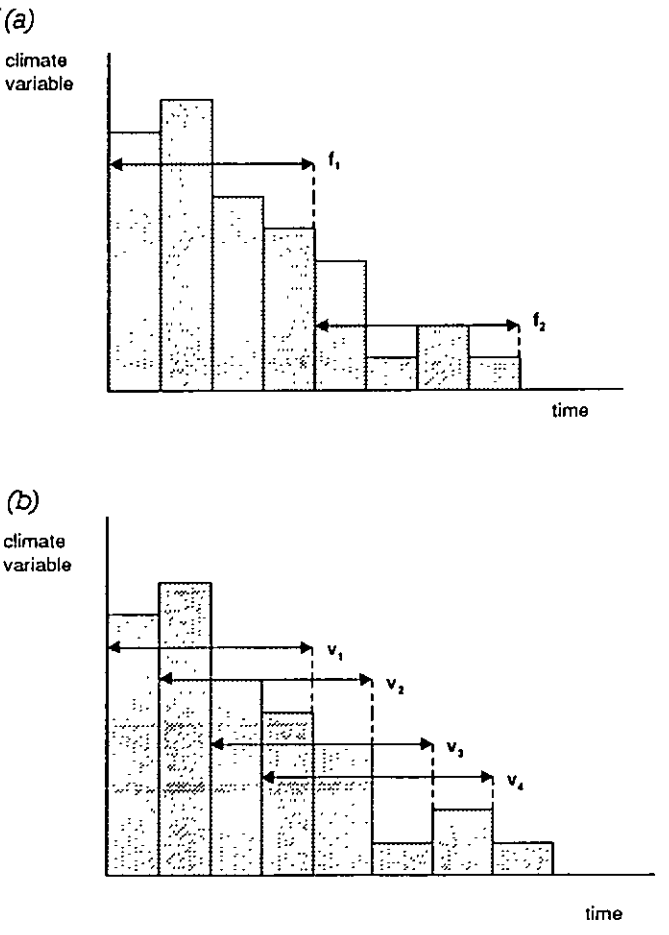


Figure 3.1 *The construction of (a) the fixed accumulations f_1, f_2 ... and (b) the sliding accumulations v_1, v_2 ..., demonstrated for duration D=4*

To obtain a whole number of fixed accumulations, the period length is required to be an integer multiple of D. Dividing the whole record into m equal periods results in a period length of $n/m=2^9$ (512), which is exactly divisible by D only when D is a power of two. Thus, so that any duration D may be considered, the period length, p, is set to the greatest multiple of D satisfying $p \leq 512$. When $p<512$, the construction of m=32 consecutive periods of equal length results in some unused data at the end of the record.

Sliding maxima

Using the notation above for the first period, the first sliding accumulation of duration D is likewise obtained from the data values x(1), x(2), ... , x(D). The second sliding accumulation, however, is obtained from x(2), x(3), ... , x(D+1) and thus overlaps with the first. The third is obtained from x(3), x(4), ... , x(D+2) and so on as illustrated in Fig. 3.1b. The maximum of the resulting accumulations is the sliding maximum for the period. The periods used for extracting sliding maxima of duration D, are identical to those used for extracting fixed maxima of duration D.

When a large storm is split across two periods, its full significance may go unnoticed if the two periods are analysed separately. Therefore, rather than having the last sliding accumulation coincide with the last fixed accumulation, the sliding 'window' is allowed to straddle across the two periods. A convention is adopted whereby an accumulation which straddles two periods is neither lost nor 'counted twice' in the sense of it contributing to the sliding maximum for both periods (see Appendix B).

3.3 Calculating sample correction factors

The above describes how, for each duration D, fixed and sliding maxima are extracted from each of the m periods of a given data record. Denoting these by $F_i(D)$ and $V_i(D)$ (i=1,...,m) respectively, the fixed and sliding *mean* period maxima (of duration D) are then

$$\bar{F}(D) = \frac{1}{m}\sum_{i=1}^{m} F_i(D) \qquad 3.1$$

and

$$\bar{V}(D) = \frac{1}{m}\sum_{i=1}^{m} V_i(D) \qquad 3.2$$

The *sample* correction factor for duration D is thus defined by

$$R(D) = \bar{V}(D)/\bar{F}(D) \qquad 3.3$$

It is the multiplier required for converting the mean fixed into the mean sliding period maximum of duration D. For example, if τ=1 hour then R(24) is the multiplier required for converting the 1-day to the 24-hour mean period maximum.

In application it may be desirable to convert individual maxima rather than the mean maximum, in which case it is tempting to calculate the sample correction factor not as above but as the mean (over i) of the individual ratios $R_i(D)=V_i(D)/F_i(D)$. It turns out, however, that this is not a very satisfactory estimator as it can be seriously biased and have large mean square error, and so R(D) is to be preferred (see Barnett, 1974). Furthermore, note that

$$R(D) = \frac{\Sigma V_i(D)}{\Sigma F_i(D)} = \frac{\Sigma [V_i(D).F_i(D)/F_i(D)]}{\Sigma F_i(D)} = \frac{\Sigma [R_i(D).F_i(D)]}{\Sigma F_i(D)} \qquad 3.4$$

where the summations are all from i=1 to m. That is, R(D) is equivalent to a weighted sum of the individual ratios for each period, with greater weight given to the larger events. This is especially desirable in the analysis of hourly data since the period length is only about 21 days and so many of the maxima will not be particularly extreme.

Dropping the notation D, an estimate s(R) of the standard error of R(D) is suggested by Barnett as

$$s^2(R) = \frac{1}{m(m-1)\bar{F}^2}\sum_{i=1}^{m}(V_i - R.F_i)^2 \qquad 3.5$$

from which confidence intervals can be calculated using Normal percentage points.

Datum

It should be noted that there is some arbitrariness in the value of R in that it is relative to the *datum* (zero level) of the units of measurement. For example, the ratio of the air temperatures 10°C and 5°C is 10/5 = 2, but

expressed in degrees Kelvin the ratio becomes 283/278 = 1.02. Once a datum has been chosen, however, the ratio is independent of the scale of the units. Thus, using a datum of 0°C for air temperature, it makes no difference whether degrees Celsius or tenths of Celsius are used.

Rainfall and wind speed offer natural datum levels corresponding to 'no rain' and 'no wind'. The equivalent for temperature is that of no energy, namely 0°K (-273°C). This datum, being far below the level at which air temperatures are observed, renders very low correction factors. Later it is desirable to compare correction factors for different variables, whereupon a common datum type is required, which reflects the level at which the variable is observed. The minimum of the data record would suffice but is somewhat sample-dependent. Thus, the lower 1% quantile is used, which has the desired properties and in practice still renders datum levels of zero for rainfall and for wind speed (for daily and sub-daily data at least).

3.4 Modelling R(D)

The model

Associated with each data record, therefore, are the sample ratios R(D) for each D. A graph of R(D) against D can be constructed as in Fig. 3.2 for wind speed data at Eskdalemuir, Scotland.

The behaviour exhibited by this example is typical in that R(D) is increasing over small D (though not monotonically) and then levels off to

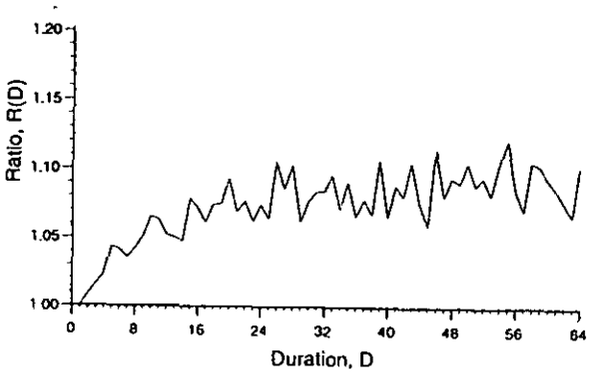


Figure 3.2 *R(D) against D for Eskdalemuir wind speed record F (τ=1 hour)*

fluctuate about a constant value. This can be explained by considering both the fixed and sliding maxima (of duration D) as approximations to the true maximum; the measured sliding maximum is only an approximation since it does not slide continuously but rather in steps of 1/D. As D increases, the sliding maximum becomes more nearly 'continuous' relative to the fixed maximum, and so an increasingly better approximation. Thus, as D increases, the ratio R(D) approaches a limiting value R^* which represents the correction of fixed maxima to true maxima.

This behaviour is modelled as exponentially diminishing growth to a limiting constant, with the constraint R(1)=1. Denoting the model function by $\rho(D)$, this gives

$$\rho(D) = 1 + a[1 - \exp\{-b(D-1)\}] \qquad 3.6$$

which can be fitted to the data R(D) by nonlinear regression, resulting in estimates of the parameters *a* and *b*. For the graph of Fig. 3.2, the regression renders estimates *a*=0.089 and *b*=0.096, the curve for which is depicted in Figure 3.3. Parameter *a* relates to the limiting value of $\rho(D)$:

$$\rho^* = \lim_{D\to\infty} \rho(D) = 1 + a \qquad 3.7$$

whence R^* is modelled by ρ^*. Parameter *b* relates to the growth rate of $\rho(D)$: its numerical value is not very enlightening, however, and so D_p is defined as the duration at which the curve attains p% of its final growth above 1, that is

$$\rho(D_p) = 1 + a\frac{p}{100} \qquad 3.8$$

giving

$$D_p = 1 - \frac{1}{b}\ln(1 - p/100) \qquad 3.9$$

In particular, D_{95} is used as the duration by which most (95%) of the growth has occurred. Thus, for the wind speed example above, ρ^*=1.089 and D_{95}=32. The maximum duration

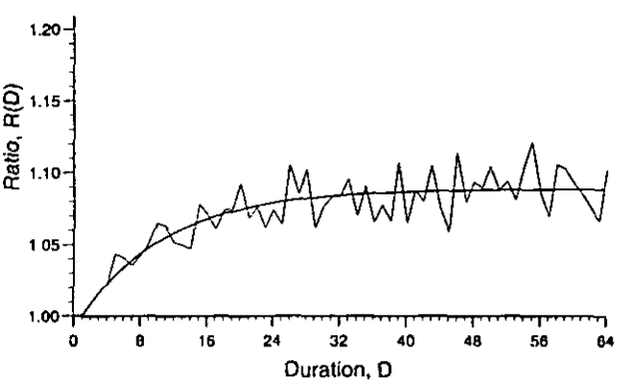


Figure 3.3 *Correction model fitted to Eskdalemuir wind speed record F*

used in modelling R(D) by $\rho(D)$ is denoted by D_{max}: it must be large enough for the limiting value to be properly estimated but not so large that is becomes a significant fraction of the period length. D_{max} equal to 32 or 64 was found to be appropriate for most records.

Regression details
Non-linear least-squares regression is used which seeks to minimize the sum (over D) of the squared differences, $[\rho(D) - R(D)]^2$. Standard subroutines from the NAG computer library (NAG, 1991) are used which also estimate the variances of the estimated parameters. Symmetric 95% confidence intervals for the estimated parameters are then calculated using the Student's t 2.5 percentage point. For example, the confidence interval for parameter *a* is given by

$$[\hat{a} - s_a t^d_{.025} \ , \ \hat{a} + s_a t^d_{.025}] \qquad 3.10$$

where $\hat{a}$ is the least-squares estimate for *a*, s_a is the square root of the estimated variance for *a* and $t^d_{.025}$ indicates the Student's t 2.5 percentage point at $d=D_{max}-2$ degrees of freedom. There is a 0.95 probability that the true value of parameter *a* lies in this interval. The confidence interval for *b* is constructed similarly. The end-points of these intervals are substituted into equations 3.7 and 3.9 to obtain the 95% confidence intervals for ρ^* and D_{95} respectively.

It is assumed that the variation about the regression curve is due to sample error. This is corroborated by an examination of a number of the data records which indicates that, for most D, the regression curve $\rho(D)$ falls within the sample 95% confidence interval for R(D) (as calculated using equation 3.5). This also makes reasonable the assumption that R(D) converges for large D.

Higher-order moments
It is possible that discretization affects not just the mean of the period maxima but also the variance and higher-order moments. The use of L-moment ratios (Hosking, 1990) is becoming popular as a measure of higher-order variation; they are more robust to outliers and sample variability than are conventional moment ratios, and are bounded in [-1,1] (simplifying comparisons).

For each data record, the L-CV and L-skewness were calculated for the set of fixed maxima $\{F_i(D): i=1, \ldots, m\}$ and sliding maxima $\{V_i(D): i=1, \ldots, m\}$ at each duration D. Graphs of the results help to discern any systematic higher-order differences between the fixed and sliding maxima. The graphs obtained show no such differences, however, as illustrated by some typical examples contained in Appendix D. Thus, it is reasonable to analyse the discretization effect with respect to the mean of annual maxima alone, without having to make corrections to the higher-order moments.

3.5 Applying the correction model

The above model can be used in four distinct ways as detailed below. An example of each is presented at the end of the section using the wind speed data pertaining to Figures 3.2 and 3.3.

Converting fixed to sliding maxima
The multiplier $\rho(D)$ may be used to convert the mean fixed maximum of duration D to the mean sliding maximum of duration D as measured at resolution τ.

Converting fixed to true maxima
The multiplier ρ^* may be used to convert the mean fixed maximum of duration D, to the mean true maximum of duration D. In doing so, however, it is assumed that fixing τ and letting D tend to infinity produces the same correction ratio as fixing D and letting τ tend to zero. That is, the correction factor for converting fixed to true maxima is assumed to be independent of the event duration D and data resolution τ. The latter is reasonable since both fixed and true maxima are independent of data resolution by definition. Both assumptions are examined for a range of durations and resolutions as part of this investigation.

Converting sliding to true maxima
It may occur that the mean sliding maximum of duration D, as measured at resolution τ, is available but, nevertheless, the mean true maximum is required. The former can be converted to the latter using the correction factor

$$\alpha(D) = \frac{\rho^*}{\rho(D)} \qquad 3.11$$

This can be seen by considering the conversion in two stages: first, convert the mean sliding maximum into the mean fixed maximum by dividing by $\rho(D)$; second, convert this mean fixed maximum into the mean true maximum using the multiplier ρ^*.

Converting sliding to sliding maxima
The mean sliding maximum of duration D, as measured at resolution τ, may be converted into the mean sliding maximum of the same duration but measured at a finer resolution $\tau'=\tau/\lambda$ ($\lambda>1$) using the multiplier

$$\beta(D) = \frac{\rho(\lambda D)}{\rho(D)} \qquad 3.12$$

This can be seen by considering the conversion in three stages: first convert the mean sliding maximum of duration D, measured at resolution τ, into the mean true maximum of the same duration by multiplying by $\alpha(D)$ (equation 3.11); second, use the assumption that ρ^* is independent of data resolution to obtain the mean fixed maximum for the same duration but expressed in terms of the resolution τ' (namely, λD) by dividing by ρ^*; lastly, convert this mean fixed maximum of duration λD into the mean sliding maximum, as measured at resolution τ', using the multiplier $\rho(\lambda D)$.

Examples
Examples of the above applications are presented below, based upon the hourly wind speed data at Eskdalemuir for which model parameters of a=0.089 and b=0.096 were estimated (as detailed earlier).

Let the mean fixed 1-day maximum be denoted by F. The mean sliding 24-hour maximum, V, is obtained from equation 3.6 as

$$\begin{aligned} V &= \rho(24)F \\ &\approx 1.079F \end{aligned} \qquad 3.13$$

The mean true maximum of duration 24-hours, T, can be obtained from equation 3.7, namely

$$\begin{aligned} T &= \rho^* F \\ &= (1+a)F \\ &= 1.089F \end{aligned} \qquad 3.14$$

If the sliding 24-hour maximum, V, is available, then T can be obtained from V using equation 3.11 as follows:

$$\begin{aligned} T &= \alpha(24)V \\ &= \frac{\rho^*}{\rho(24)}V \\ &= \frac{1.089}{1.079}V \\ &\approx 1.009V \end{aligned} \qquad 3.15$$

Suppose that the mean sliding 3-hour maximum, V1, has been calculated from the hourly data. Although it is sliding, there is still a considerable discretization effect at such a short duration (D=3). To calculate the mean sliding 3-hour maximum as if it were extracted from, say, 15-minute data, denoted by V2, use equation 3.12 to obtain

$$\begin{aligned} V2 &= \beta(3)V1 \\ &= \frac{\rho(3\lambda)}{\rho(3)}V1 \\ &= \frac{\rho(12)}{\rho(3)}V1 \\ &= \frac{1.058}{1.016}V1 \\ &\approx 1.041V1 \end{aligned} \qquad 3.16$$

4 Rainfall durations one to 32 hours

4.1 Introduction

This chapter uses rainfall measured at an hourly resolution to examine the correction of rainfall maxima of durations one to 32 hours, and in particular the correction of daily rainfall. This reflects, historically at least, where most of the applied hydrological interest lies. The aim is to compare and contrast the modelled forms, $\rho(D)$, for different sites with varying climate regimes. Six sites have been examined: Table 4.1 shows the number of standard data records (that is, of length n=16384) available at each. For example, eight records were extracted for Eskdalemuir and labelled A to H accordingly. For the UK sites the extracted records run consecutively, whereas data quality requirements prevented this for the Australian sites. Summary statistics and other details for each record can be found in Appendix C under the appropriate headings.

Table 4.1 *The number of hourly rainfall records available at each site*

Site	Number of records
Eskdalemuir	8 (A-H)
Leeming	5 (A-E)
Ringway	8 (A-H)
Brisbane	5 (A-E)
Melbourne	5 (A-E)
Sydney	6 (A-F)

4.2 Rainfall regime

All the UK sites represent cool temperate oceanic climates. Eskdalemuir, in the southern uplands of Scotland and Ringway (Manchester Airport), being west of the Pennines, are subject mainly to frontal rainfall carried in from the Atlantic Ocean. Although it rains all year round, October to January is the rainiest season. The high altitude of Eskdalemuir ensures that a significant amount of precipitation falls as snow in the winter months. Leeming, in North Yorkshire, lies east of the Pennines and is therefore more sheltered from the westerly fronts; it is subject to a greater number of convective storms in the summer, however, which swings the seasonality of heavy rainfall accordingly. Two of the three Australian sites, Melbourne and Sydney, represent warm temperate climates, whereas Brisbane veers towards the tropical. Melbourne, on the southern tip of mainland Australia, experiences moderate rainfall, evenly distributed throughout the year. Sydney, situated on the south-east coast, experiences mainly frontal rainfall, highest in autumn (May to June) and lower during spring (September to November). Brisbane, halfway down the east coast, is subject to heavy monsoon rains in its tropical summer and lighter frontal systems in the winter.

Recall that a storm split (at 0900) between two observation days is under-recorded by the fixed interval measurements. If a typical storm at a site arises from relatively few hours of intense rainfall then, by simple probability, there is less chance of this happening than for a site experiencing longer duration events, whereupon correction factors are likely to be lower. In studying correction factors for daily rainfall, it is therefore helpful to compare rainfall profiles for each site. Figure 4.1 shows average profiles derived for each site: that is, for any one site, the rainfall values forming each fixed 24-hour maximum are obtained and averaged by calculating the mean first hour value, the mean second hour value, etc. The result is converted into a percentage profile and plotted as residuals about the mean percentage, 4.17. This method of averaging preserves time-of-day features (the first hour corresponding to 0900 - 1000 hours) and weights each maximum according to its total volume. It is apparent that Brisbane has a more concentrated average profile than the other sites.

In order to compare profiles with correction factors quantitatively, it is convenient to construct a simple numerical index of profile concentration: for each site, the mean number of wet hours (N_{wet}) for the 1-day maxima is used. Table 4.2 shows the results for each site, with a 'wet' hour being one with a depth greater than 0.2mm.

4.3 Correction factors

Figure 4.2 shows examples (one record from each site) of graphs of R(D) against D complete

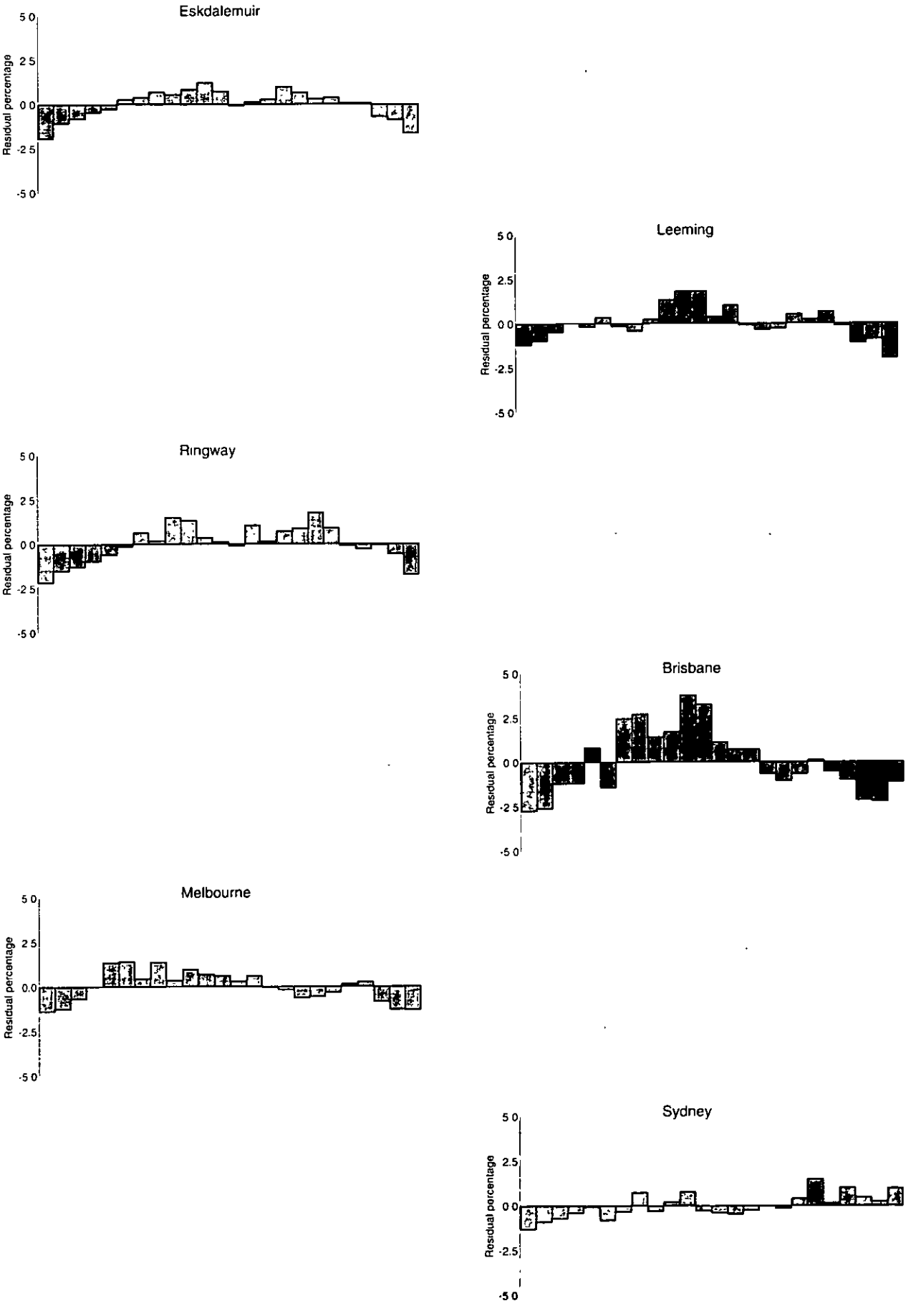


Figure 4.1 *Average storm profile for each site*

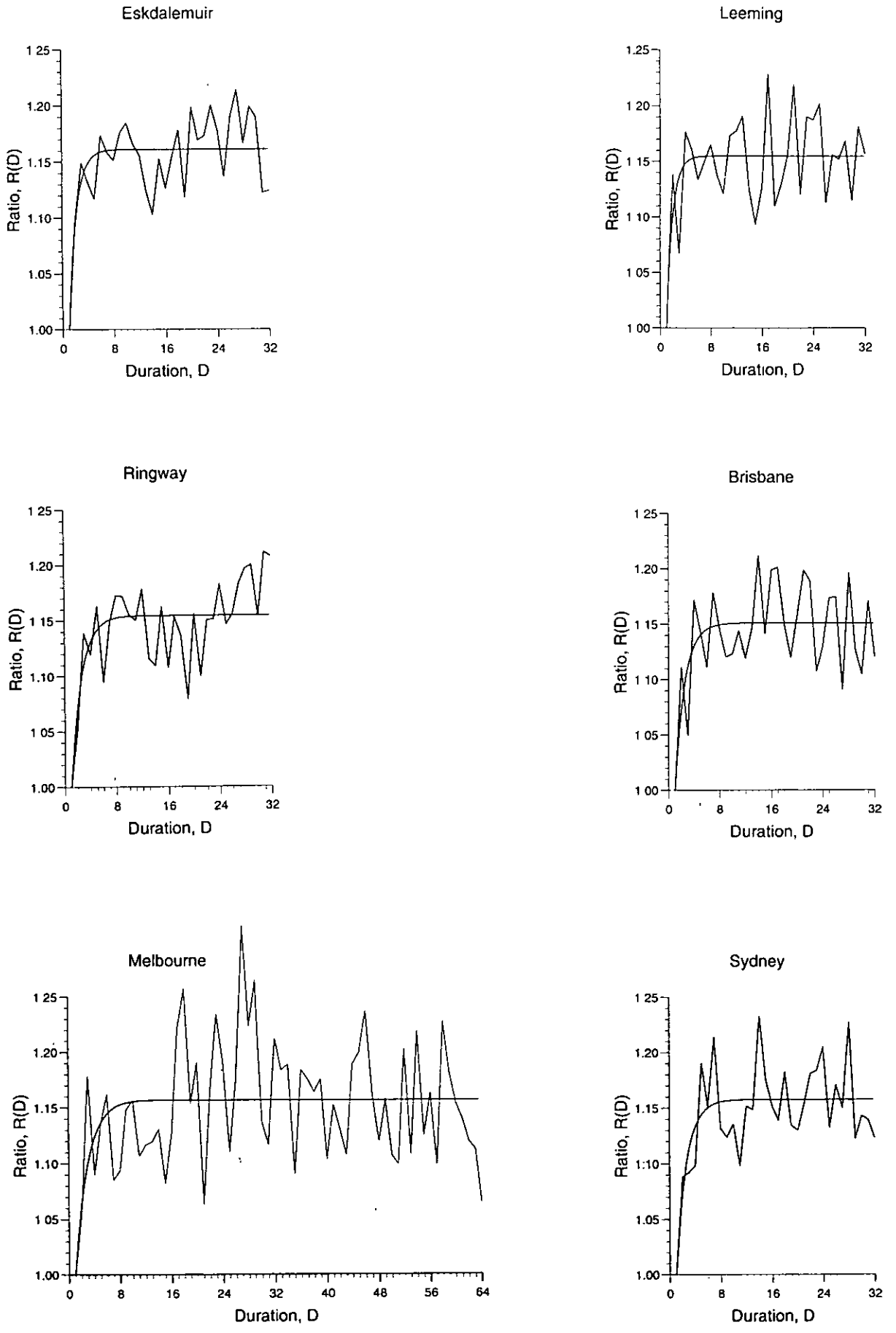


Figure 4.2 *Correction model fitted to hourly rainfall: Eskdalemuir record C, Leeming D, Ringway C, Brisbane B, Melbourne D and Sydney F*

Table 4.2 *The value of N_{wet} for each site*

Site	N_{wet} (hours)
Eskdalemuir	11.7
Sydney	8.8
Ringway	8.4
Melbourne	8.2
Leeming	8.1
Brisbane	7.2

with the fitted model. Together they illustrate the main features for rainfall (as compared to other environmental variables): convergence of $\rho(D)$ to ρ^* is rapid (rendering low D_{95} values); ρ^* is typically in the range 1.14 to 1.18; there is large sample variability in R(D) about the fitted model, resulting in large confidence intervals for the parameters *a* and *b*. The example for Melbourne shows the case, unusual for rainfall, where a D_{max} value of 64 was deemed necessary to estimate the limiting value ρ^* satisfactorily. For completeness, Figure 4.3 shows some of the less well-fitted plots: in (a), the limit ρ^* is reached before the first observation at D=2 and therefore parameter *b* cannot be properly estimated in the regression and so is effectively undefined; in (b), a great deal of sample variability leads to particularly large uncertainty in the regression curve.

No obvious inter-site differences are discernible from visual examination of the various plots and so close inspection of the parameter values, *a* and *b*, is required. Table 4.3 details the parameter values for each record at each site;

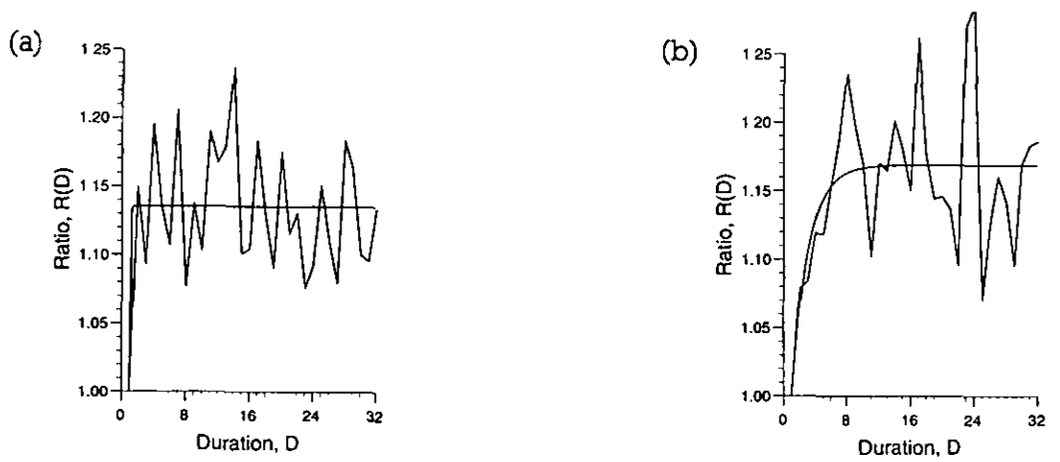


Figure 4.3 *Correction model fitted to hourly rainfall: (a) Melbourne record D. (b) Sydney F*

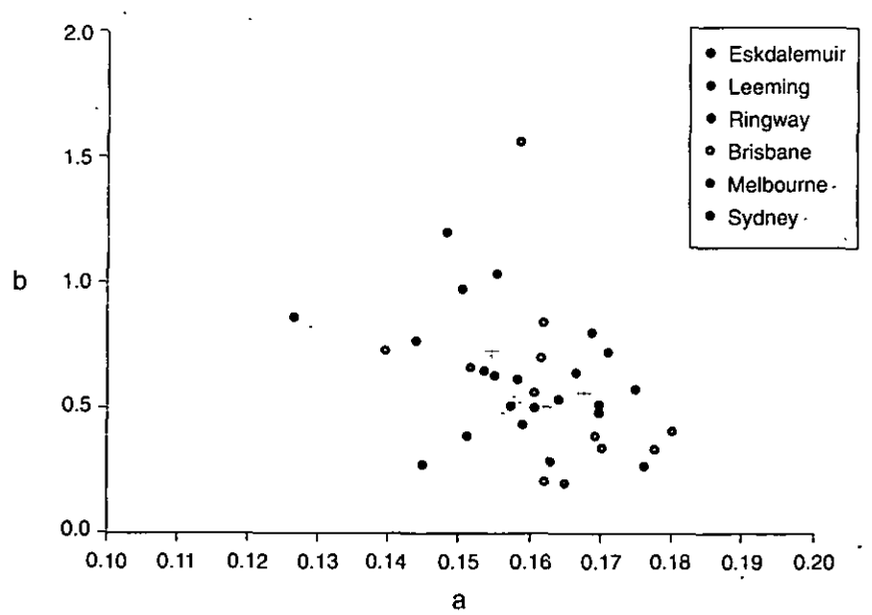


Figure 4.4 *Parameter space (a,b) for the correction model fitted to each hourly rainfall record*

Table 4.3 *Values of a, b and a′ for each hourly rainfall record at each site, plus respective site-means ("am" denotes arithmetic mean and "gm" stands for geometric mean)*

Site	Record	a		b		a′
Eskdalemuir	A	0.169		0.388		0.170
	B	0.160		0.563		0 159
	C	0.162		0.841		0.163
	D	0.180		0.402		0.180
	E	0.161		0 700		0.161
	F	0.177		0.330		0.175
	G	0.170		0.343		0.168
	H	0.158		1.564		0.158
	am:	0.167	gm	0.556	am′	0.167
Leeming	A	0.171		0.717		0.170
	B	0.168		0.796		0.168
	C	0.126		0.865		0.126
	D	0.155		1.038		0.165
	E	0.151		0.387		0.152
	am:	0.154	gm:	0.734	am:	0.154
Ringway	A	0.159		0.432		0.160
	B	0.153		0.643		0.150
	C	0.154		0.620		0.156
	D	0.148		1.198		0.150
	E	0.166		0.631		0.168
	F	0.169		0.512		0.170
	G	0.162		0.285		0.160
	H	0.144		0.269		0.148
	am:	0.157	gm:	0.517	am:	0.158
Brisbane	A	0.139		0.723		0.138
	B	0.151		0.657		0.152
	C	0.124		-		0.128
	D	0.147		-		0.154
	E	0.162		0.208		0.162
	am:	0.144	gm:	0.462	am:	0.147
Melbourne	A	0.175		-		0.172
	B	0.157		0.510		0.156
	C	0.164		0.198		0.162
	D	0.135		-		0.135
	E	0.150		0 973		0.152
	am:	0.156	gm:	0.486	am:	0.156
Sydney	A	0.144		0.771		0.145
	B	0.160		0.501		0.161
	C	0.158		0.612		0.158
	D	0.176		0.261		0 178
	E	0.164		0.529		0.164
	F	0.169		0.479		0.169
	am:	0.162	gm:	0.500	am:	0.162
Overall averages	am:	0.158	gm:	0.571	am:	0.158

site averages, as well as overall averages, are given (the arithmetic mean is used for the parameter *a* whilst the geometric mean, being less sensitive to outliers, is used for the more variable parameter *b*). Figure 4.4 displays these results for all records except those which have an undefined value for *b*. Whilst a slight negative correlation between *a* and *b* is seen, there is no discernible distinction between the various sites. Although considerable scatter is observed, the site averages (denoted by crosses) are reasonably well grouped. A separate examination of the *b* values indicates that they do not vary systematically with climate regime (as indicated by site in Figure 4.5) or sequentially (as indicated by the letters A to H in Table 4.3). Thus, it seems reasonable to fix the parameter *b* to the overall geometric mean (0.571) and refit $\rho(D)$, for each record, as a one parameter model $\rho'(D)$. A new estimate of *a*, denoted by *a'*, is obtained. If, however, this fixed value of *b* is inappropriate for a particular record, then the regression fit of $\rho'(D)$ will be poor at small D and compensated for at higher D, causing erroneous estimation of *a'*. However, the estimation of *a* depends mostly on those R(D) for which $D>D_{95}$ and so the problem can be circumvented by fitting $\rho'(D)$ to R(D) for $D \geq \max\{D_{95}, 6.2\}$, where D_{95} is estimated from the original regression of the two parameter model and 6.2 is the D_{95} value corresponding to b=0.571.

The resulting values of *a'* for each record appear in the right-hand column of Table 4.3. They are similar to the values of *a*, with the site and overall averages differing only slightly.

Thus, the estimation of *a* is generally insensitive to the value of *b* for hourly rainfall data. The values of $1+a'$ for each record (that is the correction factors) are represented in Figure 4.6, grouped by site to enable inter-site comparisons. A first inspection of the results might suggest that Eskdalemuir and Ringway have significantly different site means. More scatter is observed at the other sites, however, making comparisons more difficult. It is constructive to compare these results with those for N_{wet}.

4.4 Comparing correction factors with N_{wet}

It has been stated that sites subject to longer events, which are more likely to be divided between observational days, are expected to render higher correction factors. Figure 4.7 shows the mean correction factor for each site plotted on a linear scale and a similar plot for N_{wet} (from Table 4.2). The latter plot distinguishes Eskdalemuir as subject to relatively long events, Brisbane subject to relatively short events and the other sites as in-between and similar to one another. It is striking that the order of the sites in the two plots is preserved, in keeping with the above expectation.

The results can be used to construct particular hypotheses about the required correction factors for daily rainfall: in particular, that Eskdalemuir requires a higher correction factor than other sites, that Brisbane requires a lower

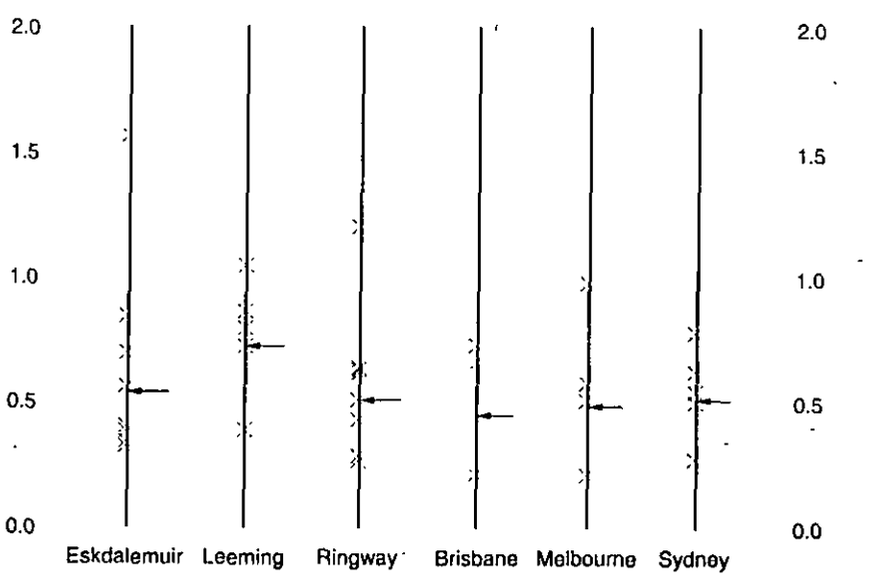


Figure 4.5 *Values of parameter b for each hourly rainfall record, grouped by site*

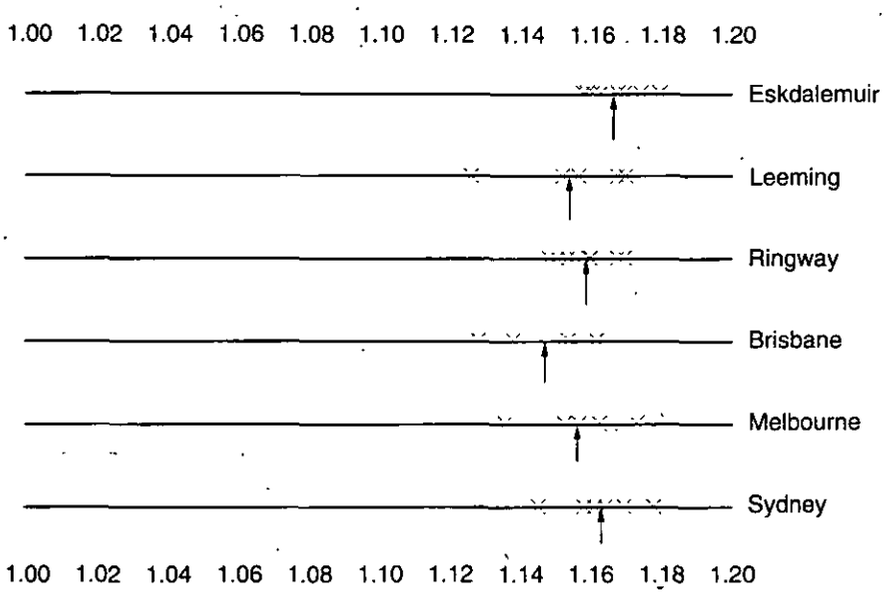


Figure 4.6 *Values of parameter 1+a' for each hourly rainfall record, grouped by site*

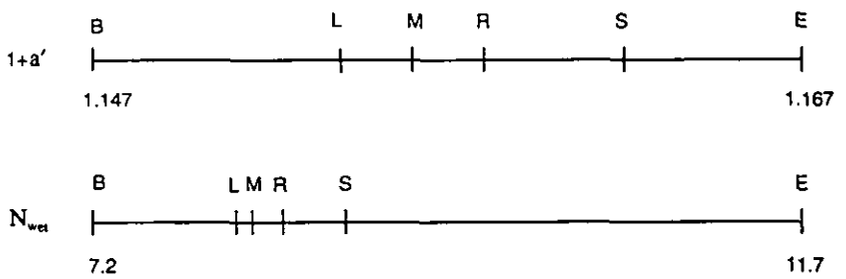


Figure 4.7 *Site-mean correction factors represented alongside site values for N_{wet} (each site is denoted by its first letter)*

correction factor than other sites and that the rest are not significantly different from one another. Based on the correction factors found for each record at each site, paired Student's t tests can be used to discern differences between the site means. A test statistic which assumes unknown but equal variances is used (Mood *et al.*, 1974). At the 95% significance level, one-sided tests indicate that Eskdalemuir has a significantly greater mean correction factor than all sites except Sydney. Similarly, Brisbane has a significantly lower mean correction factor than all sites except Leeming and Melbourne. Two-sided tests result in no significant differences between mean correction factors at the other sites. This is schematically represented in Figure 4.8 whereby sites circled together are not significantly different.

Thus, there is evidence to support the conjecture that Eskdalemuir, being subject mainly to long-duration frontal events, generally requires high correction factors for daily rainfall (with a site mean of 1.167) and Brisbane, being more prone to relatively short-lived monsoonal storms, generally requires lower correction

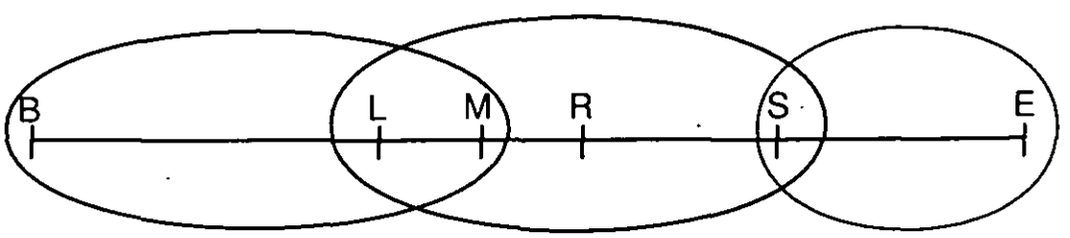


Figure 4.8 *Results of tests for significant differences between site-mean correction factors*

factors (site mean of 1.147). The other sites, and in particular Ringway, experiencing more middling duration events, generally require more moderate correction factors (overall mean of 1.158).

Before drawing conclusions and making recommendations, it is necessary to examine whether the above results are sensitive to certain aspects of the methodology.

4.5 Sensitivity of results

Timing of fixed intervals

The rainfall records, to which the above results pertain, are constructed so as to begin at 0900 on the first day. This is particularly relevant to the case D=24 because the observation day (0900 to 0900) then coincides with field practice. Nevertheless, it is interesting to examine the effect of starting the records at some other time.

The data records were reconstructed to begin at midnight. Very little difference is observed in the results for N_{wet}. Whilst changes are observed in the correction factors for individual records, the site means remain stable with no significant differences occurring. The overall mean correction factors are 1.158 (for 0900 start) and 1.160 (for midnight start).

It is concluded that, for the climates considered here, maxima extracted by either convention would suffice. In more tropical climates, where heavy rainfall is typically triggered by solar heating, the timing of the fixed intervals might be expected to be more influential.

Averaging across sites

Table 4.3 showed site-means for a'. The calculation of these means, however, does not take into account the (sometimes large) confidence intervals for the individual a'. In the case of UK sites, for which the separate records run consecutively, an alternative method for calculating the site-mean is obtained by concatenating the records for a given site to form a single large one. By maintaining the same period length, a greater number of periods is available from which to estimate the ratios R(D). In this way a single value of *a'*, as well as its confidence interval, is obtained for each site. For Eskdalemuir and Ringway, eight records are concatenated, whereupon the ratios R(D) are obtained by averaging across 256 periods; for Leeming, five concatenated records result in 160 periods. The graphs of R(D) against D, complete with the model fits, are shown in Figure 4.9 (only the R(D) which are used to fit the one-parameter model are shown).

Much less variation about the fitted model is observed than for the individual records. This demonstrates that much of the variation is due to sample variability rather than a poor model, and provides evidence for convergence to a limit. Not surprisingly, the new correction factors obtained show good agreement with those

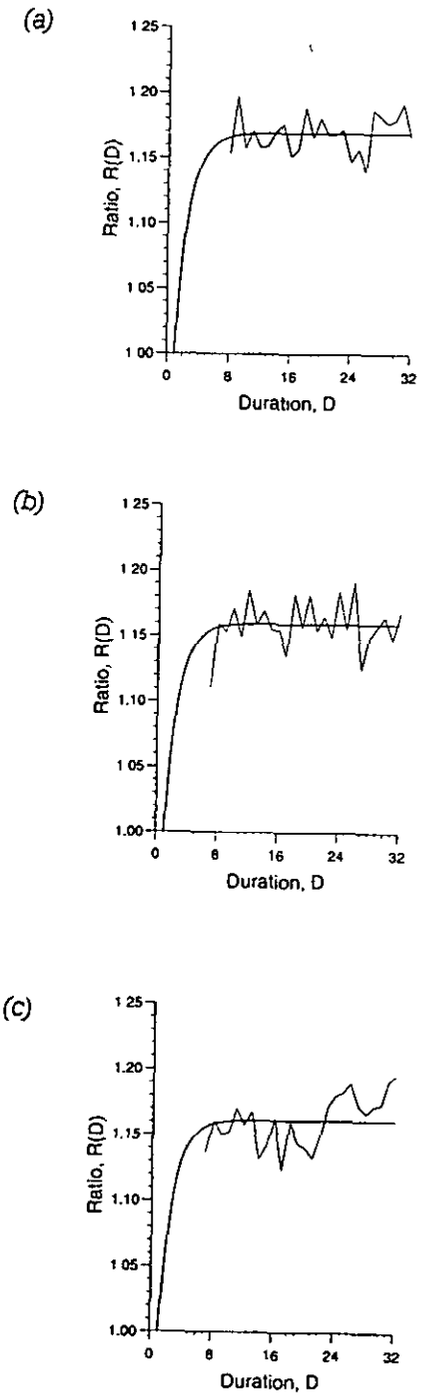


Figure 4.9 *R(D) against D, with fitted model, for concatenated hourly rainfall records (a) Eskdalemuir (b) Leeming and (c) Ringway*

Table 4.4 *Site-means for 1+a'; values of 1+a' for concatenated records including 95% confidence intervals from the regression fit*

Site	Mean of individual records	Concatenated record	95% confidence interval for concatenated record
Eskdalemuir	1.167	1.166	1.160 - 1.172
Ringway	1.156	1.156	1.149 - 1.163
Leeming	1.160	1.160	1.152 - 1.167

obtained by averaging across individual records (see Table 4.4).

Period length

Each maximum was extracted from a period of about 21 days; these represent less extreme events than do annual maxima which are more usually analysed. It is therefore necessary to investigate the sensitivity of the correction factors to changes in the period length. This is achieved by concatenating the individual records to form one long record for each UK site as described above. The ratios R(D) are then obtained for various period lengths. Table 4.5 details the results: four period lengths are considered and the resulting values of ρ^* (=1+a) and R(24) shown for each site. The number of periods pertaining to each period length is also shown for each site.

No systematic variation with period length is observed: at Eskdalemuir, ρ^* decreases as the period lengthens, whilst R(24) increases; at Leeming, ρ^* also decreases but R(24) merely fluctuates slightly; and at Ringway, ρ^*increases while R(24) is seen to decrease. Also, it should be noted that there is greater sample variability in the results for the longer period lengths due to the fewer number of periods available to estimate the ratios.

In summary, these results provide no evidence that correction factors vary systematically with period length. Visual inspection of the graphs for R(D) against D for the various period lengths at each site corroborate this conclusion.

4.6 Summary and conclusions

Hourly rainfall data were studied from three UK and three Australian sites and divided into records of fixed length. The discretization effect was investigated for each record at each site according to the methods described in Chapter 3. Thus, values for the model parameters *a* and *b* were obtained for each record and grouped by site to enable inter-site comparisons. Although there is a lot of variation in the values for *a* and *b*, the site means are in relative close proximity.

A separate analysis of the results for parameter *b* (which models the initial growth of R(D)) found no discernible inter-site differences. The value of *b* was thus fixed at the overall geometric mean (0.571) and the model re-fitted for each record to obtain comparable estimates, *a'*, of the *a* parameter.

An analysis of the *a'* values showed evidence of inter-site differences. Site mean values for the

Table 4.5 *Values of R(24) and ρ^* for concatenated hourly rainfall records using various period lengths*

Period length	Eskdalemuir			Ringway			Leeming		
	periods	R(24)	ρ^*	periods	R(24)	ρ^*	periods	R(24)	ρ^*
21 days	256	1.163	1.165	256	1.152	1.158	160	1.161	1.148
43 days	128	1.154	1.160	128	1.143	1.158	80	1.157	1.144
85 days	64	1.171	1.152	64	1.154	1.165	40	1.164	1.145
171 days	32	1.182	1.149	32	1.171	1.162	20	1.157	1.140

correction factor, $1+a'$, were found in the approximate range 1.147 to 1.167 with an overall mean value of 1.158. Concentrating on the correction of daily rainfall maxima (D=24), the results for $1+a'$ were compared to the number of wet hours (N_{wet}) at each site. A positive correlation between them, as anticipated by theoretical arguments, was observed. Values of N_{wet} were then used to construct hypotheses concerning the correction factors, $1+a'$, at each site: namely, that Eskdalemuir requires a higher correction factor than the other sites, Brisbane a lower correction factor than the other sites, and the others require broadly similar correction factors. Student's t tests were used at the 95% significance level to test differences in the site means: the resulting groupings lent support to the hypotheses.

Further analysis showed no evidence that the correction factors found are dependent upon the timing of the fixed intervals or the period length. In particular, for each UK site, it was possible to obtain a site value for a' by analysing a single long record (by the concatenation of the standard records). This resulted in very similar values for a' and a much better fit (in terms of deviation about the model), thus increasing confidence in the model as well as in the site means obtained.

In conclusion, the results suggest that the correction factors commonly used are somewhat low. Instead of 1.13 and 1.14 for example, a range of 1.15 to 1.17 is indicated. For the correction of fixed 24-hour (daily) maxima, the lower end of this range is suggested for climates prone to short-lived convective/monsoonal rainfall, the upper end for climates prone to frontal systems producing events of longer duration, and 1.16 for middling climates and as a general guideline.

5 Comparison of variables

5.1 Introduction

This chapter concerns the comparison of correction factors for different climate variables. In particular, the temporal variability of a process is investigated for its influence upon the magnitude of required correction factors.

Consider the discretization effect by imagining the time interval, corresponding to a fixed maximum, being free to slide to and fro in search of the true maximum. As the interval slides, so the accumulations found will differ. For gradually changing processes, such as that of air temperature, these differences are likely to be relatively small compared to those for erratic, intermittent processes, such as that of rainfall, for which plenty of opportunity exists for sudden and sharp changes to occur. More precisely, the slower the decay in temporal autocorrelation, the lower the expected correction factor for the data.

This conjecture is examined by comparing three climate variables - rainfall, wind speed and air temperature. The data, all taken from Eskdalemuir in the southern uplands of Scotland, are at the hourly resolution and span the years 1970 to 1989. Thus, there are ten possible data records for each variable, labelled A to J accordingly, although data quality requirements prevented analysis of some records. As a result, eight rainfall, six wind speed and seven air temperature records are examined. The rainfall records (A-H) correspond to those used for Eskdalemuir in the previous chapter. More details of the data can be found in Appendix C.

The correction factors, ρ^*, as defined in Chapter 3, are calculated for each record and compared to the measure of erraticness defined below.

5.2 Variograms and erraticness

Let a data record, of length $n=2^{14}$ values, be represented by $X(t)=\{x(1), x(2), \ldots, x(n)\}$. The calculation of autocorrelation assumes not only stationarity in the mean of X(t) but also constant and finite variance, σ^2. An alternative is the *semi-variance*, $\gamma(h)$, defined by

$$\gamma(h) = \tfrac{1}{2}E[(x(t+h)-x(t))^2] \qquad 5.1$$

where $E[\cdot]$ denotes statistical expectation and h is known as the (time) *lag*. Estimation of $\gamma(h)$, defined by the sample semi-variance

$$s^2(h) = \frac{1}{2(n-h)} \sum_{i=1}^{n-h} (x(i+h)-x(i))^2 \qquad 5.2$$

is unbiased if X(t) is stationary in the mean and $\gamma(h)$ is time-invariant. If the variance is constant and finite, however, then the autocorrelation, C(h), and semi-variance are related by

$$\gamma(h) = \sigma^2(1-C(h)) \qquad 5.3$$

The graph of $\gamma(h)$ against h is called the *variogram*, full discussion of which can be found in, for example, Webster and Oliver (1990). To enable comparisons between variograms for different data records, a non-dimensionalized version of semi-variance is defined by

$$S(h) = \frac{2s^2(h)}{s_1 s_2} \qquad 5.4$$

where s_1 and s_2 are the standard deviations of the first n-h and last n-h data points, respectively.

In the context of simply-scaling Gaussian processes, erraticness can be quantified by the Hausdorff fractal dimension, d_H, which is related to the gradient, α, of the log-log variogram at small h, by

$$d_H = 2-\alpha/2 \qquad 5.5$$

Simple scaling renders the log-log variogram linear for small h, whereupon the gradient α is well defined (see examples in Coyle *et al., 1991)*. Whilst rainfall and wind speed are non-Gaussian, this study follows in the same spirit as Constantine and Hall (1994) by defining the *effective fractal dimension*, d_E, as

$$d_E = d_H = 2-\alpha/2 \qquad 5.6$$

where α is the gradient of log(S(h)) with respect to log(h) for small h in the range over which

linearity appears to hold. Note that $1 \leq d_E \leq 2$. A more detailed discussion of effective fractal dimension, as defined above, can be found in Dwyer and Reed (1994).

5.3 Results

Some typical graphs are presented and discussed before the complete set of results is examined and comparisons made.

Variograms

The log-log dimensionless variogram for rainfall (Figure 5.1a) reveals a limited range of linearity over which simple scaling holds. The graph quickly approaches the *sill* (S(h)=2) which is reached when the lag is large enough for the data points to become uncorrelated. For some records a very rapid approach to the sill makes linearity at small h difficult to discern. For consistency, regression lines are fitted between h=1 and h=5 for all data records. For rainfall, these regression lines are of shallow gradient, whereupon the effective fractal dimensions are large (d_E = 1.77 for the example shown).

Linearity in the log-log variogram is observed over a wider range of lags for hourly mean wind speed data (e.g. Figure 5.1b). Nevertheless the regression line is fitted between h=1 and h=5 (as for rainfall) because increasing this range only increases the error in estimating d_E (Constantine and Hall, 1994). The values of d_E are smaller than for rainfall (d_E = 1.49 for the example), representing less erraticness.

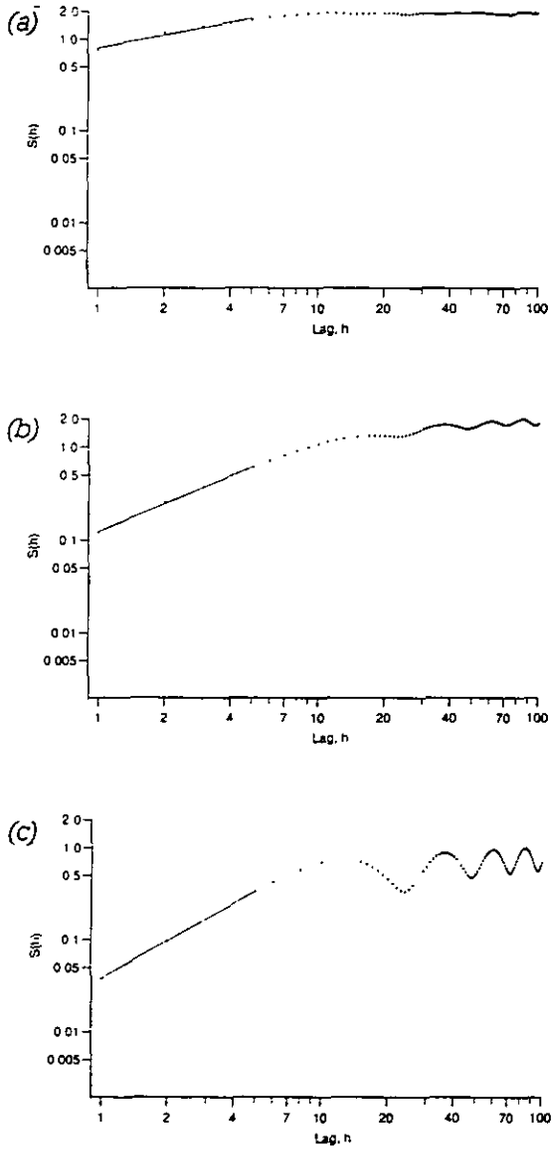


Figure 5.1 *Log-log variograms with regression lines: (a) hourly rainfall record B, (b) hourly wind speed record F and (c) hourly air temperature record G, all at Eskdalemuir*

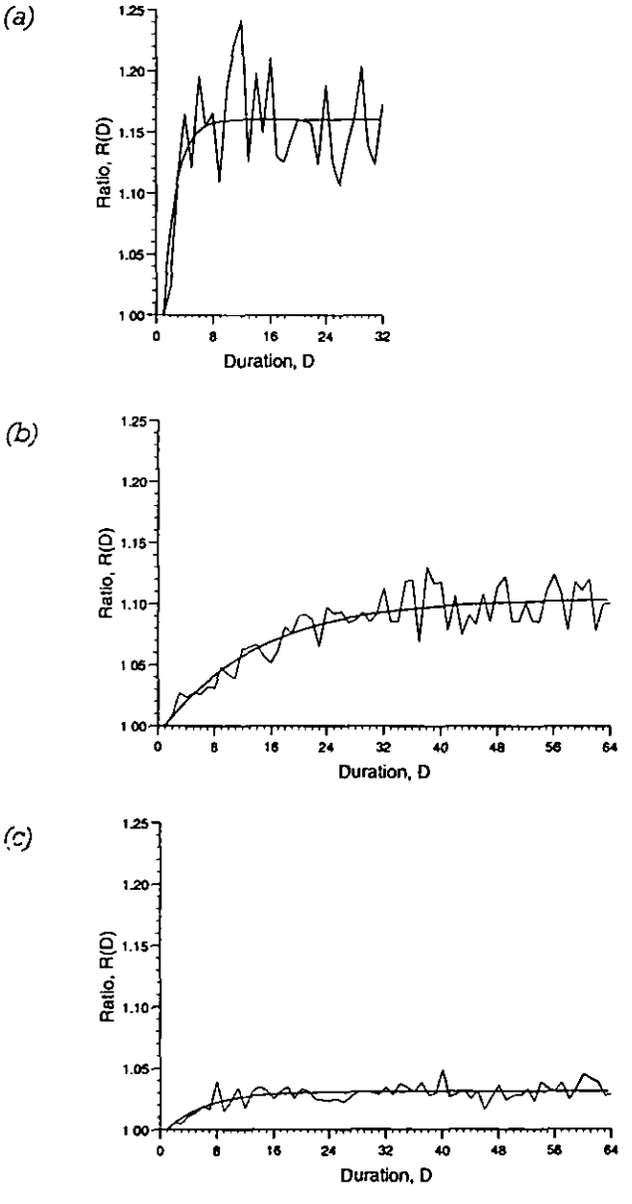


Figure 5.2 *R(D) against D with fitted model: (a) hourly rainfall record B, (b) hourly wind speed record A and (c) hourly air temperature record G, all at Eskdalemuir*

The air temperature records render the lowest values of d_E, indicating the relative smoothness of the variable. The example in Figure 5.1c (for which d_E = 1.32) clearly shows the presence of the diurnal cycle. For consistency, the regression line is again fitted between h=1 and h=5.

Correction factors

An example of a graph of R(D) against D for rainfall, complete with the fitted model, is shown in Figure 5.2a. Generally, a rapid increase in R(D) with D, and a good deal of residual variation about the fitted model were observed for rainfall. Values of ρ^* were in the range 1.16 to 1.18 for the Eskdalemuir site.

The example graph of R(D) against D for wind speed (Figure 5.2b) demonstrates less variation about the regression line, a slower increase in R(D) over D, and a lower limiting value (ρ^* = 1.105). The model $\rho(D)$ appears particularly suitable for wind speed data.

The graph of R(D) against D for the air temperature example (Figure 5.2c) results in a much lower correction factor (ρ^* = 1.031). For air temperature data, the increase in R(D) with D is not especially well described by the model. Note that, when analysing air temperature minima, the *upper* (rather than the lower) 1% quantile is used for the datum.

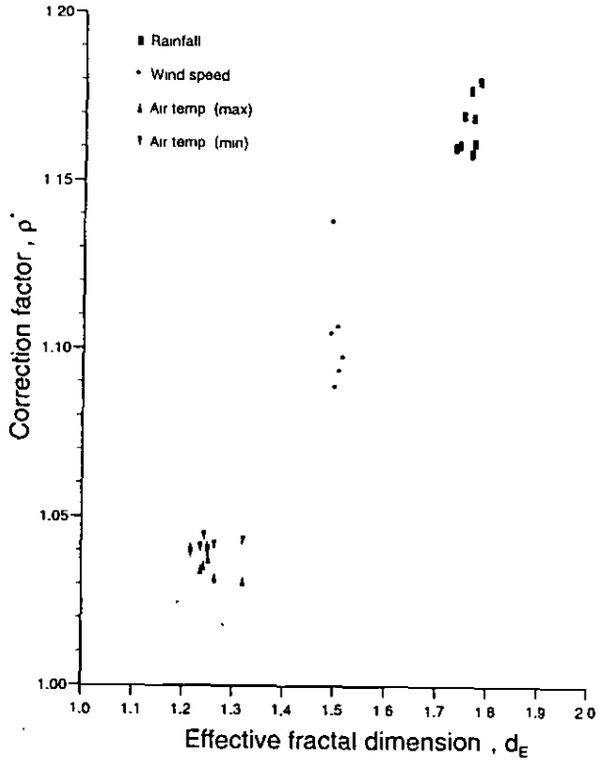


Figure 5.3 *ρ^* against d_E for different climate variables*

Table 5.1 *Means and standard deviations (s.d.) of ρ^* and d_E for each climate variable*

Variable	Mean of ρ^*	s.d. of ρ^*	Suggested range (to 2 d.p.)	Mean of d_E	s.d. of d_E	Range
Air temperature:						
maxima	1.036	0.004	1.03 - 1.04	1.26	0.032	1.19 - 1.32
minima	1.041	0.002	1.04 - 1.05	1.26	0.032	1.19 - 1.32
Wind speed	1.099	0.008	1.08 - 1.11	1.50	0.009	1.48 - 1.52
Rainfall	1.167	0.008	1.15 - 1.18	1.76	0.016	1.73 - 1.79

Comparisons

The complete results for d_E and ρ^* are summarized in Figure 5.3. There is pronounced segregation of d_E between the three variables. Although more scatter is observed in the values of ρ^*, there is also clear segregation, with rainfall requiring the highest correction factors and air temperature the lowest (with maxima requiring slightly higher correction than minima in the samples studied). Table 5.1 shows the mean values of ρ^* and d_E for each variable, their standard deviations and suggested ranges (given as two standard deviations each side of the mean). Note that the outlier for wind speed is due to a poor fit of the model; whilst shown in Figure 5.3 for completeness, it was not included when calculating the statistics for Table 5.1.

The relation that emerges between ρ^* and d_E in terms of the different climate variables affirms the expected association between erraticness and correction factors. It should be stressed, however, that the correction factors are datum dependent; here, the lower 1% quantile is used (upper 1% quantile for air temperature minima) which corresponds to the natural datums of zero for hourly rainfall and hourly wind speed.

The relationship between ρ^* and d_E does not appear to persist *within* the individual variables, however. It is interesting to examine this further using the rainfall data from the various locations taken from Chapter 4. The effective fractal dimension, d_E, is computed for each record at each site and plotted against the correction factors, 1+a', already calculated. The result is shown in Figure 5.4. Clearly, the relationship

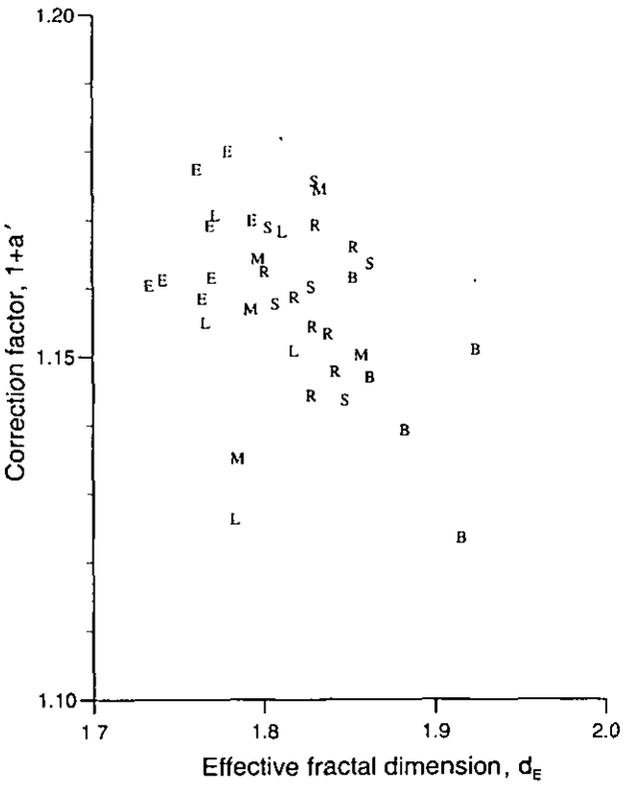


Figure 5.4 *1+a' against d_E for hourly rainfall records of Brisbane, Eskdalemuir, Leeming, Melbourne, Ringway and Sydney*

does break down at this intra-variable scale: the reasons for high or low correction factors cannot be distinguished by a single index of overall erraticness. Nevertheless, some grouping by location is observed, indicating a slight negative correlation between site average erraticness and correction factor. It is interesting that the Eskdalemuir and Brisbane sites stand out as producing the most extreme values of d_E, as they do for the correction factors 1+a' and the values of N_{wet} (see Chapter 4). Rainfall is known to be multifractal (characterized by *multiple* scaling and a spectrum of fractal dimensions) and explicit recognition of this would be warranted in a more detailed investigation.

5.4 Summary and conclusions

Using hourly data for Eskdalemuir in southern Scotland, correction factors for rainfall, wind speed and air temperature were determined based upon a lower 1% quantile datum. Air temperature minima as well as maxima were examined, for which the upper 1% quantile was used as the datum. The results clearly discriminate between the three climate variables. Note, however, that these results are specific to the Eskdalemuir site: thus, for hourly rainfall, the more comprehensive results of the previous chapter are to be preferred; for wind speed and air temperature the results are necessarily more tentative. Nevertheless, the results highlight the differences between the variables and can be used as a guideline.

Investigation of the temporal erraticness and intermittency of the hourly data, via estimation of the effective fractal dimension, clearly discriminates between the three variables. Furthermore, the dimension is independent of the datum. This quantity, and similar dimension estimates, therefore have potential as meaningful measures of the temporal character of environmental variables. In the case of rainfall, however, linearity in the log-log variograms is relatively weak and a multifractal formalism may be appropriate (see for example Rajagopalan and Tarboton, 1993).

In practice, correction factors are required when finer resolution data are not available to estimate sliding maxima, whence they are not available to estimate effective fractal dimension either. Nevertheless, d_E so clearly distinguishes between different variables that it may be estimated from hourly data collected at a different time and/or location. A broad indication of the expected correction factors for that particular variable can then be obtained using Fig. 5.3.

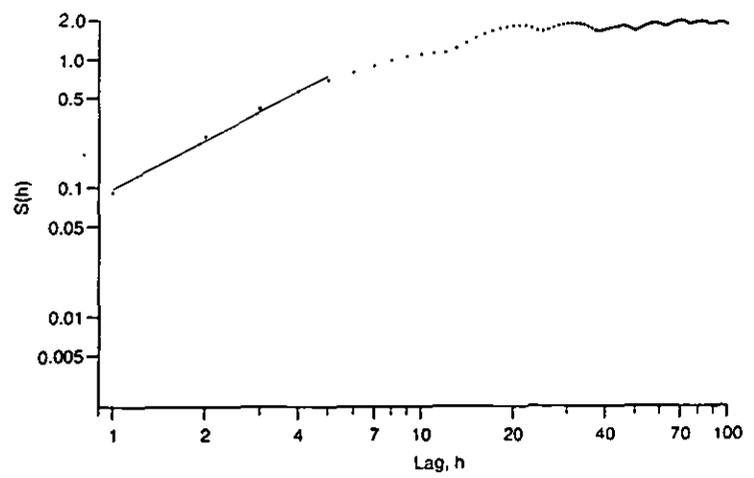


Figure 5.5 *Log-log variogram with regression line for hourly mean tide residual at Walton-on-the-Naze*

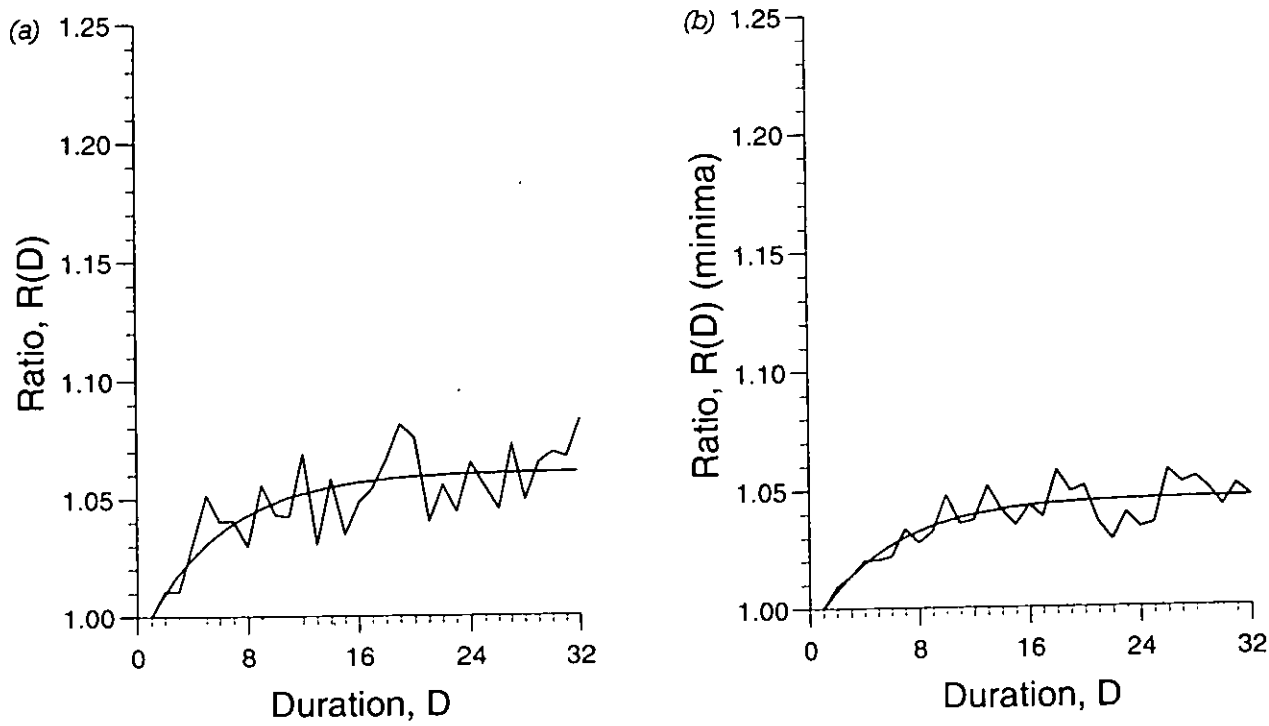


Figure 5.6 *R(D) against D, with fitted model, for (a) period maxima and (b) period minima of hourly tide residual at Walton-on-the-Naze*

The approach is illustrated using a single record of mean tide residual measured at the hourly resolution at Walton-on-the-Naze in Essex. The effective fractal dimension is calculated as d_E = 1.371 (Figure 5.5) which, according to Figure 5.3, indicates a correction factor in the region of 1.06. This compares reasonably well with the actual correction factors obtained by analysing the data (ρ^*=1.061 for period maxima and ρ^*=1.047 for period minima). The relevant graphs are shown in Figure 5.6.

6 Rainfall durations two to 32 days

6.1 Introduction

The correction of rainfall maxima of durations two to 32 days is investigated by reference to daily rainfall records. The first approach is to examine evidence for such correction factors being any different to those found so far for the shorter durations. This is done by comparing the results from hourly and daily data at the same sites (Section 6.2), and deliberately coarsening hourly data to observe the effect on the resulting correction factors (Section 6.3). Secondly, a number of daily rainfall records are analysed and mean values for the parameters *a* and *b* obtained (Section 6.4). The effect of data resolution upon the variogram is also investigated.

6.2 Comparing results from hourly and daily data

Six sites were examined in Chapter 4 with respect to hourly data. The results for Eskdalemuir and Brisbane were at opposite extremes with the other sites grouped in-between. Eskdalemuir, Brisbane and Melbourne are thus chosen as representative of the range of results and climate regimes studied. Daily data for each of the three sites are subjected to the same analysis as the hourly data, rendering correction factors, ρ^*, for each record. Details for the daily data are contained in Appendix C.

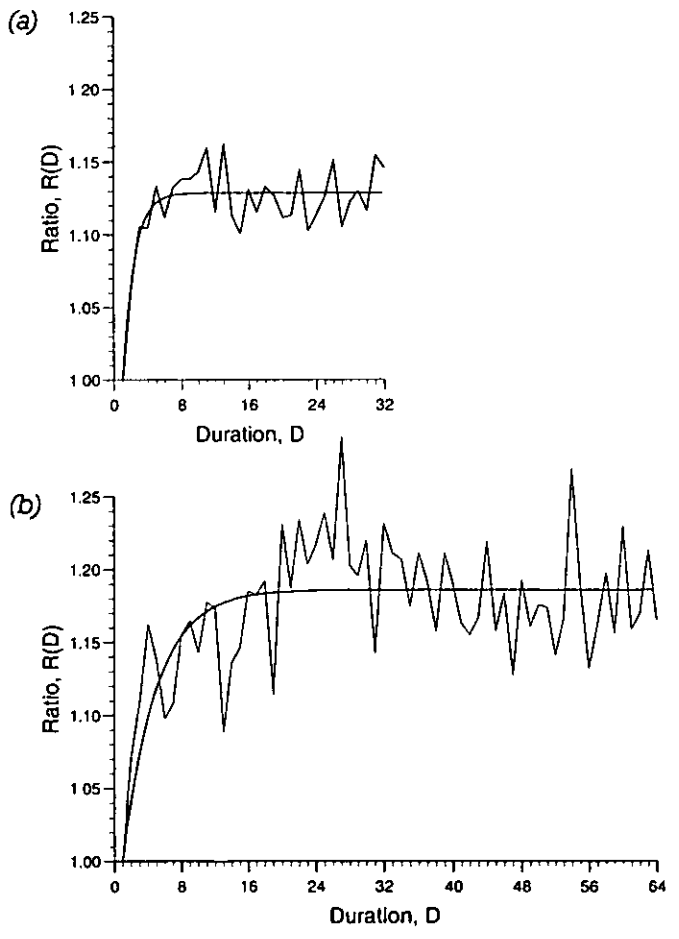


Figure 6.1 *R(D) against D, with fitted model, for daily rainfall: (a) Eskdalemuir and (b) Brisbane record 1*

Examples of the graphs of R(D) against D, with the fitted model, are shown in Figure 6.1. The complete set of results is detailed in Table 6.1. Figure 6.2 illustrates that the range of correction factors for the daily data is similar to that for the hourly data. At Eskdalemuir, the single result for daily data is notably lower than those for the hourly data; this is not repeated at the other two sites, however, and is thought to be a peculiarity of the particular daily record.

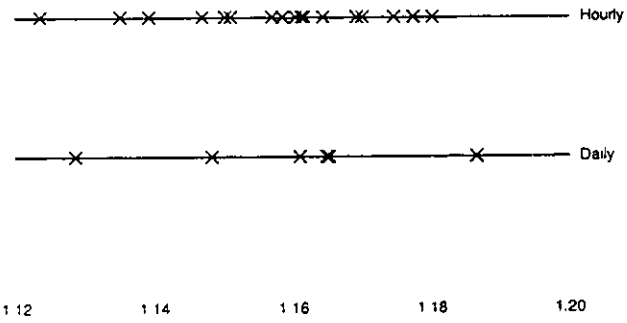


Figure 6.2 *Comparison of correction factors derived from hourly and daily rainfall data at the same sites*

6.3 Coarsening hourly data

The eight consecutive hourly records available at Eskdalemuir and Ringway enable deliberate coarsening of the data to obtain four 2-hourly, two 4-hourly and one 8-hourly record at each site. Extracting the values of ρ^* for each allows possible sensitivity of the correction factor to resolution to be explored.

The results are detailed in Table 6.2 and illustrated in Fig. 6.3. Details of the 2, 4 and 8-hourly records can be found in Appendix C. For Eskdalemuir, the correction factors appear to be decreasing up to the 4-hourly resolution, although the pattern does not extend to the 8-hourly record. No pattern emerges for the Ringway data. The two sites considered

together portray the same range of correction factors at each resolution.

Table 6.1 Values of ρ for rainfall records of varying resolutions*

Site	Hourly data		Daily data	
	record	ρ^*	record	ρ^*
Eskdalemuir:	A	1.169	1	1.129
	B	1.160		
	C	1.162		
	D	1.180		
	E	1.161		
	F	1.177		
	G	1.170		
	H	1.158		
Melbourne:	A	1.175	1	1.161
	B	1.157	2	1.165
	C	1.164	3	1.165
	D	1.135		
	E	1.150		
Brisbane:	A	1.139	1	1.186
	B	1.151	2	1.148
	C	1.124		
	D	1.147		
	E	1.162		

6.4 Results for other daily records

Table 6.3 shows the values of parameters *a* and *b* for the 21 daily rainfall records available to the study (which includes those discussed in Section 6.2 above). The arithmetic mean for parameter *a* is 0.165 as compared to 0.158 for the set of hourly rainfall records. The geometric mean for parameter *b* of 0.357, as compared to 0.571 for the hourly records, represents a slower growth rate to the limiting value. As with the hourly data, however, the values for *b* are rather variable and so the geometric mean obtained is used as a broad guideline in the implementation of the full correction model. Note that whilst the hourly data are from sites in the UK and Australia, the daily data also include some South African sites.

6.5 Variograms

Clear systematic changes with data resolution are observed with respect to the variograms and effective fractal dimension, d_E, as defined in Chapter 5. This is illustrated for the Eskdalemuir site in Figure 6.4 in which log-log variograms at each resolution are shown. As the resolution coarsens, the variograms become flatter, giving rise to higher values for d_E. At resolutions of one hour and shorter (not shown), approximate linearity over small lags enables a reasonable estimate of the gradient (and thus d_E). At coarser resolutions, however, the sill ($S(h)=2$) is reached sooner and the behaviour at small lags becomes indiscernible; for this reason, no attempt was made to extract and compare values for d_E in the examples. Analysis of other data shows the same systematic behaviour.

Table 6.2 *Values of ρ* for rainfall records of varying resolutions*

Resolution	Record	Eskdalemuir ρ^*	Ringway ρ^*
1 hour	A	1.169	1.159
	B	1.160	1.153
	C	1.162	1.154
	D	1.180	1.148
	E	1.161	1.166
	F	1.177	1.169
	G	1.170	1.162
	H	1.158	1.144
2 hours	K	1.165	1.158
	L	1.163	1.181
	M	1.166	1.172
	N	1.158	1.152
4 hours	X	1.155	1.162
	Y	1.160	1.152
8 hours	2	1.172	1.166

In the theory (Constantine and Hall, 1994; Dwyer and Reed, 1994), the fractal properties are defined by the behaviour of the variogram as the lag, h, tends to zero. Since it is not possible to measure at an infinitesimal resolution, the estimates of fractal dimension obtained using finite resolution data assume that the observed

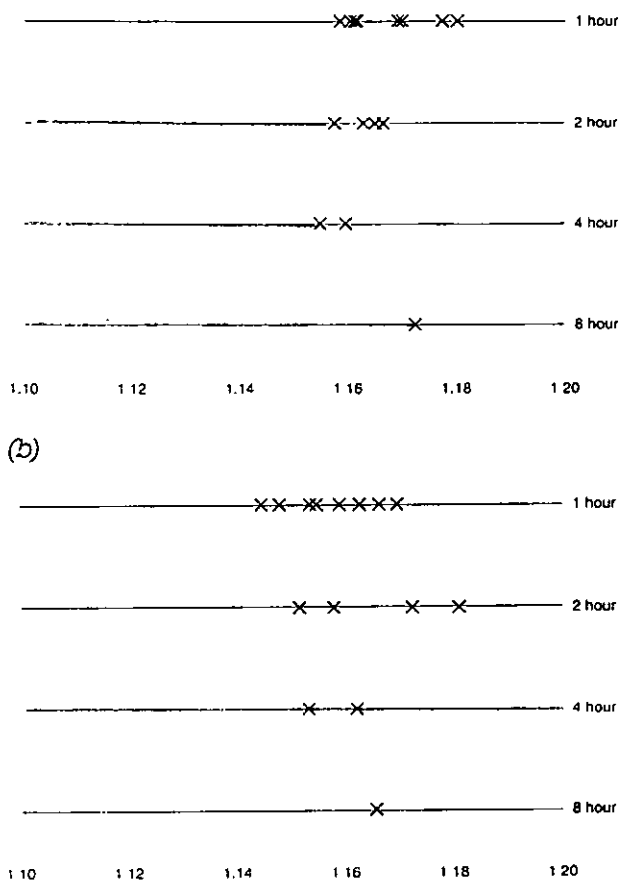


Figure 6.3 *Comparison of correction factors derived from rainfall data of varying resolutions at (a) Eskdalemuir and (b) Ringway*

linearity continues with the same gradient as h decreases to zero. Hence, the systematic changes, observed in the rainfall variograms, merely indicate that the limiting linear behaviour is poorly estimated at the coarser resolutions. Indeed, analysis of sub-hourly rainfall data suggests that a resolution of one hour is not fine enough to reveal the limiting fractal behaviour, whereupon the estimates of d_E for rainfall in Chapter 5 are particular to the one hour resolution.

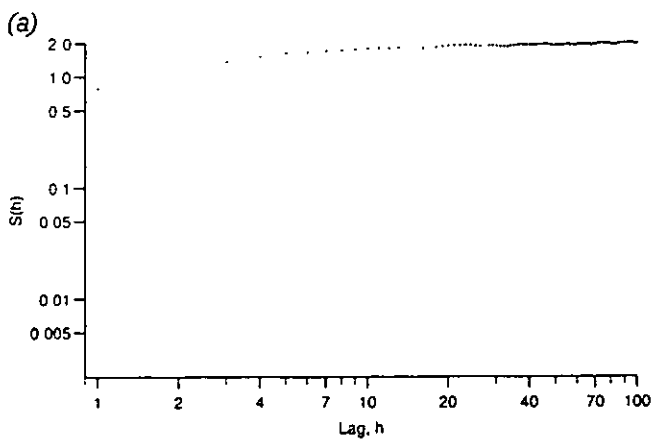


Figure 6.4 *Log-log variograms for Eskdalemuir rainfall: (a) record A (hourly) (b) L (2-hourly) (c) X (4-hourly) (d) 8-hourly and (e) daily*

Table 6.3 Values of parameters a and b for various daily rainfall records

Site	Record	a	b
UK:			
Alwen Reservoir	.	0.148	0.648
Creech Grange .	.	0.179	0.349
Edinburgh	.	0.167	0.512
Eskdalemuir	.	0.129	0.730
Etton on the Welland	.	0.174	0.190
Australia:			
Brisbane	1	0.186	0.251
	2	0.148	0.157
Hobart	1	0.142	0.757
	2	0.164	0.159
Melbourne	1	0.161	0.404
	2	0.165	0.175
	3	0.165	2.039
Perth	1	0.167	0.451
	2	0.175	0.315
Sydney	1	0.167	0.456
	2	0.166	0.323
South Africa:			
Cape Town	.	0.185	0.353
Durban	.	0.176	0.106
Johannesburg	.	0.174	0.292
Richmond	.	0.154	0.303
Upington	.	0.163	0.527
means:		am: 0.165	gm: 0.357

6.6 Conclusions

Over the range of scales studied (one hour to 32 days), there is no indication of systematic changes in rainfall correction factors. Thus, correction factors for maxima of several hours duration, are, on average, similar to those for maxima of several days duration. The qualifier "on average" should be stressed, however, for correction factors may change systematically with time scale at a particular site if, for example, the typical rainfall profile changes systematically (see Chapter 4). The analysis of several daily rainfall records suggested that the growth rate, as measured by parameter *b*, is generally lower than for hourly records but still rather variable.

In contrast, the variograms for rainfall do change systematically with time scale: as the resolution becomes finer, a better appreciation of the fractal behaviour is obtained.

7 Additional investigations

7.1 Catchment average rainfall

The rainfall results presented so far relate to point (single gauge) rainfall data. In order to investigate differences between point and catchment rainfall, the Alwen at Druid catchment was selected for study, covering an area of 185km^2 in north Wales. Catchment rainfall is calculated by averaging across 22 gauges, each weighted according to its estimated long-term average annual rainfall. A standard record of 15-minute catchment rainfall is then compared with a 15-minute point rainfall record, covering the same period, for a gauge centrally located within the catchment.

The calculation of catchment rainfall necessarily has a smoothing effect: catchment rainfall events tend to start and stop less abruptly and have more uniform profiles than point rainfall events. Evidence for this can be found in the summary statistics for the two records (Appendix C) and the values for N_{wet} (with the catchment value being the larger). Perhaps the most convincing support, however, is found by simply comparing graphs of the time series (Figure 7.1). According to the reasoning outlined in Chapter 4, longer event durations imply larger correction factors, as do more uniform hyetographs (Appendix A), whereupon

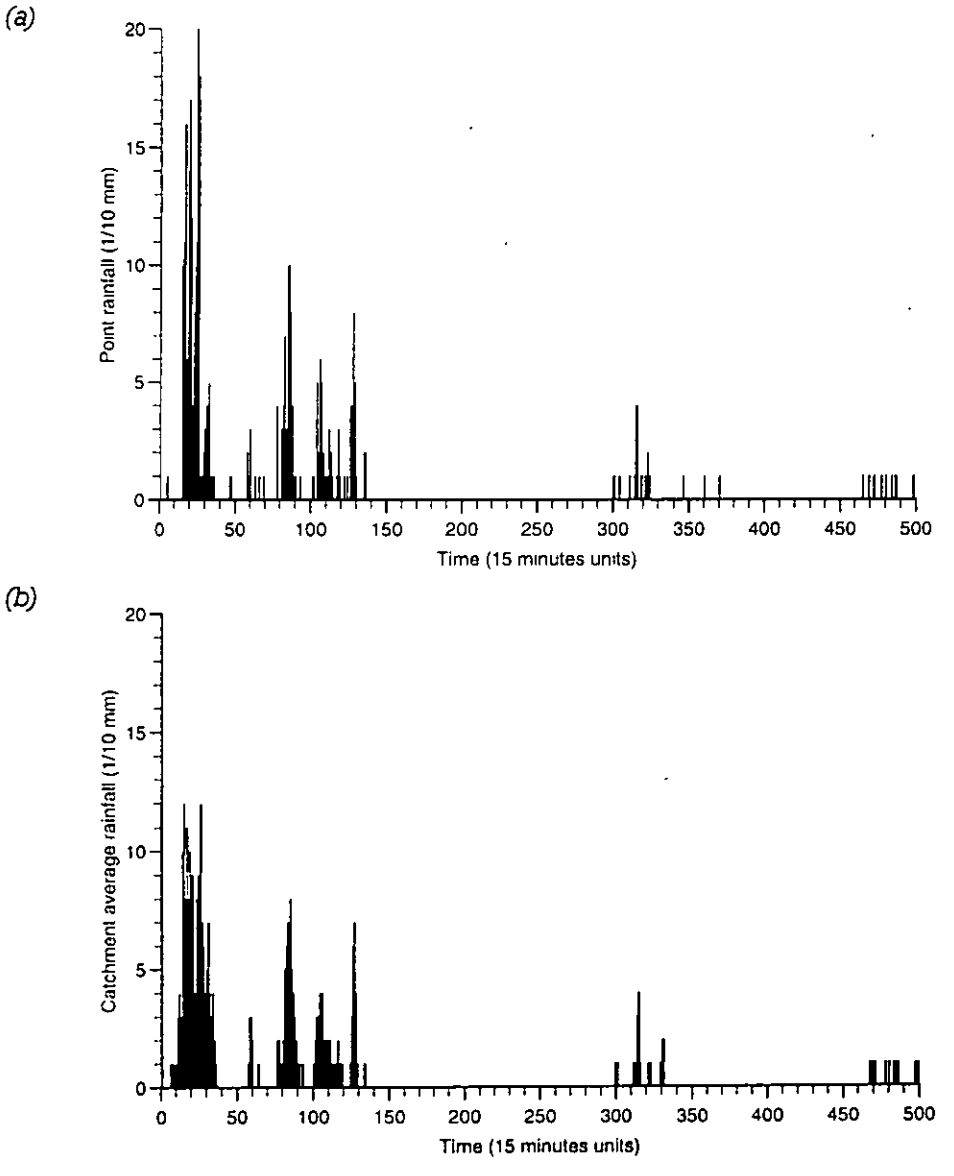


Figure 7.1 *Time series of 15-minute rainfall accumulations at Alwen at Druid: (a) point rainfall and (b) catchment rainfall*

catchment rainfall might be expected to require higher correction factors than point rainfall. Subjecting the two records to the usual analysis, catchment rainfall renders a correction factor $\rho^*=1.190$ as compared to $\rho^*=1.169$ for point rainfall. The latter is outside the former's 95% confidence interval (pertaining to the regression fit) and so it appears that the catchment rainfall, in this case, does indeed require a higher correction factor than the point rainfall.

The smoother, less intermittent nature of catchment rainfall also implies a lower effective fractal dimension ($d_E=1.657$ is observed for the catchment rainfall record as compared to $d_E=1.817$ for the point rainfall record). If they are considered as separate climate variables then the results of Chapter 5 would suggest that catchment rainfall requires lower correction factors than point rainfall, which is in contradiction to the above findings. This suggests that, in terms of the fractal analysis of Chapter 5, catchment and point rainfall cannot be regarded as distinct climate variables between which d_E can discern different correction properties. However, as only a single record has been studied, it is recognized that a more extensive investigation is required before any firm conclusion can be drawn.

7.2 Instrumentation

There are a number of types of instrument for measuring rainfall accumulations. The one chosen for a particular site will depend upon such things as cost, the accuracy and resolution demanded by the application, and the climatology of the region.

The hourly rainfall data analysed in Chapter 4 are derived from tilting syphon raingauges. Rain-water is funnelled into a small drum containing a float which operates a pen; as the water level rises, the pen traces a graph on paper attached to a rotating cylinder. When the water reaches a certain level the drum tilts, causing a syphon to operate which drains the water away. The trace is used to extract accumulations (in units of 0.1mm) for each hour.

Also popular for measuring hourly rainfall is the tipping-bucket rain gauge. A double bucket is counter-balanced by a magnet below to form an arrangement resembling that in Figure 7.2. As the bucket beneath the funnel fills, the balance is tipped, causing the pendulum to swing. The movement of the magnet triggers an electric circuit connected to a time-recording device; the tipping of the bucket causes it to empty while the second bucket takes up the position beneath the funnel until the reverse swing occurs. The size of the buckets determines the volumetric resolution of the data and the tip-times recorded can be converted into rainfall depths for each hour.

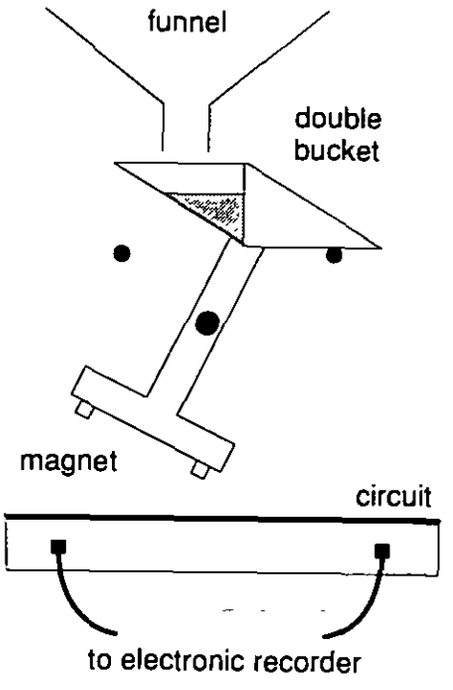


Figure 7.2 *Principle of the tipping bucket raingauge*

The sensitivity of the values of R(D) is investigated with respect to the choice between

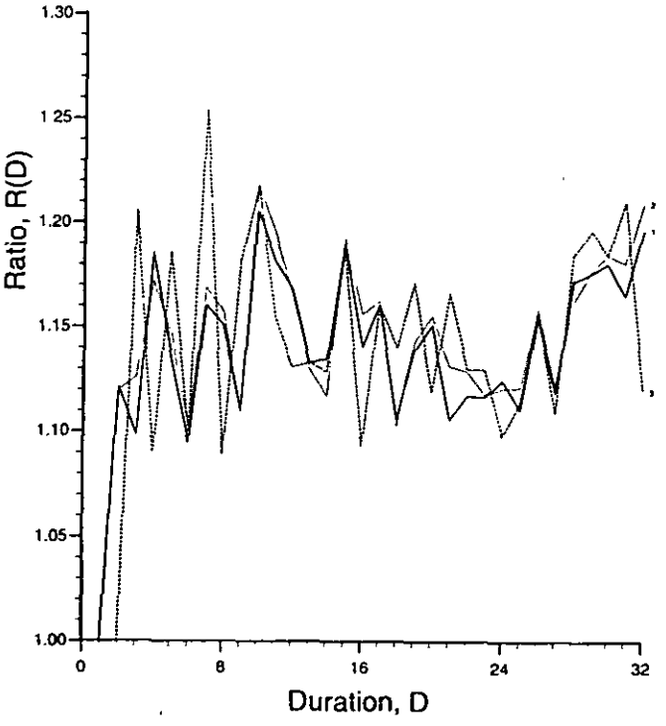


Figure 7.3 *R(D) against D for tilting syphon data (graph 1) and simulated tipping bucket data of resolutions 2.0 mm (graph 2) and 10.0 mm (graph 3) for Eskdalemuir hourly rainfall record C*

tilting syphon or tipping bucket devices. The tilting syphon data for one of the Eskdalemuir hourly rainfall records (having a resolution of 0.1mm) are used to synthesize tipping bucket data for bucket sizes corresponding to rainfall depths of 0.2, 0.5, 1.0, 2.0, 5.0 and 10.0 millimetres. This is achieved by calculating a running total for the original data until the bucket size is exceeded: the total, rounded down to the nearest multiple of the bucket size, is assigned to the hour in which the exceedance occurred and the remainder carried over to a restarted running total. As the bucket size is increased, the rainfall profiles become concentrated into fewer, higher peaks.

For each bucket size simulated, the graph of R(D) against D is compared with that for the original (tilting syphon) data. Up to a bucket size of about 2.0 mm, only slight differences are observed; larger differences become apparent, particularly at small D, for the 5.0 and 10.0 mm buckets. This is illustrated in Figure 7.3 which compares results for the original, the synthesized 2.0 mm and the synthesized 10.0 mm data. Of course, 10.0, 5.0 and even 2.0 mm buckets are too large for discriminating hourly rainfall in the UK. Thus, it appears that R(D) is insensitive to the choice between tilting syphon and tipping bucket devices provided the instrument resolution is apt for the purpose. This conclusion is supported by similar results for other data records.

7.3 Instantaneous data

The correction factors discussed so far arise because the data are discretized in the form of averages (or accumulations) measured over some time interval. Another form of measurement discretization is that of "spot sampling" whereby *instantaneous* readings are taken at regular (or irregular) time intervals. A lack of knowledge of the process between sampling times means that the instantaneous maximum (over some sampling period) is, in general, missed. The maximum of the spot-sample readings is therefore an under-estimate of the true instantaneous maximum. This problem is quite different from the one discussed up to now in terms of both the reason for under-estimation and the applications (those concerned with instantaneous rather than accumulated maxima). Consequently, the discussion here is brief, recognising that the problem deserves a full, separate investigation.

Attention was drawn to the problem from river water quality standards in general, and a particular application where the Hydrogen ion concentration in upland streams is of interest (Robson, 1993). Maximum H-ion concentration may be relevant, for example, to the survival of living organisms in the stream. Typical spot-sampling intervals for the measurement of stream chemistry can be as long as one week. Such infrequent measurement can severely under-appreciate extremes. To investigate the extent of under-estimation, one record, containing 2^{13} data values of 15-min instantaneous H-ion concentration in the Upper Hore is analysed. The stream drains a 1.78km^2 sub-catchment of the Wye basin of Plynlimon in mid-west Wales. Concurrent flow data, again sampled at 15-minute intervals, are also examined.

For each record, the maximum recorded value, max(1), is extracted. By discarding every other data point, the record can be artificially degraded to simulate a sampling interval of 30 minutes. Two such degraded, or *thinned*, records can be obtained: the first by discarding even-numbered data points, the second by discarding odd-numbered data points. Each is said to have been thinned by a factor f=2 since there are 1/f as many data points as in the original record. The maximum of each record is denoted by $\max_1(2)$ and $\max_2(2)$. In general, thinning by a factor f results in f records with corresponding maxima $\max_1(f)$, $\max_2(f)$, ... , $\max_f(f)$. Note that one of these maxima will necessarily coincide with the maximum of the original record, max(1). The ratios

$$R_i(f) = \frac{\max_i(f)}{\max(1)} \qquad 7.1$$

(for i=1,...,f) measure the degree to which the maximum max(1) is under-estimated due to thinning by the factor f. For positive data, $0 \le R_i(f) \le 1$.

This analysis system, of utilizing all the possible thinned records obtainable for a given thinning factor f, yields a distribution of ratios which is representative of the range of ratios to be expected. These distributions are depicted in Figure 7.4 for various values of f for the H-ion and flow records. Each box and whisker symbol indicates the mean (cross), median (horizontal bar), lower and upper quartiles (bottom and top of box) and the maximum and minimum (top and bottom whisker ends); these summarize the shape of the ratio distribution for each value of f. The maximum value for f of 672 corresponds to a sampling interval of one week. For flow, average ratios decrease rapidly and the distributions become more skewed as f increases. For H-ion concentration, the average

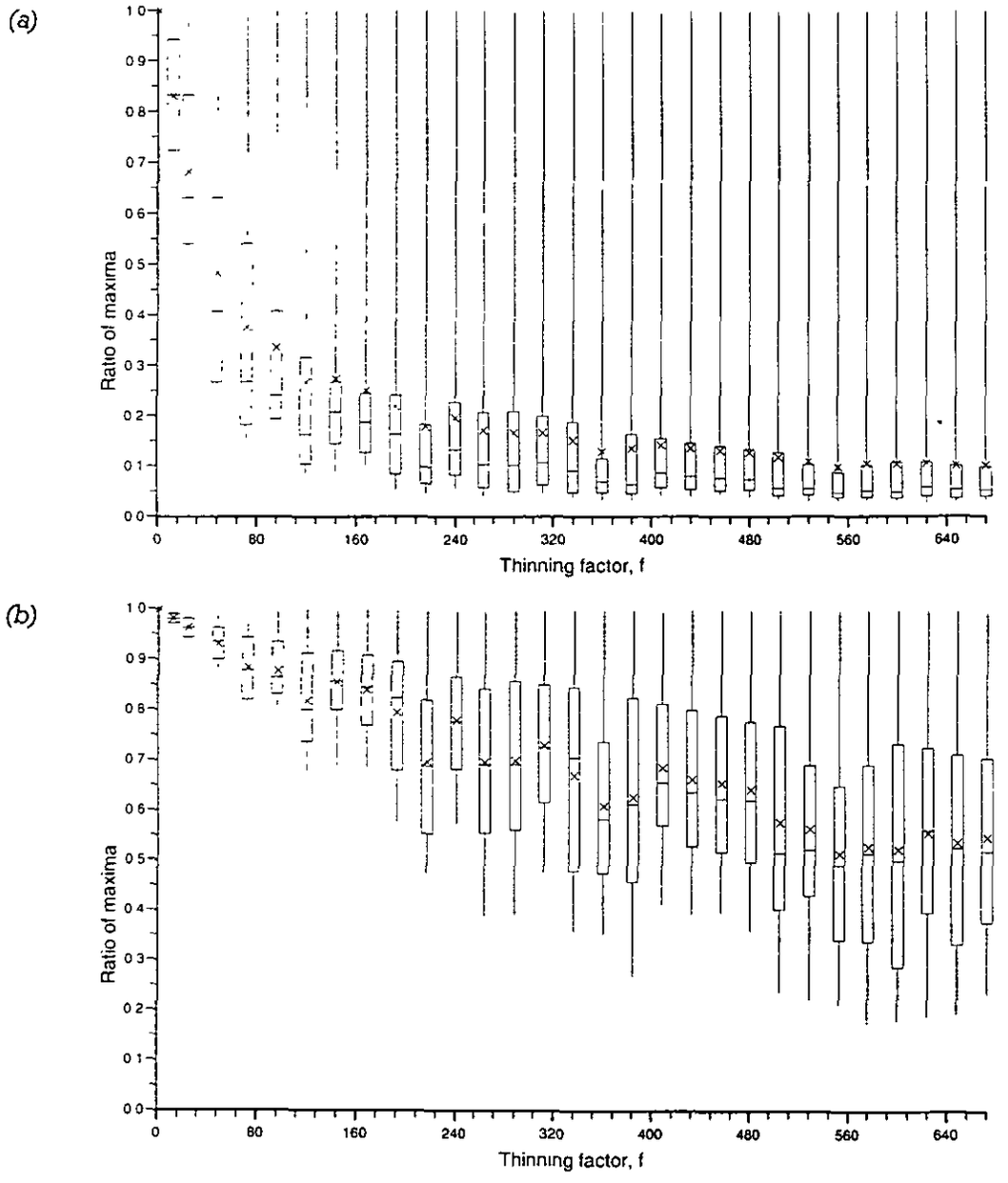


Figure 7.4 *Box and whisker plots illustrating distribution for ratios $R_i(f)$ for (a) flow and (b) H-ion concentration*

ratios decrease more steadily and the distributions remain relatively symmetrical.

The median values, denoted by $R_m(f)$, are examined with respect to their variation with f. The median is favoured over the mean since, for distributions with long thin tails — such as those for the flow ratios — it is a more robust measure of location. Various plots were used to uncover an appropriate relationship for both flow and H-ion concentration. Linearity in semi-log plots suggests the exponential relationship

$$R_m(f) = \exp[\alpha(f-1)] \qquad 7.2$$

where $\alpha<0$ is a parameter. Only points for which $R_m(f)\geq 0.5$ are examined; otherwise a correction factor greater than 2 would be required, suggesting that the data are simply inadequate for the purpose of estimating extremes, and that measurement practice should be revised.

The semi-log plots, fitted with least-squares regression lines obtained by the regression of $\ln(R_m)$ on f, are shown in Figure 7.5. The regression fit for flow is good ($r^2=0.987$, where r is the correlation coefficient) and extends to a thinning factor f=36 (corresponding to a measurement interval of nine hours) before the median ratio falls below 0.5. The regression coefficient, which corresponds to the parameter α in equation 7.2, is estimated as $\alpha=-0.0179$. For H-ion concentration the fit is less good ($r^2=0.918$) but extends as far as f=528 (corresponding to a measurement interval of 5.5 days). The regression yields $\alpha=-0.0012$. Both graphs appear to display a slight cyclic pattern of residuals about the regression line.

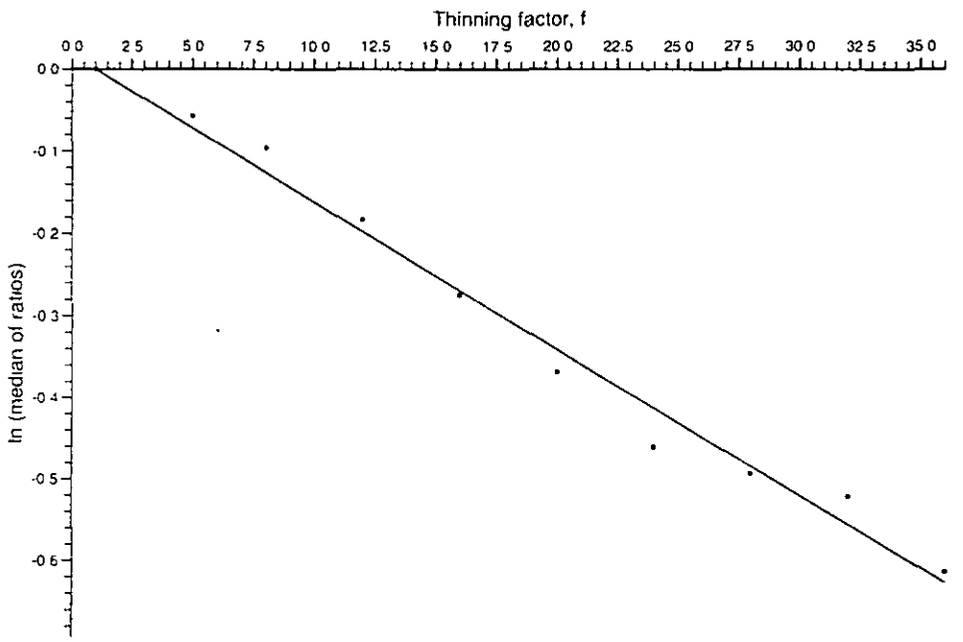
These results might be applied as follows. Given instantaneous data for flow or H-ion concentration measured at regular intervals of T>15 minutes, the maximum, m_T, is obtained. The maximum as if measured from 15-minute data, m_{15}, can then be estimated using the equation

$$m_{15} = \frac{m_T}{R_m(f)} \qquad 7.3$$

where the denominator, $R_m(f)$, is calculated from equation 7.2 with f=T/15 and the appropriate value for α(-0.0179 for flow and -0.0012 for H-ion concentration) . If $R_m(f)<0.5$ then the practice of applying a correction factor in this way is not to be recommended.

The above formulae and method have not been generalized to other catchments or hydrological variables and the sensitivity of the results to period length, data resolution etc. have not been examined. It is therefore suggested that a more comprehensive study of this problem may be required. Nevertheless, the relationships found are promising in their limited context and may provide a starting point for a more exhaustive investigation. The "data thinning" problem may be amenable to other analytical techniques.

(a)



(b)

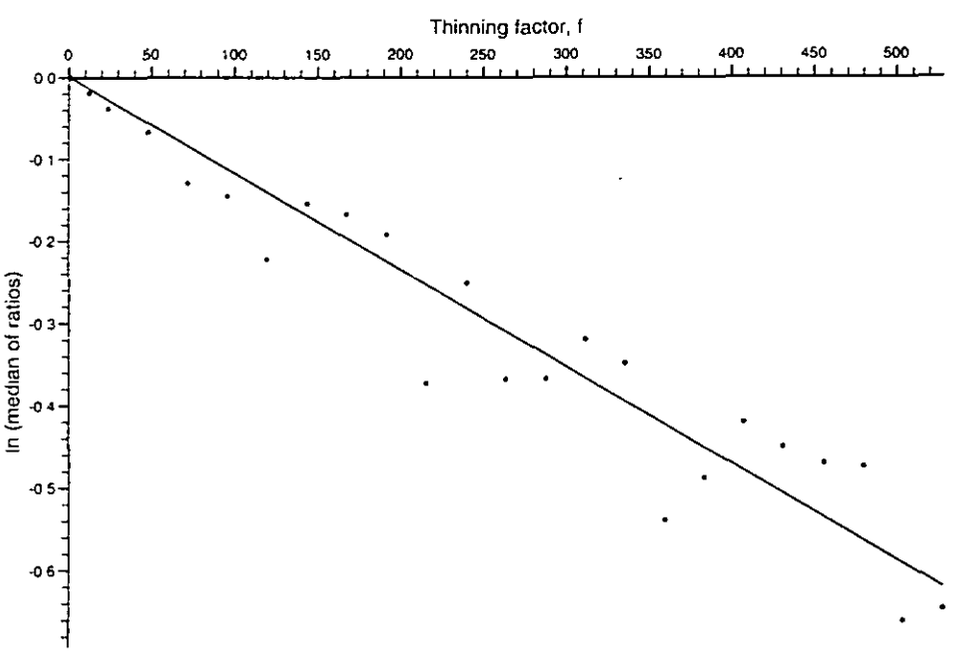


Figure 7.5 *ln(R_m(f)) against f, with regression lines, for (a) flow and (b) H-ion concentration*

8 Recommendations

The measurement of an environmental process often produces data which are discretized. Using such data to analyse extremes will, in general, under-appreciate the true extremes of the process; estimates of extremes based upon the data therefore require correction.

Where applicable, the conclusions of each chapter are consolidated into practical recommendations for making corrections to mean period maxima (or minima). Moments of higher order than the mean do not, in general, require correction.

8.1 Correction factors for point rainfall

The following advice relates to rainfall time-series data for a single gauge, measured as accumulations over a fixed time interval.

Maxima of duration one to 32 hours, measured from sub-daily data
For data in this category, the full correction model as described in Section 3.5 may be applied. Using parameter values of a=0.16 and b=0.57, a number of types of correction can be performed:

- converting fixed to sliding maxima using the multiplier $\rho(D)$ for $2 \leq D \leq 32$;
- converting fixed to true maxima using the multiplier ρ^* for $1 \leq D \leq 32$;
- converting sliding to true maxima using the multiplier $\alpha(D)$ for $2 \leq D \leq 32$;
- converting sliding to sliding maxima, from resolution τ to resolution $\tau' < \tau$, using the multiplier $\beta(D)$ for $2 \leq D \leq 32$ and $1 \leq \tau < D$.

Where

$$\rho(D) = 1 + a[1 - \exp\{-b(D-1)\}] \qquad 8.1$$

$$\rho^* = \lim_{D \to \infty} \rho(D) = 1 + a \qquad 8.2$$

$$\alpha(D) = \frac{\rho^*}{\rho(D)} \qquad 8.3$$

$$\beta(D) = \frac{\rho(\lambda D)}{\rho(D)} \quad , \quad \lambda = \tau/\tau' \qquad 8.4$$

For a full explanation refer to the model description in Section 3.5 which is supplemented by numerical examples.

The special case of daily rainfall
Special guidelines apply when converting fixed to true maxima for daily rainfall. As well as the correction factor $\rho^*=1.16$, as suggested above, correction factors of 1.15 and 1.17 may also be employed: the former for sites with rainfall regimes which tend to generate concentrated events of short duration (such as tropical storms), the latter for sites with rainfall regimes tending to generate longer events (such as those produced by large frontal systems).

Maxima for longer durations, measured from daily data
There is no evidence to suggest that the parameter *a* is, on average, any different for daily data than for hourly data. The parameter *b* does, however, appear to be typically higher for daily than for hourly data. Thus, for rainfall durations between 32 hours and 32 days, it is recommended that the full correction model is used with parameter values a=0.16 and b=0.36.

8.2 Correction factors for other variables

The correction factor ρ^* was investigated for other environmental variables. It was found that the roughness (intermittency and erraticness) of the variable, as measured by the effective fractal dimension d_E, bears a relation to the magnitude of ρ^* for that variable as indicated by Figure 5.3. These results can be utilized as follows:

- For wind speed maxima of duration one to 64 hours, use the correction factor $\rho^*=1.10$ for converting fixed to true maxima.
- For air temperature maxima of duration one to 64 hours, first estimate the lower 1% quantile, L, for the mean *hourly* air temperature at the site (see below). Then calculate the true maximum T from the fixed maximum F using

$$T = \rho^* F - (\rho^* - 1)L \qquad \textbf{8.5}$$

where $\rho^* = 1.04$.

- For air temperature minima of duration one to 64 hours, use the same correction procedure as for air temperature maxima but replace L by the *upper* 1% quantile, U.

Since it is the absence of hourly data that necessitates the above correction procedure, the quantiles L and U cannot be calculated from data directly. However, they may be estimated from air temperature statistics for the site in the following way:

- Find the long-term averages for daily mean, daily maximum and daily minimum air temperatures. Denoting these by μ, M and m respectively, estimate L and U using

$$L = \mu + 3.4(m - \mu)$$
$$U = \mu + 3.4(M - \mu) \qquad \textbf{8.6}$$

The coefficient 3.4 is based upon observation at the study site (Eskdalemuir). The statistics μ, M and m are available for various sites in the *World Survey of Climatology* series (World Meteorological Organization, 1970).

Regarding other environmental variables at a site, limited use can be made of Figure 5.3 to indicate an approximate value for the correction factor ρ^*. Some hourly data for the variable are required to enable an estimate to be made of the effective fractal dimension d_E of the variable at the hourly resolution; if transferred from another site, the data used should at least exhibit broadly similar characteristics to those expected at the subject site. The relationship between d_E and ρ^* indicated in the figure can then be assessed by eye to infer a value for ρ^*. Finally, an estimate of the 1% quantile for hourly data should be obtained and used in equation 8.5 above to perform the conversion.

8.3 Further research

Further research is warranted into the correction of extremes observed from "spot sampled" measurements. The approach detailed in Section 7.3 could be considered as a candidate although other methods might be appropriate.

A more extensive investigation into correction factors for catchment average rainfall, as distinct from point rainfall, may also be justified.

Acknowledgements

The Allowance for Discretization in Hydrological and Environmental Risk Estimation (ADHERE) project was funded by the Terrestrial and Freshwater Sciences Directorate of the UK Natural Environment Research Council. We gratefully acknowledge Andrew Coyle, Bruce Kelbe, Christine Simmonds and Lisa Stewart for their important contributions to the project. We are also grateful for refereed comments on a previously published paper, one of which prompted the analysis of the concatenated rainfall records.

The cooperation of many organizations in making high quality data series available to the project is gratefully acknowledged. In particular we thank the UK Met. Office for supplying hourly data from Eskdalemuir Observatory and allowing the use of a limited number of long-term daily rainfall records, the Computing Centre for Water Research at the University of Natal for supplying daily rainfall data, and the Australian Bureau of Meteorology for supplying extensive hourly and daily rainfall records. Some specialist data series were supplied by colleagues, sister research institutes in the UK, the Norwegian Meteorological Institute, and Professor David Stow (University of Auckland, New Zealand). These were of considerable benefit in formulating the research, although only results for the stream water quality data (Alice Robson) and the tide residual data (Proudman Oceanographic Laboratory) are presented here. Finally, we thank the Ministry of Agriculture, Fisheries and Food, the Department of the Environment, and the Office of Science and Technology for funding projects for which experimental data series were originally gathered.

References

Barnett, V. 1974. *Elements of sampling theory.* English University Press, London. Chapter 3.

Constantine, A.J. & Hall, P. 1994. Characterizing surface smoothness via estimation of effective fractal dimension. *J. R. Statist. Soc. B*, **56**, 97-113.

Coyle, A.J., Kelbe, B.E., Reed, D.W. & Stewart, E.J. 1991. A temporal look at hydrological extremes. *Proc. British Hydrological Society 3rd Nat. Hydrol. Symp., Southampton.* British Hydrological Society, London. pp. 6.51-6.59.

Dwyer, I.J., & Reed, D.W. 1994. Effective fractal dimension and corrections to the mean of annual maxima. *J. Hydrol.* **157**, 13-34.

Harishara, P.S. & Tripathi, N. 1973. Relationship of the clock-hour to 60-min and the observational day to 1440-min rainfall. *Indian J. Met. Geophys.* **24**, 279-282.

Hershfield, D.M. & Wilson, W.T. 1958. Generalizing of rainfall-intensity-frequency data. *IUGG/IAHS publication no. 43*, 499-506.

Hosking, J.R.M. 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J. R. Statist. Soc. B*, **52**, 105-124.

Kerr, R.L., McGinnis, D.F., Reich, B.M. & Rachford, T.M. 1970. Analysis of rainfall-duration-frequency for Pennsylvania. Institute for Research on Land and Water Resources, The Pennsylvania State University, research publication 70.

Mood, A.M., Graybill, F.A. & Boes, D.C. 1974. *Introduction to the theory of statistics.* McGraw-Hill International Book Company, Singapore. pp 432-435.

NAG. 1991. *NAG Fortran library, Mark 15.* The Numerical Algorithms Group Ltd., vol. 8.

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

Rajagopalan, B. & Tarboton D.G. 1993. Understanding complexity in the structure of rainfall. *Fractals*, **1**, 606-616.

Robson, A.J. 1993. The use of continuous measurement in understanding and modelling the hydrochemistry of the uplands. PhD thesis. University of Lancaster.

Van Montfort, M.A.J. 1990. Sliding maxima. *J. Hydrol.* **118**, 77-85.

Webster, R. & Oliver, M.A. 1990. *Statistical methods in soil and land resource survey.* Oxford University Press. Chapter 12.

Weiss, L.L. 1964. Ratio of true to fixed-interval maximum rainfall. *J. Hydraul. Div. Proc. ASCE*, **90**, 77-82.

World Meteorological Organization. 1970. *World survey of climatology.* Elsevier Publishing Company, Amsterdam.

Appendix A: Comments on Weiss (1964)

A.1 Weiss's formulation

Weiss's theoretical treatment of the discretization effect resembles the following formulation.

Let the true maximum, T, occur in a time interval XY which straddles the two fixed intervals AB and BC depicted in Figure A.1. Without loss of generality, let the interval width be of unit length ($|XY|=|AB|=|BC|=1$). Denote by x, the proportion of the interval XY overlapping BC.

The following assumptions are made (only (a) and (b) are alluded to by Weiss):

(a) on average, the rainfall profile within XY is uniform;

(b) The interval XY is randomly placed within AC and so the probability distribution of x is uniform;

(c) The fixed maximum, F, arises from one of the intervals AB or BC (whichever contains the majority of the rainfall in XY);

(d) No rainfall occurs in AC other than that in XY.

Given these assumptions, Weiss argues that the expected value of x for $x \geq \frac{1}{2}$ is

$$\int_{\frac{1}{2}}^{1} x\,dx = 3/8 \qquad \text{A.1}$$

For $x<\frac{1}{2}$, the larger proportion of XY is in the fixed interval AB. Since it is only the largest proportion that is of interest in calculating the fixed maximum, Weiss forces the expected value of x for $x<\frac{1}{2}$ to equal $\frac{1}{2}$, concluding that the expected value over the entire range is therefore 1/2 + 3/8 = 7/8. This produces an average ratio of the true to fixed maximum of 8/7 (i.e. 1.143).

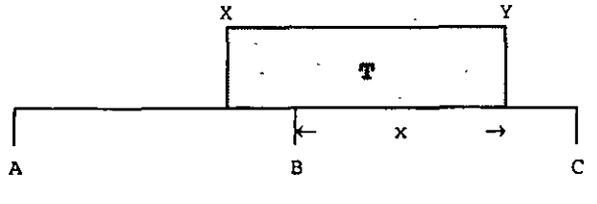


Figure A.1 *Position of true maximum T relative to the fixed intervals AB and BC*

These theoretical arguments are clearly flawed; a proper statistical formulation is required.

A.2 A proper formulation

Let the random variable X=x, as defined above, have the probability density function f(x). Defining the random variable Y as the largest proportion of T overlapping one of the fixed intervals, assumption (a) gives

$$Y = y(x) = \begin{cases} 1-x & \text{for } x<1/2 \\ x & \text{for } x \geq 1/2 \end{cases} \qquad \text{A.2}$$

The expected value of Y is

$$E(Y) = \int_0^1 y(x)f(x)dx \qquad \text{A.3}$$

Assumption (b) gives f(x)=1, whereupon

$$E(Y) = \int_0^{1/2} (1-x)dx + \int_{1/2}^1 x\,dx \qquad \text{A.4}$$

$$= 3/4$$

Assumptions (c) and (d) then give F=¾T. Thus, Weiss should have arrived at the correction factor T/F = 4/3 = 1.33, much higher than actually suggested.

A.3 Examination of the assumptions

Let the assumptions (a) - (d) be examined in turn.

Assumption (a)
If a large number of high intensity rainfall profiles are literally averaged, the result may well approximate uniformity. However, it is the *typical* profile that is important in the determination of the correction factor and this is unlikely to be uniform. An alternative assumption is that of a symmetric triangular distribution which, after a little simple geometry, renders

$$Y = y(x) = \begin{cases} 1-2x^2 & \text{for } x<1/2 \\ 1-2(1-x)^2 & \text{for } x \geq 1/2 \end{cases} \qquad \text{A.5}$$

Assumption (b) then gives

$$E(Y) = \int_0^{1/2} (1-2x^2)dx + \int_{1/2}^{1} [1-2(1-x)^2]dx$$

$$= \left[x-\frac{2}{3}x^3\right]_0^{1/2} + \left[x+\frac{2}{3}(1-x)^3\right]_{1/2}^{1} \qquad \textbf{A.6}$$

$$= 5/6$$

Under this alternative model, assumptions (c) and (d) imply F=(5/6)T which renders a correction factor of T/F = 6/5 = 1.20. This is still higher than Weiss's suggestion but much closer than 1.33 to empirical results. Other forms for the distribution of rainfall within XY will result in different correction factors.

Assumption (b)
This demands that no trends or cycles exist in the rainfall over time-scales comparable to the duration of the maxima. This is reasonable for the UK but not for other locations, such as the tropics, where rainfall experiences strong diurnal cycles.

Assumption (c)
This assumption ignores the possibility of the fixed maximum arising from a separate event which is better synchronized with the fixed intervals (as illustrated in the introduction by Fig. 1.1). This is a likely occurrence when the true maximum is divided almost equally between fixed intervals (i.e. for Y close to ½). Therefore, less weight should be given to small values of Y, whereupon E(Y) is increased and the theoretical correction factor, T/F, further reduced.

Assumption (d)
The presence of rainfall in AX or YC will tend to increase the fixed maximum F and thus also reduce the correction factor.

Appendix B: Sliding intervals that straddle borders

To ensure that events occurring across the border of two periods are not missed, a convention is adopted whereby a sliding interval which is split across the border is assigned to the period in which it is mostly contained; if it is equally split between them then it is assigned to the first period (Fig. B.1).

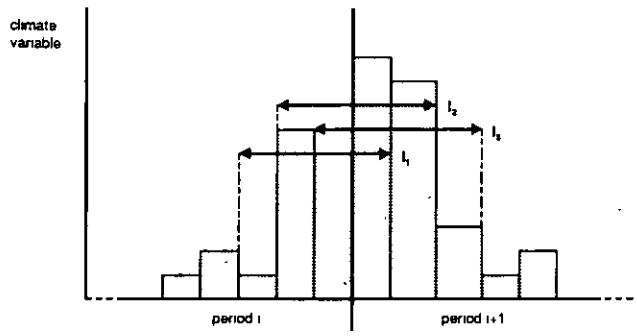


Figure B.1 *Border intervals for duration D=4. In searching for the sliding maximum for each period, the accumulations I_1 and I_2 are assigned to sample i wheras I_3 is assigned to sample i+1*

Equally, a border event must not be 'counted twice' in the sense of it contributing to the sliding maximum for both periods. Thus, if the sliding maximum for the first period straddles the border then the sliding intervals assigned to the second period must not include those which overlap with this maximum (Fig. B.2).

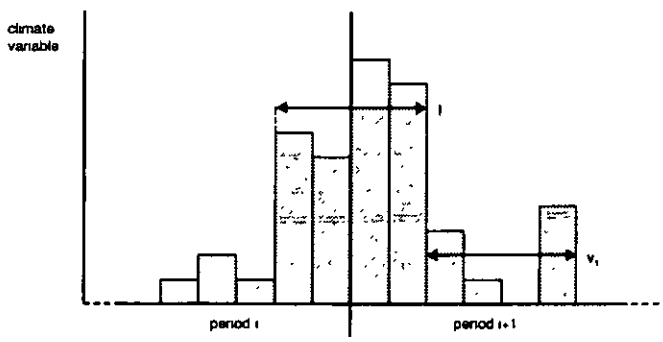


Figure B.2 *If, for duration D=4, accumulation I is the sliding maximum for period i, then the first sliding accumulation to be assigned to period i+1 is v_1 since it does not overlap with accumulation I.*

Appendix C: Data summary

Each record is found under its appropriate *site-name*, these being arranged alphabetically. The exception is for sites where only daily rainfall records are detailed - these are grouped together at the end under the title *Other daily rainfall records*.

For each site, the records are divided according to resolution and variable-type; for example, *hourly rainfall, daily rainfall, hourly wind-speed* etc. Various details are given for each record as explained below in the key. Where a particular detail is not calculated, a period (.) is entered. Each record consists of 16384 (2^{14}) data values.

Key

record: where more than one record exists under a particular heading, a distinguishing label has been assigned to each (as referred to in the main text).

dates: the time-period covered by the record. Format: DDMMYY for the case where YY refers to the 20th Century, and DDMM8YY for the 19th Century.

mean: arithmetic mean of all data values. Rainfall is given in millimetres, wind speed in knots (1 knot ≈0.514 $m.s^{-1}$ and air temperature in degrees Celsius.

cv and **skew:** respectively, the coefficient of variation and skewness of all data values.

$d_s(\pm)$: effective fractal dimension with error pertaining to the 95% confidence interval derived from the regression analysis.

a(±): value of parameter *a* in the correction model with error pertaining to the 95% confidence interval derived from the regression analysis.

b, a', D_{max}: respectively, the values of parameters *b*, *a'*, and D_{max} in the correction model.

L1, L2, L3: L-mean and L-moment ratios (L-cv and L-skewness respectively) of period maxima/minima for duration D=1. See Hosking (1990). L1 is expressed in the same units as **mean**.

Abbreviations

Alt. - altitude
AAR - average annual rainfall
SAAR - standard period (1941-70) average annual rainfall

Site-name: Alwen at Druid (catchment 67006)
Location: North Wales
Particulars: Catchment area 185km², SAAR 1260mm; gauge 59 centrally located , Alt. 350m, SAAR 1250mm
Climate: Cool temperate; mainly frontal rainfall

Record & dates	*mean*	*cv*	*skew*	$d_s(\pm)$	*a(±)*	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
15-minute catchment rainfall											
. 010772-181272	.028	4.39	7.5	1.850(±0.051)	0.190(±0.017)	0.111	.	64	0.3	.491	.212
15-minute point rainfall (gauge 59)											
. 010772-181272	.029	4.94	11.7	1.817(±0.074)	0.169(±0.033)	.	.	32	0.5	.491	.421

Site-name: Brisbane
Location: East coast of Australia
Particulars: 168 missing values among hourly series; Alt. 42m, AAR 1092mm
Climate: Warm temperate / tropical; susceptible to tropical cyclones in summer season

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	$a(\pm)$	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 120216-251217	0.12	8.47	19.9	1.883(±0.083)	0.139(±0.016)	0.723	0.138	32	10.1	.517	.365
B 040434-150236	0.11	9.49	22.3	1.924(±0.047)	0.151(±0.009)	0.657	0.152	32	13.8	.452	.331
C 190340-300142	0.08	12.01	42.1	1.915(±0.050)	0.124(±0.016)	.	0.128	32	9.2	.585	.503
D 020847-140649	0.13	8.08	15.5	1.862(±0.070)	0.147(±0.015)	.	0.154	32	11.6	.411	.144
E 061153-100955	0.16	6.99	15.0	1.853(±0.059)	0.162(±0.022)	0.208	0.162	32	12.3	.427	.335
Daily rainfall											
1 0101887-111131	2.98	3.65	11.1	.	0.186(±0.104)	0.251	.	64	120.5	.276	.445
2 011231-091076	3.11	3.69	8.6	.	0.148(±0.135)	1.157	.	32	133.6	.222	.278

Site-name: Eskdalemuir
Location: Southern Uplands, Scotland
Particulars: Tilting syphon raingauge, Alt. 250m; AAR 1527mm; cup anemometer
Climate: Cool temperate; frontal rainfall with autumn/winter bias

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	$a(\pm)$	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 010170-141171	0.15	3.63	6.6	1.769(±0.043)	0.169(±0.012)	0.388	0.170	32	5.2	.252	.181
B 141171-270973	0.14	3.74	6.6	1.733(±0.056)	0.160(±0.014)	0.563	0.159	32	4.9	.243	.057
C 270973-100875	0.16	3.52	6.5	1.770(±0.062)	0.162(±0.011)	0.841	0.163	32	5.2	.284	.285
D 110875-230677	0.15	3.53	6.0	1.779(±0.058)	0.180(±0.015)	0.402	0.180	32	5.2	.207	.105
E 230677-070579	0.18	3.58	7.3	1.741(±0.102)	0.161(±0.011)	0.700	0.161	32	5.5	.272	.220
F 070579-190381	0.19	3.32	5.9	1.761(±0.043)	0.177(±0.018)	0.330	0.175	32	5.4	.215	.112
G 200381-310183	0.21	3.41	6.0	1.794(±0.027)	0.177(±0.018)	0.330	0.175	32	6.6	.278	.087
H 310183-141284	0.16	3.73	7.6	1.764(±0.052)	0.158(±0.011)	1.564	0.158	32	5.5	.307	.235
2-hourly rainfall											
K 010170-270973	0.29	3.32	6.0	.	0.165(±0.014)	0.526	.	32	9.7	.190	.032
L 270973-230677	0.32	3.18	5.6	.	0.163(±0.013)	0.817	.	32	9.0	.179	.074
M 230677-190381	0.37	3.13	6.2	.	0.166(±0.020)	0.423	.	32	10.3	.240	.173
N 200381-141284	0.37	3.24	6.1	.	0.158(±0.012)	1.309	.	32	10.8	.216	.004
4-hourly rainfall											
X 010170-230677	0.61	2.87	5.0	.	0.155(±0.011)	0.380	.	32	15.6	.155	.033
Y 230677-141284	0.74	2.83	5.2	.	0.160(±0.013)	0.486	.	32	18.1	.155	.083
8-hourly rainfall											
. 010170-141284	1.35	2.44	4.5	.	0.172(±0.015)	0.233	.	32	28.6	.163	.198
Daily rainfall											
. 010111-101155	4.24	1.70	2.9	.	0.129(±0.007)	0.730	.	32	53.1	.138	.245
Hourly wind speed											
A 010170-141171	8.89	0.75	0.9	1.490(±0.034)	0.105(±0.007)	0.072	.	64	29.1	.126	.002
B 141171-270973	8.69	0.77	0.8	1.514(±0.034)	0.098(±0.007)	0.066	.	64	28.3	.121	-.037
C 270973-100875	9.70	0.74	1.0	1 494(±0.022)	0.138(±0.024)	0.032	.	64	30.6	.139	.187
D 110875-230677	8.77	0.76	1.0	1.504(±0.027)	0.107(±0.008)	0.063	.	64	29.1	.156	.070
E 230677-070579	9.41	0.74	0.9	1.507(±0.026)	0.094(±0.006)	0.072	.	64	30.3	.131	.083
F 070579-190381	8.42	0.74	0.9	1.499(±0.027)	0.089(±0.005)	0.096	.	64	27.2	.125	.196
Hourly air temperature	- maxima										
	- minima										
B 141171-270973	7.19	0.74	0.2	1.251(±0.015)	1.042(±0.004)	0.211	.	64	20.5	.160	.104
					1.040(±0.003)	0.210	.	64	-23.1	-.104	-.118
E 230677-070579	6.48	0.92	-0.1	1.264(±0.007)	1.032(±0.002)	0.193	.	64	22.1	.159	.099
					1.042(±0.002)	0.200	.	64	-23.5	-.116	-.089
F 070579-190381	6.94	0.76	-0.1	1.243(±0.015)	1.036(±0.003)	0.236	.	64	20.2	.162	.064
					1.044(±0.004)	0.288	.	64	-22.2	-.098	.056
G 200381-310183	7.38	0.78	-0.3	1.320(±0.004)	1.031(±0.002)	0.177	.	64	24.8	.122	.094
					1.043(±0.003)	0.178	.	64	-23.6	-.113	-.144
H 310183-141284	7.88	0.74	0.3	1.217(±0.025)	1.041(±0.003)	0.135	.	64	21.1	.179	.143
					1.040(±0.002)	0.197	.	64	-25.1	-.086	-.044
I 141284-261086	6.35	0.89	-0.1	1.253(±0.015)	1.038(±0.003)	0.217	.	64	22.2	.162	.018
					1.040(±0.004)	0.321	.	64	-22.6	-.124	-.099
J 271086-080988	6.79	0.79	0.0	1.237(±0.020)	1.034(±0.002)	0.174	.	64	20.5	.152	.122
					1.041(±0.003)	0.178	.	64	-21.0	-.109	.021

Site-name: Leeming
Location: North Yorkshire, east of Pennines
Particulars: Alt. 32m; SAAR 611mm; tilting syphon raingauge
Climate: Cool temperate; frontal rainfall predominates in winter, convective in summer

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	$a(\pm)$	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 010178-141179	0.07	4.65	8.7	1.772(±0.044)	0.171(±0.018)	0.717	0.170	32	3.6	.276	.221
B 141179-260981	0.08	5.08	10.8	1.812(±0.067)	0.168(±0.015)	0.796	0.168	32	4.0	.361	.356
C 270981-100883	0.07	4.69	8.2	1.783(±0.061)	0.126(±0.018)	0.865	0.126	32	3.2	.282	.185
D 110883-230685	0.07	5.31	12.4	1.766(±0.068)	0.155(±0.014)	1.038	0.156	32	3.8	.346	.298
E 230685-060587	0.07	5.14	10.7	1.766(±0.068)	0.151(±0.016)	0.387	0.152	32	3.8	.311	.290

Site-name: Melbourne
Location: South east tip of Australia
Particulars: 24 missing values in record B; Alt. 35m, AAR 691mm
Climate: Warm temperate; exposed to cyclonic fronts all year with only slight seasonal variation

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	$a(\pm)$	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 211253-031155	0.09	7.35	17.9	1.834(±0.077)	0.175(±0.016)	0.567	0.172	32	7.3	.431	.371
B 241162-061064	0.08	6.14	12.5	1.793(±0.089)	0.157(±0.014)	0.510	0.156	64	5.0	.368	.220
C 070544-200346	0.07	6.57	17.6	1.797(±0.044)	0.164(±0.018)	1.198	0.162	32	4.5	.370	.446
D 090952-230754	0.08	6.38	13.5	1.785(±0.060)	0.135(±0.016)	.	0.135	32	5.5	.346	.324
E 100757-230559	0.08	5.83	13.7	1.857(±0.040)	0.150(±0.010)	0.973	0.152	32	5.5	.381	.305
Daily rainfall											
1 0104855-080200	1.75	2.75	5.7	.	0.161(±0.083)	0.404	.	64	48.2	.183	.126
2 010300-080145	1.76	2.82	6.2	.	0.165(±0.098)	0.175	.	64	50.4	.185	.119
3 010245-111289	1.79	2.85	6.5	.	0.165(±0.074)	2.039	.	64	54.8	.212	.235

Site-name: Ringway
Location: Manchester airport, North-West England
Particulars: Alt. 75m; SAAR 819mm; tilting syphon raingauge
Climate: Cool:temperate; frontal rainfall with autumn/winter bias

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	$a(\pm)$	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 010176-131177	0.09	4.49	8.5	1.818(±0.062)	0.159(±0.019)	0.432	0.160	32	4.1	.234	.212
B 131177-260979	0.08	4.41	10.7	1.838(±0.040)	0.153(±0.011)	0.643	0.154	32	3.7	.325	.322
C 270979-090881	0.10	5.49	41.0	1.828(±0.062)	0.154(±0.013)	0.619	0.156	32	5.7	.453	.534
D 100881-230683	0.10	4.45	8.6	1.842(±0.070)	0.148(±0.012)	1.198	0.150	32	4.9	.294	.140
E 230683-050585	0.09	4.84	14.2	1.854(±0.052)	0.166(±0.013)	0.631	0.168	32	4.2	.349	.336
F 060585-190387	0.09	4.66	9.0	1.831(±0.056)	0.169(±0.013)	0.512	0.170	32	4.1	.313	.146
G 200387-300189	0.11	4.45	9.4	1.801(±0.062)	0.162(±0.014)	0.285	0.160	32	5.2	.257	.291
H 300189-131290	0.09	4.67	10.2	1.827(±0.062)	0.162(±0.014)	0.285	0.160	32	4.6	.279	.272
2-hourly rainfall											
K 010170-270973	0.17	3.85	7.1	.	0.158(±0.015)	0.758	.	32	6.7	.207	.329
L 270973-230677	0.21	4.31	23.5	.	0.181(±0.012)	0.527	.	32	9.2	.347	.547
M 230677-190381	0.17	4.08	7.9	.	0.172(±0.016)	0.267	.	32	7.1	.235	.272
N 200381-141284	0.19	3.97	8.6	.	0.152(±0.012)	0.402	.	32	8.1	.250	.379
4-hourly rainfall											
X 010170-230677	0.38	3.50	14.8	.	0.162(±0.029)	0.540	.	32	12.7	.299	.464
Y 230677-141284	0.37	3.48	7.4	.	0.153(±0.020)	0.587	.	32	13.0	.238	.347
8-hourly rainfall											
. 010170-141284	0.74	2.85	7.7	.	0.166(±0.012)	0.691	.	32	21.3	.233	.432

Site-name: Sydney
Location: South-east coast of Australia
Particulars: 133 missing values among hourly records; Alt. 42m, AAR 1205mm
Climate: Warm temperate; predominantly frontal rainfall, slightly higher in autumn

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	*a(±)*	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly rainfall											
A 200643-020545	0.12	8.35	28.0	1.847(±0.078)	0.144(±0.018)	0.771	0.145	32	11.3	.452	.417
B 020947-150749	0.14	6.40	16.8	1.828(±0.058)	0.160(±0.012)	0.501	0.161	64	9.4	.436	.269
C 300552-120454	0.14	6.55	13.1	1.807(±0.076)	0.158(±0.014)	0.612	0.158	32	9.6	.451	.267
D 030561-160363	0.16	7.13	22.3	1.830(±0.113)	0.176(±0.019)	0.261	0.178	32	10.3	.483	.417
E 110663-230465	0.12	7.29	15.5	1.863(±0.112)	0.164(±0.019)	0.529	0.164	32	8.9	.426	.302
F 250465-080367	0.14	6.44	12.4	1.803(±0.028)	0.169(±0.021)	0.479	0.169	32	8.7	.413	.176
Daily rainfall											
1 0107858-110503	3.36	3.32	7.1	.	0.167(±0.082)	0.456	.	64	125.8	.188	.193
2 010603-090448	2.95	3.30	7.7	.	0.166(±0.081)	0.323	.	64	114.2	.206	.306

Site-name: Walton-on-the-Naze
Location: Essex coast, UK
Particulars: Data supplied by Proudman Oceanographic Laboratory, Bidston
Climate: Cool temperate

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	*a(±)*	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Hourly tide residual (metres)			- maxima								
			- minima								
. 010169-141170	-0.73	-9.43	0.5	1.371(±0.100)	0.061(±0.008)	0.172	.	32	40.4	.154	.172
					0.047(±0.005)	0.172	.	32	-39.9	-.106	-.227

Other daily rainfall records

Record & dates	*mean*	*cv*	*skew*	$d_g(\pm)$	*a(±)*	*b*	*a'*	D_{max}	*L1*	*L2*	*L3*
Alwen Reservoir, Dee, Wales: Alt. 362m, SAAR 1282mm											
. 010125-101169	3.60	1.64	2.9	.	0.148(±0.008)	0.648	.	32	42.8	.120	.275
Cape Town, South Africa: Alt. 40m, AAR 626mm											
. 011004-100849	1.61	2.93	5.6	.	0.185(±0.014)	0.353	.	32	44.9	.224	.328
Creech Grange, Dorset, UK: Alt. 69m, SAAR 947mm											
. 010130-101174	2.57	2.09	3.8	.	0.179(±0.011)	0.349	.	32	45.2	.172	.209
Durban, South Africa: Alt. 91m, AAR 1020mm											
. 010632-100477	2.80	3.46	8.0	.	0.176(±0.011)	0.106	.	64	111.1	.238	.260
Edinburgh, Scotland: Alt. 134m, SAAR 673mm											
. 010108-091152	1.86	2.26	5.2	.	0.167(±0.013)	0.512	.	32	40.3	.211	.156
Etton on the Welland, UK: Alt. 11m, SAAR 536mm											
. 010108-091152	1.38	2.41	4.7	.	0.174(±0.010)	0.190	.	48	29.5	.196	.294
Hobart, Tasmania: Alt. 54m, AAR 668mm											
1 0101895-111139	1.68	2.87	7.1	.	0.142(±0.010)	0.757	.	32	54.0	.204	.227
2 011239-091084	1.71	3.09	8.8	.	0.164(±0.008)	0.159	.	64	61.8	.250	.362
Johannesburg, South Africa: Alt. 1737m, AAR 844mm											
. 0109893-120738	2.31	3.19	7.7	.	0.174(±0.013)	0.292	.	32	75.2	.250	.425
Perth, Australia: Alt. 60m, AAR 889mm											
1 0101880-101124	2.34	2.65	4.4	.	0.167(±0.009)	0.451	.	32	61.8	.250	.362
2 011224-101069	2.44	2.66	4.4	.	0.175(±0.012)	0.315	.	32	56.2	.160	.098
Richmond, South Africa: Alt. 1417m, AAR 323mm											
. 010540-100385	0.85	4.56	7.4	.	0.154(±0.012)	0.442	.	32	45.1	.195	.244
Upington, South Africa: Alt. 794m, AAR 151mm											
. 011139-090984	0.41	6.72	11.2	.	0.163(±0.015)	0.527	.	32	29.9	.383	.058

Appendix D: L-moment ratio diagrams for R(D)

Graphs showing L-CV and L-skew for fixed maxima (solid lines) and sliding maxima (dotted lines) for various data records.

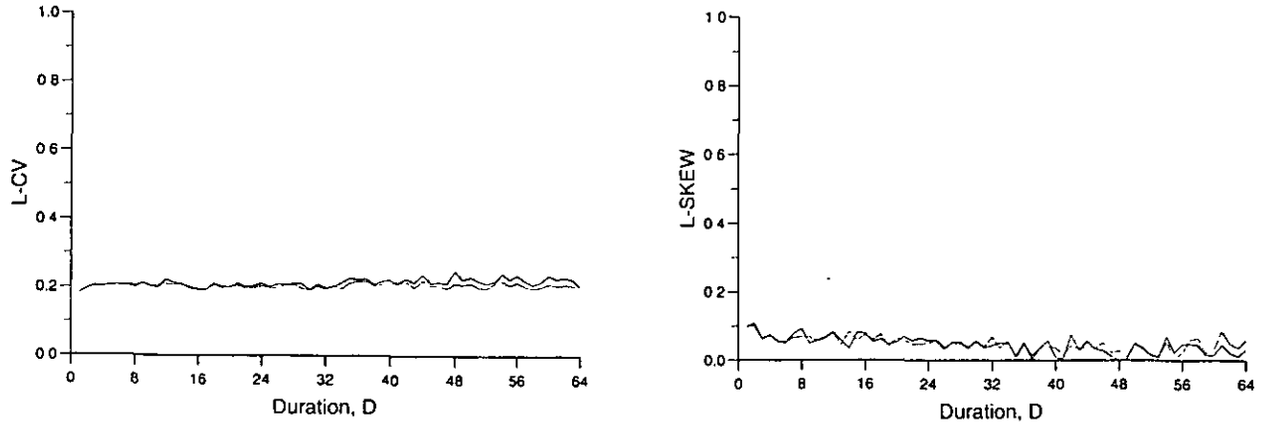


Figure D.1 *Eskdalemuir hourly air temperature record G*

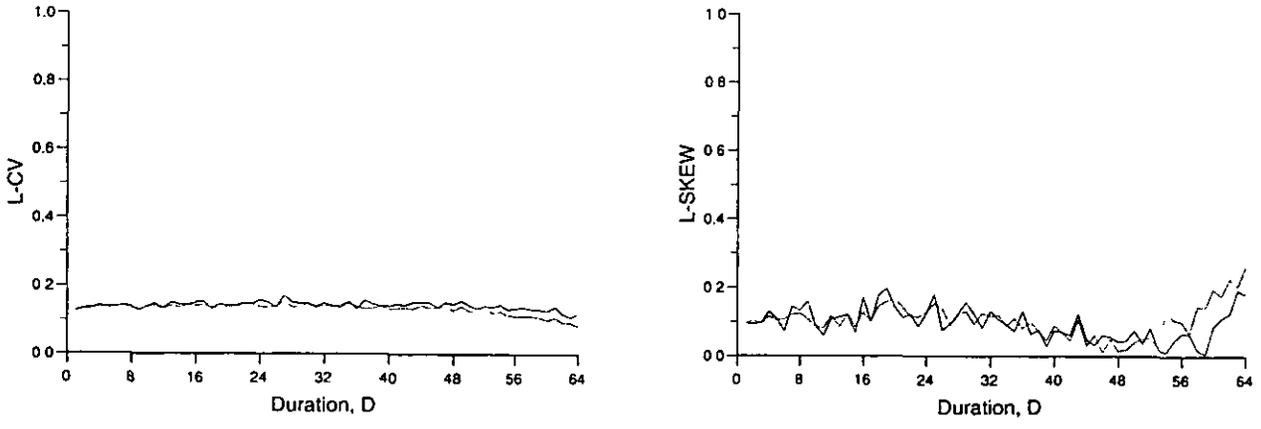


Figure D.2 *Eskdalemuir hourly wind speed record F*

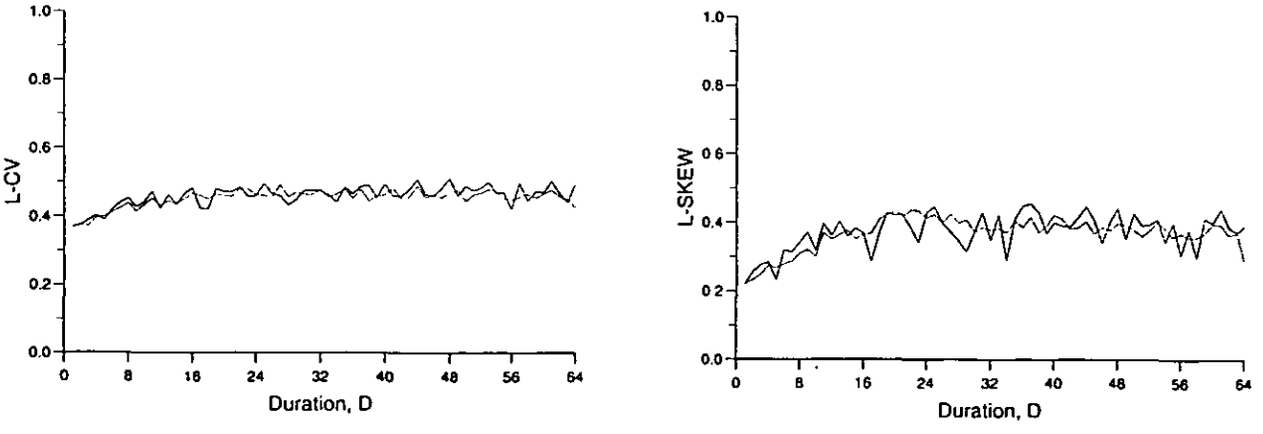


Figure D.3 *Melbourne hourly rainfall record B*

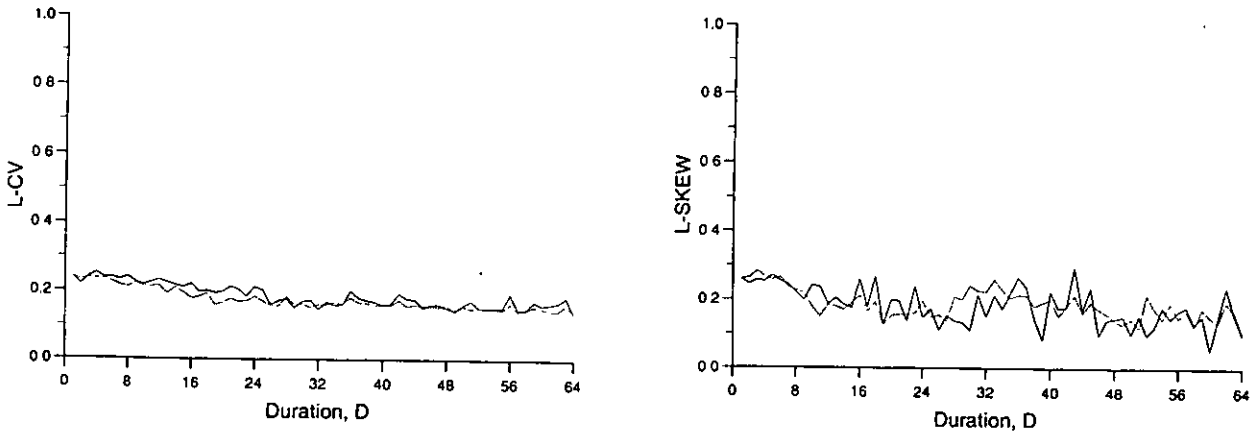


Figure D.4 *Durban daily rainfall record*