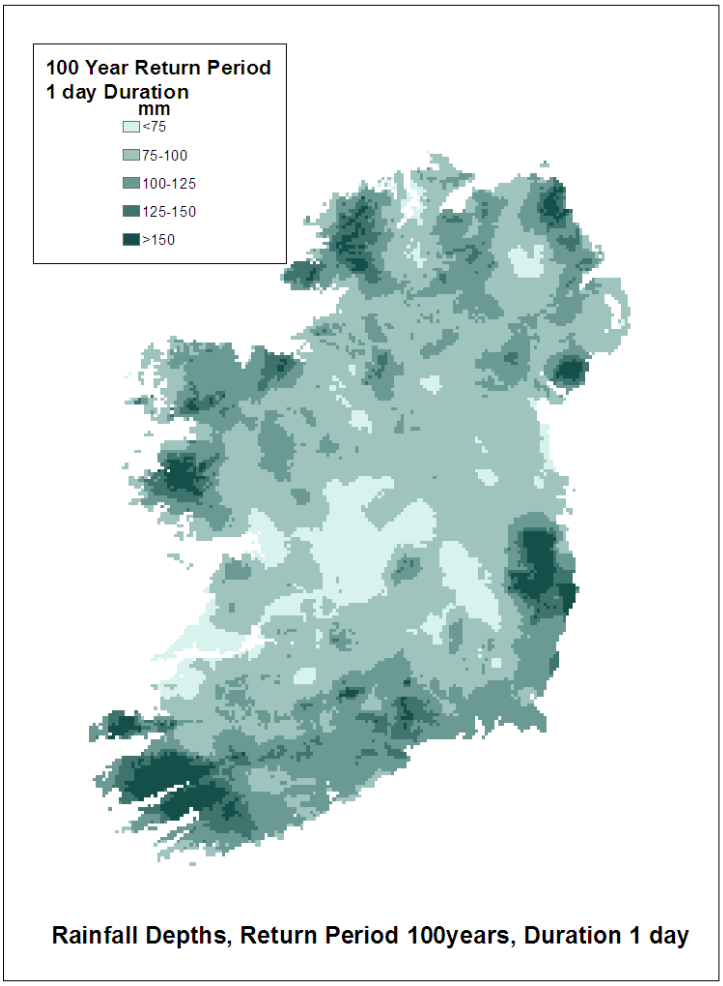
Met Éireann

Irish Meteorological Service

Technical Note 61

Estimation of Point Rainfall Frequencies

D.L. Fitzgerald



Met Éireann, Glasnevin Hill, Dublin 9, Ireland
UDC: 551.577.37 (415) October, 2007 ISSN 1393-905X

ESTIMATION OF POINT RAINFALL FREQUENCIES

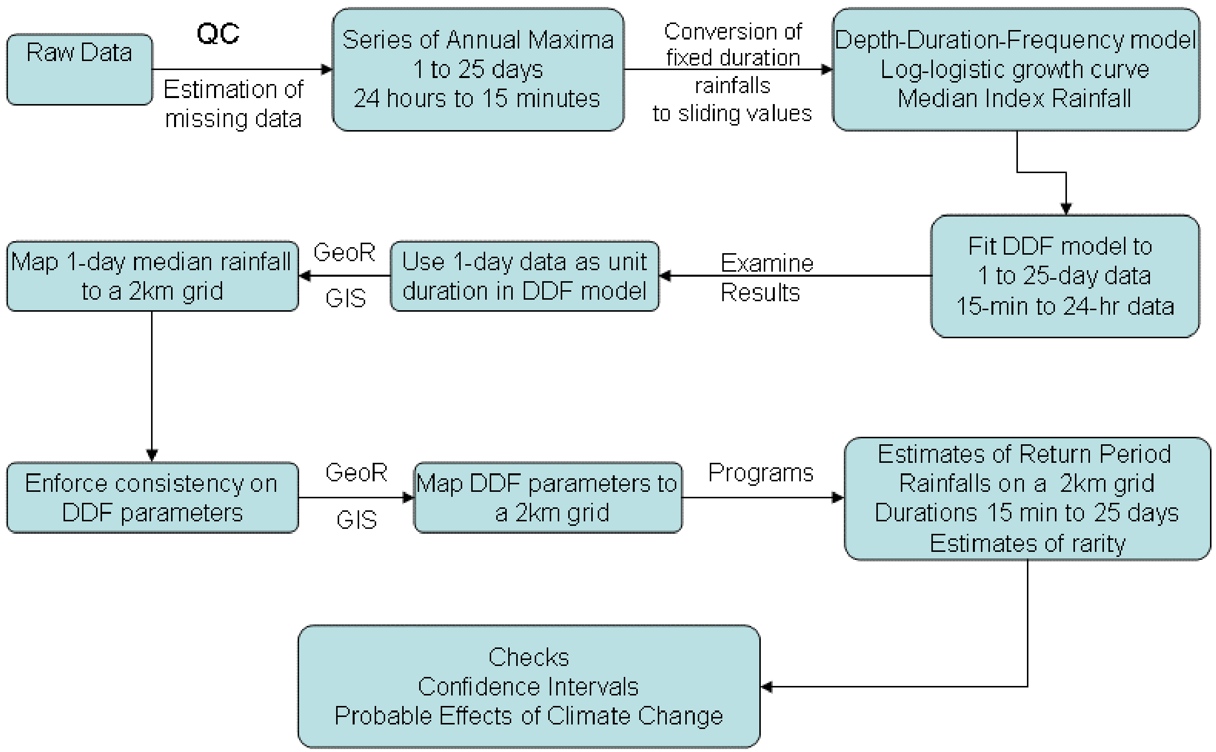
WorkPackage 1.2, Flood Studies Update

A study undertaken by Met Éireann for the Office of Public Works (OPW)

Table Of Contents

Executive Summary

A depth duration frequency model is developed which allows for the estimation of point rainfall frequencies for a range of durations for any location in Ireland. The model consists of an index (median) rainfall and a log-logistic growth curve which provides a multiplier of the index rainfall. Rainfall station data were analysed and an index rainfall extracted, interpolated and mapped on a 2km grid. The model was fitted to series of annual maxima and the growth curve parameters were determined; these were also interpolated and mapped on a 2km grid. Computer applications were written to apply the model and produce gridded outputs of the return period rainfalls which can easily be mapped; an application for deriving rarity estimates was also developed. An account is also given of the reliability and probable accuracy of the model and the probable effects of Climate Change on extreme rainfalls.



Work Plan for Study

Design Rainfall Estimates for the Irish Flood Studies Update Project

Introduction

The work was undertaken by Met Éireann using funding provided by the Office of Public Works (OPW) and is a module of the Flood Studies Update (FSU) Project.

The requirement was to a produce a gridded set of parameter values summarising the rainfall Depth-Duration-Frequency (DDF) relationship, and thereby enable the production of consistent estimates of point rainfall frequencies over durations ranging from 15 minutes to 25 days. The estimates were to supersede those provided in Logue (1975) in which the methods of FSR (Flood Studies Report ,1975) were adapted to Irish conditions. All the **design rainfall outputs** are for sliding durations e.g. an 8-day estimate is for 192 consecutive hours and may start at any hour of day; this contrasts with the raw data which are mostly for fixed durations e.g. daily values read at 0900UTC.

The body of this document describes the **model**, the **data**, the **conversion** of daily data **from fixed to sliding durations**, **the methods of spatial interpolation**, and contains an assessment of **reliability levels and confidence intervals for the gridded estimates**. More technical descriptions of some of these topics are given in the appendices. Possible **effects of climate change** and **comparisons with the estimates from Logue (1975)** are also included.

The model developed enables the estimation of rainfall frequencies at any location. As series of annual maxima were employed throughout, the rainfall frequencies are expressed in terms of **return periods**; this and other technical terms will be explained. Some guidance going beyond the requirements e.g. estimation of 5-minute and 10-minute return period rainfalls is given and the matter of converting return periods to average recurrence intervals is treated.

An underlying assumption is that the 1941-2004 data used in this study will reasonably represent the upcoming rainfall regime. Given the consensus view that we are in a period of global warming this is not a safe assumption even in the medium term. General indicators of the **effects of global warming on the precipitation regime** are available but are heavily dependent on the particular parameterisations used in the general circulation model. The indications of the latest assessment of the Irish climate modelling group, C4I (**Community Climate Change Consortium for Ireland**), are given. However, appropriate adjustments are not included in the estimates of the return period rainfalls as it appears that for quite a number of years into the future the indications of the effects of global warming on precipitation regime will change from assessment to assessment. The latest advice on the probable effects of climate change on extreme rainfalls should be sought.

1. The Depth-Duration-Frequency (DDF) Model

A DDF model consists of:
1. An index rainfall value e.g. the median or mean
2. A growth curve which provides a multiplier of the index rainfall

The model developed here applies to a single location and its mathematical form enables the estimation of return period rainfalls over a range of durations (D) and return periods (T). Appendix A gives the justification for using the log-logistic distribution as the growth curve and the median of the series of annual maxima as the index rainfall.

The growth curve is of form:

$$\frac{R(T,D)}{R(2,D)} = (T-1)^c, T = \frac{1}{(1-F)} \quad , \ T>1 \qquad (1)$$

where F is the cumulative distribution function.

R(T,D) is the rainfall of duration *D* with return period T, where *T* is the average number of years between **years** with one or more rainfalls exceeding the value *R(T,D)*. It is important to note that T is not the average recurrence interval (ARI) between rainfalls exceeding *R(T,D)*. Analysis of annual maxima leads naturally to the expression of time intervals in terms of return periods. Analysis of partial duration series (PDS), also termed peak-over-threshold (POT) analysis, leads to average recurrence intervals. How to convert a given ARI into a value of T so that rainfall for a given ARI may be estimated from annual series is discussed in the next section.

R(2,D) is the population median at duration *D* i.e. half the annual falls exceed *R(2,D)* since $T = 2$ corresponds to $F = 0.5$ and is the index rainfall ; it acts as a scaling factor in the DDF model.

A plausible form for the variation of the median with duration is (see Appendix A)

$$R(2,D) = R(2,1)D^s \qquad (2)$$

where D =1 is a suitably chosen unit duration which is 24 hours (1d) for both the 1d to 25d rainfalls (D >= 1) and for durations less than 24-hours (D < 1). Thus the 1d median rainfall, $R(2,1)$, plays a pivotal role in the DDF model.

The full DDF model combines (1) and (2) and is:

$$R(T,D) = R(2,1)D^s(T-1)^c, \ T>1 \qquad (3)$$

The exponent *s* in (2) determines the multiplier of R(2,1) yielding *R(2,D)* as duration, *D*, varies. The exponent *c* in (1) is the shape parameter of the log-logistic growth curve (see Appendix A) and determines the multiplier of *R(2,D)* which yields *R(T, D)*, the rainfall at return period *T* and duration *D*.

The final forms of the models consisted of (3) with 24 hours as the unit duration and with the exponents of form:

1.1 1 to 25 days:

$$c_D = a + b\ ln(D),\ D\ ranging\ 1\ to\ 25 \qquad (4)$$

$$s_D = e + f\ ln(D) \qquad (5)$$

Here *ln* is the natural logarithm.

Note that with $D = 1, c_1 = a = c_{24}$, the shape parameter at the 1-day (24-hour) duration.

Both exponents are taken to be functions of *D* and not of *T*.

1.2 24hours to 15 minutes:

$$c_D = c_{24} + h \ln(D),\ D < 1 \qquad (6)$$

$$s = s \qquad (7)$$

Here *D* = 0.75 at 18 hours, *D* = 0.5 at 12 hours , *D* = 0.25 at 6 hours and so on.

The shape parameter is, again, a function of duration but the duration exponent *s* is not. Only two parameters, *h* and *s* need be determined as c_{24} is available from the work on 1 to 25 days..

The reasons for these choices are discussed in Appendix A.

1.3 Implementing the DDF model

The two DDF models were fitted to the station data by a method described in Appendix A. Geostatistical methods (Kitanidis, 1997) were used to interpolate the station values of the parameters to the 2km grid. Two estimates are required:

1. *R(T,D)* the return period rainfall at duration *D*, given return period *T* and duration *D*

and

2. *T*, the return period given duration *D* and rainfall amount *R(T,D)*

Estimating R(T,D) and/or T at grid points is then straightforward. The method of interpolation between grid points is described at the end of Appendix A where the matter of the most appropriate parameter values to attach to 'the representative point' of a catchment is briefly discussed.

Programs were written to estimate *R(T,D)* over the 2km grid and also to calculate either R(T,D) or T at any location.

1.4 Conversion Factors for Partial Duration Series (PDS)

By definition annual maximum series (AMS) consist of the highest fall for each year; the second highest fall is ignored whether or not it exceeds the highest fall in other years. Partial duration series (PDS) consist of all falls exceeding a certain threshold together with their times of occurrence.

The return period T is best thought of as the inverse of a probability e.g. the rainfall corresponding to T = 50 has a probability of 0.02 of being exceeded next year. Risk can also be expressed in terms of return period rainfalls. The rainfall corresponding to T = 238 years has a probability of 0.9 of not being exceeded during the next 25 years

(Zucchini & Adamson, 1989) i.e. the risk *r* is 0.1. Assuming that the annual maxima are statistically independent and all are drawn from the same distribution

$$T = 1/(1-(1-r)^{\frac{1}{n}}) \qquad (8)$$

where the design horizon is *n* years and the risk is *r*.

The analysis of PDS gives the average period between rainfall events that exceed a particular value and is often termed the average recurrence interval (ARI) for a given duration *D*. For high values of T, values of ARI and T are nearly equal but for T less than 20 years the difference can be significant. Langbein (1949) provides a formula (see Appendix D) for the relationship between T and ARI which yields:

T = 1.16 for ARI =1/2
T = 1.58 for ARI = 1
T = 2.54 for ARI = 2

The approximation T = ARI + 0.5 improves as ARI increases.

Thus the PDS rainfall for ARI = 2 is the return period rainfall for T =2.54 years. As annual series were used in this study we can only estimate growth curves for return period rainfalls but the Langbein formula enables their conversion into PDS rainfalls with a known ARI.

2. Rainfall data for the Flood Studies Update

2.1 Irish Stations

(Note: In this section 'Irish data' is taken to mean data from rainfall stations in the Republic of Ireland)

2.1.1 Daily(0900-0900UTC) data:

The requirement was to form series of annual maxima for at least six durations ranging from one to twenty five days. On assessing the amount of quality control needed to treat each duration it was decided that the production of complete daily series was the better option.

The Met Éireann archive for periods within 1941-2004 was used as this had already undergone extensive quality control. However, dry months and daily falls in excess of 75.0mm were re-examined and some faulty data corrected before forming the initial table of daily rainfall. High-quality stations were picked out by examining the number of accumulated totals or missing days in each month of record. For these stations accumulations were broken up into daily values and missing days estimated to give a complete set of daily values.

For most stations estimations were made by using up to six neighbouring stations with similar (within about 10%) average annual rainfall (AAR); these were ranked in order of preference and the first to have a complete record for a period requiring estimation used as this had the perceived advantage of using a total actually recorded in the general area rather than a weighted mean.

If the AAR of the station requiring estimation differed considerably from its nearest neighbours then three neighbours with complete daily records were chosen and the global monthly ratios of target station over neighbour determined. For missing months the total was taken as a weighted mean of the three estimated totals and then reduced to daily values by reference to the daily falls at the nearest neighbour.

Months with a total but no daily values were treated by forming the weighted mean of the three neighbours and giving it equal weight to the target station total; the agreement between the two was usually good. Again daily values were found from the estimated monthly total by reference to the nearest neighbour.
The remaining missing or accumulated days were treated by multiplying the daily values at the nearest available neighbour by the monthly ratio.

Annual maxima extracted from the original and treated data showed that the differences were usually small as the quality of the chosen stations was high.

Using these methods daily values for 474 stations were extracted; their average period of record was 41.2 years, with a range of 20 to 64 years. Annual maxima for 1, 2, 3, 4, 6, 8, 10, 12, 16, 20 and 25 days together with the starting date were put into a table of 214,918 rows. The location of these stations plus 103 stations from Northern Ireland are shown in Figure 1.

2.1.2 Short-duration falls

Data for nine (sliding) durations between 15 minutes and 24 hours are available from 39 stations for periods ranging from 15 to 55 years but 37 of the stations have 30 or more years of record. The locations of the 39 stations used are shown in Figure 2.

Mostly the data were extracted from Dines recorder charts but since the mid-nineties these have been replaced by tipping bucket recorders (TBR) at some stations. In the periods of overlap the differences between the Dines and TBRs were found to be generally small and no adjustment for the transition was made.

The maxima of all falls attaining or exceeding at least one of a set of thresholds were extracted. The thresholds are:

Duration	15m	30m	1h	2h	3h	4h	6h	12h	24h
Threshold (mm)	4 .0	5.0	6.0	10.0	12.5	15.0	20.0	25.0	30.0

The date assigned is that of the day on which most, or all, of the total for the rainfall event was recorded. Quality control consisted of:

1. Checking the dates of occurrence and values of the
2. 0900-0900UTC totals of 30mm or more against those of the 24h values.
3. Examining the rows of the table for consistency.

Doubtful or missing values detected were estimated by reference to the nearest stations.

For the 16 synoptic stations a more extensive quality control was possible as all 1 clock- hour to 24 clock-hour totals reaching or exceeding the thresholds and their dates of occurrence were checked against the corresponding absolute maximum values.

To extract annual series we must deal with the problem of years with no value exceeding the threshold. In the case of synoptic stations good estimates can be made from the clock-hour values, if desired.

The strategy applied to all stations was to trust in the quality of the data and assume that the missing values were below the appropriate thresholds and hence unimportant for fitting a probability distribution (log-logistic) to extremes. The sample could be regarded as censored and methods that take account of this developed to determine the parameters (Appendix B). In the event, only the ordered set of values greater than or equal to the sample median was needed to fit the distribution; fortunately, at all stations these series were complete.

2.2 Data from Northern Ireland

Through the good offices of the UK Met Office and the Centre for Ecology & Hydrology (Wallingford) the daily and hourly series of annual maxima from Northern Ireland used in the Flood Estimation Handbook (1999) were made available. Of the daily stations 103 with at least 20 years of record at durations of 1, 2, 4 and 8 days were used in this study. The hourly data consisted of 8 stations with lengths of record between 11 and 19 years and with data at durations of 1, 2, 4, 6, 12, 18 and 24 hours.

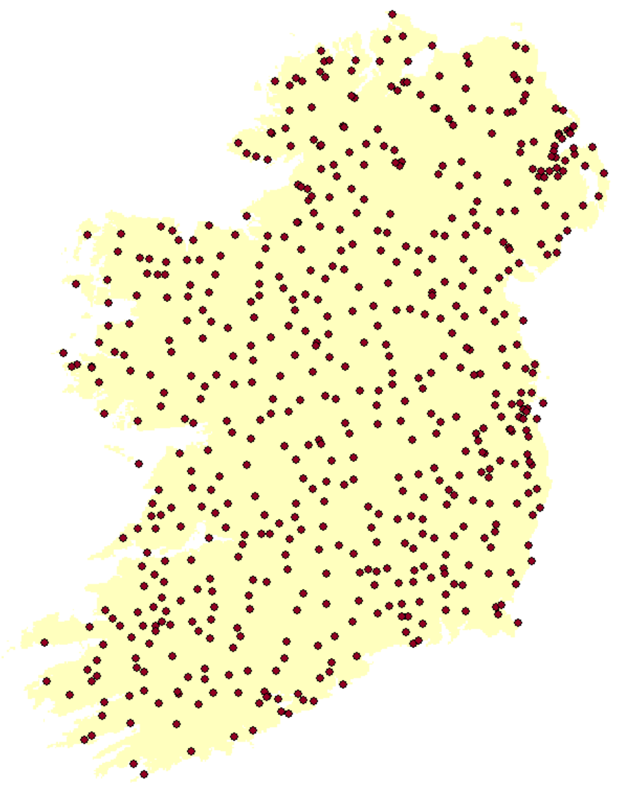


Figure 1. 577 Stations used in the Daily Rainfall Analysis.

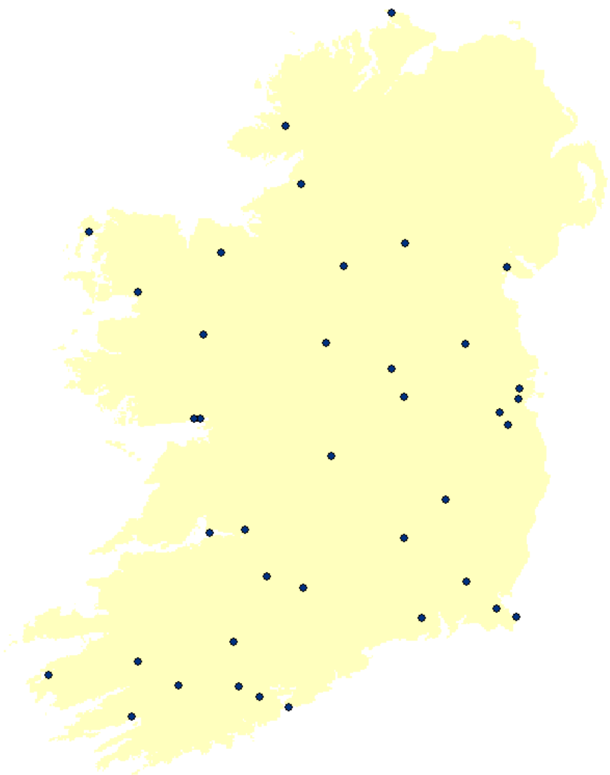


Figure 2. 39 Stations used in the Short Duration Analysis

2.3 Conversion from fixed to sliding durations

The data available were the 1-day, 2-day…..25-day annual maxima derived from daily totals read at 0900UTC. However, the aim was to estimate the 24-hour, 48-hour……600-hour return period rainfalls where 24-hour is the annual maximum for any 24-hour period within the year. How to do this from the 0900-0900UTC data involves conversion factors from fixed to sliding durations. The matter was examined in two ways:

1. At the 14 longer-term synoptic stations the log-logistic distribution was fitted to the 1, 2, 3, 4, 6, 8, 10, 12, 16, 20 and 25-day annual maxima of the 0900-0900UTC rainfall totals and also to their 24, 48, 72, 96, 144, 192, 240, 288, 384, 480 and 600 clock-hour totals. Return period rainfalls of 2, 5, 10, 20, 50, 100, 250, 500 and 1000 years were calculated for each and the ratios examined. No allowance was made for the transition from clock-hour values to absolute values as inspection showed that, even for 24 hours, the factor was very close to one. Use of the year January-December revealed significant end-of-year effects especially at the longer durations. As a check the April-March period was used and it was found that the values of the return period rainfalls were little changed and so **April-March was used as the rainfall year**.

2. For the eleven 0900-0900UTC durations 1, 2, 3, 4, 6, 8, 10, 12, 16, 20 and 25-days thresholds of 23, 30, 35, 40, 48, 55, 60, 65, 70, 75 and 80mm were set and the *n*-day exceedances compared with the corresponding clock-hour values; corresponding was taken as starting within a certain interval that had the starting date of the 0900-0900UTC accumulations as a fairly central value. Various intervals were tried and the general effect of narrowing the interval was to slightly decrease the ratios. Sample average ratios $= \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{y_i}$ were derived from the individual events, this in preference to the ratio of the sample averages $= \frac{avg\,(x_i)}{avg\,(y_i)}$ which gave slightly lower values. As pointed out in Dwyer & Reed (1995) the ratio of the sample averages is equivalent to a weighted sum of the individual ratios, with greater weight given to the larger events.

The results of fitting the log-logistic distribution, presented below in **Table A**, suggest that for the median rainfall the conversion factors should be :

	1d	**2d**	**3d**	**4d**	**6d**	**8d**	**10d**	**12d**	**16d**	**20d**	**25d**
RP =2	1.15	1.08	1.06	1.05	1.04	1.04	1.03	1.03	1.02	1.02	1.01

The work on events gives :

Events	1.11	1.06	1.04	1.04	1.03	1.03	1.02	1.02	1.01	1.01	1.01

Note that the factors based on means of actual events are generally lower. This may be attributed in fair measure to the fact that for events the times of commencement of the 0900-0900UTC and hourly accumulations are fairly close while, especially for the 24 and 48-hour durations, the time of start of the 0900-0900UTC and hourly annual maxima may, on occasion, be separated by months i.e. there is an extra source of variation. If you extract the mean of 1.11 as a characteristic multiplier for the 1-day events it is from a positively skewed distribution with a median of 1.05 but with 312 of a total 2285 cases in excess of 1.25 and a range from 1.0 to 2.0.

In Table A the range of the RP2 (median) factor is from 1.086 to 1.257 with a mean of 1.153, a median 1.148 and a standard deviation of about 0.03.

Since it is the index rainfall, the values for the 2-year return period rainfall are most important but it is of interest to see how the factors vary with increasing return period. Table A gives the means over fourteen synoptic stations of the ratios of the clock-hour estimates of return period rainfalls to the corresponding fixed-duration values for each of the eleven durations. In general the ratios decrease with return period but are nearly constant for the one-day rainfalls and for durations of 16 days or more, while the 12-day falls increase with return period. As expected the decrease of the ratio with duration is consistent over all return periods.

Table A
Fixed v Sliding Durations
Mean of Adjustment Factors

Days	RP 2	RP5	RP10	RP20	RP50	RP100	RP250	RP500	RP1000	Events
1	1.153	1.148	1.147	1.146	1.145	1.146	1.148	1.149	1.152	**1.110**
2	1.076	1.064	1.058	1.051	1.044	1.038	1.032	1.026	1.022	**1.060**
3	1.062	1.048	1.040	1.032	1.023	1.017	1.009	1.002	0.996	**1.044**
4	1.044	1.036	1.032	1.028	1.023	1.020	1.016	1.013	1.011	**1.037**
6	1.041	1.029	1.022	1.015	1.007	1.002	0.994	0.989	0.984	**1.027**
8	1.042	1.032	1.026	1.021	1.015	1.010	1.004	1.000	0.996	**1.024**
10	1.034	1.027	1.022	1.018	1.013	1.010	1.005	1.002	0.999	**1.021**
12	1.026	1.031	1.034	1.037	1.040	1.043	1.048	1.051	1.054	**1.017**
16	1.023	1.022	1.021	1.021	1.021	1.021	1.021	1.021	1.021	**1.014**
20	1.017	1.015	1.015	1.014	1.013	1.012	1.012	1.011	1.011	**1.011**
25	1.016	1.011	1.008	1.006	1.003	1.001	0.999	0.995	0.993	**1.010**

Since the conversion from 1-day 0900-0900UTC to 24 hour median values is crucial to the short-duration model, the medians of annual maximum series at the 39 stations for which absolute 24-hour maxima were available were compared with the medians of the 0900-0900UTC annual maximum series <u>over the same years</u>. The quartile summary of the 39 ratios is:

Minimum	Q1	Median	Mean	Q3	Maximum
1.0	1.091	1.123	1.124	1.164	1.331

Only 3 of the 39 values exceeded 1.20 and only 1 exceeded 1.25. The interquartile range yields an estimate of about 0.05 for the standard deviation.
Linear regression gives 1.128 as the ratio. The competing mean value is 1.15, the conversion factor for the log-logistic estimate of the median ratio and this was adopted as it compares with 1.16 used in FEH (1999) and 1.14 used in the New Zealand HIRDS system.

2.4 Notation

When it is necessary to distinguish between the two, the sliding-duration rainfalls obtained by multiplying the fixed-duration falls by the appropriate conversion factor will be referred to as 1d, 2d, 3d…25d rainfalls; the fixed duration falls will be labelled 1d09, 2d09,….25d09.

Sub-daily durations are always sliding durations. The Irish short-duration data consist of absolute maxima extracted from rainfall events for durations up to and including 24 hours. In what follows the latter value is referred to as *abs24.* The 1d value is a close approximation to *abs24* and has the advantage of being much more widely available. The 1d rainfall will be labelled *slide24* when considering durations of 24h to 15m i.e. it is being used as a substitute for *abs24*.

3. Mapping

3.1 Mapping the Index Rainfall

The pivotal value is the median one-day rainfall, *RMED*1d which is closely approximated by:

$$RMED1d = 1.15RMED1d09 \qquad (9)$$

The number of values of *RMED1d* available was 577 of which 103 were in Northern Ireland. This did not seem adequate to produce values on a 2km grid. Fortunately, there is a strong linear relationship between AAR (average annual rainfall) and *RMED1d*. In the case of the 1961-1990 averages (*AAR6190*) the relation

$$RMED1d = 0.0366(AAR6190) \qquad (10)$$

has a coefficient of determination $R^2 = 0.978$ that can be increased to 0.993 by the addition of location co-ordinates e.g. easting and northing(in metres) for each station:

$$RMED1d = 0.03331AAR + 0.000062easting - 0.0000353northing \qquad (11)$$

To exploit the strong relationships (10) or (11) values of AAR6190 on a grid were needed. The stations for which *AAR6190* were available numbered 946 of which 242 were in Northern Ireland. Drawing on considerable experience of estimating *AAR*, 101 values were added in data-sparse areas. The 1047 data points were used to produce values of A*AR6190* on a 2km grid using geostatistical methods (Kitanidis, 1997) and the R package, geoR. The gross dependence of AAR on elevation and location was removed by linear regression and the residuals interpolated to the grid using ordinary kriging with a moving neighbourhood i.e. weighted linear combinations of nearby values. The values of the regression equation at each grid point were added to the gridded residuals to produce the final result. Comparison with previously mapped values of *AAR6190* showed very good agreement.

Kriging was again used to produce gridded values of *RMED1d*. The number of stations having values of both *RMED1d* and *AAR6190* was 468. The residuals from the regression of *RMED1d* on *AAR6190* were interpolated to the 2km grid and the regression estimates added to produce the final mapping, Figure 3.

3.1.1 Comparison of recorded and interpolated values of the median rainfall

There were 109 stations having a value of *RMED1d* but no value of *AAR6190*. Using gridded *RMED1d* as the data values, the 109 kriging estimates (*KMED1d*) of *RMED1d* were obtained together with their kriging standard errors (*ksd*) .

The ratio $\dfrac{RMED1d}{KMED1d}$ is approximately normally distributed with mean 0.995 and standard error of 0.084. The percentage differences are less than 15% in 102 of the 109 cases, 66 are less than 7.5%, 50 less than 5% and 30 less than 2.5%. It indicates that *KMED1d* is an appropriate estimator of *RMED1d* and justifies the assumption that the mapped value is a good estimator of the actual median rainfall at a site.

The kriging standard error, *ksd* ,of *KMED1d* reflects the variability of the estimates which are weighted means of the surrounding grid-point values and so the standard error will be higher where the median rainfall changes more rapidly with distance e.g. in and near mountainous areas. It is reasonably well approximated by:

$$ksd = 0.001(KMED1d^2), \quad R^2 = 0.81 \qquad (12)$$

This kriging standard error increases monotonically with *KMED1d* and ranges from about 1mm to 9mm as *KMED1d* varies from 31 to 94 mm. As the higher values of *KMED1d* occur in the mountains where the density of the raingauge network is lowest, the uncertainty attaching to the interpolated values is highest there. The kriging standard error can be used to estimate the error of interpolation and hence is of interest as a measure of the reliability of *KMED1d*.

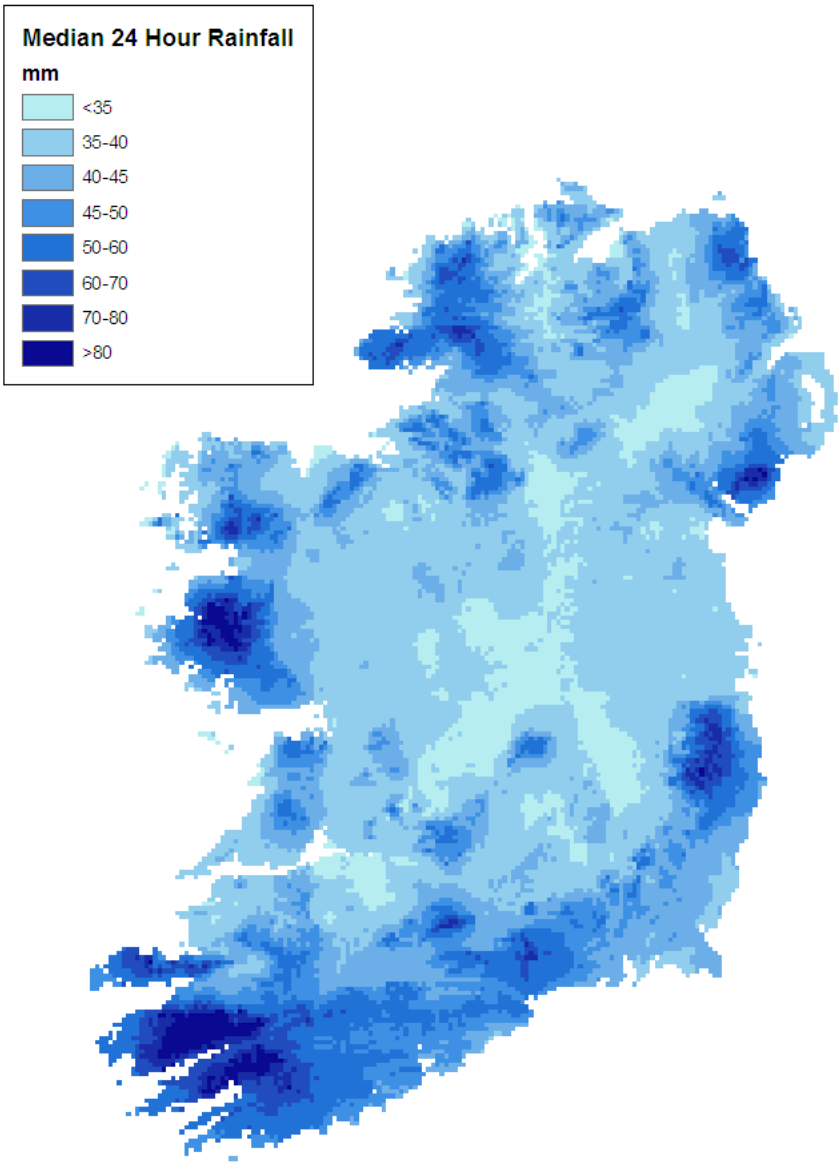


Figure 3. Index (Median 24hr) Rainfall.

3.2 Mapping the model parameters

3.2.1 1d to 25d (sliding) durations

No useful relationships were found between the four model parameters and median rainfall, AAR or easting and northing. Hence, all parameters were interpolated to the 2km grid using ordinary kriging with a moving neighbourhood i.e. a weighted mean with weight dependent on distance.

3.2.2 24h to 15m (sliding) durations

The model adopted meant that no new mappings were required. The gridded 1d shape parameter $c_1 = c_{24}$ and the grid of the median one-day rainfall were needed but these had been mapped as part of the work on the 1d to 25d model. The two model parameters *h* and *s* for the short-durations falls can then be calculated at any point by means of the scheme outlined in Appendix A.

3.2.3 Enforcing Consistency on the model parameters

For the 4-parameter model

$$c_D = a + b\,ln(D) \text{ and } s_D = e + f\,ln(D) \qquad (13)$$

To meet the requirement that, for the same return period T, the estimate should increase with duration we must have:

$$e + 2f\,ln(D) + b\,ln(T\text{-}1) > 0 \qquad (14)$$

Near 1000 years and near 1 day this becomes $e + 7b > 0$

For a given duration, to have a positive rate of change with increasing return period T requires:

$$a + b\,ln(D) > 0 \qquad (15)$$

At about 30 days this requires $a + 3.5\,b > 0$

3.2.4 1d to 25d Irish model

In nearly 7% of cases there occurs the problem that at high return periods (~1000years) that the 2-day estimate may be slightly less than the 1-day estimate. The basic reason is that the shape parameter *c* is decreasing too sharply. This can be simply corrected by enforcing the condition :

$$e + 7b > 0 \qquad (16)$$

For a given duration the return period rainfall estimates must increase with T but this presented no problem. An example of the output is shown in Figure 4.

3.2.5 Consistency of the model for the Northern Ireland data

The parameter values for the Irish data for 1d to 8d had similar quartiles to the 1d to 25d data (see Table 1, Appendix A). Hence, it was decided to use the Northern Ireland data as if it were for periods up to 25d, a convenient but decidedly risky assumption.

The condition $e+7b > 0$ had to be imposed in nearly 17% of cases and this adjustment also remedied the failure of 8% of the stations to meet the requirement $a + 3.5\,b > 0$.

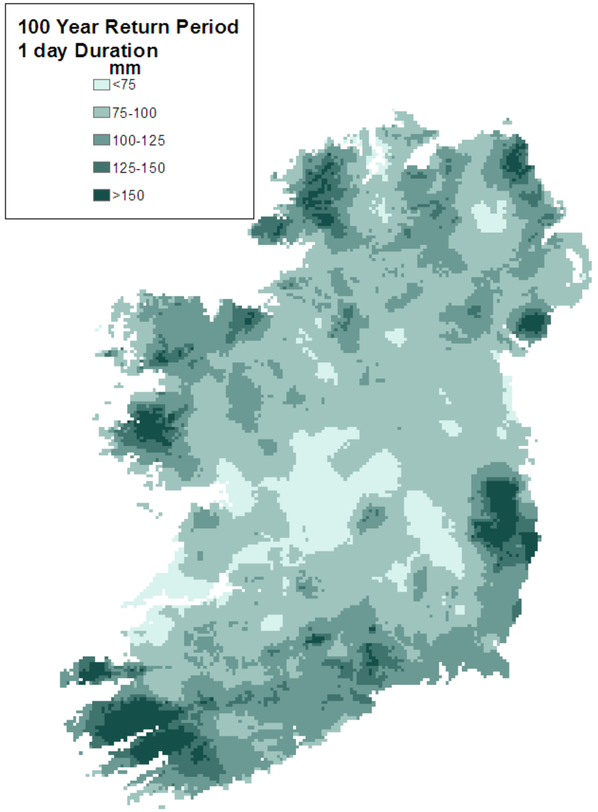


Figure 4.
Rainfall Depths, Return Period 100years, Duration 1 day

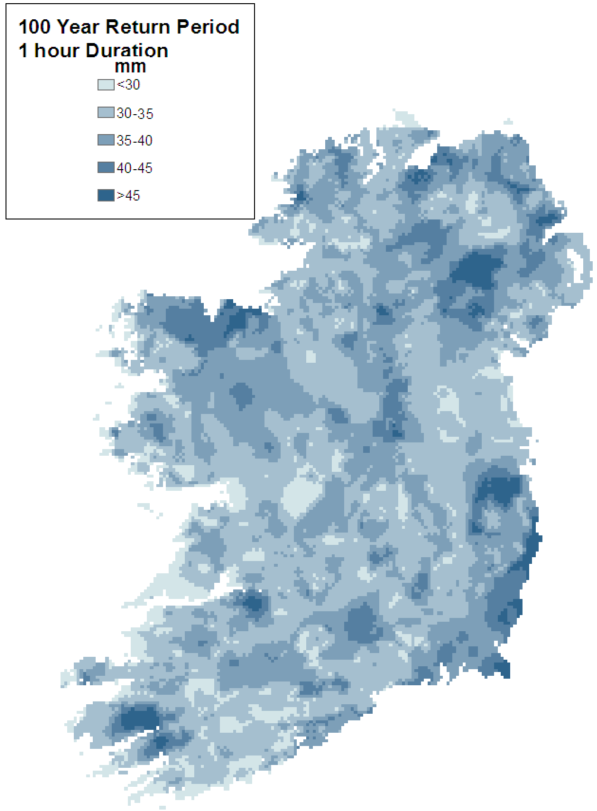


Figure 5.
Rainfall Depths, Return Period 100years, Duration 1 hour

3.2.6 Model for 24 hours to 15 minutes

Recall that the model is of form:

$$R(T,D) = R(2,1)D^{s}(T-1)^{c_{24}+h\ln(D)} \qquad (17)$$

At a given duration the estimate increases with return period T if:

$$c_{24} + h\ln(D) > 0 \qquad (18)$$

At a given return period the estimate increases with duration if:

$$s + h\ln(T-1) > 0 \qquad (19)$$

These conditions were kept in mind when assigning values to *h* and so no consistency problems arose with the short-duration model as $-0.03 \leq h \leq 0$, $\ln(D) \leq 0$ and $s \geq 0.33$. An example of the model output is shown in Figure 5.

3.3 Falls of durations less than 15 minutes

Estimates of rainfall depths for durations less than 15min are considered in Appendix E. The conclusion is that it is probably best to employ 15-minute estimates and apply the given formula for the mean fraction as a function of the fractional duration. In the case of ten and five minute durations the fractions of the fifteen minute depths are 0.85 and 0.61 respectively. Estimation for intervals of less than 5 minutes duration should be regarded as highly speculative.

4. Reliability and Probable Accuracy of the Return Period Rainfalls

These matters are discussed in more detail in Appendix C where the tentative conclusions are:

1. The 24-hour to 600-hour estimates may be used with fair confidence for return periods up to about 500 years
2. The estimates for durations of less than 24 hours may be used with fair confidence for return periods up to about 250 years.

The statistical analysis was done on the assumption that the data are representative of the future rainfall regime. Given concerns about probable changes in precipitation climate in the short to medium term due to global warming this is by no means assured. How, or even if, to adjust the estimates is not a question to which there is a good answer at present but some general, perhaps even useful, observations are made in the next section.
Accepting the current consensus on the high likelihood of changes in the precipitation climate, there seems to be little sense in estimating 500-year return period rainfalls. However, equation (8) shows that the very practical matter of estimating a 10% risk for a time horizon of 50 years requires a return period rainfall for 475 years.

Also in Appendix C a method is given for attaching a standard error to any estimated return-period rainfall from knowledge of the values of the median rainfall and the shape parameter plus a very large assumption about the effective sample size. This statistical measure of spread gives an idea of the probable accuracy of the value. It has to be regarded as merely a rough guide.

Information on the rainfall extremes of the last 100 years or so may be had in Rohan (1986) or Hand et al .(2004).

5. Probable Effects of Climate Change on Extreme Rainfalls

The most recent IPCC report on regional climate projections(Christensen et al., 2007) states that over Northern Europe between 1980-1999 and 2080-2099 the median change in precipitation was a 15% increase for the months of DJF, 12% for MAM, 2% for JJA and 8% for SON. Relative to the wettest period in 1980-1999 there was a 43% increase in wet events in DJF. However these were over a large area and the model resolution at ~ 200km was rather coarse.

The latest assessment from C4I (Community Climate Change Consortium for Ireland) states that over Ireland by mid-century there may well be:

1. An increase of about 15% in winter rainfall amounts
2. Drier summers with 20% lower precipitation in some areas, most likely the east and southeast.
3. A 20% increase in the two-day extreme rainfalls, especially in northern areas and smaller increases in the one and five day extremes.
4. More frequent rainfall extremes in autumn.

All this would suggest an increase in extreme rainfalls for durations of 24 hours or more, especially in autumn-winter. Drier summers suggest an increase in the frequency of droughts. The breakdown of droughts is sometimes the occasion of heavy short-duration rainfalls.

The general suggestion of most of the scenarios is that safety factors of maybe as much as 20% on rainfall depth might be incorporated as an attempt at a ‘no regrets policy’ in the face of uncertainty.

A purely statistical exercise in Fitzgerald (2005) comes up with safety factors for 1-day rainfalls at Phoenix Park(Dublin) of about 11% for a 20-year return period rainfall, 19% for the 100-year value and 33% for 1000-year rainfall, this based on 122 years of daily data.

Given the wide variation in predictions from assessment to assessment or between statistical exercises it would be wise always to seek the latest advice on the probable effects of climate change on extremes of precipitation before considering an adjustment to the model estimates of return period rainfall.

6. Comparisons with the estimates of Technical Note 40 (TN40)

Since 1975 the methods discussed in TN40 (Logue,1975) have been used to supply design rainfalls in Ireland. The methods were those of the British Flood Studies Report(FSR,1975) adapted to Irish conditions. The methodology was to fit a set of growth curves based on the generalised extreme value distribution (Hosking & Wallis, 1997) to series of annual maxima. The index rainfalls were the 2-day 0900-0900UTC rainfall with a five year return period and the 1-hour rainfall with five year return period. Growth curves were defined mostly in terms of duration and average annual rainfall (AAR).

The present study uses the log-logistic distribution as the growth curve and the median as the index rainfall. For an AAR of 1100mm the tables of TN40 can be used to characterize its growth curves in terms of the <u>shape parameter of the log-logistic distribution</u> i.e. the values of the 50-year return period rainfall over the median rainfall extracted from Table III of TN40 were inserted in the equation $M50/M2 = (T-1)^c = 49^c$.

Since 1100mm is near the national average rainfall these values are taken are roughly corresponding to the mean values of the parameters of the DDF model i.e. the mean value of 0.19 for c_{24} was used together with rate of change parameters (with the logarithm of duration) of -0.028 for 24hours or longer and of -0.010 for durations of less than 24 hours to yield the following:

Table C

Comparison of TN40 and FSU in terms of their Mean (log-logistic) Shape Parameters

Study	15m	30m	1 h	2h	4 h	6 h	12h	1 d	2 d	4 d	8 d	16d	25d
TN40	0.240	0.230	0.220	0.200	0.185	0.180	0.170	0.160	0.155	0.130	0.120	0.110	0.105
FSU	0.235	0.230	0.220	0.215	0.210	0.200	0.195	0.190	0.170	0.150	0.130	0.110	0.100

TN40 employed calendar month values and these were taken as equivalent to the FSU 25d falls.

Agreement is surprisingly good with the main differences centered on the 24-hour mean. More significant are the differences in the <u>ranges</u> of the shape parameters e.g. the 24-hour shape parameter has a range from 0.19 to 0.10 in TN40 and the 1-hour shape parameter ranges 0.26 to 0.16, the highest values in areas of low AAR and the lowest values in mountainous areas. In the FSU study the ranges of the shape parameters are wider e.g. the 1d shape parameter varies between 0.31 and 0.11 and Figure 6 shows that quite a number of the lowest values are in lowland areas. This helps to explain the wide ranges of the spot value percentages for the 24-hour duration in Figure 7 and for the 1-hour duration in Figure 8 The main point to note is that most of the values are in the range 85% to 125% in both cases.

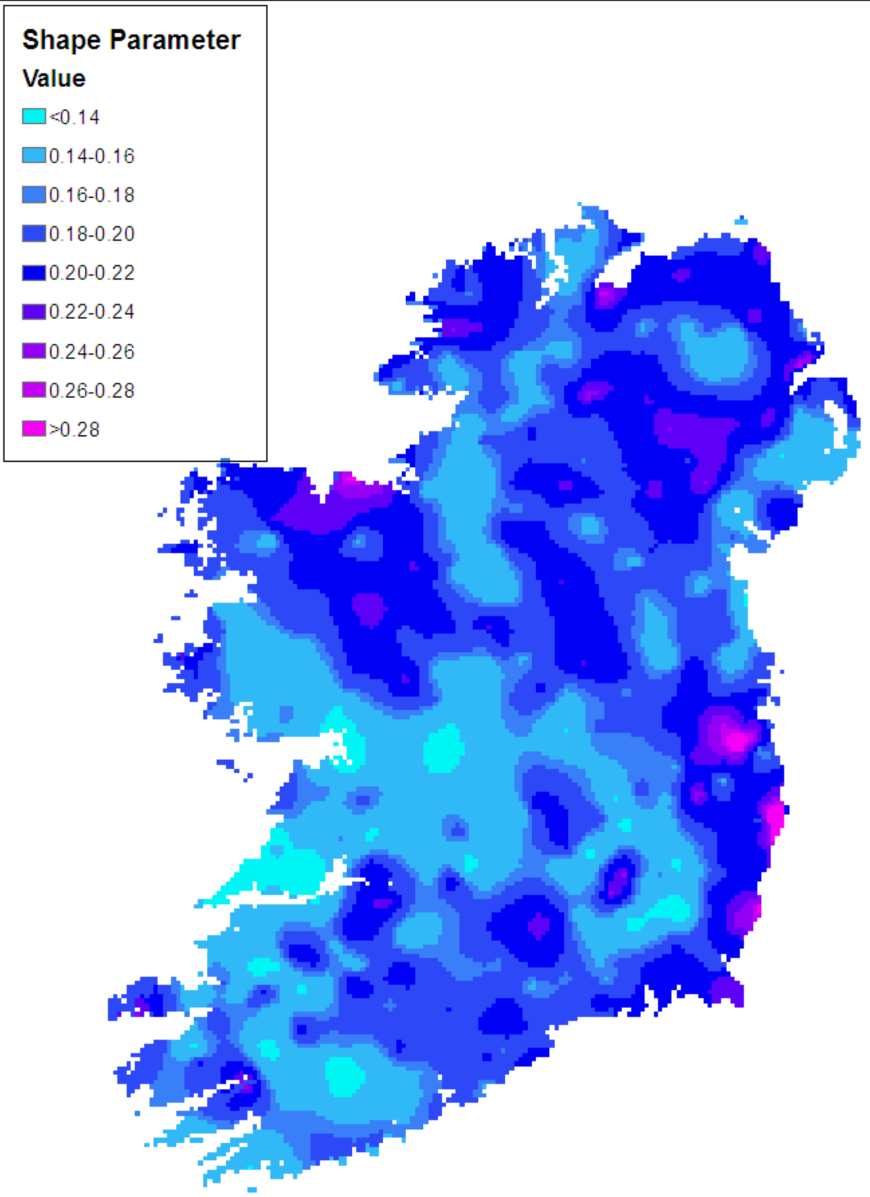


Figure 6.
Map of the Shape Parameter

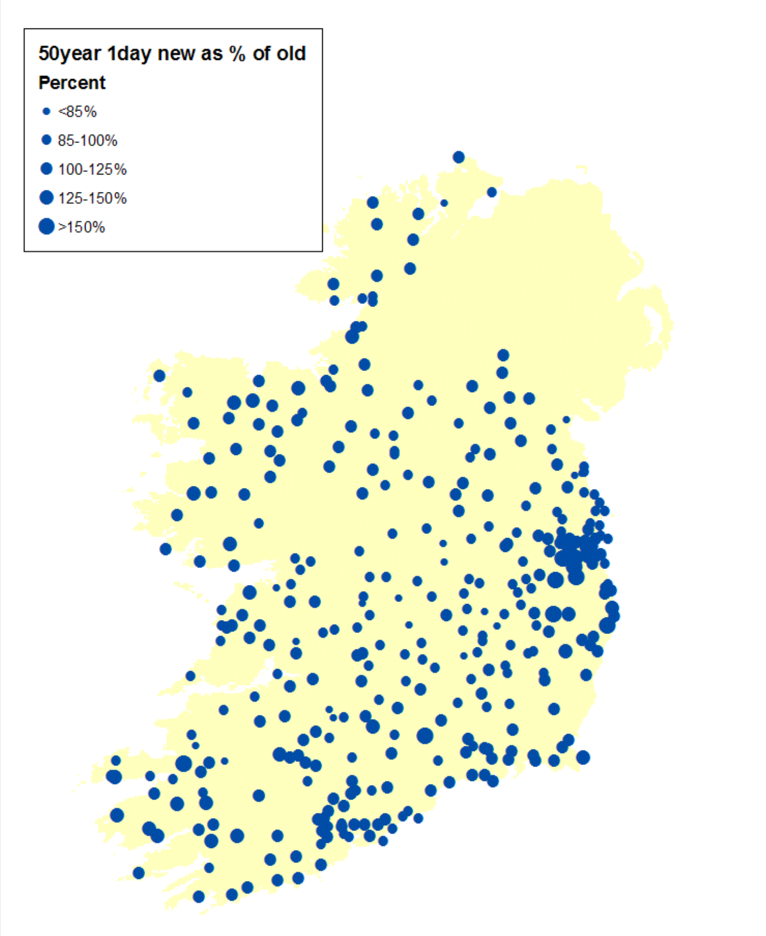


Figure 7.
New Values as Percentage of Old
50year Return Period, Duration 1 day.

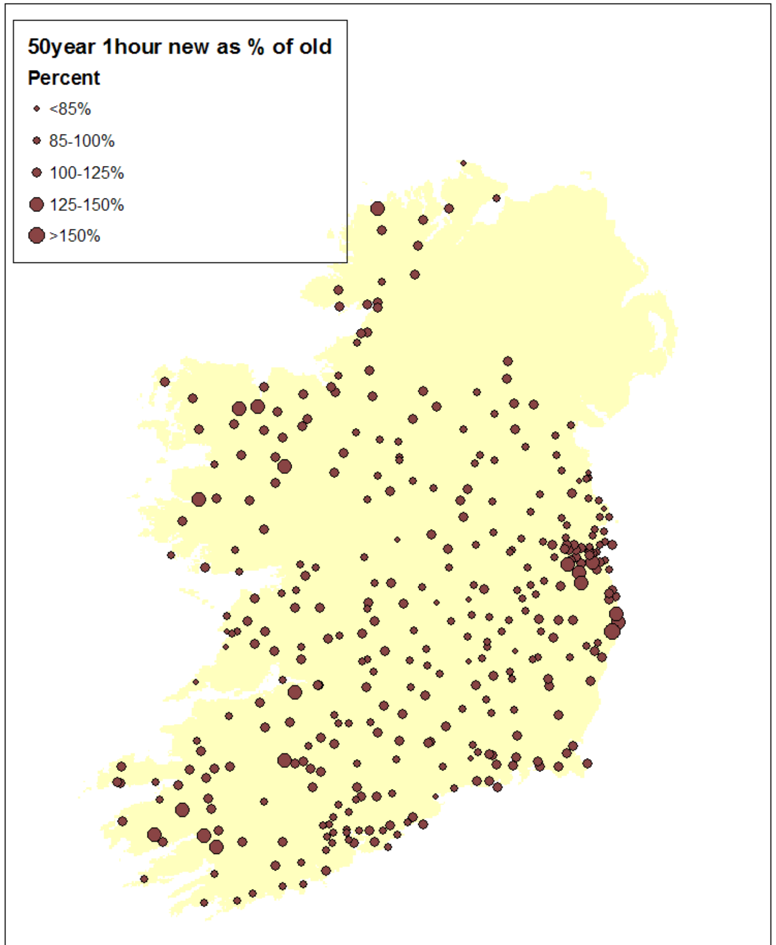


Figure 8.
New Values as Percentage of Old
50year Return Period, Duration 1 hour.

7. References

Abramowitz, M. & Stegun I. A. (1965), Handbook of Mathematical Functions, Dover, New York

Christensen et al. (2007), Regional Climate Change Projections, Climate Change 2007: The Physical Science Basis, Contributions of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon et al. eds.], CUP

C4I (Commumity Climate Change Consortium for Ireland) – http://www.c4i.ie

David, H. A. (1981), Order Statistics, Wiley, New York

Dwyer, I.J. & Reed, D.W. (1995), Allowance for discretization in hydrological and environmental risk estimation, IH Report No. 123, Centre for Ecology and Hydrology, Wallingford

FEH (1999), Flood Estimation Handbook, Vol. 2 (Rainfall Frequency Estimation), Centre for Ecology & Hydrology, Wallingford

Feller, W. (1968), An Introduction to Probability Theory and its Applications, Vol. 1, Wiley, New York

Feller, W. (1971), An Introduction to Probability Theory and its Applications, Vol. 2, Wiley, New York

Fitzgerald, D. L. (2005), 'Analysis of extreme rainfall using the log logistic distribution', Stochastic Environmental Research and Risk Assessment (SERRA), Vol. 19, No. 4, pp 249-257, Springer, Berlin

FSR (1975), United Kingdom Flood Studies Report, Vol. 2 (Meteorological Studies), Natural Environment Research Council (NERC), London

Hand W. H. et al. (2004), ' A case study of twentieth-century extreme rainfall events in the United Kingdom with implications for forecasting', Meteor. Appl., Vol. 11, pp15-31

HIRDS (2002) 'High Intensity Rainfall Design System', National Institute of Water and Atmospheric Research (NIWA), Wellington, NZ

HMSO (1981) Handbook of Meteorological Instruments, Her Majesty's Stationery Office, London

Hosking, J. R. & Wallis, J. R. (1997), Regional Frequency Analysis, CUP

Kitanidis, P. K. (1997) An Introduction to Geostatistics, CUP

Kuotsoyiannis, D. et al. (1998), ' A mathematical framework for studying rainfall intensity-duration-frequency relationships', Jnl. of Hydrol., Vol. 206, pp 118-135, Elsevier

Langbein W. B. (1949), 'Annual floods and the partial duration method', Trans. Amer. Geophys. Union, Vol. 30, pp 879-881

Logue, J. J. (1975), Extreme Rainfalls in Ireland, Technical Note No. 40, Met Eireann, Dublin

NOAA (2004), Precipitation Frequency Atlas 14, Vol. 2

R Development Core Team (2006) R: A Language and Environment for Statistical Computing

Ribeiro, P. J. & Diggle, P. J. (2006), **geoR (** version1.6-9), A package for Geostatistical Analysis

Rohan P.K. (1986), Climate of Ireland, Government Sales Publication Office, Dublin

Wang, Q. J. (1990), ' Estimation of the GEV Distribution from censored samples by the method of Partial Probability Weighted Moments, Jnl. of Hydrol., Vol. 120, pp 103-114

Zucchini, W. & Adamson, P.T. (1989), ' Bootstrap confidence intervals for design storms from exceedance series', Hydrological Sciences Jnl., Vol. 34, pp 41-48

Appendix A

Development and Implementation of the Depth-Duration-Frequency(DDF) Relationships

The DDF model applied is based on the median, R(2,D), as the index rainfall; this makes the log-logistic distribution (Fitzgerald,2005) a natural choice for the growth curve since its usual 3-parameter form in terms of F, the cumulative distribution function, is:

$$\frac{R(T,D)-g}{a} = \left(\frac{F}{1-F}\right)^{c} = (T-1)^{c} \quad , a > 0, c > 0, g >= 0 \qquad \text{(A1)}$$

R(T,D) is the rainfall at return period T and duration D, with T =2 (F= 0.5) being the median R(2,D).
Hence *a + g = R(2,D)* i.e. the sum of the location and scale parameters is the median.
For positive random variables, such as rainfall, the location parameter may be taken as the minimum.

There are two sets of annual maxima to consider:

1. 0900-0900 UTC rainfall accumulations for 11 durations between 1 and 25 days

2. the short-duration series for 9 durations ranging 15 minutes to one day, all being sliding durations.

Unlike the 0900-0900UTC daily series, the short-duration data may have missing years as values were recorded only if above preset thresholds. For analysis of individual durations, methods of parameter estimation for left-censored samples were developed. Details are given in Appendix B.

Initially, the 3-parameter log-logistic distribution was fitted to the daily 0900-0900UTC series of annual maxima using the methods outlined in Fitzgerald(2005). This work revealed that for both datasets the probability weighted moment (PWM) and maximum likelihood (ML) estimates of the log-logistic parameters exhibited too irregular a variation over the ranges of duration to be suitable for directly fitting a DDF relation at most stations.
For the short-duration dataset the average over some 40 stations showed the mean of the log-logistic shape parameter increasing, albeit somewhat unsteadily, from 24 hours to a maximum at 1hour, with slightly lower values at 30 and 15 minutes. The scale parameter expressed as a fraction of the median showed little pattern beyond a tendency for the lowest values to occur in the 1 to 4-hour durations. Another feature was that the values of the shape parameter were sometimes undesirably high (>= 0.4) because there were high growth rates at low values of the scale parameter.

For the 474 daily stations the average shape parameter decreased slowly from 1day or 2 days to 25 days. The scale parameter expressed as a fraction of the median showed its highest values at the long durations, with the lowest values in the 1 to 3-day durations but, overall, the variation with duration was erratic.

The first attempt to overcome the problems was to put the log-logistic distribution in growth-curve form:

$$\frac{R(T,D)}{R(2,D)} = 1 + \frac{a(D)}{R(2,D)}\left[(T-1)^{c_D} - 1\right],\ T > 1 \qquad \text{(A2)}$$

The parameters $K_D = \dfrac{a(D)}{R(2,D)}$ and c_D were expressed as functions of duration, typically

$c_D = a + b\ln(D)$ and $\mathrm{K_D} = \mathrm{constant, e + f\ D}$ and $\mathrm{D^s}$.

The sample median was used for R(2, D). Maxima greater than or equal to the median were extracted from the full datasets, put in the form $\dfrac{R(T,D)}{R(2,D)}$ and ordered samples formed. The median plotting position(Flood Studies Report, Vol. 2, 1975) was used for T, with T = 2 for the first row. A table with duration varying across the rows and return period $T \geq 2$ down the columns was then formed. All series of maxima greater than or equal to the median were complete and there was no further need of methods of parameter estimation for censored samples. As suggested by Kuotsoyiannis et al. (1998) the parameters were estimated simultaneously.

In general there was good agreement between the PWM or ML solutions for individual durations and the portmanteau solution, with $c_D = a + b\ln(D)$ and $K_D = e + fD$. However, as with the PWM and ML solutions, the values of c_D and K_D fluctuated markedly between neighbouring stations making it difficult to detect any pattern.

It was decided use a growth curve with the median as scale parameter i.e. to try log-linear relations of form:

$$\frac{R(T,D)}{R(2,1)} = D^s (T-1)^c \quad ,\mathrm{T} > 1 \qquad \text{(A3)}$$

where D = 1 is a suitably chosen unit duration

For variation across the rows this implies

$$\frac{R(T,D)}{R(T,1)} = D^s \qquad \text{(A4)}$$

and for variation down the columns a log-logistic median growth curve:

$$\frac{R(T,D)}{R(2,D)} = (T-1)^c \qquad \text{(A5)}$$

Final form of the 1 to 25-day model

For the shape parameter *c* the earlier work had indicated a slow variation with duration ; this was assumed to be of form

$$c_D = a + b\ ln(\ D) \qquad \text{(A6)}$$

Assuming *s* is constant in equation (A4) in conjunction with (A5) and (A6) gave good results but there was sufficient reduction in the range of the residuals to make it worthwhile to assume *s* to also be of form

$$s_D = e + f\ ln(D) \qquad \text{(A7)}$$

The four parameters were estimated simultaneously using the R routine *lm* and the results were excellent with a coefficient of determination(R^2) in excess of 0.99.
In addition agreement with the log-logistic PWM/ ML quantile estimates and those derived from (A3), A(6) and A(7) were tested for return periods up to 250 years and found good. All four parameters were highly significant in the sense of being much greater than their standard errors.
Note than for $D = 1$ we get $c_1 = a$ in equation A(6) and $s_1 = e$ in (A7).

Use of the daily 0900-0900UTC data for Northern Ireland
Daily data for 1, 2, 4 and 8 days were available and 103 stations with more than 20 years of record were used. Table A1 compares the NI 1-day *c* and *s* values with those derived for the Met Eireann daily stations, using a fitting of equation (A3) to 1,2, 3,4,6,8 days and 1,2,3,4,6,8,10,12,16,20 and 25days, 1 day being the unit duration.

Table A1
Values of the 1-day parameters of the DDF model

	N I (1,2,4,8)		Irl (1,2,3,4,6,8)		Irl (1 to 25 days)	
	c_1	s_1	c_1	s_1	c_1	s_1
Minimum	**0.123**	**0.253**	*0.081*	*0.169*	**0.079**	**0. 176**
Q1	**0.171**	**0.362**	*0.154*	*0.333*	**0.156**	**0.311**
Median	**0.198**	**0.411**	*0.180*	*0.370*	**0.183**	**0.348**
Mean	**0.200**	**0.409**	*0.181*	*0.372*	**0.184**	**0.347**
Q3	***0.227***	***0.453***	*0.200*	*0.410*	***0.205***	***0.377***
Maximum	**0.315**	**0.554**	*0.341*	*0.561*	**0.347**	**0.508**

It is a mild surprise that the central values for NI are the higher. However, the standard deviations of c_1 and s_1 are roughly 0.04 and 0.07 respectively, the durations and periods of record different and so the differences are not highly significant. It is interesting that the 1-day parameter estimates for the Irish 1 to 8 and 1 to 25-day datasets are so similar. This lends some justification to using the NI estimates for the 4-parameter model on the same basis as those of the 4-parameter model derived from the 1 to 25-day dataset but, clearly, estimates of return period rainfalls for durations longer than about 10 or 12 days for locations in NI should be treated with caution.

Effects on the parameter estimates of converting from fixed to sliding durations
Table A1 is based on fixed-duration 0900-0900UTC data. On shifting to sliding durations by multiplying the 0900-0900UTC data by the appropriate conversion factors and refitting the DDF model **the shape parameter *c* of the growth curve is unchanged** and the duration exponent *s* decreases by 0.09.
Implementing the DDF model at grid points for (sliding) durations of 1d to 25d
As described in the main text, station values of the four parameters required by equations (A6) and (A7) and also station values of the median 1d rainfall were extrapolated to a 2km grid. At grid points equation (A3) could then be used to derive the **return period rainfalls** for any duration between 1d and 25d.
Note that all model outputs are for sliding durations. If fixed-duration design rainfalls are needed it is necessary to divide the sliding-duration rainfall by the appropriate factor. The factors are:

1d	**2d**	**3d**	**4d**	**6d**	**8d**	**10d**	**12d**	**16d**	**20d**	**25d**
1.15	1.08	1.06	1.05	1.04	1.04	1.03	1.03	1.02	1.02	1.01

Final form of the 24-hour to 15-minute model

The first matter to resolve concerned the pivotal 24-hour sliding value, *slide24* which is merely a relabelling of the median 1d rainfall. This change of notation is emphasise the near equivalence of the median 1d rainfall to the median of the 24-hour maximum, *abs24,* available at the short-duration stations. To test the effect of using *slide24* instead of *abs24,* the entire column of the ordered 24-hour values and their median plotting positions available at the 39 short-duration stations were extracted from the table. The set for comparison was generated using equation (A5) with:

1. duration D = 1
2. *slide24* instead of *abs24* as R(2,1), the median rainfall
3. 1d or 24-hour shape parameter $c = c_1 = c_{24}$

Fortunately, inspection showed that agreement was good. Testing for one to one correspondence over all the 24-hour data at all the short-duration stations the mean regression coefficient was 1.037 and the coefficient of determination $R^2 = 0.994$. The larger differences were usually at the higher return periods, with the generated values exceeding the recorded values.

The residuals were:

Minimum	1st quartile	Median	3rd Quartile	Maximum
-17.2	-2.3	-0.2	+2.1	28.6

Slide24 and the associated shape parameter c_{24} were available at all the daily stations. *Slide24* was used instead of *abs24* even at the short-duration stations. However, in setting up the short-duration model *abs24* was used as a check on the results obtained using *slide24.*

For the 24-hour (*abs24*) to 15-minute dataset, assuming $c_D = a + b \ln(D)$ and also $s_D = d + e \ln(D)$ again gave the highest coefficient of determination (in excess of 0.97 and often over 0.99). Unfortunately, at some stations, *b* and/or *e* were not well determined. However, assuming *c* or *s* or both constant yielded only slightly lower R^2. Indeed, 2, 3 and 4-parameter models all performed well in comparisons with quantile estimates from (A1) or from the PWM/ML estimates for individual durations. Estimating *c* from the data was to ignore the parameter c_{24} obtained from the daily series; c_{24} had proved successful in generating the series successfully substituted for the recorded values *abs24.*

Using the coefficient of determination as a measure of fit, it was found that nothing was lost by:

1. Using c_{24} = the 1-day shape parameter from the daily data, as the starting value of the shape parameter at the unit (24-hour) duration.
2. The substitution of *slide24*, the median sliding 24-hour value, for the median absolute value *abs24*
3. Assuming the duration exponent *s* constant i.e. not a function of duration giving a DDF model :

$$R(T, D) = R(2,1) D^s (T-1)^{c_{24} + h \ln(D)} \qquad \text{(A8)}$$

Equation (A8) can be regarded as going across the top row using $R(2,D) = R(2,1)D^s$ and then down the column to R(T,D) with shape parameter $c_{24} + h\ln(D)$ applying at duration D . Since $D^s(T-1)^{h\ln(D)} = D^{s+h\ln(T-1)}$ it is fully equivalent to :

$$R(T,D) = R(2,1)(T-1)^{c_{24}} D^{s+h\ln(T-1)} = R(T,1)D^{s+h\ln(T-1)} \qquad \text{(A9)}$$

Equation (A9) corresponds to going down the column (D = 1) to the appropriate value of T and then across the row to get *R(T,D).* Both equations account for nearly all of the variance of the short-duration data (R^2 in the range 0.974 to 0.998 with a median value of 0.995).

The parameter *s* was well-determined and indeed the parameter *h* was statistically significant at most stations. Over the country *s* was well-defined and varied with the value of *slide24;* there did not seem to be any pattern to the variation of *h.*

Linearising equation (A9) the log-log regression was performed with:

1. The *slide24* series in the unit duration column and the recorded values for the eight durations 12 hours to 15 minutes.
2. The recorded values in all 9 columns i.e. *abs24* as the median at unit duration.
3. Using the *slide24* series at the unit duration and just the three shortest durations (1hour, 30 minutes and 15 minutes); this was to examine the ability of the model to express the shorter durations in terms of the 24-hour.

Naturally, this coefficient of determination was lowest but still substantial, with a range over the stations from 0.89 to 0.93.

The parameter *s*

The results were :

Log-log regression estimates of parameter *s*						
	minimum	**1st quartile**	**Median**	**Mean**	**3rd quartile**	**maximum**
slide24	0.315	0.374	0.409	0.412	0.443	0.501
abs24	0.305	0.376	0.402	0.407	0.432	0.496
<=1 hr	0.296	0.369	0.391	0.398	0.425	0.486

The closeness of the three sets of values is highly satisfactory.

The mean values also agree well with the exponent derived from averages of the short-duration data. Considering all the data and the values greater than or equal to the median we get:

Average fraction of 24-hour total for 8 durations (24h = unit duration)								
	12h	**6h**	**4h**	**3h**	**2h**	**1h**	**0.5h**	**0.25h**
All data	0.828	0.648	0.526	0.448	0.361	0.239	0.190	0.158
>= median	0.812	0.613	0.517	0.454	0.380	0.275	0.203	0.150

Assuming *fraction* = D^t , D < 1, we get *t* = 0.411 with $R^2 = 0.993$ for all the data. For the data greater than or equal to the median, *t* = 0.403 with $R^2 = 0.998$.

The parameter *h*

A feature of the short duration data is that at stations such as Birr, having a low median 24-hour rainfall (32mm), the median 1-hour fall is about one third of the 24-hour value; at high return periods this increases to nearly one half implying that the 1-hour fall as a fraction of the 24-hour increases with return period i.e. in (A9) *h* is negative; from (A8) this also implies that the shape parameter increases as duration decreases. This is a key feature to which the model must give expression.

At stations such as Valentia Observatory(median 24-hour = 51mm) the median one hour fall was about one quarter of the median 24-hour but was somewhat less than one quarter at the higher return periods. This can be catered for by having values of *h* which are positive. However, consistency would be very difficult to maintain if positive values of *h* were allowed. It is preferred to impose a value of zero in such cases. In effect this is to regard the highest 1–hour values as too low for the corresponding values of their return periods or to consider the growth rate of the 24-hour maxima as high, as may be seen from the ratio of the sample maximum to the sample median for the full range of durations:

Sample Maximum/Sample Median at Valentia Observatory								
24hrs	**12hrs**	**6hrs**	**4hrs**	**3hrs**	**2hrs**	**1hr**	**0.5hr**	**0.25hr**
2.3	2.0	1.8	1.9	2.0	2.1	1.9	2.3	2.7

The value of *s* at Valentia is 0.43 and at Birr 0.3. Pursuing such considerations the country could be readily divided according to ranges of median 24-hour values, each with a typical value of *s*. Getting a corresponding typical value of *h* proved difficult as it ranged from positive to negative within each *s*-category and was generally small, ranging from +0.034 to -0.058. However, as the multiplier *ln*(D) can be as high as 4.6 its value is important in determining the fraction of the 24-hr rainfall that applies at lower durations:

RP	***s***	***h***	**12hr**	**1hr**	**15min**
50 years	0.38	-0.02	0.81	0.38	0.27
50 years	0.38	0.0	0.77	0.30	0.175

The summary statistics for *h* over all the short-duration stations using both *slide24* and *abs24* values at the unit duration plus the results of using only the 24-hour, 1-hour, 30-minute and 15-minute durations instead of the usual 9 are:

Log-log regression estimates of the parameter *h*						
24h	**minimum**	**1st quartile**	**Median**	**Mean**	**3rd quartile**	**maximum**
slide24	-0.053	-0.018	-0.005	-0.007	+0.005	+0.034
abs24	-0.058	-0.021	-0.009	-0.011	-0.002	+0.017
<=*1 hr*	-0.051	-0.024	-0.011	-0.014	-0.005	+0.014

The standard deviation is about 0.015 in all three cases and so the means are unstable.

For individual durations, values of *h* can be found from the ratios of two return period rainfalls by applying either (A8) or (A9). The return period rainfall values used were $\frac{RP(5)}{RP(2)}, \frac{RP(10)}{RP(2)} \textit{and} \frac{RP(10)}{RP(5)}$ since these could be estimated with reasonable accuracy from the data. Both the method of quartile means (Logue, 1975) and the median plotting position were used to determine the values. The results showed that only for durations of 1 hour or less was *h* consistently negative. Mean and median values of *h* increased slowly with duration between 2 and 12 hours. However, the most strongly negative values of h over all durations occurred at 6, 4 and 3 hours respectively. For durations of 1 hour or less the summary statistics are:

Summary statistics of the parameter *h* for durations of 1hour or less						
Method	**minimum**	**1st quartile**	**Median**	**Mean**	**3rd quartile**	**Maximum**
Quartile means	-0.053	-0.023	-0.010	-0.011	+0.005	+0.030
Plotting position	-0.090	-0.024	-0.009	-0.011	+0.004	+0.044

The value of h applying is fuzzy rather than crisp but a set of values giving satisfactory mappings has been derived using the following guidelines:

1. h depends on median rainfall, being mostly zero or positive at the higher values.
   However there were some negative values of h in this category (>= 47mm). Positive values are not allowed in the model for calculating h at grid points as they imply a decrease in growth rate as duration decreases.
2. Within each category of median rainfall the value of h depends on c_{24}, being low when the shape parameter is high. There was one exception.
3. h also depends on the details of the convective or frontal/convective activity at the station, especially if there is a marked jump between the highest and second highest rainfalls recorded at the shorter durations but not at 24 hours. Examination of individual stations revealed this as the factor causing the more extreme negative values of h.
4. As $s + h\ ln(T-1)$ must be greater than zero, the rule $s > 10h$ was applied.
5. High values of the shape parameter c (>= 0.40) should be avoided.

Attempts to express the dependencies 1 and 2 by regression equations were not successful and the criteria for fixing h were eventually chosen by inspection of mappings of the parameters h, s and c_{24}.

In choosing the final values of h the short-duration stations were divided according to ranges of values of the median 24-hour rainfall: >= 60mm, >= 47mm and < 60mm, >= 35mm and < 47mm and < 35mm. Each of these categories had a characteristic value of the parameter s. 'Best' values of h were then selected by examining its values at stations within the category, keeping in mind both the guidelines and the results in the above tables.

The final scheme is:

slide24 >= 60mm : s =0.48
if(c_{24} < 0.15) h = -0.01, otherwise h = 0

slide24 >= 47mm & *slide24* < 60mm: s = 0.43
if(c_{24} < 0.16) h = -0.015, otherwise h = 0

slide24 >= 35mm & *slide24* < 47mm: $s = 0.375$
if(c_{24} >= 0.25)h = -0.01
if(c_{24} > 0.16 & $c24$ < 0.25) h = -0.015
if(c_{24} <= 0.16) h = -0.023

slide24 < 35mm $s = 0.33$
if(c_{24} >= 0.25) h = -0.015
if(c_{24} > 0.16 & $c24$ < 0.25) h = -0.023
if(c_{24} <= 0.16) h = -0.030

Implementing the DDF model for durations of 24 hours to 15 minutes

The DDF model is given by equation (A8) i.e.
$R(T, D) = R(2,1) D^{s} (T - 1)^{c24 + h \ln(D)}$, with R(2,1) = *slide24*
As grids of *c24* and *slide24* were available, **grids of return period rainfalls** for any duration could be generated using the above scheme to supply values of *s* and *h.* The mappings produced in this way were mostly satisfactory but, especially at durations of 1 hour or less, there were a few anomalous spots indicating sudden jumps in the rainfall values. Both the number and extent of these anomalies were small but to deal with them further smoothing was used. The simple scheme applied at each grid point was to add in the values of the four nearest neighbours, weigh all five equally, and take the mean as the grid point value. While this may seem like over-smoothing, especially at the longer durations, the only visible effect on the mappings was to remove the anomalous spots. In view of this the matter of the degree of smoothing was not pursued further.

Implementation of the DDF model at any location for durations of 25d to 15m
The form of the model considered was as in equation (A5):

$R(T, D) = R(2, D)(T - 1)^{c(D)}$ where $R(2, D) = R(2,1)D^{s}$

For 1d to 25d durations *s*, like *c*, is a function of duration but for sub-daily durations it is not. At grid points *R(2,D*) and *c(D)* can be found for any duration.

Two tasks were addressed:
1. Given duration *D* and return period T, estimate *R(T,D),* the return period rainfall
2. Given duration *D* and amount *R(T,D),* estimate the return period *T*

For a location on an easting or northing gridline the two nearest grid points were used, while for a point within a grid box, the four nearest grid points were used in the interpolations. Weights for the grid points were according to the inverse of the square of their distances from the location.

At a grid point and for durations >= 1d only the grid point itself was used. For sub-daily durations the grid point and its four nearest neighbours were used on an equal footing for the reason indicated in the paragraph above.

To get the estimate of $R(T,D)$ or T at a location two methods of interpolation were examined:

1. Estimate $R(T,D)$ or T at each grid point and take a weighted mean
2. Estimate $R(2,D)$ and $c(D)$ at each grid point and take the weighted means, Then estimate $R(T,D)$ or T from equation (A5)

The second option proved the more satisfactory. Estimates of $R(T,D)$ were practically the same for both but the estimate of T provided by the second option was the more satisfactory and is the scheme underlying the programs.

FEH(1999) discusses the idea of a 'representative point rainfall' for a catchment which raises the question of the most appropriate DDF parameters R(2,D) and c_D. Some averaging of their values over grid points in the catchment is obviously required but whether the averaging should be simple or weighted would depend on the hydrological context.

Appendix B

Estimation of the Parameters of the log-logistic distribution for left-censored samples

For short duration falls only values above preset thresholds were recorded. As a result there was no value in some years. To estimate the parameters it is then necessary to consider that over *n* years *n-r* annual maxima were recorded and *r* were known to be below the threshold. Maximum likelihood(ML) theory lends itself quite well to this type of censoring and the method of partial probability weighted moments (Wang,1990) can also be applied.
For the log-logistic distribution take the usual 3-parameter form (Fitzgerald, 2005). For convenience write the probability distribution function in terms of the cumulative distribution function (CDF) i.e.

$$f(x) = \frac{\frac{1}{ac}\left(\frac{x-g}{a}\right)^{\frac{1}{c}-1}}{\left(1+\left(\frac{x-g}{a}\right)^{\frac{1}{c}}\right)^2} = \frac{1}{ac}F(x)^{1-c}(1-F(x))^{1+c}, \text{where } F(x) = \text{CDF} \qquad \text{(B1)}$$

<u>Maximum Likelihood Solution</u>
For *n* years of record with *n-r* values above the threshold t_0 the likelihood L is:

$$L \propto F(t_0)^r \prod_{i=1}^{n-r} f(t_i) = \frac{1}{(ac)^{n-r}} F(t_0)^r \prod_{i=1}^{n-r} F(t_i)^{1-c}(1-F(t_i))^{1+c}$$

Now $\frac{t_i - g}{a} = F^c(1-F)^{-c}$ and so $l = \ln L$ can be written:

$$l = r\ln F(t_0) - (n-r)\ln a - (n-r)\ln c + \sum \ln F(t_i) + \sum \ln(1-F(t_i)) - \ln\left(\frac{t_i - g}{a}\right)$$

Using

$$\frac{\partial F}{\partial a} = -\frac{1}{ac}F(1-F)\,; \quad \frac{\partial F}{\partial c} = -\frac{1}{c}F(1-F)\ln\left(\frac{F}{1-F}\right); \quad \frac{\partial F}{\partial g} = -f = -\frac{1}{ac}F^{1-c}(1-F)^{1+c}$$

We get from $\frac{\partial l}{\partial a} = 0$:

$$F(t_0) = \frac{n}{r} - \frac{1}{r}\sum_{i=1}^{n-r} 2F(t_i), F(t_i) = \frac{\left(\frac{t_i - g}{a}\right)^{\frac{1}{c}}}{1+\left(\frac{t_i - g}{a}\right)^{\frac{1}{c}}} \qquad \text{(B2)}$$

From $\frac{\partial l}{\partial c} = 0$ $\quad c = \frac{1}{n-r}\sum_{i=1}^{n-r}(2F(t_i)-1)\ln\left(\frac{t_i - g}{a}\right) - \frac{r}{n-r}(1-F(t_0))\ln\left(\frac{t_0 - g}{a}\right)$ (B3)

From $\frac{\partial l}{\partial g} = 0$ $\quad 0 = \sum_{i=1}^{n-r}\frac{1}{t_i - g}\left[1 + \frac{2F(t_i)-1}{c}\right] - \frac{r}{c(t_0 - g)}(1-F(t_0))$ (B4)

Equations (B3) and (B4) can both be used to determine the shape parameter *c*, raising the possibility that, on occasion, it may not be possible to reconcile them i.e. the ML solution may not converge.

This proved to be the case and the method of partial probability weighted moments provided a more robust method of parameter estimation.

Partial Probability Weighted Moment(PPWM) Solution:

Consider the ordered sample of *n-r* values greater than the threshold t_0 from *n* years of record.

The truncated cumulative distribution function is $G(x \mid x > t_0) = \dfrac{F(x) - F(t_0)}{1 - F(t_0)}$.

$E(xG(x)^s)$ is a regular probability weighted moment (Hosking & Wallis, 1997) of the truncated distribution for a sample of size *n-r*. In terms of the data:

$$p(s+1) = E(tG(t)^s) \approx \sum_{i=s+1}^{n-r} \frac{(i-1)(i-2)...........(i-s)}{(n-r)(n-r-1)...(n-r-s)} t_i$$

But by definition $E(tG(t)^s) = \int_{G(t)=0}^{G(t)=1} tG(t)^s dG(t)$

Substituting for *t* and *G(t):*

$$p(s+1) = \int_{F(t_0)}^{1} \frac{(g + a(F(t)/(1-F(t)))^c)}{(1-F(t_0))^{s+1}} (F(t) - F(t_0))^s dF(t) \qquad \text{(B5)}$$

In this form *p(s+1)* is termed a partial probability weighted moment (Wang, 1990). Taking $F(t_0)$ as known, the PPWMs are expressed in terms of incomplete beta functions (Abramowitz & Stegun,1965).

Considering

$$\frac{d}{dt}(t^p(1-t)^q) = pt^{p-1}(1-t)^q - qt^p(1-t)^{q-1} = pt^{p-1}(1-t)^{q-1} - (p+q)t^p(1-t)^{q-1}$$

and taking $\int_u^1$ on both sides we get:

$$-u^p(1-u)^q = pB(p,q,u) - (p+q)B(p+1,q,u) \qquad \text{(B6)}$$

From (B5) and (B6) we can now get from the expressions for *p(1), p(2)* and *p(3):*

$$p(1) - g - \frac{a}{1-F(t_0)} B(1+c, 1-c, F(t_0)) = 0$$

$$cp(1) - cg - K1 = 0$$

$$2c(p(2) + \frac{F(t_0)}{1-F(t_0)} p(1)) - cg(1 + \frac{2F(t_0)}{1-F(t_0)}) - K2 = 0$$ where

$$K1 = 2(1-F(t_0))p(2) - p(1)(1-2F(t_0)) - t_0F(t_0)$$ and

$$K2 = 6(1-F(t_0))p(3) - 4(1-3F(t_0))p(2) - 2(2-3F(t_0))p(1)\frac{F(t_0)}{1-F(t_0)} - 2\frac{t_0F(t_0)^2}{1-F(t_0)}$$

The three equations give:

$$c = (K2 - K1(1+\frac{2F(t_0)}{1-F(t_0)}))/(2p(2)-p(1)) \quad \text{(B7)}$$

$$g = (cp(1)-K1)/c \quad \text{(B8)}$$

$$a = \frac{(p(1)-g)(1-F(t_0))}{B(1+c,1-c,F(t_0))} \quad \text{(B9)}$$

Setting $F(t_0) = 0$ gives the l-moment solution (Hosking & Wallis,1997) in PWM form.

Setting $F(t_0) = 1/2$ implies $g = 0$, $a = t_0 = smed$, the (sample) median. The equation for the shape parameter *c* reduces to:

$$c = \frac{3p(3) - p(2) - p(1) + (smed/2)}{(2p(2) - p(1))} \quad \text{(B10)}$$

Serving as checks we also have:

$$c = \frac{p(2) - (smed/2)}{p(1)} \text{ and } smed = \frac{p(1)}{B(1+c,1-c,1/2)}$$

Appendix C

Checks and Confidence Intervals for the gridded rainfall estimates

General

Here the gridded estimates of return period rainfall are checked against estimates provided by fitting the log-logistic distribution to the data. Further, a crude method of estimating the standard error of quantile estimates for the gridded data is tested.

1d09 to 25d09

Please note that in this section data are fixed duration e.g. daily 0900-0900UTC.

It is quite feasible to write down expressions for the standard errors of quantiles of the 4-parameter model. The correlation matrix is remarkably constant over the stations. However, the values were obtained using $R(2,1)$, the median 1-day rainfall, as if it were a known constant in:

$$\frac{R(T,D)}{R(2,1)} = D^{e+f\ln D}(T-1)^{c_{24}+b\ln D} \tag{C1}$$

where c_{24} is the shape parameter at 24hours i.e. D =1

Since the sample median was used, estimates of the standard error of $R(2,D)$ should allow for the standard error of $R(2,1)$ and the covariance of $R(2,1)$ with each of the four parameters above e.g. $\mathrm{cov}(R(2,1),c_{24}) = -0.0381$ while $\mathrm{cov}(R(2,1),e)$ =+0.0513 over the 577 stations. What this means for a single location is not quite clear but it is assumed that the sign of the inter-station covariance applies. The covariances are small but so are the terms based on $\mathrm{var}(a)$ *and* $\mathrm{var}(e)$ and so strings of small terms are obtained.

Having low confidence in this approach a crude but direct method was adopted. The basic assumption made is that since the model produces quantile estimates similar to those obtained by fitting the log-logistic distribution to individual durations, the model standard error for a given duration is also similar to that derived from log-logistic theory for the individual duration. Now, results of bootstrapping exercises and those of maximum likelihood (ML) theory (Fitzgerald, 2005) can be used. For a given duration rewrite (C1) as:

$$R(T,D) = R(2,D)(T-1)^{c_D} \tag{C2}$$

To scale down R(2,D), the sliding-duration gridded median to its fixed-duration value the following factors from Table A were used:

1d	**2d**	**3d**	**4d**	**6d**	**8d**	**10d**	**12d**	**16d**	**20d**	**25d**
1.15	1.08	1.06	1.05	1.04	1.04	1.03	1.03	1.02	1.02	1.01

ML theory gives:

$$\frac{\mathrm{var}(R(T,D))}{R(T,D)^2} = \frac{\mathrm{var}(R(2,D))}{R(2,D)^2} + \mathrm{var}(c_D)(\ln(T-1))^2 \qquad \text{and}$$

$$\mathrm{cov}(R(T,D),c_D) = 0$$

The negative value of $\mathrm{cov}(R(2,1),c_{24})$ indicates that $\mathrm{cov}(R(T,D),c_D)$ may well be negative, making the assumption of a zero value acceptable in that it increases the variance.

Further for any duration D:

$$\mathrm{var}(c_D) = \frac{9c_D{}^2}{(\pi^2+3)n} \text{ and } \frac{\mathrm{var(R(2,D))}}{\mathrm{R(2,D)}^2} = \frac{3c_D{}^2}{n} \qquad \text{(C3)}$$

Plunging on, it is now assumed that despite the very different methods of arrival at the values of $R(2,D)$ and c_D, the above ML formulae apply to the gridded values with *n*, the number of years of record, set at 41 years i.e. the average over the 577 stations. Now the standard errors (*se(ML))* of the following cases can be examined:

a. The 2p/3p l-moment solution for the series of annual maxima for individual durations with the location parameter regarded as fixed but initially unknown (Fitzgerald, 2005)
b. The solution based on the annual maxima >= the sample median (4-p data) that uses the data for the eleven durations to simultaneously estimate the four parameters.
c. The solution based on interpolated gridded values of the four parameters, assuming 41 years of record (4p-grid)

Note that for cases **b** and **c** the location parameter is zero and the scale parameter is the median. In case **a** the sum of the scale and location parameters is the median.

For the Phoenix Park 122 years of daily data were available and this gave the opportunity to compare it with the period 1941-2004 used in this study.

From the annual maximum series and from the grids of the parameters we get:

Case	Shape parameter	Scale parameter	Location parameter
2p/3p 1881-2002	0.274	22.2	12.0
2p/3p 1941-2004	0.300	20.0	14.1
4p-data	0.224	34.6	0.0
4-p grid	0.239	33.0	0.0

The following table, C1, gives the return period rainfalls and their standard errors derived from annual maximum series (AMS) of 1-day (0900-0900UTC) data for the period 1941-2004. at Phoenix Park. The unit of rainfall is the millimetre.

Table C1						
1-day Return Period Rainfalls and their standard errors						
	Model 2p/3p		**Model**		**Model**	
RP	**Data 1941-2004**		**4p-data 1941-2004**		**4p-gridded**	
years	**Rainfall**	***se(ML)***	**Rainfall**	***se(ML)***	**Rainfall**	***se(ML)***
2	34.1	1.3	34.6	1.7	33.0	2.1
5	44.3	2.4	47.0	2.8	45.9	3.5
10	52.8	3.7	53.6	4.0	55.7	5.2
20	62.5	5.5	66.9	5.6	66.6	7.4
50	78.4	8.9	82.8	8.6	83.5	11.4
100	93.5	12.5	96.9	11.4	98.7	15.4
250	118.8	19.3	119.1	16.4	123.0	22.5
500	143.1	26.4	139.2	21.4	145.2	29.5
1000	172.9	35.8	162.7	27.5	171.4	38.4

The agreement between the three estimates of the return period rainfalls is excellent. For the2p/3p and 4-p models the number of years of record is 64 while for the 4p-grid it is taken as 41. Many of the surrounding stations are full period and so in the case of Phoenix Park n = 41 is likely a bit low. Nonetheless the estimate of the standard error by interpolation from the gridded parameter values is reasonable.
To get an idea of how realistic the quantiles and ML estimates of the standard error in Table C1 are the next table compares bootstrap estimates of the standard error with *se(ML)*, both derived from 122 years of data.

Table C2			
Model 2p/3p			
1-day (0900-0900UTC) data 1881-2002			
RP years	**Rainfall (mm)**	***sd(bootstrap)***	***se(ML)***
2	34.3	1.0	1.0
5	44.4	1.7	1.7
10	52.4	2.5	2.5
20	61.4	3.6	3.6
50	75.8	5.7	5.7
100	89.2	8.2	8.2
250	111.3	13.0	12.1
500	132.1	18.2	16.3
1000	157.2	25.1	21.7

The agreement between the bootstrap and maximum likelihood estimates of the standard error is good. The indications from Table C2 are that even the 500 and 1000-year return period rainfalls derived from the gridded values of the parameters in Table1 are OK even if it does take some 'acclimitisation' to think of 171 ± 38 as being consistent with 157 ± 25. The value of the standard error is that it gives some notion of the local variation e.g. for the gridded values in Table C1, making the usual normal distribution assumption, you might regard the 1000-year return period rainfall for 1-day event as being between 133mm and 209mm in 66% of cases.

At 10 days agreement between the three quantile estimates continues to be very good, with values of the 1000-year return period rainfalls of $283 \pm 45, 260 \pm 32$ and 281 ± 39 for the three estimates. However, between 12 and 20 days and especially at 16 days the individual 2p/3p parameter estimates appear somewhat anomalous. It was decided then to test the 14-day estimates.

For the 14-day (0900-0900UTC) falls during the period 1941-2004 at Phoenix Park the parameter values are:

Case	**Shape parameter**	**Scale parameter**	**Location parameter**
2p/3p 1941-2004	0.216	55.3	38.0
4p-data 1941-2004	0.159	94.3	0.0
4-p gridded	0.175	92.7	0.0

The return period rainfalls and their standard error estimates are:

Table C3							
14-day Return Period Rainfalls and their standard errors							
	Model			**Model**		**Model**	
RP	**2p/3p-data 1941-2004**			**4p-data1941-2004**		**4p-gridded**	
years	**RR**(mm)	***se(b'strap)***	***se(ML)***	**RR(mm)**	***se(ML)***	**RR(mm)**	***se(ML)***
2	93.3	3.0	2.6	94.3	3.2	92.7	4.4
5	112.6	4.4	4.2	117.6	4.9	118.2	6.7
10	126.9	5.9	6.0	133.8	6.7	136.2	9.4
20	142.5	8.0	8.4	150.7	9.0	155.2	12.7
50	166.3	12.1	12.6	175.3	12.9	183.3	18.4
100	187.3	16.6	16.7	196.1	16.5	207.3	23.8
250	220.3	25.0	23.7	227.1	22.3	243.7	32.7
500	249.8	33.7	30.6	253.7	27.7	275.2	41.1
1000	284.1	44.9	39.1	283.4	33.5	310.8	51.1

At the higher return periods the gridded estimates of both the return period rainfalls and standard errors appear a little high but are nearly all within one standard error of the other estimates.

For the 25-day values we have:

Case	**Shape parameter**	**Scale parameter**	**Location parameter**
2p/3p 1941-2004	0.182	98.6	29.1
4p-data	0.144	122.8	0.0
4-p grid	0.161	125.0	0.0

Note that the l-moment solution for 25-days has a lower location parameter than that of 14-days. As the location parameter of the log-logistic distribution bears the interpretation of the absolute minimum, an annual 25-day value of some 29mm appears unrealistically low. The data used so far has been based on a year of April-March. Using January-December annual 25-day maxima we get quite different as the values of the parameters: 0.224, 75.0 and 51.8 respectively. However, Table C4 shows that even the 500-year return period rainfall 2p/3p and grid estimates are not seriously dissimilar given the size of the standard errors.

The sets of parameter values give the following, Table C4, with the Jan-Dec figures in brackets:

Table C4							
25-day(0900-0900UTC) return period rainfalls and their standard errors							
	Model			**Model**		**Model**	
RP	**2p/3p-data 1941-2004**			**4p-data 1941-2004**		**4p-grid**	
years	**RR**(mm)	**se(b'strap)**	**se(ML)**	**RR(mm)**	**se(ML)**	**RR(mm)**	**se(ML)**
2	127.7(126.8)	4.6 (4.4)	3.9 (3.6)	122.8	3.8	125.0	5.4
5	156.0(154.1)	6.2 (6.3)	6.0 (6.0)	149.9	5.7	156.3	8.2
10	176.2(174.5)	7.7 (7.8)	8.4 (8.7)	168.5	7.7	178.1	11.3
20	197.6(196.9)	9.7 (9.7)	11.5(12.2)	187.6	10.3	200.8	15.2
50	229.3(231.2)	13.9 (13.6)	16.8(18.5)	215.1	14.4	233.9	21.6
100	256.7(261.9)	18.4 (18.2)	25.8(24.8)	238.0	18.2	261.9	27.7
250	298.2(310.1)	27.0 (27.3)	30.2(35.7)	271.8	24.3	303.9	37.6
500	334.5(353.6)	35.0 (37.2)	38.2(46.3)	300.4	29.9	339.9	46.7
1000	375.7(404.5)	47.4 (50.5)	47.6(59.6)	332.0	36.3	380.0	63.7

Now the 4p-data model seems to yield low values of both return period rainfalls and of the standard error. The gridded values are in good agreement with the 2p/3p l-moment estimates, especially those derived from the April-March data.

Conclusions for durations of 1d09 to 25d09

For durations of one day or greater the grids of the median 1-day rainfall and of the parameter values can provide estimates of the return period rainfalls together with rough estimates of the standard error; the latter is to provide confidence intervals. Making the usual normal-type assumption you might hopefully say that the 25-day rainfall with a return period of 500 years is 340 ± 47mm i.e. has about a 66% chance of lying between 293mm and 387mm.

Given the broad assumptions made in deriving its values it would hardly be wise to go beyond one standard error in this case. Following established statistical practice any quantile estimate that was not, say, five (even four) times the standard error should be treated with reserve. **All conclusions apply equally well to the sliding-duration estimates and standard errors i.e. 1d to 25d estimates**.

Taken as what they are i.e. annual exceedance probabilities based on the assumption that the annual series 1941-2004 adequately represent the future, the estimates may be used with reasonable confidence for return periods of up to about 500 years. Climate change considerations might be dealt with by adjusting the amounts using safety factors.

Sliding Durations less than 24 hours

The same ideas can be applied to short-duration falls for which the DDF model is:

$$R(T,D) = R(2,D)(T-1)^{c_{24}+h\ln(D)}, D<1, h<=0, c_D = c_{24} + h\ln(D)$$
$$R(2,D) = R(2,24\text{h})D^s, \quad \text{24 hours is the unit duration; R(2,24h) = median 24-hour fall.}$$

For 12-hour falls $D = 0.5$ and so on. The assumption that h is zero or negative implies that the rate of growth increases as duration decreases. Imposing this condition in deriving the parameters *s* and *h* over the grid (see Appendix A) meant that the usual practice of following the data as closely as possible had, in some cases, to be violated. This, taken together with the strong local variability at short durations (Rohan, 1986 & Hand et al. 2004) makes the estimates of the standard errors arrived at for durations shorter than 24 hours even more tentative than those for durations of 1 to 25 days.

The data consisted of 39 Irish stations with a mean record length of n = 44 years and this value was used in determining standard errors since it did not differ much from the value of n = 41 for the daily stations. The eight short-duration stations available from Northern Ireland had much shorter periods of record than the Irish stations and were used mainly to check on the parameter values obtained from the Irish data.

The effect of using grid values is shown in the following three cases for Mullingar, Waterford(Tycor) and Claremorris:

Mullingar

At Mullingar the scheme making the parameters h and s functions of the median 24-hour rainfall had an unusually strong effect on both the shape parameter c_D and on the exponent s and is the most extreme case encountered.

This is illustrated by comparing the grid estimates of the shape and scale parameters for 1 hour with those estimated from the data i.e. PPWM or PWM solution(see Appendix B) for the annual series of 1-hour maxima and the log-log regression solution of the DDF model based on data >= the median rainfall for durations between 24 hours and 15 minutes.

	c_{24}	$median_{24}$	h	s	c_{1h}	scale	location
DDFgrid	0.224	33.9	-0.023	0.330	0.297	11.9	0.0
DDFdata	0.234	34.3	-0.0011	0.367	0.237	10.7	0.0
PPWM	-------	--------	---------	-------	0.298	06.4	4.6

The grid has high values of both the shape parameter c_{1h} and of the scale parameter, R(2,1h), the median 1-hour fall. The three sets of estimates are:

1-hour Return Period Rainfalls at Mullingar (mm)				
RP(years)	**PPWM**	**DDFdata**	**DDFGrid**	**Grid se(ML)**
2	11.0	10.7	11.9	0.9
5	14.3	14.9	18.0	1.7
10	16.9	18.0	22.9	2.6
20	19.8	21.5	28.5	3.8
50	24.9	26.9	37.8	6.2
100	29.7	31.8	46.6	8.8
250	37.5	39.6	61.3	13.5
500	43.4	42.8	75.3	16.6

At the 100-year return period the PPWM and DDF-data estimates agree well and the grid estimate appears high. In 48 years of recording short-duration rainfalls the 1-hour maximum at Mullingar was 34.5mm for which the median plotting position is over

68 years. While the 100-year return period rainfall estimate of 46.6mm is high the other two estimates appear a bit low. Also, if the rule of thumb of being chary of estimates that are not at least five times as large as the estimate of the standard error, then the grid estimate of 61mm for 250 years is suspect. Notwithstanding, it is more comfortable to consider a 66% chance of the 250-year return period rainfall for a sliding duration of 1 hour as lying between 48mm and 75mm than the 30.5mm to 44.5mm suggested by the PPWM estimate.

Waterford(Tycor)

	c_{24}	$median_{24}$	h	s	c_{1h}	scale	location
DDFgrid	0.180	44.4	-0.015	0.375	0.228	13.5	0.0
DDFdata	0.185	45.7	-0.025	0.433	0.265	11.5	0.0
PPWM	-------	--------	---------	-------	0.286	10.3	1.4

The grid value of the 1-hour median rainfall is high but, by way of compensation, the shape parameter is low and brings the quantile estimates into line with those based directly on short-duration data, as may be seen from the following table:

1-hour Return Period Rainfalls at Tycor (mm)				
RP(years)	**PPWM**	**DDF-data**	**DDF-Grid**	**Grid se(ML)**
2	11.7	11.5	13.5	0.8
5	16.7	16.6	18.5	1.3
10	20.7	20.6	22.3	2.3
20	25.3	25.1	26.4	2.7
50	32.7	32.3	32.8	4.2
100	39.7	38.9	38.5	5.6
250	51.3	49.6	47.5	8.0
500	62.3	59.7	55.7	10.5

The estimate from the gridded values of the DDF parameters of 5.6 as the standard error of the 100-year return period rainfall appears low but then the PPWM estimate is only 7.4 while the DDF-data estimate is 6.5.

Claremorris

	c_{24}	$median_{24}$	h	s	c_{1h}	scale	location
DDFgrid	0..200	37.5	-0.015	0.375	0.248	11.4	0.0
DDFdata	0.202	38.8	-0.009	0.394	0.231	11.1	0.0
PWM	-------	--------	---------	-------	0.369	5.2	6.1

Claremorris is interesting because the PPWM solution has a high value of the shape parameter. The 1-hour data is complete for the 47 years of record and so the solution is a PWM and a not a PPWM solution. The effect of the grids has been to increase both the median 1-hour rainfall and the shape parameter over the values suggested by the DDF-data model.

The table of estimates is:

1-hour Return Period Rainfalls at Claremorris (mm)				
RP(years)	**PWM**	**DDF data**	**Grid**	**Grid se(ML)**
2	11.3	11.1	11.4	0.7 (0.5)
5	14.8	15.3	16.8	1.3 (1.0)
10	18.0	18.4	19.7	1.9 (1.6)
20	21.5	21.9	23.7	2.7 (2.4)
50	28.0	27.3	29.9	4.1 (4.3)
100	34.4	32.1	35.6	5.6 (6.4)
250	45.9	39.7	44.8	8.3 (10.8)
500	57.6	46.6	53.2	10.9 (15.8)

The two highest values recorded at Claremorris were 34.6mm and 24.9mm with median return periods of 68.4 years and 28 years respectively.

The agreement between the estimates is quite good. Again, using the rule of thumb for the standard error on the gridded values, the 500-year year value is suspect even though the agreement between the estimates is good. In brackets is the bootstrap estimate of the standard error from the 1-hour data which looks the more realistic at the 250 and 500-year return periods.

<u>Conclusions for durations between 24 hours and 15 minutes:</u>

Grid estimates of short duration falls may be used with reasonable confidence up to about the 250-year return period rainfall. However, it does not seem sensible to consider an amount that should occur on average once in the 250 years when the precipitation climate 50 years hence is unknown. It makes more sense to interpret the 250-year value as the amount that has nearly a 10% chance of being exceeded at least once during the next 25 years. The standard error estimates must be considered tentative.

Appendix D

Langbein's Formula

For rainfall events above a certain (high) threshold assume that:

1. they occur on average once in *n* years i.e. the average recurrence interval (ARI) is *n*
2. the process of occurrences is Poisson (Feller, 1968)

Then the probability of at least one event in a given year $= 1 - e^{-\frac{1}{n}}$

Average interval between years with at least one event = $\dfrac{1}{1 - e^{-\frac{1}{n}}}$ = T

where T is the return period. Rewriting in terms of ARI we get:

$$T = (1 - \exp(-ARI^{-1}))^{-1} \qquad \text{(D1)}$$

ARI =10 is equivalent to T =10.51 and so when only series of annual maxima are available we can approximate the partial duration series rainfall for ARI = 10 by the return period rainfall for T = 10.51.

To get an idea of the size of the effects of using equation (D1) assume that the growth curve of the return period rainfall is log-logistic and then we have:

$$\frac{R(T,D)}{R(2,D)} = (T-1)^c \quad ,T = \textit{ return period}, T = 2 = \textit{median}, D = \textit{duration}$$

The exponent *c* is positive and the range of values encountered increases as duration decreases. Choosing a typical value for each duration we get:

Table D1

Factor converting Return Period rainfall to PDS rainfall with ARI = T

Duration	**Typical Growth Curve Exponent**	**T = Return Period(years)**					
		2	**5**	**10**	**25**	**50**	**100**
25d	0.100	1.044	1.012	1.006	1.002	1.001	1.001
8d	0.130	1.058	1.016	1.007	1.003	1.001	1.001
4d	0.150	1.067	1.018	1.008	1.003	1.002	1.001
1d	0.190	1.085	1.023	1.010	1.004	1.002	1.002
6hr	0.200	1.090	1.025	1.011	1.004	1.002	1.002
1hr	0.220	1.100	1.027	1.012	1.005	1.002	1.002
1/4hr	0.235	1.107	1.029	1.013	1.005	1.002	1.002

Some credence is lent to these rather speculative computations by the results in NOAA Atlas 14, Vol. 2 which deals with the Ohio River basin. Analysing both PDS and AMS data separately they find values for the 24-hour duration that are quite close to those in the table.

Short Average Recurrence Intervals(ARI)

For AMS the median corresponds to T = 2; values of T between 1 and 2 years are not clearly meaningful. However, ARI values even those less than one year are meaningful. As noted above, Langbein's formula can be used for fractional ARI values and provides a means of estimating PDS rainfalls for ARI less than 2.

When AMS data are available Jenkinson's quartile method (FSR, 1975, Vol.2) provides another method of estimating PDS rainfall for low ARI. ARI = ½ corresponds to the mean of first quartile(mean of the lowest 25%) of the ordered annual series, while the mean of the second quartile(25% to 50%) serves for ARI = 1. The geometric mean is usually employed in preference to the arithmetic mean in quartile analysis of rainfall (Logue, 1975).

Appendix F

Falls of durations less than 15 minutes

Dines recorders provided the bulk of the data for falls of durations lower than 24 hours. These recording rain gauges are not accurate below 15 minutes (Logue, 1975). Nonetheless, estimates of the daily maximum rates over a period of about 5 minutes have been extracted at Irish synoptic stations for periods of 30 to 50 years. Maximum values at the stations range 8mm to 12mm.

The short-duration model properly applies to durations between 24 hours and 15 minutes. It can be extrapolated to lower durations but estimates for durations shorter than 5 minutes might well be unrealistic. Indeed, as the shape parameter controlling the rate of growth increases as duration decreases there is a need to check that the 10-minute and 5-minutes values are realistic. Employing the hydrological concept of a rainfall profile of given duration (FEH(1999), Vol. 2,ch. 4), regard the 5-minute and 10-minute falls as sub-periods of a 15-minute rainfall event with total rainfall set to one i.e. take our 15-minute estimate as correct. Statistics can then supply reasonable estimates for the expected maxima of sub-periods by way of the distribution of the maxima of *n* random divisions of the unit interval (Feller, 1971; David, 1980). The central values of this distribution are:

Table E1									
Mean and Median of the distribution of maxima of *n* divisions of a unit interval									
N	2	3	4	5	6	8	10	12	15
Median	0.75	0.59	0.50	0.44	0.39	0.32	0.28	0.25	0.21
Mean	0.75	0.61	0.52	0.46	0.41	0.34	0.29	0.26	0.22

The mode of the distribution is less than the median and the standard errors for $n = 2$, 8 and 12 are 0.14, 0.09 and 0.07 respectively. The equation

$$\text{mean} = D^{0.37515-0.0683\ln(D)}, \text{ D } = \text{fraction of unit duration}$$

is almost exact.
For comparison, NOAA Atlas14, Vol. 2, Table 4.1.4 gives figures for sub-periods of 60-minute accumulations, based on data with a resolution of 5 minutes, that agree well with the theoretical values above.

For unimodal events the equation provides a profile e.g. for a 15-minute event with total fall = 1, the mean maximum 10-minute fall (D= 0.666) is 0.85 and the 5-minute (D = 0.333) is 0.61 and is independent of return period. Also the matter of where in the unit period the maxima are located is sidestepped.
By way of comparison, directly using the DDF model to generate summary statistics for return periods between 2 and 250 years gives mean values of 0.856 to 0.875 for the 10-minute fall and 0.66 to 0.70 for the 5-minute fall. The mean of the DDF model values on the 2km grid at the 50-year return period is 12.1mm while the 15-minute mean multiplied by 0.61 is 10.6mm which more in accord with the 8 to 12mm maxima recorded over 30 to 50 years. Hence, for durations of less than 15 minutes the indication is that it is probably better to employ 15-minute estimates and apply the formula for the mean fraction as a function of the fractional duration. Estimation for intervals of less than 5 minutes duration should be regarded as highly speculative.

Appendix F

Glossary of terms used

α-level confidence interval – Bounds within which an estimate lies with percentage probability of α %.
AAR - Average annual rainfall total; the period 1961-1990(**AAR6190**) was used in this study.
AEP (**Annual exceedance probability**) – for a given duration it is the probability of exceeding a given rainfall amount at least once in any year.
AMS (Annual Maximum Series) – Time series containing the largest value in each year (12-month period) of record for a particular duration.
ARI(Average Recurrence Interval) – for a given duration it is the average time span between exceedances of a preset rainfall total.
CDF (Cumulative Distribution Function) - The CDF, $F(x)$, is the probability of a value of the random variable X being less than or equal to a number x.
Convective activity- In meteorological terms it refers to the generation of showers in air caused to rise by heat transfer from the surface of the earth.
C4I (Community Climate Change Consortium for Ireland) – based in the headquarters of Met Eireann, the Irish National Meteorological Service, this project may be accessed at: http://www.c4i.ie
Dines Rainfall Recorder – A tilting siphon rainfall recorder of British Meteorological Office design that produces a record of rainfall over time.
DDF(Depth-Duration-Frequency) Model – Method of estimation of a rainfall amount as a function of duration and of frequency; frequency is usually expressed in terms of return period T (see below).
The basic components of a DDF model are an index rainfall and a growth curve (see below) that provides the estimate as some multiple of the index rainfall.
Easting and northing – co-ordinates of a location expressed as the distance eastwards and the distance northwards from a fixed datum.
Fixed-Duration Rainfall - Rainfall accumulation between fixed hours e.g. 24-hour total read at 0900UTC each day.
Frontal Activity- Rainfall caused by air motions induced by contrasts across zones dividing air masses of differing characteristics.
General Circulation Model – Computer model of the interactions of the atmosphere with the earth-sun system used for the prediction of climate change.
Geometric mean - n^{th} root of the product of a sample of n values of a positive variable such as rainfall.
Growth Curve – A formula giving the increase of rainfall with return period (see below) for a known duration. It provides the factor by which the index rainfall is multiplied in DDF relationship.
Index Rainfall – A suitably chosen value, usually a central value such as the median or the mean rainfall, that is multiplied by a growth factor in a DDF relationship.
Interpolation- Any method of computing new data points from a set of existing data points.
Kriging – method of inferring a value at a location with no data as a distance-weighted average of the data at neighbouring locations.
Met Eireann- Irish National Meteorological Service http://www.met.ie

NOAA- Hydrological Design Studies Centre http://www.nws.noaa.gov/oh/hdsc
OPW- Office of Public Works http://www.opw.ie
HIRDS- High Intensity Rainfall Design System (New Zealand) www.niwa.co.nz/ncc/tools/hirds
L-moments- moments computed from linear combinations of the ordered sample values that provide summary statistics such as dispersion and skewness and are often more efficient than ordinary moments in parameter estimation of probability distributions.
L-moments are intimately connected to probability weighted moments (see below).
Location parameter - A value subtracted from or added to the variable *x* to translate the graph of its probability distribution along the *x*-axis.
PDS(Partial Duration Series) - For a given duration it is the series of all events during the period of record that exceed a preset threshold together with their times of occurrence. Also known as **POT (**Peak over threshold) series.
Poisson process- The particular process of occurrences for which the occurrence of an event does not affect the probability of occurrence of the next event.
POT (Peak over Threshold) Analysis – Statistical analysis of partial duration series
Probability distribution function- For a continuous random variable *x* it yields the relative frequency or probability of occurrence of *x* over all subsets of its range of values.
PWMs(Probability weighted moments) - Certain weighted linear functions of the ordered sample data that statistical theory shows as both useful and efficient for parameter estimation of probability distributions. L-moments (see above) are a development of PWM theory. **Partial probability weighted moments(PPWMs)** are another development of PWMs and are used for censored samples.
Quantiles - Values taken at regular intervals from the CDF(see above) of a continuous random variable
For an ordered sample $(x_1, x_2, ...x_n)$ the *f*-quantile is the data value *x* with the fraction *f* of the data less than or equal to *x;* it may correspond to a sample value x_i or it may be interpolated between two successive sample values. An *f*-quantile divides an ordered sample into *f* (approximately) equal sets of values.
Quartiles – For an ordered sample they are the quantiles for which *f* assumes values 25%, 50% (the sample median) and 75% i.e. they are the three points dividing the sample values into four groups, with 25% of the observations in each group. The **interquartile range** is the difference between the values of the 3rd quartile and the 1st quartile and is a measure of dispersion.
Residual – Observed value minus the value estimated by a model.
Return Period (T) - Average number of years between years with rainfalls exceeding a certain value. It is the inverse of the AEP defined above e.g. a 50-year return period corresponds to an AEP of 0.02. Its importance stems from its being a basic component of the DDF model used to calculate the return period rainfall.
Return Period Rainfall –The estimate supplied by the DFF relation when the return period T, the duration D and the model parameters are supplied. It is, perhaps, best thought of in risk terms e.g. the 100-year return period rainfall has nearly a 10% chance of being exceeded in a 10-year period.
R^2 **(Coefficient of Determination) -** The proportion of the sum of squares variability accounted for by a (regression) model.
Scale parameter- Divisor controlling how spread out the distribution is e.g. the median is the scale parameter of a two-parameter log-logistic distribution.

Shape parameter - Power law exponent whose variation changes the shape of a distribution e.g. for the log-logistic distribution the shape parameter determines the position of the peak probability and the skewness.
Skewness – A measure of the departure from symmetry of a distribution
Sliding duration rainfall – Term used for the rainfall total for a given duration when extracting maxima from (effectively) continuous rainfall records; it often exceeds and cannot be less than the corresponding fixed duration rainfall, as for the latter the hour of reading is preset.
Standard deviation- Measure of dispersion or spread of values about their mean
Standard error – Estimated standard deviation of a sample statistic such as the mean i.e. the standard deviation of the sampling distribution of the mean
Tipping Bucket Recorder- A cylinder funnels precipitation on to one of two balanced buckets of fixed capacity. On reaching capacity the receiving bucket tips, sending a signal to a recorder while the second bucket becomes the receiver.
Unimodal- Having one maximum on its probability distribution curve i.e a single peak
UTC - Coordinated Universal Time or Universal Time Coordinated (UTC) is the international time standard. It is the current term for what was commonly referred to as Greenwich Meridian Time (GMT).

List of Met Éireann Technical Notes

44. Lynch, Peter, 1984: DYNAMO---A one dimensional primitive equation model.

45. Lynch, Peter, 1984: Initialization using Laplace Transforms.

46. Lynch, Peter, 1984: Initialization of a barotropic limited area model using the Laplace Transform technique.

47. Leech, L~S, 1985: A provisional assessment of the recreational quality of weather in summer, in terms of thermal comfort and the adverse effect of rainfall.

48. Lynch, Peter, 1986: Numerical forecasting using Laplace transforms: Theory and application to data assimilation.

49. Hamilton, J, P Lennon and B O'Donnell, 1987: Objective analysis of climatological fields of temperature, sunshine and rainfall.

50. Lynch, Peter, 1987: The Slow Equations. Part I: Derivation and Properties of the System. Part II: Application to Continuous Data Assimilation. rainfall.

51. Logue, J J, 1989: The estimation of extreme wind speeds over standard terrain in Ireland.

52. Lynch, Peter, 1991: Filtered Equations and Filtering Integration Schemes.

53. McDonald, A, 1991: Semi-Lagrangian methods.

54. Lynch, Peter, 1996: The elastic pendulum: a simple mechanical model for atmospheric balance.

55. McDonald, A, 1998: The origin of noise in semi-Lagrangian integrations.

56. Lynch, Peter, 1999: Resonant Motions of the Swinging Spring.

57. McDonald, A, 2001: A step toward transparent boundary conditions for meteorological models.

58. Lynch, Peter, 2001: Resonant Rossby Wave Triads and the Swinging Spring.

59. Sheridan, Tom, 2001: Analysis of trends at some Irish rainfall stations.

60. McDonald, A, 2004: Transparent lateral boundary conditions for baroclinic waves I: the foundations.

61. Fitzgerald, D L, 2007: Estimation of Point Rainfall Frequencies.