Regional Estimation from Spatially Dependent Data

R.L. Smith
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599-3260, USA

December 4 1990

Summary

Regional estimation methods are used in such fields as hydrology and meteorology, for estimating parameters such as return values (or quantiles) of a distribution, when data are available at many sites in a region. It is assumed that the desired parameters are either common to all the sites, or else dependent on a set of covariates through a model which is common to all the sites. However, even if we have the correct model for the *marginal* distribution of the data at each site, conventional methods of estimation will not take into account the dependence between the sites. In the past, this problem has been recognised but no clear-cut solution has emerged. In the present paper, the method of maximum likelihood is used to estimate the parameters, the likelihood function being constructed as if the sites were independent, but a new method is proposed for obtaining standard errors which allows for spatial dependence. The method is extended to obtain critical values of likelihood ratio tests. An example is given for a network of rainfall data in South-East England. The validity of the results is confirmed by a small simulation study.

1. Introduction

Consider the following problem. The data consist of a large number of time series, though each series may contain only a relatively small number of observations. We are interested in some parameter of the marginal distributions of the series. This parameter could be estimated from any one of the series but, because the individual series are short, is unlikely to be estimated accurately in this way. Therefore, it is proposed to estimate the parameter by combining the data in different series, under a model which assumes that the parameter is either common to all the series, or is related to covariates through a common regression relationship. However, the series are dependent. How should we take account of this dependence in estimating the parameter?

This encapsulates the problem of *regional estimation* in hydrology. The parameter of interest is often taken to be the N-year *return level* of a variable at a given site, i.e. the $(1 - 1/N)$-quantile of the distribution of annual maximum, and the method of estimation may be based on fitting one of the extreme value distributions (Leadbetter, Lindgren and Rootzén, 1983; Galambos, 1987) to the annual maxima at each site. However, it is common for less than 25 years' data to be available at each site, while the N-year return value is required for N of the order of 50 or 100. To get around this difficulty, *regional methods* have been developed which combine all the data from a region. The regions may be defined in a number of ways – they may be geographical regions or they may be groupings of stations based on common catchment characteristics – but the idea is to define the regions so that sites within each region are approximately homogeneous. The Flood Studies Report (NERC 1975) is a standard reference for this technique, though there are numerous more recent ones including Moore (1987), Arnell and Gabriele (1988), Dales and Reed (1989) and Cunnane (1989) to name but a few.

Buishand (1990) has proposed a model based on combining the data

at all sites into a single likelihood function and estimating the parameters by maximum likelihood. This method, like most others in current use, ignores inter-site dependence. Stedinger (1983) examined the effect of inter-site dependence, and concluded that the main effect was to increase the variance of estimates of margin parameters, compared with the independent case. Buishand (1984) was the first to propose a *model* for spatial dependence, loosely based on bivariate extreme value theory, but his method was confined to two sites and was not entirely addressed to the problem we are considering here. Hosking and Wallis (1988) conducted a simulation study which confirmed Stedinger's results, but also concluded that the effect of inter-site dependence is less important than that of heterogeneity of the margin parameters from site to site. However, this result appears to depend on the model being used to simulate dependence, and the conclusion has not been universally accepted. Buishand (1990) proposed a method of correcting for spatial dependence in the distribution of the number of exceedances over a high threshold, but did not consider how parameter estimation would be affected by dependence. Dales and Reed (1989) proposed a method of estimating spatial dependence based on the notion of an equivalent number of independent sites, and Reed and Stewart (1989) proposed a variant on standard graphical methods of estimation which incorporated this concept. There have also been more theoretical attempts to model spatial dependence, for example by Coles and Tawn (1990) and Smith (1990). However, none of these techniques leads to an explicit method of estimation taking into account the spatial dependence.

The method proposed in this paper is in principle applicable for any problem in which it is the *marginal* distributions which are of interest. In this context a viable approach to obtain point estimates is to calculate maximum likelihood estimates under the artifical assumption that the series are independent. The standard errors obtained under the independence assumption are likely to be far too small. The method outlined in Section 2 allows

us to correct for that and to quote realistic standard errors. An extension of the methodology allows it to be applied also to likelihood ratio tests. In Section 3, the method is applied to rainfall data consisting of up to 21 data points at each of 405 sites. In this example there is heavy spatial dependence and standard error estimation under an assumption of independence is completely misleading. Finally, a simulation study in Section 4 suggests that the method still tends to underestimate standard errors to some extent, but is nevertheless an enormous improvement on assuming independence of the series.

2. Accounting for spatial dependence

Estimation is based on maximising a likelihood function constructed under an assumption of independence among series. However, this is not the true model, so the maximum likelihood estimates do not have the usual properties. It is nevertheless possible to think of the likelihood equations as a set of estimating equations for the unknown parameters. This idea underlies the proposed method and in particular equations (2.1) and (2.9) below. As a side remark, we note that the same method could be used to treat other estimating-equation approaches, such as the probability-weighted moments method proposed by Hosking, Wallis and Wood (1985).

Suppose there are n years' data at each station and the log likelihood function from all stations combined, assuming independence of the stations, is of the form $\ell_n(\theta) = \sum_i h_i(\theta)$, where h_i is the contribution to ℓ_n from year i and θ is the vector of unknown parameters. Let $\widehat{\theta}$ denote the maximum likelihood estimate and θ_0 the true value. Taylor expansion of the log likelihood yields the standard approximation

$$\widehat{\theta} - \theta_0 \approx \{- \bigtriangledown^2 \ell_n(\theta_0)\}^{-1} \bigtriangledown \ell_n(\theta_0)$$

where $\bigtriangledown$ and $\bigtriangledown^2$ denote gradient and hessian respectively. Approximating the components of the hessian by their expected values (denoted E), we

have

$$\text{cov } \widehat{\theta} \approx H^{-1} V H^{-1} \tag{2.1}$$

where $H = -E \bigtriangledown^2 \ell_n(\theta_0)$ and $V = \text{cov}\{\bigtriangledown \ell_n(\theta_0)\}$.

Now, if the assumed model were correct (i.e. if the series really were independent) then we would have $V = H$ and we would be left with the conventional approximation

$$\text{cov } \widehat{\theta} \approx H^{-1}. \tag{2.2}$$

in which H may be approximated by the *observed information matrix*

$$H \approx - \bigtriangledown^2 \ell_n(\widehat{\theta}). \tag{2.3}$$

which is computed directly from the data. In practice we use a quasi-Newton routine which does not calculate the hessian exactly but gives a good approximation to it.

Now, however, suppose that the series contributing to ℓ_n are not independent but the contributions h_i from separate years are. Suppose also the h_i have a common distribution. Writing

$$\bigtriangledown \ell_n(\theta) = \sum_i \bigtriangledown h_i(\theta)$$

expresses $\bigtriangledown \ell_n$ as a sum of n independent terms. We estimate the covariance matrix of $\bigtriangledown h_1(\theta_0)$ using the empirical covariance matrix of the observed $\bigtriangledown h_i(\widehat{\theta})$ $(i = 1, ..., n)$ and then use

$$V = \text{cov}\{\bigtriangledown \ell_n(\theta_0)\} = n \text{ cov}\{\bigtriangledown h_1(\theta_0)\}. \tag{2.4}$$

Combined with the observed information matrix to estimate H, we have a numerical approximation to (2.1) itself. This allows us to compute approximate standard errors.

This calculation assumes that contributions from year to year are independent and identically distributed. In this case of hydrological data this is a reasonable assumption, since the effect of inter-site dependence is much greater than the dependence from one year to the next. Nevertheless, provided we still assume stationarity, it is possible to go beyond (2.4) by including correlations between successive h_i's. A first-order correction may be obtained by including the (estimated) correlations between the components of h_1 and h_2. A second order correction may be obtained by including also the correlations between h_1 and h_3, and so on. In the example to be discussed in Section 3, it was found that the first-order correction led to estimated standard errors that were not much different from those obtained directly from (2.4), while the second-order correction produced an estimated V that was no longer positive definite! The difficulty is presumably due to the error in estimating correlations from a short (21-year) data series. We may conclude that the method based on (2.4) is adequate for the data being considered, though with a longer time series it might be appropriate to explore this aspect more fully.

Now let us consider the consequences of the above analysis for likelihood ratio tests. Suppose $\theta = (\phi, \psi)$ where ϕ and ψ are of dimensions p and q respectively. We wish to test $H_0 : \ \phi = \phi_0$ against $H_0 : \ \phi \neq \phi_0$, ψ being a nuisance parameter. Suppose $\widehat{\theta}$ and $\widehat{\widehat{\theta}}$ are the maximum likelihood estimates of θ under H_0 and H_1 respectively. Taylor expansion of ℓ_n yields

$$2\{l_n(\widehat{\widehat{\theta}}) - l_n(\theta_0)\} \approx g^T H^{-1} g \tag{2.5}$$

where H is as before and $g = \bigtriangledown \ell_n(\theta_0)$. Suppose we partition $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$ with H_{11} $p \times p$ etc., and write the inverse $H^{-1} = \begin{pmatrix} H^{11} & H^{12} \\ H^{21} & H^{22} \end{pmatrix}$. Then it can be seen that under H_0

$$2\{l_n(\widehat{\theta}) - l_n(\theta_0)\} \approx g^T \begin{pmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{pmatrix} g. \tag{2.6}$$

Combining (2.5) and (2.6),

$$2\{l_n(\widehat{\widehat{\theta}}) - l_n(\widehat{\theta})\} \approx g^T C g \tag{2.7}$$

where

$$C = \begin{pmatrix} H^{11} & H^{12} \\ H^{21} & H^{22} - H_{22}^{-1} \end{pmatrix}. \tag{2.8}$$

From a standard formula for the inverse of a partitioned matrix, it is also possible to write $H_{22}^{-1} = H^{22} - H^{21}(H^{11})^{-1}H^{12}$ which, when substituted in (2.8), makes it clear why C is of rank at most p. However, this version has not been used in computation.

If V is the covariance matrix of g, then we may write $g = V^{1/2}z$ where z is a $(p+q)$–vector of independent standard normal variates and $V^{1/2}$ is the positive definite square root of V, and hence $g^T C g = z^T V^{1/2} C V^{1/2} z$ which has the same distribution as

$$\sum_i \lambda_i z_i^2 \tag{2.9}$$

where $\lambda_1,\ \lambda_2, ...$ are the eigenvalues of $V^{1/2}CV^{1/2}$. Since C is of rank p, we may order the eigenvalues so that only the first p terms in (2.9) are non-zero, and then we have expressed the null distribution of the likelihood ratio statistic as a weighted sum of p χ_1^2 variates. In the case of classical maximum likelihood, of course, we have $\lambda_1 = ... = \lambda_p = 1$ and the distribution reduces to the usual χ_p^2.

All of this is quite easy to implement computationally. The matrix V has been diagonalised using Jacobi's method to calculate eigenvalues and eigenvectors, and this has been used to compute $V^{1/2}$. Then the same routine is used to compute the λ_i's and the percentage points of the likelihood ratio statistic may be computed by the algorithm of Farebrother (1984).

3. An example

We consider 21 years' data on annual maximum of one-day rainfall (in units of 0.1 mm.) at each of 405 sites in South-East England. Dales and Reed (1989) have discussed these and related data in some depth. Allowing for some missing data, there are in all 8259 data points. In addition to the annual maxima of rainfall, also recorded are the E-W and N-S coordinates of the site of the rain gauge and the SAAR (Standard Annual Average Rainfall) and altitude at each site.

A widely used model for annual maxima is the Generalised Extreme Value model

$$F(y) = \exp\left[-\left\{1 + \frac{\xi(y-\mu)}{\sigma}\right\}_+^{-1/\xi}\right] \tag{3.1}$$

in which μ, σ and ξ represent location, scale and shape parameters. The simplest model is to assume that these three parameters are common to all the data sites. This, however, seems likely to be too simple an assumption, so we also consider models in which μ, σ and ξ depend on covariates. Dales and Reed (1989) showed that, of all the known covariates, SAAR has by far the highest correlation with annual maximum rainfall, so initially we consider just this. The SAAR has been standardised to have mean 0 and variance 1 over the 405 stations, and is then denoted x. A general model is of the form

$$\mu = \sum_{j=0}^{n_\mu} \mu_j x^j, \quad \sigma = \sum_{j=0}^{n_\sigma} \sigma_j x^j, \quad \xi = \sum_{j=0}^{n_\xi} \xi_j x^j, \tag{3.2}$$

where n_μ, n_σ and n_ξ represent model orders which must be determined. Six models were fitted with the following values of negative log likelihood (NLLH):

Model	n_μ	n_σ	n_ξ	No. Pars.	NLLH
1	0	0	0	3	50094.7
2	1	0	0	4	49919.0
3	1	1	0	5	49918.1
4	1	1	1	6	49917.9
5	2	0	0	5	49917.1
6	3	0	0	6	49911.6

Initial inspection of these figures suggests that Model 2 is a great improvement on Model 1, that Models 3-5 do not improve significantly on Model 2, but that Model 6 should be examined further.

In fact, a more detailed analysis by the method of Section 2 confirms that models 1, 3, 4 and 5 need not be considered further, so we concentrate on models 2 and 6. For model 2, the parameter estimates are $\widehat{\mu}_0 = 316.1$, $\widehat{\mu}_1 = 17.6$, $\widehat{\sigma}_0 = 84.0$, $\widehat{\xi}_0 = 0.065$. The standard errors using the conventional method based on observed information are 1.0, 0.9, 0.7 and 0.007 but those calculated via the new method are 9.5, 2.8, 6.3, 0.044. The sharp increase, when dependence is taken into account, is evident. The first– and second–order corrections for dependence between years, mentioned in Section 2, were also calculated and produced standard errors of 10.0, 2.4, 6.8 and 0.051 (first-order), 10.9, 2.8, 5.9, 0.052 (second-order). These are somewhat different but not greatly so, when compared with the difference from the conventional standard errors.

Under model 6 the parameter estimates are are $\widehat{\mu}_0 = 316.0$, $\widehat{\mu}_1 = 12.7$, $\widehat{\mu}_2 = -0.85$, $\widehat{\mu}_3 = 2.30$, $\widehat{\sigma}_0 = 83.8$, $\widehat{\xi}_0 = 0.066$ with standard errors 1.4, 1.6, 1.13, 0.68, 0.8, 0.008 (conventional method), 9.5, 3.8, 1.78, 0.95, 7.9, 0.046 (new method). The critical parameter μ_3 seems clearly significant if assessed by the conventional standard error, but much less so if the new method is used. A likelihood ratio test of the difference between the two models yields a likelihood ratio statistic 2(49919.0-49911.6)=14.8 which would be significant at 0.1% under the conventional χ^2_2 approximation. The new approximation yields $\lambda_1 = 2.47$ and $\lambda_2 = 1.50$ in (2.9), and the p-value of the observed

value 14.8 is 2.6% - still apparently significant, but less obviously so. In fact a plot of the two μ functions (Fig. 1) shows that over the observed range of x (roughly, -2 to 2) there is virtually no difference between the two curves. Numerical estimates of return values confirmed that, in comparison with the standard errors, there is no practical difference over this range in the results produced by the two models, and for this reason, it was decided to use Model 2 in preference to Model 6.

The next thing to consider is the use of the model to obtain estimates of return values. The N-year return value is defined in terms of μ, σ and ξ by

$$q_N = \mu + \sigma \frac{\{-\log(1 - 1/N)\}^{-\xi} - 1}{\xi}. \tag{3.3}$$

A point estimate may be obtained by substituting the maximum likelihood estimates of the unknown parameters, and the approximate variance of this estimate may be represented as $(\bigtriangledown q_N)^T \mathrm{cov}\{\widehat{\theta}\}(\bigtriangledown q_N)$ where $\bigtriangledown q_N$ is the vector of partial derivatives of q_N with respect to the parameters θ evaluated at the maximum likelihood estimator $\widehat{\theta}$. For Model 2 with $x = -1$, 0 and 1, this has been calculated with the following results:

N	10	20	50	100	500
$x = -1$	502.1	573.8	671.7	749.2	942.2
s.e.(old)	2.7	3.9	6.1	8.4	15.8
s.e.(new)	22	30	43	57	99
$x = 0$	519.7	591.4	689.3	766.8	959.8
s.e.(old)	2.6	3.8	6.1	8.4	15.8
s.e.(new)	20	28	42	55	98
$x = 1$	537.3	609.0	707.0	784.4	977.4
s.e.(old)	2.7	3.9	6.2	8.4	15.9
s.e.(new)	19	27	40	54	97

Here the old and new standard errors are those calculated using the conventional observed information matrix, and by the new method.

The main message from this table is the very considerable increase in

standard error estimate as a result of adopting the new method of assessing standard errors.

So far, we have only considered SAAR as a covariate. The results in Dales and Reed (1989) suggested that the three other covariates available (E-W coordinate, N-S coordinate and altitude) would not have a significant effect on the fit. To test this, each parameter in turn was added to Model 2, as an additional linear term on μ, with the following results. From now on we quote only the new standard errors.

E-W coordinate: new parameter estimate is -0.1, standard error 4.3, likelihood ratio statistic 0.012, $\lambda_1 = 21.6$.

N-S coordinate: new parameter estimate is 5.35, standard error 5.65, likelihood ratio statistic 17.4, $\lambda_1 = 19.8$ (so the ratio 17.4/19.8 =0.88 is taken from an approximate χ^2_1 distribution, under the null hypothesis).

Altitude: new parameter estimate is 1.39, standard error 3.0, likelihood ratio statistic 4.2, $\lambda_1 = 10.0$.

It is clear that none of these is significant.

Finally, we consider the goodness of fit. One way to test this is via a chi-squared statistic after grouping the data: in this case the approximate distribution of a likelihood ratio statistic may be computed by the method that has been outlined, so extending the teachnique from parameter estimation to goodness of fit. This has not been tried here, but instead a probability plot of the residuals has been drawn in Figure 2. To construct this plot, the marginal distributions of the data were first transformed to a standard Gumbel distribution ($F(x) = \exp(-e^{-x})$), and the $m = 8259$ ordered values plotted against $-\log[\{-\log(i - 1/2)/m\}]$, $1 \leq i \leq m$, a standard diagnostic procedure with extreme value data. This plot takes no account of the dependence in the data, but it can be seen that the model is an excellent fit except for very few order statistics at the upper end. In fact there is some suggestion that the largest value is an outlier but when this was removed and the model refitted there was very little change in the

estimates. For example, with $x = 0$, the return values just quoted of 519.7, 591.4, 689.3, 766.8 and 959.8 become 519.2, 590.4, 687.6, 764.3 and 954.9, very little different when compared with the standard errors. A method of adjusting the plot to allow for dependence was proposed by Reed and Stewart (1989) but their method uses date information to identify and eliminate dependent peaks, and this information was not available to us here.

To summarise, the new method of calculating standard errors and the critical points of likelihood ratio statistics allows us to fit a regression model to the marginal distributions of the data, and to quote estimates and standard errors for quantities of interest such as return values, while making a realistic correction for the effect of spatial dependence.

It should, however, be pointed out that the method does not help us to estimate parameters that themselves depend on the dependence between sites. For instance, there is interest in *areal reduction factors*, in which the distribution of the maximum over a set of sites is compared with the distributions of individual sites. This cannot be computed without some form of model for dependence among sites. Dales and Reed (1989) and Smith (1990) have made proposals for this, but it is really a different problem from the one being studied here.

4. A simulation study

A simulation study was performed to assess the validity of the approximation for calculating standard errors. The following results are all based on 1000 replications.

For each replication, 10 years' of independent data were generated at each of 100 dependent sites. It is desirable to have a quick and simple algorithm to generate dependent data with the right marginal distribution, and the following scheme was adopted:

Step 1: Generate $V_0,\ V_1,\ ..., V_{100}$ independent unit exponential.

Step 2: Let $W_i = V_0 + V_i$, $1 \leq i \leq 100$. At this stage the W's have a

gamma distribution with shape parameter 2.

Step 3: Let $X_i = W_i - \log(1 + W_i)$, $1 \leq i \leq 100$. This transforms the marginal distribution back to unit exponential.

Step 4: Let Y_i be $-\log(X_i)$ if $\xi = 0$, otherwise $(X_i^{-\xi} - 1)/\xi$.

Step 5: Repeat Steps 1-4 10 times to get 10 independent years' data.

This was done with $\xi = 0$, 0.2 and -0.2. In addition, for comparison, the same simulation was performed with all sites independent.

The parameters were estimated by maximum likelihood and the 10-year, 50-year and 250-year return values calculated, together with their standard errors by both methods.

For the dependent data with ξ=0, the following results were obtained for the true root mean squared error, and the (average over 1000 simulations) standard errors as estimated by the old and new methods:

Parameter	True RMSE	Old s.e.	New s.e
μ	0.25	0.034	0.23
σ	0.13	0.025	0.12
ξ	0.089	0.023	0.073
q_{10}	0.45	0.076	0.40
q_{50}	0.80	0.17	0.67
q_{250}	1.39	0.34	1.17

It can be seen that the new method of computing standard errors greatly improves on the old, but is still something of an underestimate.

This is confirmed when we consider the actual coverage probability of a confidence interval for q_N constructed so as to have a half-width of two standard errors. Using the old standard errors, this was 24%, 34% and 40% respectively for $N = 10$, 50 and 250. With the new standard errors, these were improved to 88%, 83% and 82%, which are still short of the nominal value of 95%.

These figures were broadly confirmed with simulations assuming $\xi = 0.2$ and $\xi = -0.2$. The case $\xi = 0.2$ is expected to be the most troublesome of the three, because this is the longest-tailed case, and it certainly is the

case that the standard errors are the largest in this case. For example, for q_{250} the true RMSE based on 100 simulations was calculated as 3.86, compared with averages of 0.91 and 3.45 as estimated by the old and new methods. The corresponding coverage probabilities for N =10, 50 and 250 were 24%, 32% and 38% under the old method, 88%, 83% and 83% under the new. In the case of $\xi = -0.2$ the coverage probabilities were 24%, 35%, 42% and 88%, 80% and 79%. The coverage probabilities do not seem very sensitive to ξ.

For comparison, the same calculations were carried out with the stations independent. In this case the old and new methods of calculating standard errors are asymptotically equivalent, but of course they do not give identical answers in practice. For true value $\xi = 0$ we obtain:

Parameter	True RMSE	Old s.e.	New s.e
μ	0.034	0.035	0.032
σ	0.025	0.026	0.023
ξ	0.021	0.022	0.020
q_{10}	0.073	0.078	0.072
q_{50}	0.16	0.17	0.15
q_{250}	0.30	0.32	0.29

In this case the estimated standard errors (old method) are virtually identical to those quoted in the dependent case, but here they are very close to the true values which are, of course, much smaller. Curiously, the new method still tends slightly to underestimate the true standard error. The coverage probabilities of nominal 95% confidence intervals are 96%, 96%, 96% (old method), 92%, 92%, 91% (new method). With true $\xi = 0.2$ the corresponding coverage probabilities are 96%, 96%, 96% and 92%, 93%, 92% while with $\xi = -0.2$ they are 97%, 96%, 96% and 92%, 91%, 90%.

Our conclusions are as follows. When the data are dependent, the old method of computing standard errors, which ignores dependence, leads to gross underestimates of sampling variability and very poor coverage probability for confidence intervals based on those standard errors. The new

method still seems to underestimate sampling variability and to overestimate confidence coefficients - in the simulations, nominally 95% confidence intervals had a true coverage probability of about 85%. Corresponding simulations with independent data showed the old and new methods much more comparable, but there still seems to be some tendency for the new method to underestimate standard errors and to overestimate confidence coefficients. However, even though there is still scope for improvement, it is clear that the new method is a great improvement on the old and goes a long way towards obtaining accurate interval estimates.

Acknowledgements

I am grateful to Dr. D.W. Reed of the Institute of Hydrology for the data, for copies of a number of technical reports and for much helpful discussion. I would also like to thank Dr. T.A. Buishand for a copy of the unpublished paper Buishand (1990), and additional references. The work was supported by a Wolfson Foundation personal research fellowship, and used computational facilities provided by the S.E.R.C.

REFERENCES.

Arnell, N.W. and Gabriele, S. (1988), The performance of the two-component extreme-value distribution in regional flood frequency analysis. *Wat. Resour. Res.* **24**, 879-887.

Buishand, T.A. (1984), Bivariate extreme-value data and the station-year method. *J. Hydrol.* **69**, 77-95.

Buishand, T.A. (1990), Extreme rainfall estimation by combining data from several sites. Unpublished.

Coles, S.G. and Tawn, J.A. (1990), Statistics of coastal flood prevention. *Phil. Trans. R. Soc. Lond. A* **332**, 457-476.

Cunnane, S. (1989), Statistical distributions for flood frequency analysis. *Operation Hydrology Report No. 33*, WMO, Geneva.

Dales, M.Y. and Reed, D.W. (1989), Regional flood and storm hazard assessment. Report No. 102, Institute of Hydrology, Wallingford, Oxon.

Farebrother, R.W. (1984), Algorithm AS204: The distribution of a positive linear combination of χ^2 random variables. *Applied Statistics* **33**, 332-339.

Galambos, J. (1987), *The Asymptotic Theory of Extreme Order Statistics* (2nd. edn.). Krieger, Melbourne, Fl. (First edn. published 1978 by John Wiley, New York.)

Hosking, J.R.M., and Wallis, J.R. (1988), The effect of intersite dependence on regional flood frequency analysis. *Wat. Res. Res.* **24**, 588-600.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985), Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27**, 251-261.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983), *Extremes and Related Properties of Random Sequences and Processes*. Springer Verlag, New York.

Moore, R.J. (1987), Combined regional flood frequency analysis and regression on catchment characteristics by maximum likelihood estimation. In *Regional Flood Frequency Analysis*, ed. V.P. Singh, Reidel, Dordrecht, 119-131.

N.E.R.C. (1975), *The Flood Studies Report*. The Natural Environment Research Council, London.

Reed, D.W. and Stewart, E.J. (1989), Focus on rainfall growth estimation. *Proc. 2nd. National Hydrology Symposium*, Sheffield, September 1989.

Smith, R.L. (1990), Max-stable processes and spatial extremes. Unpublished.

Stedinger, J.R. (1983), Estimating a regional flood frequency distribution. *Wat. Res. Res.* **19**, 503-510.

Figure 1: Comparison of Two Functions for Mu

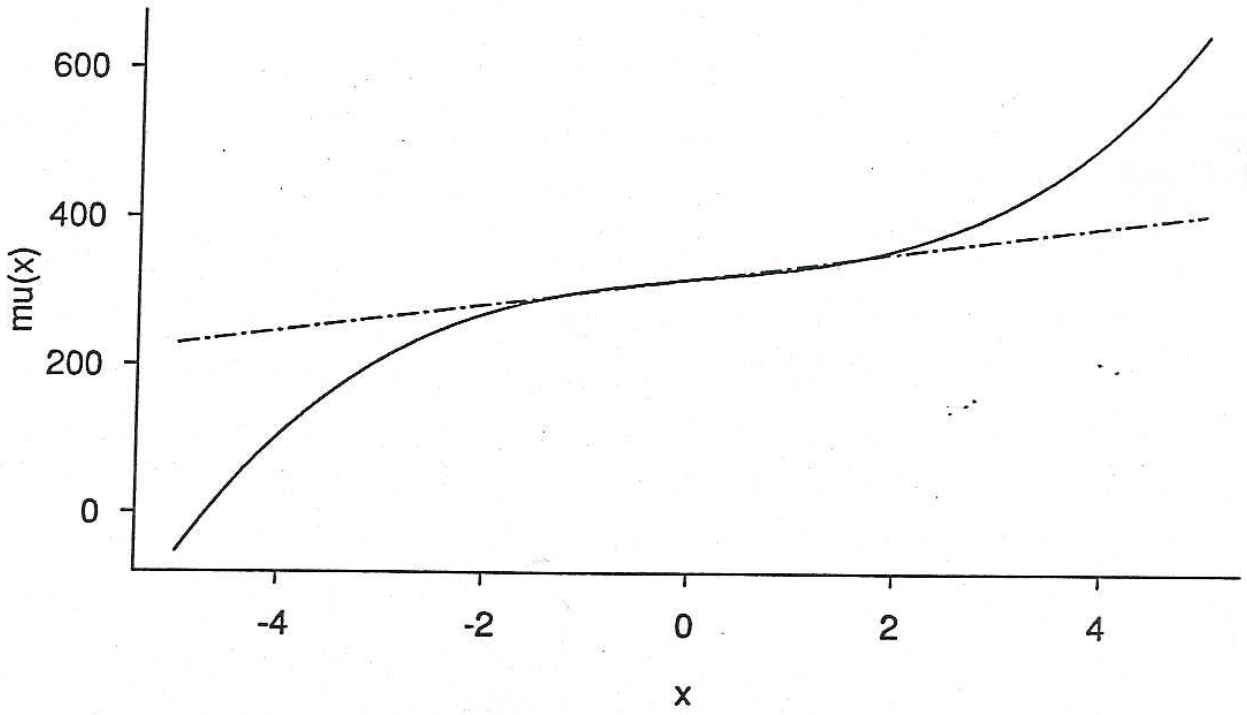

Figure 2: Residual Plot for Model 2