Mixed methods for fitting the GEV distribution

Pierre Ailliot
Université de Brest
France

Craig Thompson
National Institute of Water and Atmospheric Research
New Zealand

Peter Thomson
Statistics Research Associates Ltd
New Zealand

Outline

1. Introduction
2. GEV estimation using mixed methods
3. Computational issues
4. Numerical study and an application
5. Conclusions

1. Introduction

The distribution of extremes is of interest to many disciplines including

- Economics and Finance: VAR calculations, extreme returns, insurance loss distributions, etc

- Environmental Sciences: maximum daily rainfall, minimum hydro inflows, etc

among many others.

Focus here is on the block maximum or maximum observation in a sample (e.g. annual maximum daily rainfall for Wellington).

For an iid sequence $X_1, \ldots, X_n$, classical Extreme Value Theory (EVT) shows that, when suitably normalised,

$$M_n = \max(X_1, \ldots, X_n)$$

has an asymptotic Generalised Extreme Value (GEV) distribution

$$F(x|\boldsymbol{\theta}) = \begin{cases} \exp\{-[1 - \kappa\frac{x-\beta}{\alpha}]^{\frac{1}{\kappa}}\} & (\kappa \neq 0) \\ \exp\{-\exp[-\frac{x-\beta}{\alpha}]\} & (\kappa = 0) \end{cases}$$

where $\kappa(x - \beta) < \alpha$, $\alpha > 0$ and $\boldsymbol{\theta} = (\beta, \alpha, \kappa)$.

Notes:

- Result holds under more general circumstances;
- GEV distribution fitted more widely in practice;
- Heavy tailed when $\kappa < 0$ (Frechet), light tailed when $\kappa = 0$ (Gumbel) or $\kappa > 0$ (Weibull).

The GEV parameters are commonly fitted by

- maximum likelihood (ML) (Prescott and Walden,1980) which is asymptotically efficient if $\kappa < 0.5$ (Smith, 1985);

- method of L-moments (Hosking, 1990) giving consistent, large-sample Gaussian estimates if $\kappa > -0.5$ (finite variance).

However

- method of moments methods reported to be more accurate than ML for small samples, especially for fitted quantiles;

- ML-based methods show improved small-sample properties when parameter constraints are imposed;

where the latter include Bayesian methods (Martins and Stedinger, 2000), penalised ML (Coles and Dixon, 1999) and moment constraints (Morrison and Smith, 2002).

2. GEV estimation using mixed methods

The log-likelihood of a GEV random sample is

$$\ln L(\boldsymbol{\theta}) = -n \ln \alpha - \sum_{i=1}^{n} (1 - \kappa \frac{X_i - \beta}{\alpha})^{\frac{1}{\kappa}} + (\frac{1}{\kappa} - 1) \sum_{i=1}^{n} \ln(1 - \kappa \frac{X_i - \beta}{\alpha})$$

where

$$\alpha > 0, \quad \kappa(X_i - \beta) < \alpha \quad (i = 1, \ldots, n).$$

Further assume

$$-0.5 < \kappa < 0.5$$

to satisfy finite variance, ML regularity conditions.

Suppose now that $\boldsymbol{\theta}$ is estimated by a mixture of two methods; maximum likelihood and method of moments.

Method M1

Suppose κ is estimated by maximising $\ln L(\boldsymbol{\theta})$ with α and β given as functions of κ by the L-moments

$$\begin{aligned} \lambda_1 = E(X_1) \qquad\qquad &= \beta + \frac{\alpha}{\kappa}(1 - \Gamma(1+\kappa)) \\ \lambda_2 = \tfrac{1}{2}E(|X_1 - X_2|) &= \frac{\alpha}{\kappa}(1 - 2^{-\kappa})\Gamma(1+\kappa) \end{aligned}$$

where λ_1, λ_2 are estimated by

$$\hat{\lambda}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\lambda}_2 = \frac{1}{n(n-1)}\sum_{i<j} |X_i - X_j|.$$

This mixed estimation method was first proposed by Morrison and Smith (2002).

Notes:

- The two L-moment constraints and $\lambda_3 = \kappa$ specify a 1-1 mapping between the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$.

- Approach can be generalised to other moment constraints and associated 1-1 mappings with

$$\hat{\lambda}_1 = \hat{\lambda}_1(\mathbf{X}), \quad \hat{\lambda}_2 = \hat{\lambda}_2(\mathbf{X}), \quad \hat{\lambda}_3 = \arg\max_{\lambda_3} \ln L(\widehat{\lambda_1}, \widehat{\lambda_2}, \lambda_3)$$

  where $\widehat{\lambda_1}$, $\widehat{\lambda_2}$ are known unbiased estimators of λ_1, λ_2 and $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}(\hat{\boldsymbol{\lambda}})$ where $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{\lambda})$ specifies the 1-1 mapping.

- Method can also be extended to the case of only one moment constraint with the remaining parameters estimated by ML.

Asymptotic properties of the estimates

If $-0.5 < \kappa < 0.5$ and n is large, then $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}(\hat{\boldsymbol{\lambda}})$ is consistent and

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim N(0, J^T V J)$$

where

$$J = \partial \boldsymbol{\theta}^T / \partial \boldsymbol{\lambda}, \qquad V = BCB^T,$$

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & I_{33}^{(\boldsymbol{\lambda})} \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ I_{31}^{(\boldsymbol{\lambda})} & I_{32}^{(\boldsymbol{\lambda})} & I_{33}^{(\boldsymbol{\lambda})} \end{bmatrix}$$

with

$$\boldsymbol{I}^{(\boldsymbol{\lambda})} = J \boldsymbol{I}^{(\boldsymbol{\theta})} J^T, \qquad C_{ij} = \lim_{n \to \infty} n \, \mathrm{cov}(\hat{\lambda}_i, \hat{\lambda}_j)$$

and $\boldsymbol{I}^{(\boldsymbol{\theta})}$ is the information matrix of the GEV distribution.

Notes:

- Consistency of $\widehat{\boldsymbol{\theta}}$ follows from Smith (1985).
- Analytic form for $\boldsymbol{I}^{(\boldsymbol{\theta})}$ given by Prescott and Walden (1980).
- Analytic expressions for the C_{ij} follow from standard theory of L-moments (Hosking, 1990) and U-statistics (Lee, 1990).
- For Method M1, an alternative formula for C_{22} was derived using the symmetric Beta distribution rather than the hypergeometric function.
- Mathematics straightforward, but demanding. Theory checked by simulation.

Other examples of mixed GEV estimation methods

Method M2: Consider the mapping

$$\lambda_1 = \beta + \frac{\alpha}{\kappa}(1 - \Gamma(1 + \kappa)), \qquad \lambda_2 = \alpha, \qquad \lambda_3 = \kappa$$

with $\widehat{\lambda}_1$ given by the sample mean and λ_2, λ_3 estimated by constrained ML.

Method M3: Consider the mapping

$$\lambda_1 = \beta + \frac{\alpha}{\kappa}(1 - (\ln 2)^{\kappa}), \qquad \lambda_2 = \frac{\alpha}{\kappa}(1 - 2^{-\kappa})\Gamma(1 + \kappa), \qquad \lambda_3 = \kappa$$

which is the same as Method M1, but with the first L-moment replaced by the median. Method M3 is a robust alternative to Method M1.

Quantile estimates

GEV quantiles are given by

$$q_p \equiv q_p(\boldsymbol{\theta}) = \begin{cases} \beta + \frac{\alpha}{\kappa}(1 - (-\log p)^{\kappa}) & (\kappa \neq 0) \\ \beta - \alpha \log(-\log p) & (\kappa = 0) \end{cases}$$

where $F(q_p) = p$. A natural estimator of q_p is

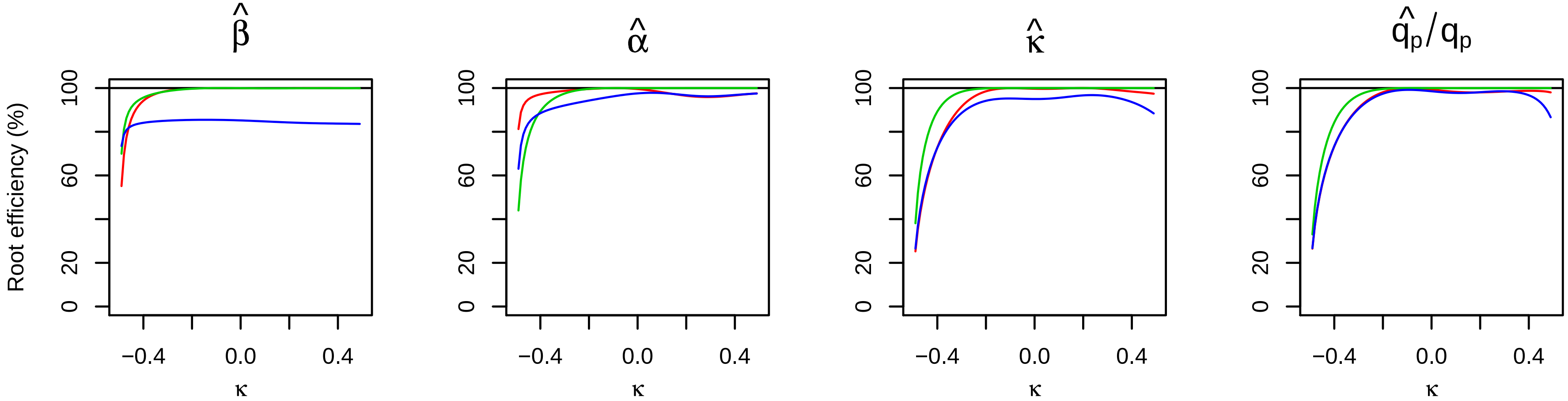
$$\hat{q}_p = q_p(\hat{\boldsymbol{\theta}})$$

where $\hat{\boldsymbol{\theta}}$ is an estimator of $\boldsymbol{\theta}$ such as one of those considered.

Then $\sqrt{n}(\hat{q}_p - q_p)$ is asymptotically Gaussian with zero mean and variance

$$\boldsymbol{\Delta}^T \Sigma \boldsymbol{\Delta}, \qquad \boldsymbol{\Delta} = (\frac{\partial q_p}{\partial \beta}, \frac{\partial q_p}{\partial \alpha}, \frac{\partial q_p}{\partial \kappa})^T$$

where $\Sigma = \text{cov}(\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}))$. This result can be used to give approximate standard errors for $q_p(\hat{\boldsymbol{\theta}})$.



Root-efficiency (%) of M1, M2 and M3 estimators $\widehat{\beta}$, $\widehat{\alpha}$, $\widehat{\kappa}$ and $\widehat{q}_p/q_p$ for $\beta = 0$, $\alpha = 1$, $-0.5 < \kappa < 0.5$.

3. Computational issues

All methods, including ML, involve constrained maximisation of the log-likelihood. The constraints fall into three groups;

- support of the GEV density;
- moment constraints;
- technical constraint $-0.5 < \kappa < 0.5$.

Estimation strategy: Optimise over all parameters except κ, taking careful account of the constraints, and form a profile log-likelihood for κ which is plotted and optimised over κ.

This procedure proves to be numerically robust, graphically informative, and computationally efficient.

For Method M1 the constraints reduce to

$$\max(\kappa^-_{M1}, -0.5) < \kappa < \min(\kappa^+_{M1}, 0.5).$$

where

$$\kappa^-_{M1} = -\ln\left(1 + \frac{\hat{\lambda}_2}{\hat{\lambda}_1 - \min_i x_i}\right) / \ln 2$$

$$\kappa^+_{M1} = -\ln\left(1 - \frac{\hat{\lambda}_2}{\max_i X_i - \hat{\lambda}_1}\right) / \ln 2$$

are determined solely from the data.

Method M1 provides simple, computationally efficient estimates that can be used in their own right, or as initial estimates to computationally intensive methods such as ML.

For maximum likelihood (ML) the constraints reduce to

$$\alpha > \alpha_{ML} = \begin{cases} \kappa(\max_i X_i - \beta) & (\kappa \geq 0) \\ -\kappa(\beta - \min_i X_i) & (\kappa < 0) \end{cases}$$

provided β, the 0.3679 quantile of the GEV distribution, satisfies

$$\min_i X_i \leq \beta \leq \max_i X_i.$$

Note that

- constraint is linear in α, β given κ;
- $P(\min_i X_i \leq \beta \leq \max_i X_i) > 1 - 10^{-5}$ for $n > 25$ so β constraint is a reasonable assumption in practice.

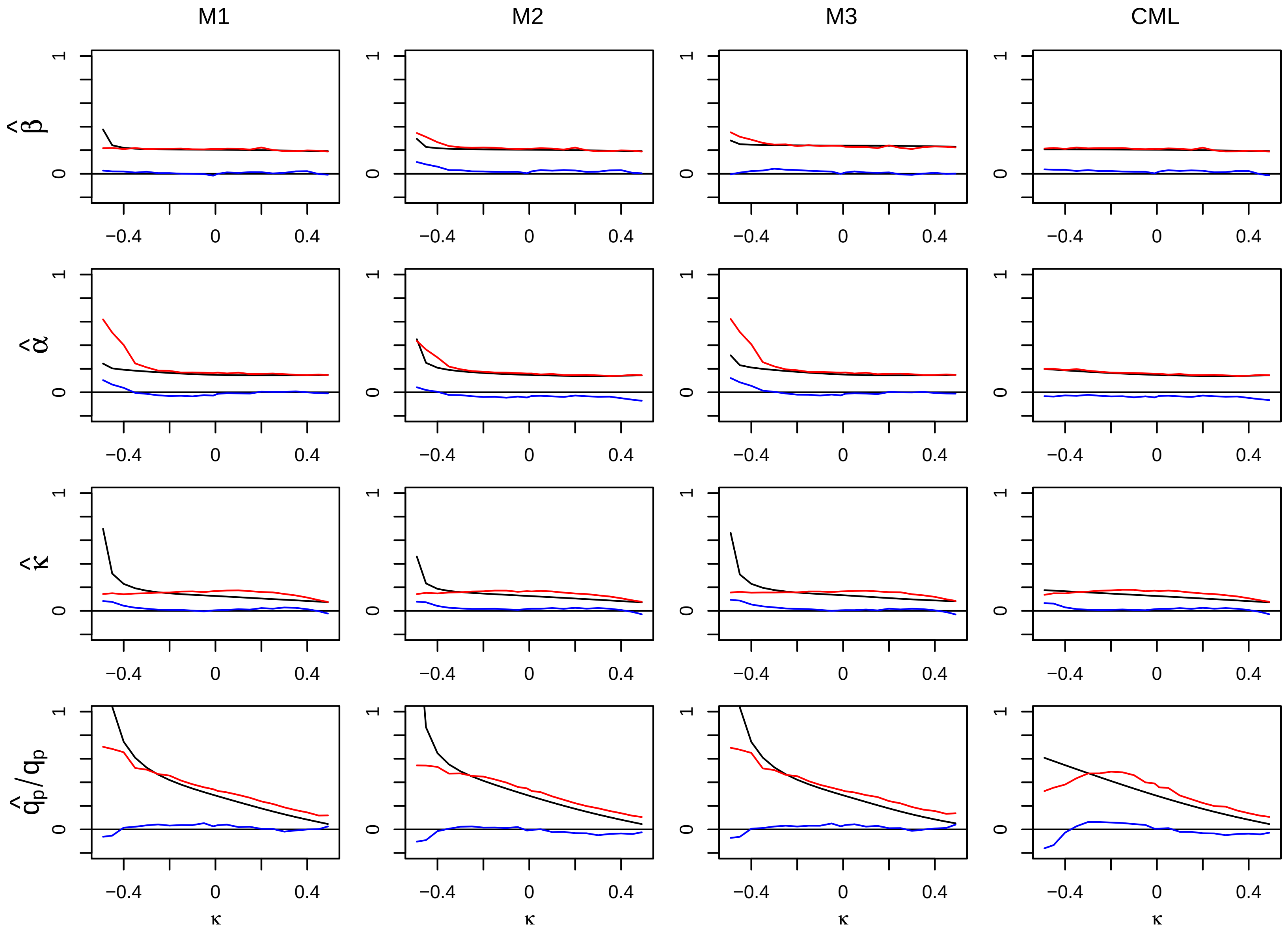
Simple constraints also hold for methods M2 and M3.

4. Numerical study and an application

A simulation study was undertaken to check how well the asymptotic results applied in practice with

- sample sizes $n = 30$, 60 and 120;
- $\beta = 0$, $\alpha = 1$ and $-0.5 < \kappa < 0.5$;
- 1000 replications for each choice of κ;
- M1, M2, M3 and ML estimates determined in each case.

The estimated bias and root-mean-squared error (RMSE) for each method was then computed as a function of κ.



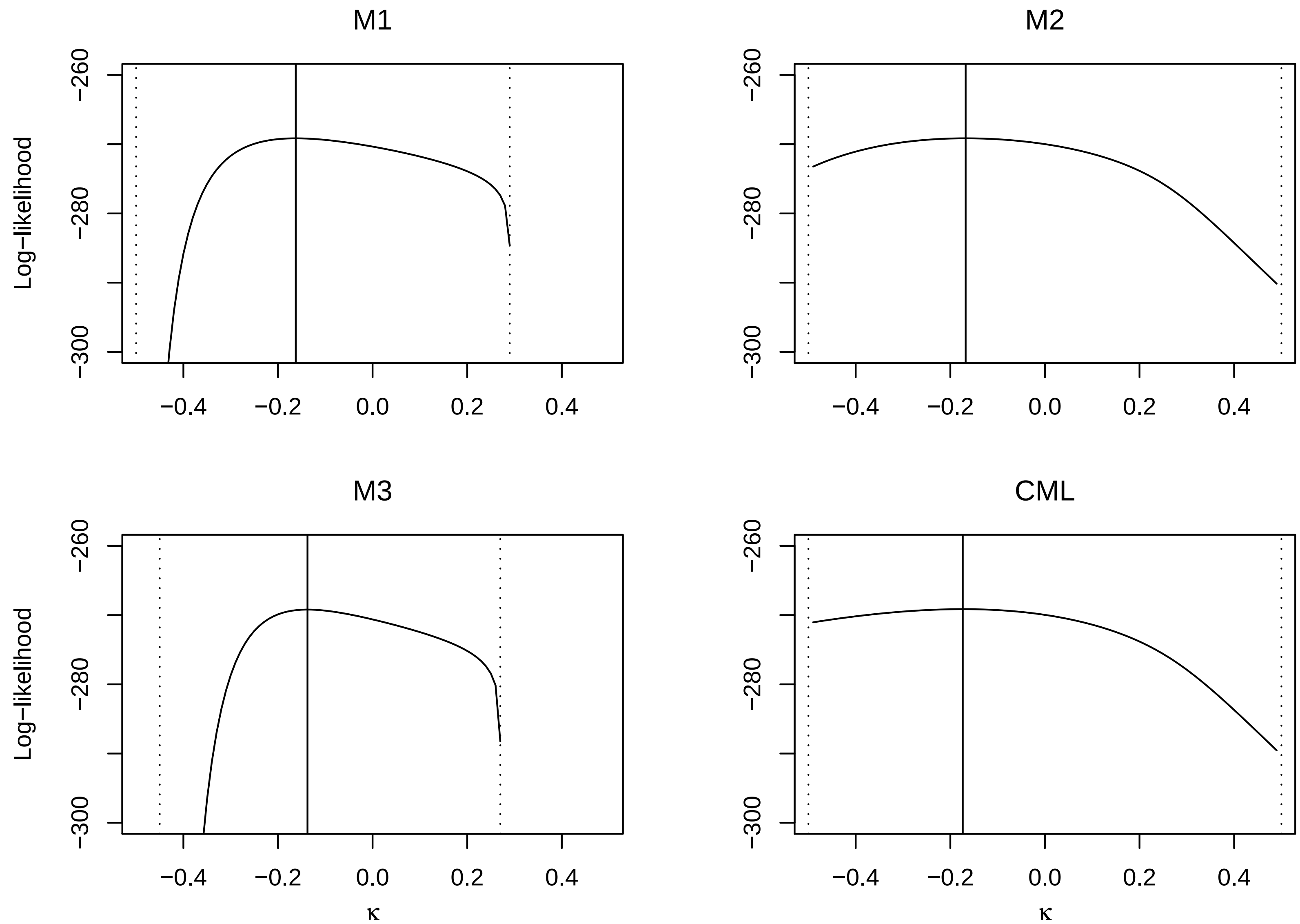
Asymptotic RMSE, **simulated RMSE** and **simulated bias** of the M1, M2, M3, CML estimators $\widehat{\beta}$, $\widehat{\alpha}$, $\widehat{\kappa}$, $\widehat{q}_p/q_p$ for $n = 30$, $\beta = 0$, $\alpha = 1$, $-0.5 < \kappa < 0.5$.

Now consider an application to a sample of 24-hour annual maximum rainfall from Wellington.

	- IPO	+ IPO	All
Sample size	31	29	60
Mean (first L-moment)	82.0	75.5	78.9
Median	73.9	70.1	73.6
Standard deviation	28.7	19.0	24.5
Second L-moment	16.3	10.4	13.6
Maximum	153.2	121.2	153.2
Minimum	47.0	51.4	47.0

Wellington 24-hour annual maximum rainfall statistics (mm) over the period 1940-1999 and for each phase of the IPO.

First analyse the entire data set.



Profile log-likelihoods for the M1, M2, M3, CML methods applied to Wellington 24-hour annual maximum rainfall (mm) over the period 1940-1999. The vertical lines mark the κ estimates (solid) and κ bounds (dashed).

Method	β	α	κ	q_p
L	66.98 (2.67)	18.22 (2.12)	-0.07 (0.10)	166.48 (23.96)
M1	66.29 (2.40)	16.45 (1.87)	-0.16 (0.10)	178.79 (30.99)
M2	66.19 (2.40)	16.43 (1.87)	-0.17 (0.10)	180.09 (31.45)
M3	67.18 (2.89)	16.94 (1.98)	-0.14 (0.10)	175.96 (28.82)
CML	66.32 (2.41)	16.54 (1.89)	-0.17 (0.10)	182.77 (32.43)

Estimates of β, α, κ, and q_p ($p = 0.99$) for Wellington 24-hour annual maximum rainfall (mm) over the period 1940-1999 using L-moments (L) and methods M1, M2, M3, CML (asymptotic standard errors in parentheses).

Now analyse the data within each phase of the IPO.

Method	- IPO			
	β	α	κ	q_p
L	68.30 (4.71)	23.24 (3.62)	-0.01 (0.14)	178.84 (33.04)
M1	66.96 (4.07)	20.05 (3.15)	-0.15 (0.14)	200.41 (50.13)
M2	66.98 (4.07)	20.05 (3.14)	-0.15 (0.14)	200.19 (49.74)
M3	66.47 (4.66)	19.65 (3.27)	-0.17 (0.15)	202.88 (53.22)
CML	67.13 (4.09)	20.16 (3.17)	-0.16 (0.14)	204.15 (51.88)

Method	+ IPO			
	β	α	κ	q_p
L	65.99 (2.77)	13.06 (2.30)	-0.14 (0.16)	149.51 (33.07)
M1	65.66 (2.55)	12.10 (2.02)	-0.20 (0.15)	156.31 (38.41)
M2	65.87 (2.55)	12.12 (2.01)	-0.18 (0.15)	153.74 (35.94)
M3	65.54 (2.94)	11.98 (2.13)	-0.21 (0.16)	157.15 (39.71)
CML	65.97 (2.56)	12.21 (2.02)	-0.19 (0.15)	155.70 (36.91)

Estimates of β, α, κ, and q_p ($p = 0.99$) for Wellington 24-hour annual maximum rainfall (mm) within each phase of the IPO over the period 1940-1999 using L-moments (L) and methods M1, M2, M3, CML (asymptotic standard errors in parentheses).

Likelihood ratio test retains hypothesis of no difference.

5. Conclusions

Mixed ML and L-moments GEV estimation methods have been generalised to include other moment constraints and

- asymptotic properties derived;
- analytic expressions given for the asymptotic covariances;
- results verified by simulation and efficiencies established;
- profile likelihoods and careful account of constraints have led to efficient and robust computational procedures.

Key findings:

- finite sample performances of all methods compare favourably with asymptotic results when $\kappa > -0.3$, $n \geq 30$;
- ML performed well over $-0.5 < \kappa < 0.5$ for $n \geq 30$.