Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

SIMULATION OF MULTIVARIATE EXTREME VALUES

S. Nadarajah

School of Mathematics and Statistics
University of Plymouth
Drake Circus
Plymouth, PL4 8AA, U.K.

ABSTRACT

A comprehensive method for simulation of bivariate extremes is introduced and a generalisation of it to multivariate extremes is outlined.

1 INTRODUCTION

The last decade has seen development of several models for multivariate extremes based on extreme value theory (Coles and Tawn 1991, 1994; Joe, Smith and Weissman 1992; Smith, Tawn and Coles 1993; Ledford and Tawn 1996). These models have attracted a great deal of attention particularly in the area of environmental extremes. Simulation of multivariate extremes is useful to study not only the statistical properties of these models but also the properties of complex processes such as environmental processes whose inputs might be modelled as multivariate extremes (Anderson and Nadarajah 1993; Dixon and Tawn 1994). Simulation of univariate extremes is straightforward since for all distributions in the univariate theory the inversion method can be applied (see e.g. Ripley 1987). However in relation to simulation of multivariate extremes there has been little work so far – the only work known to the author is by Shi, Smith and Coles (1993) who describe a method for simulation from a special class of bivariate extreme value distributions. The aim of this paper is to develop a comprehensive method for simulation of multivariate extremes.

Section 2 details a method for simulation of bivariate extremes. We first model bivariate extremes by use of bivariate extreme value theory (Section 2.1) and then show that the simulation of them can be performed by two independent steps (Sections 2.2 and 2.3), each simulating a univariate transform of the extremes. We also illustrate the performance of the simulation method for a special class of extreme value distributions (Section 2.4). Section 3 outlines a generalisation of the method of Section 2 for multivariate extremes.

2 SIMULATION OF BIVARIATE EXTREMES

2.1 The Model

Let (X, Y) denote a random vector with both X and Y having the unit Fréchet distribution, i.e. $\Pr(X \leq x) = \Pr(Y \leq x) = e^{-1/x}$, $x > 0$. Other distributions for the margins X and Y can be generated by appropriate univariate simulation techniques. Also let $P(X, Y) = (R, W) = (X + Y, X/(X + Y))$ be the radial–angular transformation of (X, Y) where R and W are the radial and angular transforms respectively. Then we say under certain regularity conditions in bivariate extreme value theory (Galambos 1978, Chapter 5; Resnick 1987, Chapter 5) that (X, Y) takes extreme values in a region $A \subset \Re_+^2 \backslash \{(0,0)\}$ sufficiently away from the origin if

$$\Pr((X, Y) \in A) = \int_{P(A)} r^{-2} dr H_*(dw) \qquad (1)$$

where H_* is a finite non–negative measure on $[0, 1]$ constrained to have unit means (i.e. $\int_{[0,1]} w H_*(dw) = \int_{[0,1]} (1 - w) H_*(dw) = 1$) and thus a total mass of 2. This section develops a method for simulation of extreme values of (X, Y) by using (1) as the model for them.

Take $A = A_0 = \{(x, y) : x + y > r_0, x \geq 0, y \geq 0\}$ with r_0 chosen sufficiently large that (1) holds. We wish to simulate extreme values of (X, Y) from A_0, see below for how simulation from other forms of A can be performed. It follows from (1) that the conditional probability density function of (R, W) over $P(A_0)$ is:

$$\begin{aligned} f_{R,W}(r, w) &= r^{-2} H_*(dw) / \int_{r_0}^{\infty} \int_0^1 s^{-2} ds H_*(dv) \\ &= r_0 r^{-2} 2^{-1} H_*(dw) \qquad (2) \end{aligned}$$

for $r > r_0$ and $w \in [0,1]$. Clearly (2) is the density of a point of a Poisson process with intensity $r^{-2}H_*(dw)$ restricted to $P(A_0)$. Accordingly simulation of (x,y) over A_0 reduces in essence to simulation of (r,w) coordinates of the Poisson process over $P(A_0)$. Further the form of A_0 ensures that the conditional density (2) is independent, so we can simulate the radial and angular coordinates (r and w) independently of each other. The procedures for these are developed in the subsequent Sections 2.2 and 2.3. To simulate (x,y) over a region A not having the form A_0 first apply the above procedures to simulate (x,y) over an A_0 for which $A \subset A_0$ and then delete those points falling in $A_0 \backslash A$.

2.2 r–step

We simulate r by the inversion principle: set $\int_{r_0}^{r} r_0 s^{-2} ds = u$ for $u \frown U(0,1)$ and invert to get $r = r_0/(1-u)$.

2.3 w–step

Associate with H_* the measure density $h(w) = \partial H_*([0,w])/\partial w$, assuming differentiability of H_* in the interior $(0,1)$, and atoms of mass $\theta_0 = H_*(\{0\})$ and $\theta_1 = H_*(\{1\})$ at the end points of $[0,1]$. Then H_* is a composition of the density h in the interior $(0,1)$ and the atoms at the end points. Thus simulation of w can be performed by the method of composition (see e.g. Ripley 1987, Section 3.2): take w as 0, belonging to $(0,1)$ or 1 with probabilities $2^{-1}\theta_0$, $1 - 2^{-1}\theta_0 - 2^{-1}\theta_1$ and $2^{-1}\theta_1$. If w falls into $(0,1)$ then simulate its value from the probability density $h^*(w) = (2-\theta_0-\theta_1)^{-1}h(w)$ by the method we present below.

We use the rejection method (see e.g. Ripley 1987, Section 3.2) to simulate from $h^*(w)$ and this involves establishing a probability density $g(w)$ and a constant M such that $h^*(w)/g(w) \leq M$ for all $w \in (0,1)$. The density g is referred to as the envelope. Under mild conditions on h^*, a form for the envelope g is easily established by the following theorem.

Theorem 1 *Suppose h^* is a continuous probability density over $(0,1)$. Suppose too that there are some constants $q_0 > -1$, $q_1 > -1$ such that $h^*(w) = O(w^{q_0})$ and $h^*(1-w) = O(w^{q_1})$ hold as $w \to 0$. Then the beta density*

$$g(w) = \frac{w^{q_0}(1-w)^{q_1}}{Be(q_0+1, q_1+1)}, \quad w \in (0,1)$$

where

$$Be(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a > 0, b > 0$$

satisfies $h^*(w)/g(w) \leq M < \infty$ for some constant M.

Proof On every compact subset of $(0,1)$ the supremum of h^*/g is bounded since h^* is continuous and g is bounded away from zero. Near the end points h^*/g is bounded by the assumed conditions on h^*. ■

Tawn (1990) and Coles and Tawn (1991) among others have developed several parametric models for H_*. Nadarajah (1994) shows that each currently known model for H_* satisfies the requirements of Theorem 1 for particular choices of q_0, q_1 and M: see Example 1 for an illustration where we choose $M = \sup_{w\in(0,1)} h^*(w)/g(w)$, the value leading to most efficiency of the rejection method. Hence by the rejection method simulation of w for each currently known model reduces to simulation from a Beta$(q_0 + 1,\ q_1 + 1)$ distribution and routines are widely available for the latter.

Example 1 The logistic model for H_* (Tawn 1990) takes the following form for density h:

$$\begin{aligned} h(w) &= (s-1)(\phi_0\phi_1)^s\{w(1-w)\}^{s-2} \\ &\quad \times\{(\phi_1 w)^s + (\phi_0(1-w))^s\}^{1/s-2} \end{aligned} \tag{3}$$

$(0 < \phi_0,\ \phi_1 \leq 1,\ s > 1)$ and has atoms $\theta_0 = 1 - \phi_1$ & $\theta_1 = 1 - \phi_0$. Limiting $w \to 0, 1$ in (3) we see $q_0 = q_1 = s - 2$. It follows then

$$\begin{aligned} h^*(w)/g(w) &= (s-1)(\phi_0\phi_1)^s(\phi_0+\phi_1)^{-1} \\ &\quad \times Be(s-1, s-1) \\ &\quad \times\{(\phi_1 w)^s + (\phi_0(1-w))^s\}^{1/s-2}. \end{aligned} \tag{4}$$

Since $s > 1$ use of calculus shows $(\phi_1 w)^s + (\phi_0(1-w))^s$ attains its global minimum at $w = \phi_0/(\phi_0+\phi_1)$. Substituting this w into (4) gives

$$\begin{aligned} M &= \sup_{w\in(0,1)} h^*(w)/g(w) \\ &= (s-1)2^{1/s-2}(\phi_0\phi_1)^{1-s}(\phi_0+\phi_1)^{2(s-1)} \\ &\quad \times Be(s-1, s-1) \end{aligned}$$

which is a finite constant for $\phi_0 > 0$, $\phi_1 > 0$ and $s > 1$. ■

2.4 Illustration

We illustrate the above simulation method graphically, the basis being the following construct. Let

(R, W) be a random vector satisfying (2). For $v \in (0,1)$ define $Z(v) = R^{-1}\max(v/W, (1-v)/(1-W))$. Then it is easily seen that for $r_0 z \leq \min(v, 1-v)$:

$$\Pr(Z(v) \leq z) = \frac{r_0}{2}\left\{\frac{1}{v}\int_{[0,v]} wH_*(dw) + \frac{1}{1-v}\int_{(v,1]}(1-w)H_*(dw)\right\} z.$$

Hence $r_0 Z(v)/\min(v,1-v) \mid r_0 Z(v)/\min(v,1-v) \leq 1 \frown U(0,1)$, the construct which we illustrate as follows using Q–Q plots. Simulate values of (R, W) over $P(A_0)$ by the r & w–steps with $r_0 = 1$, an arbitrary choice as other values for it will produce identical results given this illustration — we used the Numerical Algorithms Group routines G05CAF and G05FEF to simulate random numbers having the uniform and beta distributions. Then plot quantiles of the simulated $r_0 Z(v)/\min(v,1-v)$ not exceeding 1 versus the corresponding probabilities — we will expect the plot (Q–Q plot) to be linear with a slope of 1 and an intercept of 0. Figure 1 shows these plots for a number of choices for v and parameters of H_* when H_* is modelled as in Example 1. Clearly, in each plot, there is satisfactory agreement that the conditional distribution of $r_0 Z(v)/\min(v,1-v)$ is uniform $U(0,1)$.

3 GENERALISATION TO MULTIVARIATE EXTREMES

Here we consider simulation for a vector of more than two variables $(X_1, \ldots, X_p)$, say, with each X_i having the unit Fréchet distribution. As before the simulation is based on a model for extremes of $(X_1, \ldots, X_p)$ and involves two independent steps for simulating radial and angular transforms of the variables.

3.1 The Model

The radial–angular transform $P(X_1, \ldots, X_p)$ for p variables can be written as $(R, \mathbf{W}) = (\sum_{i=1}^{p} X_i, X_1/R, \ldots, X_{p-1}/R)$ where $\mathbf{W} = (W_1, \ldots, W_{p-1})$, the angular transform, is an element of the $(p-1)$–dimensional unit simplex $S_p = \{(w_1, \ldots, w_{p-1}) : w_1 + \ldots + w_{p-1} \leq 1\}$. By analogy to (1) we say that $(X_1, \ldots, X_p)$ takes extreme values in a region $A \subset \Re_+^p \backslash \{(0, \ldots, 0)\}$ sufficiently away from the origin if

$$\Pr((X_1, \ldots, X_p) \in A) = \int_{P(A)} r^{-2} dr H_*(d\mathbf{w}) \quad (5)$$

where H_* is a finite non–negative measure on S_p constrained to have unit means and thus a total mass of p. Following steps analogous to those in Section 2.1 we can see easily that simulation of multivariate extremes based on (5) amounts to simulation of $(r, \mathbf{w})$ coordinates of a Poisson process with intensity $r^{-2}H_*(d\mathbf{w})$. We can also see that by choosing A appropriately the r and $\mathbf{w}$ can be simulated independently of each other: specifically simulate r as in Section 2.2 by the inversion principle and simulate $\mathbf{w}$ from H_*/p. A method for the latter is developed in the following generalisation of the w–step.

3.2 w–step

As before this step requires decomposition of H_* by partitioning S_p into subspaces. For simplification of notation associate with each $\mathbf{w} = (w_1, \ldots, w_{p-1}) \in S_p$ a p–dimensional vector $\bar{\mathbf{w}} = (\bar{w}_1, \ldots, \bar{w}_p)$ where $\bar{w}_i = w_i$ for $i = 1, \ldots, p-1$ and $\bar{w}_p = 1 - \sum_{i=1}^{p-1} w_i$. Then partition S_p into $S_{j,c} = \{\mathbf{w} \in S_p : \bar{w}_k = 0, k \notin c\}$ where $c = \{i_1, \ldots, i_j\}$ is an index variable over subsets of size j of the set $c_p = \{1, \ldots, p\}$. Note $S_{j,c}$ is isomorphic to the $(j-1)$–dimensional unit simplex S_j, so, for example, if $j = 1$ then $S_{1,\{k\}}$ represent vertices of S_p and if $j = p$ then S_{p,c_p} represents the interior of S_p. Now assume H_* has atoms of mass only at the vertices of S_p and is differentiable everywhere else with $h_{j,c}$, $j > 1$ being the $(j-1)$–dimensional density of H_* over $S_{j,c}$. Then H_* is a composition of the atoms of mass at the vertices of S_p and densities $h_{j,c}$ over $S_{j,c}$. Thus we can use the method of composition to simulate $\mathbf{w}$ as belonging to $S_{j,c}$ with probability $p_{j,c}/p$ where $p_{j,c} = \int_{S_{j,c}} H_*(d\mathbf{v})$ is the total measure of $S_{j,c}$: if $j = 1$ then we take $\mathbf{w}$ as the coordinate of the vertex represented by $S_{1,c}$ and otherwise $(j > 1)$ we simulate the value of $\mathbf{w}$ from the probability density $h^*_{j,c} = p_{j,c}^{-1} h_{j,c}$ by means of the rejection method and the following generalisation of Theorem 1.

Theorem 2 *Suppose h^* is a continuous $(j-1)$–dimensional probability density over S_j. Suppose too that there are some constants $q_k > -1, k = 1, \ldots, j$ such that*

$$h^*(\mathbf{w}) = O(w_1^{q_1} \ldots w_{k-1}^{q_{k-1}} w_{k+1}^{q_{k+1}} \ldots w_j^{q_j})$$

holds as $w_l \to 0, l = 1, \ldots, j$, $l \neq k$. Then the Dirichlet density

$$g(w_1, \ldots, w_{j-1}) = \frac{w_1^{q_1} \ldots w_{j-1}^{q_{j-1}} (1 - \sum_{l=1}^{j-1} w_l)^{q_j}}{Be(q_1 + 1, \ldots, q_j + 1)}, \quad (6)$$

where

$$Be(a_1, \ldots, a_j) = \frac{\Gamma(a_1) \ldots \Gamma(a_j)}{\Gamma(a_1 + \ldots + a_j)}$$

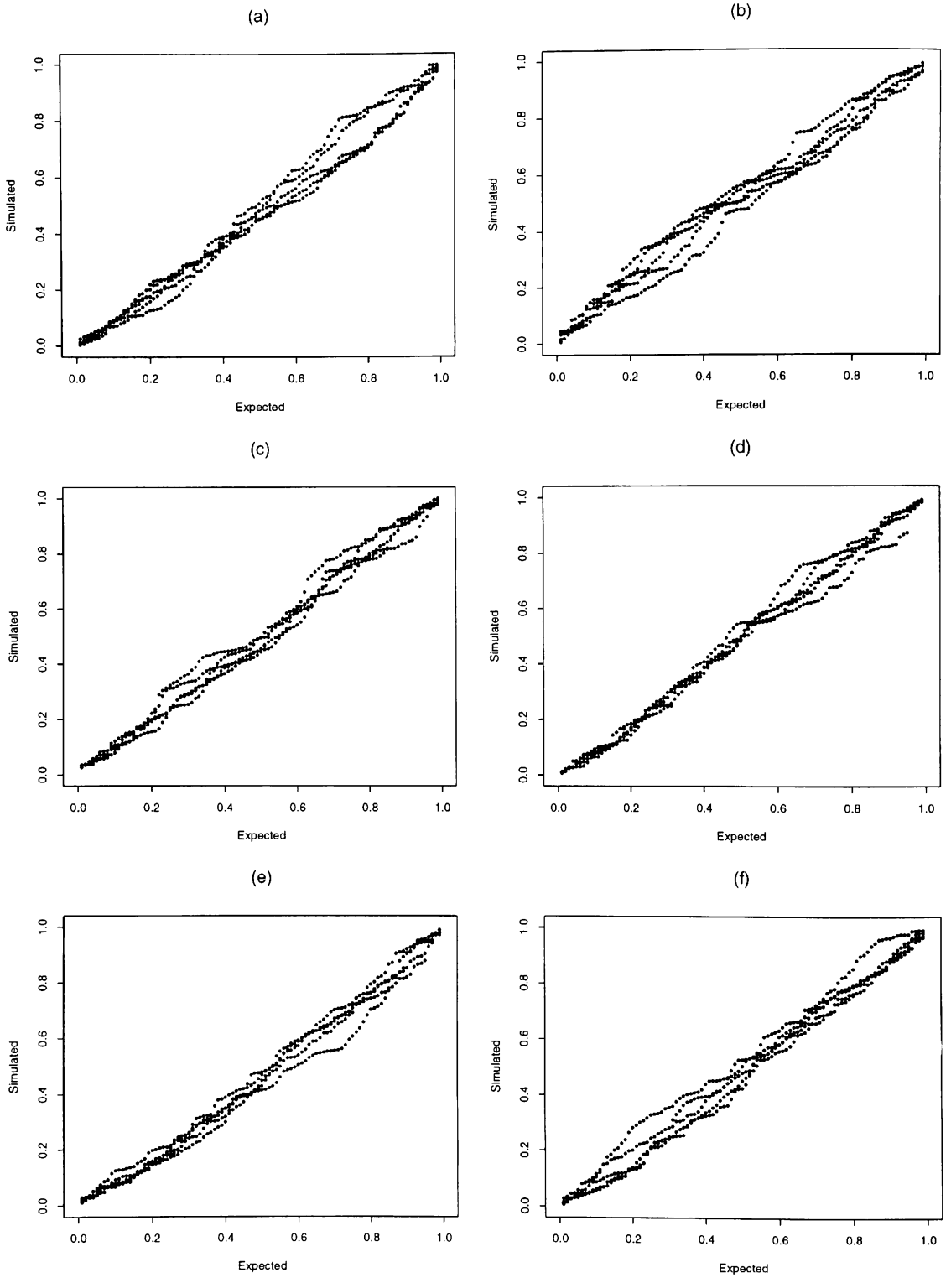

Figure 1: Q–Q plots for the logistic model (3) with parameter values: (a) $(s, \phi_0, \phi_1) = (1.5, 1, 1)$; (b) $(s, \phi_0, \phi_1) = (2, 1, 1)$; (c) $(s, \phi_0, \phi_1) = (4, 1, 1)$; (d) $(s, \phi_0, \phi_1) = (1.5, 0.2, 0.8)$; (e) $(s, \phi_0, \phi_1) = (1.5, 0.3, 0.7)$; (f) $(s, \phi_0, \phi_1) = (4, 0.4, 0.6)$. The five different curves correspond to $v = 0.1, 0.3, \ldots, 0.9$, each based on exactly 100 simulated values of $r_0 Z(v)/\min(v, 1-v)$ not exceeding 1.

$(a_1 > 0, \ldots, a_j > 0)$ satisfies $h^*(w_1, \ldots, w_{j-1})/g(w_1, \ldots, w_{j-1}) \leq M < \infty$ for some constant M.
The proof is not given as it is similar to that of Theorem 1. ■

Replacing h^* in the theorem in turn by each $h^*_{j,c}$, we see that simulation from H_*/p reduces to simulation from a Dirichlet distribution. For the latter, simulation of gamma random variables is sufficient because of the following characterisation result on a Dirichlet random vector (Wilks 1962, Section 7.7): if $G_1, \ldots, G_j$ are independent gamma random variables with means $q_1 + 1, \ldots, q_j + 1$ then $(G_1/\sum_{l=1}^{j} G_l, \ldots, G_{j-1}/\sum_{l=1}^{j} G_l)$ has the Dirichlet distribution of form (6). Gamma random variables can be simulated by standard routines.

An important assumption in Theorems 1 and 2 is that the measure density of H_* behaves as a polynomial term near the vertices of S_p. For the $p = 2$ case all currently known structures for H_* satisfy this assumption (Nadarajah 1994). However it is possible that the assumption might fail for higher dimensions, e.g. nested logistic model (McFadden 1978; Tawn 1990). A relaxed version of the assumption is to approximate the measure density near each vertex by a finite sum of polynomial terms of the form given in the theorems. Under this assumption the envelope g will take the form of a finite mixture of Dirichlet densities and simulation from it can be performed by the method of composition.

4 CONCLUSIONS

Methods for simulation of bivariate and multivariate extremes are introduced. Both simulations are reduced to simulation of bivariate/multivariate Poisson processes with independent coordinates. For the Poisson process simulation two standard techniques, method of composition and rejection method with Dirichlet–type envelope, are exploited. The bivariate method, a method applicable for all currently known structures for H_* (Nadarajah 1994), is shown to perform reasonably for a special class of distributions considered. The multivariate method is less widely applicable because of possible departure from the assumed polynomial behaviour.

ACKNOWLEDGEMENTS

The work was carried out under partial funding from the UK Department of Environment on a project with the Institute of Hydrology (IH). The author wishes to thank Jonathan Tawn (Lancaster University), Clive Anderson (University of Sheffield), Duncan Reed and Ian Dwyer (IH) for helpful suggestions and comments.

REFERENCES

Anderson, C. W. and S. Nadarajah. 1993. Environmental factors affecting reservoir safety. In *Statistics for the Environment I*, ed. V. Barnett and K. F. Turkman, 163–182. John Wiley and Sons.

Coles, S. G. and J. A. Tawn. 1991. Modelling extreme multivariate events. *Journal of Royal Statistical Society Series B* 53:377–392.

Coles, S. G. and J. A. Tawn. 1994. Statistical methods for multivariate extremes: an application to structural design. *Applied Statistics* 43:1–48.

Davison, A. C. and R. L. Smith. 1990. Model for exceedances over high thresholds. *Journal of Royal Statistical Society Series B* 52:393–442.

Dixon, M. J. and J. A. Tawn. 1994. Extreme sea levels: modelling interactions between tide and surge. In *Statistics for the Environment II*, ed. V. Barnett and K. F. Turkman, 221–232. John Wiley and Sons.

Galambos, J. 1978. *The asymptotic theory of extreme order statistics.* John Wiley and Sons.

Joe, H., R. L. Smith, and I. Weissman. 1992. Bivariate threshold methods for extremes. *Journal of Royal Statistical Society Series B* 54:171–183.

Ledford, A. and J. A. Tawn. 1996. Statistics for near independence in multivariate extreme values. *Biometrika* 83:169–187.

McFadden, D. 1978. Modelling the choice of residential location. In *Spatial Interaction Theory and Planning Models*, ed. A. Karlqvist, L. Lundquist, F. Snickers and J. Weibull, 75–96. Amsterdam: North–Holland.

Nadarajah, S. 1994. Multivariate extreme value models with applications to reservoir flood safety. Ph.D. Thesis, School of Mathematics and Statistics, University of Sheffield, Sheffield, U.K.

Resnick, S. I. 1987. *Extreme values, regular variation and point processes.* New York: Springer.

Ripley, B. D. 1987. *Stochastic simulation.* John Wiley and Sons.

Shi, D., R. L. Smith, and S. G. Coles. 1993. Joint versus marginal estimation for bivariate extremes. Submitted.

Smith, R. L., J. A. Tawn, and S. G. Coles. 1993. Markov chain models for threshold exceedances. Submitted.

Tawn, J. A. 1990. Modelling multivariate extreme value distributions. *Biometrika* 77:245–253.

Wilks, S. S. 1962. *Mathematical statistics.* John Wiley and Sons.

AUTHOR BIOGRAPHIES

SARALEES NADARAJAH is a Lecturer in the School of Mathematics and Statistics at the University of Plymouth. Previously he was a research fellow in the School of Mathematics and Statistics at the University of Sheffield. He received M.Sc. and Ph.D. degrees in Statistics from the University of Sheffield in 1991 and 1994 respectively. His research interests include statistical extreme value theory and branching processes.