

# Estimation of Rényi exponents in random cascades

BRENT M. TROUTMAN<sup>1</sup> and ALDO V. VECCHIA<sup>2</sup>

<sup>1</sup>US Geological Survey, Denver Federal Center, BOX 25046, Mail Stop 413, Lakewood, CO 80225, USA. e-mail: troutman@usgs.gov

<sup>2</sup>US Geological Survey, 821 E. Interstate Avenue, Bismarck, ND 58501, USA. e-mail:avecchia@usgs.gov

We consider statistical estimation of the Rényi exponent  $\tau(h)$ , which characterizes the scaling behaviour of a singular measure  $\mu$  defined on a subset of  $R^d$ . The Rényi exponent is defined to be  $\lim_{\delta \rightarrow 0} [\{\log M_\delta(h)\}/(-\log \delta)]$ , assuming that this limit exists, where  $M_\delta(h) = \sum_i \mu^h(\Delta_i)$  and, for  $\delta > 0$ ,  $\{\Delta_i\}$  are the cubes of a  $\delta$ -coordinate mesh that intersect the support of  $\mu$ . In particular, we demonstrate asymptotic normality of the least-squares estimator of  $\tau(h)$  when the measure  $\mu$  is generated by a particular class of multiplicative random cascades, a result which allows construction of interval estimates and application of hypothesis tests for this scaling exponent. Simulation results illustrating this asymptotic normality are presented.

*Keywords:* least-squares estimation; multifractal; multiplicative process; random cascade; Rényi exponent; scaling exponent

## 1. Introduction

We consider in this paper the estimation of the scaling exponents in a multiplicative random cascade model. These models have been applied in a range of settings, including turbulence (Novikov and Stewart 1964; Meneveau and Sreenivasan 1991; She and Waymire 1995), rainfall modelling (Gupta and Waymire 1990; Lovejoy and Schertzer 1990; Over and Gupta 1994; Olsson 1995), river flow modelling (Gupta and Waymire 1990; 1996; Marani *et al.* 1994), clouds (Davis *et al.* 1993) and nonlinear dynamics (Halsey *et al.* 1986). Cascade models are used as a way of defining singular measures with a multifractal structure, which are ideal for modelling the highly intermittent behaviour seen in many natural phenomena. If  $\mu$  is such a measure, possibly random and defined on a subset of  $R^d$ , then the singularities of  $\mu$  can be characterized by a set of *generalized dimensions*  $D_h$  introduced by Hentschel and Procaccia (1983). For  $\delta > 0$ , let  $\{\Delta_i\}$  be the cubes of a  $\delta$ -coordinate mesh that intersect the support of  $\mu$ . Then  $D_h$  is defined, for  $h \neq 1$ , by

$$D_h = \lim_{\delta \rightarrow 0} \left[ \frac{1}{h-1} \left\{ \log \left( \sum_i \mu^h(\Delta_i) \right) \right\} / \log \delta \right]$$

and for  $h = 1$  by

$$D_1 = \lim_{\delta \rightarrow 0} \left\{ \left( \sum_i \mu(\Delta_i) \log \mu(\Delta_i) \right) / \log \delta \right\}.$$

$D_0$  is the fractal dimension of the support of  $\mu$  and  $D_1$  is the *information dimension*. Because of the factor  $h - 1$  in the denominator of this definition,  $D_h$  is not a concave function of  $h$ ; so it is usually more convenient to work with the *Rényi exponents*  $\tau(h)$  defined by

$$\tau(h) = (1 - h)D_h.$$

These exponents constitute the main focus of this paper. In particular, we shall consider the question of statistical estimation of  $\tau(h)$  given a sample assumed to be generated from a random cascade process which produces the singular measure  $\mu$ .

Data which are assumed to be generated from such a process consist of observed masses  $\mu(\Delta_i)$  in intervals  $\Delta_i$  only down to a certain resolution. The main difficulty from a statistical point of view is that, because of the limited resolution of actual observations, there is sampling error in estimating  $\tau(h)$ . Log-linear regression of spatial moments versus  $\delta$  is probably the most commonly used method for estimating  $\tau(h)$  in applications. Despite its widespread use, however, there seem to be few rigorous results on the sampling properties of this least-squares estimator. In this paper, we demonstrate asymptotic normality of the (generalized) least-squares estimator under a random cascade model, which leads to a method of obtaining interval estimates for  $\tau(h)$ . Simulation results illustrating this asymptotic normality are presented.

## 2. Preliminaries

We examine a multiplicative random cascade model defined on  $[0, 1]^d$ , and without loss of generality we take  $d = 1$ . To develop this model, we begin by letting  $b$ , known as the *cascade branching number*, be an integer greater than 1, and let  $\bar{T}$  be the set of infinite sequences of integers in  $\{0, 1, \dots, b - 1\}$ :

$$\bar{T} = \{(\gamma_1, \gamma_2, \dots): \gamma_i \in \{0, 1, \dots, b - 1\}\}.$$

For  $n \geq 1$ , we use  $\gamma|n = (\gamma_1, \dots, \gamma_n)$  to denote the curtailment of a sequence  $\gamma = (\gamma_1, \gamma_2, \dots) \in \bar{T}$  after  $n$  terms, and let  $T(n)$  be the set of  $n$ -term sequences of integers in  $\{0, 1, \dots, b - 1\}$ :

$$T(n) = \{(\gamma_1, \gamma_2, \dots, \gamma_n): \gamma_i \in \{0, 1, \dots, b - 1\}\}.$$

Our interest here is in a sequence of random measures on  $[0, 1]$ ; so we associate elements of  $T(n)$ , which may be thought of as indices of the vertices at distance  $n$  from the root of a tree, with subintervals in the  $n$ th generation  $b$ -adic partition of  $[0, 1]$  by letting

$$\Delta_{\gamma|n} = \left( \sum_{j=1}^n \gamma_j b^{-j}, \sum_{j=1}^n \gamma_j b^{-j} + b^{-n} \right) \subset [0, 1].$$

With each finite sequence  $(\gamma|n) \in T(n)$  we associate a non-negative random variable

$W(\gamma|n)$  defined on a probability space  $(\Omega, \mathcal{F}, P)$ ; these random variables constitute the *cascade generator process*. We shall make the assumption throughout that the  $\{W(\gamma|n): (\gamma|n) \in T(n), 1 \leq n < \infty\}$  are independent and identically distributed with mean one, and we define a sequence of random measures on  $[0, 1]$  by

$$\mu_n(dx) = \prod_{j=1}^n W(\gamma|j) dx, x \in \Delta_{\gamma|n}.$$

This sequence of measures converges almost surely in the weak topology to a random measure  $\mu_\infty$  since, for each bounded continuous function  $f$  on  $[0, 1]$ , the sequence  $\{\int_0^1 f d\mu_n\}$  is an  $L_1$ -bounded martingale with respect to the sequence  $\{\mathcal{F}_n\}$  of sub- $\sigma$ -fields of  $\mathcal{F}$  defined by

$$\mathcal{F}_n = \sigma\{W(\gamma|j): (\gamma|j) \in T(j), 1 \leq j \leq n\}.$$

The limiting mass in the interval  $[0, 1]$  is denoted by  $Z_\infty = \mu_\infty([0, 1])$ . Kahane and Peyrière (1976) proved that this random variable is non-degenerate in the sense that  $EZ_\infty > 0$  if and only if  $\chi'_b(1-) < 0$ , where

$$\chi_b(h) = \log_b EW^h - (h - 1), \tag{1}$$

and  $Z_\infty$  has a finite moment of order  $h > 1$  if and only if

$$h < h_c = \sup\{h \geq 1: \chi_b(h) < 0\}.$$

We shall make the assumption throughout this paper that  $\chi'_b(1-) < 0$ .

In cases when  $W$  has an atom at zero, such as the beta model that is discussed in Section 4, the cascade measure is intermittent, with zero values on sets of positive Lebesgue measure. It can be shown (Gupta and Waymire 1993) that the proportion of non-zero masses at level  $n$  given by

$$p_n(0) = b^{-n} \sum_{\gamma|n} I(\mu_\infty(\Delta_{\gamma|n}) > 0)$$

satisfies

$$\frac{p_n(0)}{P[W > 0]^n} \rightarrow YP[Z_\infty > 0]$$

almost surely, where  $Y$  is a random variable with mean one. Therefore,  $p_n(0)$  converges to zero almost surely whenever  $P[W > 0] < 1$ . This illustrates that random cascade models are non-ergodic in that statistics computed from sample realizations do not converge to the corresponding ensemble statistics. As discussed below, sample moments of random cascades are also non-ergodic owing to a power-law spatial correlation structure of random cascades (see, for example, Cates and Deutsch (1987), Platt and Family (1993) and Meneveau and Chabra (1990)).

Define the  $h$ th moment  $M_n(h)$  by

$$M_n(h) = \sum_{\gamma|n} \mu_\infty^h(\Delta_{\gamma|n}),$$

where the prime indicates summation over those subintervals which meet the support of  $\mu_\infty$ . The Rényi exponent  $\tau(h)$  is defined to be the limit (when it exists; see Theorem 2.2 below) of the sequence

$$\bar{\tau}_n(h) = \frac{\log M_n(h)}{n \log b}.$$

The results in Lemma 2.1 and Theorem 2.2 are from Holley and Waymire (1992).

**Lemma 2.1.** *Let*

$$L_n(h) = \sum_{\gamma|n} \mu_n^h(\Delta_{\gamma|n}).$$

*Then*

$$\lim_{n \rightarrow \infty} \frac{L_n(h)}{(EW^h b^{1-h})^n}$$

*exists almost surely and defines a random variable, say  $Y(h)$ , such that  $EY(h) = 1$ .*

Demonstration of this lemma uses the fact that  $\{L_n(h)/(EW^h b^{1-h})^n, n = 1, 2, \dots\}$  is an  $L_1$ -bounded martingale with respect to  $\{\mathcal{F}_n\}$ .

**Theorem 2.2.** *Assume that  $h < h_c$  and  $EW^{2h}/(EW^h)^2 < b$ . Then, with probability 1,*

(a)

$$\tau(h) := \lim_{n \rightarrow \infty} \bar{\tau}_n(h) = \chi_b(h)$$

*and*

(b)

$$n \log b \{\bar{\tau}_n(h) - \tau(h)\} \rightarrow \log[Y(h)EZ_\infty^h].$$

The proof of Theorem 2.2 involves showing that, with probability one,

$$\frac{M_n(h) - L_n(h)EZ_\infty^h}{(EW^h b^{1-h})^n} \rightarrow 0$$

and using Lemma 2.1. Theorem 2.2 (a) gives strong consistency of  $\bar{\tau}_n(h)$  considered as an estimator of  $\chi_b(h)$  in (1), and Theorem 2.2 (b) gives the scaling that is required to obtain a non-degenerate difference  $\bar{\tau}_n(h) - \tau(h)$ . It would, however, be difficult to use this result to obtain a confidence interval for  $\tau(h) = \chi_b(h)$  because the distribution of  $Y(h)$  and the value of  $EZ_\infty^h$  are not in general known explicitly. This result also does not have desirable invariance properties because the distributions of  $Y(h)$  and  $Z_\infty$  depend on the distribution of  $W$ .

In practice it is more common to use  $M_n(h)$  for several values of  $n$  to estimate  $\tau(h)$  using least squares. In particular, we shall consider least-squares estimation of  $\tau(h)$  using  $k + 1$  pairs  $(\log b^{n+m}, \log M_{n+m}(h))$ ,  $m = 0, 1, \dots, k$ . Define the  $(k + 1) \times 1$  vectors

$$\begin{aligned}\tilde{M}_n(h) &= (\log M_{n+m}(h))^T, \\ \mathbf{u} &= (0, 1, \dots, k)^T, \\ \mathbf{1} &= (1, 1, \dots, 1)^T,\end{aligned}$$

and

$$C(\mathbf{V}) = \frac{1}{\log b} \frac{(1/\mathbf{1}^T \mathbf{V}^{-1} \mathbf{1}) \mathbf{u}^T \mathbf{V}^{-1} - \{\mathbf{1}^T \mathbf{V}^{-1} \mathbf{u} / (\mathbf{1}^T \mathbf{V}^{-1} \mathbf{1})^2\} \mathbf{1}^T \mathbf{V}^{-1}}{\mathbf{u}^T \mathbf{V}^{-1} \mathbf{u} / \mathbf{1}^T \mathbf{V}^{-1} \mathbf{1} - (\mathbf{1}^T \mathbf{V}^{-1} \mathbf{u} / \mathbf{1}^T \mathbf{V}^{-1} \mathbf{1})^2},$$

where  $\mathbf{V}$  is a non-singular  $(k + 1) \times (k + 1)$  matrix to be specified below. Then the least-squares estimator of  $\tau(h)$  is defined by

$$\hat{\tau}_n(h; \mathbf{V}) = C(\mathbf{V})^T \tilde{M}_n(h).$$

Our intention in this paper is to obtain a result on asymptotic normality for such an estimator. The estimate of the asymptotic variance of this distribution will be shown to depend only on observed masses  $\mu_\infty(\Delta_{\gamma|j})$ ,  $n \leq j \leq n + k$ , and not on the distribution of  $W$ . The result may therefore be readily used for interval estimation of  $\tau(h)$ . We shall present in Corollary 3.5 the asymptotic distribution of both the ordinary least-squares estimator, for which  $\mathbf{V}$  is the identity matrix, and the generalized least-squares estimator, for which  $\mathbf{V}$  is an estimator of the covariance matrix of  $\tilde{M}_n(h)$ . The generalized least-squares estimator will have minimum variance among linear unbiased estimators of  $\tau(h)$ .

### 3. Convergence of Rényi exponent least-squares estimators

The basic idea behind the results that follow can be seen by expressing  $M_n(h)$  in the form (Holley and Waymire 1992, Proposition 2.3)

$$M_n(h) = \sum_{\gamma|n} \mu_n^h(\Delta_{\gamma|n}) Z_\infty^h(\gamma|n), \tag{2}$$

where the  $b^n$  random variables  $Z_\infty(\gamma|n)$  are mutually independent, are independent of the  $\mu_n(\Delta_{\gamma|n})$  and are identically distributed as the limiting mass  $Z_\infty$ . This representation of  $M_n(h)$  may be exploited to obtain results on the asymptotic distribution of functions of these moments. In the treatment that follows, we shall consider a somewhat more general problem by looking at a multivariate version of moments such as those in (2). Let  $\mathbf{Q}_{\gamma|n} = (\mathbf{Q}_{\gamma|n}^{(0)}, \dots, \mathbf{Q}_{\gamma|n}^{(k)})^T$ ,  $1 \leq n < \infty$  be independent and identically distributed  $(k + 1) \times 1$ ,  $k \geq 0$ , random vectors with finite second moments. Define

$$\begin{aligned}\mathbf{p} &= (p^{(m)})^T = E\mathbf{Q}_{\gamma|n}, \\ \mathbf{A} &= \text{var}(\mathbf{Q}_{\gamma|n}), \\ \mathbf{X}_{\gamma|n}(h) &= (X_{\gamma|n}^{(m)}(h))^T = \mu_n^h(\Delta_{\gamma|n}) \mathbf{Q}_{\gamma|n},\end{aligned}$$

$$\mathbf{S}_n(h) = (S_n^{(m)}(h))^T = \sum_{\gamma|n} X_{\gamma|n}(h).$$

In all that follows, in addition to assuming that the  $\mathbf{Q}_{\gamma|n}$  are independent and identically distributed with finite variance, we shall assume that the  $\mathbf{Q}_{\gamma|n}$  are  $\mathcal{F}_n$  measurable, where

$$\mathcal{F}_n = \sigma\{W(\gamma|j) : (\gamma|j) \in T(j), n < j < \infty\}.$$

The result in Proposition 3.1 gives a law of large numbers for randomly weighted sums of the  $\mathbf{Q}_{\gamma|n}$ . It is proved using a method similar to that used in proving Theorem 2.7 in Holley and Waymire (1992).

**Proposition 3.1.** *Assume that  $EW^{2h}/(EW^h)^2 < b$  and  $\text{var}(\mathbf{Q}_{\gamma|n}^{(m)}) < \infty$ ,  $m = 0, 1, \dots, k$ . Then*

$$\frac{\mathbf{S}_n(h)}{L_n(h)} \rightarrow \mathbf{p} \text{ almost surely.}$$

The idea in obtaining weak convergence of Rényi exponent estimators is to find the scaling that yields a non-degenerate limit of  $L_n^{-1}(h)\mathbf{S}_n(h) - \mathbf{p}$ . Because the least-squares estimator of  $\tau(h)$  can be expressed in terms of  $\log S_n^{(m)}(h)$  for a particular choice of variables  $\mathbf{Q}_{\gamma|n}$ , we introduce a logarithmic transform as well. Let

$$\begin{aligned} \tilde{\mathbf{p}} &= (\tilde{p}^{(m)})^T = (\log p^{(m)})^T, \\ \tilde{\mathbf{S}}_n &= (\tilde{S}_n^{(m)}(h))^T = (\log S_n^{(m)}(h))^T, \\ \mathbf{D} &= \text{diag}\left(\frac{1}{p^{(m)}}\right) \end{aligned}$$

and

$$\tilde{\mathbf{A}} = \mathbf{D}\mathbf{A}\mathbf{D}^T.$$

**Proposition 3.2.** *Assume that  $EW^{2h}/(EW^h)^2 < b$ , and that there exists  $\delta > 1$  such that  $EW^{2h\delta}/(EW^{2h})^\delta < b^{\delta-1}$  and  $E|\mathbf{Q}_{\gamma|n}^{(m)} - p^{(m)}|^{2\delta} < \infty$ ,  $m = 0, 1, \dots, k$ . Then*

$$\frac{L_n(h)}{L_n^{1/2}(2h)} (\tilde{\mathbf{S}}_n(h) - \log L_n(h) \mathbf{1} - \tilde{\mathbf{p}}) \xrightarrow{d} N_{k+1}(\mathbf{0}, \tilde{\mathbf{A}}),$$

where  $N_{k+1}(\mathbf{0}, \tilde{\mathbf{A}})$  is a  $(k + 1)$  normal variate with mean  $\mathbf{0}$  and variance  $\tilde{\mathbf{A}}$ .

**Proof.** We first note that  $L_n(h)$ ,  $L_n(2h)$ , and the  $\mu_n^h(\Delta_{\gamma|n})$  are measurable  $\mathcal{F}_n$ . Thus,

$$\text{var}\left(\frac{L_n(h)}{L_n^{1/2}(2h)} \{L_n^{-1}(h)\mathbf{S}_n(h) - \mathbf{p}\} \middle| \mathcal{F}_n\right) = \frac{1}{L_n(2h)} \sum_{\gamma|n} \text{var}\{\mu_n^h(\Delta_{\gamma|n})(\mathbf{Q}_{\gamma|n} - \mathbf{p}) \middle| \mathcal{F}_n\} = \mathbf{A}$$

and

$$\begin{aligned} & \frac{1}{L_n^\delta(2h)} \sum_{\gamma|n} \mu_n^{2\delta h}(\Delta_{\gamma|n}) \mathbb{E} \|\mathbf{Q}_{\gamma|n} - \mathbf{p}\|^{2\delta} \\ &= \frac{L_n(2\delta h)}{L_n^\delta(2h)} \mathbb{E} \|\mathbf{Q}_{\gamma|n} - \mathbf{p}\|^{2\delta} \\ &= \frac{L_n(2\delta h)}{(b^{1-2\delta h} \mathbb{E} W^{2\delta h})^n} \left( \frac{L_n(2h)}{(b^{1-2h} \mathbb{E} W^{2h})^n} \right)^{-\delta} \left( b^{1-\delta} \frac{\mathbb{E} W^{2\delta h}}{(\mathbb{E} W^{2h})^\delta} \right)^n \mathbb{E} \|\mathbf{Q}_{\gamma|n} - \mathbf{p}\|^{2\delta} \rightarrow 0, \end{aligned}$$

almost surely, using Lemma 2.1. We therefore apply a multivariate version of the Lindeberg–Feller central limit theorem (see, for example, Serfling (1980, Theorem 1.9.2B)) to establish that, for every  $x \in R^{k+1}$ ,

$$\lim_{n \rightarrow \infty} P \left( \frac{L_n(h)}{L_n^{1/2}(2h)} \{L_n^{-1}(h) \mathbf{S}_n(h) - \mathbf{p}\} \leq x | \mathcal{F}_n \right) = \Phi(x; \mathbf{0}, \mathbf{A})$$

almost surely, where the limit is a  $(k + 1)$ -variate normal cumulative distribution function with mean  $\mathbf{0}$  and variance  $\mathbf{A}$ . The result in the proposition then follows using the dominated convergence theorem after a logarithmic transformation (Serfling 1980, Theorem 3.3A).  $\square$

In order to make this result useful for obtaining interval estimators of the Rényi exponent, we would like to express the estimated asymptotic variance in terms of  $\mathbf{X}_{\gamma|n}(h)$ , which will be assumed to be observable. Define the  $(k + 1) \times 1$  vector

$$\mathbf{K}_{\gamma|n}(h) = (K_{\gamma|n}^{(m)}(h))^T = \begin{pmatrix} X_{\gamma|n}^{(m)}(h) \\ S_n^{(m)}(h) \end{pmatrix}^T$$

and let

$$\mathbf{V}_n(h) = \sum_{\gamma|n} \mathbf{K}_{\gamma|n}(h) \mathbf{K}_{\gamma|n}^T(h).$$

**Proposition 3.3.** *Assume that  $\mathbb{E} W^{4h} / (\mathbb{E} W^{2h})^2 < b$  and that  $\mathbb{E} |\mathcal{Q}_{\gamma|n}^{(m)} - p^{(m)}|^4 < \infty$ ,  $m = 0, 1, \dots, k$ . Then*

$$\frac{L_n^2(h)}{L_n(2h)} \mathbf{V}_n(h) - \mathbf{11}^T \rightarrow \tilde{\mathbf{A}}$$

almost surely.

**Proof.** This follows directly from the result in Proposition 3.1. Note that  $\mathbb{E} W^{4h} / (\mathbb{E} W^{2h})^2 < b$  implies that  $\mathbb{E} W^{2h} / (\mathbb{E} W^h)^2 < b$ .  $\square$

We now return to estimation of  $\tau(h)$ . The following theorem follows directly from Propositions 3.2 and 3.3.

**Theorem 3.4.** *Assume that  $\mathbb{E} W^{4h} / (\mathbb{E} W^{2h})^2 < b$  and that  $\mathbb{E} |\mathcal{Q}_{\gamma|n}^{(m)} - p^{(m)}|^4 < \infty$ ,  $m =$*

0, 1, . . . , k. Let,  $\mathbf{c}_n$  be a random  $(k + 1) \times 1$  vector such that  $\mathbf{c}_n^T \tilde{\mathbf{p}}$  converges almost surely to a finite limit, say  $\theta(h)$ , and  $\mathbf{c}_n^T \mathbf{1} \rightarrow 0$  almost surely. Then  $\mathbf{c}_n^T \tilde{\mathbf{S}}_n(h)$  obeys

$$\frac{\mathbf{c}_n^T \tilde{\mathbf{S}}_n(h) - \theta(h)}{(\mathbf{c}_n^T \mathbf{V}_n(h) \mathbf{c}_n)^{1/2}} \xrightarrow{d} N(0, 1), \tag{3}$$

where  $N(0, 1)$  is a standard normal variate.

Application to the least-squares estimator is given in the following corollary.

**Corollary 3.5.** Assume that  $4h < h_c$  and that  $EW^{4h}/(EW^{2h})^2 < b$ . Define

$$\mathbf{K}_{\gamma|n}^*(h) = \left( \frac{M_{\gamma|n}^{(m)}(h)}{M_{n+m}(h)} \right)^T$$

and

$$\mathbf{V}_n^*(h) = \sum_{\gamma|n} \mathbf{K}_{\gamma|n}^*(h) \mathbf{K}_{\gamma|n}^{*T}(h),$$

where  $M_{\gamma|n}^{(0)}(h) = \mu_\infty^h(\Delta_{\gamma|n})$  and

$$M_{\gamma|n}^{(m)}(h) = \sum_{\gamma_{n+1} \dots \gamma_{n+m}} \mu_\infty^h(\Delta_{\gamma|(n+m)}), \quad m = 1, \dots, k.$$

Then the ordinary least-squares estimator of  $\tau(h)$  obeys

$$\frac{\hat{\tau}_n(h; \mathbf{I}) - \tau(h)}{\{\mathbf{C}(\mathbf{I})^T \mathbf{V}_n^*(h) \mathbf{C}(\mathbf{I})\}^{1/2}} \xrightarrow{d} N(0, 1)$$

and the generalized least-squares estimator of  $\tau(h)$  obeys

$$\frac{\hat{\tau}_n(h; \mathbf{V}_n^*(h)) - \tau(h)}{\{\mathbf{C}(\mathbf{V}_n^*(h))^T \mathbf{V}_n^*(h) \mathbf{C}(\mathbf{V}_n^*(h))\}^{1/2}} \xrightarrow{d} N(0, 1).$$

**Proof.** This follows directly from Theorem 3.4 if we note that  $M_{n+m}(h)$  may be written as

$$M_{n+m}(h) = \sum_{\gamma|n} M_{\gamma|n}^{(m)}(h) = \sum_{\gamma|n} \mu_\infty^h(\Delta_{\gamma|n}) \mathcal{Q}_{\gamma|n}^{*(m)},$$

where

$$\mathcal{Q}_{\gamma|n}^{*(0)} = b^{-kh} \left( \sum_{\gamma_{n+1} \dots \gamma_{n+k}} \prod_{j=n+1}^{n+k} W(\gamma|j) Z_\infty(\gamma|(n+k)) \right)^h,$$

$$\mathcal{Q}_{\gamma|n}^{*(k)} = b^{-kh} \sum_{\gamma_{n+1} \dots \gamma_{n+k}} \prod_{j=n+1}^{n+k} W^h(\gamma|j) Z_\infty^h(\gamma|(n+k))$$

and, for  $m = 1, \dots, k - 1$ ,

$$Q_{\gamma|n}^{*(m)} = b^{-kh} \sum_{\gamma_{n+1} \dots \gamma_{n+m}} \prod_{j=n+1}^{n+m} W^{h(\gamma|j)} \left( \sum_{\gamma_{n+m+1} \dots \gamma_{n+k}} \prod_{j=n+m+1}^{n+k} W(\gamma|j) Z_{\infty}(\gamma|(n+k)) \right)^h.$$

It is seen that  $Q_{\gamma|n}^{*(m)}$  defined this way are  $\mathcal{S}_n$  measurable and have finite fourth moments because  $4h < h_c$ . The conditions  $\mathbf{C}(\mathbf{I})^T \mathbf{1} = 0$ ,  $\mathbf{C}(V_n^*(h))^T \mathbf{1} = 0$ ,  $\mathbf{C}(\mathbf{I})^T \tilde{\mathbf{p}}^* = \tau(h)$ , and  $\mathbf{C}(V_n^*(h))^T \tilde{\mathbf{p}}^* = \tau(h)$ , where

$$\tilde{\mathbf{p}}^{*(m)} = \log E Q_{\gamma|n}^{*(m)} = m(\log b)\tau(h) + \log E Z_{\infty}^h$$

are easily verified. □

### 4. Application to simulated cascades

The finite-sample properties of the least-squares estimators given by Corollary 3.5 are explored in this section using simulated data from three cascade models (Table 1). The first model, for which the distribution of  $W$  is concentrated on 0 and  $\frac{3}{2}$  is often referred to as a beta cascade. It has been used to model energy dissipation in fluid turbulence (Novikov and Stewart 1964) and in the modelling of spatial rainfall rates (Over and Gupta 1994), among other applications. For the second model,  $W$  is gamma distributed and, for the third model,  $W$  is uniformly distributed on  $[0, 2]$ . Values of  $h_c$  and  $h_4$  for these models are given in Table 1, where  $h_4$  is defined by

$$h_4 = \sup \left\{ h \geq 1 : \frac{E W^{4h}}{(E W^{2h})^2} < b \right\}.$$

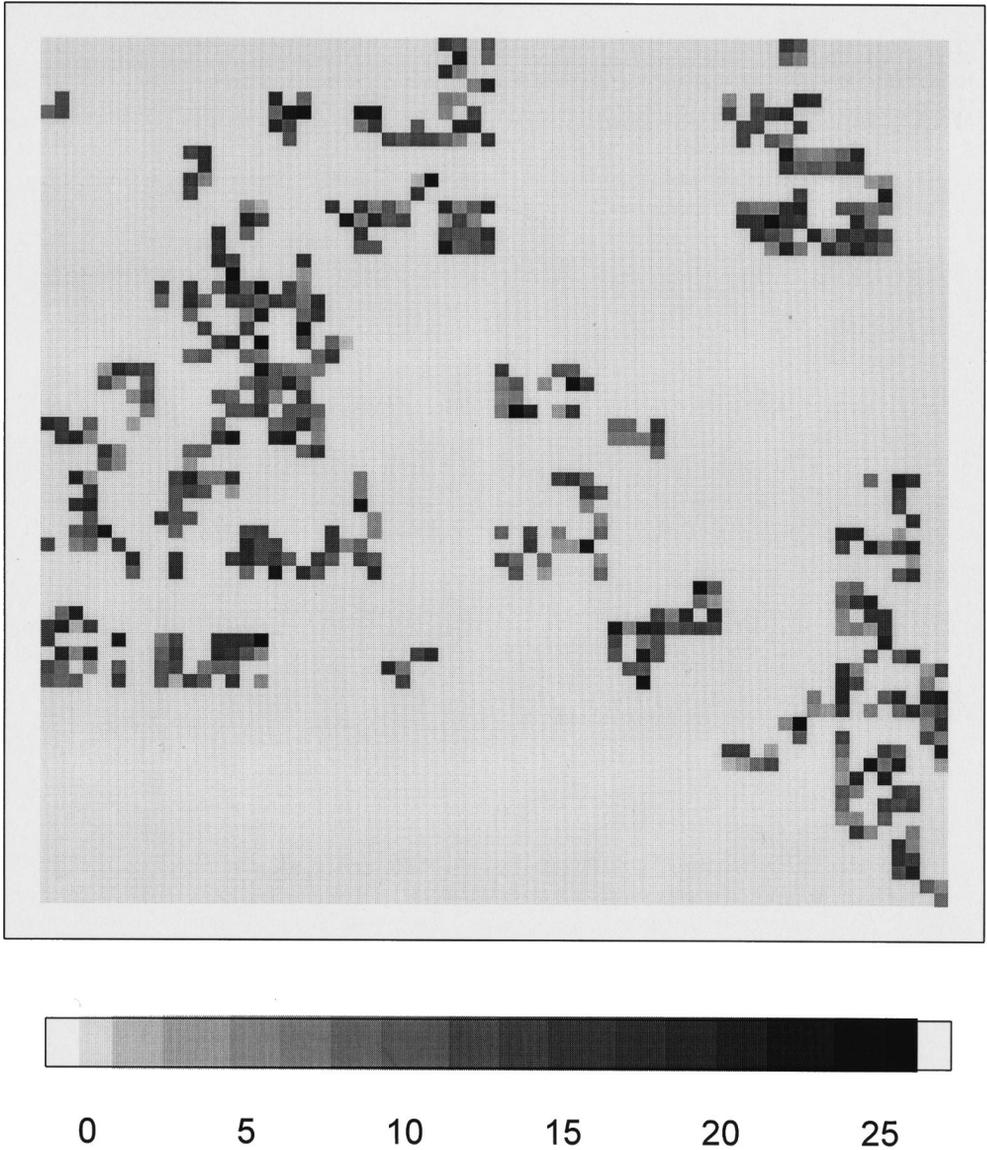
In order to give spatial structure to the data, all simulations are based on a two-dimensional cascade with  $b = 4$ . The results of Section 3 remain valid for two-dimensional cascades on the unit square by letting each cell of the two-dimensional model at level  $n$  be the Cartesian product of two one-dimensional partitions of the unit interval.

For the models in Table 1, cascade realizations were generated down to level  $n = 6$ . The high-frequency component  $Z_{\infty}$  cannot be exactly duplicated in practice. However, an adequate approximation was obtained by generating independent “mini-cascades” for each cell at level 6, down an additional seven levels, and computing the total mass of the resulting mini-cascade. This algorithm was used by Over and Gupta (1994) to generate cascade realizations.

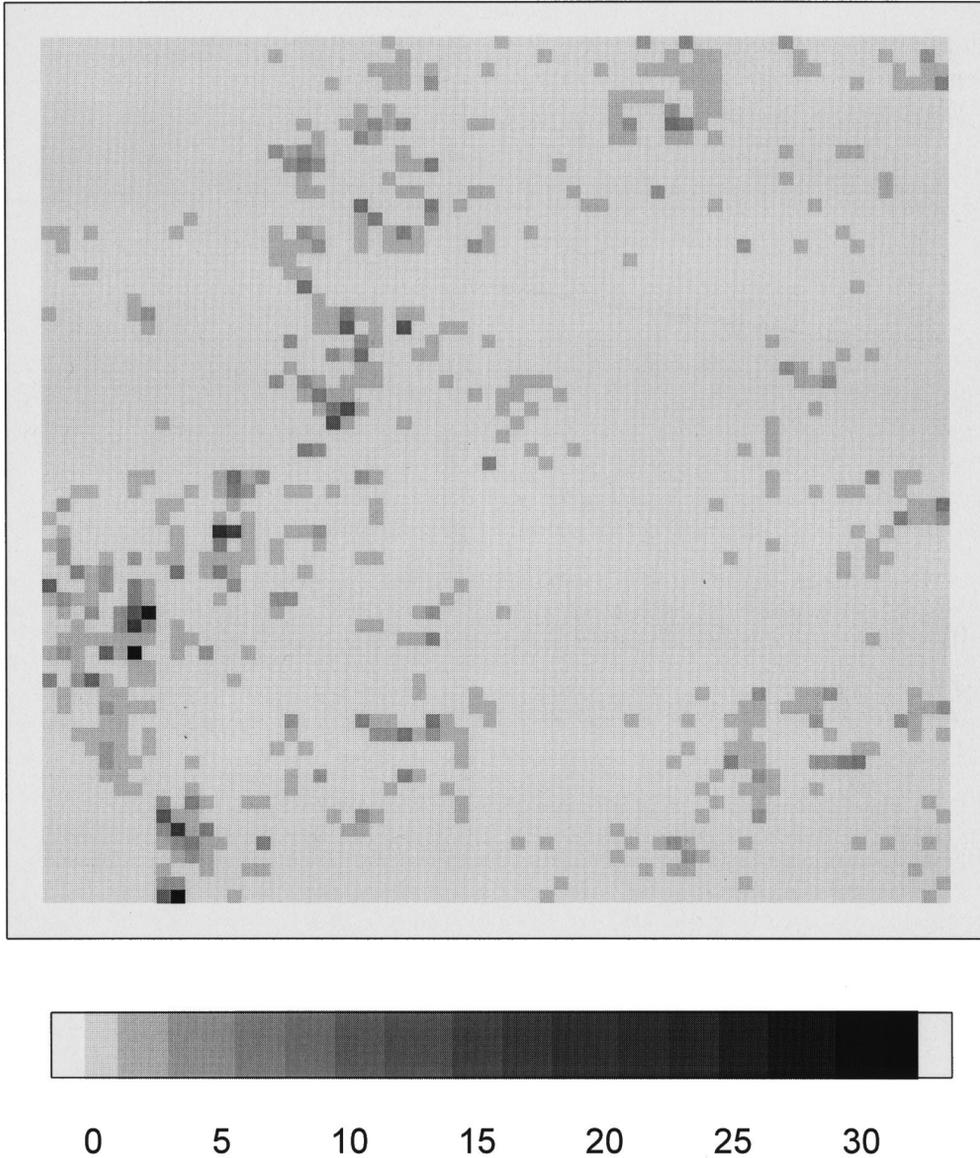
**Table 1.** Cascade generators used for simulated data sets ( $b = 4$  in all three cases)

Identifier	Probability distribution function of $W$ ( $f(w)$ )	$\chi(h)$	$h_c$	$h_4$
Beta	$\frac{1}{3}I(w = 0) + \frac{2}{3} I(w = \frac{3}{2})$	$(h - 1)(\log_4 3 - \frac{3}{2})$	$+\infty$	$+\infty$
Gamma	$(w/4) \exp(-2w)$	$1 - \frac{3}{2}h + \log_4 \Gamma(2 + h)$	13.73	1.10
Uniform	$\frac{1}{2}I(0 < w < 2)$	$1 - h/2 - \log_4(h + 1)$	$+\infty$	3.23

Examples of individual realizations are given in Figure 1 for the beta cascade and in Figure 2 for the gamma cascade. These figures give the densities (mass per unit area)  $b^6 \mu_\infty(\Delta_{\gamma|6})$  at level  $n = 6$ . Several interesting properties of random cascades are illustrated in Figures 1 and 2. For example, the support of the measure is almost surely given by a set of Hausdorff dimension  $-2\chi'_b(1)$ , which is less than 2 except in the degenerate case when



**Figure 1.** Example realization from the beta cascade with  $n = 6$ .



**Figure 2.** Example realization from the gamma cascade with  $n = 6$ .

$W = 1$  with probability one (Kahane and Peyriere 1976). Therefore, the cascade measure is singular with respect to Lebesgue measure. Mandelbrot (1974) shows that, as the scale becomes finer and finer, the measure is dominated by a small proportion of cells in which the masses are shrinking at a rate proportional to  $(b^{-n})^{-\chi_b^{(1)}}$ .

Non-ergodicity makes statistical inference for random cascade models problematic.

However, the results of Section 3 provide a methodology for placing confidence intervals on the Rényi exponents for various values of  $h$ , which in turn can be used to make inferences about the distribution of  $W$  based on a single-cascade realization. To investigate the practical usefulness of the least-squares estimators defined in Corollary 3.5, 100 independent cascade realizations were generated from each of the models in Table 1. To simplify the presentation, all the simulation results are with  $k = 1$ . Therefore, only two values ( $\log M_n(h)$  and  $\log M_{n+1}(h)$ ) are used to estimate  $\tau(h)$  for any given  $h$  and  $n$ , in which case the ordinary and generalized least-squares estimators are both given by

$$\hat{\tau}_n(h) = \frac{-\{\log M_{n+1}(h) - \log M_n(h)\}}{\log b}.$$

For these models, it was found that using only the two smallest scales ( $n + 1$  and  $n$ ) to estimate  $\tau(h)$  gives similar results to using three or more scales, and that the ordinary least-squares and generalized least-squares estimators are similar. However, using other cascade generators and considering scales finer than  $n = 6$  may result in significant differences between the estimators with  $k = 1$  and  $k > 1$ , and between the ordinary and generalized least-squares estimators. Results are presented for  $n = 4$  and  $n = 5$ . The results for  $n = 1, 2$  and  $3$  are not given because the asymptotic theory was a poor approximation for  $n$  this small. In practice, a reasonable approach would be to increase  $n$  until the estimator remained approximately constant from one  $n$  to the next. To illustrate the effect of moment order, results are given for  $h = 0.5, 2, 3$  and  $4$ , which bracket the values that are usually of most interest in practice (the Rényi exponent for  $h = 1$  is identically zero; so  $h = 1$  is not of interest). In addition, results for  $h = 0$  are presented for the beta model; in that case  $2\chi_b(0)$  can be interpreted as the fractal dimension of the support of  $\mu_\infty$ .

Table 2 gives summary statistics for the beta model computed from 100 realizations. Included are the tenth, fiftieth and ninetieth percentiles of the estimated Rényi exponents, as well as the tenth, fiftieth and ninetieth percentiles of the normalized values

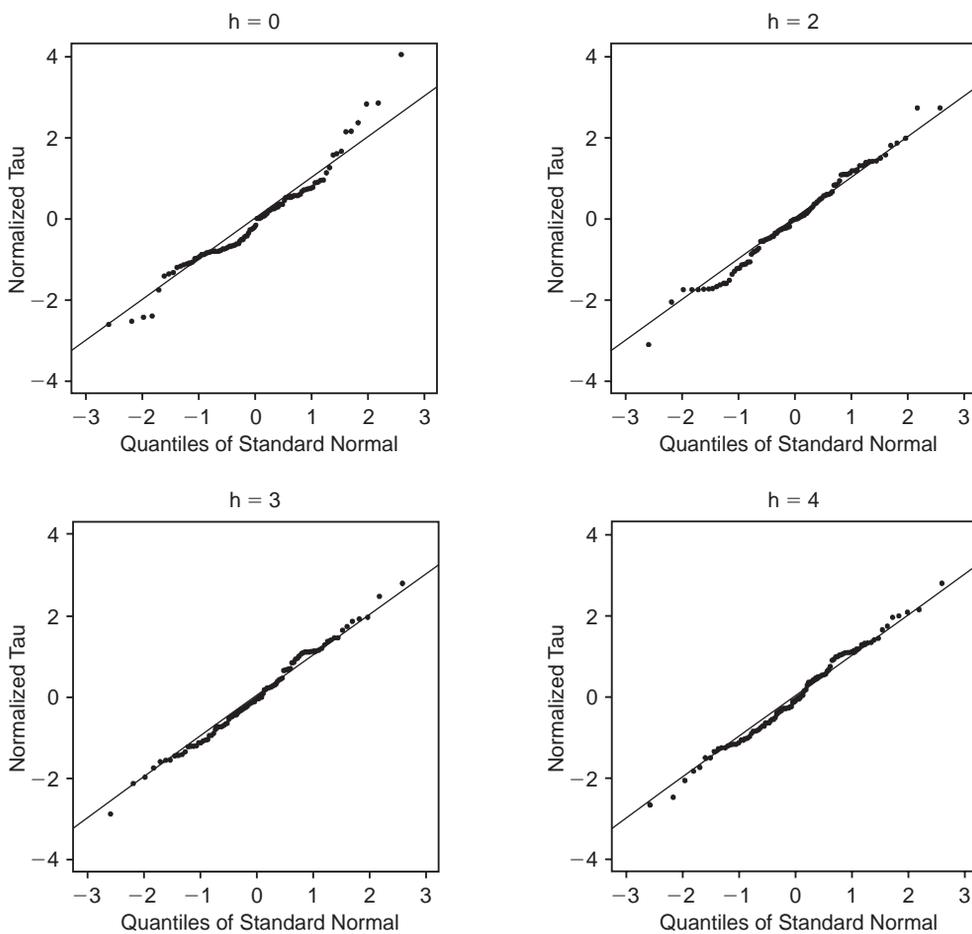
**Table 2.** Simulation results for the beta model

$n$	$h$	$\tau(h)$	$\hat{\tau}(h)$ (10%)	$\hat{\tau}(h)$ (50%)	$\hat{\tau}(h)$ (90%)	$z(h)$ (10%)	$z(h)$ (50%)	$z(h)$ (90%)
4	0	0.707	0.669	0.706	0.749	-0.901	-0.023	1.454
4	0.5	0.354	0.336	0.354	0.374	-1.085	0.034	1.480
4	2	-0.707	-0.746	-0.706	-0.659	-1.429	0.023	1.444
4	3	-1.415	-1.488	-1.414	-1.295	-1.403	-0.014	1.730
4	4	-2.123	-2.227	-2.123	-1.926	-1.303	-0.010	1.745
5	0	0.707	0.680	0.705	0.732	-1.072	-0.130	1.179
5	0.5	0.354	0.340	0.354	0.367	-1.338	-0.014	1.309
5	2	-0.707	-0.734	-0.705	-0.676	-1.497	0.070	1.462
5	3	-1.415	-1.468	-1.415	-1.353	-1.298	0.003	1.426
5	4	-2.123	-2.200	-2.125	-2.022	-1.197	-0.025	1.378

$$z_n(h) = \frac{\hat{\tau}_n(h) - \tau(h)}{\{\mathbf{C}(\mathbf{I})^T \mathbf{V}_n^*(h) \mathbf{C}(\mathbf{I})\}^{1/2}}$$

(see Corollary 3.5). Figure 3 gives normal probability plots of the  $z$  values for the case  $n = 5$ . Because  $h_c$  and  $h_4$  are both infinite (Table 1), there are no restrictions on the  $h$  values for which Corollary 3.5 holds. As seen in Table 2 and Figure 3, the asymptotic normal approximation is accurate for all  $h$  values and can be used to obtain good confidence intervals on  $\tau(h)$  for all moment orders.

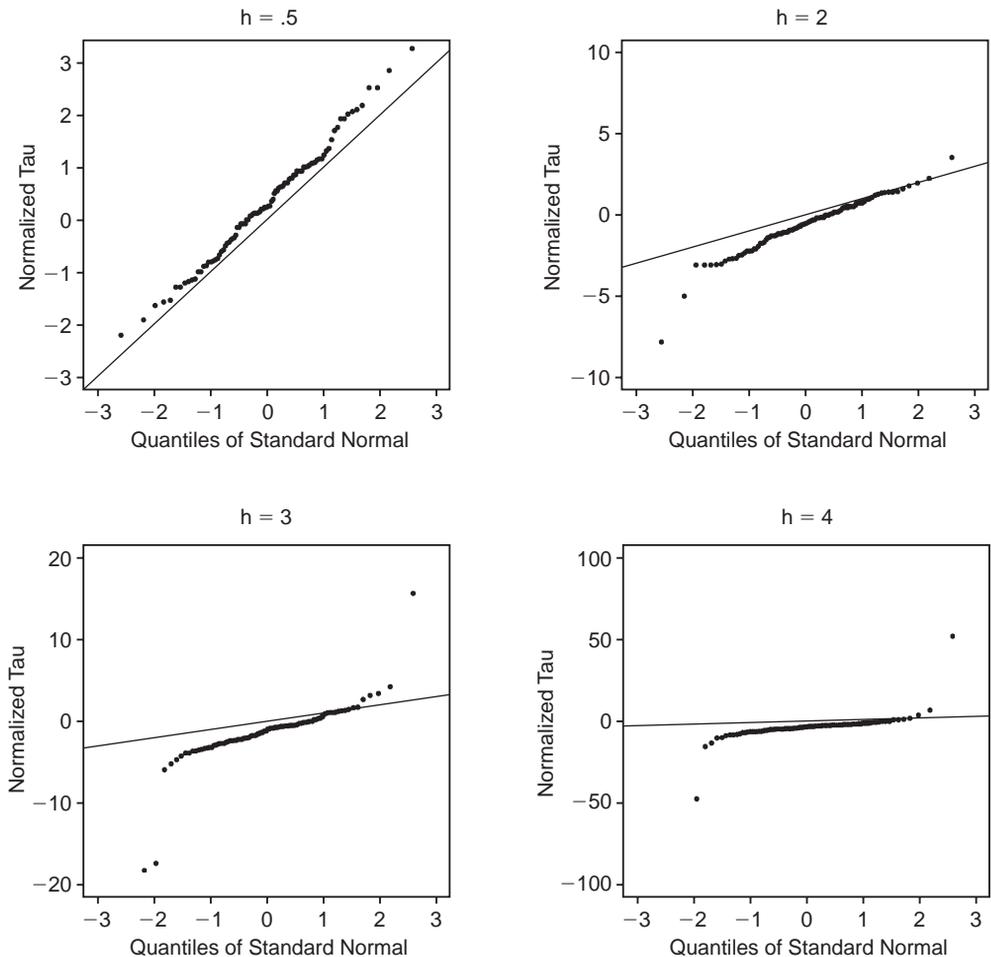
The gamma model is fundamentally different from the beta model in that  $W$  is not bounded above by  $b$ , in which case  $h_c$  is finite. Mandelbrot (1974) refers to such cases as irregular cascades. Realizations from the gamma model tend to be much more sparse and



**Figure 3.** Normal probability plots of  $z$  values based on 100 simulations from the beta cascade.

variable than realizations from the beta model, and the high-intensity cells tend to be more isolated. For this model,  $h_c/4 = 3.31$  and  $h_4 = 1.10$ ; so Corollary 3.5 holds only for  $h < 1.10$ . The estimation results in Table 3 and Figure 4 show good agreement with the asymptotic normality when  $h = 0.5$ . However, higher values of  $h$  indicate asymmetric heavy-tailed limit distributions.

Finally, for the uniform model we shall not show the probability plots but point out that agreement with normality is generally good for  $h < h_4 = 3.23$ , although there were two outliers for  $h = 3$ , indicating that larger values of  $n$  may be necessary when  $h$  is close to  $h_4$ . Results for  $h = 4$  revealed a pronounced deviation from normality, indicating that the restriction  $h < h_4$  may be a necessary as well as sufficient condition for the asymptotic results in Corollary 3.5 to hold.



**Figure 4.** Normal probability plots of  $z$  values based on 100 simulations from the gamma cascade.

**Table 3.** Simulation results for the gamma model

$n$	$h$	$\tau(h)$	$\hat{\tau}(h)$ (10%)	$\hat{\tau}(h)$ (50%)	$\hat{\tau}(h)$ (90%)	$z(h)$ (10%)	$z(h)$ (50%)	$z(h)$ (90%)
4	0.5	0.455	0.452	0.455	0.459	-1.160	0.097	1.493
4	2	-0.708	-0.771	-0.713	-0.632	-2.556	-0.155	2.040
4	3	-1.207	-1.484	-1.258	-0.970	-4.850	-0.571	2.693
4	4	-1.546	-2.222	-1.778	-1.223	-9.703	-1.378	2.486
5	0.5	0.455	0.453	0.455	0.457	-1.273	0.034	1.554
5	2	-0.708	-0.764	-0.716	-0.642	-2.507	-0.288	1.483
5	3	-1.207	-1.461	-1.262	-1.047	-3.247	-0.668	1.594
5	4	-1.546	-2.138	-1.730	-1.322	-6.124	-1.355	1.670

## 5. Conclusions

Accurate estimation of Rényi exponents is of primary concern in determining the scaling properties of random fields for resolutions that are smaller than the finest scale of measurement. For example, in hydrologic modelling of rainfall-runoff processes, measurement of average rainfall rates over large areas (say, 100 km<sup>2</sup>) are routinely available using remote sensing or output from regional-scale climate models. However, the resolution of the rainfall-runoff model may need to be much finer, of the order of 4 km<sup>2</sup>, to obtain accurate representation of runoff events in small basins. Small differences in the Rényi exponent can have a large effect on the spatial distribution of rain rates over small areas. Therefore, it is important to include uncertainty in estimated Rényi exponents when exploring the spatial variability of rainfall and runoff volumes at fine scales. For example, the uniform and gamma models discussed in Section 4 have similar Rényi exponents for small  $h$  but produce radically different realizations. Using the gamma cascade as a model for spatial rainfall over a river basin would produce much more variability in rainfall and flood peaks in small subbasins than would the uniform cascade model.

Estimation of Rényi exponents in previous studies has been primarily using ordinary least-squares regression of the logarithm of transformed moments versus the logarithm of scale, without any objective measures for assessing goodness of fit or estimation errors. Efforts to assess estimation error have focused on Monte Carlo simulation studies which are model specific (i.e. driven by a particular distribution of the cascade generator) and are therefore not readily generalizable. In this paper, we have given a general methodology for obtaining confidence intervals or testing hypotheses regarding the Rényi exponents of random cascades. The technique can be used in conjunction with the ordinary least-squares estimator as well as a generalized least-squares estimator that takes into account the correlation between moments at different scales. The main result (Corollary 3.5) depends on the distribution of the cascade generator  $W$  only through a restriction that  $h$  does not exceed an upper bound that depends on  $W$ . Therefore, to apply the result in practice requires some

assurance that  $h$  does not exceed the critical upper limit. Although we are working on ways to generalize the asymptotic results to include non-normal limit distributions for large  $h$ , there are significant difficulties to overcome before a reasonable approach can be formulated.

## References

- Cates, M.E. and Deutsch, J.M. (1987) Spatial correlations in multifractals. *Phys. Rev. A*, **35**, 4907–4910.
- Davis, A., Marshak, A. and Wiscombe, W.J. (1993) Bi-multifractal analysis and multi-affine modeling of non-stationary geophysical processes, application to turbulence and clouds. *Fractals*, **1**, 560–567.
- Gupta, V.K. and Waymire, E.C. (1990) Multiscaling properties of spatial rainfall and river flow distributions. *J. Geophys. Res.*, **95**, 1999–2009.
- Gupta, V.K. and Waymire, E.C. (1993) A statistical analysis of mesoscale rainfall as a random cascade. *J. Appl. Meteorol.*, **32**, 251–267.
- Gupta, V.K. and Waymire, E.C. (1996) Multiplicative cascades and spatial variability in rainfall, river networks, and floods. In J.B. Rundle, D.L. Turcotte and W. Klein (eds), *Reduction and Predictability of Natural Disasters*, pp. 71–96. Santa Fe Institute Proc., Vol. XXV. Reading, MA: Addison-Wesley.
- Halsey, T.C., Jensen, M.H., Kadanoff, L.P. Procaccia, I. and Shraiman, B.I. (1986) Fractal measures and their singularities: the characterization of strange sets. *Phys. Rev. A*, **33**, 1141–1151.
- Hentschel, H.G.E. and Procaccia, I. (1983) The infinite number of generalized dimensions of fractals and strange attractors. *Physica D*, **8**, 435–444.
- Holley, R. and Waymire, E.C. (1992) Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Ann. Appl. Probab.*, **2**, 819–845.
- Kahane, J.P. and Peyrière, J. (1976) Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.*, **22**, 131–145.
- Lovejoy, S. and Schertzer, D. (1990) Multifractals, universality classes, and satellite and radar measurements of cloud and rain fields. *J. Geophys. Res.*, **95**, 2021–2034.
- Mandelbrot, B.B. (1974) Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**, 331–358.
- Marani, M., Rinaldo, A., Rigon, R. and Rodriguez-Iturbe, I. (1994) Geomorphological width functions and the random cascade. *Geophys. Res. Lett.*, **21**, 2123–2126.
- Meneveau, C. and Chabra, A.B. (1990) Two-point statistics of multifractal measures. *Physica A*, **164**, 564–574.
- Meneveau, C. and Sreenivasan, K.R. (1991) The multifractal nature of turbulent energy dissipation. *J. Fluid Mech.*, **224**, 429–484.
- Novikov, E.A. and Stewart, R.W. (1964) Intermittency of turbulence and the spectrum of fluctuations of energy dissipation. *Izv. Akad. Nauk USSR, Geophys. Ser.*, **3**, 408–413.
- Olsson, J. (1995) Limits and characteristics of the multifractal behavior of a high-resolution rainfall time series. *Nonlinear Processes Geophys.*, **2**, 23–29.
- Over, T.M. and Gupta, V.K. (1994) Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large-scale forcing. *J. Appl. Meteorol.*, **33**, 1526–1542.
- Platt, D.E. and Family, F. (1993) Consistent scaling of multifractal measures: multifractal spatial correlations. *Phys. Rev. E*, **47**, 2281–2288.

Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

She, Z.-S. and Waymire, E.C. (1995) Quantized energy cascade and log-Poisson statistics in fully developed turbulence. *Phys. Rev. Lett.*, **74**, 262–265.

Received December 1996 and revised October 1997