Stochastic Environmental Research and Risk Assessment 15 (2001) 261–283 © Springer-Verlag 2001

Multifractal objective analysis: conditioning and interpolation

G. Salvadori, D. Schertzer, S. Lovejoy

Abstract. We investigate various ways of statistically estimating multifractal fields from sparse data. First, the problem is set in the general framework of conditional expectations, and the effect of (multi) fractal sampling on the statistics of the measured process is investigated, showing how analytical expressions describing the statistical properties of the phenomenon should be modified by the sampling. Then, several techniques are introduced, our goal being to estimate the intensity of a field at resolution λ, given samples of the process collected by networks at higher resolutions $\Lambda > \lambda$. The general strategy underlying all the estimating techniques presented is to approximate the unknown field values at resolution λ by means of most likely estimates conditional to the available information at resolution $\Lambda > \lambda$. Finally, the procedures are tested on simulated lognormal multifractal fields sampled by means of fractal networks, and the propagation of the errors in a scaling framework is also discussed. These techniques are necessary for estimating geophysical processes in regions where no monitoring stations are present, a scenario often encountered in practice, and may also be of great help in studying natural hazards and risk assessment.

G. Salvadori
Dipartimento di Matematica "Ennio De Giorgi"
Università di Lecce, Provinciale Lecce-Arnesano
P.O. Box 193, I-73100 Lecce, Italy
(Affiliation: DIIAR (Sezione Idraulica)
Politecnico di Milano, Milano, Italy)
e-mail: gianfausto.salvadori@unile.it

D. Schertzer (✉)
L.M.M. – Centre Nationale de la Recherche Scientifique
Université Pierre et Marie Curie
Case 162, 4 Place Jussieu, F-75252 Paris
France Cedex 05
e-mail: schertze@ccr.jussieu.fr

S. Lovejoy
McGill University, Physics Department
3600 University Street, Montréal
Canada H3A 2T8
e-mail: lovejoy@physics.mcgill.ca

The research of G. Salvadori was partially supported by Contract n. ENV4-CT97-0529 within the project "FRAMEWORK" of the European Community – D.G. XII.

1
Introduction

The practical analysis of natural processes poses severe difficulties. A major problem that plagues the geoscientists is the typically inhomogenous nature of measuring networks. Indeed, contrary to the traditional view of networks as being essentially homogeneous with occasional holes, many networks have been shown to be sparse over various scales with fractal and multifractal characteristics, i.e. the holes occur in a hierarchical scaling way at all scales. Examples where this has been demonstrated in detail include rain-gauges (Lovejoy et al., 1986; Korvin, 1992; Nicolis, 1993; Tessier et al., 1994) to environmental pollution monitoring (Salvadori et al., 1994, 1997; Belli et al., 1995).

In addition to the relevance of this question to making maps from in situ data, it is also important in risk assessment. Actually, the fractal nature of the networks introduces – in addition to that of spatial resolution – the new notion of dimensional resolution of real networks (Lovejoy et al., 1986), which quantifies the sparsest phenomenon that can be detected by a sparse network. Roughly, "hot spots", even if large, are often missed due to the gaps of the network. Thus, the problem arises of how to estimate the local intensity of a phenomenon in the unsampled locations when analysing a single realization of a process whose ensemble statistics are known.

A common goal of objective analysis techniques is to obtain information on a field sensed by in situ measuring networks. Historically, quite a few methods have been devised to deal with such a problem. To recall a few, Thiessen polygons technique dates back to the beginning of the Century (Thiessen, 1911) and is still used in hydrology. More refined methods, based on polynomial least square fits, were proposed by Panofsky (1949) and improved by Gilchrist and Cressman (1954); however, polynomials had the disadvantage of generating uncontrolled oscillations. Another class of techniques, known as weighted interpolation methods (see Goodin et al. (1979) for a survey), provide estimates of the unknown field in a given location roughly by means of weighted averages of the surrounding available measurements. Another well known and used method, performing statistical linear interpolation, is Kriging (Matheron, 1970). Objective comparisons of mapping techniques can be found in Creutin and Obled (1982); more on conventional spatial geostatistics methods can be found, e.g., in Ripley (1981) and Cressie (1992).

A common shortcoming of these standard techniques is that they generally assume a fairly regular behaviour of the quantities involved. For example, in a recent paper discussing the mathematical requirements to be satisfied by remotely sensed fields (Raffy, 1994), the author rightly criticizes the imposition of differentiability assumptions as too restrictive, but then goes on without comment to assume that remotely sensed fields are regular with respect to Lebesgue measure, even though scaling (multifractal) measures will generally be singular with respect to the latter. Other restrictive assumptions include finite decorrelation lengths, means, variances and high order moments.

The possibility of using fractals for modelling and the possibility of interpolating geometrical fractal sets (including those used to model natural objects such as clouds, mountains, landscapes, ...) has been investigated particularly by Barnsley (1988) – see also, e.g., Mandelbrot (1983), Feder (1989), Kaye (1989), Becker and Dürfler (1990), Crilly et al. (1991), Korvin (1992). For example, through the application of the Iterated Function System (IFS) technique, Barnsley (1988) showed how to generate interpolations of fractal sets "close" (in the

Hausdorff metric sense) to empirical sets having the same fractal dimension over an appropriate range of scales. However, the theoretical framework of geometric sets (monofractal functions supporting the IFS technique) is inadequate for analysing stochastic multifractal fields such as those playing a fundamental rôle in modelling scaling geophysical phenomena. This is because multifractals are resolution-dependent singular measures (Schertzer and Lovejoy, 1987, 1992), and not just resolution-independent sets of points.

In this work we outline several new possible approaches to tackling the problem of modelling and interpolating multifractal fields sensed by in situ measuring networks; the generic multifractal process is the multiplicative cascade. In the following we therefore make use of many of the properties of multiplicative processes and multifractals. We anticipate that our approach is valid even if the actual support of the observed phenomenon is fractal.

2
Universal multifractals

In this section we briefly outline the relevant mathematical background. This short introduction includes only the material strictly necessary to understand the remaining part of the work; more details can be found, e.g., in Schertzer and Lovejoy (1987, 1992). In addition, we discuss the effect of (multi) fractal sampling on the statistics of the measured process, settling the question in terms of conditional expectations, and we show how the analytical expressions describing the statistical properties of the phenomenon should be modified (Salvadori, 1993). Finally, we discuss the particular case of a fractal sampling network.

2.1
Mathematical background

For a stochastic multifractal field ε_λ at resolution $\lambda > 0$ (the ratio between the largest scale considered L and the scale of observation l, i.e. $\lambda = L/l$) the probability distribution function has the following multiscaling expression (Schertzer and Lovejoy, 1987):

$$\mathbf{P}\{\varepsilon_\lambda > \lambda^\gamma\} \approx \lambda^{-c(\gamma)} \tag{2.1}$$

where $\gamma \in \mathbb{R}$ is called order of singularity, and $c(\gamma)$ is called codimension function. Considering the statistical moments of ε_λ of order $q > 0$, an equivalent multiscaling expression can be written:

$$\mathbf{E}(\varepsilon_\lambda^q) \approx \lambda^{K(q)} \tag{2.2}$$

where $K(q)$ represents the scaling function of the moments. The functions $c(\gamma)$ and $K(q)$ form a Legendre transformation pair:

$$\begin{cases} K(q) = \max_\gamma\{q\gamma - c(\gamma)\} \\ c(\gamma) = \max_q\{q\gamma - K(q)\} \end{cases} \tag{2.3}$$

which establishes a one-to-one correspondence between moments and singularities:

$$\begin{cases} q = c'(\gamma) \\ \gamma = K'(q) \end{cases} \tag{2.4}$$

The statistical description of multifractal processes may be greatly simplified considering (conservative) universal multifractals. In this case the functions $K(q)$ and $c(\gamma)$ are given by:

$$K(q) = \begin{cases} \frac{C_1\alpha'}{\alpha}(q^\alpha - q), & \alpha \neq 1 \\ C_1\ q\ \ln(q), & \alpha = 1 \end{cases} \tag{2.5a}$$

and (via a Legendre transform):

$$c(\gamma) = \begin{cases} C_1(\frac{\gamma}{C_1\alpha'} + \frac{1}{\alpha})^{\alpha'}, & \alpha \neq 1 \\ C_1 e^{(\gamma/C_1)-1}, & \alpha = 1 \end{cases} \tag{2.5b}$$

where $1/\alpha + 1/\alpha' = 1$, $0 < \alpha \leq 2$ and $q > 0$ for $\alpha \neq 2$. The parameter α is called Lévy index (or degree of multifractality) and determines the probability class of the process. The parameter $C_1 > 0$ is the codimension of the average field and measures its sparseness and inhomogeneity. Universal multifractals are natural models for scaling geophysical processes since they are stable and attractive, and insensitive to the exact (non-linear) scale invariant dynamical mechanisms at work.

As a last point of interest we note that, independently of the specific value of α, there exists a characteristic order of singularity γ_0 given by:

$$\gamma_0 = -C_1\frac{\alpha'}{\alpha} \tag{2.6}$$

In fact, γ_0 is either (for $1 < \alpha < 2$) the lower bound of non space-filling singularities (for $\gamma \leq \gamma_0$, $c(\gamma) = 0$, i.e. space-filling singularities) or (for $0 < \alpha < 1$) the upper bound of singularities ($c(\gamma) = \infty$ for $\gamma \geq \gamma_0$, i.e. unreachable singularities).

2.2
Conditional expectations and estimates

The estimating techniques presented later assume knowledge of the parameters α and C_1 that identify the statistics of the multifractal process investigated. If these are unknown, then they need to be estimated on the basis of the available data (see, e.g., Lavallée, 1991; Lavallée et al., 1991). However, geophysical monitoring networks often show a (multi) fractal structure, which may affect the estimation of such parameters (see, e.g., Tessier et al., 1994; Salvadori et al., 1994; Belli et al., 1995).

A very general way to understand the effects of (multi) fractal sampling is to put the question in terms of the estimate of a (multifractal) field ε given a (multifractal) field ρ. This leads naturally to consider the conditional expectation of ε with respect to ρ (see, e.g., Shiryaev, 1996).

It is essential here to put on a firm theoretical ground the bases of conditional estimate in a multifractal framework, deriving general relations between the statistics of the relevant processes. As we shall see, despite the complexity, straightforward relations will be derived for both the codimension functions and the scaling functions of the moments of the interacting processes. Thus, multifractal estimates are possible in a even more general framework than that of discrete processes outlined later.

First of all, let us recall that random fields having equivalent σ-algebras are almost surely functionally related, and vice versa. In other words, the set $\mathscr{E}_\rho$ of the random fields equivalent to ρ can be defined as:

$$\mathscr{E}_\rho = \{\varrho : \exists\, f,\ \sigma(\rho)\text{-measurable}, \quad \varrho = f(\rho)\ \text{a.s.}\} \tag{2.7}$$

This has a major consequence for multiplicative processes: in fact, random fields and their generators (Schertzer and Lovejoy, 1987) have equivalent σ-algebras, as it suffices to take f as the exponential function in Eq. (2.7).

The rather general definition of conditional expectation $\mathbf{E}(\varepsilon \mid \rho)$ of ε with respect to ρ corresponds to an orthogonal projection of ε on $\mathscr{E}_\rho$. More formally:

$$\mathbf{E}((\mathbf{E}(\varepsilon \mid \rho) - \varepsilon)\varrho) = 0 \tag{2.8}$$

for all $\varrho \in \mathscr{E}_\rho$. This shows that the "error" done approximating ε by means of its conditional expectation $\mathbf{E}(\varepsilon \mid \rho) \in \mathscr{E}_\rho$ is orthogonal to $\mathscr{E}_\rho$. When the processes ε and ρ are independent, Eq. (2.8) reduces to $\mathbf{E}(\varepsilon \mid \rho) = \mathbf{E}(\varepsilon)$.

From the notion of conditional expectation $\mathbf{E}(\varepsilon \mid \rho)$ it is possible to derive (implicitly) that of conditional probability $\mathbf{P}\{\varepsilon \mid \rho\}$. A standard result states that:

$$\mathbf{P}\{\varepsilon \mid \rho\}\mathbf{P}\{\rho\} = \mathbf{P}\{\varepsilon, \rho\} \tag{2.9}$$

and if the processes ε and ρ are independent, Eq. (2.9) reduces to $\mathbf{P}\{\varepsilon \mid \rho\} = \mathbf{P}\{\varepsilon\}$.

Since for multiplicative processes the probabilities scale as $\lambda^{-c(\cdot)}$, it is possible to write (using shorthand notation):

$$\lambda^{-c_{\varepsilon|\rho}}\lambda^{-c_\rho} = \lambda^{-c_{\varepsilon,\rho}} \tag{2.10}$$

where $c_{\varepsilon|\rho}$ is the conditional codimension function of ε given ρ, and $c_{\varepsilon,\rho}$ is the joint codimension function of ε and ρ. Then the following result for the codimension functions holds:

$$c_{\varepsilon|\rho} + c_\rho = c_{\varepsilon,\rho} \tag{2.11}$$

For independent processes, $c_{\varepsilon|\rho} = c_\varepsilon$ and hence (Intersection Theorem):

$$c_{\varepsilon,\rho} = c_\varepsilon + c_\rho \tag{2.12}$$

This states that the codimension functions of interacting independent fields simply add – for instance, this could be the case of a multifractal phenomenon ε (e.g., rainfall or pollution) sampled by means of an independent multifractal network ρ.

Exploiting the above results, via a Legendre transform it is possible to derive a corresponding formula for the scaling function of the moments:

$$K_{\varepsilon|\rho}(q) = \max_{\gamma_\varepsilon}\{q\gamma_\varepsilon - c_{\varepsilon,\rho}(\gamma_\varepsilon, \gamma_\rho) + c_\rho(\gamma_\rho)\} \tag{2.13}$$

where $K_{\varepsilon|\rho}$ is the conditional scaling function of the moments of ε given ρ. Then, as before, for independent processes $K_{\varepsilon|\rho} = K_\varepsilon$, and by the same token:

$$K_{\varepsilon,\rho} = K_\varepsilon + K_\rho \tag{2.14}$$

showing how the scaling functions of the moments of interacting independent fields simply add.

Let us investigate in more detail the case of a fractal monitoring network. Assuming independence between the phenomenon ε and the network ρ sampling it, the moment scaling function $K_\mu(q)$ of the process μ actually measured

(conditionally to the presence of a sparse independent network) and that of the "true" unknown process ε, $K_\varepsilon(q)$, are related by (See Eq. (2.14)):

$$K_\mu(q) = K_\varepsilon(q) + K_\rho(q) \tag{2.15}$$

where $K_\rho(q)$ is the moment scaling function of the network density. Therefore, proper corrections should be applied when estimating the multifractal parameters (see, e.g., Tessier et al., 1994; Salvadori et al., 1994). In case of a fractal network (i.e., taking the limit $\alpha \to 0^+$ in Eq. (2.5a)):

$$K_\rho(q) = C_{1,\rho}(q-1) \tag{2.16}$$

where $C_{1,\rho}$ is the codimension of the network. Note that introducing a simple scaling network is equivalent to multiplication by a monofractal field. The codimension $C_{1,\rho}$ of the network is easy to estimate: for instance, considering data distributed on a regular mesh (e.g., analysing satellite pictures) the box-counting algorithm may be used (see, e.g., Mandelbrot, 1983; Falconer, 1988, 1990; Feder, 1989). Then, the empirical data may be properly filtered before being used.

3
Multifractal estimate

In this section we present some techniques to provide estimates of a multifractal field both at resolution λ and at resolution $\Lambda > \lambda$, given some samples at resolution Λ. The underlying idea is to first calculate the probability structure of the field of interest, and then to try and approximate the field itself by means of the most likely value of the process. For the sake of clarity, in Table 1 we summarize the notation used throughout the paper; see also Fig. 1 for a graphical illustration.

The basic assumption is that the parent field $\varepsilon_\lambda = \lambda^\gamma$ at resolution λ is generated by means of a sequence of k cascade steps from resolution 1 to λ (each single step having a scale ratio $\tilde{\lambda}$, for a total scale ratio $\lambda = \tilde{\lambda}^k$). Then, the offspring field $\varepsilon_\Lambda = \Lambda^\Gamma$ at resolution Λ is obtained by means of a single cascade step (scale ratio $\tilde{\lambda}$) from resolution λ to Λ. In Figs. 1 and 2 we illustrate the cascade algorithm and the purposes of the estimating techniques introduced below.

Given the stated assumptions, we may write the following relations:

$$\lambda^\gamma = \prod_{i=1}^{k} \tilde{\lambda}^{\tilde{\gamma}_i} = \tilde{\lambda}^{\sum_{i=1}^{k} \tilde{\gamma}_i} \tag{3.1a}$$

$$\Lambda^\Gamma = \prod_{i=1}^{k+1} \tilde{\lambda}^{\tilde{\gamma}_i} = \tilde{\lambda}^{\sum_{i=1}^{k+1} \tilde{\gamma}_i} = \lambda^\gamma \tilde{\lambda}^{\tilde{\gamma}_{k+1}} \tag{3.1b}$$

Table 1. Illustration of the notation used throughout the paper

·	Explanation
λ	Resolution of the parent field (scale $1/\lambda$)
γ	Parent order of singularity
Λ	Resolution of the offspring field (scale $1/\Lambda$)
Γ	Offspring order of singularity
$\tilde{\lambda}$	Scale ratio ($\tilde{\lambda} = \Lambda/\lambda$)
$\tilde{\gamma}$	Order of singularity of multiplicative factors
k	Number of cascade steps ($k = \ln(\lambda)/\ln(\tilde{\lambda})$)

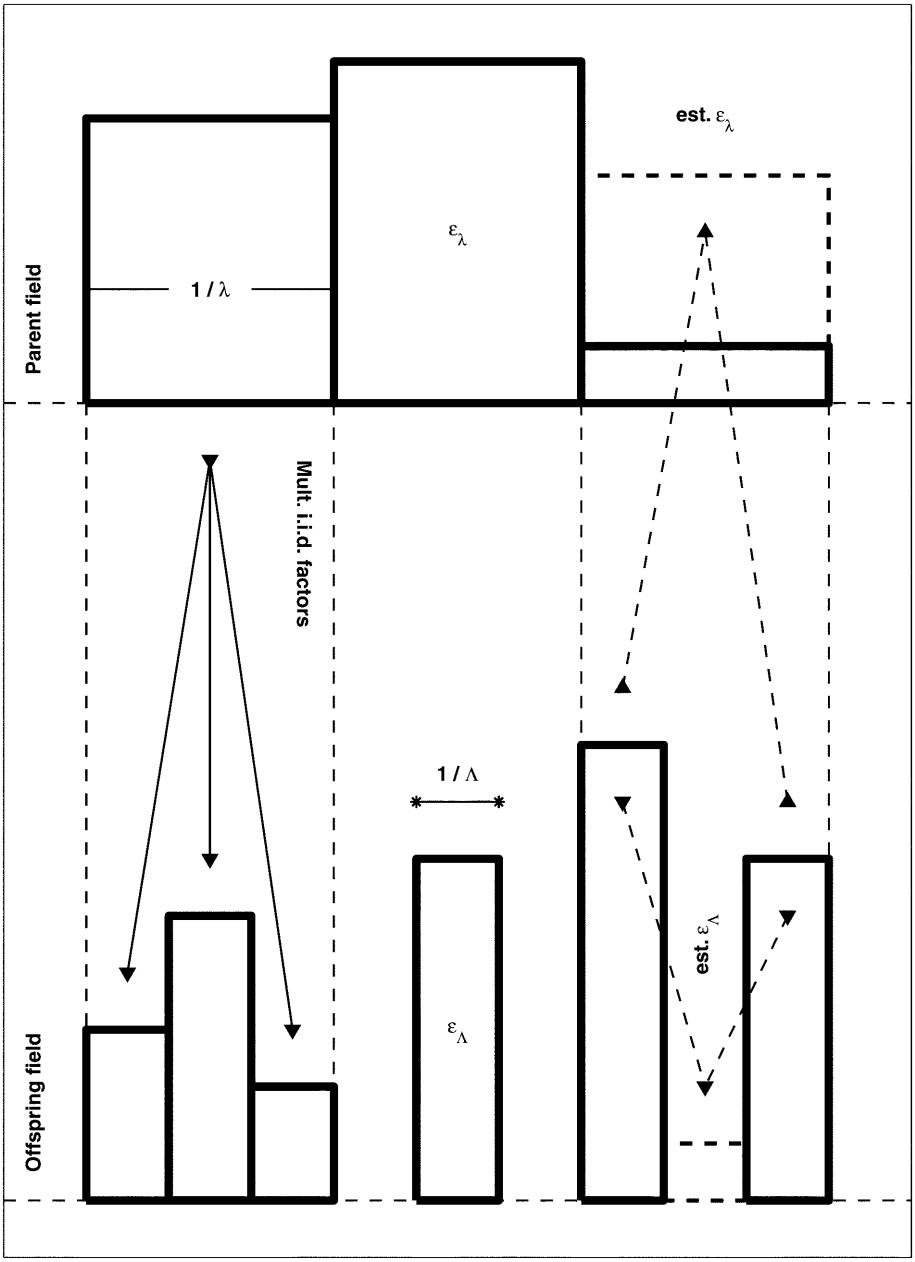

Fig. 1. Graphical illustration of the purposes of the estimating techniques explained in the text. First, it is supposed that an offspring field ε_Λ at resolution Λ (scale $1/\Lambda$) is generated starting from a parent field ε_λ at resolution λ (scale $1/\lambda$) by a single step of a multiplicative cascade mechanism (here with scale ratio $\tilde{\lambda} = 3$): the downwards continuous arrows represent the i.i.d. factors $\tilde{\lambda}^{\vec{\gamma}}$. Then, a portion of the offspring field is missed in some locations due to the sampling by a network with gaps, here indicated by empty boxes. Finally, estimates of the parent field or of the offspring field are provided (thick dashed lines), using only the available offspring data (upwards dashed arrows for the parent estimate, downwards dashed arrows for the offspring one)

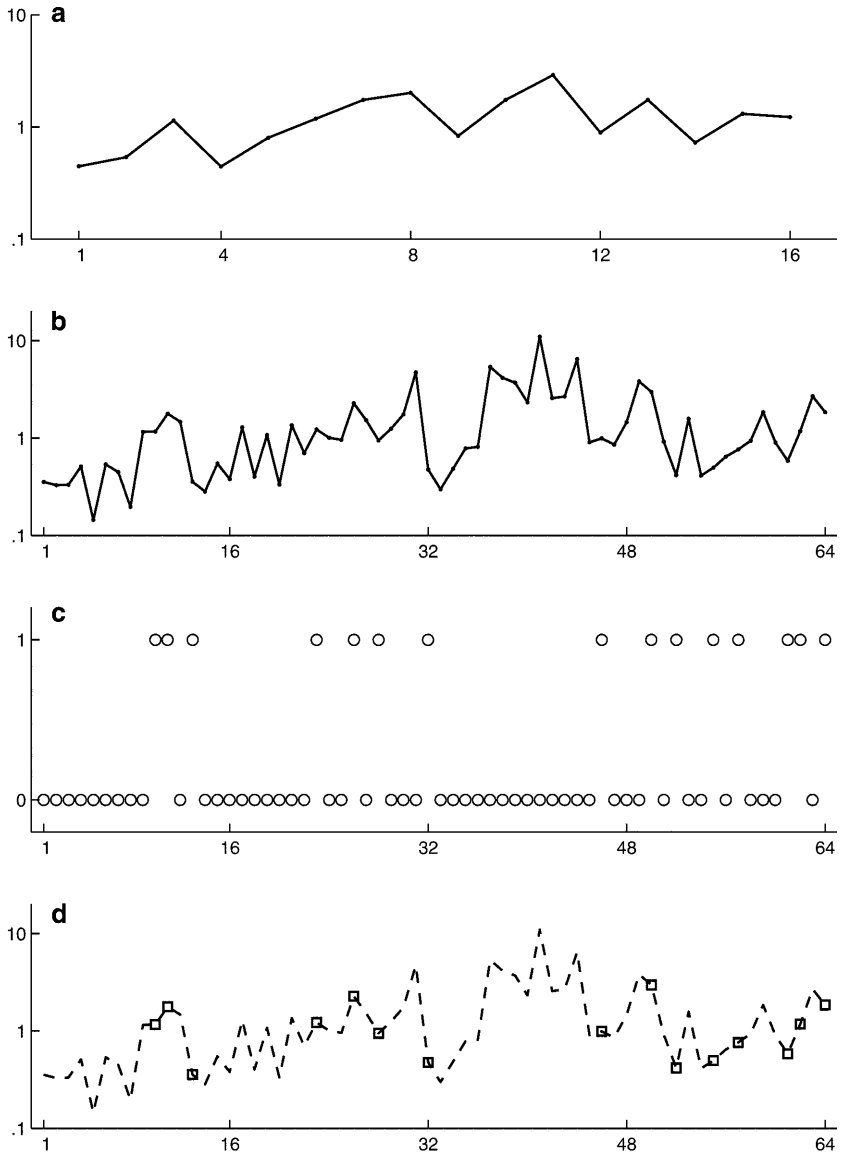

Fig. 2a–d. Graphical illustration of the simulation procedure. **a** First a parent field $\lambda^{\gamma(i)}$, $i = 1, \ldots, \lambda$, at resolution $\lambda = 16$ is simulated ("forward" direction). **b** Then, for each parent value $\lambda^{\gamma(i)}$, $\tilde{\lambda} = 4$ i.i.d. multiplicative factors $\tilde{\lambda}^{\tilde{\gamma}(i,j)}$, $j = 1, \ldots, \tilde{\lambda}$, are generated, and an offspring field $\Lambda^{\Gamma(m)} = \lambda^{\gamma(i)}\tilde{\lambda}^{\tilde{\gamma}(i,j)}$, $m = 1, \ldots, \Lambda$, at resolution $\Lambda = \lambda\tilde{\lambda} = 64$ is obtained ("forward" direction). **c** Finally, a network at resolution Λ covering only 25% of the offspring field is simulated, being equal to one when the local value $\Lambda^{\Gamma(m)}$ can be used by any estimating technique, zero otherwise. **d** The resulting spotty field (markers) provides only partial information about the original field (dashed line): for each parent box, a random number of offspring values (varying from zero to $\tilde{\lambda}$) is available to run some procedure for estimating either the parent field ("backward" direction) or the offspring field itself

where the random variables $\{\tilde{\lambda}^{\tilde{\gamma}_i}\}$ – equivalently, via the invertible transformation $\tilde{\gamma} \leftrightarrow \tilde{\lambda}^{\tilde{\gamma}}$, the random singularities $\{\tilde{\gamma}_i\}$ – are i.i.d., and have (within non-exponential prefactors) a probability density given by $\tilde{\lambda}^{-c(\tilde{\gamma})}$. In turn, since $\lambda = \tilde{\lambda}^k$, we may deduce:

$$\gamma = \frac{1}{k}\sum_{i=1}^{k} \tilde{\gamma}_i \tag{3.2a}$$

$$(1+k)\Gamma = k\gamma + \tilde{\gamma}_{k+1} \tag{3.2b}$$

From Eq. (3.2) we see that the parent singularity γ at step k is simply the average of the singularities $\{\tilde{\gamma}_i\}$ corresponding to the i.i.d. multiplicative factors. Also:

$$\tilde{\gamma} = (1+k)\Gamma - k\gamma \tag{3.3}$$

which expresses the order of singularity $\tilde{\gamma}$ of the last multiplicative factor as a linear function of both the corresponding parent order of singularity γ and the offspring order of singularity Γ.

3.1
Estimating at the parent resolution

The first technique presented is called Most Likely Parent (MLP). The goal of this method is to estimate the intensity λ^γ of a field ε_λ at resolution λ given some samples $\{\varepsilon_{\Lambda,n} = \Lambda^{\Gamma_n}\}$, $n = 1, \ldots, N$, at resolution Λ. For a schematic illustration, see Fig. 1.

Suppose that N offspring values are measured (the non-empty cells in Fig. 1). Let us denote by $\underline{\tilde{\gamma}}$ the vector $\underline{\tilde{\gamma}} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_N)$ of singularities corresponding to the N multiplicative factors, and by $\underline{\Gamma}$ the vector $\underline{\Gamma} = (\Gamma_1, \ldots, \Gamma_N)$ of the offspring singularities actually observed, all sharing the same parent singularity γ (see Fig. 1). Since the multiplicative factors $\{\tilde{\gamma}_n\}$ are i.i.d. and independent of the (common) parent singularity γ, the joint density $\mathbf{p}(\gamma, \underline{\tilde{\gamma}})$ of $\{\gamma, \underline{\tilde{\gamma}}\}$ is easily computed. Considering the (linear) invertible N-dimensional transformation $\underline{\tilde{\gamma}} \leftrightarrow \underline{\Gamma}$, i.e. the N-dimensional version of Eq. (3.3), the joint density $\mathbf{p}(\gamma, \underline{\Gamma})$ of $\{\gamma, \underline{\Gamma}\}$ is:

$$\mathbf{p}(\gamma, \underline{\Gamma}) = |J|\mathbf{p}(\gamma, \underline{\tilde{\gamma}}(\gamma, \underline{\Gamma})) = (1+k)^N \mathbf{p}(\gamma, \underline{\tilde{\gamma}}(\gamma, \underline{\Gamma})) \tag{3.4}$$

where J is the Jacobian of the N-dimensional transformation matrix.

Now, the conditional density $\mathbf{p}(\gamma \mid \underline{\Gamma})$ of the parent singularity γ given $\underline{\Gamma}$ can be written as:

$$\mathbf{p}(\gamma \mid \underline{\Gamma}) = \frac{\mathbf{p}(\gamma, \underline{\Gamma})}{\mathbf{p}(\underline{\Gamma})} \tag{3.5}$$

where the marginal density $\mathbf{p}(\underline{\Gamma})$ of $\underline{\Gamma}$ does not depend upon γ. Mimicking the Maximum Likelihood strategy, we may think of estimating the (common) parent singularity γ by means of that value $\hat{\gamma}$ which maximizes the probability $\mathbf{p}(\gamma, \underline{\Gamma})$ – and hence also $\mathbf{p}(\gamma \mid \underline{\Gamma})$ – of observing the offspring singularities $\underline{\Gamma}$, viewed as a function of γ. In a few cases, this can be done analytically (e.g., when $\alpha = 2$ – see Sect. 4), otherwise it can be carried out using numerical routines.

The MLP method is rather interesting from a computational point of view: in fact, only a one dimensional maximization (with respect to the single variable γ) is required. Although the approximation provided by the MLP technique might

look somewhat crude, it has the advantage of being easy to compute and it suffices for many practical purposes.

Let us consider the codimension function $c(\gamma, \Gamma_n, k) = c(\Gamma_n + k(\Gamma_n - \gamma))$ of the multiplicative factor $\tilde{\gamma}_n$ rewritten in terms of $\{\gamma, \Gamma_n\}$ by means of Eq. (3.3). Then, from Eq. (3.1) we may deduce:

$$\begin{aligned}\mathbf{p}(\gamma, \underline{\Gamma}) &\propto \lambda^{-c(\gamma)} \tilde{\lambda}^{-\sum_{n=1}^{N} c(\Gamma_n + k(\Gamma_n - \gamma))} \\ &= \lambda^{-c(\gamma) - \frac{1}{k}\sum_{n=1}^{N} c(\Gamma_n + k(\Gamma_n - \gamma))}\end{aligned} \tag{3.6}$$

which shows that $\mathbf{p}(\gamma, \underline{\Gamma})$ can be split into two components, one corresponding to a positive sign of γ, and the other to a negative sign. Therefore, while one component of the exponent is increasing with γ, the other is decreasing. Note that the maximization of the function $\mathbf{p}(\gamma, \underline{\Gamma})$ is equivalent to a minimization of the exponent in Eq. (3.6).

Let us define the characteristic order of singularity $\gamma_{0,n}$, $n = 1, \ldots, N$ – see also Eq. (2.6) – associated with the codimension function $c(\gamma, \Gamma_n, k)$:

$$\gamma_{0,n} = \Gamma_n + \frac{1}{k}(\Gamma_n + C_1 \frac{\alpha'}{\alpha}) \tag{3.7}$$

Using Eq. (2.5b) we may easily derive the following result.

Theorem 1 (Behaviour of $\mathbf{p}(\gamma, \underline{\Gamma})$).
As α ranges in $(0, 2]$, the function $\mathbf{p}(\gamma, \underline{\Gamma})$ behaves as follows.

- $\mathbf{0 < \alpha < 1}$**:** over the interval

  $I = (\max_n\{\gamma_{0,n}\}, \gamma_0)$

  the function $\mathbf{p}(\gamma, \underline{\Gamma})$ is positive and well behaved; hence, provided that $I \neq \emptyset$, a maximum exists in I;
- $\boldsymbol{\alpha = 1}$**:** the function $\mathbf{p}(\gamma, \underline{\Gamma})$ is defined over the entire γ-axis, where it is positive and well behaved, and it vanishes for $\gamma \to \pm\infty$; hence a maximum exists;
- $\mathbf{1 < \alpha \leq 2}$**:** over the interval

  $I = (\gamma_0, \min_n\{\gamma_{0,n}\})$

  the function $\mathbf{p}(\gamma, \underline{\Gamma})$ is positive and well behaved; hence, provided that $I \neq \emptyset$, a maximum exists in I. □

Then, the quantity $\hat{\varepsilon}_\lambda = \lambda^{\hat{\gamma}}$ will represent the MLP estimate of the field ε_λ at resolution λ given the singularities $\underline{\Gamma}$ at resolution Λ.

3.2
Estimating at the offspring resolution

The same reasoning underlying the previous technique, designed to work at the parent resolution λ, can be exploited to provide estimates of the field at the offspring resolution Λ. Clearly, this might be of interest in some specific situation, e.g. for the local guess of the intensity of a natural process. For a schematic illustration, see Fig. 1.

Let $\overline{\Gamma}$ indicate the offspring singularity corresponding to some unknown field value $\overline{\varepsilon}_\Lambda$ at resolution Λ. Recalling Eq. (3.4), and using the notation already introduced, we may write:

$$\mathbf{p}(\gamma, \overline{\Gamma}) \propto \mathbf{p}(\gamma, \tilde{\gamma}(\gamma, \overline{\Gamma})) \tag{3.8}$$

Now, substituting the MLP estimate $\hat{\gamma} = \hat{\gamma}(\underline{\Gamma})$ for γ in Eq. (3.8), we may provide an estimate of $\overline{\Gamma}$ looking for that value $\widehat{\Gamma}$ maximizing $\mathbf{p}(\Gamma, \overline{\Gamma})$, which is an implicit function of the offspring singularities $\underline{\Gamma}$ actually observed. In other words, we look for the offspring singularity $\widehat{\Gamma}$ maximizing the probability of observing $\underline{\Gamma}$.

The above technique is called Most Likely Offspring-Parent (MLOP). The approximation provided by the MLOP algorithm is based on the MLP value $\hat{\gamma}$, which is not too difficult to be computed. However, the "error" committed calculating $\hat{\gamma}$ may be amplified, and the eventual estimate of $\widehat{\Gamma}$ may turn out to be somewhat spoiled.

Following a more rigorous approach, we may consider the conditional probability $\mathbf{p}(\overline{\Gamma} \mid \underline{\Gamma})$ of $\overline{\Gamma}$ given $\underline{\Gamma}$. In principle, its calculation requires the knowledge of the joint density $\mathbf{p}(\overline{\Gamma}, \underline{\Gamma})$ of $\{\overline{\Gamma}, \underline{\Gamma}\}$, which is not known, $\overline{\Gamma}$ being unknown. However, recalling Eq. (3.4), we may first consider the joint density $\mathbf{p}(\overline{\Gamma}, \gamma, \underline{\Gamma})$ of $\{\overline{\Gamma}, \gamma, \underline{\Gamma}\}$, and then derive the desired function as a marginal probability by "integrating out" γ. In fact, introducing and then eliminating the variable γ, we may eventually obtain the following expression:

$$\mathbf{p}(\overline{\Gamma} \mid \underline{\Gamma}) = \int \mathbf{p}(\overline{\Gamma}, \gamma \mid \underline{\Gamma}) \mathrm{d}\gamma = \frac{1}{\mathbf{p}(\underline{\Gamma})} \int \mathbf{p}(\overline{\Gamma}, \gamma, \underline{\Gamma}) \mathrm{d}\gamma \tag{3.9}$$

where $\mathbf{p}(\overline{\Gamma}, \gamma, \underline{\Gamma})$ is the joint density of $\{\overline{\Gamma}, \gamma, \underline{\Gamma}\}$, which is a function of the unknown $\overline{\Gamma}$. Thus, we may think of estimating $\overline{\Gamma}$ by means of the most likely value $\widehat{\Gamma}$ maximizing $\mathbf{p}(\overline{\Gamma} \mid \underline{\Gamma})$, i.e. just the integral in Eq. (3.9); clearly, the marginal density $\mathbf{p}(\underline{\Gamma})$ of $\underline{\Gamma}$ can be factored out. As before, the domain of integration is the interval I discussed in Theorem 1. This technique is called Most Likely Offspring (MLO); it generally requires the repeated evaluation of the integral in Eq. (3.9), which may represent a heavy numerical task.

Overall, considering either the MLOP or the MLO technique, the quantity $\hat{\varepsilon}_\Lambda = \Lambda^{\widehat{\Gamma}}$ will represent an estimate of the field ε_Λ at resolution Λ given the singularities $\underline{\Gamma}$ at resolution Λ.

4
Results and discussion

In this section we test some of the estimating techniques outlined in Sect. 3 on simulated fields, and we discuss their usefulness in practical applications. More particularly, we illustrate both multifractal interpolation over a single cascade step and over multiple cascade steps, and we also discuss the propagation of the errors in a scaling framework.

Although the estimating procedures proposed are valid for any admissible value of $\alpha \in (0, 2]$, here we only use simulated lognormal fields (i.e., $\alpha = 2$), which are easy to generate and whose statistical features are simple to handle from a theoretical point of view: in fact, in this particular case, we may write explicitly the analytical expressions of both the MLP and the MLOP estimators (see Theorem 2), thus limiting the use of numerical approximations. We believe this may suffice for the illustrative purposes of the paper. The offspring fields are then sampled by means of fractal networks of known codimension, so that only a small percentage of the fields is used as input source for the interpolation: this imitates the sampling procedures involving real networks. Finally, the estimating techniques are run, and the results compared to the simulated fields. In the following

we shall use the terms "forward" to indicate the direction in which the simulation (i.e., the multiplicative cascade) develops – from small to large resolutions, and "backward" for the direction in which the estimating procedures work – from small to large scales.

In Fig. 2 we sketch the whole procedure. (*a*) First a parent field $\lambda^{\gamma(i)}, i = 1, \ldots, \lambda$, at resolution $\lambda = 16$, is simulated ("forward" direction). (*b*) Then, for each parent value $\lambda^{\gamma(i)}$, $\tilde{\lambda} = 4$ i.i.d. multiplicative factors $\tilde{\lambda}^{\tilde{\gamma}(i,j)}$, $j = 1, \ldots, \tilde{\lambda}$, are generated, and an offspring field $\Lambda^{\Gamma(m)} = \lambda^{\gamma(i)}\tilde{\lambda}^{\tilde{\gamma}(i,j)}$, $m = 1, \ldots, \Lambda$, at resolution $\Lambda = \lambda\tilde{\lambda} = 64$, is obtained ("forward" direction). (*c*) Then, a network at resolution Λ covering only 25% of the offspring field is simulated, being equal to one when the local value $\Lambda^{\Gamma(m)}$ can be used by any estimating technique, zero otherwise. (*d*) Finally, for each parent box, a random number of offspring values (varying from zero to $\tilde{\lambda}$) is available to run some procedure for estimating either the parent field ("backward" direction) or the offspring field itself.

As already mentioned above, the use of lognormal fields gives us the possibility of providing analytical solutions (instead of numerical approximations) to the estimation problem, since the codimension functions $c(\cdot)$ involved in the probability densities $\mathbf{p}(\cdot)$ of interest are quadratic forms (see Eq. (2.5b)). The proof of the results shown below is elementary, though lengthy and tedious, and will not be shown here: after all, it is just a matter of minimizing the quadratic in the exponent of Eq. (3.6), obtained from a quadratic $c(\gamma)$, and then taking expectations of the resulting expressions as presented in Theorem 2. Overall, we may state the following result; the notation is the same as in Sect. 3.

Theorem 2 (MLP and MLOP estimators).
For a lognormal multifractal field, the following estimators are unbiased and consistent:

- **MLP:** considering the MLP technique (see Sect. 3.1), the value $\hat{\gamma}$ maximizing the function $\mathbf{p}(\cdot \mid \underline{\Gamma})$ is given by:

$$\hat{\gamma} = \frac{(N-1)C_1 + (1+k)\sum_{n=1}^{N} \Gamma_n}{1+Nk}$$

- **MLOP:** considering the MLOP technique (see Sec. 3.2), the value $\widehat{\Gamma}$ maximizing the function $\mathbf{p}(\cdot \mid \hat{\gamma}(\underline{\Gamma}))$ is given by:

$$\hat{\Gamma} = \frac{-C_1 + k\sum_{n=1}^{N} \Gamma_n}{1+Nk} \qquad \square$$

The two estimators are similar: essentially, they are averages of the known singularities, the major difference being an additive constant (i.e., a multiplicative factor within the cascade formalism) accounting for the different resolutions of the parent and the offspring fields. It is worth pointing out that, if $\widehat{\Gamma}$ is used to provide further estimates of the offspring field, the new values will not differ from $\hat{\varepsilon}_\Lambda$.

To check whether or not the estimating procedures are able to reproduce the statistical features of the simulated fields, we use the two-sample Kolmogorov–Smirnov test (KS) – see, e.g., Rohatgi (1976), Kottegoda and Rosso (1997). Accordingly, we calculate the value of the test statistics:

$$D_{\mathrm{KS}} = \max_{x \in \mathrm{R}} \{ |F_{\mathrm{S}}(x) - F_{\mathrm{E}}(x)| \}$$

where F_S and F_E represent the distribution functions of, respectively, the Simulated and the Estimated fields. In turn, we compute the significance level p of the test. These parameters provide a quantitative indication about the possibility that two fields are drawn from the same population or, more precisely, that they have the same distribution function: practically, the closer p is to 1, the higher is the probability that the two fields have the same distribution.

In the examples below, we always compare the MLP estimate $\hat{\varepsilon}_\lambda = \lambda^{\hat{\gamma}}$ to the actual value of the parent field. Also, since for lognormal fields the most likely value corresponds to the median, the MLOP estimate $\hat{\varepsilon}_\Lambda = \Lambda^{\hat{\Gamma}}$ is always compared to the sample median calculated using all the offspring singularities sharing the same parent box (the one corresponding to the MLP estimate $\hat{\varepsilon}_\lambda$ used by the MLOP technique). Hereafter "PF" and "MOF" will indicate, respectively, the Parent Field and the Median of the Offspring Field.

4.1
Interpolating over a single cascade step

As a first application, we use here a lognormal offspring field ($\alpha = 2$, $C_1 = 0.5$) at resolution $\Lambda = 2^{16}$ (not shown) and a fractal network sampling only $\approx 25\%$ of the offspring field; the parent field considered is at resolution $\lambda = 2^8$, corresponding to a scale ratio $\tilde{\lambda} = 256$. Such a large scale ratio may be illustrative of a practical situation when, e.g., starting from a few sparse daily measurements of rainfall, an estimate of the annual precipitation is sought, i.e. at a (temporal) scale much larger than that of the observations. The visual inspection of Fig. 3 shows that the estimates provide a good reproduction of the "true" fields (wherever possible), including the fluctuating behaviour, the trend of the field, and the values of the extrema: this is noteworthy, since both the PF and the MOF range over about six orders of magnitude; in addition, the estimates are based only on the knowledge of about 25% of an offspring field spanning about 13 orders of magnitude (from $\approx 10^{-8}$ to $\approx 10^5$). Therefore, given such an extreme variability and the "spotty" nature of the network, the results should indeed be considered reasonable. Also, according to the KS test, the statistical features of both the PF and the MOF are matched: in fact we find, respectively, $D_{KS} \approx 0.034$ and $D_{KS} \approx 0.042$, with a significance level p always greater that 99%. Thus, even the KS test indicates an overall ability to reproduce the statistics of both the PF and the MOF.

As a further application, we use a lognormal offspring field ($\alpha = 2$, $C_1 = 0.5$) at resolution $\Lambda = 2^{10}$ (not shown) and a fractal network sampling only 25% of the offspring field; the parent field considered is at resolution $\lambda = 2^8$, corresponding to a scale ratio $\tilde{\lambda} = 4$. Such a small scale ratio may be illustrative of a practical situation when, e.g., starting from sparse daily measurements of rainfall, an estimate of the weekly precipitation is sought, i.e. at a (temporal) scale comparable to that of the observations; it should be noted that, if small scale ratios are used, then only a very limited amount of offspring information can be used to run any estimating technique. Once again, the visual inspection of Fig. 4 shows that the estimate of the "true" fields (wherever possible) is empirically acceptable: the fluctuating behaviour is fairly well reproduced, as well as the trend of the field. The estimate of the extreme values is more difficult, especially in the case of the MLOP technique: however, this was expected, since both the PF and the MOF range over about six orders of magnitude, and at most four offspring samples ($\tilde{\lambda} = 4$) are available in each parent box to provide an estimate of the field of interest. Therefore, we may consider the results as valuable. The KS test returns, respectively, $D_{KS} \approx 0.078$ and $p \approx 85\%$ for the MLP technique, and $D_{KS} \approx 0.125$ and $p \approx 32\%$ for the MLOP technique. While in the former case the result looks

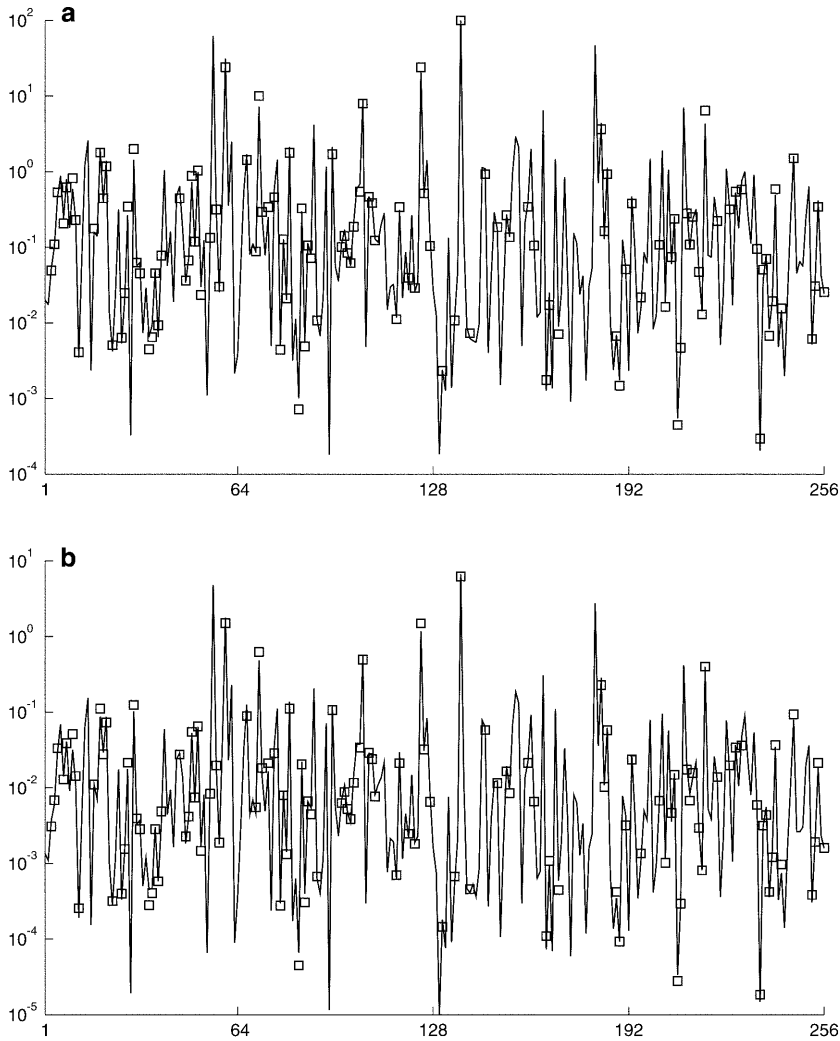

Fig. 3. **a** Example of application of the Most Likely Parent technique. Here PF indicates the Parent Field (line) and MLP the Most Likely Parent estimate (marker). The parent resolution is $\lambda = 2^8$ and the offspring resolution is $\Lambda = 2^{16}$. **b** Example of application of the Most Likely Offspring–Parent technique. Here MOF indicates the sample median of the Offspring Field (line) and MLOP the Most Likely Offspring–Parent estimate (marker). The parent resolution is $\lambda = 2^8$ and the offspring resolution is $\Lambda = 2^{16}$

acceptable, in the latter case it should not be considered as a failure of the algorithm, since the test itself is not quite meaningful given the actual sampling conditions: in fact, the estimator of the offspring sample median is based on only $\tilde{\lambda} = 4$ values, and hence its reliability is questionable (but these are the actual sampling conditions); thus, most of the "uncertainty" is hidden in the estimate of the sample median and not in the MLOP method itself.

4.2
Interpolating over multiple cascade steps

A further point that is worth investigating is the behaviour of the estimating techniques when the scale ratio is as small as possible (i.e., $\tilde{\lambda} = 2$ in the present

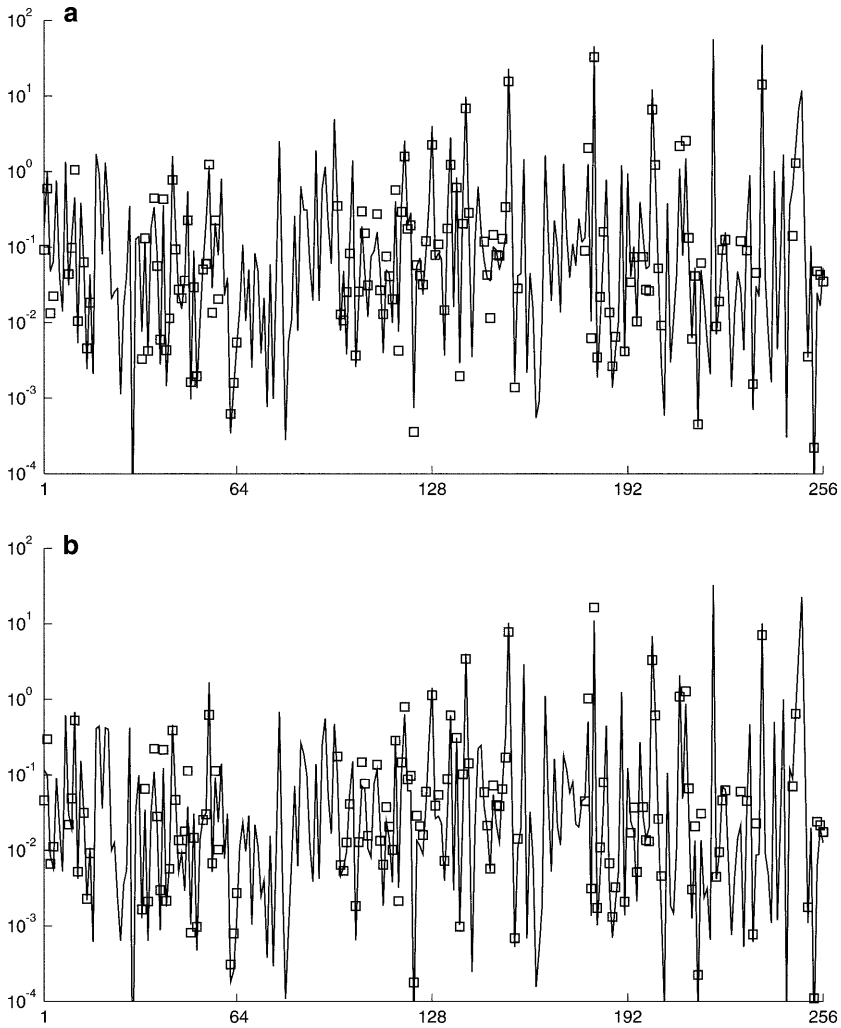

Fig. 4. **a** Example of application of the Most Likely Parent technique. Here PF indicates the Parent Field (line) and MLP the Most Likely Parent estimate (marker). The parent resolution is $\lambda = 2^8$ and the offspring resolution is $\Lambda = 2^{10}$. **b** Example of application of the Most Likely Offspring–Parent technique. Here MOF indicates the sample median of the Offspring Field (line) and MLOP the Most Likely Offspring–Parent estimate (marker). The parent resolution is $\lambda = 2^8$ and the offspring resolution is $\Lambda = 2^{10}$

discrete case) and the algorithms are systematically run backwards to provide estimates at larger and larger scales. Such a test is needed in order to understand the actual performance of the procedures when using continuous cascades (Schertzer and Lovejoy, 1987), which represent the closest multifractal "approximation" to natural phenomena. By the same token, it is also possible to check what happens when large scale estimates (necessarily based on smaller scale estimates) are calculated, a situation which may frequently arise in practical applications. This example may be illustrative of a practical situation when, e.g., starting from sparse daily measurements of rainfall: (*a*) first an estimate of the weekly precipitation $\hat{\varepsilon}_W$ is sought; (*b*) then, using the field $\hat{\varepsilon}_W$ as input data, an estimate of the monthly precipitation $\hat{\varepsilon}_M$ is sought; (*c*) further, using the field $\hat{\varepsilon}_M$

as input data, an estimate of the seasonal precipitation $\hat{\varepsilon}_S$ is sought; (*d*) finally, using the field $\hat{\varepsilon}_S$ as input data, an estimate of the annual precipitation $\hat{\varepsilon}_A$ is sought. It should be noted that such a scheme involves a small scale ratio of the order of about four, and then only a very limited amount of input information can be used at any step to run the backward algorithm; in addition, only at the first step the actual direct observations of the offspring field are used to provide estimates at a larger scale.

As an example, we simulate a lognormal offspring field ($\alpha = 2$, $C_1 = 0.1$) at resolution $\Lambda = 2^{10}$ (not shown) and a fractal network sampling only $\approx 25\%$ of the offspring field. Then, a sequence of backward parent estimates is performed: in Fig. 5 we show a set of six successive backward estimates using a scale ratio $\tilde{\lambda} = 2$, with a parent resolution λ decreasing from 2^9 to 2^4. The results presented are somewhat typical, as observed in many other tests. The visual inspection of Fig. 5 shows that, in general, the backward estimates provide a fairly good reproduction of the "true" fields (wherever possible), including both the fluctuating behaviour and the trend. The agreement between simulated and estimated fields is measured by the significance level p of the KS test. As a general rule, p may fluctuate (especially considering the very first steps of the backward procedure), whereas it seems to converge to one in the successive steps: such a typical behaviour may be explained noting that, on the one hand, at larger and larger scales the fields become less and less extreme, and consequently the magnitude of the errors decreases and the agreement between simulated and estimated fields improves; on the other hand, it should be pointed out that using a scale ratio as small as $\tilde{\lambda} = 2$ means that at most two small-scale values can be used to provide a large-scale estimate (i.e., only a minimum amount of information is available for each estimate). Thus, the results shown in Fig. 5 are interesting.

4.3
Error propagation

A final point that is worth developing is the analysis of the error occurring in the backward estimate procedure. A natural way to define it and to study its statistics is to introduce an error function ϑ as follows:

$$\vartheta_{\lambda,\lambda_e}(q) = \mathbf{E}(|\varepsilon_{\lambda,\lambda_e} - \hat{\varepsilon}_{\lambda,\lambda_e}|^q) \tag{4.1}$$

where $q > 0$ is the order of moment, $\varepsilon_{\lambda,\lambda_e}$ and $\hat{\varepsilon}_{\lambda,\lambda_e}$ are, respectively, the "true" field and the estimated field at resolution λ, and λ_e is the extrapolation scale ratio: for instance, starting with an offspring field at resolution Λ and estimating at the parent resolution λ yields $\lambda_e = \Lambda/\lambda$.

The function $\vartheta_{\lambda,\lambda_e}(q)$ depends upon both λ and λ_e: in a way, we may think of the backward estimated field $\hat{\varepsilon}_{\lambda,\lambda_e}$ at resolution λ, starting with an offspring field at resolution Λ, as being the product of the "true" field ε times an intermediate independent (multiplicative) "cascade of errors" ε_e which develops from resolution λ up to resolution Λ, i.e. spanning from 1 to λ_e in resolution. This factorization is taken into account as follows:

$$\vartheta_{\lambda,\lambda_e}(q) \approx \lambda^{K(q)} \lambda_e^{K_e(q)} \tag{4.2}$$

where $K(q)$ is the standard moment scaling function of ε_λ and $K_e(q)$ is a function tuning the (scaling of the) error. On the one hand, the term $\lambda^{K(q)}$ results from the fact that, introducing more and more cascade steps while keeping λ_e fixed – i.e. taking larger and larger λ and Λ – forces both ε_λ and $\hat{\varepsilon}_\lambda$ to be multiplied by

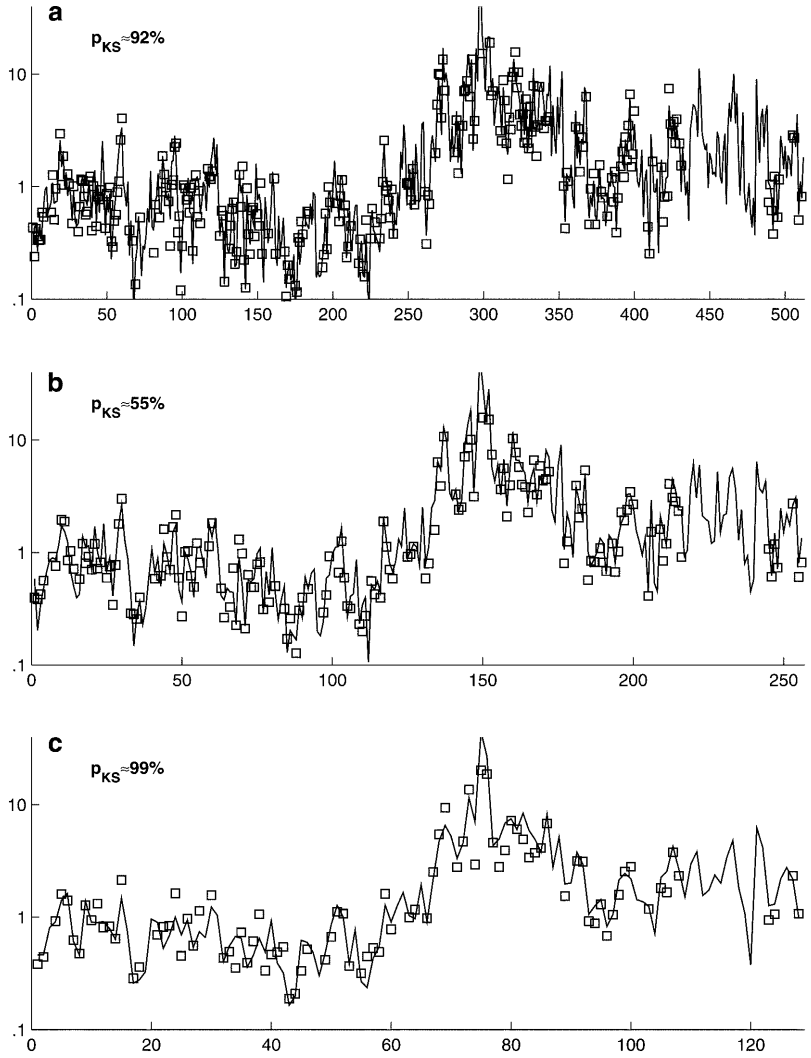

Fig. 5a–f. Example of iteration of the Most Likely Parent technique to calculate backwards estimates. The continuous line indicates the simulated values, and the markers the estimated ones; also reported is the significance level p of the two-sample Kolmogorov–Smirnov test. From **a** to **f**, the parent resolution λ is systematically reduced from 2^9 to 2^4 using a scale ratio $\tilde{\lambda} = 2$. The estimates in (a) are calculated using the actual offspring values at resolution $\Lambda = 2^{10}$ (not shown), whereas all the others are based on parent estimates obtained at the previous step

additional random factors, which can be factored out in Eq. (4.1); on the other hand, the term $\lambda_e^{K_e(q)}$ rules the (scaling) dynamics of the error as a function of the extrapolation scale ratio. Thus, we may define a new error function θ as follows:

$$\theta_{\lambda_e}(q) = \frac{\mathbf{E}(|\varepsilon_{\lambda,\lambda_e} - \hat{\varepsilon}_{\lambda,\lambda_e}|^q)}{\mathbf{E}(\varepsilon_\lambda^q)} = \frac{\vartheta_{\lambda,\lambda_e}(q)}{\mathbf{E}(\varepsilon_\lambda^q)} \approx \lambda_e^{K_e(q)} \tag{4.3}$$

which depends only upon λ_e, and gives the possibility to study $K_e(q)$ directly.

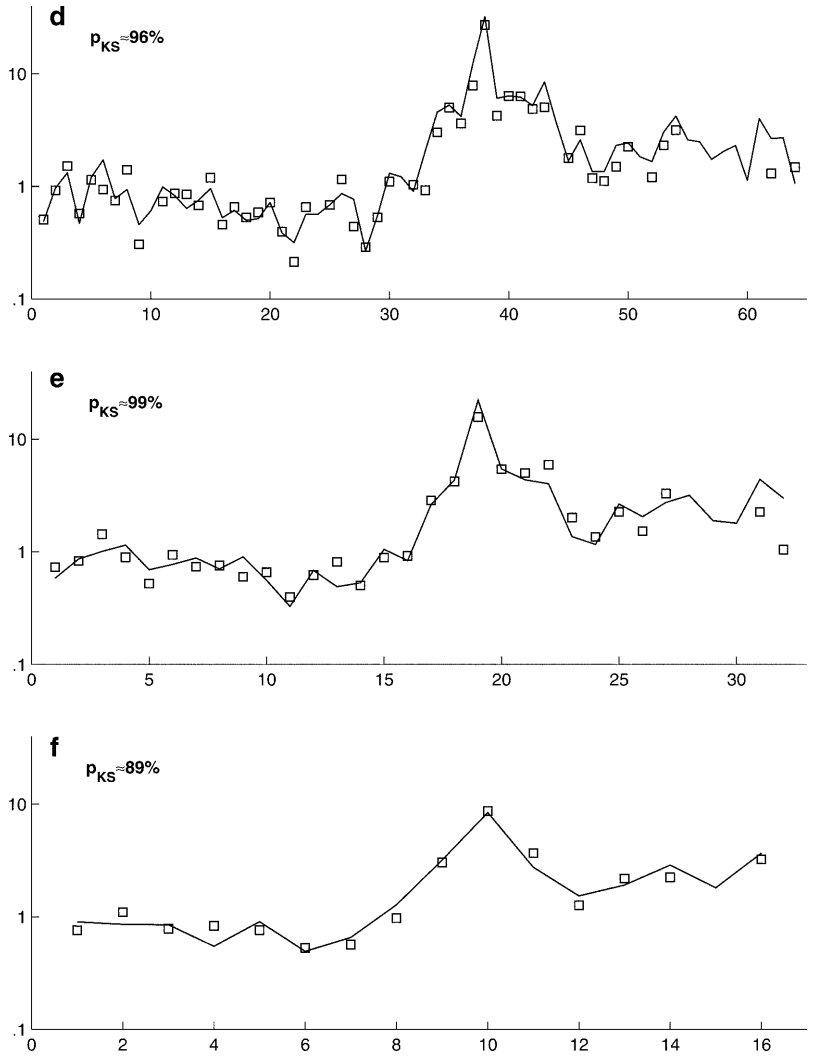

Fig. 5a–f. (Continued)

A further way to get information about the relationship between "true" and estimated fields is to calculate the double-indices generalized correlation coefficients given by:

$$C(q_1, q_2) = \frac{\mathbf{E}(\varepsilon^{q_1}\hat{\varepsilon}^{q_2})}{\mathbf{E}(\varepsilon^{q_1+q_2})} \tag{4.4}$$

where the dependence upon λ and λ_e is momentarily dropped. In our case, we find it easier to work with the single-index correlation function defined as:

$$C(q) = C(q, q) = \frac{\mathbf{E}(\varepsilon^{q}\hat{\varepsilon}^{q})}{\mathbf{E}(\varepsilon^{2q})} \tag{4.5}$$

Since $\hat{\varepsilon}$ is supposed to be given by ε times an independent "cascade of errors" from resolution λ up to resolution Λ, we may infer that:

$$C_{\lambda_e}(q) \approx \lambda_e^{-K_c(q)} \tag{4.6}$$

where the function $K_c(q)$ rules the correlation dynamics and should be non-negative and should equal zero for $q = 0$, so that $C_{\lambda_e}(0) = 1$. Clearly, if the two fields ε and $\hat{\varepsilon}$ are (almost surely) identical then $K_c(q) = 0$. Otherwise, if they are statistically independent, we obtain:

$$C_{\lambda_e}(q) \approx \lambda_e^{K(2q)-2K(q)} \tag{4.7}$$

to within a constant factor, which represents an upper bound. In the lognormal case (i.e. $\alpha = 2$), we may thus establish the theoretical bounds:

$$0 \leq K_c(q) \leq 2C_1 q^2 \tag{4.8}$$

Below we present an empirical analysis of the error. The study is based on 200 independent simulations of an offspring field at resolution $\Lambda = 2^{15}$, with parameters $\alpha = 2$ and $C_1 = 0.4$; in fact, since the scaling is only an "ensemble" property which can be badly broken on any single realization (this is why the operator $\mathbf{E}(\cdot)$ is used), we need to check it on a large data base. The extrapolation scale ratio λ_e ranges from 2 to 1024.

In Fig. 6a we show $C_{\lambda_e}(q)$ vs. λ_e on a log–log scale, for several values of q: a scaling behaviour is fairly well present over all the investigated scales; then, Fig. 6b shows the corresponding estimate of $K_c(q)$. The function $K_c(q)$ satisfies, within the numerical approximation, the requirements previously described: in particular, the theoretical bounds of Eq. (4.8) hold. In addition, for $q < 1$ (corresponding to weak events giving negligible contribution to the mean), the error is roughly maximal. This is presumably a consequence of the fact that the interpolation is mostly sensitive to the larger singularities corresponding to $q > 1$.

The interpretation of the function $\theta_{\lambda_e}(q)$ is more involved and delicate, since it may be given by the superposition of different scaling behaviours. Actually, Fig. 7a shows $\theta_{\lambda_e}(q)$ vs. λ_e on a log–log scale, for several values of q, and a scaling regime can be identified roughly for small λ_e. The corresponding estimate of $K_e(q)$ is plotted in Fig. 7b: as expected, it is positive and increasing, meaning that the error gets larger as the extrapolation scale ratio λ_e grows; apparently the amplification of $\theta_{\lambda_e}(q)$ is non-linear, and therefore the quality of the estimate is also expected to be increasingly worse at larger and larger scales. For the sake of comparison, in Fig. 7b the function $K(q)$ of the simulated field is also plotted: for q small $K_e(q)$ is comparable (in magnitude) to $K(q)$, and then it grows much slower than the latter function. Thus, it is true that the quality of the estimate is increasingly worse, but the rate at which it gets bad is apparently moderate.

5
Conclusions

The present research was motivated by the need to understand and model natural phenomena showing multifractal features. Several facets were investigated: first, the problem has been posed in the general framework of conditional expectations, and the effects of (multi) fractal sampling on the statistics of the measured process were analysed; then, several techniques were introduced to estimate the

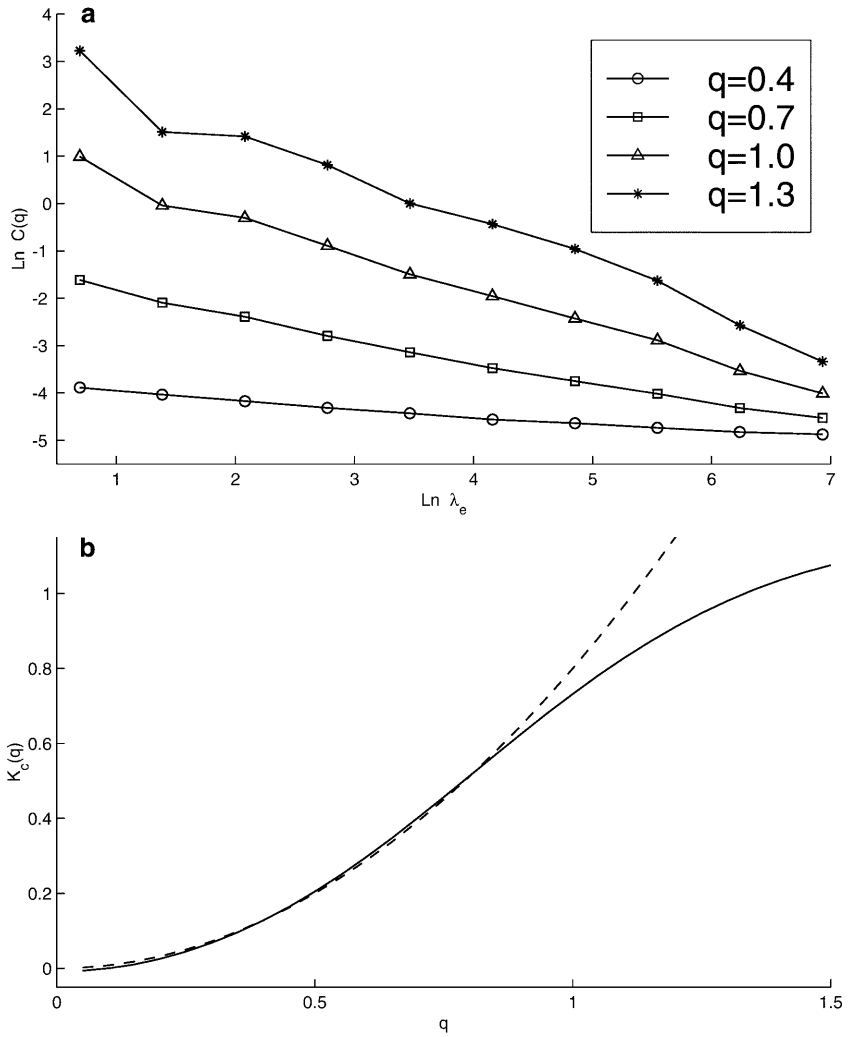

Fig. 6. **a** Plot of the function $C_{\lambda_e}(q)$ vs. λ_e on a log–log scale for several values of q. **b** The corresponding estimate of the function $K_c(q)$ (continuous line); also indicated is the theoretical bound of Eq. (4.8) (dashed line)

intensity of a multifractal process in locations not sampled by a monitoring sparse network, all revolving around the idea of approximating the unknown values by means of statistical most likely estimates of the process.

Numerical tests of the theory look encouraging: the structure of the model is simple and the basic assumptions are not too strict or as demanding as in some conventional techniques. Moreover, the proposed techniques seem to provide an empirically acceptable reproduction of the "true" fields, including their fluctuating behaviour, their trend, and the values of the extrema – even in cases when the simulated fields range over many orders of magnitude. This is noteworthy, since the estimates are usually based on the knowledge of only a small percentage of the input fields. In addition, the KS tests seem to indicate an overall ability of the estimation techniques to reproduce also the statistics of the simulated fields.

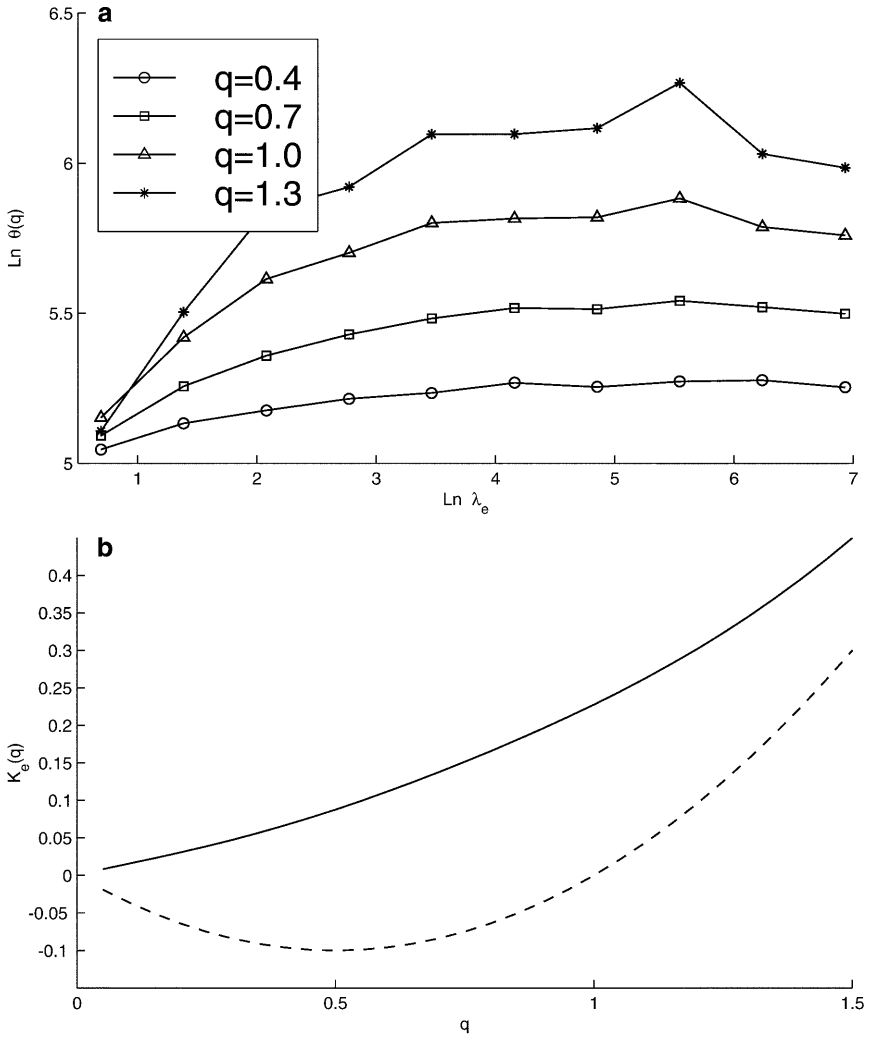

Fig. 7. **a** Plot of the function $\theta_{\lambda_e}(q)$ vs. λ_e on a log–log scale for several values of q. **b** The corresponding estimate of the function $K_e(q)$ (continuous line); also indicated is the function $K(q)$ of the simulated field (dashed line)

Furthermore, the systematic backward iteration of multifractal interpolation provides interesting information on the behaviour of the techniques when large scale estimates (necessarily based on smaller scale estimates) are calculated – a situation which may frequently arise in practical applications. It is worth stressing that such tests are also needed in order to understand the actual performance of the procedures when using continuous (in scale) cascades, the closest multifractal "approximation" to natural phenomena.

Finally, the scaling analysis of the errors shows how the (rate of the) error grows and the (rate at which the) quality of the estimate gets worse when increasing the extrapolation scale ratio, and also the way the correlation between the "true" and the estimated fields fades away.

We are still far from exploiting all the potentialities hidden in the mathematics of multifractals. Several refinements are still needed, both of a theoretical and of a

computational nature, in order to improve the techniques introduced and their performance. For example, stronger tests are needed to evaluate to what extent the algorithms are able to (locally) reproduce the fields of interest. Furthermore, an essential point in practical applications is the computational performance of the algorithms; this might be of interest for, e.g., risk assessment in case of environmental pollution accidents, when estimates at large scales (i.e. at the parent resolution) are needed in real-time, relying only on extremely spotty measurements (see, e.g., Salvadori et al., 1994, 1997; Belli et al., 1995; Schertzer et al., 1995). In our case the whole estimation procedure is not computationally too demanding and time consuming. Indeed, since natural phenomena are often sampled by means of (multi) fractal networks, the above techniques may represent a useful contribution to the ongoing investigations concerning the estimate of multifractal fields, and may have fruitful applications in many areas of geophysical and environmental sciences.

References

Belli G, Cambiaghi M, Chigirinskaya Y, Kwak D, Lanza A, Marsan D, Naud C, Quinto E, Ratti SP, Ravera A, Salvadori G, Scannicchio D, Schertzer D, Schmitt F (1995) Radiation Protection Research Action, Commission of the European Communities (D.G. XII), Contract FI3P-CT93-0077, Final Report, (available upon request to the Commission of the European Communities – D.G. XII)

Barnsley M (1988) Fractals Everywhere. Academic Press, San Diego

Becker KH, Dürfler M (1990) Dynamical Systems and Fractals. Cambridge University Press, Cambridge

Cressie N (1992) Statistics for Spatial Data. J. Wiley & Sons, New York

Creutin JD, Obled C (1982) Objective analyses and mapping techniques for rainfall fields: an objective comparison. Wat. Res. Res., 18(2): 413

Crilly AJ, Earnshaw RA, Jones H (1991) Fractals and Chaos. Springer-Verlag, New York

Falconer KJ (1988) The Geometry of Fractal Sets. Cambridge University Press, Cambridge

Falconer KJ (1990) Fractal Geometry. J. Wiley & Sons, Chichester

Feder J (1989) Fractals. Plenum Press, New York-London

Gilchrist B, Cressman G (1954) An experiment in objective analysis. Tellus, 6: 309

Goodin WR, McRae GJ, Seinfeld JH (1979) A Comparison of Interpolation Methods for Sparse Data. J. Appl. Meteo., 18: 761

Kaye BH (1989) A Random Walk through Fractal Dimensions. VCH, Weinheim, FRG

Korvin G (1992) Fractal Models in the Earth Sciences. Elsevier, New York

Kottegoda NT, Rosso R (1997) Probability, Statistics, and Reliability for Civil and Environmental Engineers. McGraw-Hill, New York

Lavallée D (1991) Multifractal Analysis and Simulation Techniques and Turbulent Fields. PhD Thesis, McGill University, Physics Dept., Montréal, Canada

Lavallée D, Lovejoy S, Schertzer D (1991) On the determination of the codimension function. In: Schertzer D, Lovejoy S (eds.) Non-Linear Variability in Geophysics. Kluwer Academic Publishers, p. 99

Lovejoy S, Schertzer D, Ladoy P (1986) Fractal characterization of inhomogeneous measuring networks. Nature, 319: 43

Matheron G (1970) Random functions and their application in geology. In: Merriam DF (ed.) Geostatistics. Plenum, New York, p. 79.

Mandelbrot BB (1983) The Fractal Geometry of Nature. Freeman, S. Francisco

Nicolis C (1993) Optimizing the global observational network: a dynamical approach. J. Appl. Meteo., 33: 1751

Panofsky H (1949) Objective weather-map analysis. J. Appl. Meteo., 6: 386

Raffy M (1994) The role of spatial resolution in quantification problems: spatialization method. Int. J. Remote Sensing, 15: 2381

Ripley BD (1981) Spatial Statistics. J. Wiley & Sons, New York

Rohatgi VK (1976) An Introduction to Probability and Mathematical Statistics. J. Wiley & Sons, New York

Salvadori G (1993) Multifrattali Stocastici: Teoria e Applicazioni. PhD Thesis, Diss. 93/1316, Biblioteca Nazionale, Roma (Italy); TDR.1993.1278, Biblioteca Nazionale, Firenze (Italy); (in Italian)
Salvadori G, Ratti SP, Belli G, Lovejoy S, Schertzer D (1994) Multifractal objective analysis of seveso ground pollution. Tox. Envir. Chem., 43: 63
Salvadori G, Ratti SP, Belli G (1997) Fractal and multifractal approach to environmental pollution. Environ. Sci. Poll. Res., 4(2): 91
Schertzer D, Lovejoy S (1987) Physical modelling and analysis of rain and clouds by anisotropic scaling multiplicative processes. J. Geophys. Res., 92: 9693
Schertzer D, Lovejoy S (1992) Hard and soft multifractal processes. Physica A, 185: 187
Schertzer D, Chigirinskaya Y, Lovejoy S, Ratti SP, Salvadori G, Belli G (1995) Multifractal analysis of chernobyl fallout: self-organized criticality and hot spots. Proc. "Mathematics and Computations, Reactor Physics, and Environmental Analyses", Portland (Oregon), April 30–May 4, 1995, Amer. Nucl. Soc. Publisher, Illinois, p. 743
Shiryaev AN (1996) Probability, 2nd edn., Springer-Verlag, New York
Tessier Y, Lovejoy S, Schertzer D (1994) The multifractal global rain-gauge network: analysis and simulation. J. Appl. Meteo., 33: 1572
Thiessen AH (1911) Precipitation averages for large areas. Mon. Wea. Rev., 39: 1082