On the determination of the codimension function

by D. Lavallée, D. Schertzer, S. Lovejoy

In « *Nonlinear Variability in Geophysics* »,

Eds. D. Schertzer, S. Lovejoy, Kluwer, 1991, p. 99=109.

Note, we have restored the appendix which the publisher omitted by error.

ON THE DETERMINATION OF THE CODIMENSION FUNCTION

D. Lavallée, D. Schertzer* and S. Lovejoy
Department of Physics,
McGill University,
3600 University st., Montréal, Québec
Canada, H3A 2T8

ABSTRACT. Motivated by the necessity of developing new multifractal analysis techniques to characterize empirical fields by scale invariant (sensor resolution independent) codimension functions, we introduce a new method, PDMS, to directly estimate the codimension of the singularity spectrum $c(\gamma)$ and we also indicate the theoretical (or practical) limits of this method as well as its consequences for the determination of highest values of $c(\gamma)$. These properties also have implications for the behaviour of $K(h)$ – related to $c(\gamma)$ by a Legendre transformation – in particular for large h. The characteristic behaviour of $c(\gamma)$ and $K(h)$ are illustrated respectively by the estimation of the scaling properties of the probability distribution and of statistical moments of simulated fields, obtained with multiplicative self similar cascade processes.

1. INTRODUCTION

For several years now, the concepts of fractals and multifractals have increasingly served as new tools in investigating and understanding the behaviour of atmospheric and other geophysical fields. It has been particularly important in characterizing and mathematically linking two fundamental properties of turbulent flows: their intermittency (high variability of the field) and scale invariance (observed over a wide range of time and/or space scales). By scaling (or scale invariance) we mean that quantities associated with the fields at different scales are related by transformations involving only the scale ratios. When all the statistical properties of these quantities can be described by a unique exponent of the scale ratio, we have simple scaling; see Schertzer and Lovejoy (1988), Lovejoy and Schertzer (this volume) for brief reviews. In general however the complex behaviour of the fields could not be reduced to simple scaling (as already pointed out by Kolmogorov (1962) and Obukhov (1962) since they found multiple scaling in their log-normal[1] model for intermittency), which leads recently to the concept of multifractal dimensions, introduced by Hertschel and Procaccia (1983), Grassberger (1983) and Schertzer and Lovejoy (1983). Multiple scaling takes into account that region of varying intensities scale with different exponents and are characterized by different (fractal) dimensions (a dimension function). The latter describe how the various intensities level of the fields are distributed - or projected - on a given space of observation. Subtracting the dimension of the space of observation from the dimension we obtain the codimension, a quantity not only scale invariant but also independent of the dimension of the space in which the field is embedded, and could be then regarded as the fundamental function characterizing the system.

The codimension function is usually determined by examining the statistical properties of the fields. This could be done by estimating directly the probability distribution of the field 's intensities or by looking at the behaviour of their statistical moments. The codimension functions obtained in both case are related to each other by a Legendre transformation as discussed in Frisch and Parisi (1985) and Halsey et al. (1986). Until now, the different attempts to estimate the codimension functions of experimental data

*EERM/CRMD, Météorologie Nationale, 2 Ave. Rapp, Paris 75007, France

[1] Corresponding to the case $\alpha = 2$ discussed below.

D. Schertzer and S. Lovejoy (eds.), Non-Linear Variability in Geophysics, 99–109.

belong to one of these approaches. The codimension function of the radar rain reflectivity fields (Schertzer and Lovejoy, 1987 a,b), of the energy dissipation fields (Meneveau and Sreenivasan, 1987) and of lidar rain drops (Lovejoy and Schertzer, this volume) have been obtained by studying the scaling behaviour of their statistical moments. On the other hand Box Counting – used recently by Hubert and Carbonnel (1988) to analyze rain gauge data – or Functional Box Counting – introduced by Lovejoy et al. (1987) and used by Gabriel et al. (1988) for satellite data – are techniques developed to determine the codimension function from the probability distributions of the fields. Unfortunately the two last methods, which are designed to work on sets, can only be applied indirectly to physical fields.

In section 2, we propose a new method to estimate the codimension function, the PDMS – Probability Distribution-Multiple Scaling – essentially based on the scaling properties of the probability distribution. We indicate how the finite number of samples impose upper limits on the measured codimension function . In section 3, its consequences for the behaviour of the (highest) statistical moments of the fields, are also discussed in parallel with the spurious (or pseudo-) scaling.

2. PDMS: PROBABILITY DISTRIBUTION-MULTIPLE SCALING

It is now well established (in other paper in this book but also in Halsey et al. (1986), Meneveau and Sreenivasan (1987), etc.) that the multiple scaling properties of the field intensities ε_λ are correctly describe by the following two equations:

$$\varepsilon_\lambda \propto \lambda^\gamma , \quad \lambda > 1 \tag{2.1}$$

$$Pr(\varepsilon_\lambda \geq \lambda^\gamma) \propto \lambda^{-c(\gamma)} \tag{2.2}$$

where λ is ratio of the largest scale of interest to the scale of homogeneity, and γ is positive it is the order of the singularity: $\varepsilon_\lambda \to \infty$ when $\lambda \to \infty$. The second equation indicates that the probability distribution of singularities of order higher than a value γ, is related to the fraction of the space occupied by them, as determined by their codimension function $c(\gamma)$. This equation holds to within slowly varying functions of λ (e.g. logs). The function $c(\gamma)$ and $K_D(h)$, used in this paper, are related to the $f(\alpha)$ singularity spectrum and the $\tau(q)$ introduced by Frisch and Parisi (1985) and used by Halsey et al. (1986), by these simple relations $c(\gamma) = d - f(\alpha)$ and $K_D(h) = -\tau(q)$; when $\alpha = (d - \gamma)$, h=q and d is the dimension of the space in which the set is embedded. It could be show that the h^{th} statistical moments of the field intensities will generally have a scaling exponent $K(h) = (h-1)C(h)$ and then using the definition for the statistical moments we obtain the following relation:

$$\langle \varepsilon_\lambda^h \rangle \propto \int \lambda^{h\gamma} \lambda^{-c(\gamma)} \, dPr(\gamma) \propto \lambda^{(h-1)C(h)} \tag{2.3}$$

In the limit $\lambda \to \infty$, the scaling exponents are related by the the following Legendre transformation:

$$K(h) = (h - 1)\, C(h) = \max_\gamma (h\gamma - c(\gamma)) \tag{2.4}$$

$$c(\gamma) = \max_h (h\gamma - K(h))$$

where C(h) (a monotically increasing function) is the codimension function associated to the statistical moments of the field intensities, as $c(\gamma)$ is the one associated to the probability distribution.

Let us now consider how these properties arise and how they could be understood in terms of cascade processes, firs postulated by Richardson. Without losing generality, we will restrict ourselves studying the dynamical cascade (i.e. energy flux). Although we discuss only this case, the cascade process could also be used to describe passive scalars (e.g. scalar variance) or pseudo scalars (the helicity flux, discussed in

Levich and Shtilman (this volume)) and have been also used to study showers of cosmic rays (Bialas and Peschanski, 1986). Here, the scale invariant quantity is the energy flux density ε whose ensemble spatial average is constant (i.e. scale independent) but nevertheless highly intermittent. This condition on the ensemble spatial average of the energy flux density corresponds to the canonical case. In this process the flux of the field at large scale multiplicatively modulates the various fluxes at smaller scales, the mechanism of flux redistribution is repeated at each cascade step (self similarity). Figure 1 gives a schematic illustration of the discrete cascade process in two dimensions: a large eddy of characteristic length l_0 with an initial energy flux density ε_0, is broken up (via non-linear interaction with other eddies or internal instability) into smaller sub eddies of characteristic length $l_1 = l_0/\lambda$ ($\lambda = 2$, is the scale ratio between to consecutive cascade step), transferring in the process to each sub-eddy a fraction of its energy flux density $\mu\varepsilon_0$.

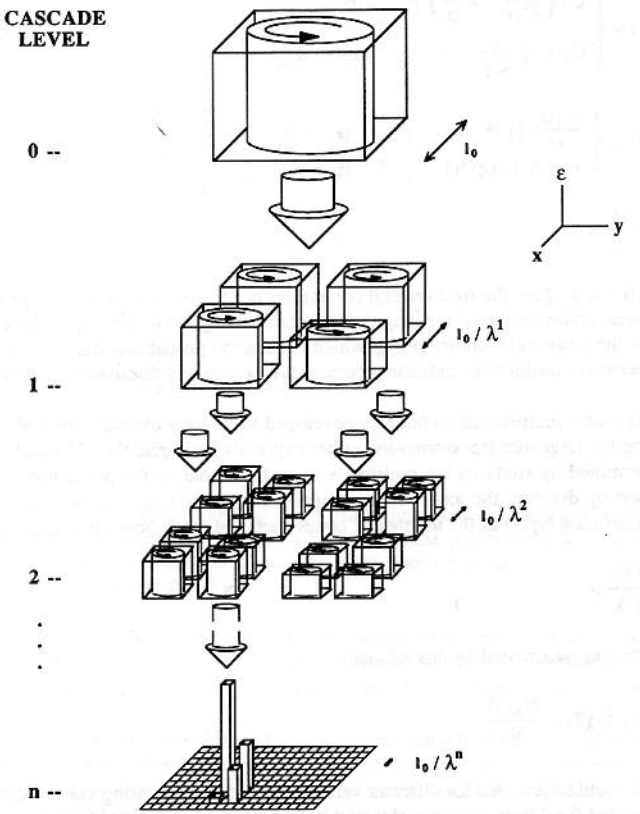

Fig. 1 : A schematic diagram showing a two-dimensionnal cascade process at different levels of its construction to smaller scales. Each eddy breaks up into four sub-eddies, transferring a part or all its energy flux.

The process could be repeated, and after n iterations (or n cascade steps) we obtain sub-eddies with length of homogeneity $l_n = l_0/\lambda^n$ and energy flux density ε_n:

$$\varepsilon_n = \mu\varepsilon_1 \, \mu\varepsilon_2 \, \mu\varepsilon_3 \ldots \mu\varepsilon_n \, \varepsilon_0$$

when $\mu\varepsilon_i$ is the fraction of energy flux density to sub-eddy of i^{th} cascade step. The last quantity is a typical "bare" quantity as long as n is finite, conversely "dressed" quantity are obtained by integrating a completely developed cascade processes, experimental data are then best approximated by the second one; see Schertzer and Lovejoy (1987a, b). As the cascade proceeds to smaller scale, high values of ε_n appears, concentrated on smallest volume l_n^d, which is habitually recognized as a basic characteristic of intermittency. In the limit of the cascade scale going to zero ($l_n \to 0$), the energy flux densities ε_n take singular values of all orders. Such behaviour is characteristic of multiple scaling, and could be expressed by relation (2.1).

For a generalization of discrete cascade processes to the continuous cascade process and its applications see Wilson et al. (this volume). In particular Schertzer and Lovejoy (1987a, b) shows that in this case the codimension functions belongs to universal classes:

$$c(\gamma) = \begin{cases} C_1\left(\frac{\gamma}{C_1\alpha'} + \frac{1}{\alpha}\right)^{\alpha'} & \alpha \neq 1 \\ C_1 \exp\left(\frac{\gamma}{C_1} - 1\right) & \alpha = 1 \end{cases} \tag{2.5}$$

$$K(h) = \begin{cases} \frac{C_1\alpha'}{\alpha}(h^{\alpha} - h) & \alpha \neq 1 \\ C_1\, h\, \mathrm{Log}(h) & \alpha = 1 \end{cases} \tag{2.6}$$

$$\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$$

where C_1 and α ($0 \leq \alpha \leq 2$) are the fundamental parameters needed to characterize the processes. The first one, C_1 is the fractal codimension of the singularities contributing to the average values of the field, and the second one, α the Lévy index indicating to which classes the probability distribution belongs. These fundamental properties of multifractal cascade processes are extensively discussed in Schertzer and Lovejoy (this volume).

The PDMS is a new multifractal techniques developed to directly estimate the codimension function $c(\gamma)$, without using the Legendre transformation either explicitly or implicitly. The codimension function is essentially determined by studying the multiple scaling behaviour of the probability distribution . The last one is obtained by dividing the total volume, support of the field ε_λ, into $N_\lambda = \lambda^d$ disjoint boxes of volume λ^{-d}, then defining $N_\lambda(\gamma)$ as the number of boxes such that the following inequality is verified:

$$\frac{\log \varepsilon_\lambda}{\log \lambda} \geq \gamma$$

then equation (2.2) is approximated by this relation:

$$\Pr(\varepsilon_\lambda \geq \lambda^\gamma) \propto \frac{N_\lambda(\gamma)}{N_\lambda}$$

The operation could be repeated for different values of γ, and for decreasing values of scale ratio λ. The corresponding rescaled field intensities are obtained by averaging over the field intensities ε_λ at the finest resolution; and are therefore "dressed" quantities. The slope of the curve of $N_\lambda(\gamma)/N_\lambda$ as a function of the scale ratio λ on a log-log graph, for a given value of γ, will give the corresponding codimension function $c(\gamma)$. This procedure bypassed the problem of the correct but non trivial normalization of eq. (2.2).

Using this techniques, we estimate the codimension function of unidimensional ($d = 1$) simulated fields generated with the help of discrete log-normal ($\alpha = 2$) multiplicative cascade process. The field has been simulated over 10 cascade steps ($n = 10$), i.e. the scaling ratio between the largest and smallest length is 2^{10}.

The eq. (2.5) with $\alpha = 2$ gives the theoretical expression of the codimension function:

$$c(\gamma) = \frac{C_1}{4}\left(\frac{\gamma}{C_1} + 1\right)^2 \qquad (2.7)$$

The free parameter C_1, $0 \le C_1 \le d$[1], is chosen to equal to 0.125 and the number of independent samples N, used to estimate to codimension function, is equal to 10 in the first example and to 1000 in the second. The curves of the $\log(N_\lambda(\gamma)/N_\lambda)$ as function of $\log(\lambda)$ (λ decreasing by a factor of 1/2, from 1024 to 2) are given in fig. 2 for different values of γ. The straightness of the curves indicates that the scaling is accurately respected.

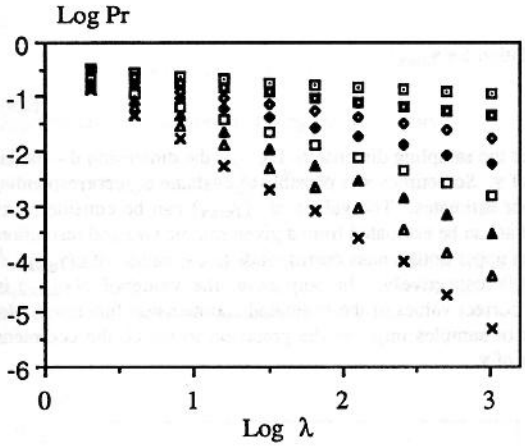

Fig. 2 : Illustration of the scaling behaviour of the probability distribution of 1000 independent samples of density fields induced by log-normal ($\alpha = 2$) multiplicative cascade process over a scale range $2^{10} = 1024$. From bottom to top, each curve corresponds to an order of singularity γ going from 0.1 to 0.8, in increments of 0.1. Note that the scaling is accurately followed.

In fig. 3, the curves of the estimated codimension functions, with N = 10 (black symbols) and N = 1000 (white symbols), are compared to the theoretical functions (continuous curves) given by eq. (2.7). They are in good agreement up to a given value of γ, which obviously increases with the samples size.

This behaviour could be understood if we recall that in the canonical case, a finite number of samples, each of finite scale ratio, will impose an upper bound on the highest values of the order of singularities γ that could be generated by the multiplicative cascade processes, and conversely the accessible values of γ will be limited by the finite sample size used to evaluated it. This implies that the estimate of the frequency of occurrence of the largest value of γ, from which we estimate the codimension function, could not be performed on only one sample, as they will only appear in some samples events. Obviously, increasing the number of samples will increase the probability of their observation. The codimension function $c(\gamma)$ is then a measure of the fraction of the space, formed by the total number of samples of dimension d (corresponding.roughly to the hyper space of the probability distribution)), occupied by the singularities of order equal or superior to γ.

[1] The case $C_1 > d$ yields degenerate cascades discussed by Schertzer and Lovejoy (1987 a, b).

Following this and noticing that $c(\gamma)$ is an increasing function for positive values of γ (indicating that the largest singularities are the rarest) one will find that the maximum value of γ, denoted γ_{max}, observed at least once in the N independent samples of volume d (with λ^d sub-boxes on each sample) is approximated, within a normalization constant, by:

$$N \lambda^d \lambda^{-c(\gamma_{max})} \sim 1 \tag{2.8}$$

Introducing the definition of the dimension of sampling D_S for N samples:

$$\lambda^{D_S} = N$$

$$D_S = \frac{\log N}{\log \lambda}$$

Using eq. (2.8) we obtain the following relation for γ_{max}:

$$c(\gamma_{max}) \approx d + D_S \tag{2.9}$$

The last equation shows that the larger the sampling dimension D_S – or the dimension d – the larger will be the spectrum of accessible values of γ. Sometimes it is possible to evaluate $c(\gamma)$ corresponding to γ greater than γ_{max}, but they are then poor estimates. The values of $c(\gamma_{max})$ can be considered as an indication of the maximum values of $c(\gamma)$ that can be estimated from a given sample size and resolution λ.

Going back to fig. 3, the lower and the upper dotted lines corresponds to the values of $c(\gamma_{max})$, for a number of samples N = 10 and N = 1000 respectively. In both case, the value of $c(\gamma_{max})$ is an appropriate upper limits for the maximum correct values of the estimated codimension functions. Notice also in fig. 3, that increasing the number of samples improve the precision to which the codimension function is determined even at small values of γ.

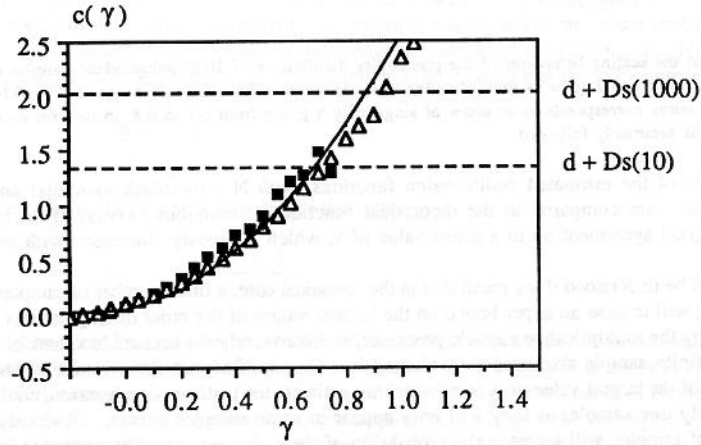

Fig. 3 : The curves of the estimated $c(\gamma)$, using the PDMS technique, with 1000 independent samples (white symbols) and 10 independent samples (black symbols) compared to their theoretical values given by the continuous curve. The horizontal dashed lines, indicate the estimated upper limits of $c(\gamma)$ imposed by the finite values of the dimension of sampling D_s. In both cases, the estimated $c(\gamma)$ is in good agreement with the theoretical curve when $c(\gamma) < d+D_s$.

3. UPPER LIMITS OF THE ESTIMATED CHARACTERISTIC FUNCTION K(h)

We have seen at the preceding section that a finite number of samples restricts the breath spectrum of the accessible γ, and hence of $c(\gamma)$, to those smaller than the maximum values γ_{max} and $c(\gamma_{max})$ respectively. Now we will see its consequences on the determination of statistical moments, in particular for the trace-moments, and how it will affect the behaviour of the codimension function C(h) associated to them. Following Schertzer and Lovejoy (1989) we define the h^{th} trace moment of the energy flux densities over a set A of dimension d by these relations:

$$Tr_{A_\lambda} \varepsilon_\lambda{}^h = < \int_{A_\lambda} \varepsilon_\lambda{}^h \, d^{hd}x > \propto \lambda^{K_D(h)} \tag{3.1}$$

with:

$$K_D(h) = (h - 1)\,[C(h) - d] \tag{3.2}$$

and approximated for discrete fields by this summation:

$$Tr_{A_\lambda} \varepsilon_\lambda{}^h \approx < \sum_{A_\lambda} \varepsilon_\lambda{}^h \, \lambda^{-hd} > \tag{3.3}$$

Using this we estimate the trace moments of fields, generated the by log-normal ($\alpha = 2$) multiplicative cascade process already discuss in section 2. The statistical ensemble averages are replaced by sums over N independent samples of the fields.

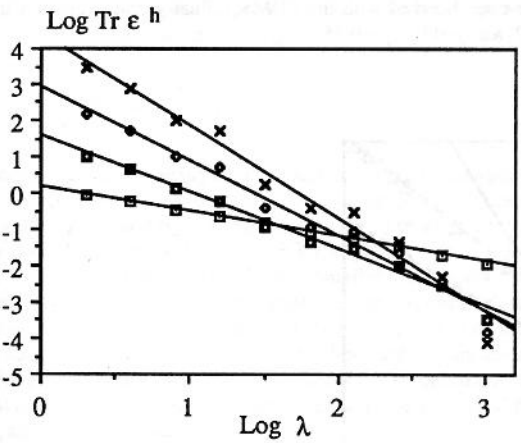

Fig. 4 : The estimated h^{th} trace moments of 1000 independent samples of density fields induced by log-normal ($\alpha = 2$) multiplicative cascade process with scale range equal to 2^{10}. The curve correspond at h = 2, 5, 7 and 9 going from the bottom to the top. Here also the scaling is accurately followed.

In fig. 4, the curves of the log ($Tr_{A\lambda} \varepsilon_\lambda{}^h$) against log($\lambda$) are illustrated for several values of h (for a number of sample N = 1000); here also the scaling is well respected for all the range of λ (from 1 to 1024). The slopes of these curves gives the estimated codimension functions, with N = 10 (black

symbols) and N = 1000 (white symbols), are compared in fig. 5 to the theoretical curve (continuous line) of K(h), obtained with eq. (2.6) and $\alpha = 2$. We see that for small h the three curves are in good agreement, but for large h the estimated values of K(h) no longer correspond to the theoretical parabola, but are asymptotically straight. This behaviour could be understood as a failure in the Legendre transformation, given by eq. (2.4), essentially caused by the finite sample size. The Legendre transformation implies a dual correspondence between the orders of moments h and the orders of singularities γ (given by the following relation $h = c'(\gamma)$ and $\gamma = K'(h)$). So as long as γ is smaller then γ_{max}, the corresponding h and K(h) are in agreement with the theoretical curve given by the Legendre transformation. The finite number of sample implies a γ_{max}, but we have no restriction of this kind on the order of the moments h, that could be increase indefinitely (in practice its value stay finite), corresponding to γ greater then γ_{max}, and as for $c(\gamma)$ the estimated K(h) is no longer in agreement with the codimension function of the system. It is clearly not possible to estimate correctly the codimension function of singularities never encountered.

The asymptotic linear behaviour of K(h) could then be approximated by the following relation:

$$K(h) \approx h\,\gamma_{max} - c(\gamma_{max}) \tag{3.4}$$

(This relation could be deduced from eq. (2.3), using the asymptotic expansion for general Laplace integrals when the maximum values of the exponent $(h\,\gamma - c(\gamma))$ lies at one of the endpoints of the integral and its derivative is no longer equal to zero; see, e.g. Bender and Orszag (1978).) The asymptotic slopes and the y axis intercept, estimated for the 5 last points K(h) in fig. 5, are equal to 0.54 and 1.01 with sample size N = 10, and to 0.74 and of 1.24 when N = 1000. Using PDMS to determine the $c(\gamma)$ of the same fields, one found that $c(0.55) \approx 1.1$ when N = 10 and $c(0.75) = 1.3$ when N = 1000. So the values given by the estimation of the asymptotic slopes of K(h) are in good agreement with that one obtained with PDMS. However the values of γ_{max} observed for both examples are respectively 0.75 and 0.95 for N equal to 10 and 1000, and are slightly greater than the values given by the asymptotic slopes of K(h). But – as already pointed out – these highest values of the order of singularity are extremely rare events and their probability of occurrence too low to contribute significantly to the statistical moments, which is why the values of γ_{max} in eq. (3.4) will be smaller than the one observed with the PDMS. Finally, the slope of C(h) is evaluated for values of h between 0.1 and 2, and yields $C_1 \approx 0.15$.

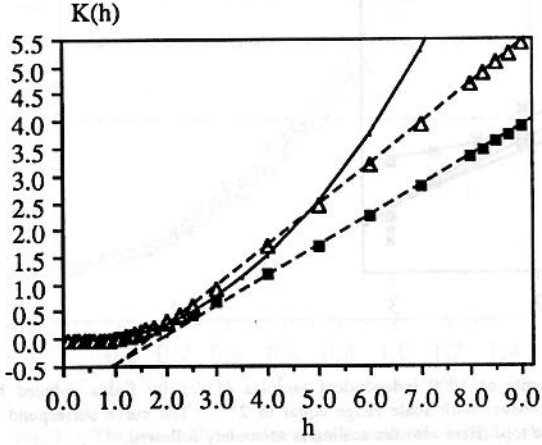

Fig. 5 : The curves of the estimated moment scaling function K(h) for 1000 independent samples (white symbols) and 10 independent samples (black symbols) compared to theirs theoretical values given by the continuous curve. The dashed lines are the asymptotic slopes of K(h) evaluated for the five last values of h.

This underestimation of the codimension function should not be confused with that resulting from the divergence of the trace moments, which leads to spurious scaling. Since C(h) is a monotonically increasing function, the scaling exponents $K_D(h)$ of the trace moments given by eq. (3.2) have two zero: $h = 1$ implies the scale independence of the averaged fields, and $h = h_d > 1$, defined by the following equality:

$$C(h_d) = d \tag{3.5}$$

So when h is greater than this critical value, the scaling exponent $K_D(h)$ is positive, and in the limit of $\lambda \rightarrow \infty$ the trace moment, given by eq. (3.1), will diverge. The divergence of trace moments and its interpretation as a Hausdorff measure was first pointed out in Schertzer and Lovejoy (1987 a,b). However statistical moments of the "bare" quantities always remains finite. But for the dressed fields, resulting from the integration of a completely developed multiplicative cascade processes (i.e. in the limit $\lambda \rightarrow \infty$), the divergence of moments implies a break down in the law of large numbers, and then the usual procedure of estimating their statistical properties fails for $h > h_d$. This divergence occurs because the dimension d of the volume over which the field is integrated is not large enough to smooth out rare high order singularities, and rescaling the field by averaging it at larger scales will not "kill" these singularities, or attenuated them enough, since the dimension d is not affected by this procedure. (The behaviour of the statistical moments of dressed fields for some special cases of discrete multiplicative cascade processes and h integer is derived in Lavallée (1990).)

As soon as the number of samples is large enough so that h_{max} (i.e. the order of moments corresponding, by the Legendre transformation to γ_{max}) is greater than h_d, it implies that rare high singularities are encountered, and spurious scaling will be observed. However it is important to notice how to distinguish the spurious scaling that underestimation of K(h) due to the under sampling. In the latter case, the estimated $K_D(h)$ becomes more and more negative, as can be seen on the graph of the $\log(Tr_{A_\lambda} \varepsilon_\lambda{}^h)$ versus $\log\lambda$ in fig. 4 where their slopes decreases to lower and lower negative values, as h increase. On the contrary spurious scaling will be observed when the slope of $\log(Tr_{A_\lambda} \varepsilon_\lambda{}^h)$ versus $\log\lambda$ goes to zero for $h > 1$, that implies $K_D(h)$ became less negative as h goes to h_d.

The following relation for h_{max}:

$$h_{max} = \left(\frac{d + D_s}{C_1} \right)^{1/2} \tag{3.6}$$

is obtained for $\alpha = 2$, using eq. (2.9) and the Legendre transformation.

Using discrete log-normal ($\alpha = 2$) multiplicative cascade process with 10 cascade steps and $C_1 = 0.72$, we found $h_d = 1.388$ using eq. (2.7) to solve eq. (3.5) with $d = 1$, and the number of samples $N = 2000$, is chosen so that $h_{max} = 1.7$ is greater than h_d. In the preceding example, with $C_1 = 0.125$, and for $N = 1000$, $h_{max} = 4$ which is smaller than $h_d = 8$. We then use eq. (3.3) to estimate the trace moments of the fields at different scale ratio, and as usual to estimate the scaling exponents by their slopes. The estimated codimension function C(h) is compared to its theoretical values (continuous curve) in fig. 6. For large values of h, the estimates stay bounded and are therefore poor estimator of the theoretical C(h) which grows linearly with h. This behaviour is recognized as the spurious scaling. For small values of h, C(h) is in agreement with the theoretical curve, the slope of C(h) evaluated for values of h between 0.1 and 0.9, and gives C_1 equal 0.75.

So to increase the range of values of h for which one can expect to correctly determine the scaling exponents, we must first increase N, until $h_{max} > h_d$, and then measure (or project) the fields on spaces of increasing dimension d. In practice this means that experimental data will have to be taken with instruments of the largest dimension possible (e.g. $d = 4$ for isotropic processes varying in space and time). Finally we would like to point out that while the spurious scaling can only be observed on "dressed" fields, the finite number of samples will modified the behaviour of both the "bare" and of the "dressed" fields and can be observed on either. One can find other examples of the application of PDMS as well as trace moments to geophysical fields in Lovejoy and Schertzer (this volume).

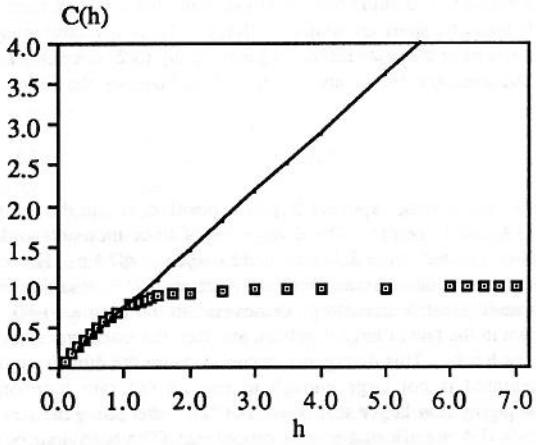

Fig. 6 : Illustration of the consequences of spurious scaling of the "dressed" quantities for the estimated C(h), that stay bounded for large values of h, in disagreement with the linear behaviour of the theoretical C(h) given by the continuous curve. The estimated C(h) is obtained by trace moment analysis of 2000 independent samples of density fields induced by log-normal ($\alpha = 2$) multiplicative cascade process, the scale ratio $1 = 2^{10}$.

4. CONCLUSIONS

The results obtained in this paper underline the importance of phenomenological stochastic models - in particular multiplicative cascade processes - not only in understanding turbulence, but also in verifying the relevance and limits of conceptual tools developed for analyzing their scaling properties. We introduce a new method to directly estimate the codimension function ($c(\gamma)$) which is directly related to the probability distribution. The PDMS method avoids using Legendre transformations and their implicit assumption of the converge of all statistical moments. Since we generally expect the observed "dressed" moments to diverge for sufficiently high orders, this method is a significant improvement: it is insensitive to spurious scaling. We also discuss how a finite number of samples reduces the range of accessible γ, and how to characterize it with the dimension of sampling D_S. Its effects on K(h), the Legendre transform of $c(\gamma)$, are also studied in parallel with spurious scaling. This clarifies their differences and shows how the asymptotic behaviour of the real scaling exponents are modified.

We obtained specific estimates of the maximum order of singularities that can be evaluated with a given sample size and resolution. These limits give precise information about the reliability of exponent estimates and suggests that many $K_D(h)$ (or equivalently $\tau(q)$) values cited in the literature should be critically re-examined, especially in the (common) case where straight-line behaviour is observed at large h. This asymptotic linearity could also introduce a discontinuity, or a sudden change in the derivative of $K_D(h)$. Since there is a formal analogy between thermodynamics and "flux dynamics", it has often been claimed on the basis of observed discontinuities that phase transitions occur. We discuss several mechanisms by which discontinuities will arise even when no "phase transition" is present. This suggests that the analogy with thermodynamics may be largely formal in nature.

APPENDIX A

In this section, the fundamental distinction between the "bare" and "dressed" is discussed and analytical expressions for the latter are derived. The behaviour is studied for both h smaller and greater than h_D. The theoretical expression for the "bare" h^{th} trace moment of the energy flux densities over a set A of dimension D is straightforward and given by the eq. (3.1). For simplicity the discrete cascade processes will be the only ones considered, generalizations to continuous cascade processes are possible following the same steps but are not essential for the discussion here.

From here to the end of this section $\lambda = l_n/l_{n+1}$ ($\lambda > 1$) will be the constant scale ratio between two consecutive cascades steps (or between the length of homogeneity of two consecutive sub-eddies) (see fig. 1). The expression for the h^{th} statistical moments of the energy flux densities is then given by the following relation:

$$\langle \varepsilon_n{}^h \rangle \propto \lambda^{n(h-1)C(h)} \qquad n >> 1 \tag{A.1}$$

The index n indicates the number of cascade steps used to iterate the cascade process to obtain the field ε_n distributed over the volume $l_n{}^D = \lambda^{-nD}$ (when l_0 and ε_0 are put equal to 1 for simplicity). It also refers to "bare" quantities (as long as it is finite).

Using the same fields, they are divided in λ^{mD} disjoint boxes of equal volume λ^{-mD}, covering the entire field. When the index m takes values between 0 and n, the average energy flux density on each boxes of volume λ^{-mD} is given by the following expression:

$$\bar{\varepsilon}_m = \frac{\sum \varepsilon_n \lambda^{-nD}}{\sum \lambda^{-nD}} \tag{A.2}$$

for h, h-1, ... , 2, 1 ≠ h_D and the coefficients a_h are independent of the number of cascade steps n. Then for n >> 1, the h^{th} moment of the energy flux and the h^{th} trace moment – using eq. (A.4) – of averaged fields have respectively the following behaviour:

$$\langle \Pi(\varepsilon_n)^h \rangle \sim \begin{cases} a_1 & 1 < h < h_D \quad K_D(h) < 0 \\ a_h \lambda^{nK_D(h)} & h > h_D \quad K_D(h) > 0 \end{cases}$$

$$Tr_{Am} (\bar{\varepsilon}_m)^h \sim \begin{cases} a_1 \lambda^{mK_D(h)} & 1 < h < h_D \quad K_D(h) < 0 \\ a_h \lambda^{nK_D(h)} & h > h_D \quad K_D(h) > 0 \end{cases} \qquad (A.6)$$

where we keep only the dominant term, and we use the fact that $K_D(h) > K_D(h-1)$ for $h > h_D$ because C(h) is a monotonic increasing function.

The values of the h^{th} trace moment of averaged fields given by eq. (A.4) and (A.5) – without the approximation used to obtain eq. (A.6) – are compared to the estimated trace-moments in figs. 7 and 8. In the first case the trace moments have been estimated using 1000 independent samples of density fields produced by log-normal cascade processes with scale range equal to 2^{10}. The free parameter $C_1 = 0.08$ is chosen such that $h_D > h_{max} > 4$. Both values are in good agreement, and the approximation used in eq. (A.6) – corresponding to the first case, $1 < h < h_D$ – that the trace moment of averaged fields scale like their "bare" counterpart is more accurate when the volume of integration of the boxes (or averaging volume) λ^{mD} is much larger then the volume of homogeneity λ^{nD}.

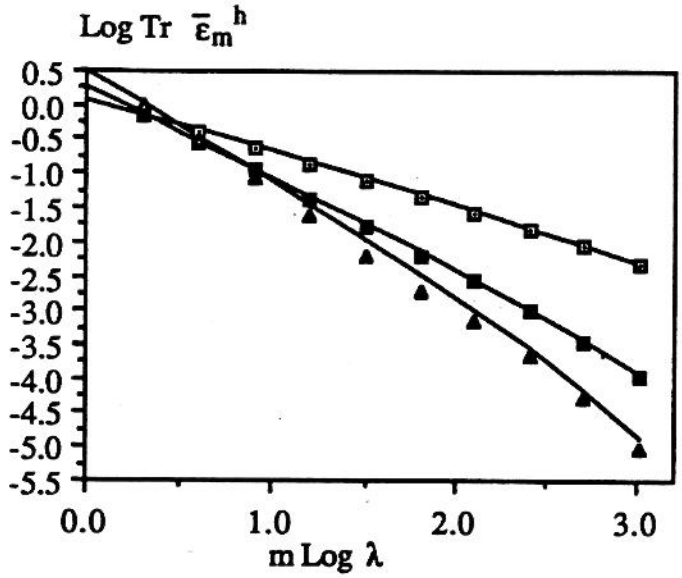

Fig. 7 : The estimated h^{th} moments of the average energy flux densities as a function of the number of cascade steps, for h = 2, 3 and 4 < h_D from bottom to top, are compares to the values of the expression (A.4) given by the continuous curve. Both of them are in good agreement. The energy flux density fields are produced by a log-normal ($\alpha = 2$) multiplicative cascade process over a scale range $2^{10} = 1024$, and the number of independent samples N = 1000.

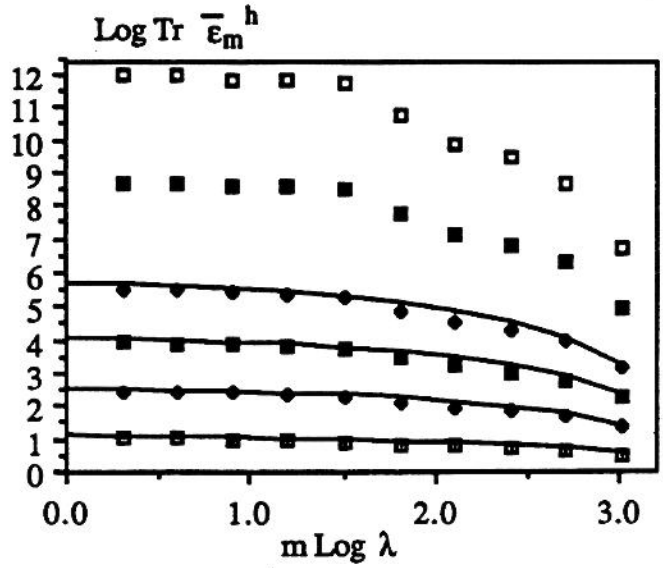

Fig. 8 : Same as figure 7, but for for h = 2, 3, 4 and 5, $h_D < 2$ from bottom to top. The estimated trace moment for h = 7 and 9 are given alone, observe that their behaviour becomes constant as m decreases. The energy flux density fields are produced by the "α-model". The parameters were the following; C = 1.43, α = 1.1 (not to be confused with α, the Lévy index used in the preceding section) and the number of independent samples N = 1500.

In fig. 8 the estimated trace moments have been evaluated using the α-model (see Schertzer and Lovejoy (1987a, b) and Schertzer and Lovejoy (1990). In these α-model simulation the parameters take respectively the following values: c = 1.43, α = 1.1 (not to be confused with α the Lévy index used in the preceding section), the number of samples N = 1500 and $h_D \approx 1.14 < h_{max}$. Again both values are in good agreement, and the behaviour of the trace moment given by eq. (A.6) – corresponding this time to the second case when $h > h_D$ – indicating that the trace moments go to constant values as expected when λ^{mD} is much larger then λ^{nD}. The same behaviour is also observed for larger values of h in fig. 8. The trace moment of fields generated by log normal cascade processes exhibit also the same behaviour for $h > h_D$, but the number of samples needed to obtain convergence to the value given by eq. (A.6) is larger than in the case of the α-model, especially for large values of h. This is essentially due to the fact that the latter has a bounded maximum value of γ.

The "dressed" quantities are usually defined for $n \rightarrow \infty$, and therefore the behaviour of the "dressed" trace moment is well described by the last equation. As long as $K_D(h)$ is negative, i. e. $C(h) < D$, the "dressed" trace moments are scaling, and their estimates at different scales of integration λ^{-mD}, allows the scaling exponents to be estimated by the slope of $Tr_{Am} (\overline{\varepsilon}_m)^h$ as a function of the scale ratio λ^m on a log-log graph. The (multiple) scale invariance of the "dressed" fields observed at a given scale, is only the consequences of the same behaviour of the "bare" fields at the corresponding length scale as indicated by eq. (A.4). When $K_D(h)$ is positive, i. e. $C(h) > D$, the "dressed" trace moment have no more dependance on the scale of integration λ^{-mD}, as illustrated in fig. 8, and the methods of determination of the scaling exponent suggested for $K_D(h) < 0$ is of no use here. This behaviour is a consequence of the divergence of statistical moments for $n \rightarrow \infty$ when $h > h_D$. It is related to the impossibility of filtering out or smoothing out the largest order singularity of the fields without increasing the dimension of the space in which the field is embedded.

where the sum is over all $\lambda^{(n-m)D}$ sub-boxes of volume λ^{-nD}. The numerator is simply the energy flux over the boxes of volume λ^{-mD}, and the denominator its volume. Their ratio gives the (spatial) average values of the energy flux densities (indicated by a bar over ε_m) We obtain therefore a rescaled field at larger resolution λ^{-m}.

Taking into account the fact that the ε_n appearing in eq. (A.2) results from a multiplicative cascade process it can be then factored into two parts: $\varepsilon_n = \varepsilon_m \varepsilon_{n-m}$. When ε_m is the flux density distributed over the sub-eddy of the m^{th} generation before each of them breaks into $\lambda^{(m+1)D}$ sub-eddies, and is then the common ancestor of the ε_{n-m} distributed over sub-eddies of volume λ^{-nD} as the cascade progresses from the m^{th} to the n^{th} cascade step. The prefactor ε_m is then independent of the sum in eq. (A.2). Averaging the process over disjoint boxes of a given size corresponds to a translation in the number of cascade steps: equivalent to the evaluation of energy flux of fields resulting from a cascade processes developed over (n - m) cascade steps, with a particular distribution of the initial energy flux densities over eddies of volume λ^{-mD}. Following this the average becomes:

$$\bar{\varepsilon}_m = \varepsilon_m \sum \varepsilon_{n-m} \lambda^{-(n-m)D}$$

and, using eq. (3.3), the h^{th} trace moment at the same resolution λ^{-mD} is then:

$$Tr_{Am} (\bar{\varepsilon}_m)^h = < \sum \bar{\varepsilon}_m{}^h \lambda^{-mhD} >$$

The multiplicative increments operating at a given cascade step are statistically independent of those of the previous steps so that ε_m is not correlated to ε_{n-m}, even though the ε_n of the same generation are correlated. This essentially means that the redistribution of energy flux density from eddy to sub-eddies is independent of the preceding one. The h^{th} trace moment of an averaged field is then given by these expressions:

$$Tr_{Am} (\bar{\varepsilon}_m)^h = \lambda^{-mhD} \sum <\varepsilon_m{}^h> <(\sum \varepsilon_{n-m} \lambda^{-(n-m)D})^h > \qquad (A.3)$$

where the first sum is made over the λ^{mD} boxes of the set A_m, and the second over $\lambda^{(n-m)D}$ sub-boxes included in each box of volume λ^{-mhD}. Using eq. (A.1) and the fact that the second sum can be identified as the usual energy flux $\Pi(\varepsilon_{n-m})$, for h = 1, transferred from the eddy of the m^{th} generation to that of the $(n-m)^{th}$ generation as the cascade proceeds to smaller scales, the last equation could then be rewritten:

$$Tr_{Am} (\bar{\varepsilon}_m)^h \propto \lambda^{mK_D(h)} <\Pi(\varepsilon_{n-m})^h> , \quad K_D(h) = (h-1)\ (C(h) - D) \qquad (A.4)$$

The first term on the right hand side of the proportionality sign is the expression for "bare" trace moment and the second is the usual h^{th} statistical moments of the fields.

The factorization used to obtain eq. (A.4) could also be used to obtain a recursion formula for the h^{th} statistical moments $<\Pi(\varepsilon_n)^h>$ of a cascade process developed to the n^{th} cascade step, when h is a positive integer. A detailed derivation of this formula can be found in Lavallée (1990); it consists mainly in first factorizing the energy flux density of the first generation ε_1 from the energy flux $\Pi(\varepsilon_n)$ and recognizing that the term multiplying each ε_1 are proportional to $\Pi(\varepsilon_{n-1})$. Finally using the statistical independence of the multiplicative increments and multinomial expansion leads to the recursive formula sought. In particular for D = 1 and λ = 2, we can establish, using the relations (A1) and (A.4), that for positive integer h the h^{th} moment of the energy flux is given by:

$$<\Pi(\varepsilon_n)^h> = a_h \lambda^{nK_D(h)} + a_{h-1} \lambda^{nK_D(h-1)} + ... + a_2 \lambda^{nK_D(2)} + a_1 \qquad (A.5)$$

5. ACKNOWLEDGMENTS

We acknowledge fruitful discussions with A. Davis, P. Gabriel, J.P. Kahane, P. Ladoy, E. Levich, A. Saucier, G. Sarma, G. Sèze, Y. Tessier, R. Viswanathan, and J. Wilson.

6. REFERENCES

Bialas, A., R. Peschanski, 1986: Moments of rapidity distributions as a measure of short-range fluctuations in high-energy collisions. Nucl. Phys. B, B273, 703-718.

Bender, C.M., S.A. Orszag, 1978: Advanced mathematical methods for scientists and engineers. McGraw-Hill, New York, NY.

Frisch,U., G. Parisi, 1985: A multifractal model of intermittency. in Turbulence and predictability in geophysical fluid dynamics and climate dynamics, Eds. Ghil, Benzi, Parisi, North-Holland, 84-88.

Gabriel, P., S. Lovejoy, D. Schertzer, and G.L. Austin, 1988: Multifractal analysis of resolution dependence in satellite in satellite imagery. Geophys. Res. Lett., 15, 1373-1376.

Grassberger, P., 1983: Generalized dimensions of strange attractors. Phys. Lett., A 97, 227.

Halsey, T.C., M.H. Jensen, L.P. Kadanoff, I. Procaccia, B. Shraiman, 1986: Fractal measures and their singularities: the characterization of strange sets. Phys. Rev., A 33, 1141-1151.

Hentschel, H.G.E., I. Proccacia, 1983: The infinite number of generalized dimensions of fractals and strange attractors. Physica, 8D, 435-444.

Hubert, P., J. P. Carbonnel, 1988: Caractérisation fractale de la variabilité et de l'anisotropie des précitations tropicales. C. R. Acad. Sci. Paris, 2-307, 909-914.

Kolmogorov, A. N., 1962: A refinement of previous hypothesis concerning the local structure of turbulence in viscous incompressible fluid at high Reynolds number. J. Fluid Mech., 13, 82-85

Lavallée, D., 1990: Ph. D. Thesis, University McGill, Montréal, Canada.

Lovejoy, S., D. Schertzer, A.A. Tsonis, 1987a: Functional box-counting and multiple elliptical dimensions in rain. Science, 235, 1036-1038.

Lovejoy, S., D. Schertzer, 1990: Multifractal analysis techniques and the rain and cloud fields from 10^{-3} to 10^{-6}m. (this volume).

Levitch, E., I. Shtilman, 1990: Helicity fluctuations and coherence in developed turbulence. (this volume).

Mandelbrot, B., 1974: Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. J. Fluid Mech., 62, 331-350.

Obukhov, A., 1962: Some specific features of atmospheric turbulence. J. Geophys. Res, 67, 3011-3014

Meneveau, C., K. R. Sreenivasan, 1987: Simple multifractal cascade model for fully developed turbulence. Phys. Rev. Lett., 59(13), 1424-1427.

Schertzer, D., S. Lovejoy, 1983: On the dimension of atmospheric motions. Preprint Vol., IUTAM Symp. on Turbulence and Chaotic phenomena in Fluids, Kyoto, Japan, IUTAM, 141-144.

Schertzer, D., S. Lovejoy, 1984: On the dimension of atmospheric motions. Turbulence and chaotic phenomena in fluids, Ed. Tatsumi, Elsevier North-Holland, New York, .505-508.

Schertzer, D., S. Lovejoy, 1987a: Physical modelling and analysis of rain and clouds by anisotropic, scaling multiplicative processes. J. Geophys. Res., 92(D8), 9693-9714.

Schertzer, D., S. Lovejoy, 1987b: Singularités anisotropes, divergences des moments en turbulence: invariance d'échelle généralisé et processus multiplicatifs. Annales Math. du Qué., 11, 139-181

Schertzer and Lovejoy 1988: Multifractal simulation and analysis of clouds by multiplicative process. Atmospheric Research, 21, 337-361

Schertzer, D., S. Lovejoy, 1990a: Scaling nonlinear variability in geodynamics: Multiple singularities, observables and universality classes. (this volume).

Schertzer, D., S. Lovejoy, 1990b: Nonlinear variability in geophysics: multifractal simulations and analysis. in Fractals: physical origins and properties, edited by L. Pietronero, Plenum, New York, NY, pp.49-79.

Wilson, J., D. Schertzer, S. Lovejoy, 1990: Physically based cloud modelling by multiplicative cascade processes. (this volume).