Scaling and multifractal fields in the solid earth and topography

S. Lovejoy, D Schertzer

Nonlin. Processes Geophys., 14, 465–502, 2007
www.nonlin-processes-geophys.net/14/465/2007/

Nonlinear Processes in Geophysics

Scaling and multifractal fields in the solid earth and topography

S. Lovejoy[1,2] **and D. Schertzer**[3,4]

[1]Physics, McGill University, 3600 University st., Montreal, Que. H3A 2T8, Canada
[2]GEOTOP_UQAM/McGill Center, Montreal, Qc, Canada
[3]CEREVE, Ecole Nationale des Ponts et Chaussées, 6–8, avenue Blaise Pascal, Cité Descartes, 77455 Marne-La-Vallee Cedex, France
[4]Meteo France, 1 Quai Branly, 75007 Paris, France

Received: 29 March 2007 – Revised: 4 July 2007 – Accepted: 4 July 2007 – Published: 2 August 2007

Abstract. Starting about thirty years ago, new ideas in nonlinear dynamics, particularly fractals and scaling, provoked an explosive growth of research both in modeling and in experimentally characterizing geosystems over wide ranges of scale. In this review we focus on scaling advances in solid earth geophysics including the topography. To reduce the review to manageable proportions, we restrict our attention to scaling fields, i.e. to the discussion of intensive quantities such as ore concentrations, rock densities, susceptibilities, and magnetic and gravitational fields.

We discuss the growing body of evidence showing that geofields are scaling (have power law dependencies on spatial scale, resolution), over wide ranges of both horizontal and vertical scale. Focusing on the cases where both horizontal and vertical statistics have both been estimated from proximate data, we argue that the exponents are systematically different, reflecting lithospheric stratification which – while very strong at small scales – becomes less and less pronounced at larger and larger scales, but in a scaling manner. We then discuss the necessity for treating the fields as multifractals rather than monofractals, the latter being too restrictive a framework. We discuss the consequences of multifractality for geostatistics, we then discuss cascade processes in which the same dynamical mechanism repeats scale after scale over a range. Using the binomial model first proposed by de Wijs (1951) as an example, we discuss the issues of microcanonical versus canonical conservation, algebraic ("Pareto") versus long tailed (e.g. lognormal) distributions, multifractal universality, conservative and nonconservative multifractal processes, codimension versus dimension formalisms. We compare and contrast different scaling models (fractional Brownian motion, fractional Levy motion, continuous (in scale) cascades), showing that they are all based on fractional integrations of noises built up from singularity basis functions. We show how anisotropic (including stratified) models can be produced simply by replacing the usual distance function by an anisotropic scale function, hence by replacing isotropic singularities by anisotropic ones.

Correspondence to: S. Lovejoy
(lovejoy@physics.mcgill.ca)

1 Introduction

The surface topography and lithosphere are extremely variable over huge ranges of scales displaying structures smaller than millimeters, and as large as the planet; a ratio of at least 10^{10}. The mathematical modeling of this variability has long stimulated mathematicians and physicists. For example Perrin (1913) considered the problem of differentiability: "Consider the difficulty in finding the tangent to a point of the coast of Brittany... depending on the resolution of the map the tangent would change. The point is that a map is simply a conventional drawing in which each line has a tangent. On the contrary, an essential feature of the coast is that ... without making them out, at each scale we *guess* the details which prohibit us from drawing a tangent...". The converse problem – integrability ("rectifiability") was considered by Steinhaus (1954): "... The left bank of the Vistula when measured with increased precision would furnish lengths ten, hundred, and even a thousand times as great as the length read off a school map. A statement nearly adequate to reality would be to call most arcs encountered in nature as not rectifiable. This statement is contrary to the belief that not rectifiable arcs are an invention of mathematicians and that natural arcs are rectifiable: it is the opposite which is true...". Richardson (1961) quantified integrability by considering the empirical scaling of the coast of Britain and of several frontiers using the "Richardson dividers" method. In his paper Mandelbrot (1967) "How long is the coast of Britain?" Richardson's scaling exponent was interpreted in terms of a fractal dimension. Also among the early pioneers, we could cite Vennig-Meinesz (1951) who argued that the spectrum $E(k)$

Published by Copernicus Publications on behalf of the European Geosciences Union and the American Geophysical Union.

of the earth's topography was of the "remarkable" scaling form $k^{-\beta}$ with spectral exponent β=2; close to the modern value $\beta \approx 2.1$ (k is a wavenumber, see below for discussion), and de Wijs (1951) who – for the distribution of ores – first suggested an explicit cascade model.

Until Mandelbrot (1977)'s seminal "Fractals: form, chance and dimension" these pioneering papers were isolated. However, the 1970s were a period of explosive growth of nonlinear dynamics, particularly after the discovery by Feigenbaum (1978) and others of quantitative universality in deterministic chaos: the geoscience community was primed for new ideas. In this context, and riding on the back of the computer graphics revolution, Mandelbrot's proposal that fractals are ubiquitous in nature struck a responsive chord. It promised to characterize and model many of the messy problems of geocomplexity using unique fractal dimensions.

When it came to geophysical applications, this audacious idea turned out to have serious limitations: most geofields of interest are mathematical fields (i.e. they have a value at each space-time point such as the atmospheric temperature or rock density), and – in spite of many attempts – they cannot be reduced to geometric sets of points. They therefore cannot generally be characterized by unique fractal dimensions. Furthermore, the proposed fractal sets were only scale invariant under isotropic scale changes or occasionally under the slightly more general "self-affine" scale changes in which different exponents act in different orthogonal directions.

Since Mandelbrot's original proposal of applying fractal geometry to natural systems, geoscience applications – especially in turbulence – played an important role in stimulating advances. There are four key developments on which we focus here. The first is the realization that a generic consequence of scale invariant dynamics – where the same basic mechanism repeats scale after scale from large to small – are multifractal fields i.e. it requires the transition from fractal geometry to multifractal processes. In these "cascades", the variability is built up scale after scale; the generic result is that the extremes are particularly singular, they display "divergence of moments" or equivalently algebraic/power law ("Pareto") distributions (also called the "multifractal butterfly effect" (Lovejoy and Schertzer, 1998)). Since Bak et al. (1987, 1988), the combination of fractals combined with algebraic probabilities has been termed "Self-Organized Criticality" (SOC), therefore cascades can be said to provide an alternative nonclassical "multifractal phase transition" route to SOC (Schertzer and Lovejoy, 1994). The third advance was the realization that when – over a finite range of scales – such a scaling process interacts with many others or is iterated enough, that the resulting behaviour is stable and attractive. This implies that it doesn't depend on many of the details of the dynamics; i.e. that there exist "universality classes" for multifractal processes. This essentially reduces the number of exponents from infinity to only three and finally allows multifractals to be manageable. This is a kind of multiplicative central limit theorem. The fourth key advance was the recognition that scale invariance is a very general (although nonclassical) symmetry principle. The development of this "Generalized Scale Invariance" (GSI) effectively extended scale invariance from the restrictive and unrealistic isotropic ("self-similar") fractals and multifractals to highly anisotropic systems. In scaling but anisotropic systems, as one "zooms" into a structure, one finds that the "blown up" structure is (statistically) equivalent to the starting structure only if in addition to the magnification, one "squashes" and/or rotates the structures by an amount which depends on a scale invariant rule. When viewed using traditional (isotropic, Euclidean) notions of scale, one finds that structures at different scales and possibly different locations – can be quite different. GSI thus demonstrates the "phenomenological fallacy": that one is not justified in infering dynamics from phenomenological appearance. More concretely, the common geophysical approach of making a hierarchy of different dynamical models to cover different ranges of scales is often unjustified.

In this review, we focus on the scaling of geophysical quantities that can best be represented as mathematical fields, i.e. having a value (e.g. altitude) or intensity (e.g. rock density) at each point. From these intensive variables, various extensive quantities can be derived. For example, the distribution of islands (the "Korcak law" (Korcak, 1938)), the size of ore deposits (e.g. Barton and Scholz, 1995; Crovelli and Barton, 1995) are geometric sets which can be derived from the fields (the topography, ore concentration fields in these examples) and will be outside our scope. Similarly, we will not discuss the literature on the scaling of rock fractures (e.g. Barton, 1995; Leary, 2003a) nor on rock fragment distributions (e.g. Turcotte, 1989; Kaminski and Jaupart, 1998); many examples can be found in Turcotte (1989). Finally, the burgeoning literature on scaling in seismology (starting with the famous Omori, 1895, and then Gutenberg and Richter, 1944, laws) generally treat earthquakes as sets of points and only considers the distribution of intensities their without reference to their locations (hence not as fields or measures) and are also outside our scope (see however Hooge et al., 1994). Earthquakes are also fertile ground for classical SOC type models which build upon the classical slider-block model (Burridge and Knopoff, 1967; see Carlson et al., 1994, for a review, and Weatherley and Abe, 2004, for a recent example).

This paper is organized as follows. In Sect. 2 we use spectral analysis and many examples, to argue that scaling in the solid earth and topography cover huge ranges of scale in both the horizontal and vertical directions. We argue that the scaling is systematically different in the two directions and that this is a symptom of the vertical lithospheric stratification. In Sect. 3, we show – again with examples – that the scaling of the intense and weak field regions is different, that the fields are typically multifractal not monofractal, and we examine some of the consequences for classical geostatistics. In Sect. 4 we consider the generic multifractal process

– cascades – concentrating on aspects which require particular clarification, especially the issues of the type of scale by scale conservation, singularity localization, divergence of moments, universality, the dimension and codimension multifractal formalisms. Finally in Sect. 5 we compare and contrast various scaling models and show how to take into account anisotropic scaling. In Sect. 6 we conclude.

2 Wide range scaling

2.1 Anisotropic scaling and vertical stratification of structures

Scale invariance – no matter how theoretically appealing – would not be of general geophysical interest were it not for the basic empirical fact that geofields display wide range scaling. Perhaps the most straightforward way to show this is by using spectral analysis which is both fairly familiar to geoscientists and has the advantage of being very sensitive to scale breaks. In addition, it can also be useful for studying anisotropy. Consider the geophysical field $I(\boldsymbol{r})$ where $\boldsymbol{r}$ is a position vector. We define the spectral density $P(\boldsymbol{k})$:

$$P(\boldsymbol{k}) = \langle |\tilde{I}(\boldsymbol{k})|^2 \rangle; \quad \tilde{I}(\boldsymbol{k}) = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}} I(\boldsymbol{r}) d\boldsymbol{r} \tag{1}$$

where $\boldsymbol{k}$ is a wavevector. Since P is quadratic in I, it is a second order statistic. In the definition, we have included the theoretically motivated ensemble average (denoted "<.>") although in fact, often P is estimated from a single realization using a fast Fourier algorithm on gridded (finite resolution) data; in this case, it is more properly called a "periodogram". We may then define the "isotropic" spectrum $E(k)$ obtained by angle integrating P:

$$E(k) = \int\limits_{|\boldsymbol{k}'|=k} P(\boldsymbol{k}') d\boldsymbol{k}' \tag{2}$$

where k is the modulus of the wavevector (the notation indicates angle integration in Fourier space). If the statistical properties of $I(\boldsymbol{r})$ are both isotropic and scaling then E is of the power law form:

$$E(k) \propto k^{-\beta} \tag{3}$$

where β is the "spectral exponent". Note that sometimes angle averaging (rather than integration) is performed; in 2-D, the corresponding exponent is $\beta-1$. The advantage of using the present (turbulence based) definition is that if the process is isotropic, then β is independent of the dimension of space so that 1-D sections will have the same exponent.

The exponents of isotropic spectra are invariant under the scale change $\boldsymbol{k}\rightarrow\lambda\boldsymbol{k}$ (corresponding in real space to the scale reduction $\boldsymbol{r}\rightarrow\lambda^{-1}\boldsymbol{r}$); the spectra – which keeps its form but which changes by the factor $\lambda^{-\beta}$ – is called "scaling". In physics the term "scaling" is generally reserved for invariance under scale transformations in space, time or space-time, although occasionally it is also used to describe the tails of algebraic probability distributions, (in this case it refers to scaling in a probability space; see the discussion of SOC below). In the geosciences there is an unfortunate tendency to use "scaling" to denote the general problem of changing from one scale to another even if there are no conserved properties; below we reserve the term for the more precise physics sense which implies some invariant properties under (possibly anisotropic) scale changes.

In the following we will be interested in the vertical stratification which – if scaling – will manifest itself in different horizontal and vertical spectral exponents although for simplicity, we will assume isotropy in the horizontal plane. In this case, we have $P(\boldsymbol{k})\, d\boldsymbol{k}=P(K,\theta,k_z)\, K dK d\theta dk_z=2\pi P(K,k_z)\, K dK dk_z$ where (K,θ) are the horizontal polar coordinates. We therefore have the following 1-D spectra:

$$\begin{aligned} E(K) &= 2\pi \int K P(K,k_z)\, dk_z; \\ E(k_z) &= 2\pi \int P(K,k_z)\, K dK; \\ K^2 &= k_x^2 + k_y^2 \end{aligned} \tag{4}$$

where $\boldsymbol{k}=(k_x,k_y,k_z)$ is a wavevector. In order to model the horizontal stratification – the fact that the 1-D spectra will typically be different in the horizontal and vertical directions – we take P to be of the general anisotropic scaling form:

$$\begin{aligned} &P(K,\theta,k_z) \propto \|(\boldsymbol{K},k_z)\|^{-s}; \\ &\|(\boldsymbol{K},k_z)\| = \left(\left(\frac{K}{k_s}\right)^2 + \left(\frac{k_z}{k_s}\right)^{2/H_z} \right)^{1/2} \end{aligned} \tag{5}$$

where $\boldsymbol{K}=(k_x,k_y)$ is a horizontal vector, and we have introduced the (Fourier) scale function $\|(\boldsymbol{K},k_z)\|$, the "sphero wavenumber" k_s at which Fourier structures are roughly spherical, the spectral exponent s, and the anistropic exponent H_z. From Eqs. (4), (5) we obtain the horizontal and vertical scaling exponents (β_h, β_v):

$$E(K) \approx K^{-\beta_h}; \; E(k_z) \approx k_z^{-\beta_v} \tag{6}$$

with exponents β_h, β_v satisfying:

$$\begin{aligned} \beta_h &\approx s - H_z - 1; \; s > H_z \\ \beta_v &\approx \tfrac{s-2}{H_z}; \qquad s > 2 \end{aligned} \tag{7}$$

i.e.:

$$H_z = (1-\beta_h)/(1-\beta_v); \; s = \beta_h + H_z + 1 \tag{8}$$

Equation (5) assumes horizontal isotropy and we have deliberately given the very simplest possible scale function which leads to 1-D horizontal/vertical spectra with different exponents (Eq. 6). In fact, in order for Eqs. (6), (7), (8) to follow

from Eq. (5, top), the scale function need only be a solution of the functional "scale" equation:

$$\|T_\lambda(\boldsymbol{K}, k_z)\| = \lambda^{-1}\|(\boldsymbol{K}, k_z)\|;\ T_\lambda = \lambda^{-\mathbf{G}^T} \quad (9)$$

where T_λ is an operator which reduces the scale of a vector by a ratio λ, and $\mathbf{G}$ is the (real space) generator of the anisotropy ("T" means "transpose", necessary for the Fourier generator). In this case:

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & H_z \end{pmatrix}$$

leading to stratification. In this section, we will only need this simple case with diagonal matrix $\mathbf{G}=\mathbf{G}^T$ corresponding to "self-affine" scale changes (when H_z=1, $\mathbf{G}$ is the identity and the system is invariant under isotropic scale changes, i.e. it is "self-similar"). However, the basic formula Eq. (9) ("Generalized Scale Invariance", GSI developed in the context of atmospheric turbulence (Schertzer and Lovejoy, 1985b)) is valid when $\mathbf{G}$ has off-diagonal elements (corresponding to differential rotation as well as squashing of structures) and for generalizations in which the anisotropy depends on location ("nonlinear" GSI, $\mathbf{G}$ is then a nonlinear operator); see Sect. 5. We may note that here it is the real space vertical cross-sections (i.e. the lines $I(x,z)$=*constant*) which are self-affine. This is quite different from the self-affinity of monofractal functions such as fractional Brownian Motion (fBm) which have self-affine *graphs*. For example, consider an fBm model of the ore concentration $c(x,y)$, i.e. the concentration in two dimensional (x, y) space. The graph of c – which is the surface defined in the $(x, y, c(x, y))$ space – has self-affine $(x, c(x, 0))$ sections (i.e. for constant y=0), yet the real-space iso-concentration lines defined by $c(x,y)$=*constant* will be self-similar.

The connection between the Fourier and real space structures can be established by using structure functions. Consider the q^{th} order "structure function" S_q defined by:

$$S_q(\boldsymbol{\Delta r}) = \left\langle |\Delta I(\boldsymbol{\Delta r})|^q \right\rangle;\ \Delta I(\boldsymbol{\Delta r}) = I(\boldsymbol{r}+\boldsymbol{\Delta r}) - I(\boldsymbol{r}) \quad (10)$$

where $\boldsymbol{\Delta r}$ is a displacement vector in (x,y,z) space: $\boldsymbol{\Delta r}$=($\boldsymbol{\Delta R}$, Δz) where $\boldsymbol{\Delta R}$=(Δx,Δy) is a horizontal displacement vector. Equation (10) assumes that I is statistically independent of $\boldsymbol{r}$. Note that the fluctuations can more generally be defined by wavelets; the ΔI in Eq. (10) is in fact a "Haar" or "poor man's" wavelet generally adequate for our purposes. If the field $I(\boldsymbol{r})$ is scaling, then:

$$S_q(\boldsymbol{\Delta r}) \propto ||\boldsymbol{\Delta r}||^{\xi q} \quad (11)$$

where $\xi(q)$ is the structure function exponent, and $||\boldsymbol{\Delta r}||$ is the scale function (the real space counterpart of $\|(\boldsymbol{K}, k_z)\|$ of Eq. (5):

$$\left\|\lambda^{-G}(\boldsymbol{\Delta R}, \Delta z)\right\| = \lambda^{-1}\|(\boldsymbol{\Delta R}, \Delta z)\| \quad (12)$$

(in spite of the notation, the real space and Fourier space scale functions are not the same, and the Fourier generators $\mathbf{G}^T$ is the transpose of the real space generator $\mathbf{G}$). If we consider horizontal and vertical displacements $\boldsymbol{\Delta R}$, Δz, respectively), then this reduces to:

$$\begin{array}{l} \left\langle |\Delta I(\boldsymbol{\Delta R})|^q \right\rangle \propto |\boldsymbol{\Delta R}|^{\xi_h(q)}; \\ \Delta I(\boldsymbol{\Delta R}) = I(\boldsymbol{r}+\boldsymbol{\Delta r}) - I(\boldsymbol{r}); \\ \boldsymbol{\Delta r} = (\boldsymbol{\Delta R}, 0) \end{array}$$
$$\begin{array}{l} \left\langle |\Delta I(\Delta z)|^q \right\rangle \propto |\Delta z|^{\xi_v(q)}; \\ \Delta I(\Delta z) = I(\boldsymbol{r}+\boldsymbol{\Delta r}) - I(\boldsymbol{r}); \\ \boldsymbol{\Delta r} = (0, 0, \Delta z) \end{array} \quad (13)$$

where:

$$\begin{array}{l} \xi_h(q) = \xi(q) \\ \xi_v(q) = H_z\xi(q) \end{array} \quad (14)$$

H_z is thus the ratio of the horizontal and vertical structure function exponents. If we take q=2, this reduces to the usual structure function exponent (applied to a single realization, i.e. without ensemble averaging, this is termed the "semi-variogram"). Due to the Wiener-Khintchin theorem (the spectrum of a homogeneous process is the Fourier transform of the autocorrelation function), we then obtain:

$$\begin{array}{l} \xi(2) = s - 2 - H_z \\ \beta_h = \xi_h(2) + 1 \\ \beta_v = \xi_v(2) + 1 \end{array} \quad (15)$$

We therefore see that the ratio $(\beta_h-1)/(\beta_v-1)$ is the ratio of the variances in the horizontal and vertical directions so that if we define the extent of typical structures by their variances, then a structure of horizontal extent $|\boldsymbol{\Delta R}|$ has a corresponding vertical extent: $\Delta z = l_s\left(\frac{|\boldsymbol{\Delta R}|}{l_s}\right)^{H_z}$ where l_s is the real space counterpart of k_s; it is the scale at which the real space structures are roughly "roundish": $\Delta z \approx |\boldsymbol{\Delta R}| = l_s$; Fig. 11 shows a vertical cross-section of the magnetic susceptibility (see below) showing how structures start out very flat/stratified at small scales becoming less and less stratified at larger scales.

2.2 Horizontal structures

We now attempt to demonstrate that many geofields, including some of the most important such as the topography and rock density – are scaling over wide ranges of scale. We first consider the topography which is of particular importance since it is not only relatively well measured, but is a fundamental geophysical field. Starting with Vennig-Meinesz (1951), many spectral analyses of the earth's topography have been made (Balmino et al., 1973; Bell, 1975; Berkson and Matthews, 1983; Fox and Hayes, 1985; Gibert and Courtillot, 1987; Balmino, 1993; Mareschal, 1989; Lavallée et al., 1993; Tchiguirinskaia et al., 2000; Gagnon et al., 2003) all finding approximately power law (isotropic) spectra (also

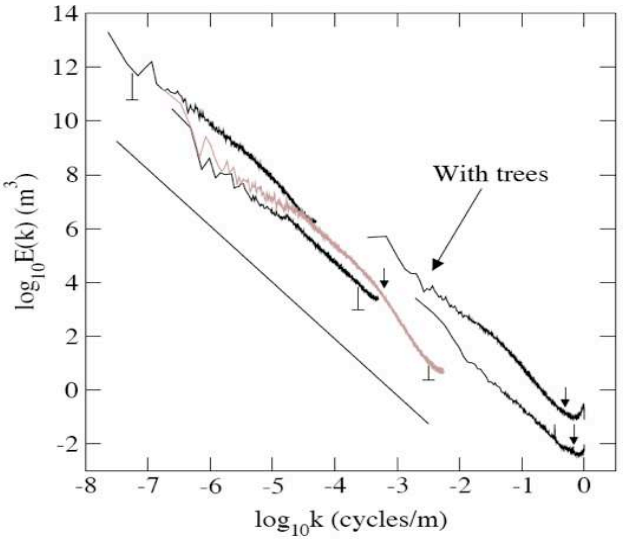

Fig. 1. Log/log plot of the spectra for four Digital Elevation Maps (DEMs). From right to left: Lower Saxony (with trees, top), Lower Saxony (a sub-region without trees, bottom), the U.S. at 90 m (in grey), at 30″ (about 1 km, GTOPO30) and the earth (including bathymetry) at 5′ (about 10 km), ETOPO5. A reference line of slope −2.10 is on the graph to show the overall slope of the spectra. The small arrows show the frequency at which the spectra are not well estimated due to the inadequate dynamical range of the data; see Gagnon et al. (2006) for this theoretical estimate (for ETOPO5, it is well estimated over the whole range). The "semi error bar" symbols indicate the amount of offset due to the resolution dependent factor $\lambda^{K(2)}$ (see Gagnon et al., 2006) for this necessary resolution dependent correction). Reproduced from Gagnon et al. (2006).

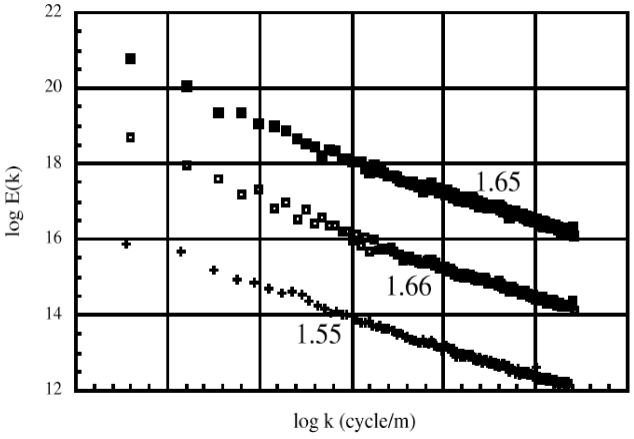

Fig. 2a. Spectra from three images near the Puu Oo volcanic vent. Spectra from the visible.

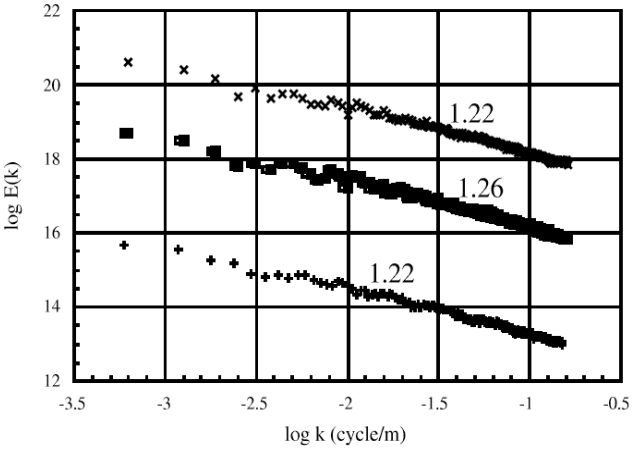

Fig. 2b. Same but spectra from thermal infrared images. All spectra were shifted vertically for clarity. Corresponding values of β are indicated. Reproduced from Harvey et al. (2002).

relevant are similar results on Venus topography (Kucinkas et al., 1992)). Although the exponents are somewhat variable from region to region (see the discussion in Sect. 3.4), the values of β are not so different; the overall conclusion of Gagnon et al. (2006) is that it varies from about 1.6 for oceans to 2.1 for continents. Figure 1 shows a recent spectral analysis of topography covering the range 1 m to 20 000 km, showing the excellent scaling over at least planetary scales down to about 40 m where vegetation starts becoming an issue.

Others surface fields – especially from remote sensing have also been shown to have wide range spectral scaling. Some – such as the reflected visible radiances and thermal infra red emissions from volcanoes over the range of roughly 50 cm to 2 km (Fig. 2) (Harvey et al., 2002; Laferrière and Gaonac'h, 1999) have implications for the subsurface, while others primarily reflect soil, vegetation and other surface characteristics. Using remote sensing many surface fields and their surrogates have been shown to exhibit wide range scaling, for example soil moisture (Dubayah et al., 1997) and humidity indices (Lovejoy et al., 2007b[1]).

The horizontal scaling of the topography and other surface fields is significant because the geophysical processes responsible for them (including orographic, erosional, hydrological etc.) are strongly nonlinearly coupled so that the scaling in one is strong evidence for scaling in another. We can be fairly confident of this because scale invariance is a symmetry principle and one generally assumes that symmetries are respected unless specific symmetry breaking mechanisms can be found. Another way of viewing the same argument is to consider a dynamical process which generates structures over a wide range of scales and then to decompose it into a finite number of different scaling regimes each valid over various sub ranges. The principle of parsimony demands that

[1]Lovejoy, S., Tarquis, A., Gaonac'h, H., and Schertzer, D.: Single and multiscale remote sensing techniques, multifractals and MODIS derived vegetation and soil moisture, Vadose Zone J., submitted, 2007b.

Table 1. A comparison of estimates of the stratification exponent H_z from horizontal and vertical spectra. Since $H_z=(\beta_h-1)/(\beta_v-1)$, it can be very sensitive to small errors in the the β's. Although theoretically, the simplest model involves the same degree of stratification for all the fields, this is not strictly necessary. From these limited data, we conclude that H_z is likely to be in the range 1.5 to 3.

Quantity	β_h	β_v	H_z
Mean of rock density, gamma emission, seismic velocity (Leary, 1997)	1.34±0.12	1.10±0.12	3.4
Carbonate rock density (Tubman and Crane, 1995) Figs. 7a, b	0.86	0.78	1.57
Susceptibility (Figs. 4a, b, KTB, Fig. 9)	1.32	1.22 (1.2, KTB)	1.45
Susceptibility inferred from regional magnetic anomalies, (two regions)	0.6, 1.4	0.8, 1.2	2
Rock Density inferred from high wavenumber surface gravity	1.3	1.1	3
Hydraulic conductivity (Tchiguirinskaia, 2002)	1.66	1.3	2.22

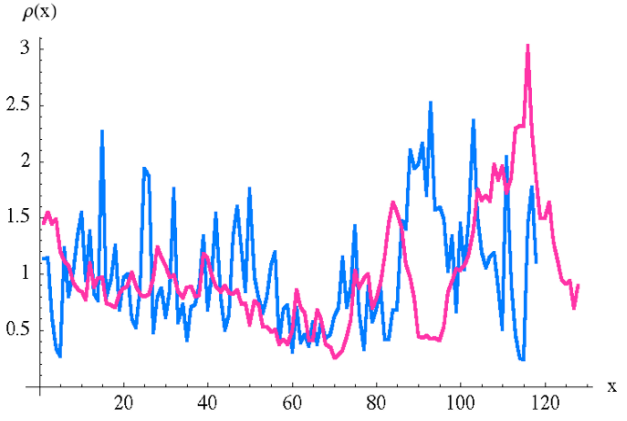

Fig. 3a. de Wijs Zinc concentration data from the Pulacayo mine, Bolivia, with x the horizontal distance in units of 2 m data (blue) simulation (pink with parameters α=1.8, C_1=0.03, H=0.090), both normalized to unity (the mean concentration is 15.6%).

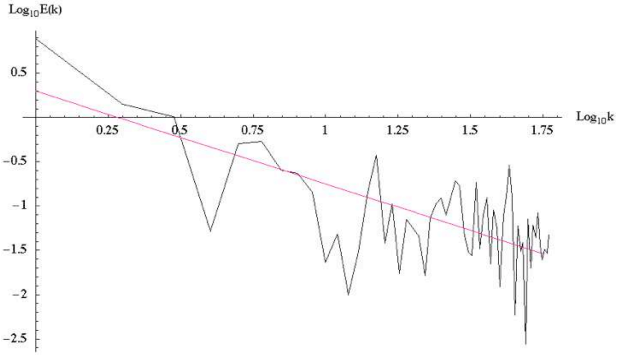

Fig. 3b. de Wijs spectrum: Red line is theory: $1-K(2)+2H$ with $K(2)$=0.05 (trace moments), H=0.090 (first order structure functions), hence $\beta_h\approx$1.12.

we start with the assumption of a single regime and then only add new additional regimes when absolutely necessary.

Nevertheless, it is still important to directly verify the scaling on as many geophysically significant fields as possible; we discuss in particular the rock density, magnetic susceptibility, ore concentrations. Unfortunately, these generally require in situ measurements so that the corresponding horizontal fields are only known over sparse (possibly fractal (Lovejoy et al., 1986)) sets of sample locations. In principle this demands special multifractal interpolation techniques (Salvadori et al., 2001), but an operational method is still lacking. Cheng et al. (1994) has proposed a partial solution to this sparse measurements problem; the "Integrated spatial-spectrum Analysis" method. The first step is to use traditional Kriging methods to obtain a 2-D field on a uniform grid. If the data are not too sparse (essentially they must be 2-D but with perhaps uniformly distributed "holes"), this may be adequate. The Kriging is followed this by spectral analysis. However instead of plotting the spectral density as a function of the modulus of the wavenumber (after integrating in circles in Fourier space, see Eq. 2 above), the (Fourier space) log areas exceeding a log spectral density is plotted. If the process is isotropic in 2-D space, with spectral exponent β, then the result will be linear but with slope $-1/(\beta-1)$ (the reciprocal because of the interchange of the ordinate and abscissa; the -1 because of the cumulation of all the values below a spectral density threshold). The method has the usual advantage that integrating smoothes the statistics, but has the added attraction of being insensitive to anisotropies (as long as the latter are scaling; it doesn't involve integration over circles). Finally the method can be used to design new kinds of anisotropic filters useful for prospecting.

As a first example of an in situ horizontal analysis, we consider the famous de Wijs (1951) zinc concentration series which has been discussed in the literature many times and also reanalyzed many times including via spectral techniques (Agterberg, 1974; although not on a log-log plot to test the scaling); we return to this series in Sect. 4.10 structure functions and other analyses). In Fig. 3a we show the original de Wijs series of Zn concentrations and in Fig. 3b the corresponding spectrum with $\beta_h\approx$1.12. Although the number of points is very small, the general power law form of the spectrum is visible. A somewhat higher quality horizontal

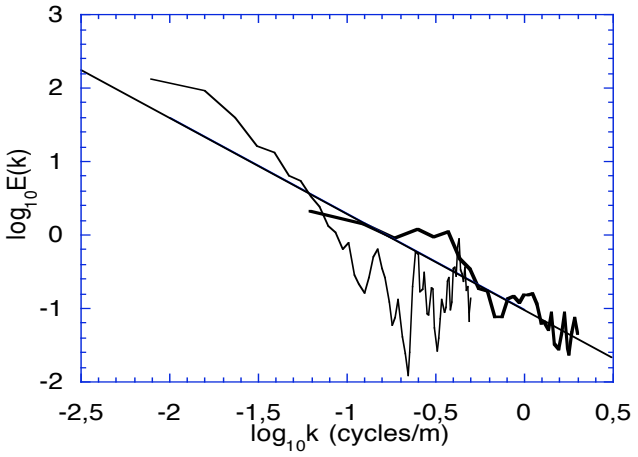

Fig. 4a. Magnetic susceptibility spectra in the horizontal: Power spectra for two sets of magnetic susceptibilities in the horizontal obtained by Pilkington and Todoeschuck (1993, 1995). The straight line shows the theoretical slope β_h=1.32.

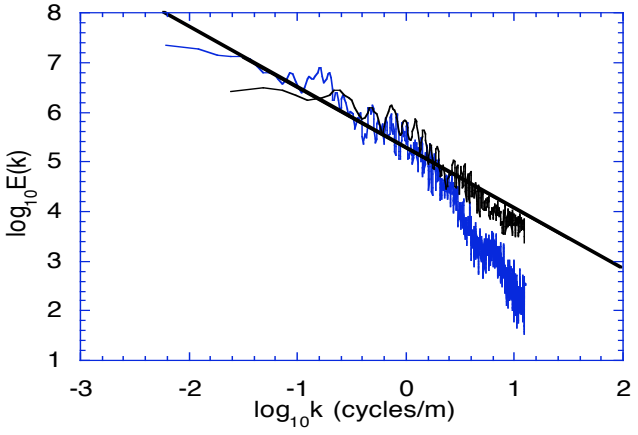

Fig. 4b. Magnetic susceptibility power spectra from vertical borehole logs in sedimentary (top) and igneous (bottom) rock from the same region as Fig. 4a (Pilkington and Todoeschuck, 1995). The straight line has the slope of 1.22. As discussed in Lovejoy et al. (2001) a values of $\beta_h \approx 1.4$ and $\beta_v \approx 1.2$ gives a good explanation for the observed surface gravity anomalies in the same (Canadian shield) region. The high wavenumber fall-off for the igneous series is probably due to slight oversampling.

spectrum is shown in Fig. 4a (Pilkington and Todoeschuck, 1995) which was obtained after Hankel transforming the radial autocorrelation function from a sample of several thousand in situ susceptibility measurements. Due to the inadequate sampling, the spectrum is not perfectly scaling, but coupled with a corresponding vertical (borehole) spectrum Fig. 4b, it turns out to be roughly what is required to explain magnetic surface anomaly spectra discussed in Sect. 2.4 below. Perhaps the most convincing of the horizontal in situ spectra are the 1-D "horizontal borehole" spectra of Leary

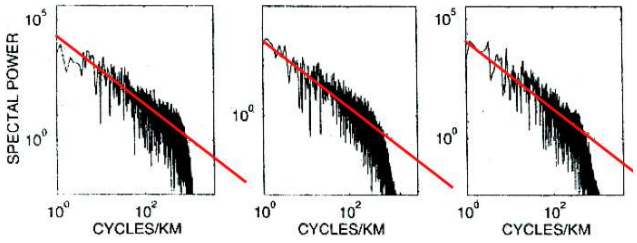

Fig. 5a. Horizontal borehole species: left to right gamma emission, rock density and seismic velocity absolute reference slopes = β_h=1.4, adapted from Leary (1997).

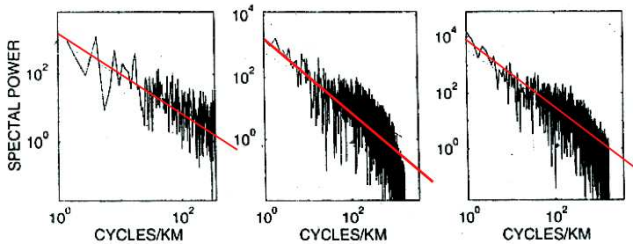

Fig. 5b. Vertical borehole analyses for the same quantities and from the same region as Fig. 5a, the absolute reference slopes have β_v=1.2. Adapted from Leary (1997).

(1997) (Fig. 5a), for gamma emission, rock density and seismic velocity over the range of about 10 m to 1 km. Other examples of horizontal analyses of in situ fields are hydraulic conductivity (see Fig. 6a), (Tchiguirinskaia, 2002) and carbonate concentration (see Figs. 7a and 8a) (Tubman and Crane, 1995). These figures provide some of the rare examples where both horizontal and vertical exponents from essentially the same regions have been analyzed; we could also mention the horizontal and vertical spectra in Shiomi et al. (1997). In Table 1, we summarize some of these results and we return to their implications for the stratification in Sect. 2.4.

2.3 Vertical scaling

We started out our survey of evidence for wide range scaling in the solid earth by considering the horizontal direction; with the exception of the topography and remotely sensed radiances, surprisingly little is known about the horizontal scaling due to the difficulty in obtaining the necessary large quantities of in situ data. Although the geopotential fields (geomagnetism, geogravity) are relatively well measured (at least in certain regions) and do give us information about the horizontal structure, they also depend on the vertical structure and for their interpretation require anisotropic scaling models of rock susceptibility and density respectively, see Sect. 2.4.

We now turn our attention to evidence for scaling in the vertical. It is perhaps surprising that for many geophysical

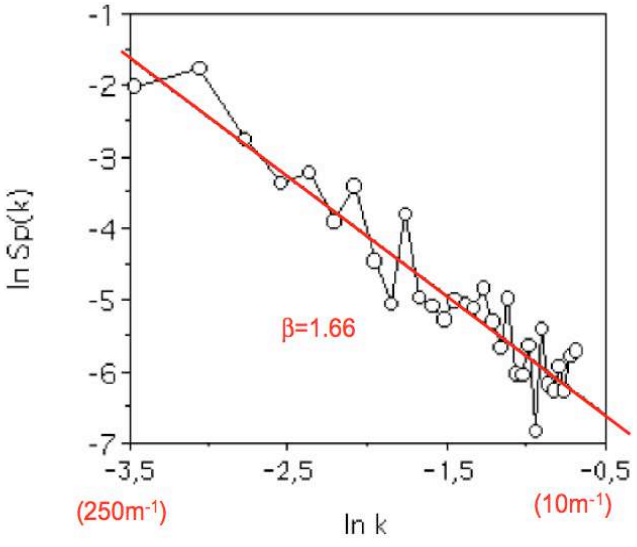

Fig. 6a. Ensemble power spectra (25 samples, from the MADE site, Tennessee); horizontal measurements, a straight line indicates β_h=1.66, units are such that the lowest wavenumber is about 250 m, highest about 10 m.

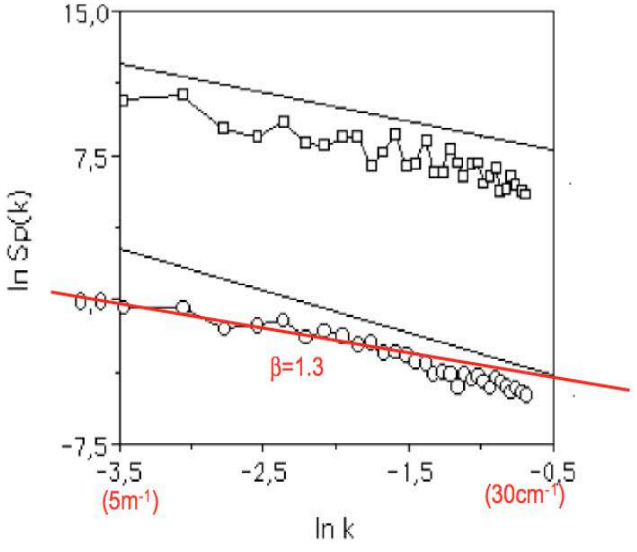

Fig. 6b. Same, but vertical measurements, straight black lines indicate β_v=2.2 for hydraulic conductivity (bottom points) and β_v=1.5 for the logarithm conductivity data (top points), however the red line shows that the lower value β_v=1.3 (corresponding to H_z=0.66/0.3=2.22) is a better fit for all except the highest wavenumbers. Units are such that the lowest wavenumber is about 5 m, highest about 30 cm. Adapted from Tchiguirinskaia (2002).

parameters the vertical structure is better known than the horizontal due to the large number of borehole analyses. Examples of scaling spectra from boreholes (gamma emission, rock density, magnetic susceptibility, sonic velocity, porosity, electrical resistivity) are Pilkington and Todoeschuck (1990),

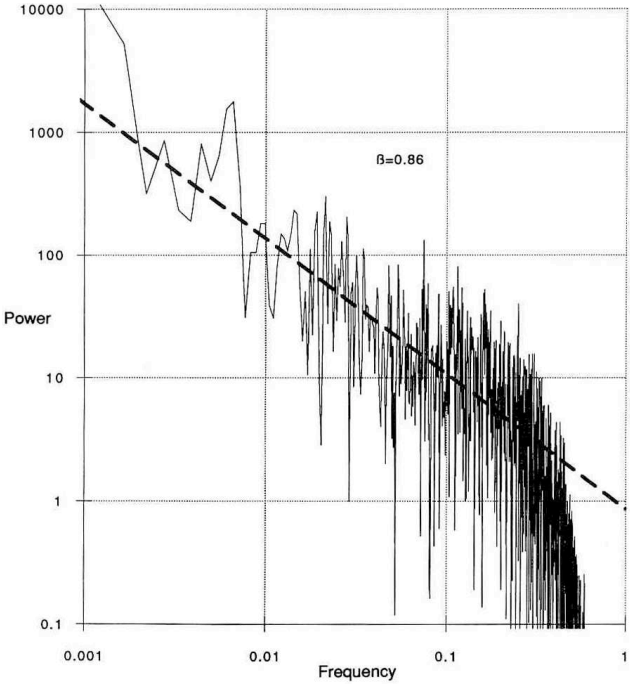

Fig. 7a. Horizontal power spectrum of the density of carbonate rock well the last factors of 2 high frequency are a bit too smooth due to limitations of the data (no units given in the original). Reproduced from Tubman and Crane (1995), $\beta_h \approx 0.86$.

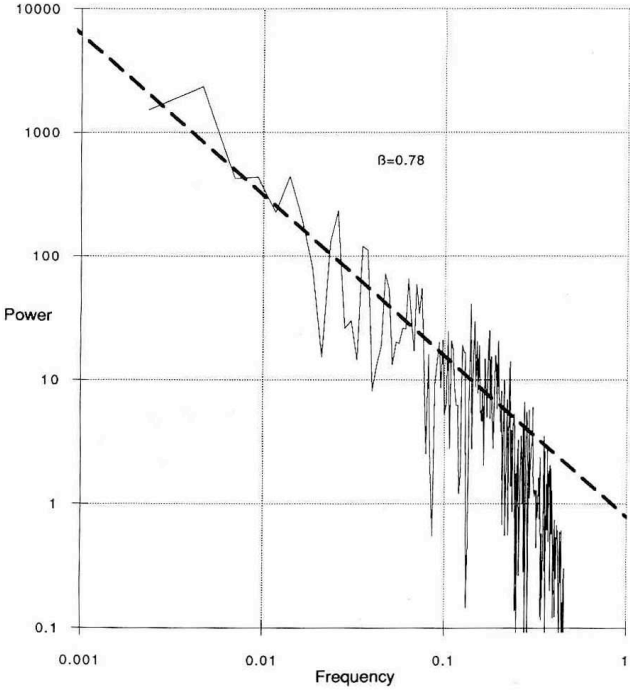

Fig. 7b. Same as (a) except for vertical spectrum. Reproduced from Tubman and Crane (1995), $\beta_v \approx 0.78$. Together with (a), this implies H_z=1.57.

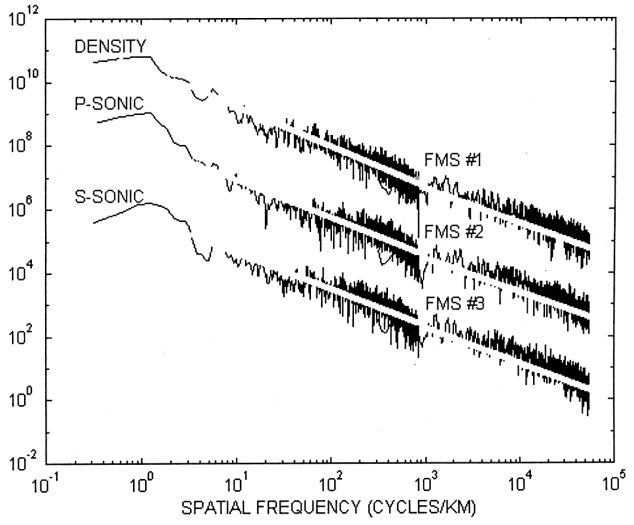

Fig. 8. Power law scaling of 3 well log and 3 small-scale resistivity (FMS) power spectra over 5 decades of spatial frequency length (0.3 cycles/km to 60 cycles/km). Power law scaling exponents β_v are 1.06 (P-wave sonic), 1.16 (S-wave sonic), 1.26 (density), 1.08 (FMS 1), 1.06 (FMS 2), 1.12 (FMS 3). Reproduced from Leary (2003b).

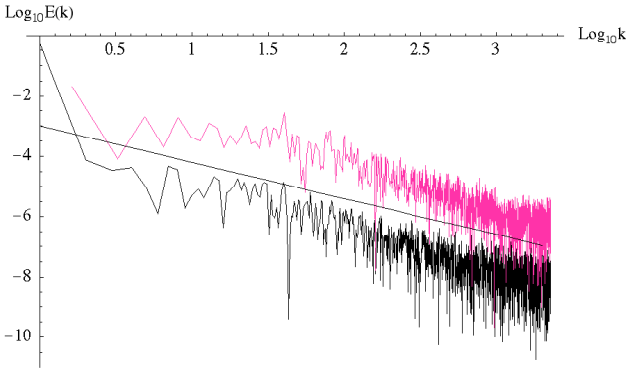

Fig. 9. Power spectrum of the KTB susceptibility (top) and density (bottom) over the top 5596 m and 9098 m depths, respectively (2m resolution; wavenumber k in units of $(2\,\mathrm{m})^{-1}$). The reference slope has β=1.2 (authors' analyses).

Todoeschuck et al. (1990), Todoeschuck and Jensen (1991), Bean and McCloskey (1993), Molz and Boman (1993); Molz and Liu (1997), Wu et al. (1994), Hollinger (1996), Leary (1997), Dolan et al. (1998), Leonardi and Kümpel (1999), Tchiguirinskaia (2002), Leary (2003a), Marsan and Bean (1999, 2003), Dimri (2005). Other parameters such as thermal conductivities (Dimri and Vedanti, 2005) have also been shown to be scaling using other analysis techniques. Figures 4b, 5b, 7b show some of the rare cases where both vertical and horizontal statistics can be compared allowing us to deduce the stratification exponent H_z (Eq. 8). Figures 8, 9, 10 are shown because they are particularly striking examples: Fig. 8 is a composite, but the spectra collectively cover a range of scales from centimeters to several kilometers, and Figs. 9, 10 show spectra of various parameters from the deep (KTB) borehole.

Two aspects of these analyses are particularly worth mentioning. The first – widely recognized – is the proximity of many of the β values to 1, hence the term "1/f noise". This term originates in the ubiquitous noise in electrical circuits (due for example to contacts) with similar spectra. Indeed, Leary (1997) has argued that the β's of sonic velocities, rock densities, H_2 density (porosity), gamma activity and resistivity, porosity and permeability are all approximately unity (although with fluctuations of order 0.2–0.4) and he has argued that this could best be understood from a phase transition type mechanism such as percolation (for an introduction, see Stauffer, 1985, for applications to rock conductivity see Bahr, 2005, and references therein). The obvious problem with this as a general explanation is that in phase transitions, unless one happens to be exactly at the critical point there will only be scaling over a finite range of scales with a drastic breakdown for larger scales. In addition, this critical scale diverges at the critical point so that one would expect to see scale breaks whose value depends sensitively on some physical parameter such as porosity (indeed the possibility of such sensitive dependence of magma strength on porosity due to percolation of bubbles has been suggested as a mechanism for volcanic eruptions (Gaonac'h et al., 2003, 2007[2]). Although there are large fluctuations these are expected in scale invariant systems and the well logs all demonstrate wide range scaling with no obvious systematic or strong breaks. From the perspective of scale invariant dynamics (multifractal cascades, see below), the scaling can be roughly explained as follows: the scaling is due to the absence of a strong scale breaking mechanism and the value β close to one due to the fact that the observed processes are close to "conserved" multifractal processes which generically give spectra with β a little below one (if not too intermittent). The empirical values slightly above 1 are due to small degree of non scale by scale conservation; a parameter $H>0$, see below.

The second important point about the vertical βs (as mentionned by Leary, 1997) is that the vertical and horizontal exponents are somewhat different. Indeed, the vertical exponents are systematically a bit closer to 1. As pointed out by Schertzer and Lovejoy (1985a) in the context of atmospheric stratification (and Lavallée et al., 1993) in the context of topography); if the horizontal and vertical scalings are different, then the corresponding structures will exhibit differential stratification; the key quantity is the ratio H_z (Eq. 8) which we saw is the ratio of the horizontal to vertical structure function/variogram exponents. From Figs. 4–8 we see that the stratification exponent $H_z \approx 1.7$–2 for magnetic susceptibility

[2]Gaonac'h, H., Lovejoy, S., Nunes-Carrier, M., Schertzer, D., and Lepine, F.: Percolating magmas in three dimensions, Nonlin. Processes Geophys., in review, 2007.

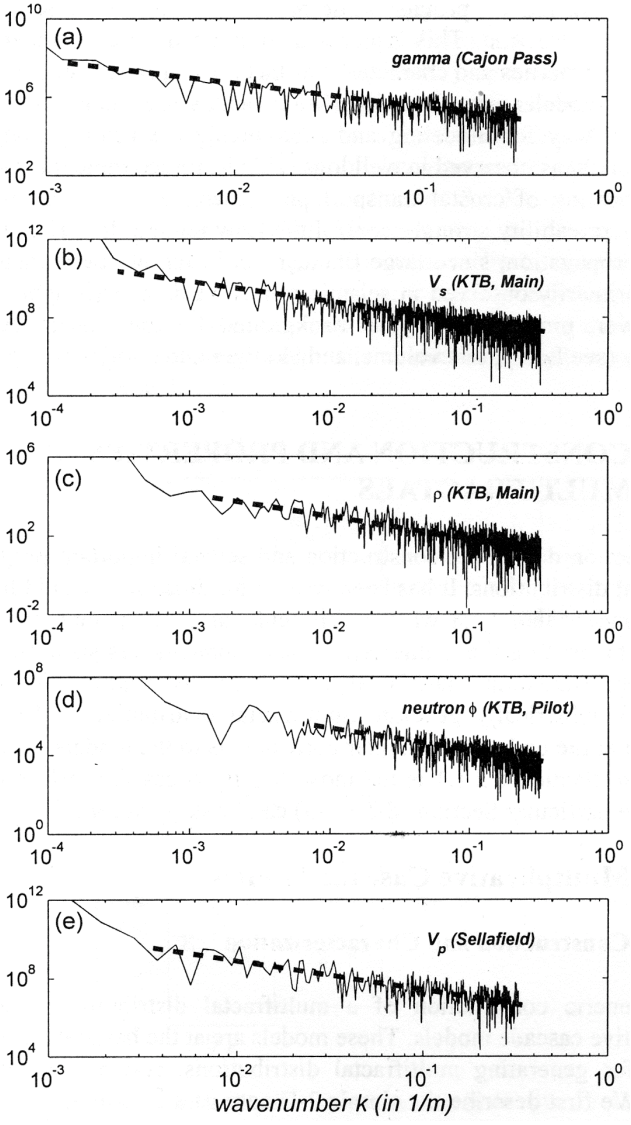

Fig. 10. Power spectra of five logs from various boreholes, from top to bottom: **(a)** gamma log from the Cajon Pass borehole; **(b)** S-wave sonic log from the KTB main borehole, Germany; **(c)** resisitivity log from the KTB main borehole; **(d)** neutron porosity log from the KTB borehole; **(e)** P-wave sonic log from the Nirex 1 borehole at Sellafield UK. The dashed black lines give power law fits $k^{-\beta_v}$ of the spectrum decay, with spectral exponent β_v equal to (a) 1.22, (b) 0.98, (c) 1.31, (d) 1.37, (e) 1.4. All four boreholes probe the crystalline part of the upper crust. Reproduced from Marsan and Bean (2003).

(Lovejoy et al., 2001) and $H_z \approx 3$ for rock density; (see Table 1). A value $H_z > 1$ means that while the rock strata are very thin (highly flattened structures in vertical sections), that they nonetheless become progressively rounder at larger and larger scales (see Figs. 11, 14). In these examples, the sphero-scale is typically found to be quite large, thousands of kilometers. This is the opposite of the atmosphere where

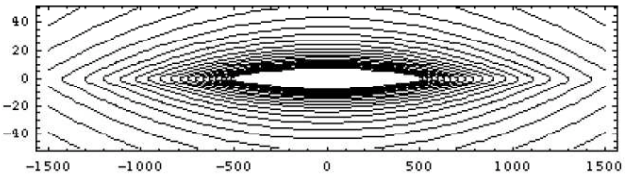

Fig. 11. Vertical cross-section of the magnetization scale function assuming H_z=2 and a spheroscale of $(40\,000\,\text{km})^{-1}$. The scale is in kilometers and the aspect ratio is 1/4. Reproduced from Lovejoy et al. (2001).

the value $H_z \approx 5/9$ is found both theoretically and empirically (see e.g. Schertzer and Lovejoy 1985a; Lilley et al., 2004) and the sphero scale is typically <1 m so that atmospheric structures become more and more stratified at larger and larger scales.

2.4 Combining Horizontal and vertical statistics: geopotential fields

Although over huge ranges of scale the processes which produce variations in the lithospheric properties are undoubtedly highly nonlinear, some are sources for geopotential fields (notably geogravity, geomagetism) and are related to them by purely linear relations (Poisson's equation, Maxwell's equations). Indeed, the relations are particularly convenient to deal with in Fourier space so that we can obtain very simple relations between the magnetic susceptibility spectrum P_M, and the spectrum of the surface magnetic field P_B or between the rock density and geogravity spectra P_ρ and P_g.

The example of magnetism and susceptibility has been studied in particular detail in Lovejoy et al. (2001) and Pecknold et al. (2001). With various reasonable assumptions (that there is a scalar magnetic potential, that over the limited region of the study that the magnetic anomaly (B) and susceptibility (M) have roughly constant directions so that only their magnitudes are variable), one obtains (see e.g. Blakely, 1995):

$$P_B(K) = \int_{k_c}^{\infty} \frac{K^2}{K^2 + k_z^2} P_M(K, k_z)\, dk_z \tag{16}$$

The integration in the above is over all wavenumbers higher than the Curie wavenumber ($k_c \approx 2\pi/z_c$ where z_c is the Curie depth at which all magnetization ceases due to high temperatures; $z_c \approx 30$–80 km). In order to model the horizontal stratification, we take P_M to be of the general anisotropic scaling form (Eq. 5). Using the susceptibility as a surrogate for the magnitude of the magnetization M, from the data in Fig. 4a, b (see also Fig. 9), the values $\beta_M \approx 1.2$, $\beta_{hv} \approx 1.4$ can be used to determine $s \approx 4.4$, $H_z \approx 2$ (Eq. 8) (Lovejoy et al., 2001).

Using Eq. (5) for an anisotropic scaling M field in Eq. (16) for P_B we can see that due to both the Curie depth and the sphero-scale, there will be breaks in the horizontal B

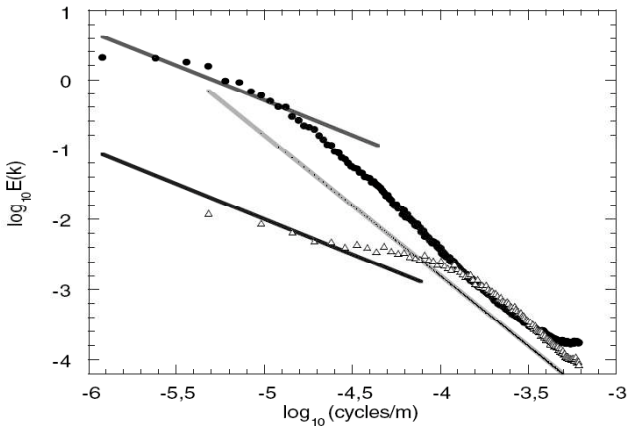

Fig. 12. Power spectra of aeromagnetic anomaly fields from two regional studies over the Canadian shield (triangles and circles). Superposed reference are lines with the theoretical high and low-wave number slopes, β_{Bh}=2, β_{Bi}=1 (see Eq. 17). The spectra have been normalized so that the high wave number regions roughly coincide. Reproduced from Lovejoy et al. (2001).

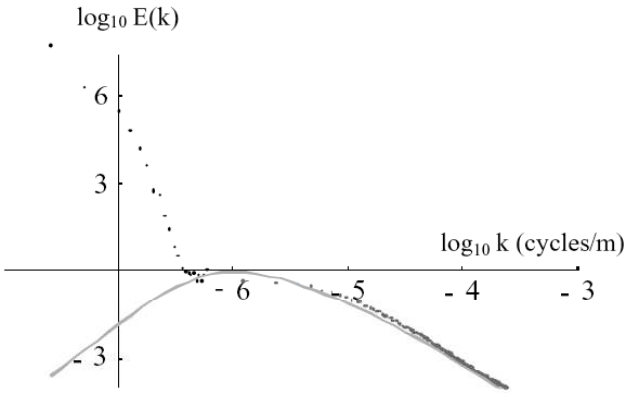

Fig. 13. Theoretical and experimental power spectra of surface magnetic fields. The high wavenumber points are from data set 2 (circles) of Fig. 12, the high wavenumber points are from the global Magsat determined spherical harmonics (n=1 taken as $(40\,000\,\mathrm{km})^{-1}$, from Langel and Estes, 1982). Reproduced from Lovejoy et al. (2001).

spectrum; Fig. 12 shows that this is indeed the case on regional B anomaly fields. The various relevant regimes are:

$$\begin{aligned} &\beta_{Bh} = s-2;\ K > k_c \\ &\beta_{Bi} = s-3;\ k_c > K > K_{ic} \\ &\beta_{Bl} = -3;\ K < K_{ic} \end{aligned} \tag{17}$$

where K_{ic}=$k_s \left(\frac{k_c}{k_s}\right)^{1/H_z}$ is the horizontal wavenumber corresponding to the vertical Curie wavenumber k_c, and the β_B's are the horizontal spectral exponents of the anomaly surface magnetic field and β_{Bh}, β_{Bi}, β_{Bl} are the high. intermediate and low wavenumber spectral exponents. Since in the region studied it was found that $k_s \approx 10^{-5}\,\mathrm{km}^{-1}$, $k_c \approx (30\,\mathrm{km})^{-1}$ and $H_z \approx 2$, this implies $K_{ic} \approx (1000\,\mathrm{km})^{-1}$ so that the low wavenumber regime is masked by the contribution from the core, hence we only expect to see the β_{Bh}, β_{Bi} regimes with a break near k_c. Figure 13 shows that with $s \approx 4.4$, we can explain both: the same wide range but anisotropic scaling can explain the large scale earth magnetic anomalies up to several thousand kilometers (at larger scales it is dominated by the main dynamo component form the liquid core). Figure 14 shows how stratified multifractal simulations (using the empirically determined universal multifractal parameters; see Sect. 5 below) can be used to simulate the magnetization, and Fig. 15 shows the corresponding B fields. See also Tennekoon et al. (2005) for scaling analyses of geomagnetic fields and Fedi (2003) for multifractal analysis of borehole susceptibilities.

Essentially the same type of relations hold between the vertical component of the surface gravity field (g) and the density of the rock (ρ):

$$P_g(K) = \int \frac{P_\rho(K, k_z)}{K^2 + k_z^2} dk_z \tag{18}$$

Maus and Dimri (1995, 1996) used this relation but with an isotropic (unstratified) P_ρ in order to model high wavenumber surface gravity fields, Bourlon et al. (1998) proposed using anisotropic scaling. See also Bansal and Dimri (2005) which includes scaling analyses of the horizontal anisotropy of gravity anomalies. As in the case of the susceptibility/magnetic anomaly relation, there are complications in the vertical so that there appear to be three regimes in the surface gravity field; essentially they are due to a) the mantle (low wavenumbers), b) the variable lithospheric thickness coupled with the strong mantle/lithosphere density gradient (intermediate range), c) the high wavenumber regime dominated by vertical and horizontal lithospheric heterogeneities (scales smaller than a hundred kilometers or so). While a detailed analysis of these contributions to the integral (Eq. 18) is in a forthcoming paper (Lovejoy et al., 2007a), the parameters s=5.3, H_z=3 are roughly compatible with the high wavenumber regime and the horizontal and vertical density spectra published in Shiomi et al. (1997) and Leary (1997) (see Figs. 5, 9). Also, the mantle regime has been briefly discussed in Lovejoy et al. (2005) and on the basis of an analysis of the equations of mantle convection, the parameters s=3, H_z=3 were proposed.

Before leaving the topic of geopotentials, we could mention recent work by Bahr (2005) to link rock conductivity with electric fields. A final example of scaling geofields linked by linear equations is the thermal properties of the earth's crust: the thermal conduction coefficient, the temperature and the distribution of radioactive sources of heat, see Dimri and Vedanti (2005).

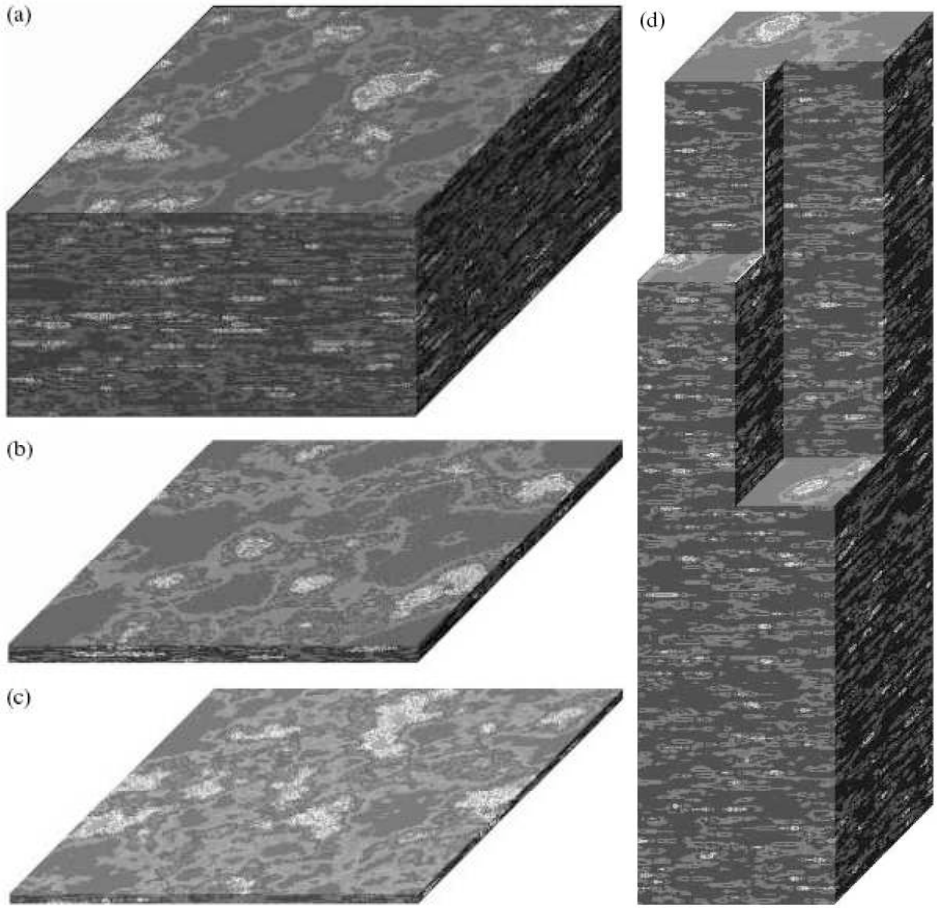

Fig. 14. (a) Simulated M field for horizontally isotropic crustal magnetization. The vertical anisotropy has H_z=1.7, s=4 and the universal multifractal parameters H=0.2, C_1=0.08, α=1.98. The sphero-scale was taken to be only ≈2500 km; the simulation region is 32×32×16 km with resolution 250 m. This is a reasonably realistic crustal section, although the sphero scale was taken to be a bit too small in order that strata may be easily visible. The direction of M is assumed to be fixed in the z direction. **(b)** Simulated M for horizontally isotropic crustal magnetization; same parameters as (a). The simulation is 128×128×32 km; the resolution is 1 km and only the portion above the Curie depth of 10 km is shown. **(c)** Simulated M field for horizontally isotropic crustal magnetization; same parameters as (a). The simulation is 512×512×16 km; the resolution is 4 km. **(d)** Simulated M field the simulation is 4×4×16 km, resolution is 62.5 m. The cut-out shows the stratification and the presence of anomalies at all depths. Reproduced from Pecknold et al. (2001).

3 From fractal sets to multifractal fields, the limitations of classical geostatistics

3.1 Box counting, functional box counting

Using Fourier spectra, we have seen that many solid earth fields display wide range scaling in both horizontal and vertical directions. Spectra were first widely used to characterize turbulence, and in the early 1970s in conjunction with the development of quasi-gaussian statistical closure models, the theoretical or empirical determination of the spectral exponent became a key task. During the same period, Mandelbrot (1977) proposed using fractal geometry with its appealing promise of simplifying the description and modeling of geoprocesses; in topography and geomorphology by quantifying complexity (roughness, sparseness) by a unique fractal dimension. Many of the seductive early fractal simulations were of precisely quasi-gaussian processes (essentially the "fractional Brownian motions" generalizations of Brownian motion), where there is a single basic exponent so that β is simply related to the (unique) fractal dimension D of exceedance sets (the set of points exceeding a fixed threshold) by the simple formula: $D=(7-\beta)/2$ (valid for the fractal dimension of monofractal surfaces with $1<\beta<3$). Indeed a great many papers were published which simply assumed that fields were monofractal and estimated the supposedly unique D from β (for many examples of this, see Scholz and Mandelbrot, 1989; Turcotte, 1989; Takayasu, 1990; Korvin, 1992).

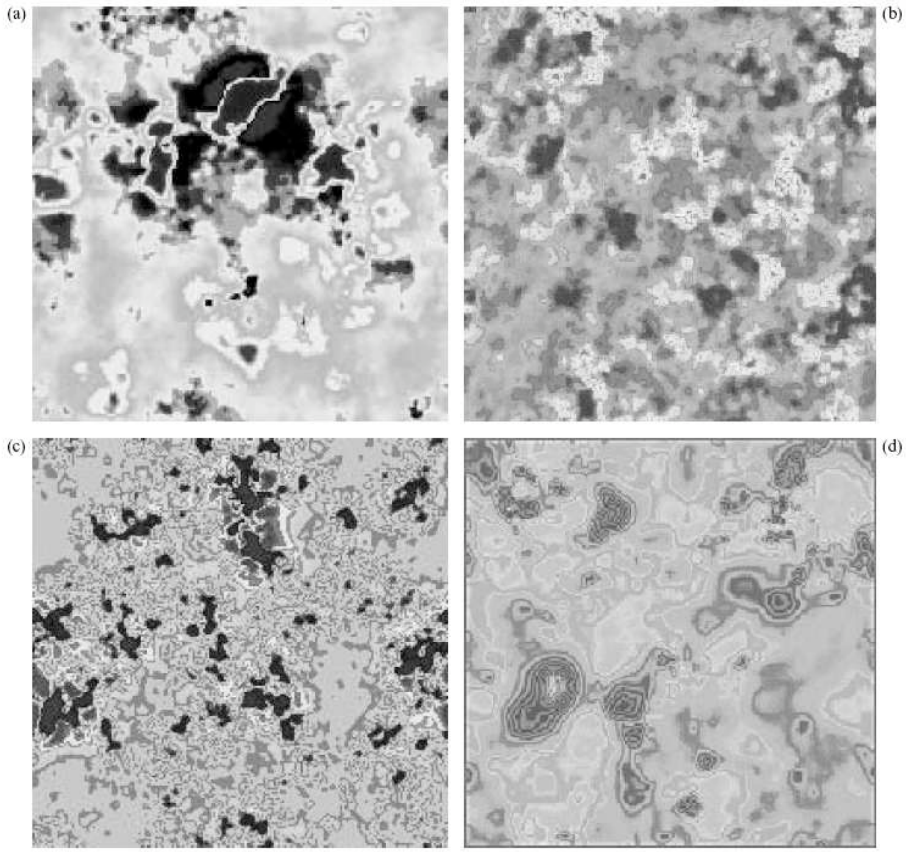

Fig. 15. (a) The surface B field from simulations shown in Fig. 14a. The Curie depth=16 km so that nearly the entire field shown is in the smooth, high wave number regime β_h=2. **(b)** The surface B field corresponding to Fig. 14b. Since the entire region simulated is 128 km across and the Curie depth is 10 km, the transition from high to intermediate wave number regime is in the middle of the range shown; the high wave number structures are noticeably smoother than the lower ones. **(c)** The B field corresponding to Fig. 14c; the entire simulation represents a region 512 km across, the Curie depth is 16 km so that most of the field shown with the exception of the very highest wave number structures is in the (rough) intermediate wave number regime with β_i=1. **(d)** The same but B for Fig. 14d, the entire field is in the smooth high wave number regime. Reproduced from Pecknold et al. (2001).

However, by the early 1980s, the development of cascade models to study turbulent intermittency lead to the realization that in general an infinite number of dimensions were needed. The generic result of a cascade process (see Sect. 4 below) is that the cascade quantity at resolution ε_λ has the statistics:

$$\langle \varepsilon_\lambda^q \rangle = \lambda^{K(q)} \tag{19}$$

where $K(q)$ is (convex) the moment scaling function and λ is the ratio of the largest (outer) cascade scale and the scale of observation. The symbol "ε" is used for the turbulent (scale by scale) energy flux. Below, we discuss the link between these scaling exponents and $\xi(q)$ introduced earlier for the q-th order structure function and the spectral exponent β.

Viewed from the point of multifractals, spectra are second order statistics so that the spectrum provide only a very partial statistical description. A more complete and direct description follows from the use of thresholds (T) to convert fields $\varepsilon(\boldsymbol{x})$ into exceedance sets ($\boldsymbol{x}$ is a position vector), and then the use of box-counting to systematically degrade the resolution of the sets, determining the fractal dimension using the formula:

$$N_T(L) \propto L^{-D(T)}; \quad P_T(L) \approx N_T(L)/L^{-d} \approx L^{c(T)};$$
$$c(T) = d - D(T) \tag{20}$$

where $N_T(L)$ is the number of $L \times L$ sized boxes needed to cover the set of points satisfying $\varepsilon(\boldsymbol{x}) > T$. Since L^{-d} is the total number of boxes in the space at resolution L, P_T is the probability that a box (size L) placed at random on the set will cover part of the set. $D(T)$ is the dimension and $c(T)$ defines the statistical codimension function; a probability exponent. Since probability exponents can be defined without reference to the embedding space of the process (i.e. whether it occurs in a 1-D, 2-D... or for stochastic processes, in infinite D probability spaces), codimensions are generally needed for stochastic fractals and multifractals (see below for the popular $f(\alpha)$, $\tau(q)$ dimension formalism

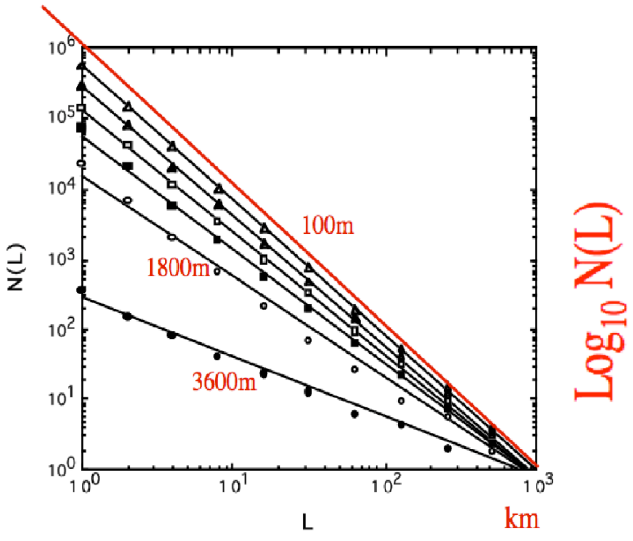

Fig. 16a. Functional box-counting on French topography data at 1 km resolution. For each threshold, the scaling is quite accurate, but as the threshold increases, the slope systematically decreases so that the topography is apparently not monofractal. The line with slope −2 is shown since this is the theoretical assumption of classical geostatistics. Adapted from Lovejoy and Schertzer (1990).

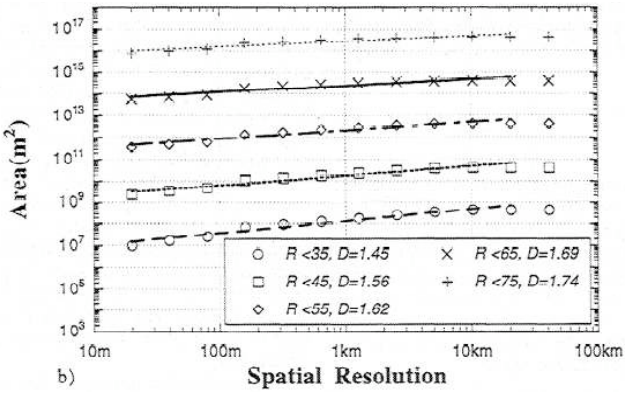

Fig. 16b. A log-log plot of the areas ($A_T(L) \approx L^2 L^{-D(T)}$) of SPOT satellite radiances of Mauna Lao volcano (visible, 20 m resolution) exceeding a radiance threshold $T=R$ (in digital counts), with corresponding fractal dimensions indicated. Each line has been offset by 2 orders of magnitude for clarity. Reproduced from Laferrière and Gaonac'h (1999).

of multifractals). Indeed whenever $D(T)<d$, the stochastic codimension function $c(T)$ defined by Eq. (20) is equal to the geometric codimension function $d-D(T)$; however, in general $c(T)$ will be unbounded. For the corresponding extreme events, if one uses $D(T)=d-c(T)$, one would obtain the geometrically impossible values $D(T)<0$. By using the statistical codimension we thus avoid the paradox negative or "latent" dimensions (Mandelbrot, 1983).

When this "functional" box-counting (Lovejoy et al., 1987) was applied to the topography (Fig. 16a) it was found that the scaling was excellent: the power law Eq. (20) was accurately obeyed for all T, L. However – as expected for a multifractal – $D(T)$ systematically decreases with threshold, it is not constant as assumed in the monofractal models. Indeed, from the point of view of multifractals, it would have been a miracle if for each threshold T, each (different) set had exactly the same fractal dimension. Figure 16b shows the results of functional box-counting on reflected visible radiances from lava flows, showing both the excellent wide range scaling of the flows and also the systematic decrease of $D(T)$ with T. In this figure we directly see a consequence: the areas of lava flows exceeding a threshold depend in a power law way on the resolution: $A_T(L) \approx L^2 L^{-D(T)}$, we return to this important point below.

If the topography could be adequately modeled as a geometrical fractal set, then many different techniques (including spectral analysis) could be used to estimate its unique dimension D. However, due to the multifractality evidenced in the functional box-counting (Fig. 16), on the contrary, when different analysis techniques were applied to different data sets commonly gave different values of D. In particular the empirical topography spectral exponent $\beta \approx 2$ (Fig. 1) would imply D=2.5 for monofractal surfaces (1.5 for monofractal vertical sections) whereas the (rare) direct estimates (Goodchild, 1980; Aviles et al., 1987; Okubo et al., 1987; Turcotte, 1989) commonly gave a diversity of values (see the reviews Klinkenberg and Goodchild, 1992; Maliverno, 1995).

The use of simplistic monofractal ideas had consequences beyond a failure to reach consensus on a supposedly "unique" fractal dimension of the topography. Due to their random singularities, multifractals have such strong variability that they violate many conventional geostatistical assumptions so that normal multifractal variability can easily be misinterpreted in terms of spurious scale breaks, spurious nonstationarity etc. The loss of interest in scaling was encouraged by the extensive use of (low variability) fractional Brownian motion (fBm) models of topography. As argued in Gagnon et al. (2006), the topography in fact has excellent multiscaling (multifractal) properties (see Figs. 1, 16a, 18) – but an infinite hierarchy of fractal dimensions; this requires new analysis techniques.

An unfortunate consequence of this reliance on simplistic monofractal models was that by the end of the 1990s the mainstream surface geomorphology community had "moved on", relegating fractals to narrow ranges of scale and to very technical applications. This near abandonment of scaling occurred in spite of the fact that entire fields of research such as surface hydrology are riddled with scaling laws which virtually require the topography to respect some form of scaling (see e.g. the review Rodriguez-Iturbe and Rinaldo, 1997). Classical examples include power law relations between river basin size area and stream length, basin area to discharge, the relation between velocity, width and depth to discharge (Leopold and Maddock, 1953). Lack of an

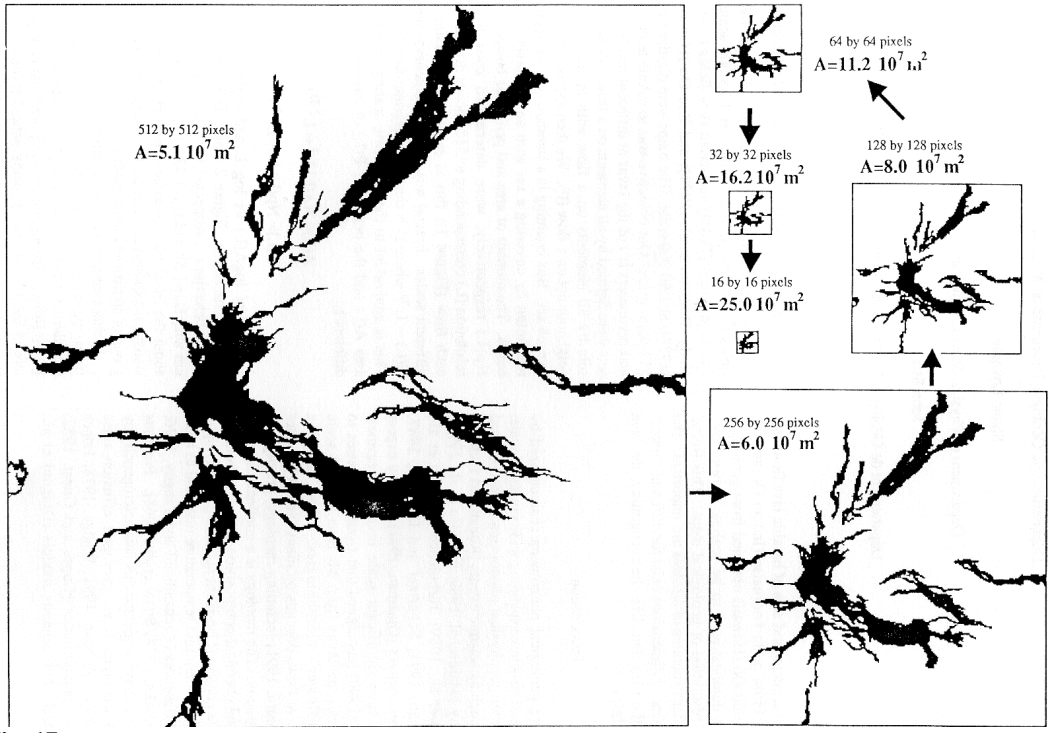

Fig. 17. Lava flows from mount Etna 1900–1974 taken from a geological map at 43 m resolution. The resolution is then successively degraded by factors of 2 using box counting. Reproduced from Gaonac'h et al. (1992).

adequate theoretical framework for scaling has thus led the baby to be thrown out with the bathwater.

3.2 Consequences for classical geostatistics

It is worth mentioning that the functional box-counting results (Fig. 16) have direct and important consequences for classical geostatistics (e.g. Matheron, 1970) which assume (explicitly or implicitly) that geomeasures such as the areas of the topography exceeding a threshold are regular with respect to Lebesgue measures. If this assumption were true, then the areas above a given threshold T would be well-defined independently of the resolution L, i.e. the expression $L^2 N_T(L)$ would be independent of L for small L; however since $D(T)<2$ we see that generally it vanishes as $L \to 0$. Ultimately at small scales – probably millimeters or less – the scaling will break down yielding a finite limit of $L^2 N_T(L)$. However the area estimated $L^2 N_T(L)$ will depend on the very small scale details; at any larger resolutions the result will be subjective depending on the observing resolution L. While Fig. 16b shows this directly on various sets defined by radiance thresholds on volcanoes, Fig. 17 shows the same effect visually, using step by step degradation of the resolution of lava flow maps determined by geological mapping techniques. The usual box-counting method is used to successively degrade the resolution of the flows. As the resolution improves by a factor of 512/16=32, we see (moving in the direction opposite the arrows) more and more fine details. Over this scale factor, the area decreases by a factor of about 5 corresponding to a fractal dimension of the areas of about 1.58; the fractal dimension of the perimeter set is 1.42 so that it is a little bit sparser.

If we express the field values as powers of the resolution with random exponents γ, i.e. if we write $T \propto \lambda^{\gamma}$ then we obtain:

$$\Pr(\varepsilon_\lambda > T) = \Pr(\varepsilon_\lambda > \lambda^{\gamma}) \propto \lambda^{-c(\gamma)} \tag{21}$$

where "Pr" indicates "probability". For cascade processes, we derive this result directly in Sect. 4.3. Since the moments (Eq. 19) are integrals over the probability density (dPr), $c(\gamma)$ determines $K(q)$; we discuss this link in Sect. 4.7.

3.3 Evidence for multiscaling of statistical moments

We have shown that for many geophysical fields in both the horizontal and vertical, that the spectrum shows evidence of power law behaviour – scaling – over wide ranges of scale. However, (functional) box-counting on exceedance sets defined by higher and higher field thresholds showed that many exponents were needed to characterize the scaling of the low and high regions of the topography, weak and strong radiances etc. Although for the topography, the multiscaling is demonstrated by functional box counting (Fig. 16a), we can directly test the multiscaling of the moments (Eq. 19) and

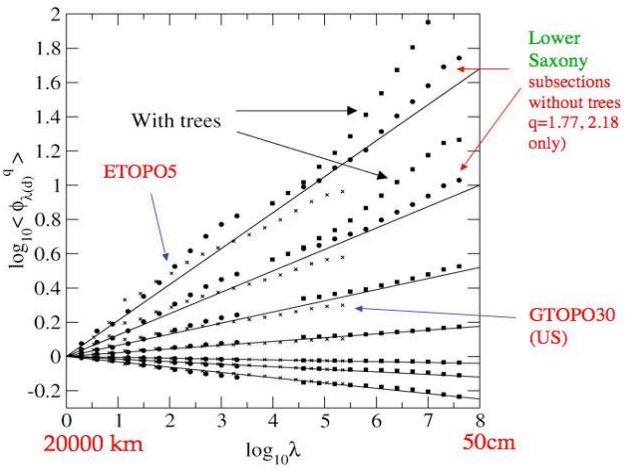

Fig. 18. Log-log plot of the normalized moments versus the scale ratio $\lambda = L_{\text{outer}}/l$ (with L_{outer}=20 000 km) for the three DEM's (circles correspond to ETOPO5, X's to U.S. (GETOPO30), and squares (Lower Saxony). The solid lines are there to distinguish between each value of q (from top to bottom, q=2.18, 1.77, 1.44, 1.17, 0.04, 0.12, 0.51) The trace moments of the Lower Saxony DEM with rees for q=1.77 and q=2.18 are on the graph (indicated by arrows). The theoretical lines are computed with the global $K(q)$ function discussed with universal multifractal parameters α=1.79, C_1=0.12. At scales <40 m, in this Lower Saxony data set, the effect of trees becomes important, apparent increasing the variability at the smallest scales. Reproduced from Gagnon et al. (2006).

determine the moment scaling exponent $K(q)$. To do this, we take for the multifractal field ε_Λ the absolute gradients of the topography at the finest resolution of the data set $\Lambda = L/l$ where L is the external scale (taken as 20 000 km here) and l is the pixel scale (see Sect. 4.10 for more discussion of this). The result of degrading the high resolution ε_Λ to intermediate scale ratios λ is shown in Fig. 18 (using the same data sets as in Fig. 1). We can see that the multiscaling holds very well over a factor of more than 10^5 in scale. Indeed, Gagnon et al. (2006) estimates that the "reduced moments" $<\varepsilon_\lambda^q>^{1/q}$ for all $q \leq 2$ can be reproduced to within $\pm 45\%$ using just a 2 parameter "universal multifractal" fit to the $K(q)$ function (Eq. 45; see Sect. 4.6). Other relevant examples of multifractal analysis are soil moisture (Dubayah et al., 1997), LANDSAT TM channels (Cheng, 1999), sonic velocities (Marsan and Bean, 1999) and neutron porosity (Marsan and Bean, 2003); Figs. 25a, b (the latter two in the KTB borehole). In Sect. 4.10 we perform various multifractal analyses on the de Wijs (1951) Zn concentration series.

3.4 Multifractality and spurious breaks

In spite of the systematic finding of scaling or near scaling statistics, many geophysicists instinctively reject all wide range scaling; they consider a priori that the scaling is broken. However conclusions about broken scaling are frequently unwarranted. Perhaps the most important source of misinterpretation is the fact that scale invariance is a statistical symmetry which is almost surely broken on every single realization, hence it is important to have a large data base (i.e. large range of scales, many realizations) to average fluctuations and to approximate the theoretically predicted ensemble average scaling. In fact, due to the singularities of all orders (see the previous section) the variability of multifractals is much greater than that of classical stochastic processes; for example, rare (extreme) singularities are produced by the process yet they are almost surely absent on any given realization. This means that multifractal processes generally do not have the property of "ergodicity". What may be nothing more than normal multifractal statistical variability can thus easily be interpreted as breaks in the scaling. A second reason for unwarranted rejections of scaling is the assumption that the scaling is isotropic. If the scaling is anisotropic, there may be breaks in the scaling on 1-D subspaces (e.g. transects) but not for the full process in the higher dimensional space in which they evolve. A third reason discussed in more detail in Gagnon et al. (2006) is that there can be systematic biases due to the use of conditional statistics such as studying transects that just happen to pass through special features (such as high mountains).

There are also nonclassical statistical effects which can lead to yet other misinterpretations of the data. One of these is a consequence of the fact that the strong singularities in multifractals leads to apparent nonstationarities: e.g. to quite different morphologies which can often be found in close proximity. This is often interpreted in terms of nonstationarities/spatial inhomogeneities – different processes at work in different regions or at the very least, variations in the parameters of a single basic model. However, with multifractals such interpretations would be unwarranted: the basic multifractal processes are statistically stationary/homogeneous in the strict sense that over the region over which they are defined (which is necessarily finite), the ensemble multifractal statistical properties are independent of the (space/time) location (and this – contrary to certain affirmations in the literature – for any spectral slope β). Rather than discussing this at an abstract level, let us see what happens when we analyse a self-similar 1024×1024 multifractal simulation (Fig. 19a).

In the simulation, consider the "regional" variability in the spectral exponent β by dividing it into 8×8 squares, each with 128×128 pixels. Figure 19b shows the histogram of the 64 regression estimates of the spectra compensated by the theoretical behaviour i.e. $E(k)/(k^{-\beta_{\text{theory}}})$ with β_{theory}=2.17. As expected, the mean is close to zero but we see a large scatter implying that there are some individual regions having β as low as 1.2, some as high as 2.7; the standard deviation is ± 0.3. As we shall see later, this would imply a random variation in local estimates of the nonconservation parameter H of $\pm 0.3/2 = \pm 0.15$. Although it is of the order of the difference observed between continents and oceans, this spread in β, H will decrease as the size of the data set increases. In

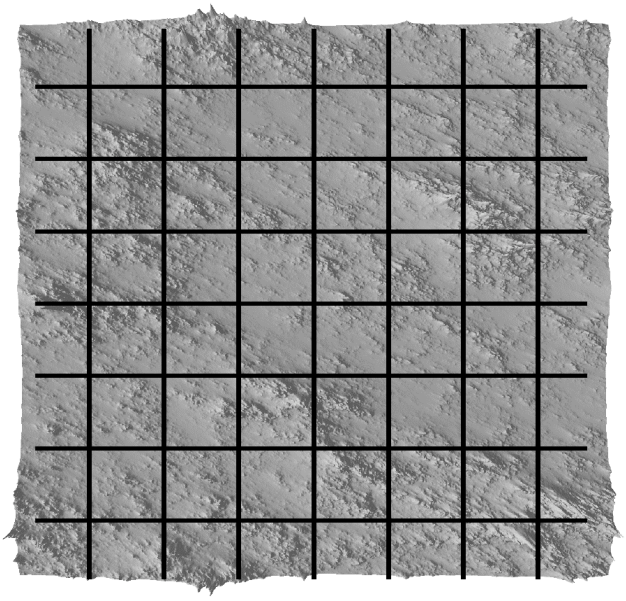

Fig. 19a. A self-similar multifractal (with some trivial anisotropy) simulated on a 10–24×1024 poihnt grid with observed universal multifractal parameters (H=0.7, C_1=0.12, α=1.9); the spectral exponent is β=1+2$H-K$(2)=2.17. Adapted from Gagnon et al. (2006).

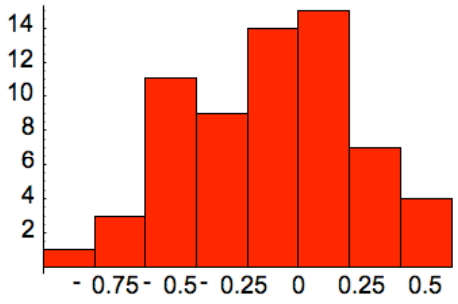

Fig. 19b. After dividing Fig. 19a into 64 128×128 squares, we calculated the isotropic spectrum in each, and fit the slope to the lowest factor 16 in scale (we remove the highest factor 4 due to numerical artifacts at the highest wavenumbers). The resulting $\Delta\beta$ is given in the left; it is twice the ΔH, showing that H can vary by 0.5 over a single region. From Gagnon et al. (2006).

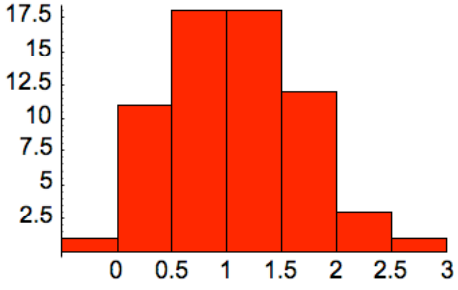

Fig. 19c. A histogram of the $\log_{10}E_1$ (E_1 is the spectral prefactor: $E(k)=E_1k^{-\beta}$) showing variation of 1000 from the smoothest to roughest subregion. From Gagnon et al. (2006).

Fig. 19c, we can also see the large variations in the log prefactors ($\log_{10}$E$_1$; $E(k)=E_1k^{-\beta}$). If E_1 is interpreted in terms of roughness, the roughest of the 64 regions has about 10^3 times the variance of the smoothest. While it would obviously be tempting to give different interpretations to the parameters in each region, this would be a mistake. The physical interpretation of such a model is that the roughest and the smoothest are associated with huge variations in the corresponding erosional, orographic and other processes; this would follow if these processes are also scaling and would have correlated variations.

4 Cascades and multifractals

4.1 Ore distributions, the de Wijs binomial cascade, the lognormal versus Pareto debate

We have seen that there is much evidence for the wide range scaling of various geophysical fields in both the horizontal and vertical directions from sub metric to the largest scales probed by the deepest boreholes (several kilometers) in the vertical and from sub metric to planetary scales in the horizontal. So far, we have not made a serious attempt to explain these results except to comment that since scale invariance is a symmetry principle, the nonlinear dynamics which are responsible for the wide range heterogeneity must repeat scale after scale in a cascade like manner. We now turn to the generic cascade process.

Cascades were first proposed as a dynamical mechanism in order to explain atmospheric dynamics by Richardson (1922) in his celebrated poem; "Big whorls have lesser whorls that feed on their velocity and lesser whorls have smaller whorls and so on to viscosity (in the molecular sense)". While Richardson had the idea of structures breaking up into smaller and smaller structures, an independent idea going back somewhat earlier, was the "law of proportional effect" (Kapteyn, 1903), in which a random variable is the product of other random variables leading under certain assumptions to the log-normal distribution, see e.g. Aitchison and Brown (1957). Although they weren't explicitly mentioned, both the cascade idea and the law of proportional effect provided the impetus for proposals for log-normal distributions of rock fragments (Kolmogorov, 1941a) and energy dissipation in turbulence (Kolmogorov, 1962). Starting with the monofractal "pulse in pulse" model (Novikov and Stewart, 1964), explicit cascade models began to appear systematically in the turbulence literature in the 1960s; notably Yaglom (1966)'s lognormal model followed by Mandelbrot (1974)'s criticism concerning the divergence of statistical moments. In spite of these later developments it is interesting that they were all anticipated by a mining engineer (de Wijs, 1951, 1953).

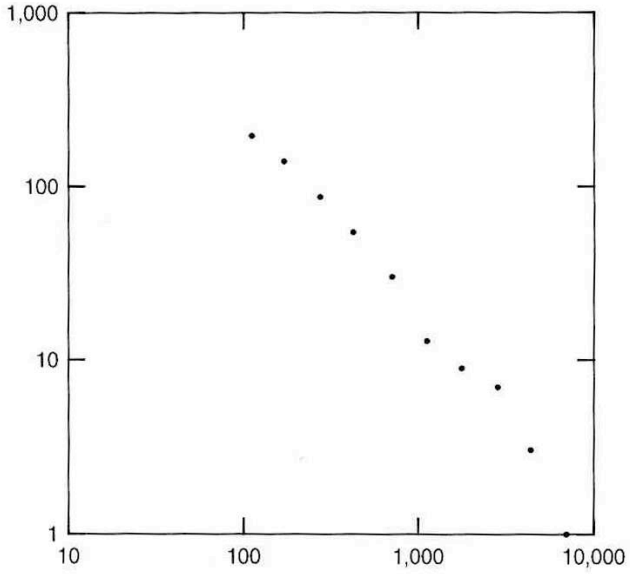

Fig. 20. Ore valuations of a South African Mine; "cumulated Paretian graphs" i.e. doubly logarithmic plots of the numbers of valuations exceeding a given valuation. The distribution is nearly hyperbolic (linear on this plot) with exponent q_D near one (Krige, 1960) (reprinted in Mandelbrot, 1995).

De Wijs was interested in the concentration of ores and was debating the form of their probability distributions. Along with Lasky (1950), he defended the idea that the probabilities were log-normal, criticizing (Van Tongeren, 1950) who on the contrary defended algebraic (power law) distributions. In order to help prove his point he proposed a simple cascade model which he called the "binomial" model. At the time, some of the proponents of log-normality even went so far as to propose it as a the first law of geochemistry (Ahrens, 1953). The debate about lognormality versus power law (often called "Pareto" in this context) continued through the 1960s to the 1980s, with notably Matheron (1962) siding with the lognormal camp; see also Cargill et al. (1981) and Agterberg (2007). In the 1980s, this binomial model was rediscovered by and applied in geology to the distribution of fossils (Plotnick and Prestegaard, 1995), while Turcotte (1986) made a drastic modification to the cascades so as to generate a Pareto distribution (see Sect. 4.8 below). In the turbulence literature the binomial model became known as the "p model" (Meneveau and Sreenivasan, 1987); we shall see below that it is actually a microcanonical restriction on the "α model" (Schertzer and Lovejoy, 1985a).

To put the debate in perspective, we show Fig. 20 which is an example of the distribution of ore grades indicating that empirically, they can be far from log-normal (Krige, 1960; reprinted in Mandelbrot, 1995); see Fig. 3 and below for a re-examination of the de Wijs data). Note that Cheng et al. (1994), and Cheng (2000a) proposed a variant of the method of plotting in Fig. 20 called the "Concentration-area fractal method" involving plotting the logs of the areas of metal bearing ores exceeding various concentration thresholds, the latter also plotted on a log scale.

4.2 The binomial/p model and the α model

In order to demonstrate how a roughly log-normal distribution of ores might arise, de Wisj considered a 1-D section which he successively divided into two equal halves; he then reasoned that various processes might concentrate the material in the left segment by a factor $(1+d)$ reducing the concentration on the right segment by the factor $(1-d)$ where $0 \leq d \leq 1$ is the "dispersion index"; empirically for many ore concentration series, de Wijs found $d \approx 0.2$ (typical of iron and zinc deposits), although for precious metals, values as high as 0.45 were obtained (see also Agterberg, 2007, for more examples). He then considered the effect of repeating this multiplicative construction to smaller and smaller scales (but without considering the nontrivial mathematical limit).

In order to understand this, let's change the notation and generalize this slightly. Denote by λ the division ratio, and the multiplicative factors by $\mu\varepsilon_i$ (in analogy with the symbol "Δx" for an additive increment), where "i" indexes the factors (left or right which can be chosen randomly); and D the dimension of the space. de Wijs's model thus corresponds to λ=2, D=1 and the values $\mu\varepsilon_+=(1+d)$, $\mu\varepsilon_-=(1-d)$ were always chosen together (left or right) so that they satisfy:

$$\frac{1}{\lambda^D}\sum_{i=1}^{\lambda^D}\mu\varepsilon_i = 1 \tag{22}$$

where $\mu\varepsilon_i=\mu\varepsilon_+$ or $\mu\varepsilon_-$. The sum ensures that the increase (decrease) in ore in the left half is exactly compensated by a decrease (increase) in the right half. If $\lambda>2$ and/or if $D>1$ then there can be several states but at each step, each "parent" and "daughter" structures satisfy the restriction Eq. (22). In analogy with statistical mechanics, this strict scale by scale conservation is called "microcanonical" (Mandelbrot, 1974).

To obtain the more general "canonical" cascade, it suffices to replace the microcanonical conservation by:

$$\langle \mu\varepsilon \rangle = 1 \tag{23}$$

where "<.>" indicates ensemble averaging; in the canonical cascade the left and right hand factors are thus chosen independently of each other. The resulting two state model (in any dimension D) was called the "α model" (Schertzer and Lovejoy, 1985a) (see Fig. 21) in order to distinguish it from the pure fractal "β model" (Frisch et al., 1978). To understand the statistics of these binomial processes, write the probabilities of the model states as follows:

$$\begin{aligned} &\Pr(\mu\varepsilon = \lambda^{\gamma_+}) = \lambda^{-c} (> 1 \Rightarrow \text{ increase}) \\ &\Pr(\mu\varepsilon = \lambda^{\gamma_-}) = 1 - \lambda^{-c} (< 1 \Rightarrow \text{ decrease}) \end{aligned} \tag{24}$$

with parameters γ_+ and γ_- corresponding to the maximum and minimum singularities that the model can produce. The

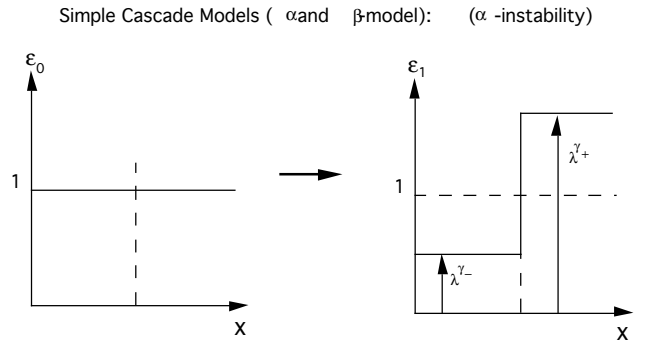

Fig. 21a. Illustration of the α model: here $\varepsilon_n = \mu\varepsilon \cdot \varepsilon_{n-1}$. The *weak* sub-eddies have an associated probability $\Pr(\mu\varepsilon = \lambda^{\gamma_-}) = 1 - \lambda^{-c}(\gamma_- < 0)$ whereas the *strong* sub-eddies have $\Pr(\mu\varepsilon = \lambda^{\gamma_+}) = \lambda^{-c}(\gamma_+ > 0)$.

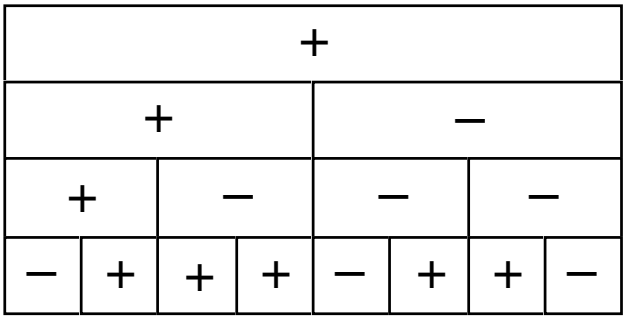

Fig. 21b. Schema of tree of singularities for a one-dimensional α model the "+" indicates a choice ofγ_+, the "–",γ_- , with probabilities as above. The microcanonical "de Wisj" (or "binomial" or "p" model) would have "+" and "–" always occurring in pairs so that at each scale and each location, the total ore amount is rigidly conserved.

de Wijs model is recovered with the parameters λ=2, c=1, $\gamma_+ = \log\mu\varepsilon_+/\log\lambda = \log_\lambda(1+d)$, $\gamma_- = \log\mu\varepsilon_-/\log\lambda = \log_\lambda(1-d)$ and with the additional condition that the only randomness is to choice of which of the two is left or right.

In the α model, the canonical conservation condition implies:

$$\lambda^{\gamma_+} \cdot \lambda^{-c} + \lambda^{\gamma_-} \cdot (1 - \lambda^{-c}) = 1 \quad (25)$$

because of this constraint out of c, γ_+ and γ_-, there are really only two free parameters, this is valid for any λ, D. For the microcanonical model, the conservation condition on the contrary depends not only on λ, but also on D. A purely "all or nothing" process called the "β-model" (Frisch et al., 1978) is obtained with $\gamma_- = -\infty$; this is the monofractal limit; the nonzero region is a fractal set with codimension c.

Whenever $\gamma_- > -\infty$ and the process is iterated, the pure orders of singularity γ_- and γ_+ lead to the appearance of mixed orders of singularity, (the "α model"). Mixed singularities of different orders $\gamma(\gamma_- \le \gamma \le \gamma_+)$ are built up step by

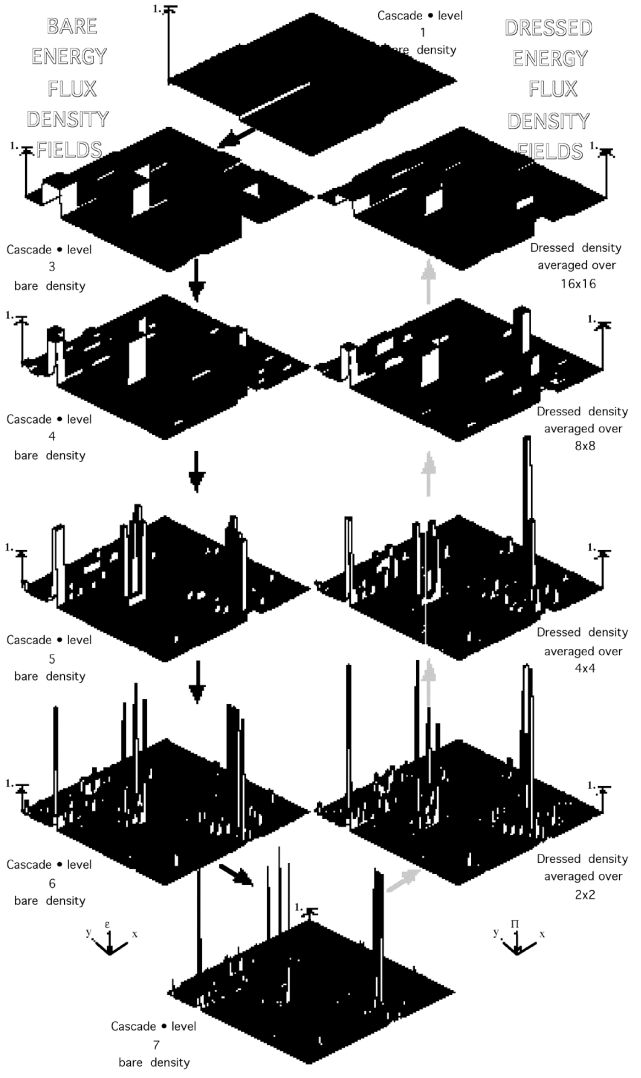

Fig. 21c. The α model in 2-D showing both the bare (left) and dressed cascades (right). Reproduced from Wilson (1991).

step through a complex succession of γ_- and γ_+, as illustrated in Fig. 21b. Figure 21c shows a 2-D example of the α model which we will study in more detail in the next section. In other words, leaving the simplistic alternative dead or alive ("β model") for the alternative weak or strong ("α model") leads to the appearance of a full hierarchy of levels of survival, hence the possibility of a hierarchy of dimensions.

4.3 Renormalizing discrete cascades

What is the behavior as the number of cascade steps, $n \to \infty$? Consider two steps of the process, the various probabilities and random factors are:

$$\begin{aligned} &\Pr(\mu\varepsilon = \lambda^{2\gamma_+}) = \lambda^{-2c} \text{ (two boosts)} \\ &\Pr(\mu\varepsilon = \lambda^{\gamma_+ + \gamma_-}) = 2\lambda^{-c}(1 - \lambda^{-c}) \text{ (one boost and one decrease)} \\ &\Pr(\mu\varepsilon = \lambda^{2\gamma_-}) = (1 - \lambda^{-c})^2 \text{ (two decreases)} \end{aligned} \quad (26)$$

This process has the same probability and amplification factors as the three-state α model with a new scale ratio of λ^2, i.e.,

$$\begin{aligned}&\Pr(\mu\varepsilon = (\lambda^2)^{\gamma_+}) = (\lambda^2)^{-c}\\ &\Pr(\mu\varepsilon = (\lambda^2)^{(\gamma_+ + \gamma_-)/2}) = 2(\lambda^2)^{-c/2} - 2(\lambda^2)^{-c}\\ &\Pr(\mu\varepsilon = (\lambda^2)^{\gamma_+}) = 1 - 2(\lambda^2)^{-c/2} + (\lambda^2)^{-c}\end{aligned} \quad (27)$$

Iterating this procedure, after $n{=}n^+{+}n^-$ steps we find:

$$\begin{aligned}&\gamma_{n^+,n^-} = \tfrac{n^+\gamma_+ + n^-\gamma_-}{n^+ + n^-},\ n^+ = 1, ..., n\\ &\Pr(\mu\varepsilon = \lambda^{\gamma_{n^+,n^-}}) = \binom{n}{n^+}\lambda^{-c\,n^+}(1-\lambda^{-c})^{n^-}\end{aligned} \quad (28)$$

where $\binom{n}{k}$ is the number of combinations of n objects taken k at a time. This implies that we may write:

$$\Pr(\varepsilon_{\lambda^n} \geq (\lambda^n)^{\gamma_i}) = \sum_j p_{ij}\,(\lambda^n)^{-c_{ij}} \quad (29)$$

The p_{ij}'s are the "submultiplicities" (the prefactors in the above), c_{ij} are the corresponding exponents ("subcodimensions") and λ^n is the total ratio of scales from the outer scale to the smallest scale. Notice that the requirement that $\langle\mu\varepsilon\rangle{=}1$ implies that some of the λ^{γ_i} are greater than one (boosts) and some are less than one (decreases), that is some $\gamma_i{>}0$ and some $\gamma_i{<}0$. Note also that the α-model will have bounded singularities:

$$\gamma_- \leq \gamma_i \leq \gamma_+ \quad (30)$$

(i.e., the maximum attainable singularity $\gamma_{\max}$ is equal to γ_+). The final step in "renormalizing" the cascade is to replace the above n-step (ratio λ), 2-state cascade by a single λ^n step cascade with $n{+}1$ states. Note that we are not saying that there is absolutely no difference between the n-state α-model with ratio λ and the corresponding $(n{+}1)$-state model with $\lambda'{=}\lambda^n$; however their properties will be identical for integral powers of λ'. Finally, doing this and making the replacement $\lambda^n{\to}\lambda$, and the limit $\lambda{\to}\infty$, one of the terms in the sum will dominate (that with the smallest c_{ij}). Hence defining

$$c_i = \min\left\{c_{ij}\right\} = c(\gamma_i) \quad (31)$$

yields for $\lambda{\to}\infty$:

$$\Pr(\varepsilon_\lambda \geq \lambda^{\gamma_i}) = p_i \cdot \lambda^{-c_i} \quad (32)$$

where c_i is the codimension and p_i is the multiplicity. If we now drop the subscripts "i" (this allows for the possibility of a continuum of states, e.g., the original process being defined by a uniform or other continuous distribution) then we obtain:

$$\Pr(\varepsilon_\lambda \geq \lambda^{\gamma_i}) \sim \lambda^{-c(\gamma)} \cdot p(\gamma);\ \frac{dc}{d\gamma} > 0 \quad (33)$$

This is a basic multifractal relation for cascades. We now simplify this using the "$\sim$" sign which absorbs the multiplicative ($p(\gamma)$) as well as taking into account the logarithmic number of terms in the sum (which can lead to logarithmic prefactors corresponding to "sub-codimensions"). With this understanding about the equality sign, we may write

$$\Pr(\varepsilon_\lambda \geq \lambda^\gamma) \sim \lambda^{-c(\gamma)} \quad (34)$$

Each value of ε_λ corresponds to a singularity of order γ and codimension $c(\gamma)$. Note that strictly speaking the expression "singularity" applies to $\gamma{>}0$ (for $\lambda{\to}\infty$), when $\gamma{<}0$ it is a "regularity".

In the geophysics literature, there has also been a variant on the microcanonical cascade called the "bounded cascade" (Cahalan, 1994) in which the cascade is progressively killed as the cascade proceeds by multiplying each $1{-}\mu\varepsilon_i$ by r^n where $0{<}r{<}1$ and n is the number of cascade steps from the beginning of the cascade. In this way, the dispersion coefficients d algebraically decreases: $d_{n+1}{=}rd_n$ so that rapidly all the $\mu\varepsilon_i{\approx}1$. This has the drastic effect of essentially destroying the multiplicative nature of the cascade at the small scales, effectively turning it into an additive process (Lovejoy and Schertzer, 2006). In the small scale limit we obtain essentially a truncated Brownian motion with only trivial multifractality. Another variant on the basic microcanonical model has been proposed by Cheng (2005). In this 2-D model, there are 4 different weights $\mu\varepsilon$ which are chosen deterministically always in the same 2X2 pattern. The resulting cascade is generally anisotropic. Although Cheng notes that no scale by scale conservation property generally holds on 1-D sections (this presumably leads to nontrivial problems of convergence in the small scale limit), this model is proposed for anisotropic multifractal fields.

4.4 Unlocalized versus localized singularities

Note that while the above form of the probability distributions/histogram Eq. (34) is valid at every step of the cascade process (every finite λ), this in no way implies that there is convergence of γ at a given mathematical point $\boldsymbol{x}$. Indeed for canonical cascades, in general we have $\lim_{\lambda\to\infty} \mathrm{Log}_\lambda(\varepsilon_\lambda(\boldsymbol{x}))$ does *not* converge to a well defined (point singularity) value $\gamma(\boldsymbol{x})$. Viewed from a single mathematical point, the local resolution λ singularities; $\gamma_\lambda{=}\mathrm{Log}_\lambda(\varepsilon(\boldsymbol{x}))$ perform random walks as λ increases, they are therefore generally *not* "Hölder exponents". In contrast, many multifractal papers (including the original proposal by Parisi and Frisch (1985) and the dimension formalism of multifractals, Halsey et al. (1986), simply assume a priori that the singularities are localized. Since then, wavelets have been extensively applied to multifractals. However, since wavelets are tools of functional analysis whereas multifractals generated by cascades are not mathematical functions but rather densities of singular measures – the applications of wavelets (Bacry et al., 1989)

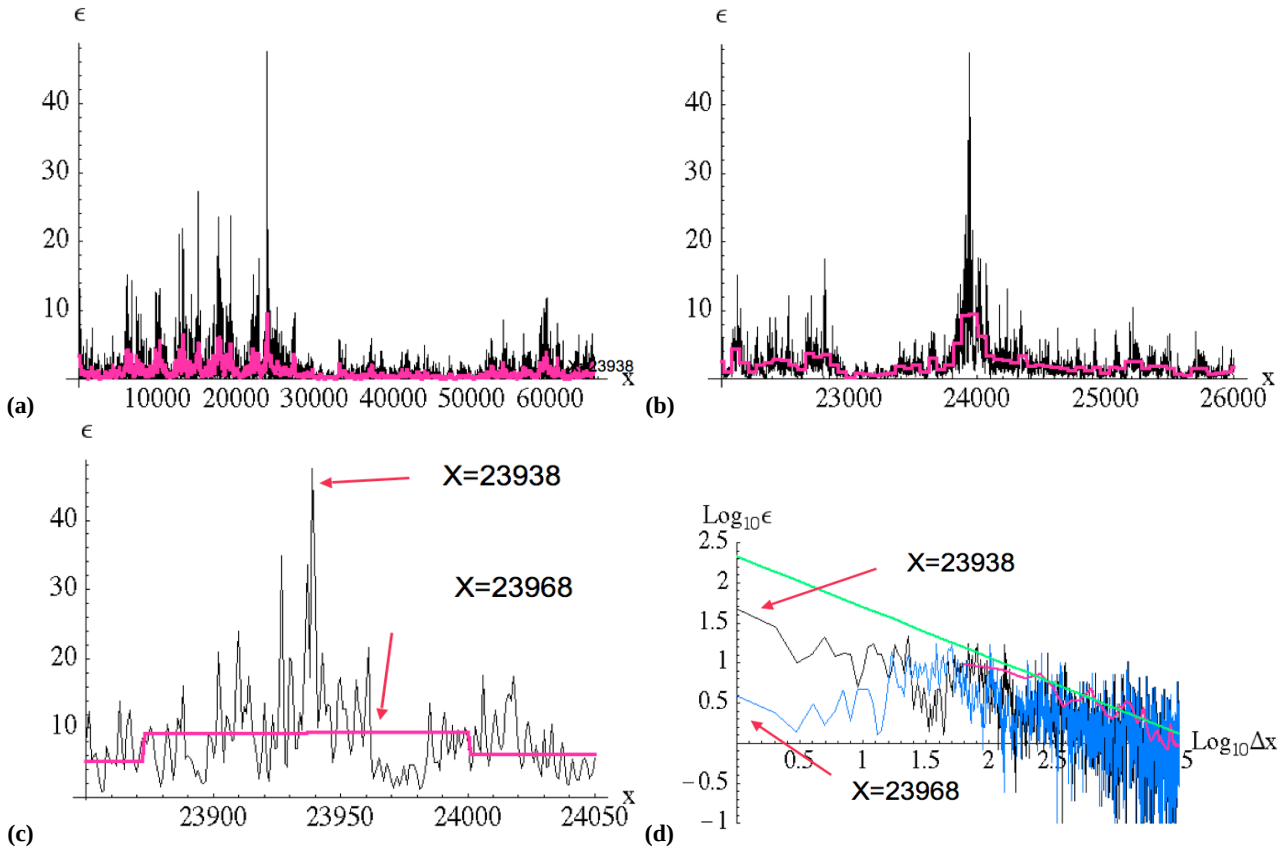

Fig. 22. The anatomy of a singularity. **(a)** Upper left is full (continuous in scale) simulation 2^{16} long with universal multifractal parameters α=1.8, C_1=0.05 (close to de Wijs data values for Zn ore). Black is the full resolution data, pink is low resolution, degraded by a factor of 64. **(b)** The upper right shows a zoom (factor 16) into the section with the maximum value, **(c)** the lower left is a zoom by a further factor 20 near the maximum; the arrows show the position of the high resolution maximum as well as the centre of the low resolution maximum. **(d)** The lower right is the same but on log-log plot using distance from the maximum. The pink shows the approach to the maximum low resolution singularity on the low resolution series with the green being the rms fit giving the estimate 0.63 (the absolute slope) for the maximum singularity, showing it's extrapolation to the full resolution of the process; the low resolution tends to over estimate the high resolution singularity. The black is the actual approach to the maximum singularity, while the blue is the approach to the centre of the low resolution singularity showing that the latter is a poor estimator of the position.

may not always be justified. This is particularly true of certain complex derived analysis methods such as the "modulus maximum" analysis technique which is designed to "zoom" into point singularities. While these methods may work to some extent, they are not fully justified for multifractals generated from general cascade processes. For an example of the use of wavelets in this way in boreholes, see Fedi et al. (2005); the local singularity estimates he obtains is very similar to that those described in the next paragraph.

In order to graphically appreciate the difficulties of zooming into singularities, we refer the reader to Fig. 22a which shows a universal multifractal continuous in scale multifractal simulation with parameters close to those obtained for de Wijs for his binomial model of ores (Sect. 4.10). A low resolution (degraded by factor 64) curve is superposed; this simulates an empirical transect which would be at lower resolution than the actual ore concentration series. Figure 22b shows a blow up by a factor 16 and Fig. 22c by a further factor of 20. This is a 1-D simulation of the idea proposed by Cheng (2006, 2007) of prospecting by zooming in on major ore deposits by extrapolating singularities. In Fig. 22d we have plotted on a log-log plot the distance from the maximum of the low resolution series (pink) as well its log-log regression line (green). Also shown is the log-log plot of the approach to the actual maximum singularity at high resolution as well as the centre of the low resolution singularity; this is presumably the best estimate of the singularity location if only low resolution data were available. Note also that the extrapolation of the low resolution regression yields an estimate of the maximum which is about 5 times too large, while the centre value is about 10 times smaller.

In this case, we zoomed into the centre of the low resolution singularity, other choices would have given different small distance behaviour, but since the multiplicative factors are statistitically independent, the low resolution data gives little indication of where within the low resolution element that the highest singularity is located. This graphically indicates that the low resolution approach to a singularity lacks predictive power for locating the position of singularities at high resolution within low resolution elements. Of course, singularities in real ore concentration fields may have more localized singularities than those generated by cascades, and Cheng (2006)'s method uses sparse 2-D data (rather than regular 1-D data as here), so that it is possible that the method may nevertheless prove useful; indeed, Cheng (2007) shows some success for finding Sn, Cu, As and other deposits in China.

4.5 The statistical moments, cumulants, second characteristic functions

The simplest way to investigate the statistical properties of a cascade developed over a scale range λ, is to consider their "moment scaling function exponent" $K(q)$:

$$K(q) = \mathrm{Log}_\lambda \langle \varepsilon^q \rangle \tag{35}$$

$K(q)$ is the "second (base λ, Laplace) characteristic function", or "cumulant generating function" of the random variable $\mathrm{Log}_\lambda \mu\varepsilon$. This is valid for both canonical and microcanonical conservation; the differences will appear when we consider the integrals of completed cascades over finite sets; the "dressed properties" (Sect. 4.7). Due to the statistical independence of the cascade factors $\mu\varepsilon$, it is enough to consider a single step of the cascade process. To see this, note that the $K_n(q)$ which is the $K(q)$ for n cascade steps is:

$$K_n(q) = \mathrm{Log}_{\lambda^n} \left\langle \prod_{i=1}^{n} \mu\varepsilon_i^q \right\rangle = \frac{1}{n} \mathrm{Log}_\lambda \left[\prod_{i=1}^{n} \langle \mu\varepsilon_i^q \rangle \right] = K(q) \tag{36}$$

We now note that for the α model, binomial/p model we have the expansion:

$$\begin{aligned} K(q) &= \log_\lambda \left(\lambda^{q\gamma_+} \lambda^{-c} + \lambda^{q\gamma_-} \left(1 - \lambda^{-c}\right) \right) \\ &= A_1 q + A_2 q^2 + O\left(q^3\right) \end{aligned} \tag{37}$$

we have used the fact that the normalization of the probability density implies quite generally for nonzero processes that $K(0)$=0 (if the cascade is only nonzero on a fractal support, then $K(0)$=$-C_s$ where C_s is the codimension of the support). If we keep only the terms up to second order, we have a log normal cascade; de Wijs realized that his model was only approximately log-normal, but this was sufficient for his purpose. Agterberg (2007) numerically studied the difference between the binomial model and the lognormal model and introduced a variant, the "random cut" model. However de Wijs didn't realize that this argument against algebraic probabilities was not as convincing as it seemed; whereas his binomial/p model does have finite moments of all orders; the same is not true of the superficially similar α model which has algebraically distributed extremes; we return to this question of "divergence of moments" in Sect. 4.7.

4.6 Universality

As presented above, the problem with cascades is that we need an entire (nearly arbitrary) convex function $K(q)$ for their specification. Yaglom (1966) sensed the problem and already argued – essentially on the basis of the law of proportional effects – for approximate "log-normality" on the basis of the usual central limit theorem applied to $\log\varepsilon$ after a large number of cascade steps (see also Venugopal et al., 2006) for similar arguments on the smallness of the high order terms of the cumulant generating function $K(q)$). The problem is that the cascade requires a scale by scale conservation principle, otherwise there are no well defined small scale cascade limits, and it turns out that this normalization is in contradiction with the normalization required for central limit convergence. In other words, as seen above (Eq. 37, due to the $O(q^3)$ terms) an α or binomial/p model remains only "approximately" log normal even after an arbitrary number of cascade steps. This lead notably (Mandelbrot, 1989) to declare that "in the strict sense, there is no universality whatsoever... this fact about multifractals is very significant in their theory and must be recognized...". However multifractal universality classes *do* exist; two different routes to universality have been proposed; both consider a cascade developped only over a finite range of scales. Only after central limit theorem convergence has been achieved does one consider the small scale limit. The first route to universality (Schertzer and Lovejoy, 1987) relies on a "densification" of the cascade, adding more and more intermediate scales in a cascade defined over a finite range; an "infinitely divisible" or continuous (in scale) cascade. An easier to analyze route – the nonlinear "mixing" of cascade processes – was proposed by Schertzer et al. (1991); indeed this very practical question of multifractal universality was the the subject of debate during the 1990s (Gupta and Waymire, 1993; Schertzer and Lovejoy, 1997); see also Brax and Pechanski (1991), Kida (1991) for the closely related issue of Log-Levy cascades. We should also mention that a weaker "log-Poisson" universality has also been proposed by She and Levesque (1994) but this is only "infinitely divisible" (continuous in scale), not stable nor attractive.

In order to obtain an exactly log-normal cascade we may consider ε which is the result of nonlinear (renormalized, multiplicative) interaction of N (generally non-lognormal) discrete cascades over a cascade with a total range of scale λ:

$$\varepsilon = \left(\prod_{j=i}^{N} \frac{\varepsilon_i}{a_N} \right)^{1/b_N}$$

$$K_N(q)=Nq\left(A_1-\frac{\mathrm{Log}_\lambda a_N}{b_N}\right)+A_2N\frac{q^2}{b_N}+A_3N\frac{q^3}{b_N}+\ldots \quad (38)$$

here, i indexes the N independent cascade processes, and a_N, b_N are recentring and renormalizing constants which must be chosen so that the limit of many interacting processes $N\rightarrow\infty$ is well defined. Indeed, for the α model it suffices to choose a_N, b_N such that:

$$b_N = N^{1/2}; \quad A_1 = \frac{\mathrm{Log}_\lambda a_N}{b_N} \quad (39)$$

and we obtain:

$$K_\infty(q) = \lim_{N\rightarrow\infty} K_N(q) = A_2q^2 \quad (40)$$

i.e. the higher order terms disappear, $K_\infty(q)$ is a pure quadratic, it is the moment scaling function of a pure lognormal multifractal. Once the central limit theorem convergence has been achieved, one then considers the small scale limit (Sect. 4.7); here we must normalize the pure log-normal process so that the small scale cascade limit is well behaved, this is easily performed by noting that an unnormalized ε may be normalized by $\varepsilon\rightarrow\varepsilon/\langle\varepsilon\rangle$ so that $K(q)\rightarrow K(q)-qK(1)$ and we obtain:

$$K(q) = C_1(q^2 - q) \quad (41)$$

where we have used the notation C_1 for the constant A_2 (see below).

The above argument explains how using only a small variant, de Wijs could have argued for a pure (rather than approximate) lognormal multifractal process. However, the above argument is apparently much more general than simply a binomial process (p or α model). Indeed, it simply relies on the fact that $K(q)$ is analytic at the origin and then uses the Taylor expansion (Eq. 37). Unfortunately – as pointed out by Levy (1925) in the context of sums of independent random variables – this does not exhaust the possibilities. Indeed, more generally we must allow for the possibility of nonanalytic $K(q)$ with the following small q expansion:

$$K(q) = A_\alpha q^\alpha + A_1qA_2q^2 + O(q^3) \quad (42)$$

if the new nonanalytic term $A_\alpha q^\alpha$ has $\alpha<2$, then, repeating the above universality argument, with the choice:

$$b_N = N^{1/\alpha}; \quad A_1 = \frac{\mathrm{Log}_\lambda a_N}{b_N} \quad (43)$$

we obtain:

$$K_\infty(q) = A_\alpha q^\alpha; \quad 0 \leq \alpha < 2 \quad (44)$$

(note that when α=1, the nonanalytic term must be taken as $q\log q$). As a technical point, $K_\infty(q)=A_\alpha q^\alpha$ corresponds to a random $\log\varepsilon$ which follows an "extreme asymmetric" Levy distribution, sufficient for cascade processes (the "multiplicative central limit theorem"); see Samorodnitsky and Taqqu (1994) for the more general Levy's needed for the complete treatment of random sums. The final normalization step needed for small scale convergence (following the log-normal derivation) leads to:

$$K(q) = \frac{C_1}{\alpha-1}(q^\alpha - q);\ 0 \leq \alpha \leq 2 \quad (45)$$

(Schertzer and Lovejoy, 1987); for α=1 we have $C_1q\log q$, (see Fig. 23 for simulations). The constant has been written this way so that $K'(1)=C_1$; see below. As a final comment, when $\alpha<2$, and $q<0$, then $K(q)=\infty$; this is a consequence of the extreme Levy tail on the negative (but not positive) fluctuations of $\log\varepsilon$. The possibility (even likelihood) of $\langle\varepsilon_\lambda^q\rangle\rightarrow\infty$ for $q<0$ means that extreme caution should be used when analyzing negative moments of empirical data. While finite data sets will always have finite negative moments their values would sensitively depend on the data and would yield spurious scaling properties.

4.7 Bare/dressed, SOC, divergence of moments

Up until now, we have only considered cascades constructed over a finite range of scales, and de Wijs's log normal argument looks superior to Van Tongergen's argument for hyperbolic behaviour (even with a few caveats about possible log-Levy generalizations). However, in the limit, $\lambda\rightarrow\infty$ it is not obvious that there will be any small scale convergence properties at all since the moments $\lambda^{K(q)}$ generally diverge as $\lambda\rightarrow\infty$. Indeed, to obtain any convergence properties, we must consider integrals over finite sets; "dressed" cascade properties as opposed to the "bare" ones discussed up until now. This corresponds to the right hand side of Fig. 21c. To see this, define the partially dressed "flux":

$$\Pi_\Lambda(B_\lambda) = \int_{B_\lambda} \varepsilon_\Lambda d^D\boldsymbol{x} \quad (46)$$

which represents the spatial integral over a D-dimensional "ball" B_λ of resolution λ of a cascade constructed down to scale ratio Λ. B_λ is a finite scale set of scale ratio λ; e.g. a segment in 1-D, a square or circle in 2-D etc. This allows us to define the dressed flux density:

$$\varepsilon_{d,\lambda} = \lim_{\Lambda\rightarrow\infty} \frac{\Pi_\Lambda(B_\lambda)}{\mathrm{vol}B_\lambda} \quad (47)$$

which is a λ resolution average over a completed cascade. Since measurements are typically made at scales much larger than the true inner scale of the process, $\varepsilon_{d,\lambda}$ corresponds to a typical empirical quantity, whereas the bare ε_λ is purely theoretical. Figure 21c shows that even though $\varepsilon_{d,\lambda}$ takes into account smaller scale cascade steps whereas ε_λ only takes into account the larger scale ones, that most of the time for a fixed λ the bare and the dressed fluxes are roughly the same. Nevertheless, there are a few spikes in the $\varepsilon_{d,\lambda}$ fields which are much stronger than in the corresponding ε_λ. In order to understand this, we can use the factorization property of the cascade, to derive the following relation between $\varepsilon_{d,\lambda}$ and ε_λ.

$$\varepsilon_{d,\lambda} = \varepsilon_\lambda\Pi_\infty(B_1) \quad (48)$$

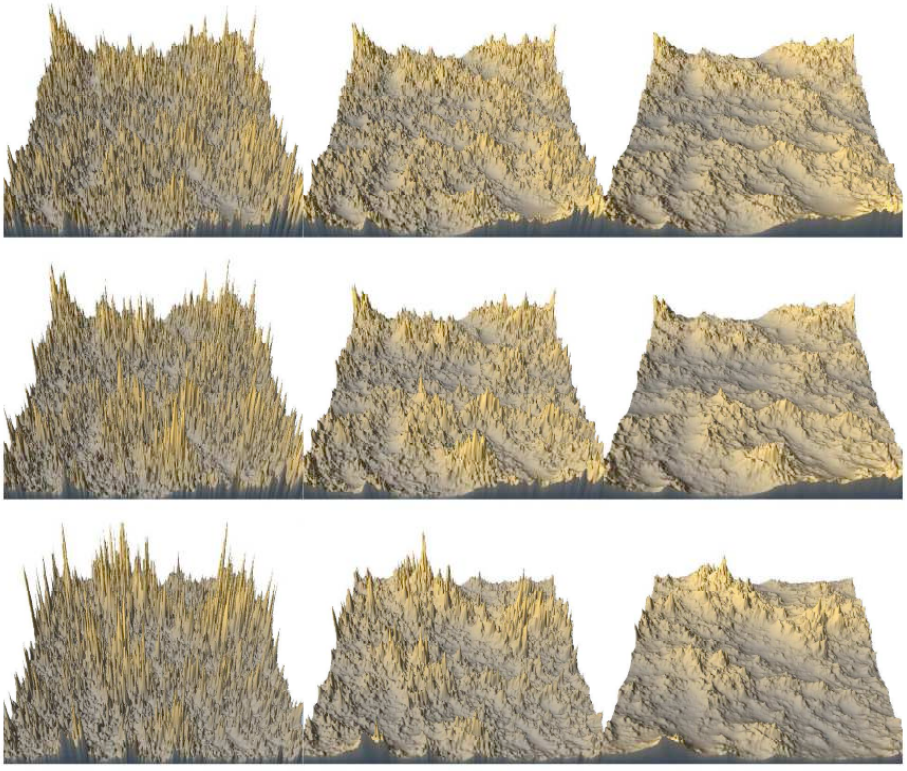

Fig. 23. Isotropic (i.e. self-similar) multifractal simulations showing the effect of varying the parameters α and H (C_1=0.1 in all cases). From left to right, H=0.2, 0.5 and 0.8. From top to bottom, α=1.1, 1.5 and 1.8. As H increases, the fields become smoother and as α decreases, one notices more and more prominent "holes" (i.e. low smooth regions). The realistic values for topography (α=1.79, C_1=0.12, H=0.7) correspond to the two lower right hand simulations. All the simulations have the same random seed. Reproduced from Gagnon et al. (2006).

where the factor $\Pi_\infty(B_1)$ represents the complete spatial integral of a completed cascade. In a generalization of an argument first given by Mandelbrot (1974), it was shown with the help of "trace moments" (Schertzer and Lovejoy, 1987) that:

$$\langle \Pi_\infty(B_1)^q \rangle \rightarrow \infty; \quad q \geq q_D \tag{49}$$

where q_D is a critical order of divergence given by:

$$C(q_D) = D; \quad C(q) = \frac{K(q)}{q-1} \tag{50}$$

where $C(q)$ is a new codimension function (not to be confused with $c(\gamma)$). Using this result, we find:

$$\begin{aligned} &\langle \varepsilon_{d,\lambda}^q \rangle \approx \lambda^{K_d(q)} \\ &K_d(q) = K(q); \quad q < q_D \\ &K_d(q) = \infty; \quad q \geq q_D \end{aligned} \tag{51}$$

where K_d is the dressed moment scaling function.

In order to understand the implications for the probabilities, we can use the link between the statistical moments and the probabilities. Using the method of steepest descents, Parisi and Frisch (1985) showed that for large enough λ, this relation reduces to a Legendre transform between the exponents $K(q)$, $c(\gamma)$:

$$\begin{aligned} K(q) &= \max_\gamma (q\gamma - c(\gamma)) \\ c(\gamma) &= \max_q (q\gamma - K(q)) \end{aligned} \tag{52}$$

Equation (52) (valid for large enough λ) proves that there is a one to one relation between the singularities and moments:

$$q = c'(\gamma); \quad \gamma = K'(q) \tag{53}$$

using the Legendre transform on the dressed $K_d(q)$, we find that the relation between bare and dressed codimensions (and hence probabilities) is:

$$\begin{aligned} &c_d(\gamma) = c(\gamma); \quad \gamma < \gamma_D \\ &c_d(\gamma) = q_D\gamma - K(q_D); \quad \gamma \geq \gamma_D \end{aligned} \tag{54}$$

where γ_D=K$'$(q_D) is the critical singularity contributing to the divergence of moments. In terms of $\varepsilon = \lambda^\gamma$, we find that the extreme behaviour of the ε distribution is:

$$\Pr(\varepsilon_\lambda > s) \approx s^{-q_D}; \quad s \gg 1 \tag{55}$$

i.e. we have the algebraic distributions proposed by Van Tongeren! In other words, a rather minor modification in de

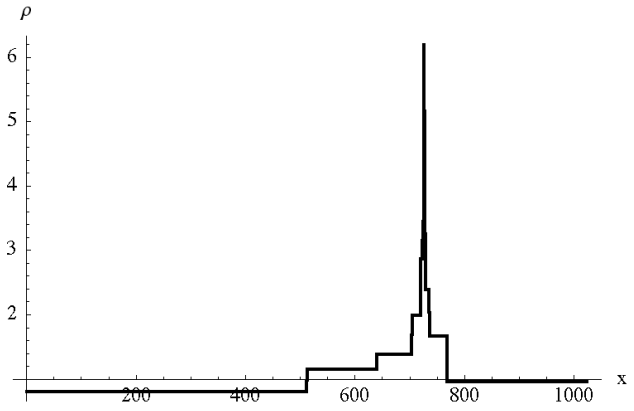

Fig. 24a. The Turcotte cascade for (de Wijs parameters d=0.2 hence $\mu\varepsilon_+$=1.2 hence γ_+=0.27).

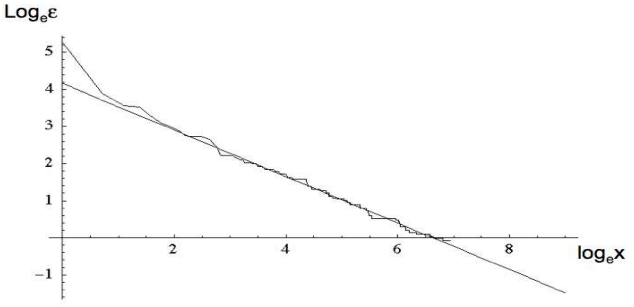

Fig. 24b. The singularity in the Turcotte model with singularity shifted to the origin, x is the distance from the peak, logs in base e (the of mean of 16 simulations with 2^{13} each were used with d=0.5 hence $\mu\varepsilon_+$=1.5 hence γ_+=0.58, c=1 fit slope=0.62 compared to 0.58 theory.

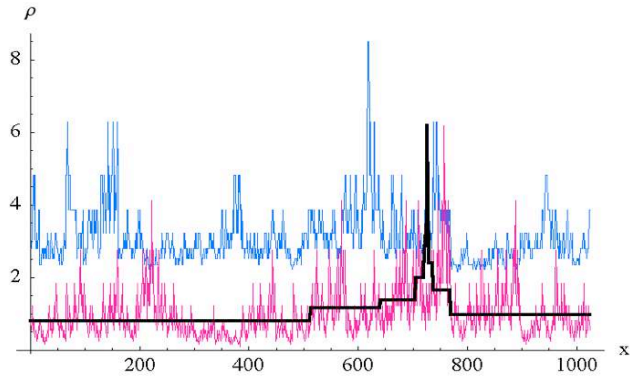

Fig. 24c. Turcotte model (black), binomial/p model (pink), α model (blue) (offset by 2 units for clarity). All models have the same parameters: the de Wijs values: d=0.2 hence $\mu\varepsilon_+$=1.2 (hence γ_+=0.27 for alpha model).

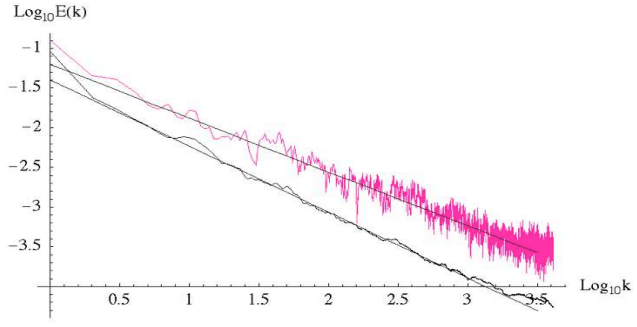

Fig. 24d. Comparisons of ensemble spectra for 16 realizations of the Turcotte (black) model and 16 realizations of the binomial/p model (pink) with $\mu\varepsilon_+$=1.5 along with theoretical slopes (0.83=2(1–log$_2$ $\mu\varepsilon$), 0.678=1–$K(2)$=1–log$_2$ $<\mu\varepsilon^2>$, respectively).

Wijs's model (the use of canonical conservation combined with the consideration of spatial integrals of the process) gives algebraic tails, not the near log-normals that de Wisj promoted.

We can now understand why microcanonical cascades do not display divergence of moments. From Eq. (22) we notice that the largest singularity that a microcanonical cascade can produce in a D dimensional space occurs when all the λ^D "daughter" multipliers have $\mu\varepsilon_i$=0 except for a single one of them which has $\mu\varepsilon_i=\lambda^D$, i.e. it has $\gamma_i=\log_\lambda \mu\varepsilon_i=D$ so that in general we see that $\gamma\leq D$. However, Eq. (54) shows that the singularities which are responsible for the divergence of moments are $\gamma\geq\gamma_D$=K$'(q_D)$=(q_D-1) $C'(q_D)+C(q_D)$. But $C(q_D)=D$ (Eq. 50) and since $C(q)$ is an increasing function, $C'(q)>0$, hence $\gamma_D>D$ i.e. it is beyond the range of microcanonical cascades.

4.8 Turcotte's cascade and divergence of moments

Turcotte (1986, 1989) argued in favour of power law distributions which he termed "fractal" using the symbol "D" rather than q_D for the probability exponent. Rather than relaxing the microcanonical constraint and obtaining "Pareto"/hyperbolic distributions via the above dressing mechanism on canonical cascades, he kept the microcanonical constraint and profoundly modified the cascade mechanism. In these modified cascades, at each step it is only the single most concentrated region ("daughter cell") which participates in the next cascade step. In this way most of the region participates in only a few steps, it is little affected by the cascade. This model is essentially an adaptation of a rock fragmentation model presented in Turcotte (1989) and can produce any $q_D<D$. Figure 24a shows the result of such a cascade. Figure 24c compares it with the binomial/p model and α model with the same parameters. From the construction, around each side of the unique maximum, the ore concentration is a nondecreasing function, in fact, it is very close to the nonrandom singularity $x^{-\gamma+}$ where $\gamma_+=\log\mu\varepsilon/\log\lambda$ (see Fig. 24b); in this model $q_D=\gamma_+$. In fact, the density profile is so close to this nonrandom singularity that the spectral

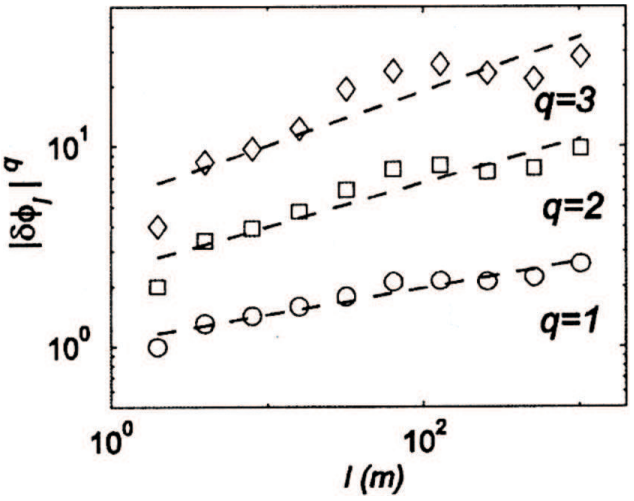

Fig. 25a. Structure function analysis of the first three moments of the neutron porosity estimated from the KTB main borehole. Reproduced from Marsan and Bean (2003).

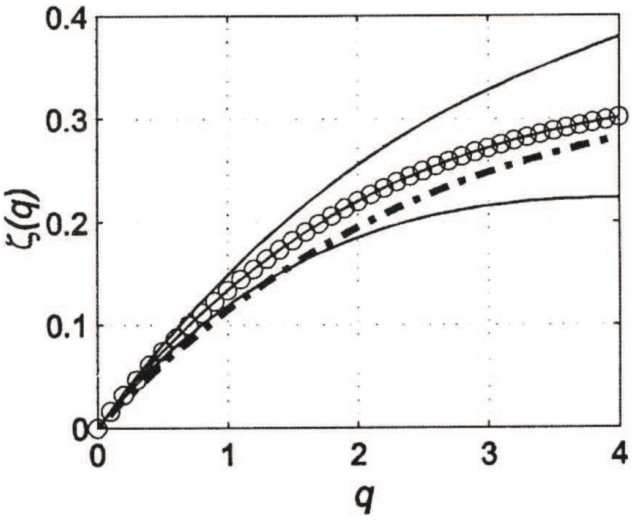

Fig. 25b. ξ(q) from the slopes of Fig. 24a (and for other q values). From the value $\xi(1)=H\approx 0.13$, we see that this is close to the value for the de Wijs Zn ore series Sect. 4.10, and from $\xi'(1)=H-C_1\approx 0.10$ we see that $C_1\approx 0.03$ (see Eq. 58). Also shown is the envelope of the errors (continuous curves, and well as the $\xi(q)$ estimated by the method of Extended Self-Similarity (this plots one of the moments, usually the third against the others, see Benzi et al., 1993). Reproduced from Marsan and Bean (2003).

exponent β is well reproduced by the theoretical value of a randomly positioned singularity $x^{-\gamma+}$:

$$\beta = 2(1-\gamma_+) \tag{56}$$

(see Fig. 24d).

In spite of the term "fractal" it is clear that – contrary to all the cascades and their variants proposed elsewhere in the literature – this is the only cascade whose exceedance sets are all strictly *nonfractal* (a consequence of the fact that on each side of the unique maximum, it is nondecreasing). Even ignoring the unrealistic nonfractal spatial variability of the model, even as a model for power law probabilities it is not so attractive since it is limited to $q_D<D$ and many examples exist with $q_D>D$ which are therefore outside of its range. The dressed mechanism presented above has the advantage of being quite general and is also more flexible; it can easily have $q_D>D$. As a final note on the debate on log-normal versus algebraic distributions, we could mention that Mandelbrot (1995) suggests that ore concentrations and distributions of other geological quantities, are in fact Levy distributions (i.e. with exponent restricted to values $q_D<2$, not log-Levy); hence presumably it is the result of additive rather than multiplicative cascade processes.

4.9 Observables and Nonconservative multifractals

The results of behavior described by Eq. (19) is called "multiscaling" because each statistical moment is scaling with a different exponent; it is the generic result of a scale by scale conservative multiplicative cascade. However, there is no reason to assume a priori that ore concentrations should be the direct result of a multiplicative cascade. The classic example is turbulence where it is the energy flux which is conserved by the nonlinear terms of the dynamical equations; it is thus the cascaded quantity. This leads to the famous Kolmogorov (1941b) law for isotropic turbulence which relates the energy flux to velocity gradients (Δv) as follows:

$$\Delta v_\lambda = \varepsilon_\lambda^a \lambda^{-H}; \quad H = 1/3; \quad a = 1/3 \tag{57}$$

The usual interpretation of this equation is that the equality is in the sense of scaling laws so that, taking the q-th powers of both sides and ensemble averaging, we obtain:

$$\langle \Delta v_\lambda^q \rangle = \langle \varepsilon_\lambda^{q/3} \rangle \lambda^{-qH} = \lambda^{\xi(q)}; \quad \xi(q) = Hq - K(q); \quad K(q) = K_\varepsilon(aq) \tag{58}$$

We can see that the typical observables have an extra linear scaling term Hq and where the (generalized) structure function exponent $\xi(q)$ from Sect. 2, Eq. (11) has been used. H thus characterizes the distance from the (conserved) pure multiplicative process ε; it is the degree of non (scale by scale) conservation of the process. Note that since the power spectrum is the Fourier transform of the autocorrelation function, we have the following relation with the spectral exponent:

$$\beta = 1 + 2H - K(2) \tag{59}$$

In Fig. 25a, we show an example of the multiple scaling of various moments of the KTB porosity estimates taken from Marsan and Bean (2003), other examples of more classical semi variograms (q=2) are more common, see e.g. Cheng (2000a) for an application to Cu concentrations.

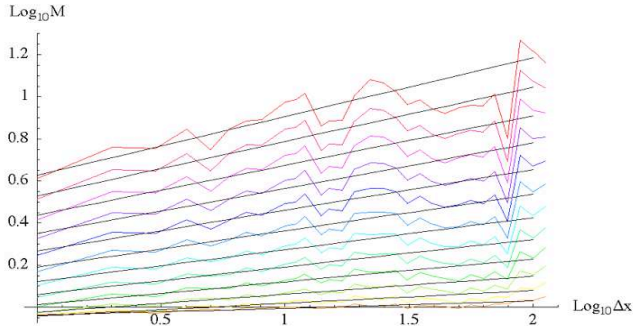

Fig. 26a. Log-log plot of the first order structure function: for the de Wijs data for q=0.25, 0.5, . . . 3.

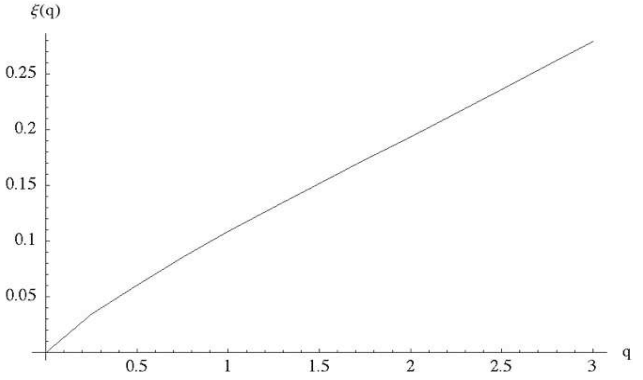

Fig. 26b. The slopes of the structure functions, $\xi(q)$. The slope and value near q=1 yield H=0.090, C_1=0.018.

4.10 Multifractal analysis: the example of the de Wisj data

In order to demonstrate the various ideas discussed above, we analyzed the original de Wijs data set: 118 horizontally spaced values of Zinc concentration in the Pulacayo Mine, Bolivia, spaced at 2 m intervals (data shown in Fig. 3a, the spectrum Fig. 3b). Although this spatial series is very small for our purposes, it is freely available (in the de Wijs paper) and has the interest that it has been re-analyzed by many authors since including Matheron (1962), Agterberg (1974), Cheng and Agterberg (1996), Cheng (1997, 2000b).

To begin, we note that the spectrum (Fig. 3b) has $\beta \approx 1.12$ which is consistent with a small value of $K(2)$ and H. The first step is to calculate the structure functions (Fig. 26a); we note that the scaling is quite good considering the smallness of the sample size. On the right (Fig. 26b), we see the structure function exponent $\xi(q)$; it is surprisingly linear. From the slopes at q=1, we can estimate the value of H=0.090, and C_1=0.018 ($H = \xi(1)$; $H - C_1 = \xi'(1)$; Eq. 58). The relatively small value of C_1 with respect to H indicates that the multifractality is weak enough that the deviation from conservation (H) will be dominant except for quite high moments. This means that pure multiplicative models will not be too accurate for the lower order moments (they assume H=0).

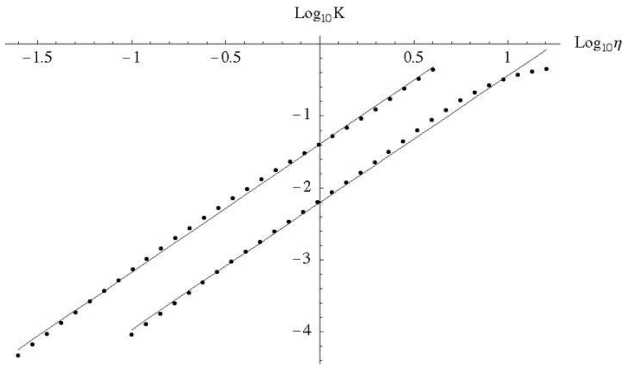

Fig. 27. Double Trace Moment analysis of the de Wijs Zn data q=2, 0.5 (left, right), slopes, α=1.76, 1.78, C_1=0.023, 0.022, respectively.

It is worth noting that we have used the usual structure function, based on the statistics of fluctuations defined as the differences in the concentration. Defining the fluctuations as differences in this way is sometimes called the "Haar" or "poor man's wavelet"; other choices of definition are possible; wavelets provide a systematic framework for this (see e.g. Holschneider, 1995). However even in 1-D there are an infinite number of possible wavelets and there is usually no compelling reason to use one rather than another. In practice, the use of differences is usually adequate, the main restriction being that it is only appropriate when $0 \leq H \leq 1$, a condition which is usually (although not always) satisfied in geophysical applications (here we found $H \approx 0.090$). For example, when $H > 1$, one must measure fluctuations with respect to a local linear trend; this can be done either by fractionally differentiating the process (power law filtering, Schertzer and Lovejoy, 1987), using appropriate wavelets (Bacry et al., 1989) or using the "Multifractal Detrended Fluctuation Analysis" technique (Kantelhart et al., 2002; see Telesca and Lapenna, 2005) for application to self-potential fields associated with seismic areas. Note that using the Wiener-Khintchin theorem, we obtain a simple relation between the second order structure function exponent and the spectral exponent: $\beta = 1 + \xi(2)$; this is indeed approximately verified. Finally, we could mention another scaling analysis technique which has been used on occasion in geophysics; the rescaled range method ("R/S analysis"), Mandelbrot and Wallis (1969) (applied for example to borehole thermal conductivities in Dimri and Vedanti, 2005). The difficulty with this method is not so much that it determines only a single exponent, but more that the unique value is only easily interpretable for quasi-gaussian processes. While the R/S exponent is also denoted "H" in honour of Hurst (1951), it is not generally the same as the H discussed here; indeed as far as we know, its relation to the basic parameters of a multifractal process are not at present known (Schmitt et al., 1995).

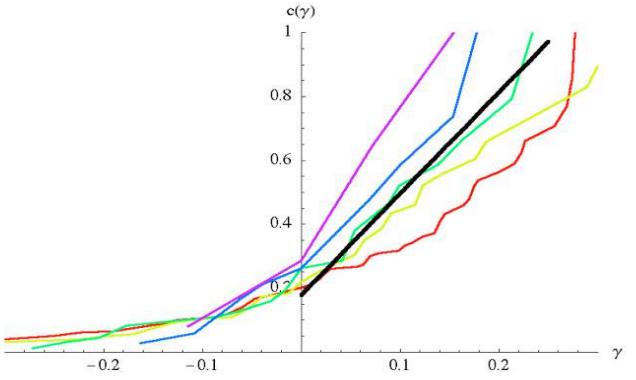

Fig. 28. $c(\gamma)$ for fluxes estimated from the de Wijs Zn data at resolutions of 2, 4, 8, 16, 32 m reference slope is 3.2.

We can see that (as expected) $\xi(q)$ is concave downwards. We have already quantified the behaviour of $\xi(q)$ near the mean (q=1; we used $\xi'(1)=H-K'(1)$ where $C_1=K'(1)$ is the "codimension of the mean"); it remains to characterize the remaining nonlinear part $K(q)$. Although this value appears small, typical values in turbulence are only a bit larger e.g. $C_1\approx 0.07$ for the horizontal wind in the horizontal (Schmitt et al., 1992, 1994), and $C_1\approx 0.04$ for passive scalars in the horizontal (Lilley et al., 2004). The corresponding values in the vertical are about $1/H_z$=9/5 times larger; this is as predicted by the 23/9D model of scaling stratification (Schertzer and Lovejoy, 1985a). Finally, topography has $C_1\approx 0.12$ (pretty much the same for both continents and oceans). For an early review of these and other results, see Lovejoy and Schertzer (1995).

To more fully characterize the $\xi(q)$ function, we can test whether it belongs to a universality class and attempt to estimate the Levy index α of the generator (see above). One convenient way is to use the Double Trace Moment technique (DTM, Lavallée et al., 1993). We have seen that the generic statistical properties of processes which are repeated scale after scale are characterized by a nonlinear exponent $K(q)$, and that the observables will generally have an extra linear scaling term qH. Since at least for low q; the linear term qH is often larger than the nonlinear $K(q)$, in analyses, it can mask the latter. It is therefore advantageous to first estimate the conserved flux ε from the observed v, and the estimate $K(q)$ directly. From Eq. (57), we see that in principle, this can be done by removing the λ^{-H} scaling. To do this, note that if we start with a field ε^q and fractionally integrate it by H; (a power law filter k^{-H}, see Sect. 5), that the resulting field will have the fluctuation statistics indicated by Eq. (58) (see Marsan et al., 1996). This suggests that in order to obtain a flux from v (i.e. a conserved field with H=0), that it suffices to invert the power law filter, i.e. to fractionally differentiate it by an order H. It turns out that a finite difference approximation to an integer order $H'>H$ differentiation followed by taking the absolute value of the result is typically sufficient (the absolute value is necessary since the multiplicative cascade flux respects $\varepsilon>0$; see however for complex and vector cascades Schertzer and Lovejoy, 1995). Since usually $0\leq H\leq 1$, a first order finite difference is typically sufficient. One simply takes the absolute differences at the finest available resolution ε_Λ and degrades them:

$$\varepsilon_\lambda = (\varepsilon_\Lambda)_\lambda; \quad \varepsilon_\Lambda \approx |\Delta\rho_\Lambda| \tag{60}$$

where again the notation $(\varepsilon_\Lambda)_\lambda$ indicates the average of the finest resolution data ε_Λ over the intermediate resolution λ. Once ε has been estimated by the absolute value of the differences at the finest resolution, the DTM uses the following:

$$\langle (\varepsilon_\Lambda^\eta)_\lambda^q \rangle = \lambda^{K(q,\eta)} \tag{61}$$

i.e. one degrades the η power of the flux at the highest resolution Λ down to an intermediate resolution λ and then determines the scaling of the qth power result. The advantage of this method is that the new exponent $K(q,\eta)$ is related to $K(q)$=$K(q,1)$ as:

$$K(q,\eta) = K(q\eta) - qK(\eta) \tag{62}$$

so that for universal multifractals, (c.f. Eq. 45) we have:

$$K(q,\eta) = \eta^\alpha K(q,1) \tag{63}$$

which is a convenient power law of η. Figure 26 shows the result for two values of q with varying η. We find q=2, 0.5 (left, right), the slopes α=1.76, 1.78, respectively, and from the intercepts, C_1=0.023, 0.022, respectively (close to the estimate above C_1=0.018 from the structure function). Our structure function and DTM analyses have thus shown that contrary to the previous analyses, that the ore concentration process is neither purely multiplicative, nor is its generator a Gaussian.

Finally, we can estimate the $c(\gamma)$ function and check the behaviour of the extremes (are these hyperbolic as expected for extreme enough events?). A simple way to estimate $c(\gamma)$ is simply:

$$c(\gamma) \approx -\mathrm{Log}_\lambda \varepsilon_\lambda \tag{64}$$

(see Eq. 34). In this approximation, we ignore the (slowly varying) prefactors. Lavallée et al. (1991) discuss this Probability Distribution Multiple Scaling PDMS technique as well as ways of improving the approximation Eq. (64). In Fig. 28 we do this for the fluxes estimated at resolutions of 1, 2, 4, 8, 16 units. Due to the very small sample, the results are variable for larger γ. However, we see a hint of linear behaviour at large γ (black line, absolute slope = q_D=3.2). Certainly if the origin of the divergence of moments is the "dressing" of the ore over a set of dimension 1 discussed above, then for α near 2, we have $C(q)\approx C_1 q$ (see Eq. 50), so that the solution to $C(q_D)=D$ is simply $q_D\approx D/C_1$. In this case, since C_1 is small, we expect q_D to be too large to be observed except perhaps on huge data sets. However caution should be used

in applying Eq. (50) to determine q_D: experience in turbulent systems has shown that q_D can be sufficiently small so that – if dressing is the correct explanation – it would imply that D is less than one; in turbulence a D of the order 0.3 would be necessary (see Lazarev et al., 1994). At this stage in our undertanding of scaling processes, it is perhaps best to simply regard q_D as an empirical parameter characterizing the system.

4.11 Other multifractal formalisms: singular densities (γ, c(γ)) versus singular measures (α, f(α))

At the same time as multifractal cascades were being developed in turbulence, they were also being developed for applications in strange attractors and chaos. In strange attractors, one has a large number of points N which are generated either by mappings or flows with a finite (usually small) number of degrees of freedom (equal to the dimension of the phase space, denoted D). The density of points in the limit of the number of points tending to infinity is interpreted as being a realization of a multifractal probability measure. The probability density p_λ at resolution λ is estimated by the "partition function method" covering the space with a λ^{-D} sized grid (boxes) and using $p_\lambda = n_\lambda/N$ for each box (n_λ is the number of points in the box). The probability measure is a geometric multifractal since although it represents the probability of finding the system at a point in the phase space, it is not itself random at all!

Halsey et al. (1986) wrote an influential paper proposing a notation for dealing with these "geometric attractor" multifractals. Rather than considering the density of the multifractal measure p_λ (the non-random analogue of the cascade ε_λ), they considered the measure itself integrated over a ball (box) size λ^{-1} i.e. B_λ.

We use the symbol $\Pi_\infty(B_\lambda)$ since it is the analogue of the energy flux Π_∞ and the volume of B_λ is λ^{-D}. The p_λ defined this way is really a dressed quantity, but for these multifractals the bare/dressed distinction is irrelevant. Halsey et al. (1986) then defined the order of singularity α_D of the flux rather than the density of the flux:

$$\Pi_\infty(B_\lambda) = \int_{B_\lambda} p\, d^D \boldsymbol{x} = p_\lambda\, \lambda^{-D} \sim \lambda^{-\alpha_D} \tag{65}$$

(the subscript "D" was not used in the original; we have added it to underscore the dependence on the dimension of the system). In cascade/turbulence notation, we may write $p_\lambda \sim \lambda^\gamma$; we thus obtain:

$$\alpha_D = D - \gamma \tag{66}$$

Each box can thus be indexed according to α_D. The number of boxes at each resolution corresponding to α_D can then be used to define the (box counting) dimension $f_D(\alpha_D)$:

$$\text{Number}\left[\Pi_\infty(B_\lambda) = \lambda^{-\alpha_D}\right] \sim \lambda^{f_D(\alpha_D)} \tag{67}$$

since "Number"=λ^D probability, we obtain:

$$f_D(\alpha_D) = D - c(\gamma);\ \alpha_D = D - \gamma \tag{68}$$

Finally, the partition functions can be used to define scaling exponents for the moments:

$$\sum_A \Pi_\infty{}^q(B_{\lambda,i}) \sim \lambda^{-\tau_D(q)} \tag{69}$$

where the sum is over all the i balls $B_{\lambda,i}$ size λ^{-1} needed to cover the D-dimensional phase space region A. Since p is non random, we have:

$$\left\langle \sum_A (p_\lambda\, \lambda^{-D})^q \right\rangle = \lambda^D \lambda^{-qD} \left\langle p_\lambda^q \right\rangle = \lambda^{-(q-1)D+K(q)} = \lambda^{-\tau_D(q)} \tag{70}$$

therefore we obtain:

$$\tau_D(q) = (q-1)D - K(q) = (q-1)D(q);\quad D(q) = D - C(q) \tag{71}$$

As long as we deal with strange attractors and study the full D-dimensional phase space, the α_D, $f_D(\alpha_D)$, $\tau_D(q)$ (and $D(q)$) notation is adequate. However, if we are interested in random multifractals (involving probability spaces; $D \to \infty$), or if we are interested in looking at subspaces with dimension smaller than D, the D dependence is respectively a fundamental limitation or an unnecessary complication. The turbulent γ, $c(\gamma)$, $K(q)$ and $C(q)$ notation always has the advantage of being intrinsic to the process (it is D independent).

5 Models

5.1 Singularities and morphology

In Sect. 2 we discussed the fact that the horizontal and vertical scaling in the lithosphere has different exponents and that this is associated with scale dependent vertical stratification. In Sect. 4, we discussed the fact that in general scaling is characterized by an infinite hierarchy of exponents (the $K(q)$ or $c(\gamma)$ function) and that this can be modeled with multiplicative cascades. However, we presented only unrealistic (discrete scale ratio) cascades and indicated that typical observables are not scale by scale conserved, that they have an extra qH in their moment scaling exponent (Eq. 58). In this section, we briefly discuss how to make continuous in scale and nonconservative ($H > 0$) multifractal processes, and we compare these with other scaling models.

It is natural to model scaling processes using combinations of scale invariant basis functions i.e. mathematical singularities. For the topography, an early model is the Turcotte and Oxburgh (1967) model for the variation of altitude as a function of distance from mid-ocean ridges, mathematically the form is indicated in Table 2. Mandelbrot (1975) proposed

Table 2. An intercomparison between various models of the topography showing the essential similarities and differences in their mathematical structure, statistical properties. Here D=2 for horizontal planes and the dimension D_F is the fractal dimension of lines of constant altitude in the horizontal. The deterministic mid-ocean ridge model is represented here by a fault in unit direction vector $\boldsymbol{r}$ through the point $\boldsymbol{x}_0$. Here the variables are nondimensionalized and the height of the fault is normalized to one. Note that δ is a Dirac delta function. The model of Turcotte and Oxburgh uses H=1/2. The monofractal fractional Brownian motion (fBm) model involves a fractional integration of order H' with a flux $\phi_2(\boldsymbol{x})$ which is simply a ("δ correlated") gaussian white noise with variance σ^2. Note that the symbol $a=^d b$ indicates equality in probability distributions, i.e. $a=^d b \Leftrightarrow \Pr(b>s)=\Pr(a>s)$ for all s, "Pr" indicates "probability". It results in altitude fluctuations with gaussian statistics, linear structure function exponent $\xi(q)$ and altitude independent surface codimension c (or dimension D_F). The value H=1/2 is compatible with the commonly cited value D_F=1.5 for the level sets or D=1+1.5=2.5 for the dimension of points on the surface. The fLm is the generalization obtained by replacing Gaussian variables by stable Levy variables with index α (fBm is obtained in the case α=2). These have diverging moments q for $q \geq \alpha$. Finally, the multifractal Fractionally Integrated Flux (FIF) model has the same structure, except that the white noises are replaced by multifractal noises ϕ_λ where λ is the resolution. The multifractal noise ϕ_λ is the result of a continuous in scale multiplicative cascade, mathematically it is given by

$$\Gamma_\lambda(\boldsymbol{x} = \int_{|}^{\lambda} \frac{\gamma_\alpha(\boldsymbol{x}')d\boldsymbol{x}'}{|\boldsymbol{x}-\boldsymbol{x}'|D-H}; \quad \phi_\lambda(\boldsymbol{x} = e^{\Gamma_\lambda(\boldsymbol{x})}; \quad H' = D(1-1/\alpha)$$

$\gamma_\alpha(\boldsymbol{x})$ = independent Levy noise, index α, amplitude depends on C_1, Γ_λ is an fLm process called the "generator". It is multiplicative because of the exponentiation of the additive process Γ: ϕ=e$^\Gamma$. We again find $H \approx 1/2$, although now there are an infinite number of codimensions $c(\gamma)$ (or dimensions $D_F(\gamma)$) that depends on the threshold given by λ^γ. (γ is an order of singularity; not to be confused with the subgenerator γ_α). In all cases, H can in principle be determined by dimensional analysis so that the Turcotte- Oxburgh exponent H=1/2 may be valid for all the models, c.f. Lovejoy (1995). To generalize fBm, fLm and FIF to anisotropic topographies, we must replace the distances in the fractional integration denominators by anisotropic scale functions as discussed in the text. Reproduced from Gagnon et al. (2006).

Model	Altitude (and noise statistics)	Altitude increments (and statistics)	Codimension (c) of level sets
Mid-ocean ridge (deterministic)	$h(\mathbf{x})=1-\int d\mathbf{x}' \frac{\delta(\mathbf{x}'-\mathbf{x}_0)}{\lvert\mathbf{r}\cdot(\mathbf{x}-\mathbf{x}')\rvert^{-H}}$ (No noise statistics)	$\Delta h \propto \lvert\Delta\mathbf{x}\rvert^H$ (No altitude statistics)	$c=D-D_F$ $D_F=1$
Monofractal fBm (stochastic)	$h(\mathbf{x})=\int d\mathbf{x}' \frac{\phi_2(\mathbf{x}')}{\lvert\mathbf{x}-\mathbf{x}'\rvert^{D-H'}}$ $\phi_2(\mathbf{x})$=Gaussian white noise $H'=H+D/2$	$\Delta h \stackrel{d}{=} \phi_2\lvert\Delta\mathbf{x}\rvert^H$ $\langle\lvert\Delta h\rvert^q\rangle \propto \lvert\Delta\mathbf{x}\rvert^{\xi(q)}$ $\xi(q)=qH$	$c=H$ $D_F=D-c$
Monofractal fLm (stochastic)	$h(\mathbf{x})=\int d\mathbf{x}' \frac{\phi_\alpha(\mathbf{x}')}{\lvert\mathbf{x}-\mathbf{x}'\rvert^{D-H'}}$ $\phi_\alpha(\mathbf{x})$=Levy noise ($0\leq\alpha\leq 2$) $H'=H+D/\alpha$	$\Delta h \stackrel{d}{=} \phi_\alpha\lvert\Delta\mathbf{x}\rvert^H$ $\langle\lvert\Delta h\rvert^q\rangle \propto \lvert\Delta\mathbf{x}\rvert^{\xi(q)}$ $\xi(q)=\begin{cases} qH & \text{for } q<\alpha \\ \infty & \text{for } q\geq\alpha \end{cases}$	$c=H$ $D_F=D-c$
Multifractal FIF (stochastic)	$h_\lambda(\mathbf{x})=\int d\mathbf{x}' \frac{\phi_\lambda(\mathbf{x}')}{\lvert\mathbf{x}-\mathbf{x}'\rvert^{D-H}}$ $\langle\phi_\lambda^q\rangle=\lambda^{K(q)}$	$\Delta h=\phi_\lambda\lvert\Delta\mathbf{x}\rvert^H$ $\langle\lvert\Delta h\rvert^q\rangle \propto \lvert\Delta\mathbf{x}\rvert^{\xi(q)}$ $\xi(q)=qH-K(q)$	$c(\gamma)=\max_q\{qH-K(q)\}$ $D_F(\gamma)=D-c(\gamma)$

a model based on the idea of making singular faults the basic shapes by summing large numbers of faults with random centers and orientations with Gaussian amplitudes; he produced a Gaussian process with long range (power law) correlations. Due to the central limit theorem (the gaussian special case), a process with the same statistical properties can be produced by using singularities of a quite different shape; Table 2 (second row) indicates a model with point rather than line singularities; in this form the mathematics is more convenient for comparison with the other singular topography models summarized in Table 2. In this case, in the limit of many faults, because all of the singularities have nearly the

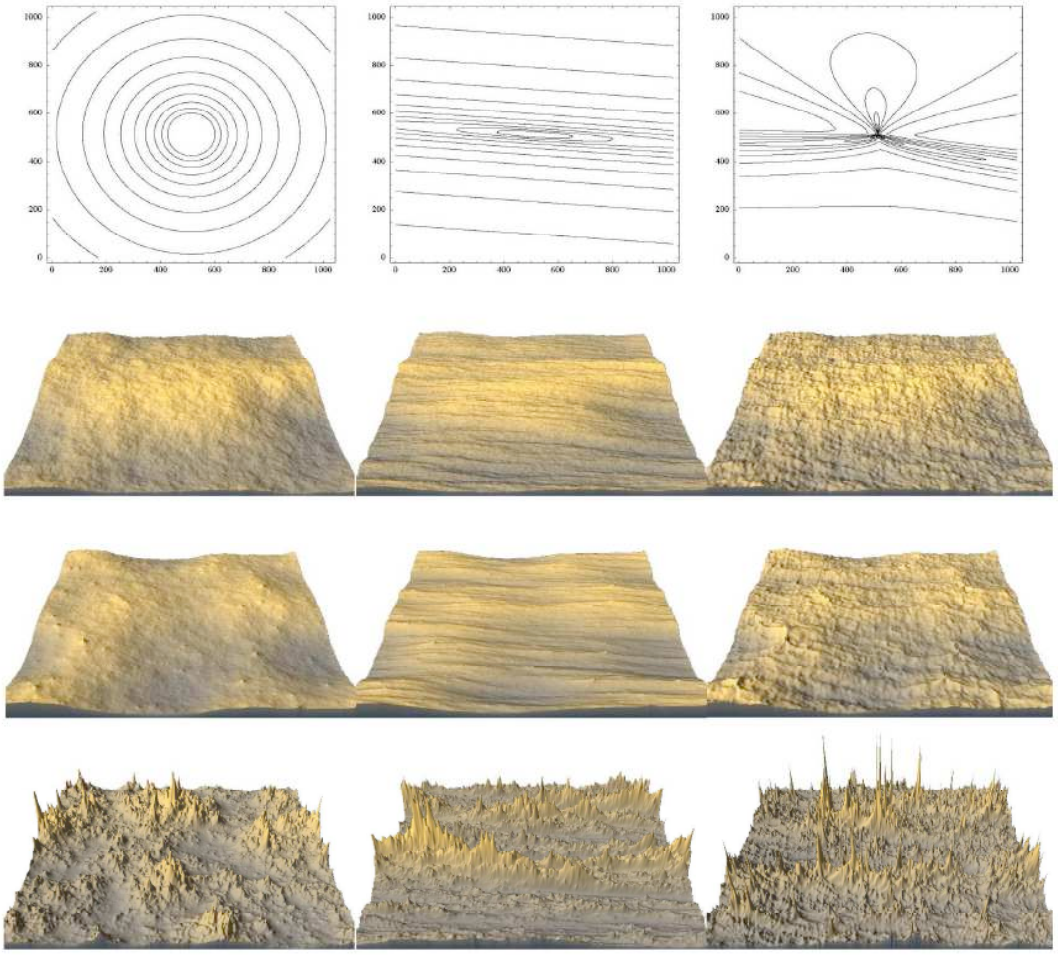

Fig. 29. Comparison of isotropic versus anisotropic (with symmetric scale functions) simulations for three different scaling models. Top row shows scale functions. From left to right, we change the anisotropy: the left column is self-similar (isotropic) while the middle and right columns are anisotropic and symmetric with respect to $\mathbf{G}=\begin{pmatrix} 0.8 & -0.05 \\ 0.05 & 1.2 \end{pmatrix}$. The middle column has unit ball circular at 1 pixel, while for the right one it has the form $\Theta(\theta')=1+0.65\cos\theta'$ (in polar coordinates in the nonlinearly transformed space, see Eq. 74). Second, third and fourth rows show the corresponding fBm (with H=0.7), fLm (α=1.8, H=0.7) and multifractal (α=1.8, C_1=0.12, H=0.7) simulations. We note that in the case of fBm, one mainly perceives textures, there are no very extreme mountains or other morphologies evident. One can see that the fLm is too extreme, the shape of the singularity (particularly visible in the far right) is quite visible in the highest mountain shapes. The multifractal simulations are more realistic in that there is a more subtle hierarchy of mountains. When the contour lines of the scale functions are close, we change the scale $\|\boldsymbol{r}\|=\lambda$ rapidly over short (Euclidean) distances. For a given order of singularity γ, λ^{γ} will therefore be larger. This explains the strong variability depending on direction (middle bottom row) and on shape of unit ball (right bottom row). Indeed, spectral exponents will be different along the different eigenvectors of **G**.

same amplitude (Gaussian variables are rarely more than a few standard deviations from the mean), the basic singularity shape is not important, we end up a rough texture but without any more interesting morphologies.

In Table 2 we see that all the stochastic models are obtained by convolutions with singularities, such convolutions are "fractional integrations" of order H' (if $H'<0$, there are differentiations; the difference between H, H' for fractional Brownian motion (fBm), fractional Levy motion (fLm) are necessary to take into account the scaling of the basic noises ϕ_2, ϕ_α). The lesson from fBm is that if we are to explain real topography by such a singular model, then the statistics of the singularities must be more extreme than gaussian so that the basic singularity shape may remain important in the limit of a large number of large singularities (i.e. after integration over the noise). One way to make some of the singularities always stand out is to use the fractional Levy motion model obtain by replacing the Gaussian noise by a Levy noise index α. The Levy random variables can be regarded as a generalization of the Gaussian variables to the case where the variance (second

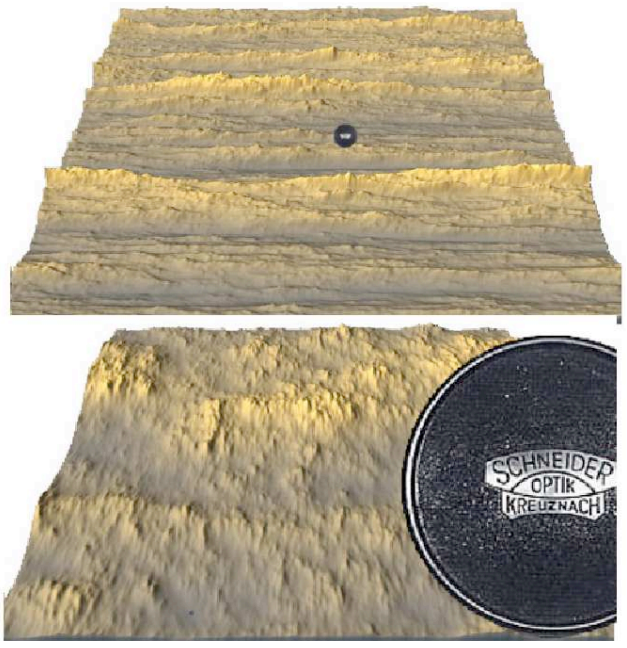

Fig. 30. This self-affine simulation illustrates the "phenomenological fallacy" since both the top and bottom look quite different while having the same generators (**G** is diagonal with elements 0.8, 1.2), same (anisotropic) statistics at scales differing by a factor of 64 (top and bottom blow-up). The figure shows the proverbial geologists' lens cap at two resolutions differing by a factor of 64. Seen from afar (top), the structures seem to be composed of left to right ridges, however closer inspection (bottom) shows that in fact this is not the case at the smaller scales (adapted from Lovejoy and Schertzer, 2007).

moment) is infinite; they have long probability tails such the statistical moments q order $q \geq \alpha$ and higher diverge. Due to the (generalized) central limit theorem, sums of independent (possibly weighted) Levy variables are still Levy variables. Figure 29 shows a comparison with the corresponding fBm; several strong mountain peaks stand out; in fact, the strong peaks are too strong – although far from Gaussian – real topography empirically seems to have finite variance so this cannot be a good model. Finally, we note that the continuous in scale conservative multifractal process is obtained by using an additive Levy process for the log, and the extra qH needed to obtain the nonconservative multifractal process is modeled by an extra fractional integration order H; Fig. 23 in Sect. 4.7 showed the effect of varying the H, α parameters.

5.2 Modelling, Anisotropic fractals, multifractals, the scale function

Let's consider the singularity shape in more detail. The shape of line (fault-like) and point singularities depends on powers of distances from either a line or a point; in order to generalize this it turns out to be sufficient to replace the standard Euclidean distances by scale functions. We therefore digress a moment to discuss scale functions. In order to change the shape of the singularities while conserving the basic statistical properties of the process, it turns out to be sufficient to make the replacement everywhere in Table 2:

$$|\Delta \boldsymbol{r}| \rightarrow \|\Delta \boldsymbol{r}\| \; ; \; D \rightarrow D_{el} \tag{72}$$

i.e. to replace the usual distance ("||") by a "scale function" ("‖‖") and usual dimension of space by an "elliptical dimension" D_{el} which satisfies the following basic equation scaling:

$$\|T_\lambda \boldsymbol{r}\| = \lambda^{-1} \|\boldsymbol{r}\| \; ; \; T_\lambda = \lambda^{-\mathbf{G}}; \; D_{el} = Trace\mathbf{G} \tag{73}$$

where T_λ is a scale changing operator which reduces the scale of a vector by a factor λ. In order for the scale function to be scaling (i.e. have no characteristic scale), it must satisfy group properties, hence it must admit a generator **G** as indicated. Once all the unit vectors $\boldsymbol{r}_1$ are specified the scale Eq. (73) uniquely specifies the scale of all vectors; all the nonunit vectors ($\|\boldsymbol{r}_\lambda\| = \lambda$; $\lambda \neq 1$) are then generated by the action of T_λ: $\boldsymbol{r}_\lambda = T_\lambda \boldsymbol{r}_1$ (see Schertzer and Lovejoy, 1985b, for technical details on this Generalized Scale Invariance, GSI). The set of all vectors $\|\boldsymbol{r}\| \leq \lambda$ is called a "ball", denoted B_λ; for physical scale functions, B_λ must be strictly decreasing (i.e. $B_{\lambda'} \subseteq B_\lambda$; $\lambda' < \lambda$). We can see that if the replacements $|\boldsymbol{r} - \boldsymbol{r}'| \rightarrow \|\boldsymbol{r} - \boldsymbol{r}'\|$; $D \rightarrow D_{el}$ are made in the denominators of the models in Table 2, with scale functions satisfying the scale Eq. (72) (in fact they then define the notion scale) then the convolutions will have power law dependencies under "zooming", i.e. the models will be scaling as long as the noises are also scaling (hence the special choices of Gaussian or Levy noise, or in the multifractal case, multifractal noise).

When scale functions are used as the basic singularities, the shapes can be extremely varied, hence demonstrating the possibility of modeling geomorphologies in this way. First consider **G** = the identity: the resulting models will be "self-similar" in the sense that their statistics will vary in power law ways under isotropic "zooming" (blow-ups). When the unit ball is a circle (or more generally a D dimensional sphere), then we obtain $\|\boldsymbol{r}\| = |\boldsymbol{r}|$. However when the unit ball is not circular (spherical), then there will still be preferred directions. These preferred directions will be the same at all scales, the anisotropy is "trivial" (see Fig. 23 for examples). Things become more interesting as soon as **G** is no longer the identity. If **G** is a diagonal matrix, then the singularities order γ: $\|\boldsymbol{r}\|^{-\gamma}$ are quite different in different directions, the resulting fractals/multifractals are "self-affine". The case where **G** is nondiagonal and the eigenvalues are real is a generalization in which the main stretching/shrinking occurs along nonorthogonal eigendirections; Fig. 29, 30 shows the resulting differential stratification. When the eigenvalues are complex, then the eigenvectors rotate continuously as functions of scale. Finally,

we can consider noncircular/nonspherical unit balls, Fig. 29 shows how the basic singularity shapes are clearly visible in the fLm. In the multifractal case, the effect of the singularity morphology is still important for the result; but things are more subtle. Figure 29 shows the effect of changing the scale function while maintaining the **G**. Finally – outside our present scope but presumably important for realistic topography modelling – we can consider **G** as a nonlinear operator (rather than a matrix). In this case, the anisotropy depends not only on scale but also on the location. This allows for spatially varying morphologies. In this case, the linear GSI discussed above in simply a local approximation.

To understand the relation between usual distances and generalized scales, consider a (real) 2-D **G** matrix which in a diagonal frame is: $\mathbf{G}=\begin{pmatrix} H_x & 0 \\ 0 & H_y \end{pmatrix}$; the nonlinear coordinate transformation $x'=sgnx\,|x|^{1/H_x}$; $y'=sgny\,|y|^{1/H_y}$ transforms the problem into a problem with $\mathbf{G}'=1$= identity; this shows that the basic scale function is:

$$||(x, y)|| = r(\theta')(x^{2/H_x} + y^{2/H_y})^{1/2} \quad (74)$$

where $r(\theta')$ is an arbitrary function of the polar angle θ' (in the nonlinearly transformed space i.e tan $\theta'=y'/x'$). Figure 29 shows the effect of varying the unit balls and G matrices for various topography models.

5.3 The phenomological fallacy

Geophysicists commonly derive their models from phenomenological classifications based largely on classical (scale bound) notions of scale and shape. Once a phenomenon has been defined - often involving somewhat subjective criterion – models are constructed to explain them. However we have seen that scaling processes – if based on sufficiently strong anisotropic singularities, can lead to quite different looking structures at different scales even though the basic underlying mechanism is scale invariant; see Fig. 30 for an example. This possibility demonstrates what we call the "phenomological fallacy" i.e. the danger of inferring process from appearance.

6 Conclusions

Central problems in the geosciences are those of resolution and scale. It is quite typical for fields and structures to have variability ranging over factors of over 10^{10} in scale. Starting in the early 20th century, there were isolated insights into the nature of such wide range variability. However, it wasn't until the 1970s that new ideas of deterministic chaos and fractals began to spark wide interest in the problem. Deterministic chaos is essentially a low number of degrees of freedom paradigm whereas (stochastic) fractals provide an attractive large number of degrees of freedom alternative. In the original form of fractal geometry of self-similar (and occasionally self-affine) sets (Mandelbrot, 1977, 1983), fractals were arguably more inspirational than practical. However in the 1980s two developments were made which were essential for geoscience applications: the generalization from scaling geometric sets to scaling fields (multifractals), and the generalization from isotropic (self-similar) to quite general anisotropic scaling ("Generalized Scale Invariance").

Partly in order to limit our scope, and partly because of their fundamental importance, in this review we concentrated on scaling geofields in the solid earth and topography. Using the (somewhat) familiar method of power spectra, we reviewed evidence that many geofields including the topography, ore concentrations, rock density, magnetic susceptibility and others were scaling over considerable ranges of horizontal and vertical scale. From the relatively small number of studies where proximate vertical and horizontal data were available, we argued that generally the scaling was anisotropic with the exponent ratio H_z in the range $\approx$1.5–3. This – combined with estimates of the "sphero-scale" near planetary scales – implies that structures in lithospheric vertical cross-sections typically start off very stratified but that at larger and larger scales they become less and less so (exactly the opposite of the behaviour observed in the atmosphere). However, power spectra are only second order moments; using "functional box-counting" to systematically determine the fractal dimensions of sets exceeding higher and higher thresholds – we argued that geofields are in fact multifractal. Since this implies that areas above thresholds are power law functions of resolution (box size), this contradicts assumptions of classical geostatistics which assume that the relevant Lebesgue measures are well defined lengths, areas, and volumes (i.e. that they have no significant resolution dependencies).

Having argued that the lithosphere is largely the product of scaling processes involving dynamical mechanisms repeating scale after scale, we concentrated on the resulting cascade processes. The history of the development of explicit cascade models has many lessons for geoscientists. They were first developed by de Wijs (1951) in an attempt to refute the idea that mineral ore concentrations had algebraic ("Pareto") distributions and to support the notion that they were at least roughly lognormal. However, later developments of cascades (in the 1960s, 1970s for applications in turbulence) showed that they could only at best give approximately log-normal distributions; Mandelbrot (1974) showed that the "dressed" cascades generally displayed the phenomenon of "divergence" of moments – i.e. precisely the algebraic behaviour that de Wijs had fought against.

But the issue of algebraic versus lognormality continued to be deepened. On the one hand, Bak et al. (1987) made another connection between fractals and algebraic probabilities: Self-Organized Criticality, the prototypical model being the sandpile in which grains are added one at a time provoking avalanches with algebraic distributions. In this context, the multifractal phase transition route to algebraic probabilities is sometimes called "nonclassical SOC". Finally, in

the 1980s it was demonstrated that cascade processes have stable, attractive behaviours so that most of the details of the dynamical mechanism are irrelevant, only three parameters will generally be enough to determine the behaviour (Schertzer and Lovejoy, 1987). This was a kind of "multiplicative central limit theorem"; it shows that even approximate log-normality compatible with the divergence of moments (which only affects the extreme tail of the probability distributions) is a special case; real world systems in fact had approximately log-Levy distributions and with the Levy parameter α typically <2; the log-normal value α=2 being special. Indeed we saw that α estimates for topography and ore concentration (using de Wijs's original data) have values of $\alpha \approx 1.8$, close to values for the susceptibility and magnetic field anomalies. The existence of universality classes turns out to be essential for the application of cascades and multifractals to the real world: without it, every scaling process would require an infinite number of parameters either to empirically characterize or to model.

In order to clarify these ideas, we included a long discussion of cascade models. This was partly pedagogical, but was also aimed at highlighting areas where clarification was necessary. Key additional points were a) the important distinction between microcanonical and canonical conservation, b) the distinction between "bare" and "dressed" cascade properties c) the nonlocal nature of cascade singularities (they are generally *not* Holder exponents), d) the codimension (γ, $c(\gamma)$) versus dimension (α, $f(\alpha)$) formalism for multifractals and e) the fact that typical observables are generally *not* quantities which are conserved scale by scale, they are generally *not* the direct result of cascade processes but involve an extra linear term qH in their moment scaling exponent $K(q)$ (requiring fractional integrations of cascades to model). In Sect. 5 we intercompared various scaling models and discussed how to make (realistic) continuous in scale cascades. We also show how – by introducing an (anisotropic) scale function in the place of the usual (isotropic) distance function, we can model anisotropic multifractals, and we gave several examples of topography simulations. The recognition that real world scaling systems have both scale and location dependent anisotropy is fundamental in geophysics since without it we could not explain the coexistence of diverse geomorphologies and scaling. Indeed, GSI demonstrates the "phenomenological fallacy" i.e. the fact that mechanism cannot be phenomenologically inferred from form: the same process at small and large scales can have drastically different phenomenologies yet be produced by the same scale invariant mechanism.

The problem of structures within structures occurring over enormous ranges of scale is a unifying geoscience problematic and advances over the last 30 years have shown that such systems have many – sometimes surprising – commonalities over vastly disparate scales. This is possible because it is now known that scale invariance is a symmetry principle of great generality encompassing systems with both scale by scale (differential) anisotropy and also those with anisotropy varying from place to place.

In many ways we are living in a "golden age" of geophysical data; to fully take advantage of this manna requires scaling techniques. Unfortunately, many of the necessary techniques and notions are still little known sometimes even among practioners. An unfortunate consequence is that mainstream geoscientists have found many of the applications of scaling too restrictive or simplistic leading them to discard scaling ideas altogether. An example of this is the debate in the 1990s about the value of the supposedly unique fractal dimension of the topography. Lack of agreement between disparate analysis techniques on diverse data sets lead many to conclude that scaling only held over narrow ranges of scale. In actual fact, as modern multifractal analyses amply show – the topography displays excellent (multi) scaling over more than 10^5 in scale. If the geofields really do respect some (generalized) scaling principle, then this fact will transform the geosciences.

Acknowledgements. This research was performed purely for scientific purposes, it was unfunded.

Edited by: A. Tarquis
Reviewed by: F. Agterberg and V. Dimri

References

Agterberg, F.: Geomathematics, 596 pp., Elsevier, 1974.

Agterberg, F.: New Applications of the Model of de Wisj in regional geochemistry, Math. Geol., 39, 1–25, doi:10.1007/s11004-006-96063-7, 2007.

Ahrens, L. H.: A fundamental law of geochemistry, Nature, 172, 1148–1152, 1953.

Aitchison, J. and Brown, J. A. C.: The lognormal distribution, with special reference to its uses in economics, 176 pp., Cambridge University Press, 1957.

Aviles, C. A., Scholz, C. H., and Boatwright, J.: Fractal Analysis Applied to Characteristic Segments of San Andreas Fault, J. Geophys. Res., 92, 331–344, 1987.

Bacry, A., Arneodo, A., Frisch, U., Gagne, Y., and Hopfinger, E.: Wavelet analysis of fully developed turbulence data and measurement of scaling exponents, in: Turbulence and coherent structures, edited by: Lessieur, M. and Metais, O., pp. 703–718, Kluwer, 1989.

Bahr, K.: The route to fractals in magnetotelluric exploration of the crust, in: Fractal behaviour of the earth system, edited by: Dimri, V. P., Springer, Heidelberg, 2005.

Bak, P., Tang, C., and Weiessenfeld, K.: Self-Organized Criticality: An explanation of 1/f noise, Phys. Rev. Lett., 59, 381–384, 1987.

Bak, P., Tang, C., and Weiessenfeld, K.: Self-Organized Criticality, Phys. Rev. Lett., A 38, 364–374, 1988.

Balmino, G.: The spectra of the topography of the Earth, Venus and Mars, Geophys. Res. Lett., 20(11), 1063–1066, 1993.

Balmino, G., Lambeck, K., and Kaula, W.: A spherical harmonic analysis of the Earth's topography, J. Geophys. Res., 78(2), 478–481, 1973.

Bansal, A. R. and Dimri, V. P.: Self-affine gravity covariance model for the Bay of Bengal, Geophys. J. Inter., 161, 21–30, 2005.

Barton, C. C.: Fractal analysis of scaling and spatial clustering of fractures, in: Fractals in the Earth Sciences, edited by: Barton, C. C. and La Pointe, P. R., pp. 141–178, Plenum Press, New York, 1995.

Barton, C. C. and Scholz, C. H.: The fractal size and spatial distribution of hydrocarbon accumulations: Implications for resource assessment and exploration strategy, in: Fractals in Petroleum Geology and Earth Processes, edited by: Barton, C. C., and La Pointe, P. R., pp. 13–34, Plenum Press, New York, 1995.

Bean, C. and McCloskey, J.: Power-law random behaviour of seismic reflectivity in boreholes and its relationship to crustal deformation models, Earth Planet. Sci. Lett., 21, 2641–2644, 1993.

Bell, T. H.: Statistical features of sea floor topography, Deep Sea Res., 22, 883–891, 1975.

Benzi, R., Ciliberto, S., Tripiccione, R., Baudet, C., Massaioli, F., and Succi, S.: Extended self-similarity in turbulent flows, Phys. Rev. E., 48, R29–R32, 1993.

Berkson, J. M. and Matthews, J. E.: Statistical properties of seafloor roughness, in: Acoustics and the Sea-Bed, edited by: Pace, N. G., p. 215–223, Bath University Press, Bath, England, 1983.

Blakely, R. J.: Potential Theory in Gravity and Magnetic Applications, 441 pp., Cambridge University Press, 1995.

Bourlon, E., Mareschal J.-C., Gaonac'h H., Lovejoy S., and Schertzer, D.: Anisotropic Scaling Rock Density and the Earth's Gravity Field, paper presented at Nonlinear VAriability in Geophysics, 4 July 1998, Roscoff, France, 1998.

Brax, P. and Pechanski, R.: Levy stable law description on intermittent behaviour and quark-gluon phase transitions, Phys. Lett. B, 253, 225–230, 1991.

Burridge, L. and Knopoff, L.: Model and theoretical seismicity, Bull. Seismol. Soc. Am., 57, 341–371, 1967.

Cahalan, R.: Bounded cascade clouds: albedo and effective thickness, Nonlin. Processes Geophys., 1, 156–167, 1994, http://www.nonlin-processes-geophys.net/1/156/1994/.

Cargill, S. M., Root, D. H., and Bailey, E. H.: Estimating unstable resources from historical industrial data, Econ. Geol., 76, 1081–1095, 1981.

Carlson, J. M., Langer, J. S., and Shaw, B. E.: Dynamics of earthquake faults, Rev. Mod. Phys., 66, 657–671, 1994.

Cheng, Q.: Multifractal modeling and lacunarity analysis, Math. Geol., 29, 919–932, 1997.

Cheng, Q.: Multifractality and spatial statistics, Comp. Geosci., 25, 949–961, 1999.

Cheng, Q.: Integrated spatial and spectrum method for geochemical anomaly separation, Natural Resour. Res., 9, 43–51, 2000a.

Cheng, Q.: Interpolation of means of multifractal, kriging and moving average techniques, paper presented at GAC/MAC meeting of GeoCanada2000, Calgary, 2000b.

Cheng, Q.: Multifractal Distribution of Eigenvalues and Eigenvectors from 2D Multiplicative Cascade Multifractal Fields, Mathematical Geology, 37, 15–27, doi:10.1007/s11004-11005-19223-11001, 2005.

Cheng, Q.: GIS based fractal and multifractal methods for mineral deposit prediction, in: Geostatistics Banff 2004, edited by: Leuanthong, D. and Deutsch, C. V., pp. 1017–1026, Springer, New York, 2006.

Cheng, Q.: Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China, Ore Geology Rev., in press, doi:10.1016/j.oregeorev.2006.1010.1002, 2007.

Cheng, Q. and Agterberg, F.: Multifractal modelling and spatial statistics, Math. Geol., 28, 1–16, 1996.

Cheng, Q., Agterberg, F., and Ballantyne, S. B.: The separation of geochemical anomalies from background by fractal methods, J. Geochem. Exploration, 51, 109–130, 1994.

Crovelli, R. A. and Barton, C. C.: Fractals and the Pareto Distribution applied to Petroleum Accumulation-Size distributions, in: Fractals in Petroleum Geology and Earth Processes, pp. 59–71, Plenum, New York, 1995.

de Wijs, H. J.: Statistics of ore distribution, part I, Geologie en Mijnbouw, 13, 365–375, 1951.

de Wijs, H. J.: Statistics of ore distribution: (2) Theory of binomial distribution applied to sampling and engineering problems, Geologie en Mijnbouw, 15, 12–24, 1953.

Dimri, V. P.: Fractals in geophysics and seismology: an introduction, in: Fractal behaviour of the earth system, edited by: Dimri, V. P., pp. 1–22, Springer, Heidelberg, 2005.

Dimri, V. P. and Vedanti, N.: Scaling evidences of thermal properties in earth's crust and its implications, in: Fractal behaviour of the earth system, edited by: Dimri, V. P., Springer, Heidelberg, 2005.

Dolan, S. S., Bean, C., and Riollet, B.: The broad-band fractal nature of heterogeneity in the upper crust from petrophysical logs, Geophys. J. Int., 132, 489–507, 1998.

Dubayah, R., Wood, E. F., and Lavallée, D.: Multiscalig analysis in distriubted modeling and remote sensing: an application using soil moisture, in: Scale in Remote Sensing and GIS, edited by: Quattrochi, D. and Goodchild, M., CRC Lewis Publishers, Boca Raton, 1997.

Fedi, M.: Global and local multiscale analysis of magnetic susceptibility data, Pure and Appl. Geophys., 160, 2399–2417, 2003.

Fedi, M., Fiore, D., and La Manna, M.: Regularity analysis applied to well log data, in: Fractal Behaviour of the earth system, edited by: Dimri, V. P., Springer, Heidelberg, 2005.

Feigenbaum, M. J.: Quantitative universality for a class of nonlinear transformations, J. Stat. Phys., 19, 25–52, 1978.

Fox, C. G. and Hayes, D. E.: Quantitative methods for analyzing the roughness of the seafloor, Rev. Geophys., 23, 1–48, 1985.

Frisch, U., Sulem, P. L., and Nelkin, M.: A simple dynamical model of intermittency in fully develop turbulence, J. Fluid Mech., 87, 719–724, 1978.

Gagnon, J. S., Lovejoy, S., and Schertzer, D.: Multifractal surfaces and topography, Europhys. Lett., 62, 801–807, 2003.

Gagnon, J. S., Lovejoy, S., and Schertzer, D.: Multifractal earth topography, Nonlin. Processes Geophys., 13, 541–570, 2006, http://www.nonlin-processes-geophys.net/13/541/2006/.

Gaonac'h, G., Lovejoy, S., and Stix, J.: The resolution dependence of basaltic lava flows and their fractal dimensions, Geophys. Res. Lett., 19, 785–788, 1992.

Gaonac'h, H., Lovejoy, S., and Schertzer, D.: Percolating Magmas and explosive volcanism, Geophys. Res. Lett., 30, 1559–1563, 2003.

Gibert, D. and Courtillot, V.: Seasat altimetry and the South Atlantic geoid 1. Spectral analysis, J. Geophys. Res., 92(B7), 6235–6248, 1987.

Goodchild, M. F.: Fractals and the accuracy of geographical mea-

sures, Math. Geol., 12, 85–98, 1980.
Gupta, V. K. and Waymire, E.: A Statistica l Analysis of Mesoscale Rainfall as a Random Cascade, J. Appl. Meteorol., 32, 251–267, 1993.
Gutenberg, B. and Richter, C. F.: Frequency of earthquakes in California, Bull. Seis. Soc. Amer., 34, 185–193, 1944.
Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., and Shraiman, B.: Fractal measures and their singularities: the characterization of strange sets, Phys. Rev.A, 33, 1141–1151, 1986.
Harvey, D. A., Gaonac'h, H., Lovejoy, S., and Schertzer, D.: Multifractal characterization of remotely sensed volcanic features: a case study from Kiluaea volcano, Hawaii, Fractals, 10, 265–274, 2002.
Hollinger, K.: Fault scaling and 1/f noise scaling of seismic velcoity fluctuations in the upper crystaline crust, Geology, 24, 1103–1106, 1996.
Holschneider, M.: Wavelets: an analysis tool, Oxford: Clarendon Press, New York, 1995.
Hooge, C., Lovejoy, S., Schertzer, D., Schmitt, F., and Malouin, F.: Multifractal phase transitions: the origin of self-organized criticality in earthquakes, Nonlin. Processes Geophys., 1, 191–197, 1994,
http://www.nonlin-processes-geophys.net/1/191/1994/.
Hurst, H. E.: Long-term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers, 116, 770–808, 1951.
Kaminski, E. and Jaupart, C.: The size distribution of pyroclasts and the fragmentation sequence in explosive volcanic eruptions, J Geophys Res., 103, 29 759–29 779, 1998.
Kantelhart, J. W., Zscchegner, S. A., Koscielny-Bunde, K., Havlin, S., Bunde, A., and Stanley, H. E.: Multifractal detrended fluctuation analysis of nonstationary time series, Physica A, 316, 87–114, 2002.
Kapteyn, J. C.: Skew frequency curves in biology and statistics, Astronomical Laboratory, Noordhoff., Groningen, 1903.
Kida, S.: Log stable distribution and intermittency of turbulence, J. Phys. Soc. Japan, 60, 5–8, 1991.
Klinkenberg, B. and Goodchild, M. F.: The fractal properties of topography: A comparison of methods, Earth surface Proc. and Landforms, 17, 217–234, 1992.
Kolmogorov, A. N.: Logarithmically normal distribution of fragmentary particle sizes, Dokl. Akad. Nauk SSSR, 311, 99–101, 1941a.
Kolmogorov, A. N.: Local structure of turbulence in an incompressible liquid for very large Reynolds numbers, English translation: Proc. Roy. Soc. A434, 9–17, 1991, Proc. Acad. Sci. URSS., Geochem. Sect., 30, 299–303, 1941b.
Kolmogorov, A. N.: A refinement of previous hypotheses concerning the local structure of turbulence in viscous incompressible fluid at high Reynolds number, J. Fluid Mech., 83, 349–353, 1962.
Korcak, J.: Deux types fondamentaux de distribution statistique, Bulletin de l'Institut Inter. de Statistique, 3, 295–299, 1938.
Korvin, G.: Fractal models in the Earth sciences, 396 pp., Elsevier Science Publishers, 1992.
Krige, D. G.: On the departure of ore value distributions from the lognormal model in South African gold mines, J. South African Institute of Mining and Metallurgy, 231–244, 1960.
Kucinkas, A. B., Turcotte, D. L., Huang, J., and Ford, P. G.: Fractal analysis of Venus topography in Tinatin Planitia and Ovda Regio, J. Geophys. Res., 97, 13 625–13 641, 1992.
Laferrière, A. and Gaonac'h, H.: Multifractal properties of visible reflectance fields from basaltic volcanoes, J. Geophys. Res., 104, 5115–5126, 1999.
Langel, R. A. and Estes, R. H.: A geomagnetic field spectrum, J. Geophys. Res., 9, 250–253, 1982.
Lasky, S. G.: How tonnage and grade relations help predict ore reserves, Eng. Mining J., 151, 81–85, 1950.
Lavallée, D., Lovejoy, S., and Schertzer, D.: On the determination of the codimension function, in: Non-linear variability in geophysics: Scaling and Fractals, edited by: Schertzer, D. and Lovejoy, S., pp. 99–110, Kluwer, 1991.
Lavallée, D., Lovejoy, S., Schertzer, D., and Ladoy, P.: Nonlinear variability and landscape topography: analysis and simulation, in: Fractals in geography, edited by: De Cola, L. and Lam, N., pp. 171–205, Prentice-Hall, Englewood, N.J., 1993.
Lazarev, A., Schertzer, D., Lovejoy, S., and Chigirinskaya, Y.: Unified multifractal atmospheric dynamics tested in the tropics: part II, vertical scaling and Generalized Scale Invariance, Nonlin. Processes Geophys., 1, 115–123, 1994,
http://www.nonlin-processes-geophys.net/1/115/1994/.
Leary, P.: Rock as a critical-point system and the inherent implausibility of realiable earthquake prediction, Geophys. J. Internat., 131, 451–466., 1997.
Leary, P. C.: Fractures and physical heteorgeneity in crustal rock, in: Heterogeneity in the crust and upper mantle, edited by: Goff, J. A. and Hollinger, K., pp. 155–182, Kluwer, New York, 2003.
Leonardi, S. and Kümpel, H. J.: Fractal variabiloity in super deep borehole – implications for the signature of crustal heterogeneties, Tectonophysics, 301, 173–181, 1999.
Leopold, L. B. and Maddock, T.: The hydraulic geometry of stream channels and some physiographic implications, U.S. Geol. Survey, 1953.
Levy, P.: Calcul des probabilites, Gautier-Villars, Paris, 1925.
Lilley, M., Lovejoy, S., Strawbridge, K., and Schertzer, D.: 23/9 dimensional anisotropic scaling of passive admixtures using lidar aerosol data, Phys. Rev. E, 70, 1–7, doi:10.1103/036307, 2004.
Lovejoy, S., Lavallée, D., and Schertzer, D.: Multifractal topography and the $l^{1/2}$ law, Nonlin. Processes Geophys., 2, 17–22, 1995,
http://www.nonlin-processes-geophys.net/2/17/1995/.
Lovejoy, S., Pecknold, S., and Schertzer, D.: Stratified multifractal magnetization and surface geomagnetic fields, part 1: spectral analysis and modelling, Geophys. J. Inter., 145, 112–126, 2001.
Lovejoy, S., Mareschal, J.-C., Gaonac'h, H., and Schertzer, D.: Anisotropic scaling models of rock density and the earth's surface gravity field, edited by: Bonham-Carter, G. Q. C., 2007a.
Lovejoy, S. and Schertzer, D.: Our multifractal atmosphere: A unique laboratory for non-linear dynamics, Phys. Canada, 46, 62–72, 1990.
Lovejoy, S. and Schertzer, D.: How bright is the coast of Brittany?, in: Fractals in Geoscience and Remote Sensing, edited by: Wilkinson, G., pp. 102–151, Office for Official Publications of the European Communities, Luxembourg, 1995.
Lovejoy, S. and Schertzer, D.: Stochastic chaos and multifractal geophysics, in: Chaos, Fractals and models 96, edited by: Guindani, F. M. and Salvadori, G., pp. 38–52, Italian University Press, 1998.
Lovejoy, S. and Schertzer, D.: Multifractals, cloud radiances and

rain, J. Hydrol., 322, 59–88, 2006.
Lovejoy, S. and Schertzer, D.: Scale, scaling and multifractals in geophysics: twenty years on, in: Nonlinear dynamics in geophysics, edited by: Tsonis, A. A. and Elsner, J., Elsevier, 2007.
Lovejoy, S., Schertzer, D., and Gagnon, J. S.: Multifractal simulations of the Earth's surface and interior: anisotropic singularities and morphology, in: GIS and Spatial Analysis, Proc. Of Inter. Assoc. Math. Geology, edited by: Cheng, Q. and Bonham-Carter, G., pp. 37–54, 2005.
Lovejoy, S., Schertzer, D., and Ladoy, P.: Fractal characterisation of inhomogeneous measuring networks, Nature, 319, 43–44, 1986.
Lovejoy, S., Schertzer, D., and Tsonis, A. A.: Functional Box-Counting and Multiple Dimensions in rain, Science, 235, 1036–1038, 1987.
Maliverno, A.: Fractals and ocean floor topography: a review and a model, in: Fractals in the Earth Sciences, edited by: Barton, C. C. and La Pointe, P. R., pp. 107–130, 1995.
Mandelbrot, B.: Fractal geometry: What is it and what does it do?, in: Fractals in the Natural Sciences, edited by: Fleischman, M., Tildesley, D. J., and Ball, R. C., p. 3–16, Princeton University Press, 1989.
Mandelbrot, B.: The statistics of Natural Resources and the law of Pareto, in: Fractals in petroleum geology and the earth sciences, edited by: Barton, C. C. and La Pointe, P. R., pp. 1–12, Plenum Press, New York, 1995.
Mandelbrot, B. B.: How long is the coastline of Britain? Statistical self-similarity and fractional dimension, Science, 155, 636–638, 1967.
Mandelbrot, B. B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, J. Fluid Mech., 62, 331–350, 1974.
Mandelbrot, B. B.: Stochastic models for the earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands, Proc. Nat. Acad. of Sci. USA, 72, 3825–3828, 1975.
Mandelbrot, B. B.: Fractals, form, chance and dimension, Freeman, San Francisco, 1977.
Mandelbrot, B. B.: The Fractal Geometry of Nature, Freeman, San Francisco, 1983.
Mandelbrot, B. B. and Wallis, J. R.: Some long run properties of geophysical records, Water Resour. Res., 5, 228–259, 1969.
Mareschal, J. C.: Fractal reconstruction of sea-floor topography, Pageoph, 131, 197–210, 1989.
Marsan, D. and Bean, C.: Multifractal modelling and analyses of crustal heterogeneity, in: Heterogeneity in the crust and upper Mantle, edited by: Goff, J. A. and Hollinger, K., pp. 207–236, Kluwer Academic, New York, 2003.
Marsan, D. and Bean, C. J.: Multiscaling nature of sonic velocities and lithography in the upper crystalline crust: evidence from the KTB main borehole, Geophys. Res. Lett., 26, 275–278, 1999.
Marsan, D., Schertzer, D., and Lovejoy, S.: Causal space-time multifractal processes: predictability and forecasting of rain fields, J. Geophys. Res., 31D, 26 333–26 346, 1996.
Matheron, G.: Traie de géostatistique appliiqué, I, Editions Technip, Paris, 1962.
Matheron, G.: Random functions and their applications in geology, in: Geostatistics, edited by: Merriam, D. F., pp. 79–87, Plenum Press, New York, 1970.
Maus, S. and Dimri, V.: Potential field power spectrum inversion for scaling geology, J. Geophys. Res., 100, 12 605–12 616, 1995.
Maus, S. and Dimri, V.: Depth estimation from the scaling power spectrum of potential fields?, Geophys. J. Int., 124, 113–120, 1996.
Meneveau, C. and Sreenivasan, K. R.: Simple multifractal cascade model for fully develop turbulence, Phys. Rev. Lett., 59, 1424–1427, 1987.
Molz, F. and Boman, G. K.: A stochastic interpolation scheme in subsurface hydrology, Water Resour. Res., 29, 3769–3774, 1993.
Molz, F. J. and Liu, H.: Fractional Brownian motion and fractional gaussian noise in subsurface hydrology: a review, presentation of fundamental properties and extensions, Water Resour. Res., 33, 2273–2286, 1997.
Novikov, E. A. and Stewart, R.: Intermittency of turbulence and spectrum of fluctuations in energy-disspation, Izv. Akad. Nauk. SSSR. Ser. Geofiz., 3, 408–412, 1964.
Okubo, P. G., and Aki, K.: Fractal Geometry in the San Andreas Fault System, J. Geophys. Res., 92, 345–355, 1987.
Omori, F.: On the After-Schocks of Earthquakes, J. Coll. Sci. Tokyo Imp. University, 7, 111–200, 1895.
Parisi, G. and Frisch, U.: A multifractal model of intermittency, in: Turbulence and predictability in geophysical fluid dynamics and climate dynamics, edited by: Ghil, M., Benzi, R., and Parisi, G., pp. 84–88, North Holland, Amsterdam, 1985.
Pecknold, S., Lovejoy, S., and Schertzer, D.: Stratified multifractal magnetization and surface geomagnetic fields, part 2: multifractal analysis and simulation, Geophys. Inter. J., 145, 127–144, 2001.
Perrin, J.: Les Atomes, NRF-Gallimard, Paris, 1913.
Pilkington, M. and Todoeschuck, J.: Stochastic inversion for scaling geology, Geophys. J. Int., 102, 205–217, 1990.
Pilkington, M. and Todoeschuck, J.: Fractal magnetization of continental crust, Geophys. Res. Lett., 20, 627–630, 1993.
Pilkington, M. and Todoeschuck, J.: Scaling nature of crustal susceptibilities, Geophys. Res. Lett., 22, 779–782, 1995.
Plotnick, R. E. and Prestegaard, K. L.: Fractal and multifractal models and methods in stratigraphy, in: Fractals in petroleum geology and earth processes, edited by: Barton, C. C. and Lapointe, P. R., pp. 73–96, Plenum, New York, 1995.
Richardson, L. F.: Weather prediction by numerical process, Cambridge University Press republished by Dover, 1965, 1922.
Richardson, L. F.: The problem of contiguity: an appendix of statistics of deadly quarrels, General Systems Yearbook, 6, 139–187, 1961.
Rodriguez-Iturbe, I. and Rinaldo, A.: Fractal River Basins, 547 pp., Cambridge University Press, Cambridge, 1997.
Salvadori, G., Schertzer, D., and Lovejoy, S.: Multifractal objective analysis and interpolation, Stoch. Environ. Resear. and Risk Analysis, 5, 261–283, 2001.
Samorodniitsky, G. and Taqqu, M. S.: Stable non-gaussian random processes: stochastic models with infinite variance, Chapman and Hall, New York, 1994.
Schertzer, D. and Lovejoy, S.: The dimension and intermittency of atmospheric dynamics, in: Turbulent Shear Flow 4, edited by: Launder, B., pp. 7–33, Springer-Verlag, 1985a.
Schertzer, D. and Lovejoy, S.: Generalised scale invariance in turbulent phenomena, Physico-Chemical Hydrodynam. J., 6, 623–635, 1985b.
Schertzer, D. and Lovejoy, S.: Physical modeling and Analysis of

Rain and Clouds by Anisotropic Scaling of Multiplicative Processes, J. Geophys. Res., 92, 9693–9714, 1987.

Schertzer, D. and Lovejoy, S.: Multifractal Generation of Self-Organized Criticality, in: Fractals In the natural and applied sciences, edited by: Novak, M. M., pp. 325–339, Elsevier, North-Holland, 1994.

Schertzer, D. and Lovejoy, S.: From scalar cascades to Lie cascades: joint multifractal analysis of rain and cloud processes, in: Space/time Variability and Interdependance for various hydrological processes, edited by: Feddes, R. A., pp. 153–173, Cambridge University Press, New-York, 1995.

Schertzer, D. and Lovejoy, S.: Universal Multifractals do Exist!, J. Appl. Meteor., 36, 1296–1303, 1997.

Schertzer, D., Lovejoy, S., Lavallée, D., and Schmitt, F.: Universal hard multifractal turbulence: theory and observation, in: Nonlinar dynamics of Structures, edited by: Sagdeev, R. Z., Frisch, U., Moiseev, A. S., and Erokhin, A., pp. 213–235, World Scientific, Singapore, 1991.

Schmitt, F., Lavallée, D., Schertzer, D., and Lovejoy, S.: Empirical Determination of Universal Multifractal Exponents in Turbulent Velocity Fields, Phys. Rev. Lett., 68, 305–308, 1992.

Schmitt, F., Schertzer, D., Lovejoy, S., and Brunet, Y.: Empirical study of multifractal phase transitions in atmospheric turbulence, Nonlin. Processes Geophys., 1, 95–104, 1994, http://www.nonlin-processes-geophys.net/1/95/1994/.

Schmitt, F., Schertzer, D., Lovejoy, S., and Tessier, Y.: R/S analysis and multifractals: application to rain and river flows, paper presented at General Assembly of the European Geophysical Society, Hambourg, Germany, 1995.

Scholz, C. H. and Mandelbrot, B. B. (Eds.): Fractals in Geophysics, Birkauser, Basel, 1989.

She, Z. S. and Levesque, E.: Universal scaling laws in fully developed turbulence, Phys. Rev. Lett., 72, 336–339, 1994.

Shiomi, K., Sato, H., and Ohtake, M.: Broad-band power-law spectra of well-log data in Japan, Geophys. J. Inter., 130, 57–64, 1997.

Stauffer, D.: Introduction to percolation Theory, Taylor & Francis, London, Philadelphia, 1985.

Steinhaus, H.: Length, Shape and Area, Colloquium Mathematicum, III, 1–13, 1954.

Takayasu, H.: Fractals in the physical sciences, Manchester University Press, Manchester and New-York, 1990.

Tchiguirinskaia, I.: Scale invariance and stratification: the unified multifractal model of hydraulic conductivity, Fractals, 10(3), 329–334, 2002.

Tchiguirinskaia, I., Lu, S., Molz, F. J., Williams, T. M., and Lavallée, D.: Multifractal versus monofractal analysis of wetland topography, SERRA, 14, 1, 8–32, 2000.

Telesca, L. and Lapenna, V.: Fractal methods in self-potential signals measured in seismic areas, in: Fractal behaviour of the earth system, edited by: Dimri, V. P., Springer, Heidelberg, 2005.

Tennekoon, L., Bouffadel, M. C., and Nyquist, J. E.: Multifractal characterization of airborne geophysical data at the Oak Ridge facility, Stoch. Environ. Res. and Risk Assess., 19, 227–240, 2005.

Todoeschuck, J. P. and Jensen, O. G.: 1/f Geology and Seismic deconvolution, in: Nonlinear variability in geophysics: scaling and fractals, edited by: Schertzer, D. and Lovejoy, S., pp. 261–269, Kluwer, Dordrecht, 1991.

Todoeschuck, J. P., Jensen, O. G., and Labonte, S.: Gaussian scaling noise model of seismic reflection sequences: evidence from well-logs, Geophys., 55, 480–484, 1990.

Tubman, K. M. and Crane, S. D.: Vertical versus horizontal well log variabiloity and application to fractal reservoir modeling, in: Fractals in Petroleum Geology and Earth Processes, edited by: Barton, C. C. and La Pointe, P. R., Plenum Press, New York, 1995.

Turcotte, D. L.: A fractal approach to the relationship between ore grade and tonnage, Econ. Geol., 81, 1528–1532, 1986.

Turcotte, D. L.: Fractals in Geology and Geophysics, Pageoph, 131, 171–196, 1989.

Turcotte, D. L. and Oxburgh, E. R.: Finite amplitude convective cells and continental drift, J. Fluid Mech., 28, 29–42, 1967.

Van Tongeren, W.: On the distribution of chemical elements in their natural occurrences, Geologie en Mijnbouw, 12, 123–128, 1950.

Vennig-Meinesz, F. A.: A remarkable feature of the Earth-s topography, Proc. K. Ned. Akad. Wet. Ser. B Phys. Sci., 54, 212–228, 1951.

Venugopal, V., Roux, S., Foufoula, E., and Arneodo, A.: Revisiting multifractality of high-resolution temporal rainfall using a wavelet-based formalism, Water Resour. Res., doi:10.1029/2005WR004489, 2006.

Weatherley, D. and Abe, S.: Earthquake statistics in a Block Slider Model and a fully dynamic Fault Model, Nonlin. Processes Geophys., 11, 553–560, 2004, http://www.nonlin-processes-geophys.net/11/553/2004/.

Wilson, J.: Physically based stochastic modelling of rain and cloudfields, M.Sc. thesis, McGill University, Montréal (Québec), Canada, 1991.

Wu, R. S., Xu, Z., and Li, X. P.: Heterogeneity spectrum and scale-anisotropy in the upper crust revealed by the German Continetal Deep-Drilling (KTB) Holes, Geophys. Res. Lett., 21, 911–914, 1994.

Yaglom, A. M.: The influence on the fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval, Sov. Phys. Dokl., 2, 26–30, 1966.