

Single- and Multiscale Remote Sensing Techniques, Multifractals, and MODIS-Derived Vegetation and Soil Moisture

S. Lovejoy, A. M. Tarquis,* H. Gaonac'h, and D. Schertzer

Scaling processes are increasingly understood to be the result of nonlinear dynamic mechanisms repeating scale after scale from large to small scales leading to nonclassical resolution dependencies. This means that the statistical properties systematically vary in strong, power-law ways with the resolution. When present in geophysical and remotely sensed fields, it implies that when classical (single-scale) remote sensing algorithms are used to determine surrogates of various geophysical fields, they can at most be correct at the unique (and subjective) calibration resolution. Scaling analysis and modeling techniques were applied to MODIS TERRA Bands 1 through 7 and to the standard derived vegetation and soil moisture indices in order to quantitatively characterize the wide range of scaling of these fields. The scaling exponents we found are not so large; however, they act across wide scale ranges and imply large effects. For example, for the statistics near the mean, the MODIS (500-m) resolution would be biased by a factor of ~ 1.52 when compared with similar results from an "ideal" sensor at 1-mm resolution. Applying the standard index algorithms on lower and lower resolution satellite data, we obtained indices with significantly different statistical properties than if the same algorithm was used at the finest resolution and then degraded to an intermediate value (a difference of a factor ~ 1.54). This shows that the algorithms can, at best, be accurate at the unique calibration scale and this points to the need to develop resolution-independent algorithms based on the scaling exponents.

SOIL MOISTURE is an important hydrologic state variable critical to successful hydroclimatic and environmental predictions. Soil moisture varies in both space and time because of spatiotemporal variations in precipitation, soil properties, topographic features, and vegetation characteristics (Corwin et al., 2006). Monitoring all these processes across multiple scales is a fundamental component of many environmental and natural resource issues and remote sensing has long been identified as a technology capable of supporting such development (Viña et al., 2004; Vereecken et al., 2007).

The geophysical and biological processes that determine vegetation, soil moisture, and other surface characteristics are

highly nonlinear; they involve interacting structures from planetary down to millimeter scales. A consequence is that the surface radiance fields are statistically interrelated across comparable ranges. Such statistical interrelations (correlations) are the basis of most remote sensing algorithms. A typical approach identifies wavebands that are particularly sensitive to the surface features of interest—such as the vegetation or soil moisture; then semiempirical algorithms are developed to relate them to the radiances. Typically, the algorithms assume that satellite data are homogenous or roughly homogenous at scales below the resolution.

The problem with this approach is that it ignores systematic scale or resolution dependencies due to strong subpixel variability and singles out a single scale for the calibration. This scale is essentially subjective; it is primarily determined by the available technology. A symptom of this scale problem is that when new, higher resolution satellite data are used in algorithms calibrated at earlier, lower resolutions, the algorithm typically must be recalibrated.

In the last 25 yr, an understanding of these strong resolution dependencies has started to emerge, and systematic techniques are now available for analyzing and modeling such behavior. Motivated by the fractal geometry of sets (Mandelbrot, 1983) and the development of (multifractal) cascade models in turbulence, the origin of this behavior has been traced to nonlinear dynamic mechanisms that repeat scale after scale from large to small scales. Paralleling this revolution in our understanding of the consequences of wide-range scaling is the generalization of the notion of scale itself to encompass systems that are scaling but highly anisotropic (generalized scale invariance, Schertzer and

S. Lovejoy, Dep. of Physics, McGill Univ., 3600 University St., Montreal, QC H3A 2T8 Canada; A. M. Tarquis, Dep. of Applied Mathematics, E.T.S.I.A. Agricultural Engineering, Univ. Politécnica de Madrid, Ciudad Universitaria s.n. 28040, Madrid, Spain, and CEIGRAM-U.P.M., Madrid, Spain; H. Gaonac'h, GEOTOP, L'Université du Québec à Montréal, Montreal, Canada; D. Schertzer, Université de Paris Est, ENPC, CERVE, 6-8, Avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex, France. Received 12 Nov. 2007. *Corresponding author (anamaria.tarquis@upm.es).

Vadose Zone J. 7:533–546
doi:10.2136/vzj2007.0173

© Soil Science Society of America
677 S. Segoe Rd. Madison, WI 53711 USA.
All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Lovejoy, 1985b; Pecknold et al., 1997; Lilley et al., 2004; Lovejoy and Schertzer, 2007). We illustrate some of these ideas using MODIS data and vegetation and soil moisture surrogates derived from them. Although we briefly review the basic multifractal notions, our aim is to perform systematic scale-by-scale analyses on the MODIS data and surrogates.

Spectral Analysis

The data we used are calibrated MODIS TERRA radiances (product level B1), with wavelengths indicated in Table 1; some of the images are shown in Fig. 1. The vegetation and soil surface moisture indices are standard products described in Rouse et al. (1973) and Lampkin and Yool (2004):

$$\sigma_{VI} = \frac{I_2 - I_1}{I_2 + I_1} \quad [1a]$$

$$\sigma_{SM} = \frac{I_6 - I_7}{I_6 + I_7} \quad [1b]$$

where σ_{VI} and σ_{SM} are the vegetation and soil surface moisture indices (for a complete list of symbols and definitions, see the appendix) and the I s are the radiances (the band number is indicated in the subscript); Eq. [1] is usually stated without reference to any particular scale. We may immediately note that, although the MODIS TERRA data have a resolution of 500 m (Bands 1 and 2 were degraded from 250 m) and the variability of the radiances and surface features continues to much smaller scales, the surrogates are defined at a single (subjective) resolution equal to that of the sensor. One of the applications of our analyses is to investigate how the relations between the surrogates and the bands used to define them change with scale: we anticipated that since the scaling properties are different, making the surrogates with data at different resolutions would produce fields with different properties.

Before using more sophisticated (multifractal) scale-by-scale analysis techniques, we first used one that is fairly familiar to geoscientists: spectral analysis. In addition to its familiarity, spectral analysis has the advantage of being very sensitive to scale breaks; it is also useful for studying anisotropy. For an intensity field I , we may define the spectral densities $P(\mathbf{k})$:

$$I(\mathbf{k}) = \int \exp(i\mathbf{k}\cdot\mathbf{r})I(\mathbf{r})d\mathbf{r} \quad [2a]$$

$$P(\mathbf{k}) = \langle |I(\mathbf{k})|^2 \rangle \quad [2b]$$

TABLE 1. Characteristics of the MODIS data, taken over Guadalajara (central Spain) over 512 by 512 pixels, each 500 by 500 m (~ 250 by 250 km²), on 29 July 2006.

Band	Wavelength
	nm
1	620–670
2	841–876
3	459–479
4	545–565
5	1230–1250
6	1628–1652
7	2105–2155

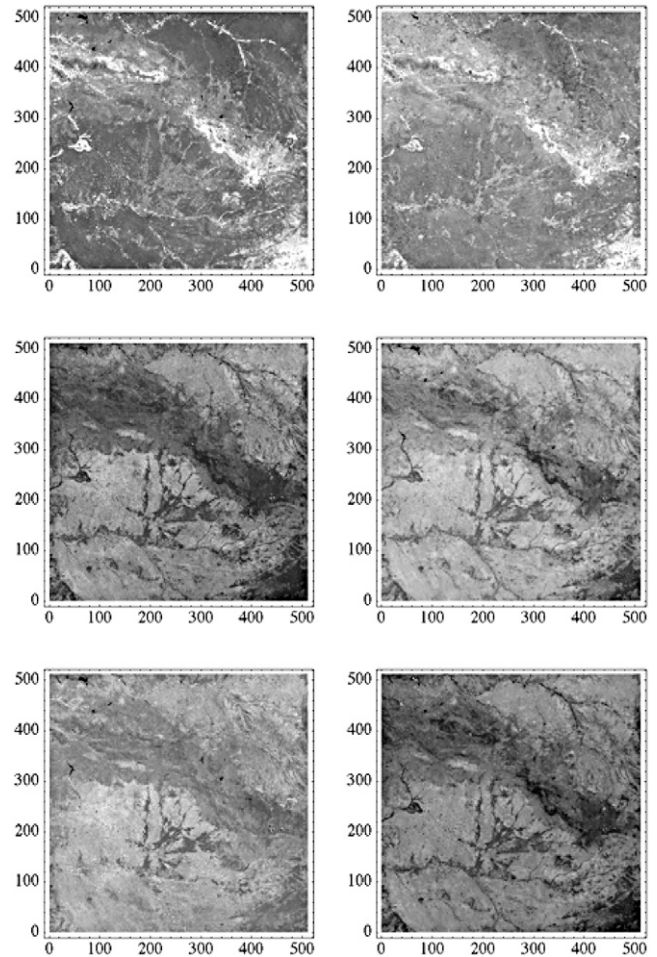


FIG. 1. Grey-shade images of Guadalajara (central Spain) corresponding to an area of 250 by 250 km. Left column: (top) vegetation index, (middle) Band 1, (bottom) Band 2; right column: (top) soil moisture index, (middle) Band 6, (bottom) Band 7.

where \mathbf{r} is a position vector and \mathbf{k} is a wave vector. Since P is quadratic in I , it is a second-order statistic. In the definition, we have included the theoretically motivated ensemble average (denoted $\langle \rangle$) although, in fact, below we estimate P from a single realization using a fast Fourier algorithm and a standard Hanning window.

In Fig. 2, we show P estimated for the vegetation and soil moisture indices. The indices were estimated from Eq. [1] at the highest resolution, i.e., they are single-scale surrogates (we discuss multiple-scale surrogates below). To make the contours clearer, we smoothed $\log_{10}P$ with a four-point Gaussian smoother. In the figure, we see that the contours are fairly roundish, indicating that, to some approximation, the second-order statistics are isotropic. Careful scale-by-scale analyses of the anisotropy would be rewarding (see, e.g., Pflug et al., 1993; Lewis et al., 1999; Beaulieu et al., 2007), but is outside our present scope. Analysis of the individual radiance bands revealed a degree of isotropy similar to the indices shown in Fig. 2.

The roundness of the P contours justifies the use of the “isotropic” spectrum $E(k)$ obtained by angle integrating P :

$$E(k) = \int_{|\mathbf{k}|=k} P(\mathbf{k})d\mathbf{k} \quad [3]$$

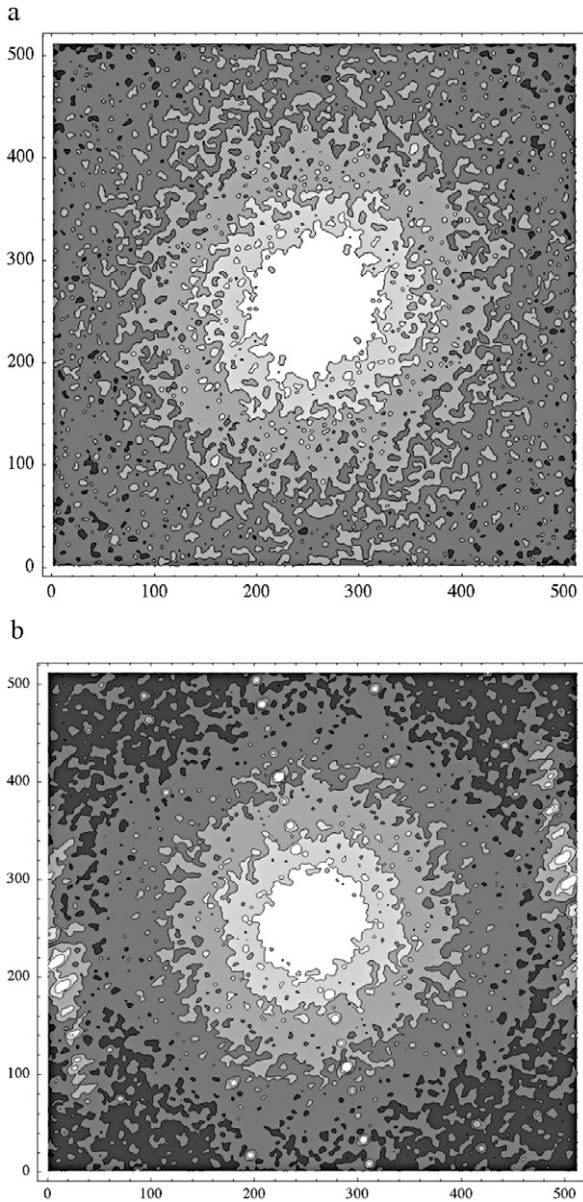


FIG. 2. A gray shade rendition of the (log) spectral densities (P) of (a) the vegetation index, and (b) the soil moisture index.

where k is the modulus of the wave vector (the notation indicates angle integration in Fourier space). In Fig. 3a, we see E on a log-log plot showing that, with the exception of the single lowest wave number $k = 1$, E is quite accurately of the power-law form:

$$E(k) \propto k^{-\beta} \quad [4]$$

where β is the spectral exponent. Note that sometimes angle averaging (rather than integration) is performed so that in two dimensions, 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 (e.g., one-dimensional sections of two-dimensional data will have the same exponent). We can also see from the figure that Band 5 (with strong artifacts, “banding”), indicated by blue in Fig. 3a nevertheless has good scaling except at harmonics of the basic banding wave number. In Table 2, we indicate the β values (estimated for $k \geq 2$), finding for all the bands $\beta \sim 1.17$

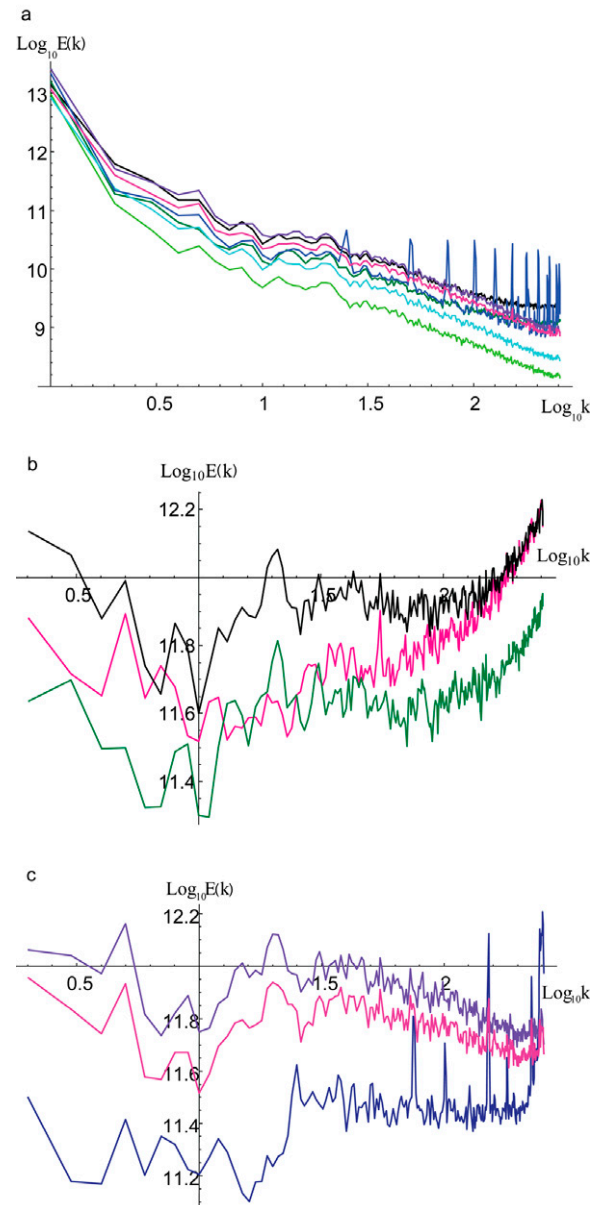


FIG. 3. (a) Spectra of all five bands $[E(k)]$ as a function of the modulus of the wave vector k (and using a Hanning window). In order from top to bottom at the point $\log_{10} k = 0.7$, the curves are: purple = Band 6, black = Band 1, magenta = Band 7, blue (with spikes) = Band 5, light green = Band 2, cyan = Band 4, dark green = Band 3; (b) log spectra compensated by $k^{1.17}$ [i.e., $\log_{10} k^{1.17} E(k)$] vegetation index (red, middle left) with Band 1 (black, top left) and Band 2 (green, bottom left) so that flat curves indicate $k^{-1.17}$ spectra (note the greatly expanded scale, the fluctuations are small); and (c) spectrum compensated by $k^{1.17}$ soil moisture (blue, bottom left) with Band 6 (purple, top left) and Band 7 (magenta, middle left).

± 0.08 , where the uncertainty indicates the band-to-band spread in exponents.

The power-law form Eq. [4] is called *scaling* because it is invariant under isotropic “zooms” ($E \rightarrow \lambda^{-1}E$, $\mathbf{k} \rightarrow \lambda\mathbf{k}$), and the exponent β is “scale invariant.” To better judge the quality of the scaling, we can “compensate” the spectrum by dividing by the mean behavior; this is shown in Fig. 3b and 3c (where we have removed the single $k = 1$ values to magnify the ordinate). Deviations from flatness indicate deviations from $k = 1.17$ scaling. In this way, we can easily see that the extreme high wave-number

TABLE 2. A comparison of the different exponents estimated in this study: β is the spectral exponent, H the basic fluctuation (“conservation”) exponent, C_1 the codimension characterizing the pure resolution dependence near the mean, α is the Levy index, which determines how quickly the resolution exponents vary with statistical moment (intensity) of the fields. TM indicates that estimates are based on the trace moments, Struc indicates that they are based on the fluctuations via the structure functions, q_D is the exponent of the probability distribution characterizing the extremes, VI is vegetation index, SM is soil moisture index, mss indicates statistics for the multiple-scale surrogate definition of index (all the other statistics are for single-scale surrogates [sss]). For all except α (which is fit in q space), fits are across the smallest spatial scale to a factor of 4 from the lowest (except TM, which also excludes a small scale factor of 2).

Band no. of index	β	Struc H	TM H	Struc C_1	TM C_1	α	TM α	q_D
1	1.15	0.186	0.109	0.031	0.036	1.86	1.88	7.2
2	1.09	0.144	0.078	0.024	0.034	1.89	1.87	5.8
3	1.24	0.226	0.160	0.031	0.039	1.88	2.02	4.0
4	1.26	0.218	0.166	0.030	0.037	1.92	1.96	5.7
5	1.05	0.141	0.064	0.054	0.039	1.95	2.00	8.5
6	1.25	0.207	0.159	0.026	0.035	1.93	1.95	6.9
7	1.18	0.199	0.130	0.026	0.037	1.95	2.14	6.6
VI sss	1.03	0.162	0.074	0.033	0.064	2.02	1.78	5.9
VI mss		0.231		0.049				
SM sss	1.03	0.144	0.073	0.010	0.053	2.23	2.29	5.5
SM mss		0.213		0.059				
Mean	1.17 ± 0.08	0.189 ± 0.034	0.124 ± 0.04	0.032 ± 0.010	0.037 ± 0.001	1.91 ± 0.03	1.97 ± 0.09	6.4 ± 1.4

† The q_D mean becomes 6.64 only for Bands 1, 2, 6, and 7.

factor of roughly 2 in scale has some extra power, presumably because of noise. Also, the single-scale surrogates (vegetation index, in red in the middle of Fig. 3b, and the soil moisture index, in blue on the bottom of Fig. 3c) are both more strongly affected at the highest wave numbers, and their β values are also lower, indicating that they also have greater variability at the lower wave numbers. Presumably, the extra high wave-number variability is because they are the results of a nonlinear operation ([Eq. 1]) at the pixel scale, which breaks the scaling. Note that variability is still present, although it is reduced by Fourier angle integration; the resulting spectrum is typical of multifractal processes. In any case, the largest (\log_{10}) deviations from power law (\log - \log linearity) are <0.1 across the entire range $k = 2$ to 128 cycles per image; since the corresponding variation of $\log_{10}E$ is ~ 2.5 , the spectral scaling holds to better than 4% of this range. Also note that the angle integration smoothes more effectively at the higher wave numbers since there are many more small-scale structures than large ones. Finally, note that the low wave-number break in the scaling (for $k = 1$) is probably caused by post-processing, which attempts to correct for atmospheric attenuation. In effect, the algorithm works as a low-pass filter, ensuring that the overall intensity across the scene is roughly constant.

Multifractal Analysis

Generalized Structure Functions

As indicated in Eq. [2], the spectrum is a second-order statistic; if we assume statistical horizontal translational invariance (*statistical homogeneity*; this is the spatial counterpart of *statistical stationarity*, which refers to statistical translational invariance in time), then the Wiener–Khinchin theorem shows

that the spectrum is the Fourier transform of the autocorrelation function. Before the advent of multifractals, it was believed that many turbulent and turbulent-like processes could be well modeled by Gaussian processes in which the statistics are completely determined by $P(\mathbf{k})$ [indeed, *isotropic* Gaussian processes are completely determined by $E(k)$]. However, to explain the extreme variability associated with the phenomenon of intermittency, cascade models were developed in which a simple mechanism repeats scale after scale in a multiplicative manner while conserving the mean of the process. The resulting statistical behavior is given by

$$\langle \varepsilon_\lambda \rangle = \lambda^{K_\varepsilon(q)}; \quad \lambda = L/l \quad [5]$$

where ε_λ is the conserved flux associated with an observable, such as an intensity, at resolution λ , which is the scale ratio across which the process has been developed, L is the outer scale, l is the inner scale, and $K_\varepsilon(q)$ is the (convex) moment scaling exponent (Kolmogorov, 1962; Mandelbrot, 1974). In remote sensing terms, Eq. [5] quantifies the rate at which spatial averaging (i.e., lower and lower resolution pixels, smaller λ) smoothes the various powers of the field ε .

The symbol ε is used for the cascade process since the prototypical cascade quantity is the turbulent energy flux, denoted ε , which is conserved from one scale to another in hydrodynamic turbulence. Note that we are discussing systems far from equilibrium and it is not the energy that is conserved but rather the ensemble mean flux of energy from large scales to smaller scales (in fact, ε is only a true flux in Fourier space, i.e., a flux from low to high wave numbers). In terms of the moments in Eq. [5], this conservation implies that the (appropriately normalized) process respects $\langle \varepsilon_\lambda \rangle = 1$ so that $K_\varepsilon(1) = 0$. Actually, although this ensemble “canonical” conservation is the most general one, it is not the only one possible. Indeed, more restrictive “micro-canonical” conservation principles are quite popular—in spite of the fact that the variability of the corresponding multifractal processes is much lower and that such processes do not have simplifying universal behaviors (see Schertzer and Lovejoy, 1992, and below).

The typical inner scales of turbulent processes in the atmosphere are the viscous dissipation scales, which are typically on the order of millimeters or less. In remote sensing applications, while the actual inner scale of the radiances may be of the same order, the satellite averages them to the larger scale l (here 500 m) so that the relevant resolution scale in Eq. [5] is the resolution of the images. In this case, the small subsensor scales average out most of the small-scale variability. In general, however, they will not completely smooth it out; this is the origin of the interesting strong variability of the extremes, which goes variously under the names multifractal butterfly effect, non-classical self-organized criticality, first-order multifractal phase transitions, and divergence of statistical moments, which are discussed below (for a review of this and other multifractal properties, see Schertzer et al., 1997).

The behavior described by Eq. [5] is called *multiscaling* because each statistical moment is scaling but with a different exponent. Continuing with the turbulence example, the famous Kolmogorov law for isotropic turbulence relates the energy flux to velocity fluctuations (Δv) across a distance Δx as follows:

$$\Delta v_\lambda = \varepsilon_\lambda^a \Delta x^H$$

$$\Delta x = L/\lambda; H = 1/3; a = 1/3 \quad [6]$$

where H is the scaling exponent of the fluctuations with respect to the flux and a is the exponent of the flux. 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:

$$\langle \Delta v_\lambda^q \rangle \propto \langle \varepsilon_\lambda^{aq} \rangle \lambda^{-qH} = \lambda^{-\xi(q)}$$

$$\xi(q) = Hq - K(q); \quad K(q) = K_\varepsilon(aq) \quad [7]$$

We can see that the fluctuations of typical observables have an extra linear scaling term Hq and we have introduced the (generalized) structure function exponent $\xi(q)$ (the usual structure function is for $q = 2$, for a single realization—no ensemble averaging—the latter is called a *variogram*). From Eq. [6], we can see that the statistics of the fluctuations across distances Δx has two components, the first because it depends of the flux ε at low resolution $\lambda = L/\Delta x$, the second because of the scaling relation between the fluctuations and the flux (Δx^H); 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.

To check this behavior on the MODIS data and surrogates, we need only estimate the gradients using:

$$\Delta I_\lambda = |I(\mathbf{x} + \Delta \mathbf{x}) - I(\mathbf{x})|; \quad \lambda = L/|\Delta \mathbf{x}| \quad [8]$$

where we have assumed statistical translational invariance and isotropy. It is worth noting that this definition of the fluctuation ΔI_λ is sometimes called the “poor man’s wavelet”; other choices of definition are possible. Wavelets provide a systematic framework for this (see, e.g., Holschneider, 1995); in practice, however, the definition of Eq. [8] is usually adequate, the main restriction being that it is only appropriate when $0 \leq H \leq 1$, a condition that is usually (although not always) satisfied in geophysical applications (here we will see that $H \sim 0.18$). 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 (Kantelhardt et al., 2002). Note that by 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 (see Table 2).

Using Eq. [8] to estimate the fluctuation ΔI_λ , we obtain the generalized structure functions shown in Fig. 4a for Bands 1, 2, 6, and 7 and in Fig. 4b for the vegetation and soil moisture indices. We can see that, except for the smallest factor of 4 to 8, the (multiple) scaling is excellent. To quantify the differences in the scaling, in Fig. 5a and 5b we show the slopes that are our estimates of $\xi(q)$. We can see that (as expected) $\xi(q)$ is concave downward. It is interesting to note that while the vegetation index has a $\xi(q)$ value intermediate between its defining band $\xi(q)$ values, the soil moisture $\xi(q)$ is somewhat lower, outside the range. It is

clear from the figures that a direct quantitative intercomparison of these continuous concave curves $\xi(q)$ is not possible; we first need to reduce the problem to a finite number of manageable parameters. Although we discuss this in much more detail below, we may already make a first quantification using the exponent H . If we assume that, due to the scale-by-scale conservation of the underlying nonlinear cascade, $K(1) = 0$, we see from Eq. [6] that $H = \xi(1)$ [actually, according to Eq. [7], $H = \xi(1) + K_\varepsilon(a)$, and, in general, we don’t know a ; however, the correction $K_\varepsilon(a)$ is generally small and will be ignored below]. In Table 2, we compare these estimates, finding that across all seven bands, $H = 0.189 \pm 0.034$ while $H_{VI} = 0.162$ and $H_{SM} = 0.144$, so that the surrogates have H values that are a little lower than those of the individual bands. To gauge the significance of this effect across the range of scales in the MODIS images (a factor of 512), we see that in the mean [$\xi(q = 1) = H$], the largest and smallest fluctuations differ by a factor $512^{0.162} = 2.74$.

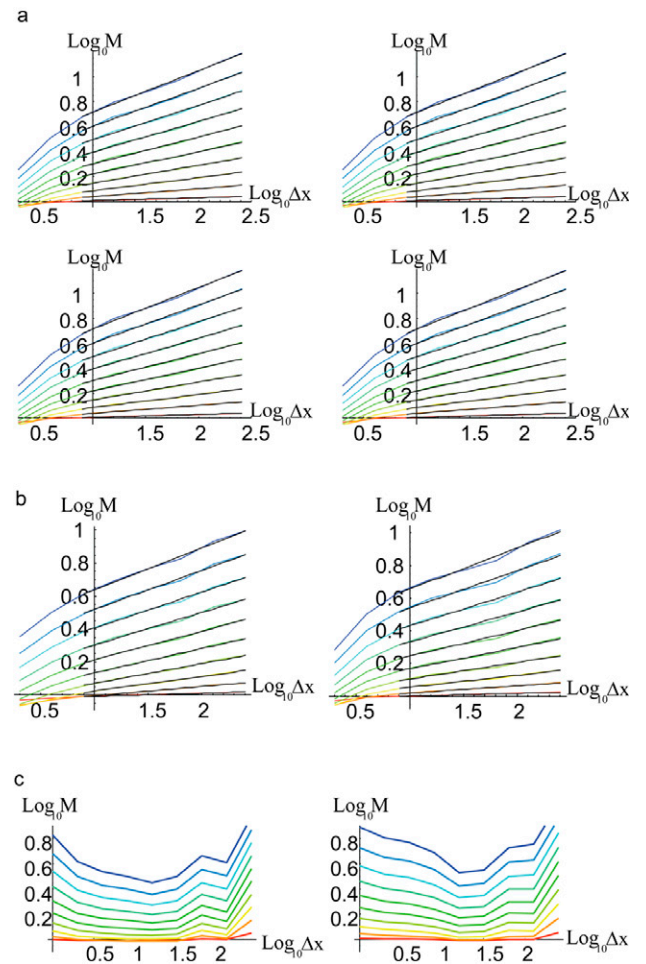


FIG. 4. (a) Structure functions for Bands 1, 2, 6, and 7 ($M = \langle \Delta I_\lambda^q \rangle = \langle |\Delta I(\Delta x)|^q \rangle$, where Δx is the number of pixels of the “lag”; see Eq. [8]) with moments of order $q = 0.1, 0.3, \dots, 1.7$, and 1.9 shown bottom to top, along with fits across the scaling range (eight pixels up to 256); (b) left, vegetation index and right, soil moisture index single-scale surrogate (sss) with moments of order $q = 0.1, 0.3, \dots, 1.7$, and 1.9 shown bottom to top; and (c) left, vegetation index and right, soil moisture index scale by scale surrogate structure functions (mss) with moments of order $q = 0.1, 0.3, \dots, 1.7$, and 1.9 shown bottom to top.

It is helpful to compare these H values with those of other geophysical fields. We have already mentioned that the classical Kolmogorov turbulent result is $H = 1/3$. Another classical theoretical H is for passive scalars in turbulence; the Corrsin–Obukhov law also has $H = 1/3$. Empirically, the topography has $H = 0.4$ to 0.7 (oceans and continents, respectively; Gagnon et al., 2006), whereas many infrared and visible signals from volcanic and other surfaces give H values quite close to those found here (e.g., Laferrière and Gaonac’h, 1999; Harvey et al., 2002; Gaonac’h et al., 2003; see also Lovejoy and Schertzer, 2006, for cloud radiances). In addition, Lovejoy et al (2008 and unpublished data) analyzed 10 radiance bands spanning the visible, near infrared,

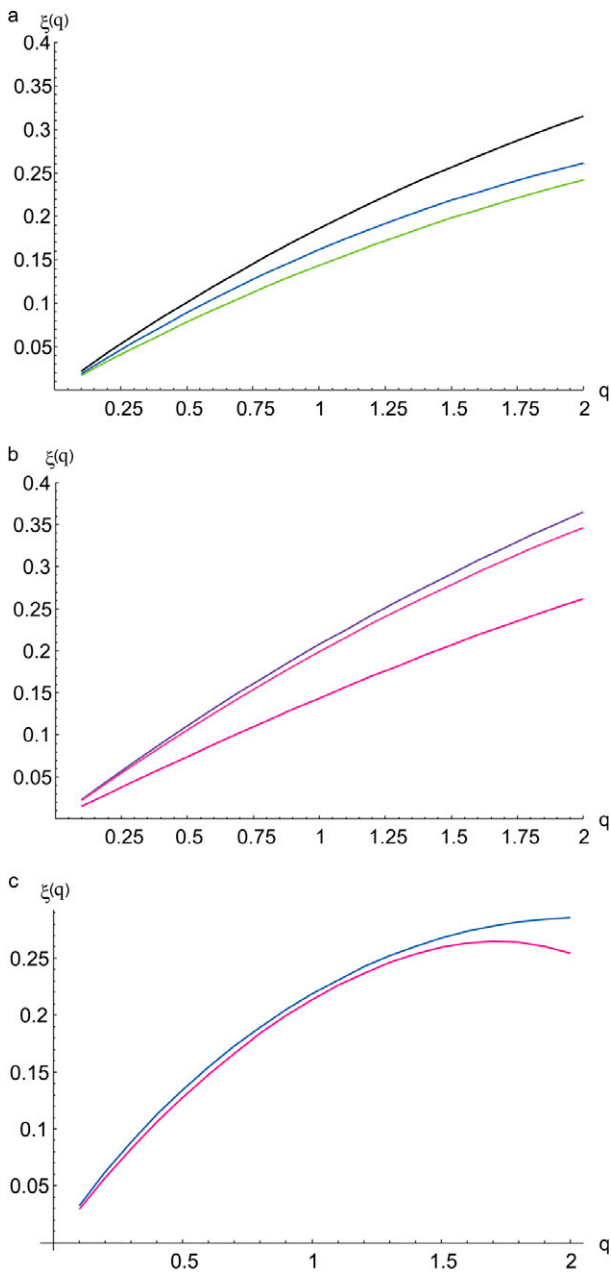


FIG. 5. Single-scale structure function exponents $[\xi(q)]$ for (a) the vegetation index (blue, middle) with Bands 1 (black, top) and 2 (green, bottom), and (b) the soil moisture index (purple, top) with Bands 6 and 7 (pink, bottom and middle, respectively); and (c) multiscale $\xi(q)$ for the vegetation index (blue, top) and the soil moisture index (red, bottom).

thermal infrared, and passive microwave wavelengths across the range ~ 8 to $20,000$ km and showed that, to within $\sim 1\%$, the latter respects the predictions of cascade models (Eq. [5]) with outer scales in the range 5000 to $18,000$ km (depending on the wavelength). Also depending on the wavelength, the corresponding H values were found to be in the range 0.2 to 0.5 .

It remains to characterize the nonlinear $K(q)$. Although we do this more fully below, we can nevertheless go a step further by characterizing the slope of the $\xi(q)$ curve near the mean, i.e., $\xi'(1) = H - K'(1)$. Defining $C_1 = K'(1)$ as the codimension of the mean (see below for the justification of this terminology), we obtained the mean for all seven bands: $C_1 = 0.032 \pm 0.010$ while $C_{1VI} = 0.033$ and $C_{1SM} = 0.010$. Although these values appear small, typical values in turbulence are only a bit larger, e.g., $C_1 \sim 0.07$ for horizontal wind (Schmitt et al., 1992, 1994) and $C_1 \sim 0.04$ for passive scalars in the horizontal (Lilley et al., 2004). The corresponding values in the vertical are about 1.8 times larger; this is as predicted by the 23/9D model of scaling stratification (Schertzer and Lovejoy, 1985a). Finally, the topography has $C_1 \sim 0.12$ (pretty much the same for both continents and oceans). For an early review of these and other results, see Lovejoy and Schertzer (1995).

Since C_1 characterizes the statistics of ε near its mean, it (partially) quantifies the effect of pure changes in satellite (pixel) resolution. For example, using the mean band value $C_1 = 0.032$, we see that for MODIS we may expect that the bias due to a 500 -m resolution with respect to an “ideal” 1 -mm-resolution satellite (i.e., assuming that the smallest scale of variability in the radiance field is 1 mm) is $(500/0.001)^{0.032} = 1.52$. This bias will be important when trying to calibrate satellite-derived quantities from in situ measurements.

Multiscale vs. Single-Scale Surrogates

We have already mentioned above that the vegetation and soil moisture surrogates are defined in an ad hoc way using the (subjective, finest available) resolution. The fact that the surrogates have different scaling $[\xi(q)s]$ implies that if the surrogates are defined at different resolutions, their statistical properties will be different; in other words, the single-scale surrogate can be (at most) correct at a single resolution. To see this more clearly, these single-scale surrogates [sss, at scale λ , $\sigma_\lambda^{(s)}$] can be contrasted with the corresponding multiscale surrogates [mss, at scale λ , $\sigma_\lambda^{(m)}$]. Mathematically, the difference can be expressed as

$$\sigma_\lambda^{(s)} = [\sigma_\Lambda^{(s)}]_\lambda; \quad \sigma_\Lambda^{(s)} = \frac{I_{i,\Lambda} - I_{j,\Lambda}}{I_{i,\Lambda} + I_{j,\Lambda}} \quad [9a]$$

$$\sigma_\lambda^{(m)} = \frac{I_{i,\lambda} - I_{j,\lambda}}{I_{i,\lambda} + I_{j,\lambda}}; \quad I_{i,\lambda} = (I_{i,\Lambda})_\lambda \quad [9b]$$

where the maximum scale ratio (satellite image scale/single pixel scale) $L/l = \Lambda$ and the notation $(I_{i,\Lambda})_\lambda$ and $[\sigma_\Lambda^{(s)}]_\lambda$ denotes averaging from this finest resolution Λ up to the intermediate resolution $\lambda < \Lambda$. The mss is the surrogate that would be obtained by applying an identical algorithm (Eq. [1]) to satellite data at the lower resolution λ , whereas the sss is the surrogate at the same resolution λ but based on the Λ scale calibration. In Fig. 4c, we show that the resulting $\xi_{mss}(q)$ obtained across the same range of scales as the single-scale surrogate $\xi_{sss}(q)$ (Fig. 4b) is quite different (it is statistically significantly larger; see below). To partially quantify this

difference, we see from Table 2 that $H_{VImss} = 0.231$ and $H_{SMmss} = 0.213$, which are larger than the sss estimates above: $H_{VIss} = 0.162$ and $H_{SMss} = 0.144$. Similarly, determining the corresponding C_1 values we find the mss values $C_{1VImss} = 0.049$ and $C_{1SMmss} = 0.059$ (c.f., the sss values $C_{1VIss} = 0.033$ and $C_{1SMss} = 0.010$, which are substantially smaller, especially the soil moisture values). Note that whenever an exponent is quoted without the indication of sss or mss, the sss value is understood; we add it explicitly here only to contrast it with the mss estimates. These statistical differences imply that if the indices are calculated at the MODIS resolution Λ and then degraded to an intermediate resolution λ , that the results would be quite different than if the same algorithms were used but directly on data at a lower resolution. The problem gets worse as the range of scales increases because the scaling exponents are different. For example, the exponent difference $H_{SMmss} - H_{SMss} = 0.213 - 0.144 = 0.069$ so that the difference in the mean index fluctuations across the range of scales of the MODIS image is about a factor $500^{0.069} \sim 1.54$. Alternatively, considering the effect of the inadequate MODIS resolution [500 m rather than, say, a scale of ~ 1 mm, where “ideal” indices might be defined, i.e., $\lambda = (500/0.001)$], then the ratio between the bias due to pixel size in the soil moisture statistics near the mean is $\lambda^{C_{1SMmss} - C_{1SMss}} = 1.90$ (with $C_{1SMmss} - C_{1SMss} = 0.059 - 0.01 = 0.049$). We therefore conclude that due to the scale dependence of the data, the surrogate indices cannot be valid across a very significant range of scales. This scale dependence must at the very least be taken into account during calibration with in situ data, which are typically at much smaller resolutions. This underscores the need to develop scale-invariant algorithms based on the scale-invariant exponents; see Lovejoy et al. (2001) for a discussion.

Trace Moment Analysis

We have seen that the generic statistical properties of nonlinear processes that 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 Hq . Since at least for low q the linear term Hq is often larger than the nonlinear $K(q)$, in analyses it can mask the nonlinear term. It is therefore advantageous to first estimate the conserved flux ε from the observed v and then estimate $K(q)$ directly. From Eq. [6], 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' and fractionally integrate it by H (a power-law filter k^{-H}), the resulting field will have the fluctuation statistics indicated by Eq. [6] (see Marsan et al., 1996). This suggests that, to obtain a flux from v (i.e., a conserved field with $H = 0$), it suffices to invert the power-law filter, i.e., to fractionally differentiate by an order H . It turns out that a finite difference approximation to this is obtained by an integer order differentiation $H' > H$ followed by taking the absolute value of the result (the absolute value is necessary since the multiplicative cascade $\varepsilon > 0$; see, however, complex and vector cascades [Schertzer and Lovejoy, 1995]). Since, usually, $0 \leq H \leq 1$, a first-order finite difference is sufficient. One simply takes the absolute differences at the finest available resolution ε_Λ and degrades them:

$$\varepsilon_\lambda = (\varepsilon_\Lambda)_\lambda; \quad \varepsilon_\lambda \propto |\Delta I_\Lambda|; \quad 0 < H < 1 \quad [10]$$

where again the notation $(\varepsilon_\Lambda)_\lambda$ indicates the average of the finest resolution data ε_Λ across the intermediate resolution λ . The

proportionality constant (l^{-H}) is unimportant (see [Eq. 6] with v replaced by I and with $a = 1$, $\Delta x = l$). In one dimension (e.g., series), one can simply use $\Delta I_\Lambda(x) = I(x+l) - I(x)$ where the inner scale is $l = L/\Lambda$, where L is again the outer scale. In higher dimensions on data on regular grids (as here), if the data are fairly isotropic, then we can simply treat each line separately and then use the one-dimensional method; however, there are more possibilities. For example, we may use finite difference moduli of gradient vectors, or finite difference Laplacians (the latter being an isotropic order 2 derivative). In two dimensions, this yields, respectively,

$$|\Delta I_\Lambda(x, y)| \sim \left\{ [I(x+l, y) - I(x, y)]^2 + [I(x, y) - I(x, y-l)]^2 \right\}^{1/2} \quad [11]$$

$$|\Delta I_\Lambda(x, y)| \sim |I(x, y) - [I(x+l, y) + I(x-l, y) + I(x, y+l) + I(x, y-l)]/4|$$

($l = L/\Lambda$ is the interpixel distance). These and other variants have been extensively tested on numerical simulations and on data (see, e.g., Tessier et al., 1993; Lavallée et al., 1993); there is generally not much difference between the various choices (or between direct fractional differentiation using Fourier techniques). Due to analogies with quantum statistical mechanics, the methods based on degrading the resolution of the fluxes and averaging across the available samples is called the *trace moment technique* (Schertzer and Lovejoy, 1987). It is distinguished from similar partition function methods (Halsey et al., 1986) in that, in addition to spatial averaging, it also involves ensemble averaging. Since here we apply the method to individual realizations (images), the two are equivalent.

In Fig. 6a and 6b, we show the results on the sss indices. The plot shows various statistical moments of ε_λ at various intermediate resolutions obtained by spatially averaging ε_Λ obtained from the modulus of the gradient vector across larger scales (smaller scale ratios) λ . Since the intermediate scales Δx are inversely proportional to the scale ratio ($\lambda = L/\Delta x$), the smallest scales are on the right, the largest scales are on the left. As a first remark, we see that the overall range of variability is quite a bit smaller than for the corresponding structure functions; this is because for the moments shown ($q < 2$), the Hq term is larger than the nonlinear $K(q)$ term. Notice that for $q > 1$, the effect of spatial averaging is to decrease the values ($K > 0$), whereas for $q < 1$, it increases them ($K < 0$). Next, we can see that, due to the nonlinear operations at the smallest scales (the definition in terms of the various bands, the modulus of the gradient operator), the scaling is not so good at the finest resolutions; the exponents $K(q)$ (the slopes on the log-log figure) were estimated across the intermediate range shown. At the very largest scales, the scaling is also poor partly because of poor statistics (there aren't very many large-scale structures), but also because of the low wave-number atmospheric corrections mentioned above.

To quantify the scale-by-scale statistical differences between the bands and the indices, we can use the scaling of the moments (the slopes in Fig. 6) to estimate $K(q)$ and compare the various moment scaling exponents; this is done in Fig. 7a, 7b, and 7c.

Again, we can quantify the differences by the behavior near $q = 1$; numerically evaluating the slope $C_1 = K'(1)$, we obtain a second series of estimates shown in Table 2. For the bands, these are very close to the previous estimates (for the seven bands: $C_1 = 0.0367 \pm 0.001$, c.f. the previous value 0.032 ± 0.010), but for the surrogates, the values are somewhat different: $C_{1VI} = 0.064$ and $C_{1SM} = 0.053$, compared with the structure function estimates of 0.033 and 0.010, respectively. The fact that the two methods agree so well for the bands but not so well for the surrogates is due to the poorer scaling of the latter, itself due to their scale-dependent definitions.

Multiscaling the Probabilities and the Probability Distribution Multiple Scaling Technique

We have concentrated on the statistical moments as a simple method of characterizing scale-invariant fields; however, the moments are integrals of the probability densities, and the multiscaling of the former implies the multiscaling of the latter. In addition, under fairly general conditions, the relationship can be inverted so that knowledge of the scaling of the moments determines the scaling of the probabilities. In particular, we obtain

$$\Pr(\varepsilon_\lambda > \lambda^\gamma) \sim \lambda^{-c(\gamma)} \quad [12]$$

which states that the probability (Pr) that a (resolution λ) value ε_λ exceeds the threshold λ^γ is a power of the resolution with exponent $-c(\gamma)$ (Schertzer and Lovejoy, 1987). The approximation sign is to within constant or slowly varying factors (e.g., including subexponential factors such as $\log \lambda$). Note that although, in the above, we used the expression *probability distribution*, it is not strictly appropriate since it represents the integral of the probability density from λ^γ to ∞ rather than the usual $-\infty$ to λ^γ (the

latter would correspond to a true cumulative distribution function). The exponent γ is called a *singularity* since it quantifies the rate at which the field values ε_λ diverge in the small-scale limit $\lambda \rightarrow \infty$ (if $\gamma < 0$, it is in fact a *regularity*). The function $c(\gamma)$ defined via Eq. [12] is called the (statistical) codimension function (Schertzer and Lovejoy, 1987) or sometimes the Cramer function (Mandelbrot, 1989). If it is low enough [$c(\gamma) < d$, where d is the dimension of the embedding space; $d = 2$ for single images], then $d(\gamma) = d - c(\gamma)$ is the (geometrical) fractal dimension of the set of points (at resolution λ) with $\varepsilon_\lambda > \lambda^\gamma$. While the geometrical $d(\gamma)$ must satisfy $0 \leq d(\gamma) \leq d$, the only restriction on $c(\gamma)$ is that it is nonnegative.

The use of c in place of d is necessary in stochastic processes since events with $c > d$ [i.e., those that would have an impossible negative geometrical dimension $d(\gamma) < 0$] correspond to events that are almost surely absent in a d -dimensional space but on the contrary are almost surely present in a large enough sample. Indeed, if we have N_s independent realizations of the process, each across a range of scales λ , then the effective or sampling dimension of the sample is $D_s = \log N_s / \log \lambda$, and one would

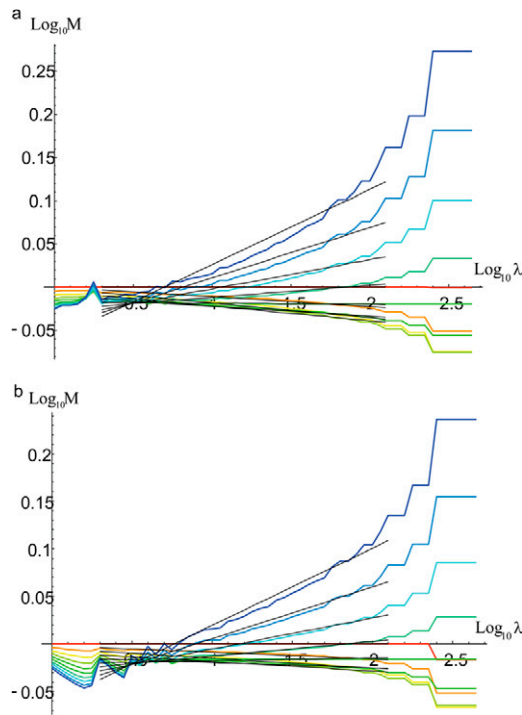


FIG. 6. Log-log plot of trace moments ($\langle \varepsilon_\lambda^q \rangle$) using 20 values of scale ratio λ per order of magnitude in scale: (a) vegetation index; (b) soil moisture index. From bottom to top, $q = 0.2, 0.4, \dots, 1.6$, and 1.8.

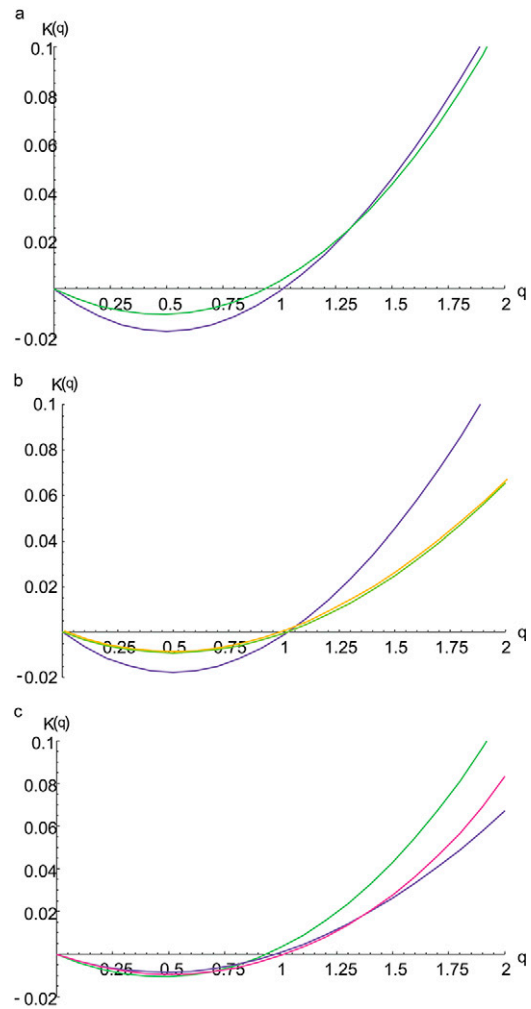


FIG. 7. Trace moment determination of the scaling exponent of the q th order moment of the field $[K(q)]$ for: (a) vegetation index (purple, top right) and soil moisture index (green, bottom right); (b) vegetation index (purple, top right) and Band 1 (orange, middle right) and Band 2 (green, bottom right); and (c) soil moisture index (green) and Band 6 (blue, bottom right) and Band 7 (purple, middle right).

expect to find all levels of activity up to extreme events such that $c(\gamma_s) = d + D_s$, where γ_s is the *sampling singularity* (Schertzer and Lovejoy, 1989). In particular, there will be a largest moment $q_s = c'(\gamma_s)$ that can be estimated from a finite sample of N_s realizations; for $q > q_s$, the $K(q)$ becomes (spuriously) linear following the tangent $K'(q_s)$.

As mentioned above, knowledge of $K(q)$ is equivalent to knowledge of $c(\gamma)$ and visa versa; indeed, the two are conveniently related via Legendre transformations (Parisi and Frisch, 1985):

$$\begin{aligned} K(q) &= \min_{\gamma} [q\gamma - c(\gamma)] \\ c(\gamma) &= \min_q [q\gamma - K(q)] \end{aligned} \quad [13]$$

This establishes a one-to-one correspondence between orders of singularity γ and statistical moments q : $\gamma = K'(q)$ and $q = c'(\gamma)$. We can now see the justification for the term $C_1 = K'(1)$ introduced above; in a precise sense, it corresponds to the singularity that gives the dominant contribution to the mean; in addition, since $K(1) = 0$, we find $C_1 = c(C_1)$ is also the codimension of the $\gamma = C_1$ singularities.

Before using Eq. [12] to directly estimate $c(\gamma)$, we recall that the approximation sign in Eq. [12] means that using the approximation $c(\gamma) \sim -\log \text{Pr} / \log \lambda$ is not so accurate. It is therefore better to calculate the histograms at a series of resolutions λ and then, for a fixed $\gamma = \log \varepsilon_{\lambda} / \log \lambda$, regress $\log \text{Pr}$ against $\log \lambda$; the slope is $-c(\gamma)$. This is the probability distribution multiple scaling technique (Lavallée et al., 1991). We show the result in Fig. 8a, 8b, and 8c. As can be seen at the larger γ (the extremes), the statistics become poor so that even though, in principle, there is information up until $c = d = 2$, it is very noisy. The extremes can be examined by determining the histograms directly at the highest resolution. This is especially important since, in general, as mentioned above, in general “canonical” cascade processes there is a fundamental distinction that must be made between the “bare” cascade, quantities $\varepsilon_{\lambda,b}$, i.e., that developed from the outer scale L down to a smaller (but finite) scale l , and the “dressed” process, $\varepsilon_{\lambda,d}$, obtained by averaging (“dressing”) a bare process down to the small-scale limit (this is a good approximation to measured fields, which typically have variability much below that of their averaging resolution). While the former has all its positive moments converge [the corresponding bare $K_b(q)$ is always finite], the dressed process is more variable; it generally has diverging moments for all moments $q > q_D$ where q_D is a critical exponent dependent on the dimension D across which the process is averaged (Mandelbrot, 1974; Schertzer and Lovejoy, 1987) ($K_d(q) = K_b(q)$ for $q < q_D$). In general,

$$\begin{aligned} \left\{ \text{Pr}(\varepsilon_{\lambda,d} > s) \sim s^{-q_D}; \quad s \gg 1 \right\} \leftrightarrow \\ \left\{ \varepsilon_{\lambda,d}^q \rightarrow \infty; \quad q > q_D \right\} \end{aligned} \quad [14]$$

In practice, the moments of a finite data set are always finite; nevertheless, the divergence of moments implies that the $K(q)$ estimated for $q > q_D$ will be dominated by a single (largest) singularity, the corresponding scaling will be spurious, and there will be discontinuity in the slope of $K(q)$ at $q = q_D$. Since there is a mathematical analogy between classical thermodynamics and multifractals, the discontinuity $K'(q)$ at $q = q_D$ is

called a *first-order multifractal phase transition*. Another term for the phenomenon is the *multifractal butterfly effect* (Lovejoy and Schertzer, 1998), a term that is justified because—just as for the classical deterministic chaos “butterfly effect”—the large-scale moments for $q > q_D$ are in fact determined by the small-scale details. Finally, the divergence of moments occurs in combination with scaling fractal structures (indeed, it is the because of the buildup of such strong variability from large to small scales that the spatial averaging across finite sets cannot sufficiently “calm” the higher order statistical moments). Elsewhere in statistical physics, the combination of fractal structures with algebraic extreme probabilities has been called *self-organized criticality* (SOC; Bak et al., 1987); hence, we see that multifractals provide a “nonclassical” route to SOC (Schertzer and

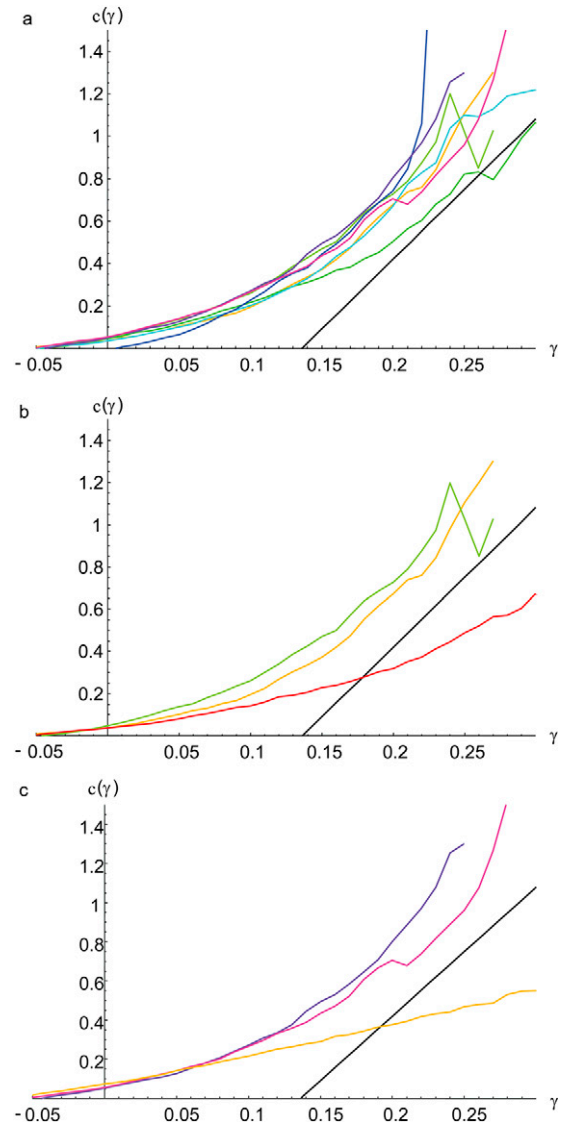


FIG. 8. The codimension function [$c(\gamma)$] of the set of points with singularities γ for: (a) the seven bands, in order from top to bottom at the point $\gamma = 0.23$ (below the large spike in the blue curve): blue (with spikes) = Band 5, purple = Band 6, dark green = Band 3, cyan = Band 4, orange = Band 1, magenta = Band 7, and light green = Band 2; (b) the vegetation index (red) and Bands 1 and 3 (yellow and green, respectively); and (c) the soil moisture index (orange, bottom line) and Band 6 (purple, top) and Band 7 (magenta, middle). The straight reference lines (black) have slopes of 6.6.

Lovejoy, 1994). Actually, this nonclassical SOC may be more generally physically relevant since it based on a (quasi-) constant flux, whereas the classical SOC is in fact only strictly valid in the (often unrealistic) zero-flux limit.

To check for the divergence of moments, we therefore calculated the probability distributions of ε_λ (see Fig. 9a and 9b). Due to the slight excess smoothing at the highest factor of 2 in scale (see the spectra, above), we averaged the (gradient-estimated) ε_λ across two by two pixels. We can see that the forms of the extreme probability tails are roughly the same for the indices and the bands from which they were derived, but it is not obvious that any asymptotic linear behavior is followed. The problem is that it is notoriously difficult to estimate the exponent of a probability tail, especially when theoretically we know neither the value of the exponent nor where the asymptotic regime begins. To quantify the tail behavior, we estimated the tail slope for the greatest factor of 2 in scale (see Table 2); for the Bands 1, 2, 6, and 7, we obtained a mean $q_D \sim 6.64$ (the mean of all the bands is 6.4 ± 1.4) and this reference slope is placed on the distributions in Fig. 9a and 9b. We see that the agreement is quite suggestive (especially when we consider that various instrumental [smoothing] effects could be responsible for artificially reducing the values of the extreme gradients used to estimate ε). The evidence for algebraic tails is apparently strongest for the soil moisture surrogate and the corresponding Bands 6 and 7. Alternatively, we note that the absolute logarithmic slope in the tails in Fig. 9 yields the maximum moments that can be accurately measured with the available sample, q_s . It may be that $q_s < q_D$ so that the data

set is simply too small to observe q_D (recall that as N_s increases, so does q_s).

A Complete, Manageable, and Physically
Motivated Characterization of the Processes:
Universal Multifractals

We have argued that a general feature of processes whose mechanism repeats across a wide range of scales is that they are scaling, characterized by convex moment functions $[K(q), c(\gamma)]$. If this was all that we could deduce, then multifractals would be quite unmanageable: theoretically, it would mean that an infinite number of parameters were important [e.g., each value of $K(q)$] or empirically it would require the determination of an infinite number of coefficients (e.g., regression slopes). In physics, however, it has generally been found that when processes are iterated sufficiently or if there is a large enough number of interactions, then only a few of their characteristics are important in the limit of many interactions or iterations. This is the idea of “universality.” Perhaps the oldest and most familiar example of universality is the central limit theorem in probability theory, which is routinely invoked to justify assumptions about Gaussianity of measurement noises. Although the Gaussian example is well known, the full result—the generalized central limit theorem (Levy, 1925) for sums of independent identically distributed random variables—is less known. It states that if one appropriately centers and normalizes sums of a sufficiently large number of independent identically distributed random variables with finite variance, the result is indeed a Gaussian; however, if the finite variance requirement is dropped (so that one allows for algebraic distributions with exponents $\alpha < 2$), then one obtains Levy distributions. This result is relevant to cascade processes because if one considers the “generator” of the process, $\log \varepsilon$, then this is the sum of the logarithms of the random cascade factors, hence one expects the generalized central limit theorem to apply to the logarithms. This would lead to logarithmic Gaussian and logarithmic Levy multifractals. Although this basic idea turns out to be correct, the nontrivial small-scale limit of the cascade process has obscured the issue, even leading to strong statements that there are no universal multifractal properties (Mandelbrot, 1989). Indeed, to obtain universal properties, one must consider the universality issue on cascades developed only across a finite range of scales, and only then—after central limit convergence has been achieved—consider the small-scale limit (see Schertzer and Lovejoy, 1997). The resulting bare universal multifractals have the following exponent functions:

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

$$c(\gamma) = C_1 \left(\frac{\gamma}{C_1 \alpha'} + \frac{1}{\alpha} \right)^{\alpha'}; \quad \frac{1}{\alpha'} + \frac{1}{\alpha} = 1 \quad [15]$$

The above Legendre transformation pairs (Schertzer and Lovejoy, 1987) are valid for $0 \leq \alpha < 1$ and $1 < \alpha \leq 2$ [when $\alpha = 1$, $K(q) = C_1 q \log q$]. Also, $0 \leq C_1 \leq d$, since if $C_1 > d$, the process cannot be normalized on the subspace. The restriction to $q \geq 0$ is necessary; since generally for $\alpha < 2$, the $q < 0$ moments diverge. This is why we did not plot them in Fig. 5 and 7 (the exception is the $\alpha = 2$ lognormal multifractal). Once again, the results for the dressed cascades apply, so that these formulae are only valid for $q < q_D$ and $\gamma < \gamma_D = K'(q_D)$. Because of this divergence, the terms

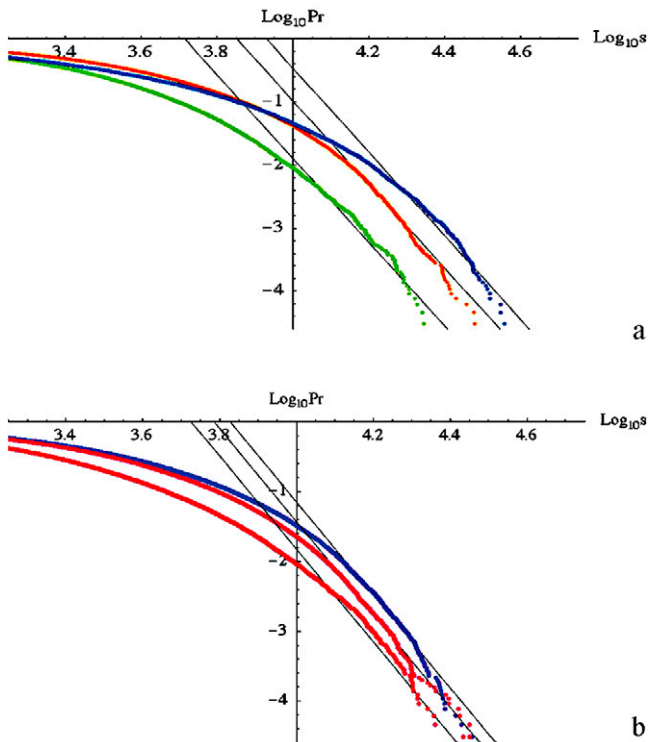


FIG. 9. Log–log plot of the probability (Pr) that a value ε_λ (indicated by the threshold s), with resolution λ , exceeds the threshold λ^γ vs. the order of singularities (γ) (see Eq. [12]): (a) vegetation index (blue, top right), Band 1 (orange, middle right) and Band 2 (green, bottom right); and (b) soil moisture index (bottom right), Bands 6 and 7 (middle and top right, respectively).

log-Levy and log-Gaussian are misnomers; the probability tails are actually somewhat stronger than those terms would imply. It could be mentioned that in the literature, a weaker form of universality has been proposed leading to log-Poisson multifractals (She and Levesque, 1994). While log-Poisson multifractals share the infinite divisibility properties of the universal (Levy-generator) multifractals (necessary for cascades continuous in scale), the corresponding processes are neither stable nor attractive so that it would be surprising for them to emerge from complex real-world processes (see Schertzer et al., 1995).

The universal form of Eq. [15] shows that only two parameters are needed to specify the conserved flux of universal multifractals (ε); for the “observables” (v), we need the third parameter (H). We have seen above that two of these parameters (H and C_1) can be estimated from $\xi(1)$ and $\xi'(1)$; the above shows that only a single additional parameter (α) is needed to characterize the entire process. This “Levy index” α can be estimated either from the radius of curvature of $\xi(q)$ [or $K(q)$] at $q = 1$ (equivalently, determined from the second derivative), or—in practice more accurately—by considering $K(q)$ or $\xi(q)$ across a wider range of q values and exploiting the dependence on q^α . We tested one of these methods (Schmitt et al., 1995) on the MODIS data (for another popular method, double trace moments, see Lavallée et al., 1993). As long as $\xi'(0)$ is finite, we may remove the linear part of $\xi(q)$ by considering the power-law “residue” $r(q)$:

$$r(q) = q\xi'(0) - \xi(q) \quad [16]$$

Applied to universal multifractals (as long as $\alpha > 1$), we have

$$r(q) = \frac{C_1}{\alpha - 1} q^\alpha \quad [17]$$

Hence we can estimate α by a linear regression of $\log_{10} r(q)$ vs. $\log_{10} q$. We show this in Fig. 10a and 10b using q values from 0.1 to 3 at intervals of 0.1 [$\xi'(0)$ is estimated from a local quadratic fit to $\xi(q)$ with $q = 0, 0.1, \text{ and } 0.2$]. As can be seen, the residue is accurately a power law; for the bands, we find (Table 2) $\alpha = 1.91 \pm 0.03$, while for the indices, we find $\alpha_{VI} = 2.02$ and $\alpha_{SM} = 2.23$, both of which are slightly outside the theoretically allowed range for universal processes (it can thus only be an approximate empirical characterization). This is presumably a further consequence of the poor scaling of the indices (but not the radiances). Finally, we compare these results to those of a nonlinear regression on the trace moment estimate $K(q)$ curve. Due to the presence of the linear term in the universal $K(q)$ (Eq. [15]), the nonlinear regression is not as accurate. We might also note that a potential problem with it is that one must only make a fit for $q < \min(q_s, q_D)$ [recall that for $q > q_s$, or $q > q_D$, $K_d(q)$ becomes linear and is spurious], and a priori, the appropriate values of q_s and q_D are unknown. Here we have seen that the most extreme q that can be used is on the order $q \sim 6$ or greater and we made our nonlinear regressions using the relatively well estimated moments $q \leq 2$.

Using these universal multifractal parameters, we can make multifractal simulations with the same statistics or resolution dependencies (Schertzer and Lovejoy, 1987; Wilson et al., 1991; Pecknold et al., 1993); see Fig. 11a and 11b for examples with roughly the same universal multifractal parameters as the radiances. For more simulations, showing the effect not only of

varying the H , C_1 , and α parameters but also of various scaling anisotropies, see Lovejoy and Schertzer (2007) and also the multifractal explorer site: www.physics.mcgill.ca/~gang/multifrac/index.htm (verified 14 Feb. 2008).

Conclusions

Due to extreme variability across huge ranges of scale, it is often difficult or impossible to obtain adequate quantities of in situ data; remotely sensed surrogates are attractive alternatives. The classical approach combines remote data from judiciously chosen radiance bands using semiempirical algorithms that are calibrated at the best available resolutions. The problem is that such algorithms are predicated on the classical geostatistical assumption that the fields of interest are sufficiently regular and smooth so that they have no significant resolution dependencies. In effect, they postulate a priori that the subpixel variability is not serious. In this study, we have argued that on the contrary, the fields are multifractal, having strong resolution dependencies (they are singular with respect to Lebesgue measures). This implies that, at best, the surrogate fields derived in this manner can be correctly calibrated at only a single resolution; they will be incorrect at other resolutions. This shows that new, resolution-independent algorithms must be developed.

In this study, we illustrated these ideas using soil moisture and vegetation indices and surrogates obtained from red and far red MODIS satellite imagery. We systematically reviewed the key notions and analysis methods that have emerged after more than 20 yr of study of cascades and other multifractal processes. First,

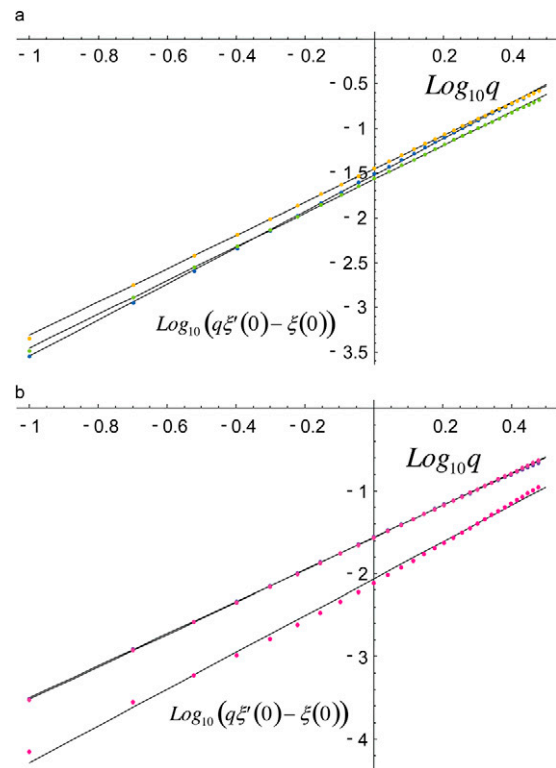


FIG. 10. The nonlinear part of the scaling exponent of the q th order moment of the field [$r(q) = q\xi'(0) - \xi(q)$] for: (a) the vegetation index (blue, bottom left) and Band 1 (orange, top left) and Band 2 (green, middle left); and (b) soil moisture (red, bottom) with Bands 6 and 7 (pink and blue, respectively, top, nearly superposed).

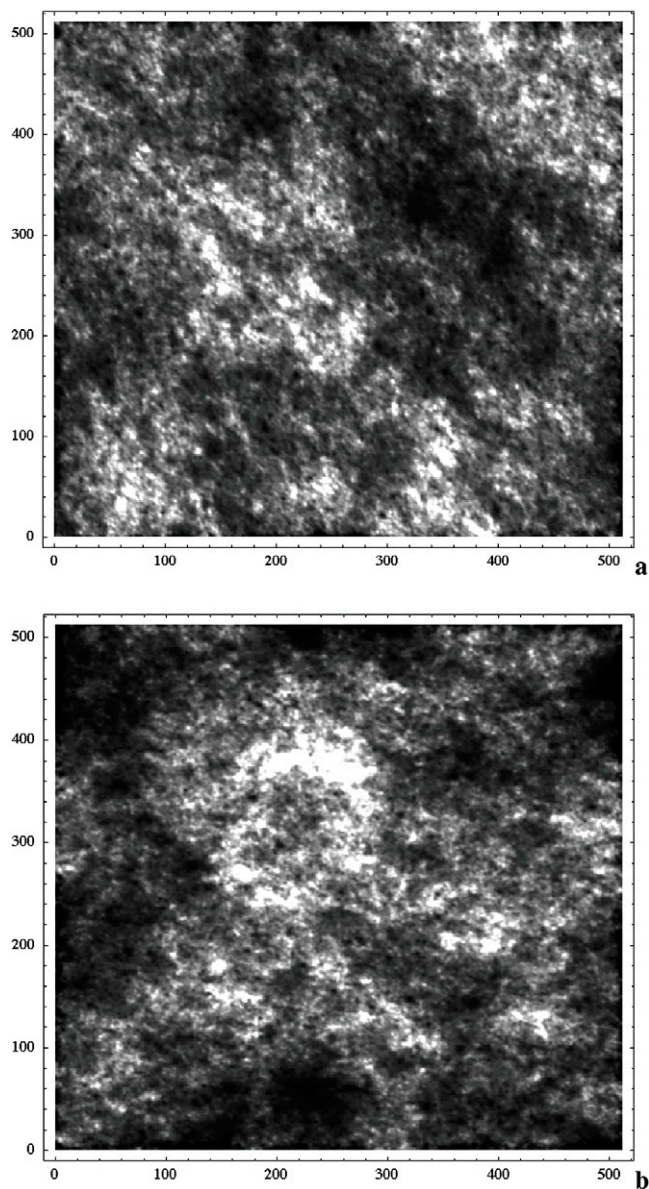


FIG. 11. Two simulations to illustrate the effect of a small amount of anisotropy with universal multifractal parameters of basic fluctuation exponent (H) = 0.18, codimension characterizing pure resolution dependence near the mean (C_1) = 0.05, and Levy index (α) = 1.9: (a) the same anisotropy at all scales, so that the scale-changing operator is a standard “zoom” but on an anisotropic initial structure, the degree and type of anisotropy is independent of scale; and (b) differential anisotropy with the direction and elongation of structures slowly changing with scale and the scale-changing operator itself anisotropic so that the degree of anisotropy changes with scale.

using the classical technique of Fourier analysis, we showed that the basic isotropic scaling was reasonably well obeyed across most of the available range of scales (here, a factor of 512). We then considered the statistics of fluctuations, which can be systematically analyzed as functions of scale and intensity (by varying the order, q) using generalized structure functions $\xi(q)$. This showed that the basic radiance bands had good multiscaling [i.e., scaling with concave $\xi(q)$], motivating the introduction of two different surrogates: the first, the classical *sss*, is defined by a nonlinear combination of radiances at the finest resolution; the second, the *mss*, is based on the same algorithm except that it is defined

on successively lower resolution satellite radiances. Although the basic algorithm is identical (the surrogates are equal to the difference in radiances at two bands divided by their sums), their resolution dependencies were found to be different. This is a consequence of the scale dependence of the radiances and the nonlinear operation defining the surrogates. The fact that the scale-by-scale properties of the *sss* and *mss* surrogates are different (a fact quantified further with the parameters H and C_1) shows that unless the finest resolution of the satellite data just happens to coincide with the inner scale of the soil moisture and vegetation indices, the surrogates cannot be more than rough approximations to the indices, valid only near the resolutions at which they were calibrated. In particular, across the MODIS range of scales (a factor 512), we showed that the amplitude of the mean statistics fluctuation in *mss* and *sss* was ~ 1.54 .

We then went on to make a more complete scale-by-scale statistical characterization of the data and surrogates, showing how to obtain and analyze the underlying conserved fluxes using a technique called *trace moments*, as well as a scale-invariant characterization of the probability distributions, called the *probability distribution multiple scaling technique*. These analysis techniques were originally developed to quantify multifractal turbulent processes and we reviewed the relevant theory while explaining the various techniques. At this level, the intercomparison of the scaling involves two exponent functions, one for the moments [$K(q)$], one for the probabilities [$c(\gamma)$]. Although these are related by a Legendre transform—so that they contain essentially the same information—they still effectively involve an infinite number of parameters (e.g., a different exponent for each K or c value). To make the problem manageable, we argued that in the real world, multifractal processes are likely to be in the basin of attraction of the three parameter “universal” multifractal processes so that analysis requires the estimation of just three parameters: in addition to H and C_1 mentioned above, the Levy index α , which characterizes the degree of multifractality. We then estimated the remaining α index. Overall (Table 2), we found the scaling parameters (H , C_1 , and α) were not too different for the different bands but were significantly different for the surrogates (which also had scaling that was less accurately followed, a consequence of the nonlinear transformation at the smallest scale). Even though the exponents H and C_1 seem small, because they act across wide ranges of scale they lead to large effects. For example, we pointed out that the difference in the mean intensity fluctuation across the MODIS image scale range (factor of 512) was about 1.47 depending on whether the indices are defined scale by scale or at a fixed scale. In contrast, there is another multifractal approach to remote sensing (Lévy-Vehel and Mignot, 1994; Lévy-Vehel, 1995) that is agnostic about the existence of wide-range scaling; in practice it only considers very small ranges of scale (e.g., factors of 2–4) so that it leads to applications that are not too different from classical remote sensing techniques.

The similitude of the scaling properties of the different radiance bands does not mean that overall they have the same statistics; the analysis techniques applied in this study are essentially isotropic. Even though by spectral analysis we showed that the anisotropy is not too large, the (unanalyzed) deviations from isotropic scaling can nevertheless lead to significant variations in the morphologies of structures. This is perhaps most graphically shown in the simulations (Fig. 11). Systematic analysis of

the scaling anisotropy is not easy, however, and is outside our present scope.

Since our conclusions are based on an analysis of a single scene (albeit at seven wavelengths with nearly 2×10^6 pixels in all), we expect to find differences when data from other sites are analyzed. Since the qualitative (multiscaling) behavior that we found is theoretically expected, and the corresponding exponents are very fundamental, we anticipate that differences will be primarily in the (possibly large, but “normal”) sample-to-sample variability of multifractal fields coupled with possible differences in the scale-by-scale anisotropy mentioned above.

Although we criticized classical scale-bound remote sensing algorithms, we did not propose a specific alternative. The general problem is to obtain a scale-invariant characterization of the statistical interrelation between the various radiances or fields. The basic approach is to consider a “state vector,” with each relevant field represented by a different component. We are thus led to scale-invariant vectors (Lie cascades, Schertzer and Lovejoy, 1995), an approach that has been discussed to some extent in Lovejoy et al. (2001); however, very little is known about vector multifractals or about the appropriate analysis techniques.

Appendix: Symbol Definitions

I_i	Intensity in band i (Eq. [1])
σ_{VI}, σ_{SM}	Vegetation index (VI), soil moisture index (SM) (Eq. [2])
P	Spectral density (Eq. [2])
E	Spectrum (angle integral of P ; Eq. [3])
\mathbf{k}	Wave vector (Eq. [2])
\mathbf{r}	Position vector (Eq. [2])
β	Spectral exponent (Eq. [4])
L	Largest scale in the system, here typically the image scale
l	Smallest scale in the system, here typically the pixel scale
λ	Scale ratio: largest/resolution scale
Λ	Maximum scale ratio: L/l , $\Lambda \geq \lambda$
ε_λ	Conserved flux associated with an observable, such as an intensity, at resolution λ (Eq. [5])
$\varepsilon_{\lambda, b}$	The “bare” ε_λ , i.e., the result of a cascade over a ratio λ that stops at scale λ
$\varepsilon_{\lambda, d}$	The “dressed” ε_λ , i.e., the result of a cascade down to infinitely small scales that is then averaged to resolution λ
$K(q)$	Scaling exponent of the q th order moment of field quantifying how the q th order statistics change with resolution λ (Eq. [5])
q	Order of statistical moments (e.g., Eq. [5])
q_s	Highest order statistic that can be reliably estimated with N_s samples at dimension d .
q_D	Critical dimension for the divergence of statistical moments, equal to the asymptotic (absolute) probability distribution exponent
Δv_λ	The velocity radiance across a distance $\Delta x = L/\lambda$ (Eq. [6])
ΔI_λ	The velocity radiance across a distance $\Delta x = L/\lambda$
Δx	Horizontal distance
a	Exponent of the flux (Eq. [6])
H	Scaling exponent of the fluctuations with respect to the flux (Eq. [6]), scale-by-scale conservation exponent
$\xi(q)$	Exponent of the q th order structure function (Eq. [7])
$c(\gamma)$	(Statistical) codimension of the set of points with singularities γ (Eq. [12])

γ	Order of singularity corresponding to a field level ε_λ ($\gamma = \log \varepsilon_\lambda / \log \lambda$; Eq. [12])
d	Dimension of space, here (for images) $d = 2$
$d(\gamma)$	Dimension of the set with singularity γ [$= d - c(\gamma)$]
D_s	Sampling dimension = $\log N_s / \log \lambda$
N_s	Number of independent fields of dimension d , resolution λ , here it = 1
Pr	Probability
s	Dummy (threshold) variable (Eq. [14])
C_1	Codimension corresponding to the mean (i.e., $q = 1$)
α	Levy index quantifying how quickly the higher and lower order statistics depart from the mean ($q = 1$) behavior
α'	Auxiliary variable (Eq. [15])
γ_D	Critical singularity corresponding to q_D
γ_s	Critical singularity corresponding to q_s
$r(q)$	Pure nonlinear part of $K(q)$ (Eq. [17])

ACKNOWLEDGMENTS

We are grateful to Carmelo Alonso for providing the images. This work has been partially supported by CICYT under Project MTM2006-15533-C02-02.

References

- Bacry, A., A. Arneodo, U. Frisch, Y. Gagne, and E. Hopfinger. 1989. Wavelet analysis of fully developed turbulence data and a measurement of scaling exponents. p. 703–718. *In* M. Lesieur and O. Metais (ed.) *Turbulence and coherent structures*. Kluwer Acad. Press, Dordrecht, The Netherlands.
- Bak, P., C. Tang, and K. Wiesenfeld. 1987. Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.* 59:381–384.
- Beaulieu, A., H. Gaonac’h, and S. Lovejoy. 2007. Anisotropic scaling of remotely sensed drainage basins: The differential anisotropy scaling method. *Nonlin. Processes Geophys.* 14:337–350.
- Corwin, C.L., J. Hopmans, and G.H. de Rooij. 2006. From field- to landscape-scale vadose zone processes: Scale issues, modeling, and monitoring. *Vadose Zone J.* 5:129–139.
- Gagnon, J., S. Lovejoy, and D. Schertzer. 2006. Multifractal earth topography. *Nonlin. Processes Geophys.* 13:541–570.
- Gaonac’h, H., S. Lovejoy, and D. Schertzer. 2003. Resolution dependence of infrared imagery of active thermal features at Kilauea volcano. *Int. J. Remote Sens.* 24: 2323–2344.
- Halsey, T.C., M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B.I. Shraiman. 1986. Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A: Gen. Phys.* 33:1141–1151.
- Harvey, D. A., H. Gaonac’h, S. Lovejoy, and D. Schertzer. 2002. Multifractal characterization of remotely sensed volcanic features: A case study from Kilauea volcano, Hawaii. *Fractals* 10:265–274.
- Holschneider, M. 1995. *Wavelets: An analysis tool*. Oxford Univ. Press, New York.
- Kantelhardt, J.W., S.A. Zschiegner, K. Koscielny-Bunde, S. Havlin, A. Bunde, and H.E. Stanley. 2002. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A* 316:87–114.
- Kolmogorov, A.N. 1962. 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.
- Laferrière, A., and H. Gaonac’h. 1999. Multifractal properties of visible reflectance fields from basaltic volcanoes. *J. Geophys. Res.* 104:5115–5126.
- Lampkin, D.J., and S.R. Yool. 2004. Monitoring mountain snowpack evolution using near-surface optical and thermal properties. *Hydrol. Processes* 18:3527–3542.
- Lavallée, D., S. Lovejoy, and D. Schertzer. 1991. On the determination of the codimension function. p. 99–110. *In* D. Schertzer and S. Lovejoy (ed.) *Nonlinear variability in geophysics: Scaling and fractals*. Kluwer Acad. Publ., Dordrecht, The Netherlands.
- Lavallée, D., S. Lovejoy, D. Schertzer, and P. Ladoy. 1993. Nonlinear variability and landscape topography: Analysis and simulation. p. 171–205. *In* L. De Cola and N. Lam (ed.) *Fractals in geography*. Prentice Hall, Englewood Cliffs, NJ.

- Levy, P. 1925. Calcul des probabilités. Gautier-Villars, Paris.
- Lévy-Vehel, J. 1995. Multifractal analysis of remotely sensed images. p. 85–101. In G.G. Wilkinson (ed.) *Fractals in geoscience and remote sensing*. Official Publ. of the Eur. Commun., Luxembourg.
- Lévy-Vehel, J., and P. Mignot. 1994. Multifractal segmentation of images. *Fractals* 2:371–377.
- Lewis, G.M., S. Lovejoy, D. Schertzer, and S. Pecknold. 1999. The scale invariant generator technique for parameter estimates in generalized scale invariance. *Comput. Geosci.* 25:963–978.
- Lilley, M., S. Lovejoy, K. Strawbridge, and D. Schertzer. 2004. 23/9 Dimensional anisotropic scaling of passive admixtures using lidar aerosol data. *Phys. Rev. E* 70(3):036307, doi:10.1103/PhysRevE.70.036307.
- Lovejoy, S., and D. Schertzer. 1995. How bright is the coast of Brittany? p. 102–151. In G. Wilkinson (ed.) *Fractals in geoscience and remote sensing*. Office for Official Publ. of the Eur. Commun., Luxembourg.
- Lovejoy, S., and D. Schertzer. 1998. Stochastic chaos and multifractal geophysics. p. 38–52. In F.M. Guindani and G. Salvadori (ed.) *Chaos, fractals, and models '96*: Proc. Int. Conf. Univ. of Pavia, Pavia, Italy.
- Lovejoy, S., and D. Schertzer. 2006. Multifractals, cloud radiances, and rain. *J. Hydrol.* 322:59–88.
- Lovejoy, S., and D. Schertzer. 2007. Scale, scaling, and multifractals in geosciences: Twenty years on. p. 311–337. In A.A. Tsonis and J.B. Elsner (ed.) *Nonlinear dynamics in geosciences*. Springer, New York.
- Lovejoy, S., D. Schertzer, and V. Allaire. 2008. The remarkable wide range spatial scaling of TRIMM precipitation. *Atmos. Res.* doi:10.1016/j.atmosres.2008.02.016.
- Lovejoy, S., D. Schertzer, Y. Tessier, and H. Gaonac'h. 2001. Multifractals and resolution-independent remote sensing algorithms: The example of ocean colour. *Int. J. Remote Sens.* 22:1191–1234.
- Mandelbrot, B.B. 1974. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.* 62:331–350.
- Mandelbrot, B.B. 1983. *The fractal geometry of nature*. Freeman, San Francisco, CA.
- Mandelbrot, B.B. 1989. Fractal geometry: What is it and what does it do? Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, Volume 423, Issue 1864p. 3–16. In M. Fleischmann et al. (ed.) *Fractals in the natural sciences*. Princeton Univ. Press, Princeton, NJ.
- Marsan, D., D. Schertzer, and S. Lovejoy. 1996. Causal space–time multifractal processes: Predictability and forecasting of rain fields. *J. Geophys. Res.* 101(D21):26333–26346.
- Parisi, G., and U. Frisch. 1985. A multifractal model of intermittency. p. 84–88. In M. Ghil et al. (ed.) *Turbulence and predictability in geophysical fluid dynamics and climate dynamics*. North Holland, Amsterdam.
- Pecknold, S., S. Lovejoy, D. Schertzer, and C. Hooge. 1997. Multifractals and the resolution dependence of remotely sensed data: Generalized scale invariance and geographical information systems. p. 361–394. In M.G.D. Quattrochi (ed.) *Scaling in remote sensing and geographical information systems*. Lewis Publ., Boca Raton, FL.
- Pecknold, S., S. Lovejoy, D. Schertzer, C. Hooge, and J.F. Malouin. 1993. The simulation of universal multifractals. p. 228–267. In J.M. Perdan and A. Lejeune (ed.) *Cellular automata: Prospects in astronomy and astrophysics*. World Scientific, Singapore.
- Pflug, K., S. Lovejoy, and D. Schertzer. 1993. Generalized scale invariance, differential rotation and cloud texture. *J. Atmos. Sci.* 50:538–553.
- Rouse, J.W., Jr., R.H. Haas, J.A. Schell, and D.W. Deering. 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Prog. Rep. RSC 1978-1. Remote Sensing Ctr., Texas A&M Univ., College Station.
- Schertzer, D., and S. Lovejoy. 1985a. The dimension and intermittency of atmospheric dynamics. p. 7–33. In B. Launder (ed.) *Turbulent shear flow 4*. Springer-Verlag, New York.
- Schertzer, D., and S. Lovejoy. 1985b. Generalised scale invariance in turbulent phenomena. *Physicochem. Hydrodyn.* 6:623–635.
- Schertzer, D., and S. Lovejoy. 1987. Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes. *J. Geophys. Res.* 92:9693–9714.
- Schertzer, D., and S. Lovejoy. 1989. Nonlinear variability in geophysics: Multifractal analysis and simulation. p. 49. In L. Pietronero (ed.) *Fractals: Physical origin and consequences*. Plenum Press, New York.
- Schertzer, D., and S. Lovejoy. 1992. Hard and soft multifractal processes. *Physica A* 185:187–194.
- Schertzer, D., and S. Lovejoy. 1994. Multifractal generation of self-organized criticality. p. 325–339. In M.M. Novak (ed.) *Fractals in the natural and applied sciences*. Elsevier, Amsterdam.
- Schertzer, D., and S. Lovejoy. 1995. From scalar cascades to Lie cascades: Joint multifractal analysis of rain and cloud processes. p. 153–173. In R.A. Feddes (ed.) *Space/time variability and interdependence for various hydrological processes*. Cambridge Univ. Press, New York.
- Schertzer, D., and S. Lovejoy. 1997. Universal multifractals do exist! *J. Appl. Meteor.* 36:1296–1303.
- Schertzer, D., S. Lovejoy, and F. Schmitt. 1995. Structures in turbulence and multifractal universality. p. 137–144. In M. Meneguzzi (ed.) *Small-scale structures in three-dimensional hydrodynamic and magnetohydrodynamic turbulence*: Proc. Worksh., Nice, France. 10–13 Jan. 1995. Springer-Verlag, New York.
- Schertzer, D., S. Lovejoy, F. Schmitt, Y. Chigirinskaya, and D. Marsan. 1997. Multifractal cascade dynamics and turbulent intermittency. *Fractals* 5:427–471.
- Schmitt, F., D. Lavalée, D. Schertzer, and S. Lovejoy. 1992. Empirical determination of universal multifractal exponents in turbulent velocity fields. *Phys. Rev. Lett.* 68:305–308.
- Schmitt, F., S. Lovejoy, and D. Schertzer. 1995. Multifractal analysis of the Greenland Ice-Core Project climate data. *Geophys. Res. Lett.* 22:1689–1692.
- Schmitt, F., D. Schertzer, S. Lovejoy, and Y. Brunet. 1994. Empirical study of multifractal phase transitions in atmospheric turbulence. *Nonlin. Processes Geophys.* 1:95–104.
- She, Z.S., and E. Levesque. 1994. Universal scaling laws in fully developed turbulence. *Phys. Rev. Lett.* 72:336–339.
- Tessier, Y., S. Lovejoy, and D. Schertzer. 1993. Universal multifractals: Theory and observations for rain and clouds. *J. Appl. Meteorol.* 32:223–250.
- Vereecken, H., R. Kasteel, J. Vanderborght, and T. Harter. 2007. Upscaling hydraulic properties and soil water flow processes in heterogeneous soils. *Vadose Zone J.* 6:1–28.
- Viña, A., A.A. Gitelson, D.C. Rundquist, G. Keydan, B. Leavitt, and J. Schepers. 2004. Monitoring maize (*Zea mays* L.) phenology with remote sensing. *Agron. J.* 96:1139–1147.
- Wilson, J., S. Lovejoy, and D. Schertzer. 1991. Physically based modelling by multiplicative cascade processes. p. 185–208. In D. Schertzer and S. Lovejoy (ed.) *Nonlinear variability in geophysics: Scaling and fractals*. Kluwer Acad. Press, Dordrecht, The Netherlands.