Fractals, Vol. 17, No. 3 (2009) 351–363

World Scientific
www.worldscientific.com

ACCURACY OF GENERALIZED DIMENSIONS ESTIMATED FROM GRAYSCALE IMAGES USING THE METHOD OF MOMENTS

E. PERFECT
Department of Earth and Planetary Sciences
University of Tennessee, Knoxville
1412 Circle Drive, TN 37996-1410, USA
eperfect@utk.edu

A. M. TARQUIS
Dpto. de Matemática Aplicada, E.T.S. Ingenieros Agrónomos
U.P.M. Ciudad, Universitaria s.n. Madrid 28040, Spain

N. R. A. BIRD
Rothamsted Research, Harpenden
Hertfordshire, AL5 2JQ, UK

Received July 7, 2008
Accepted November 6, 2008

Abstract

The moment-based box counting method of multifractal analysis is widely used for estimating generalized dimensions, D_q, from two-dimensional grayscale images. An evaluation of the accuracy of this method is needed to establish confidence in the resulting estimates of D_q. We estimated D_q from $q = -10$ to $+10$ for 23 random geometrical multifractal fields with different grid sizes, and known analytical D_q versus q functions. The fields were transformed to give normalized grayscale values between zero and one. Comparison of the estimated and analytical functions indicated the moment-based box counting method overestimates D_q by as much as 6.9% when $q \ll 0$. The root mean square error, RMSE, for the entire range of q values examined ranged from 7.81×10^{-6} to 1.35×10^{-1}, with a geometric mean of 6.50×10^{-3}. The RMSE decreased with decreasing grid size and increasing heterogeneity. These trends appear to be

largely due to the presence of zeros in the normalized grayscale fields. Variations in the slope of the log-transformed partition function, $\ln[\chi(q,\delta)]$, with box size resulted in the overestimation of D_q when $q \ll 0$. An alternative procedure for estimating D_q was developed based on the numerical first derivatives of $\ln[\chi(q,\delta)]$. Using this approach the maximum deviation in D_q values was only 1.2%, while the RMSE varied from 3.11×10^{-6} to 2.72×10^{-2}, with a geometric mean of 2.57×10^{-4}. When analyzing normalized grayscale fields, moment-based estimates of D_q should be interpreted with care. An order of magnitude increase in the accuracy of D_q can be achieved for such fields if the numerical first derivatives of $\ln[\chi(q,\delta)]$ are used in the analysis instead of standard linear regression.

Keywords: Box-Counting; Generalized Dimensions; Method-of-Moments; Multifractal; Numerical Derivatives; Truncated Binomial Distribution.

1. INTRODUCTION

Multifractal analysis of two-dimensional grayscale fields or images has become a popular method of spatial analysis in many different disciplines. This technique has been applied to a wide variety of two-dimensional data sets. A partial list of applications includes remote sensing imagery,[1,2] magnetic resonance and X-ray medical images,[3,4] soil spatial variability,[5,6] crop yield patterns,[7,8] fracture surfaces,[9] and even abstract expressionist artworks.[10] In such applications it is often desirable to compare images in terms of their estimated multifractal parameters. Thus, it is important to know the accuracy of the particular multifractal technique employed.

Geometrical multifractals can be characterized in terms of their $f(\alpha)$ spectra[11] or their generalized dimensions, D_q.[12] These parameters are theoretically related through the Legendre transformation of the mass exponents of order q, $\tau(q)$,[13] and so it is no surprise that both have been employed, either singly or jointly, to parameterize natural systems. A number of studies have compared the relative performance of different methods of estimating either $f(\alpha)$ or D_q from empirical data.[14–16] However, it is not possible to evaluate the accuracy of any particular method using this approach.

By estimating multifractal parameters for mathematical multifractals with known $f(\alpha)$ and D_q functions it is possible to assess the accuracy of different calculation methods. Chhabra and Sreenivasa,[17] Chen *et al.*,[18] and Turiel *et al.*[19] compared estimated and analytical $f(\alpha)$ spectra for synthetic one-dimensional multifractal signals. Similar analyses performed on the generalized dimensions showed that estimates of D_q can deviate significantly from the exact result, particularly for $q < 0$.[20,21]

In contrast to studies on one-dimensional signals, very few assessments of accuracy are available for multifractal methods applied to two-dimensional fields. Meisel *et al.*[22] utilized two-dimensional binary multifractals (Koch split snowflake halls and asymmetric Koch triadic snowflakes) to investigate the effects of different box counting algorithms on the estimation of the D_q versus q function. For $q < 0$ the algorithms yielded unreliable results with estimates of D_q strongly dependent on the minimum box size considered. This result can be attributed to the use of binary rather than grayscale fields. In order to define a multifractal probability measure for binary fields, the minimum box size must be much greater than the pixel size. As a result, the range of scales available for the accurate determination of multifractal parameters from binary images is often inadequate.[23]

We are aware of only two previous studies that have investigated the accuracy of multifractal parameter estimation techniques using grayscale fields simulated with two-dimensional geometrical multifractals.[4,24] In both cases accuracy was evaluated only qualitatively by visual comparison of the estimated and analytical $f(\alpha)$ spectra. Furthermore, both studies were conducted without varying the grid size of the multifractal fields, so any effects of image resolution on accuracy could not be determined.

The purpose of this paper is to provide a quantitative assessment of the moment-based box counting method for estimating the generalized dimensions, D_q, from two-dimensional grayscale fields. Originally introduced by Halsey *et al.*[25]

this method is widely used to compute multifractal parameters.[11] Method of moment (MOM) calculations were performed on random geometrical multifractal fields with known analytical D_q functions.[26,27] The influence of varying levels of grayscale heterogeneity and grid size (resolution) on the estimation of the generalized dimensions was investigated.

2. METHODS

2.1. Generation of Multifractal Grayscale Fields

Two-dimensional random geometrical multifractal grayscale fields were constructed based on a unit square subdivided into $N(\ell) = (b^i)^2$ square grid cells of length $\ell = 1/b^i$, where b is an integer scale factor > 1 and i is the iteration level. Generator mass fractions, μ_{j_1}, are computed for the $j_1 = 1$ to b^2 grid cells of the $i = 1$ field according to[26]:

$$\mu_{j_1} = \sum_{j_1=1}^{b^2} B_T(j_1, b^2, p)\frac{1}{j_1}, \tag{1}$$

where $B_T(j_1, b^2, p) = \binom{b^2}{j_1} p^{j_1}(1-p)^{b^2-j_1} / \sum_{j_1}^{b^2} \binom{b^2}{j_1} \times p^{j_1}(1-p)^{b^2-j_1}$ is the truncated binomial probability[28] for getting j_1 "successes" or parts in a b^2 field when the selected probability is p, and $\binom{b^2}{j_1}$ is the binomial coefficient. The spatial locations of the mass fractions are then randomized within the grid.[27] Next, this randomized generator is applied onto itself. At $i = 2$ there are $j_2 = 1$ to $(b^2)^2$ grid cells. The values of the mass fractions in the $i = 2$ cells are calculated from the generator mass fractions by:

$$\mu_{j_2} = \mu_{j_1} \times \mu_{j_1}. \tag{2}$$

The spatial locations of the mass fractions at $i = 2$ are then randomized within the $i = 1$ subunits. Repetition of this procedure to the ith iteration level produces a multiplicative cascade of mass fractions or geometrical multifractal.[29,30]

By way of example, consider the following $b = 3$, $p = 8/9$, $i = 1$ generator:

$$\mu_{j_1} = \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 \\ \mu_4 & \mu_5 & \mu_6 \\ \mu_7 & \mu_8 & \mu_9 \end{bmatrix} = \begin{bmatrix} 0.127 & 0.127 & 0.127 \\ 0.127 & 0.127 & 0.124 \\ 0.115 & 0.087 & 0.038 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0.124 & 0.127 & 0.127 \\ 0.087 & 0.127 & 0.127 \\ 0.127 & 0.038 & 0.115 \end{bmatrix},$$

where $\Rightarrow$ indicates the randomization process. Based on Eq. (2), at $i = 2$ the $j_1 = 1$ cell (with $\mu_1 = 0.124$) is sub-divided into nine smaller cells of length $\ell = 1/9$ containing the following mass fractions:

$$\begin{bmatrix} 0.124 \times 0124 & 0.124 \times 0.127 & 0.124 \times 0.127 \\ 0.124 \times 0.087 & 0.124 \times 0.127 & 0.124 \times 0.127 \\ 0.124 \times 0.127 & 0.124 \times 0.038 & 0.124 \times 0.115 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0.124 \times 0.038 & 0.124 \times 0.127 & 0.124 \times 0.127 \\ 0.124 \times 0.127 & 0.124 \times 0.124 & 0.124 \times 0.115 \\ 0.124 \times 0.127 & 0.124 \times 0.087 & 0.124 \times 0.127 \end{bmatrix}$$

and so on. Once all of the j_1 cells have been subdivided, the grid at the $i = 2$ level consists of 9×9 cells containing 81 mass fractions. Extension of this multiplicative process to higher iteration levels results in randomized geometrical multifractal fields with variable grid sizes.

From Eqs. (1) and (2), the maximum, $\mu_{\max}$, and minimum, $\mu_{\min}$, mass fractions at the ith iteration level are $(\mu_1)^i$ and $(\mu_{b^2})^i$, respectively. The logarithmic range, $\lambda = \ln(\mu_{\max}/\mu_{\min})$, provides a measure of the variation in the mass fractions. For a perfectly homogeneous field, $\mu_{\max} = \mu_{\min}$ and $\lambda = 0$. Heterogeneity (as indicated by $\lambda \gg 1$) increases as both b and i increase, and as p decreases (Table 1).

In digital image processing, pixel values are often normalized to give grayscale fields ranging between zero (white) and one (black).[31] To simulate such fields, the multifractal mass fractions were transformed into grayscale values using the expression: $g_{j_i} = \frac{\mu_{j_i} - \mu_{\min}}{\mu_{\max} - \mu_{\min}}$. Examples of such fields are shown in Fig. 1 for a $b = 3$, $p = 8/9$ random multifractal generator iterated to different i levels. The effects of varying the p value while b and i are held constant have been illustrated previously.[27]

2.2. Analytical Generalized Dimensions

Here, following Perfect *et al.*,[26] we derive an analytical expression for the generalized dimensions of the random geometrical multifractal fields discussed above. The generalized moments of the ith level field, $M_i(q)$, are given by[29,32]:

$$M_i(q) = \sum_{j_i=1}^{b^{2i}} (\mu_{j_i})^q = \left(\frac{1}{b^i}\right)^{(q-1)D_q}, \tag{3}$$

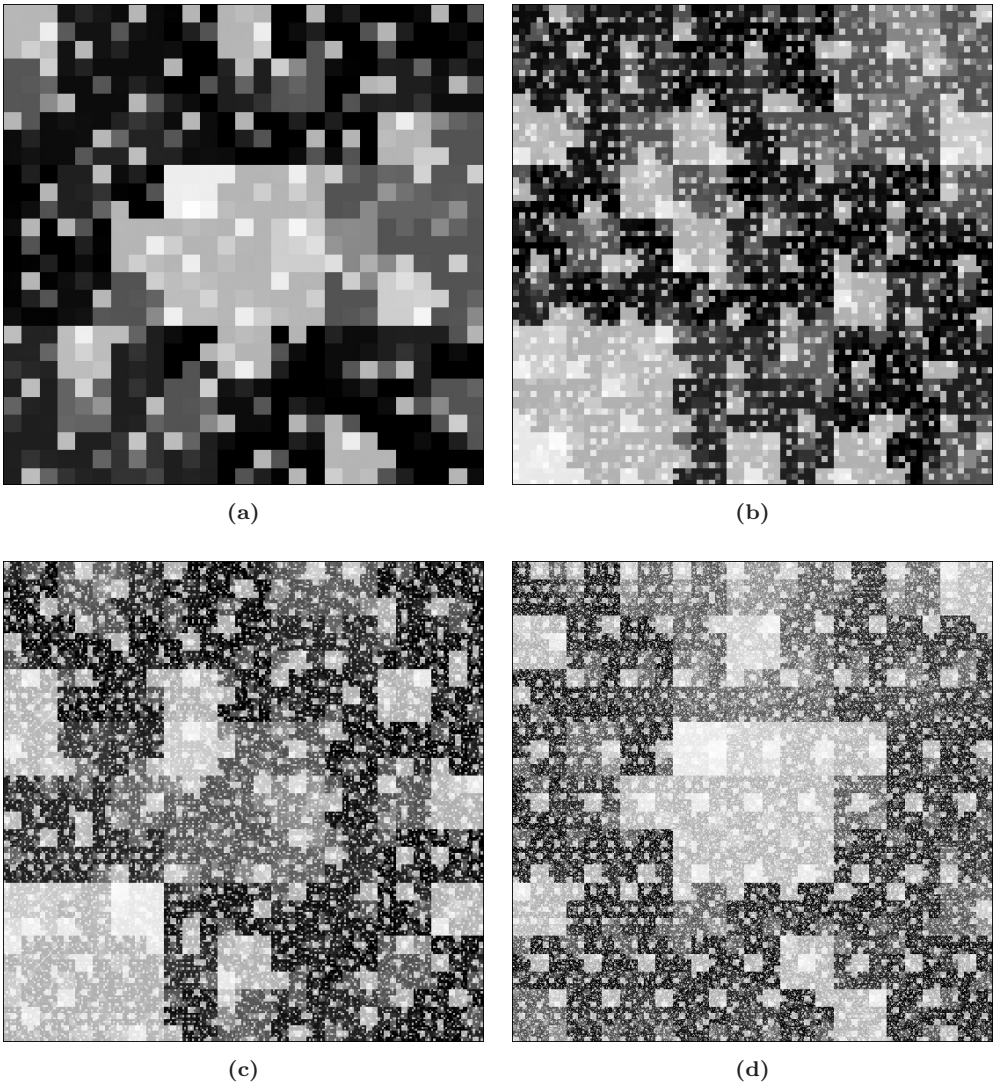


Fig. 1 Example grayscale images of random geometrical multifractal fields with $b = 3$, $p = 8/9$, and grid sizes of: **(a)** 27×27 ($i = 3$); **(b)** 81×81 ($i = 4$); **(c)** 243×243 ($i = 5$); and **(d)** 729×729 ($i = 6$).

where q is any integer between $\pm\infty$, and D_q, is the Rényi or qth order generalized dimension. Setting $i = 1$ in Eq. (3) and rearranging gives:

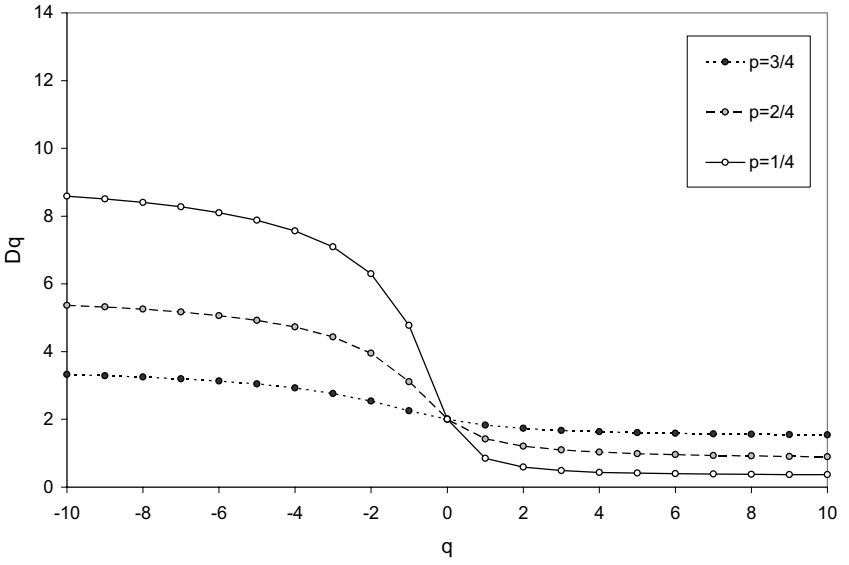
$$D_q = \frac{1}{q-1} \log \left(\left(\sum_{j_1}^{b^2} (\mu_{j_1})^q \right)^{-1} \right) \Big/ \log(b); \quad q \neq 1. \tag{4}$$

The following expression, the derivation of which can be found in Gouyet,[29] is used to compute D_q when $q = 1$:

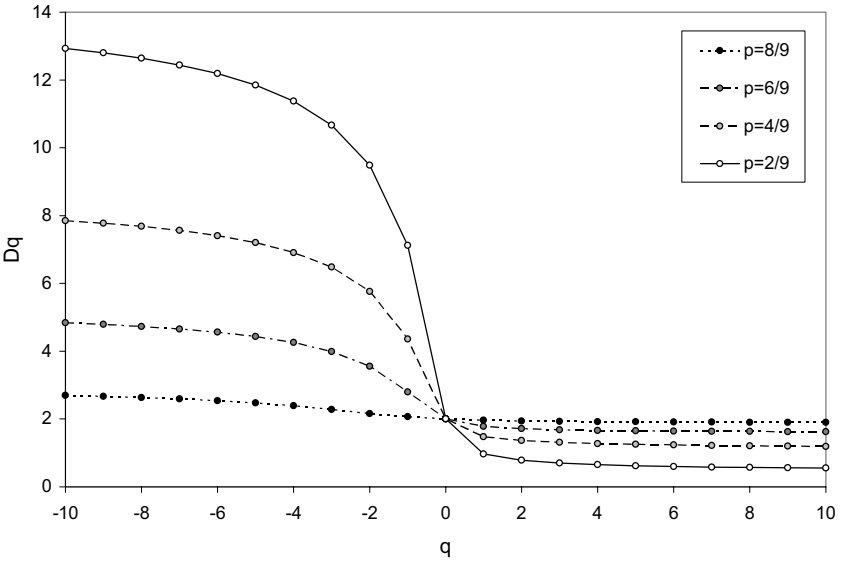
$$D_1 = -\sum_{j_1}^{b^2} \mu_{j_1} \times \log(\mu_{j_1}) / \log(b) \tag{5}$$

where D_1 is the entropy or information dimension. By substituting Eq. (1) into Eqs. (4) and (5), generalized dimensions can be calculated based on the b and p values used to generate the random geometrical multifractal fields. The resulting D_q versus q curves are independent of the iteration level.

Generalized dimensions calculated using Eqs. (3), (4) and (5) for $b = 2$ and $b = 3$ random geometrical multifractal fields with different p values are shown in Fig. 2. For all the fields considered here $D_q = 2$ when $q = 0$ because the initiator was a unit square. It can be seen that the D_q versus q functions are more sensitive to p when q is negative than when it is positive. The D_q maxima (as $q \to -\infty$) decrease with increasing p value, while the minima (as $q \to +\infty$) increase. The range (maximum minus minimum) in possible D_q values increases with increasing b value.



(a)



(b)

Fig. 2 Analytical generalized dimensions as a function of q for different values of p: **(a)** $b = 2$; and **(b)** $b = 3$.

2.3. Empirical Generalized Dimensions

Empirical estimates of the generalized dimensions for the grayscale fields listed in Table 1 were obtained using the method of moments (MOM). This box-counting-based approach has been described many times previously[5,11,13,15,16,25] and the reader is referred to these publications for details. Briefly, the normalized grayscale values were converted into probability density values using the transformation $\rho_{j_i}(\ell) = g_{j_i} / \sum_{j_i=1}^{b^{2i}} g_{j_i}$. Grids comprised of square cells (boxes) of varying length, $\ell < \delta < 1$, were superimposed on each field. The probability density value of the kth box in a superimposed grid, $\rho_k(\delta)$, was calculated as $\sum_{j_i=k}^{k+(\delta/\ell)^2} \rho_{j_i}(\ell)$, where $1 \leq k \leq \delta^{-2}$. The δ values varied with the b and i values for each field, i.e. $\delta = 1/b^{i-m}$, with $m = 1, 2, 3 \ldots i$.

A weighted summation was performed over all boxes in a particular grid yielding the partition

Table 1 Parameters Used to Generate the Random Geometrical Multifractal Grayscale Fields and Associated Grid Statistics.

Field #	b	p	i	N	ℓ	λ
1	2	3/4	3	64	1/8	4.6
2	2	3/4	4	256	1/16	6.2
3	2	3/4	5	1024	1/32	7.8
4	2	3/4	6	4096	1/64	9.2
5	2	3/4	7	16,384	1/128	10.8
6	2	2/4	3	64	1/8	10.6
7	2	2/4	4	256	1/16	14.0
8	2	2/4	5	1024	1/32	17.7
9	2	2/4	6	4096	1/64	21.2
10	2	2/4	7	16,384	1/128	24.9
11	2	1/4	3	64	1/8	18.9
12	2	1/4	4	256	1/16	25.3
13	2	1/4	5	1024	1/32	31.5
14	2	1/4	6	4096	1/64	38.0
15	2	1/4	7	16,384	1/128	44.2
16	3	8/9	3	729	1/27	3.7
17	3	8/9	4	6,561	1/81	4.8
18	3	8/9	5	59,049	1/243	6.0
19	3	8/9	6	531,441	1/729	7.1
20	3	8/9	7	4,782,969	1/2187	8.3
21	3	6/9	7	4,782,969	1/2187	28.8
22	3	4/9	7	4,782,969	1/2187	58.0
23	3	2/9	7	4,782,969	1/2187	105.5

Note: b = scale factor, p = probability in truncated binomial distribution, i = iteration level, N = number of cells, ℓ = cell length, λ = logarithmic range.

function of order q, i.e.

$$\chi(q,\delta) = \sum_{k=1}^{\delta^{-2}} (\rho_k(\delta))^q. \tag{6}$$

For a multifractal measure, the partition function scales with the box length according to:

$$\chi(q,\delta) \propto \delta^{-\tau(q)}, \tag{7}$$

where $\tau(q)$ is the mass exponent or index for q. Estimates of $\tau(q)$ were obtained for $-10 \le q \le +10$ from the slopes of linear regression analyses performed on $\ln[\chi(q,\delta)]$ versus $\ln[\delta]$. For $q \neq 1$, the generalized dimensions were calculated from the $\tau(q)$ estimates using the expression: $D_q = \frac{\tau(q)}{q-1}$. For $q = 1$, D_1 was estimated from the slope of a linear regression analysis performed on $\sum_{k=1}^{\delta^{-2}} \rho_k(\delta) \ln[\rho_k(\delta)]$ versus $\ln[\delta]$.

Use of non-linear regression analysis to estimate $\tau(q)$ in Eq. (7) was explored but not adopted because in many cases the resultant Hessian matrix was singular. Instead, numerical differentiation of $\ln[\chi(q,\delta)]$ versus $\ln[\delta]$ was employed as an alternative to conventional linear regression analysis. Estimates of $\tau(q)$ for each $\ln[\delta]$, and thus $D_q(\ln[\delta])$, were obtained based on the following three point formula:[33]

$$y'(x) = \frac{y(x+h) - y(x-h)}{2h}, \tag{8a}$$

where $y = \ln[\chi(q,\delta)]$, $x = \ln[\delta]$, and h is a small positive finite difference. The two end points were calculated using a different three point scheme:[33]

$$y'(x) = \frac{-3y(x) + 4y(x+h) - y(x+2h)}{2h}, \tag{8b}$$

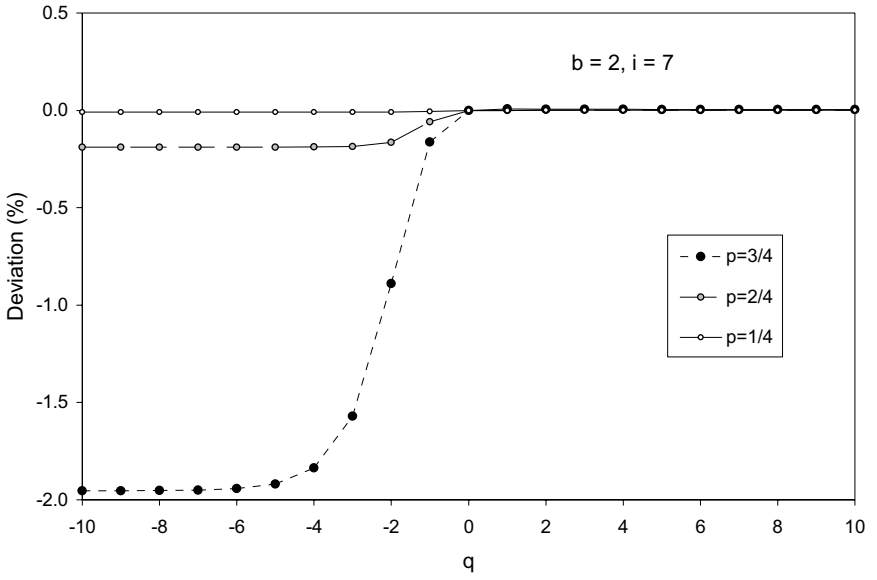
with h positive at the start and negative at the end. Second derivatives were computed by replacing y and y' in Eq. (8) with y' with y'', respectively.

3. RESULTS AND DISCUSSION

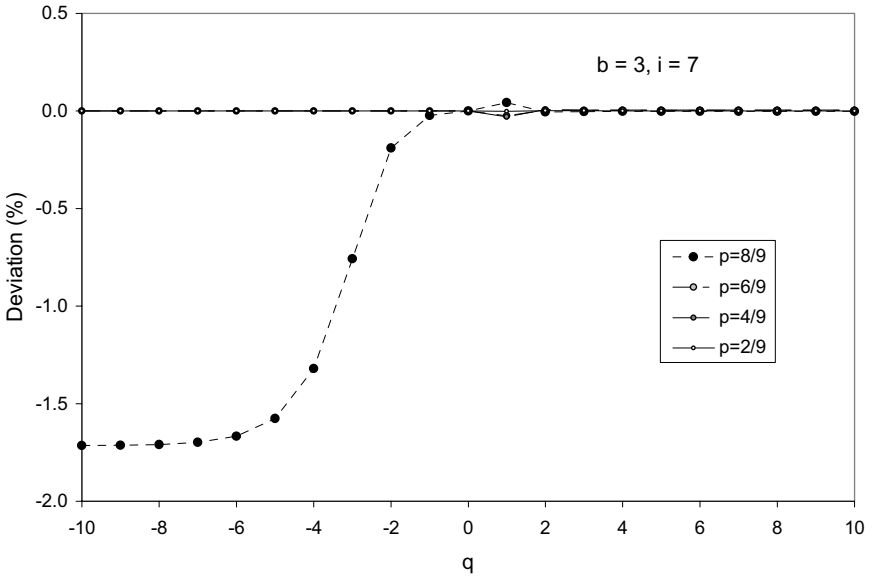
The number of box sizes (n) used to evaluate $\tau(q)$ in the linear regression analyses of $\ln[\chi(q,\delta)]$ on $\ln[\delta]$ varied from 3 to 7 (Table 2). This variation is a result of the different iteration levels employed in constructing the multifractal fields (Table 1). All of the coefficients of determination (R^2) from the regression analyses were > 0.999 signifying excellent goodness of fit regardless of the different n and q values.

Deviations of the resulting empirical D_q versus q curves from the corresponding analytical functions are shown in Fig. 3. For high p value fields there is a clear trend towards systematic overestimation (negative deviations) of the analytical function as q becomes increasingly negative. Campagna and Turchetti[20] report a similar overestimation of D_q for $q < 0$ in the case of a one-dimensional multifractal "Cantor" set. In our study, the maximum absolute deviation, $|\Delta D_q|_{\max}$, occurred when $q \le -8$ in 19 out of the 23 fields investigated (i.e. 83%) (Table 2). Overall, the $|\Delta D_q|_{\max}$ ranged from 0.002 to 6.901%, with a geometric mean value of 0.322%.

In addition to the maximum absolute deviation, a root mean squared error (RMSE) was calculated as: $\text{RMSE} = \sqrt{\sum_{q=-\nu}^{q=+\nu}((D_q)_e - (D_q)_a)^2/(2\nu+1)}$, where $\nu = 1, 2, 3, \ldots, 10$, and the subscripts "$e$" and "$a$" denote the empirical and analytical values of D_q, respectively. The RMSE statistic provides a quantitative measure of the degree correspondence between the empirical and analytical generalized dimensions integrated over the range of q values considered, i.e. $\Delta q = 2\nu$. Because of the deviations observed at negative q values (Fig. 3), the RMSE generally increased with increasing Δq. This trend



(a)



(b)

Fig. 3 Percent deviation between analytical and estimated generalized dimensions as a function of p and q for: **(a)** $b = 2$, $i = 7$; and **(b)** $b = 3$, $i = 7$ random geometrical multifractal fields.

was most pronounced for high p-value fields and $\Delta q < 10$ (Fig. 4). For $\Delta q > 10$ the RMSE was relatively stable. Therefore, only the RMSE values for $\Delta q = 20$ were investigated further. These values, simply referred to as RMSE from this point on, varied between 7.81×10^{-6} and 1.35×10^{-1}, with a geometric mean of 6.50×10^{-3} (Table 2).

The RMSE decreased with decreasing values of the ℓ and p parameters used to construct the multifractal fields (Tables 1 and 2). Grayscale heterogeneity, as quantified by λ, increased as these two parameters decreased. Figure 5 shows the resulting relationship between RMSE and λ. Greater field heterogeneity improved the degree of correspondence between the moment-based box counting estimates and the analytical values of D_q as indicated by decreased RMSE values. In terms of natural grayscale images, this result implies that the moment based box counting method should give the most accurate results for high resolution images comprised of a wide range of pixel values.

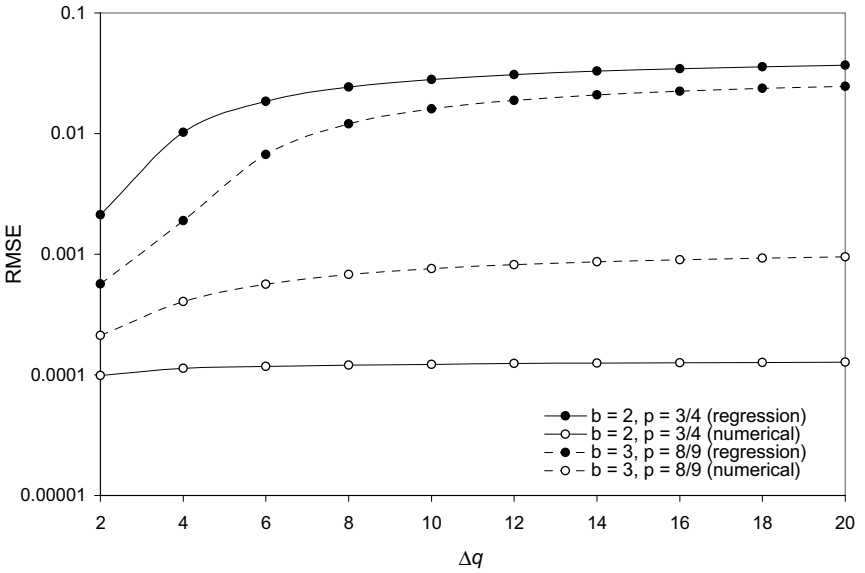


Fig. 4 Dependence of RMSE on Δq for $b = 2$, $p = 3/4$, $i = 7$; and $b = 3$, $p = 8/9$, $i = 7$ random geometrical multifractal fields. Filled and unfilled circles denote estimates based on regression analysis and numerical derivatives, respectively.

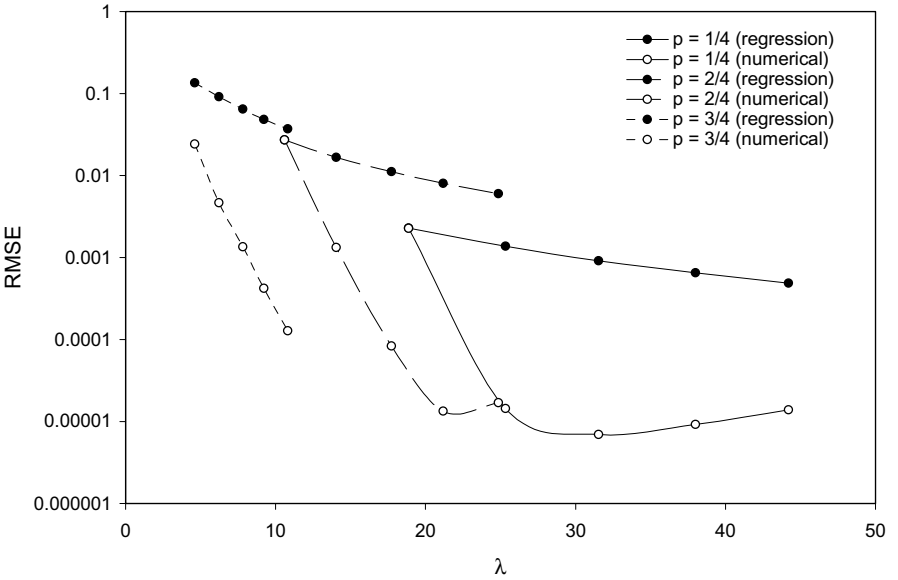
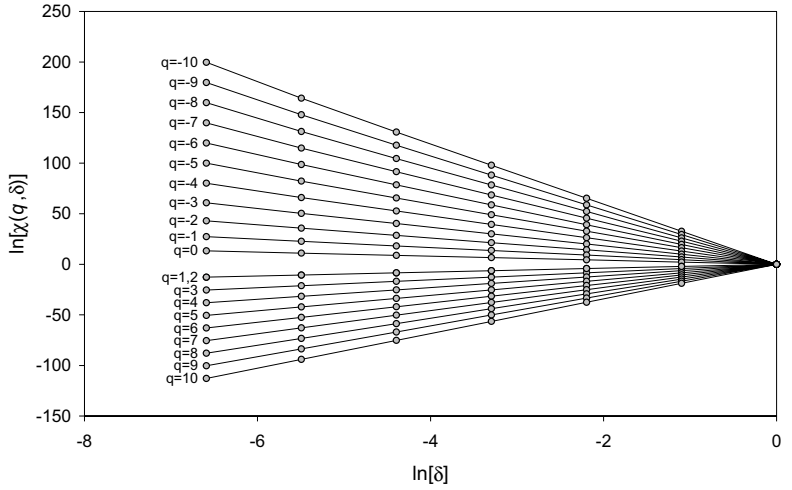


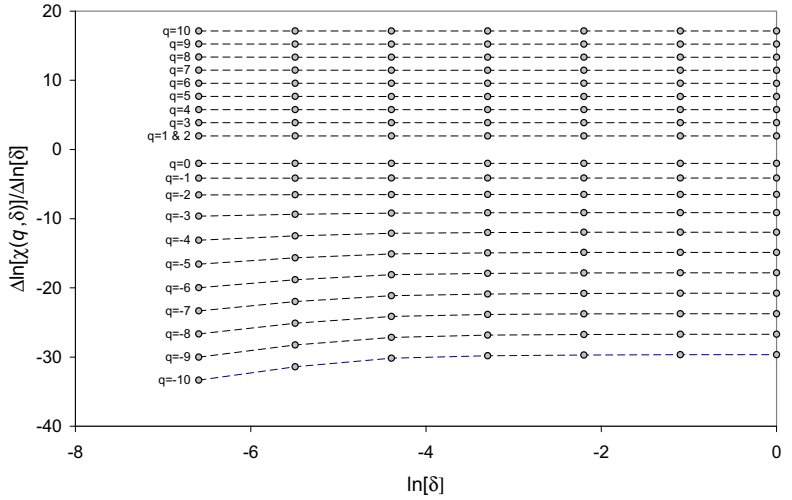
Fig. 5 Influence of heterogeneity on root mean square error (RMSE) in estimated D_q values for $b = 2$ random geometrical multifractal fields with different p values. Filled and unfilled circles denote estimates based on regression analysis and numerical derivatives, respectively.

To investigate the reasons for the poor estimation of D_q as q becomes more negative (Fig. 3) and the resulting dependency of RMSE on λ (Fig. 5), we numerically analyzed the log-transformed partition functions. A typical log-transformed partition function is shown in Fig. 6(a). Visually all of the $\ln[\chi(q, \delta)]$ versus $\ln[\delta]$ relationships appear to be linear and this explains the very high R^2 values obtained in the regression analyses. Inspection of the numerical first derivatives (Fig. 6(b)), however, indicates that the slope is not always constant. For negative q values the slope becomes more negative as $\ln[\delta] \rightarrow -\infty$. Furthermore, this trend is amplified as the q values decrease, which can be clearly seen in the plot of the second numerical derivatives as a function of $\ln[\delta]$ (Fig. 6(c)).

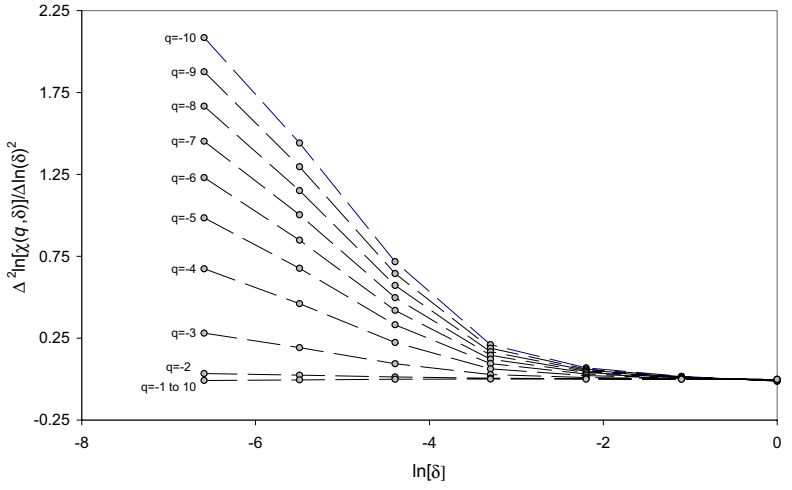
Given the behavior of the log-transformed partition function in Fig. 6, it is logical to ask the question: which of the $\ln(\delta)$ increments yields numerical first derivatives that best estimates the



(a)



(b)



(c)

Fig. 6 **(a)** Example log-transformed partition function, $\ln[\chi(q,\delta)]$ versus $\ln[\delta]$, for a random geometrical multifractal field with $b = 3$, $p = 8/9$, and $i = 7$; (**b**) first numerical derivatives of $\ln[\chi(q,\delta)]$ versus $\ln[\delta]$; and **(c)** second numerical derivatives of $\ln[\chi(q,\delta)]$ versus $\ln[\delta]$.

analytical D_q values? To answer this question we computed generalized dimensions from the numerical first derivatives for all of the $\ln[\delta]$ increments in each log-transformed partition function ($n = 3$ to 7 depending upon the field investigated). These numerical estimates of D_q consistently decreased with increasing $\ln[\delta]$. As expected from Fig. 6, the greatest changes in D_q occurred when $q = -10$. Four examples of this pronounced $\ln(\delta)$ dependency are shown in Fig. 7, which also includes the corresponding analytical D_q values for comparison. The numerically-derived D_q estimates asymptotically approach the known analytical values as $\ln[\delta] \to 1$ (i.e. as the box size approaches the field size). Considering all of the multifractal fields, the best estimates of the analytical D_q values were obtained using the numerical first derivatives computed for the largest box size increment (i.e. Eq. (8b) with h negative).

The performance of Eq. (8b) as an estimator of $\tau(q)$, and thus D_q, is summarized in Table 2. Maximum absolute differences for the numerically derived estimates of D_q ranged from < 0.001 to 1.229%, with a geometric mean $|\Delta D_q|_{\max}$ of 0.017%. In contrast to the regression-based estimates there was no clear dependency of the maximum error on q, and in only 2 cases (i.e. 8.7%) did $|\Delta D_q|_{\max}$ occur when $q \leq -8$ (Table 2). The improvement in accuracy achieved with the numerical approach, including reduced sensitivity to Δq, is readily apparent in Fig. 4. For $\Delta q = 20$ the numerical RMSE varied between 3.11×10^{-6} and 2.72×10^{-2}, with a geometric mean of 2.57×10^{-4} (Table 2). These values, along with $|\Delta D_q|_{\max}$, indicate at least an order of magnitude increase in accuracy in the estimation of D_q using the numerical first derivatives as compared to the conventional regression-based approach.

Despite the overall increase in accuracy achieved by the numerical derivative method, the resulting RMSE values still varied as a function of field heterogeneity (Fig. 5, Table 2). The least accurate estimates of D_q versus q were obtained for the most homogeneous fields (i.e. those with low b and high p and i values). As with the regression-based approach, the magnitude of the RMSE's decreased with increasing λ. However, for low p-value fields, the numerical RMSE's appear to approach a minimum beyond which there was no further improvement in the estimation of the generalized dimensions (Fig. 5).

We have observed that subtle variations in the slope of the log-transformed partition function, $\ln[\chi(q, \delta)]$, result in the systematic over estimation of D_q when $q \ll 0$. These errors appear to be largely due to the transformation used to normalize the random geometrical multifractal fields, i.e. $g_{j_i} = \frac{\mu_{j_i} - \mu_{\min}}{\mu_{\max} - \mu_{\min}}$. The impact of this transformation

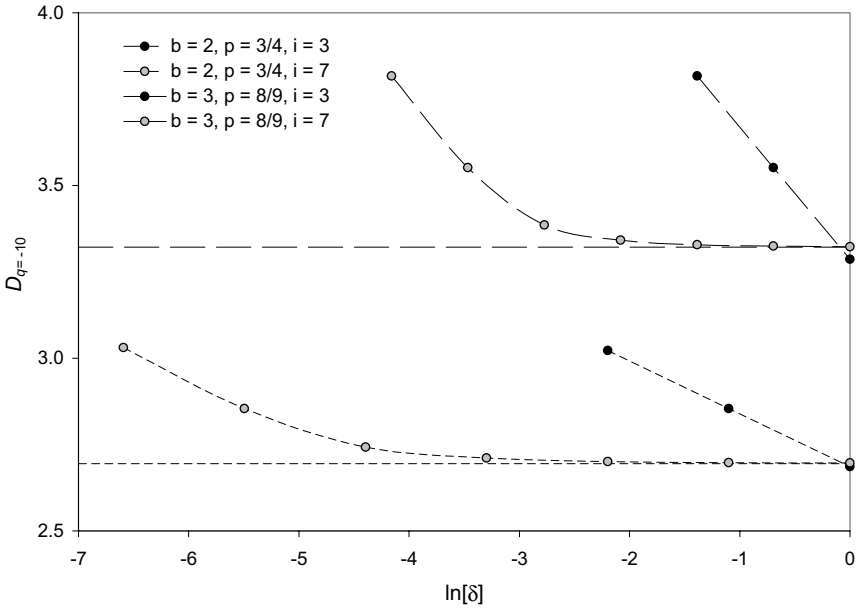


Fig. 7 Generalized dimensions for $q = -10$ computed from numerical derivatives of log-transformed partition functions for random geometrical multifractal fields with $b = 2$, $p = 3/4$, $i = 3$ and 7, and $b = 3$, $p = 8/9$, $i = 3$ and 7. Corresponding horizontal lines represent the analytical D_q values.

Table 2 Summary of Moment-Based Box Counting Analyses Performed on Random Geometrical Multifractal Grayscale Fields with $-10 \leq q \leq +10$.

Field #	n	Regression Analysis			Numerical Derivatives		
		$\lvert\Delta D_q\rvert_{\max}(\%)$	$q^{\dagger}_{\max}$	RMSE	$\lvert\Delta D_q\rvert_{\max}(\%)$	$q_{\max}$	RMSE
1	3	6.901	−10	1.35×10^{-1}	1.229	10	2.41×10^{-2}
2	4	4.696	−10	9.10×10^{-2}	0.382	10	4.65×10^{-3}
3	5	3.371	−10	6.47×10^{-2}	0.120	10	1.35×10^{-3}
4	6	2.524	−10	4.81×10^{-2}	0.038	10	4.22×10^{-4}
5	7	1.953	−10	3.70×10^{-2}	0.012	10	1.28×10^{-4}
6	3	0.839	−8	2.72×10^{-2}	0.636	−7	2.72×10^{-2}
7	4	0.516	−8	1.67×10^{-2}	0.041	−7	1.32×10^{-3}
8	5	0.348	−8	1.12×10^{-2}	0.003	−10	8.34×10^{-5}
9	6	0.250	−8	8.03×10^{-3}	0.002	8	1.34×10^{-5}
10	7	0.189	−8	6.03×10^{-3}	0.002	10	1.70×10^{-5}
11	3	0.044	−10	2.29×10^{-3}	0.043	−7	2.26×10^{-3}
12	4	0.026	−10	1.37×10^{-3}	0.002	6	1.44×10^{-5}
13	5	0.018	−10	9.15×10^{-4}	< 0.001	−1	6.93×10^{-6}
14	6	0.013	−5	6.53×10^{-4}	0.001	5	9.20×10^{-6}
15	7	0.009	−10	4.89×10^{-4}	0.003	8	1.39×10^{-5}
16	3	5.872	−10	8.93×10^{-2}	0.362	−1	7.06×10^{-3}
17	4	4.044	−10	6.05×10^{-2}	0.170	−2	1.54×10^{-3}
18	5	2.927	−10	4.32×10^{-2}	0.077	−2	8.67×10^{-4}
19	6	2.203	−10	3.21×10^{-2}	0.049	−3	7.69×10^{-4}
20	7	1.714	−10	2.47×10^{-2}	0.060	−10	9.54×10^{-4}
21	7	0.025	2	1.04×10^{-4}	0.003	10	7.48×10^{-5}
22	7	0.036	1	1.04×10^{-4}	0.004	10	1.06×10^{-4}
23	7	0.002	1	7.81×10^{-6}	< 0.001	10	3.11×10^{-6}

Note: n = number of box sizes, $|\Delta D_q|_{\max}$ = maximum absolute deviation between analytical and empirical D_q values, $q_{\max}$ = value of q corresponding to $|\Delta D_q|_{\max}$, RMSE = root mean square error.

on the MOM-based box counting analyses was most pronounced for homogenous fields and at low resolutions. This is because the greatest differences between $\rho_{j_i}(\ell) = \mu_{\min}$ (i.e. the minimum mass fraction in the untransformed state) and $\rho_{j_i}(\ell) = 0$ (i.e. minimum mass fraction in the transformed state) occur under such conditions. Normalization also explains why the D_q values were numerically best estimated using the largest box sizes. The presence of a zero value instead of $\mu_{\min}$ has a large impact on the computation of $\chi(q, \delta)$ when $\delta \to \ell$, but only a minimal impact when $\delta \to 1$. It should be noted that this behavior is completely opposite to what one might have expected based on previous experience with the box counting algorithm applied to geometrical monofractals.

We used the IEEE standard double precision ($\sim$ 16 digits) for floating point calculations. Chen *et al.*[18] have pointed to computational overflow (associated with extremely large values of $\chi(q, \delta)$ when $\delta \to \ell$ and $q \ll 0$) as a possible source of error in MOM box counting analyses. In our study, the largest values of $\chi(q, \delta)$ occurred when analyzing the most heterogeneous fields (i.e. $\lambda \gg 1$). Since these fields always resulted in the smallest estimation errors (see Tables 1 and 2), computational overflow can be discounted as a contributing factor.

Another potentially important issue relates to the choice of the box counting scale factor used in the MOM analyses. Errors in the estimation of D_q can be expected to arise from any grid mismatch caused by the use of different scale factors for the random geometrical multifractal field and the superimposed box counting grid. In this study, we always matched these two scale factors. Further research is needed to investigate the magnitude of the errors introduced by such grid mismatches. For digital images and experimentally-determined data there may be no inherent grid scale factor beyond the value suggested by the overall dimensions of the field. In this case, the box counting scale

factor is almost universally taken as two. Does this somewhat arbitrary choice influence the estimation of D_q?

An additional consideration is how to extend this research to natural grayscale fields which, regardless of the grid size, are often limited to 256 pixel values. In contrast, geometrical multifractal fields always contain as many mass fractions (or grayscales) as there are grid cells. How does this difference influence the estimation of D_q versus q? Natural systems can be expected to deviate from ideal multifractal scaling. Given the nonlinearities apparent in the log-transformed partition functions for known geometrical multifractal fields, what statistical criteria should be used to separate out multifractal scaling from monofractal or Euclidean behavior? It is clear that an R^2 value close to unity for $\ln[\chi(q,\delta)]$ regressed on $\ln[\delta]$ is not a reliable indicator of linearity and more work should be devoted to this issue.

Finally, a logical future research direction might be to explore use of the moment-based box counting method to inversely estimate the b and p parameters for natural grayscale fields and images. Such a parsimonious parameterization of multifractality might find applications in a variety of different disciplines.

4. CONCLUSIONS

Random geometrical multifractal grayscale fields provide a theoretical foundation for evaluating the accuracy of generalized dimensions estimated by the moment-based box counting method. Comparison of estimated D_q values with their analytical counterparts indicated the moment-based box counting method systematically overestimates the generalized dimensions when q is negative. Maximum absolute deviations generally increased as q became more negative. As a result, RMSE values usually increased as the range in q increased. For any selected q range, the RMSE values increased with increasing grid size and decreasing field heterogeneity. Based on all 23 multifractal fields investigated, the geometric mean RMSE was 6.50×10^{-3} for $\Delta q = 20$.

The source of the deviations in the estimated generalized dimensions when $q \ll 0$ was attributed to the transformation used to convert the mass fractions into normalized grayscale fields ranging between zero and one. The presence of cells containing zero generated nonlinearities in the log-transformed partition functions as the box size was gradually reduced to the dimensions of the underlying grid. As a result, linear regression analyses yielded biased estimates of the slope, $\tau(q)$. An alternative procedure for estimating $\tau(q)$ was proposed based on the numerical first derivatives of the log-transformed partition function as the box size approaches the field size. This approach resulted in an order of magnitude increase in accuracy in the estimation of D_q versus q when compared to the linear regression analyses.

REFERENCES

1. Q. Cheng, Multifractality and spatial statistics, *Comput. Geosci.* **25** (1999) 949–961.
2. T. Parrinello and R. A. Vaughan, Multifractal analysis and feature extraction in satellite imagery, *Int. J. Remote Sens.* **23**(9) (2002) 1799–1825.
3. T. Stojić, I. Reljin and B. Reljin, Adaptation of multifractal analysis to segmentation of microcalcifications in digital mammograms, *Physica A* **367** (2006) 494–508.
4. T. Takahashi, T. Murata, K. Narita, T. Hamada, H. Kosaka, M. Omori, K. Takahashi, H. Kimura, H. Yoshida and Y. Wada, Multifractal analysis of deep white matter microstructural changes on MRI in relation to early-stage atherosclerosis, *NeuroImage* **32** (2006) 1158–1166.
5. A. N. Kravchenko, C. W. Boast and D. G. Bullock, Multifractal analysis of soil spatial variability, *Agron. J.* **91** (1999) 1033–1041.
6. T. B. Zeleke and B. C. Si, Scaling relationships between saturated hydraulic conductivity and soil physical properties, *Soil Sci. Soc. Am. J.* **69** (2005) 1691–1702.
7. A. N. Kravchenko, D. G. Bullock and C. W. Boast, Joint multifractal analyses of crop yield and terrain slope, *Agron. J.* **92** (2000) 1279–1290.
8. T. B. Zeleke and B. C. Si, Scaling properties of topographic indices and crop yield: multifractal and joint multifractal approaches, *Agron. J.* **96** (2004) 1082–1090.
9. S. Stach, S. Roskosz, J. Cwajna and J. Cybo, Multifractal detection of overlaps based on a stereometric analysis of a fracture surface: application to fractures of sintered carbides, *Mater. Charact.* **56** (2006) 429–435.
10. J. R. Mureika, C. C. Dyer and G. C. Cupchik, Multifractal structure in nonrepresentational art, *Phys. Rev. E* **72** (2005), 046–101.
11. C. J. G. Evertsz and B. B. Mandelbrot, in *Chaos and Fractals: New Frontiers of Science*, eds. H. O. Peitgen, H. Jürgens and D. Saupe (Springer-Verlag, New York, NY, 1992) pp. 921–953.

12. H. G. E. Hentschel and I. I. Procaccia, The infinite number of generalized dimensions of fractals and strange attractors, *Physica D* **8**(3) (1983) 435–444.
13. J. Feder, *Fractals* (Plenum Press, New York, NY, 1988) pp. 66–103.
14. D. Veneziano, G. E. Moglen and R. L. Bras, Multifractal analysis: pitfalls of standard procedures and alternatives, *Phys. Rev. E* **52**(2) (1995) 1387–1398.
15. Q. Cheng, The gliding box method for multifractal modeling, *Comput. Geosci.* **25** (1999) 1073–1079.
16. J. Grau, J. V. Méndez, A. M. Tarquis, M. C. Díaz and A. Saa, Comparison of gliding box and box-counting methods in soil image analysis, *Geoderma* **134** (2006) 349–359.
17. A. B. Chhabra and K. R. Sreenivasan, Negative dimensions: theory, computation, and experiment, *Phys. Rev. A* **43**(2) (1991) 1114–1117.
18. H. Chen, X. Sun, H. Chen, Z. Wu and B. Wang, Some problems in multifractal spectrum computation using a statistical method, *New J. Phys.* **6** (2003) doi:10.1088/1367–2630/6/1/084.
19. A. Turiel, C. J. Pérez–Vincente and J. Grazzini, Numerical methods for the estimation of multifractal singularity spectra on sampled data: a comparative study, *J. Comp. Phys.* **216** (2006) 362–390.
20. S. Campagna and G. Turchetti, Dimension spectra of fractal measures from uniform partitions and correlation integrals, *J. Phys. A* **32** (1999) 7989–7999.
21. H. Wendt, P. Abry and S. Jaffard, Bootstrap for empirical multifractal analysis, *IEE Signal Process. Mag.* **24**(7) (2007) 38–48.
22. L. V. Meisel, L. V. M. Johnson and P. J. Cote, Box-counting multifractal analysis, *Phys. Rev. A* **45**(10) (1992) 6989–6996.
23. N. Bird, M. Cruz Díaz, A. Saa and A. M. Tarquis, Fractal and multifractal analysis of pore-scale images of soil, *J. Hydrol.* **322** (2006) 211–219.
24. R. Santoro, N. M. Maraldi, S. Campagna and G. Turchetti, Uniform partitions and a dimensions spectrum for lacunar measures, *J. Phys. A* **35** (2002) 1871–1884.
25. T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. I. Procaccia and B. I. Shraiman, Fractal measures and their singularities: the characterization of strange sets, *Phys Rev. A* **33**(2) (1986) 1141–1151.
26. E. Perfect, R. W. Gentry, M. C. Sukop and J. E. Lawson, Multifractal Sierpinski carpets: theory and application to upscaling effective saturated hydraulic conductivity, *Geoderma* **134** (2006) 240–252
27. S. R. Koirala, E. Perfect, R. W. Gentry and J.-W. Kim, Effective saturated hydraulic conductivity of 2-dimensional random multifractal fields, *Water Resour. Res.* **44**(8) (2008) w08410.1–w08410.9.
28. G. P. Patil, Maximum likelihood estimation for generalized power series distributions and its application to the truncated binomial distribution, *Biometrika* **49** (1962) 227–237.
29. J.-F. Gouyet, *Physics and Fractal Structures* (Springer, New York, NY, 1996) pp. 1–39.
30. A. Saucier, Scaling properties of disordered multifractals, *Physica A* **226** (1996) 34–63.
31. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd edn. (Prentice Hall, Upper Saddle River, NJ, 2006) pp. 107–120.
32. D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics*, 2nd edn. (Cambridge University Press, New York, NY, 1997) pp. 100–131.
33. R. L. Burden and J. D. Faires, *Numerical Analysis*, 5th edn. (PWS Publishing Co. Boston, MA, 1993) pp. 156–167.