Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chaos, Solitons and Fractals 28 (2006) 1337–1346
CHAOS SOLITONS & FRACTALS
www.elsevier.com/locate/chaos

Fractal methods and the problem of estimating scaling exponents: A new approach based on upper and lower linear bounds

Antoine Saucier *, François Soumis [1]

École Polytechnique de Montréal, C.P. 6079, Station centre-ville, Montréal (Que.), Canada H3C-3A7

Accepted 8 August 2005

Abstract

The characterization of irregular objects with fractal methods often leads to the estimation of the slope of a function which is plotted versus a scale parameter. The slope is usually obtained with a linear regression. The problem is that the fit is usually not acceptable from the statistical standpoint. We propose a new approach in which we use two straight lines to bound the data from above and from below. We call these lines the upper and lower linear bounds. We propose to define these bounds as the solution of an optimization problem. We discuss the solution of this problem and we give an algorithm to obtain its solution. We use the difference between the upper and lower linear bounds to define a measure of the degree of linearity in the scaling range. We illustrate our method by analyzing the fluctuations of the variogram in a microresistivity well log from an oil reservoir in the North Sea.

1. Introduction

The characterization of irregular objects with fractal methods often leads to the estimation of the slope of a function plotted as a function of a scale parameter δ. Most fractal methods assume that a function F, e.g. a variogram or a partition function, satisfies a power law behavior, i.e. $F(\delta) \propto \delta^a$, where a is an exponent that can have different interpretations, e.g. a fractal dimension. In such cases, a is usually estimated with a linear regression of $\log(F(\delta))$ versus $\log(\delta)$ using the model $\log(F(\delta)) = a\log(\delta) + b$.

The function F is typically calculated from signals or images. It is well known that the function $F(\delta)$ typically fails to be perfectly linear on a log–log plot [2]. In many cases, this departure from linearity is real, i.e. it cannot be explained by measurement errors [5,6]. In other words, the linear model is usually not acceptable from a statistical standpoint. The practical implementation of fractal methods has always been confronted with this difficulty.

In spite of its roughness, the linear model can be useful. Indeed, if the departure from linearity is not too large, then the linear regression slope a is an index which measures roughly the rate of variation of $F(\delta)$ in a given range of scales

* Corresponding author. Tel.: +1 514 340 4711x4516; fax: +1 514 340 4463.
E-mail addresses: Antoine.Saucier@polymtl.ca (A. Saucier), Francois.Soumis@gerad.ca (F. Soumis).
[1] Tel.: +1 514 340 4711x6044; fax: +1 514 340 5665.

0960-0779/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chaos.2005.08.045

$\delta_{\min} \leqslant \delta \leqslant \delta_{\max}$, often called the scaling range. If the index a is used to analyze images for which the span[2] of a is much larger than the uncertainty on a, then a can be used to distinguish and classify the images. This shows that an index that characterizes roughly the rate of variation of $F(\delta)$ in the scaling range can be useful. For example, fractal methods with approximate linear fits have been used successfully to characterize DNA sequences [1], well logs [4] and signals in general [3].

In this paper, we examine the problem of estimating fractal exponents in cases for which the log–log linearity of $F(\delta)$ is questionable. The idea of measuring a slope by fitting a straight line to a curve with a statistically unacceptable fit is not satisfactory. Indeed, a statistically unacceptable fit indicates that the straight line is not the appropriate model to use. A conceptually clearer standpoint can be obtained if one avoids the linear regression. Instead of fitting a straight line to the data, which is typically doomed to fail, we can search for two straight lines which are upper and lower linear bounds for the data. These two lines should be defined in such a way that they are unique and as close as possible to the data, in some sense. In other words, we use two straight lines to sandwich the data as tightly as possible. In the following, the two straight lines that bound the data from above and from below will be called the *upper linear bound* and the *lower linear bound* respectively.

This approach has several advantages. Firstly, statistically unacceptable fits are avoided. Secondly, the exponents (i.e. slopes) have a clear geometrical meaning: they are the slopes of the two straight lines that sandwich the data from above and from below. Thirdly, if the slopes a_{upper} and a_{lower} of the upper and lower linear bounds differ significantly, then our approach provides a more precise description of $F(\delta)$. Fourthly, if the data actually forms a straight line, then $a_{\text{upper}} = a_{\text{lower}} = a$, where a is the slope obtained by linear regression. Hence our approach reduces to the usual approach in this case. Finally, the difference between the upper and lower linear bounds provides a measure of the degree of linearity of the curve. Indeed, if the curve is straight, then the upper and lower linear bounds merge. However, if the curve is not straight, then the upper and lower linear bounds differ.

2. Upper and lower linear bounds

2.1. Definition

The data is defined by the coordinates of each point:

$$(x_i, y_i), \quad i = 1, 2, \ldots, N. \tag{1}$$

We assume that the x_is are equidistant and arranged in increasing order, i.e. $x_1 < x_2 < \cdots < x_N$. We define the upper linear bound to be the straight line of equation $y(x) = ax + b$ which is the solution of the following constrained optimization problem:

$$\begin{cases} \min\limits_{(a,b)} \sum\limits_{i=1}^{N} (ax_i + b - y_i)^2, \\ ax_i + b \geqslant y_i, \quad i = 1, 2, \ldots, N. \end{cases} \tag{2}$$

Similarly, the lower linear bound is the straight line of equation $y(x) = ax + b$ which satisfies

$$\begin{cases} \min\limits_{(a,b)} \sum\limits_{i=1}^{N} (ax_i + b - y_i)^2, \\ ax_i + b \leqslant y_i, \quad i = 1, 2, \ldots, N. \end{cases} \tag{3}$$

The problems (3) and (2) are optimization problems with linear inequality constraints. In the following, we focus on the upper linear bound problem (2). The solution of the problem (3) can be obtained from the solution method of problem (2). Indeed, it can be shown easily that using the substitutions $y_i \to -y_i$, $i = 1,2,\ldots,N$, $a \to -a$ and $b \to -b$ on (2) leads to (3). This implies that the solution of (3) is obtained by solving (2) after the substitutions $y_i \to -y_i$, $i = 1,2,\ldots,N$, and then by changing the signs of the resulting a and b.

2.2. Measuring the degree of linearity

The upper and lower linear bounds, which are defined by $y_i = a_{\text{upper}} x_i + b_{\text{upper}}$ and $y_i = a_{\text{lower}} x_i + b_{\text{lower}}$ respectively, can be used to measure the degree of linearity. The data, together with the upper and lower linear bounds, form a

[2] The span is the difference between the maximum and minimum values of a for the various images.

sandwich that has a varying width. The widths of this sandwich at $x = x_1$ and $x = x_N$ are $W(x_1) \equiv a_{\text{upper}} x_1 + b_{\text{upper}} - (a_{\text{lower}} x_1 + b_{\text{lower}})$ and $W(x_N) \equiv a_{\text{upper}} x_N + b_{\text{upper}} - (a_{\text{lower}} x_N + b_{\text{lower}})$ respectively. A possible measure of the distance between the upper and lower linear bounds is $b_* \equiv \max(W(x_1), W(x_N))$. b_* could be a good measure of the degree of linearity because $b_* = 0$ if and only if the curve is perfectly straight.

3. Solution of the optimization problem

It is intuitively easy to accept that the solution of the optimization problem (2) is a line that touches at least one of the N points. This can be shown as follows. If we assume that the solution of (2) is a straight line $y = ax + b$ that does *not* touch any of the points (x_i, y_i), then it is always possible to decrease the value of b so that all the inequalities (2)b are respected. Indeed, the substitution $b \to b - \alpha$, where $\alpha \in [0, \min_i(ax_i + b - y_i)]$, yields a solution (a, b) that satisfies the inequalities (2)b, but which has a lower value of the sum of squares $\sum_{i=1}^{N}(ax_i + b - y_i)^2$. Hence the initial solution is not the global minimum of the sum of squares, which is a contradiction. Hence the solution touches at least one point.

This result leads us to consider the set of all straight lines that have at least one contact point with the data while being above all the other points. In this set, the contact points play the role of *pivots*, i.e. a line can be rotated around a pivot as long as it remains above all the other points.

The pivots are represented in Fig. 1 as empty circles. The lines that intersect the point (x_1, y_1) and which are above all the other points can be obtained as follows. Starting from a vertical line passing through (x_1, y_1), we rotate the line clockwise around the pivot (x_1, y_1) until it touches another point, which leads to the straight line denoted by L_1. The next pivot happens here to be the third point (x_3, y_3). L_1 can be rotated clockwise around (x_3, y_3) until it reaches the line denoted by L_2. This contouring process can be continued until the last pivot is reached, which corresponds to the line L_4 in Fig. 1. This contouring process allows to scan all the lines which have at least one contact point with the data while being above it. The solution of the optimization problem (2) is one of these lines.

If the solution touches the kth point of coordinates (x_k, y_k), then its parameters a and b must be related by $ax_k + b = y_k$, which implies that $b = y_k - ax_k$. Hence the parameter b can be eliminated if we know that the solution touches the kth point. Substituting b by $y_k - ax_k$ into (2) leads to the following equivalent optimization problem:

$$\begin{cases} \min\limits_{(a,k)} \sum\limits_{i=1}^{N} (a(x_i - x_k) - (y_i - y_k))^2, \\ a(x_i - x_k) - (y_i - y_k) \geqslant 0, \quad i = 1, 2, \ldots, N, \end{cases} \tag{4}$$

where we have included k into the minimization variables because the point (x_k, y_k) leading to the best solution is not known *a priori*. The formulation (4) shows that only the slope parameter a plays a role in this problem, and reveals explicitly the role of the pivots. It is emphasized that only the pivots satisfy the constraints (4)b.

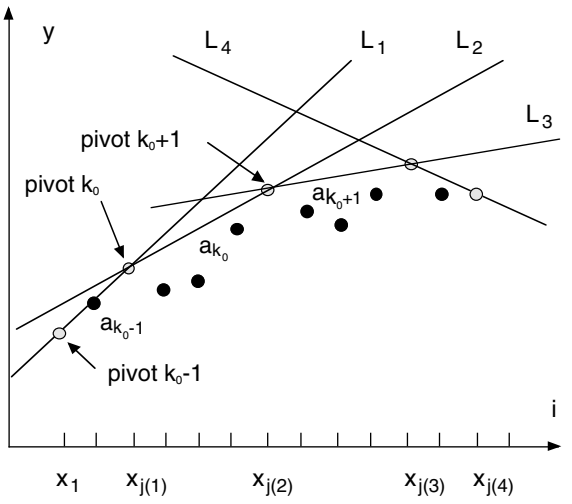

Fig. 1. The circles, black or empty, represent the data points. The empty circles are the pivots, which form a subset of the data points. a_k denotes the slope of the straight line L_k.

For the kth pivot, the inequality constraints imply that the slope parameter a must be in a range defined by the slopes a_{k-1} and a_k of two lines L_{k-1} and L_k respectively, i.e. $a \in [a_k, a_{k-1}]$. In this range, the regression sum of squares takes the form

$$U_k(a) \equiv \sum_{i=1}^{N} (a(x_i - x_{j(k)}) - (y_i - y_{j(k)}))^2, \tag{5}$$

where $j(k)$ is the coordinate index of the kth pivot. The left-to-right contouring process allows to scan all the values of a, from large to small values. This leads us to define a function $U(a)$ by

$$U(a) = \begin{cases} U_1(a) & \text{if } a \geqslant a_1, \\ U_2(a) & \text{if } a_2 \leqslant a \leqslant a_1, \\ U_3(a) & \text{if } a_3 \leqslant a \leqslant a_2, \\ \dots & \\ U_k(a) & \text{if } a_{k+1} \leqslant a \leqslant a_k, \\ \dots & \\ U_N(a) & \text{if } a \leqslant a_{k_{\max}}, \end{cases} \tag{6}$$

where $k_{\max}$ is the total number of pivots. The slope of the solution is the global minimum of U. By construction, U is *continuous* across the intervals $a_{k+1} < a < a_k$ since $U_k(a_k) = U_{k+1}(a_k)$ for all ks. More importantly, U is also *strictly convex*. Indeed, using (5), we find that

$$\frac{\partial^2}{\partial a^2} U_k(a) = 2 \sum_{i=1}^{N} (x_i - x_{j(k)})^2 > 0 \tag{7}$$

for $a \in]a_k, a_{k-1}[$ and for all ks. The strict convexity of U implies that the minimum of U is *unique*, and therefore the solution of the optimization problem (2) is also unique. Finally, it is stressed that the derivative $U'(a)$ is *piecewise continuous*. The discontinuities of $U'(a)$ occur if one changes the pivot, i.e. $U'_k(a_k)$ and $U'_{k+1}(a_k)$ are typically different.

4. Algorithms

The solution is a line that touches at least one of the N points. In general, it might touch several points simultaneously. The first step of the solution is to locate the optimal pivot, i.e. the pivot around which the global minimum of U is reached. One way to proceed is to start by locating all the pivots.

4.1. Locating the pivots

4.1.1. First pivot ($k = 1$)

Initially, we use the point (x_1, y_1) as the first pivot, which has the coordinate index $j(1) = 1$.

4.1.2. Finding the second pivot ($k = 2$)

A vertical line which passes through the first pivot (x_1, y_1) (Fig. 1) is pivoted clockwise around (x_1, y_1) until it touches one of the other points. The resulting line is denoted by L_1 in Fig. 1. It might happen that several points are touched simultaneously if they are aligned. For instance, in Fig. 1 the line L_1 touches two points simultaneously, i.e. the second and third points. The rightmost of these two points will be the second pivot, which has the coordinate index $j(2) = 3$. In general, $j(2)$ is found with

$$j(2) = \arg\max_{k>1} \frac{y_k - y_1}{x_k - x_1}. \tag{8}$$

If (8) gives several solutions, then one chooses the one having the largest x_k, i.e. the rightmost solution. We compute the value of the regression sum of squares $U(a_1)$ for the line L_1 using

$$\begin{cases} U(a_1) = \sum_{i=1}^{N} (a_1 x_i + b_1 - y_i)^2, \\ \text{where} \\ a_1 = (y_{j(2)} - y_1)/(x_{j(2)} - x_1), \\ b_1 = y_1 - a_1 x_1. \end{cases} \tag{9}$$

4.1.3. Finding the (k + 1)th pivot

Having found the kth pivot, which has the coordinate index $j(k)$, we can search for the $(k+1)$th pivot. As previously, $j(k+1)$ is found with

$$j(k+1) = \arg\max_{k>j(k)} \frac{y_k - y_{j(k)}}{x_k - x_{j(k)}}. \tag{10}$$

In Fig. 1, the third pivot is obtained by rotating L_1 clockwise around the second pivot until it touches a point, which happens to be the 7th point in this case, i.e. $j(3) = 7$. The resulting line is L_2. Next, we compute the regression sum of squares $U(a_k)$ of L_k, which is given by

$$\begin{cases} U(a_k) = \sum_{i=1}^{N} (a_k\, x_i + b_k - y_i)^2, \\ \text{where} \\ a_k = (y_{j(k+1)} - y_{j(k)})/(x_{j(k+1)} - x_{j(k)}), \\ b_k = y_{j(k)} - a_k x_{j(k)}. \end{cases} \tag{11}$$

This process can be continued until the last pivot is obtained, which is always the last point, i.e. $j(k_{\max}) = N$.

4.2. Locating the optimal pivot

The notations used in the following argument are illustrated in Fig. 1. The function U is convex and consequently the sampled values $U(a_k)$ also form a convex series. Let $k_* = \arg\min_k U(a_k)$. The optimal line is close to L_{k_*}. The convexity of U implies that $(U'(a))' \geqslant 0$, i.e. $U'(a)$ increases if a increases. If follows that the values

$$U'_k(a_{k-1}), \quad k = 2, 3, \ldots, k_{\max} \tag{12}$$

form an *decreasing series* as k increases, because a_k decreases as k increases. The $U'_k(a_{k-1})$s can be obtained by differentiating (5), which yields

$$U'_k(a) = 2\sum_{i=1}^{N} (a(x_i - x_{j(k)}) - (y_i - y_{j(k)}))(x_i - x_{j(k)}). \tag{13}$$

Two main cases must be distinguished: either the series $U'_k(a_{k-1})$ changes sign in the range $2 \leqslant k \leqslant k_{\max}$, or it does not.

4.2.1. Finding the optimal parameters if the derivative changes sign

If the series $U'_k(a_{k-1})$ changes sign, then there exists $k_0 \in [2, k_{\max} - 1]$ such that $U'_k(a_{k-1}) \geqslant 0$ for $k \leqslant k_0$ and $U'_k(a_{k-1}) < 0$ for $k > k_0$. k_0 is the optimal pivot. It follows that the minimum is located in the range $a \in [a_{k_0}, a_{k_0-1}]$. If $U'_{k_0}(a_{k_0-1}) = 0$, then the minimum occurs for $a = a_{k_0-1}$. If $U'_{k_0}(a_{k_0-1}) > 0$, then the minimum occurs for $a \in [a_{k_0}, a_{k_0-1}[$. We compute the linear regression slope for all straight lines passing through the pivot k_0, without taking into account the inequality constraints. The zero derivative condition $U'_{k_0}(\hat{a}_0) = 0$, used with (13), leads to the estimate

$$\hat{a}_0 = \frac{\sum_{i=1}^{N} (x_i - x_{j(k_0)})(y_i - y_{j(k_0)})}{\sum_{i=1}^{N} (x_i - x_{j(k_0)})^2}. \tag{14}$$

The estimate (14) may not satisfy the inequality constraints. Indeed, the function U may not have a critical point in the range $a \in [a_{k_0}, a_{k_0-1}[$, in which case U is strictly decreasing in this range. Hence two cases must be considered:

(i) if $a \in [a_{k_0}, a_{k_0-1}[$, then U is minimum for $a = \hat{a} \equiv \hat{a}_0$;
(ii) if $a \notin [a_{k_0}, a_{k_0-1}[$, then U is minimum for $a = \hat{a} \equiv a_{k_0}$.

Once the estimate $\hat{a}$ has been determined, the estimate $\hat{b}$ of b is obtained with $\hat{b} = y_{j(k_0)} - \hat{a} x_{j(k_0)}$.

4.2.2. Finding the optimal parameters if the derivative does not change sign

If all the $U'_k(a_{k-1})$s are strictly negative, then the minimum occurs for $a = a_1$, i.e. for the first line L_1, and $k_0 = 2$. If all the $U'_k(a_{k-1})$s are strictly positive, then the minimum occurs for $a = a_{j(k_{\max})}$, i.e. for the last line $L_{k_{\max}}$, and $k_0 = k_{\max} - 1$. The final estimate $\hat{b}$ of b is again obtained with $\hat{b} = y_{j(k_0)} - \hat{a} x_{j(k_0)}$.

We have illustrated in Fig. 2 the upper and lower linear bounds obtained with the above algorithm on three different data sets.

4.3. Connection with the simplex algorithm

The algorithm that we presented happens to be a specialization of the simplex algorithm for convex problems. The problem (2) can be rewritten in standard form with equality constraints

$$ax_i + b = y_i + s_i, \quad i = 1, 2, \ldots, N, \tag{15}$$

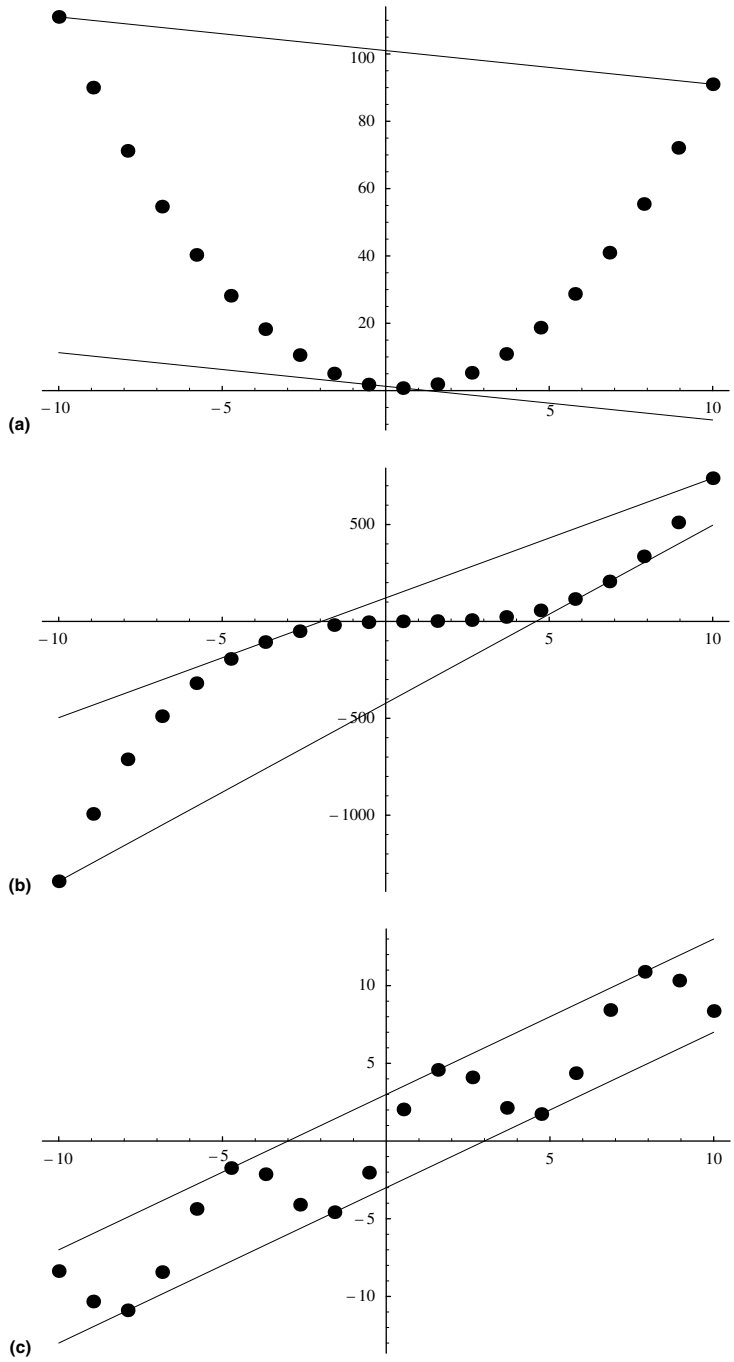

Fig. 2. Illustration of the upper and lower linear bounds for three different data sets.

where s_i is a non negative slack variable. A basic solution of this problem with N equality constraints and $N+2$ variables contains two non basic slack variables equal to zero. A feasible basic solution passes through two points (x_i, y_i) and satisfies the constraints (2)b. It is one of the L_k lines. A pivot of the simplex algorithm moving from a base to an other corresponds to pivoting the line L_k to reach L_{k+1}.

The main advantage of our algorithm is that it allows to solve efficiently our optimization problem without resorting to a general purpose optimization toolbox.

5. Application to the analysis of variograms derived from microresistivity well logs

The characterization of facies using well logs is a useful aid to the interpretation of the depositional setting of a geological area. It has been recently recognized that dipmeter microresistivity signals (DMS) bear not only information about the stratigraphic dip angle in geological formations but may also contain important information regarding the textural contents of these formations. Indeed, DMS have a high sampling rate (about one measurement every 2.5 mm) and a relatively high vertical resolution (1–2 cm). DMS is one of the only logs that can provide information about small-scale geological events.

This information can be extracted by a texture analysis of DMS in which texture parameters are calculated locally (i.e. in an interval) at different depths in the well [4]. Texture parameters characterize the variability of DMS in each interval. Plotted as a function of depth, these parameters form so-called texture logs that provide a compact representation of the DMS texture variations as a function of depth. Texture logs can be used in subsequent clustering analyses that will provide a well zonation based on the DMS.

In the following, we examine the variability of the DMS log of the well 30/9-6 from the Oseberg field in the North Sea, which is plotted in Fig. 4 (center). More precisely, we examine the variations of the variogram as a function of depth. For a signal defined by its values $S(i)$, $i = 1,2,\ldots,N$, the variogram is defined as usual by

$$V(k) = E((S(i+k) - S(i))^2), \tag{16}$$

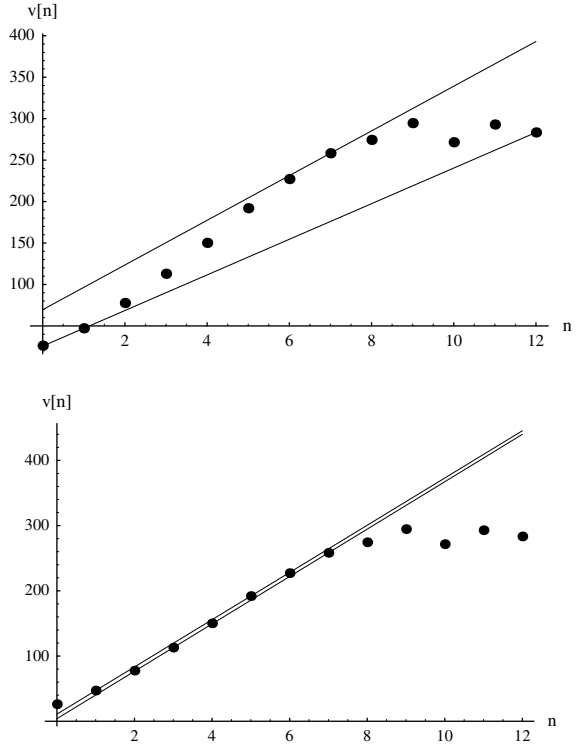

Fig. 3. $v(n) \equiv \sqrt{V(2^n)}$ versus n derived from the well log displayed in Fig. 4 (center), together with the upper and lower linear bounds. *Top*: the linear bounds are computed over the whole range. *Bottom*: the linear bounds are computed in the range $2 \leqslant n \leqslant 7$.

where $E(\cdots)$ denotes an expectation value. For a stationary signal containing N points, the estimator $\widehat{V}(k)$ of $V(k)$ is given by

$$\widehat{V}(k) = \frac{1}{N-k}\sum_{i=1}^{N-k}(S(i+k)-S(i))^2. \tag{17}$$

We first computed the variogram for the whole microresistivity well log at the scales $k = \delta_n \equiv 2^n$, $n = 0,1,\ldots,M$ with $M = 12$, and we plotted in Fig. 3 the values of $v(n) \equiv \sqrt{V(2^n)}$ as a function of n. It is seen that the points $(n, v(n))$ are not aligned particularly well over the whole range. However, they are aligned fairly well in the range $n \in [2,7]$, as indicated by the upper and lower linear bounds obtained in this range. This fit corresponds to a logarithmic behavior of the form $v(n) = a\log(\delta_n) + b$, which happens in this case to give a much better fit than the power law model $v(n) \propto \delta_n^H$.

To examine the local variability of the DMS log, we used a sliding window processing to compute $v(n)$ in consecutive windows of the form $[1 + i\delta L_w, 1 + i\delta L_w + L_w - 1]$, $i = 0,1,2,\ldots$, where $L_w = 512$ points is the window size and $\delta L_w = L_w/8 = 64$ points is the step size. For each window, $v(n)$ was computed in the range $[2,7]$ and the parameters (a,b) of the upper and lower linear bounds were obtained. We then plotted the resulting parameters as a function of the window center in Fig. 4.

We observe in Fig. 4 (top) that the slope parameter a is quite consistent, i.e. the values a_{upper} and a_{lower} of the upper and lower linear bounds respectively are very close to each other. More precisely, the difference $|a_{upper} - a_{lower}|$ is much

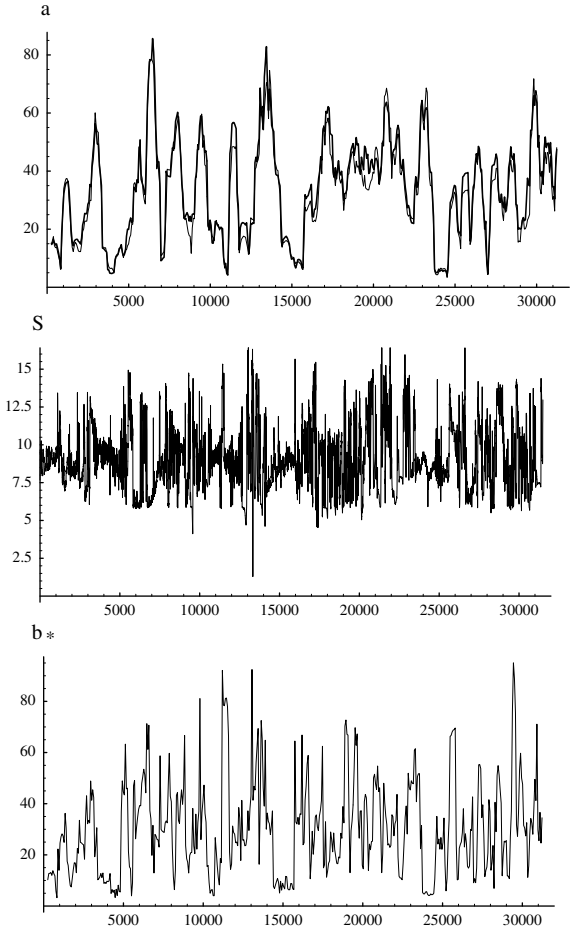

Fig. 4. Parameters a_{upper}, a_{lower} and b_* derived from the upper and lower linear bounds of the variogram (i.e. $v(n)$ versus n), plotted as a function of the window center position (in unit of points). *Top*: slope parameters a_{upper} (bold curve) and a_{lower} (thin curve). *Center*: DMS log. *Bottom*: parameter b_*, representing the width of the sandwich composed of the upper and lower linear bounds. b_* is a measure of the degree of linearity: $b_* = 0$ if and only if the curve is perfectly straight.

smaller than the amplitude of the variations of a_{upper} over the whole DMS log. This shows that the slope a is a potentially useful characterization parameter for DMS logs.

The parameter b_*, which was previously defined in Section 2.2, is a measure of the degree of linearity of the curve $v(n)$ versus n. We observe in Fig. 4 (bottom) that b_* fluctuates in the range $[2, 90]$, which indicates that the degree of linearity of $v(n)$ is quite variable. We also observe that b_* and the slope parameter a appear to be correlated. In particular, it can be seen that low values of b_* correspond to low values of a.

6. Conclusions

The characterization of irregular objects with fractal methods often leads to the estimation of the slope of a function which is plotted versus a scale parameter. The slope is typically obtained with a linear regression. The problem is that the fit is usually not acceptable from the statistical standpoint. We proposed a new approach in which we use two straight lines to bound the data from above and from below. We call these lines the upper and lower linear bounds. We defined these bounds as the solution of an optimization problem. We discussed the solution of this problem and we gave an algorithm to obtain its solution. We used the difference between the upper and lower linear bounds to define a measure of the degree of linearity of the scaling range.

Our definition of the upper and lower linear bounds is based on least squares with inequality constraints. Other definitions could also be considered. For instance, the distance between the data points and the straight line could be measured in a direction perpendicular to the line. In our implementation, the distances are measured in the vertical direction, which is a simple yet special choice.

We have shown that the upper and lower linear bounds are typically defined by only one or two points, the so called pivots. It follows that the upper and lower linear bounds can be quite sensitive to the modification of a single point, especially if this point happens to be a pivot. Let us consider for instance the points (x_k, y_k) defined by

$$\begin{cases} x_k = k, & k = 0, 1, \ldots, 10^6, \\ y_0 = 10^6, & \\ y_k = 0, & 1 \leqslant k \leqslant 10^6. \end{cases} \tag{18}$$

In this case, it is easy to see that $a_- = 0$ and $a_+ = -1$. The point $(0, 10^6)$, which behaves like an outlier, is responsible for a large perturbation of a_+ because $(0, 10^6)$ happens to be a pivot. a_+ and a_- can therefore be sensitive to outliers. This lack of robustness would be problematic if the function $F(\delta)$ actually contained outliers. However, $F(\delta)$ is often continuous and smooth, by construction. Consequently, the occurrence of extreme outliers, as in the above example, may not be a problem. Nevertheless, this sensitivity issue would deserve a closer study.

The definitions of a fractal exponent are typically given in terms of limits (inf or sup) as $\delta \to 0$. For instance, the upper box dimension of a set S is given by $d_{\text{B}}(S) = -\limsup_{\delta \to 0} \log(N_{\text{B}}(\delta))/\log(\delta)$, where $N_{\text{B}}(\delta)$ is the number of boxes of diameter δ needed to cover S. In practical data analysis, e.g. in the analysis of DMS logs, the limit $\delta \to 0$ is not necessarily important. If our goal is to characterize the signal variability, then what matters most is the determination of a range of scales *as wide as possible* in which the function $F(\delta)$ is approximately linear. This range does not necessarily have to extend to the smallest scales. From this standpoint, our approach is legitimate.

Acknowledgements

A. Saucier and F. Soumis acknowledge financial support from individual NSERC grants. A. Saucier would like to thank Norsk Hydro for providing the dipmeter data, and Charles Audet for stimulating discussions.

References

[1] Audit B, Vaillant C, Arnéodo A. Wavelet analysis of DNA bending profiles reveals structural constraints on the evolution of genomic sequences. J Biol Phys 2004;30:33–81.

[2] Avnir D, Lidar D, Malcai O. On the abundance of fractals. In: Novak MM, Dewey TG. editors. Fractals in the natural and applied sciences. Singapore; 1997.

[3] Lévy Véhel J, Legrand P. Signal and image processing with FRACLAB. In: Novak MM, editor. Thinking in patterns, fractals and related phenomena in nature. Singapore, London, New Jersey, Hong Kong: World Scientific Publishing Co.; 2004, ISBN 981-02-3593-3. p. 321–2.

[4] Saucier A, Huseby OK, Muller J. Electrical texture characterization of dipmeter microresistivity signals using multifractal analysis. J Geophys Res 1997;102:10327–37.
[5] Saucier A, Muller J. Multifractal approach to textural analysis. In: Novak MM, editor. Fractals and beyond complexities in the sciences. Singapore, London, New Jersey, Hong Kong: World Scientific Publishing Co; 1998, ISBN 981-02-3593-3. p. 161–71.
[6] Saucier A, Muller J. A generalization of multifractal analysis based on polynomial expansions of the generating function. In: Dekking M, Lévy Véhel J, Lutton E, Tricot C, editors. Fractals: theory and applications in engineering. London: Springer-Verlag; 1999. p. 81–91.