

Statistical Analysis of Time Series with Scaling Indices

HARALD ATMANSPACHER^{a,b,*}, WERNER EHM^a, HERBERT SCHEINGRABER^b
and GERDA WIEDENMANN^{b,c}

^a*Institut für Grenzgebiete der Psychologie, Wilhelmstr. 3a, D-79098 Freiburg;*

^b*Max-Planck-Institut für extraterrestrische Physik, Giessenbachstr., D-85740 Garching;*

^c*Nabios GmbH, Donnersberger Str. 41, D-80634 München*

(Received 19 May 2000)

Statistical techniques based on scaling indices are applied to detect and investigate patterns in empirically given time series. The key idea is to use the distribution of scaling indices obtained from a delay representation of the empirical time series to distinguish between random and non-random components. Statistical tests for this purpose are designed and applied to specific examples. It is shown that a selection of subseries by scaling indices can significantly enhance the signal-to-noise ratio as compared to that of the total time series.

Keywords: Time series analysis; Pattern detection; Scaling indices

PACS number: 05.45.Tp

1. INTRODUCTION

If the behavior of a system evolving in time shows features that are “irregular” in some sense, the observed data series invites statistical modeling as a stochastic or random process. In such cases it is often useful and possible to conceptually distinguish between some *null mode* of behavior on one side and deviations from that null mode on the other. The notion of a null mode refers to some knowledge about the basic physical (null)

characteristics of the system, which can be understood in terms of random or deterministic mechanisms. If there is no such knowledge, the situation is more difficult insofar as one has to start with assumptions as to model classes that may be appropriate to reproduce the observed data.

There are many approaches that have been developed in this spirit for many different purposes. Their success generally depends on both the nature of the deviations and that of the null mode. Deviations of a pre-conceived sort, such as a shift

*Corresponding author.

or an abrupt change in the mean or trend of the process, can be detected efficiently if one uses a correspondingly tailored procedure, and if the actual deviation is of the anticipated sort. On the other hand, if the goal is to detect *any* deviation, or deviations one has only a vague notion of, then less specific procedures of a wider scope may be more appropriate, even at the risk of losing precision with respect to the distinction between different types of specific deviations.

In many practical cases of pattern detection and pattern recognition, the basic problem is to distinguish a (more or less) faint regular pattern, the deviation, in front of a random background, the null mode. In this framework, the most ambitious situation for modeling purposes is met if neither the type of regularity nor the type of randomness are known. This is the kind of situation for which a data-analytic method called *scaling index analysis* is particularly helpful. It aims at detecting non-random deviations from a random null mode and is applicable if the random process is (approximately) stationary in the null mode. Non-random contributions can be due to some (weak) regular signal or pattern superimposed on or replacing the null process, or due to increased correlations between successive measurements reducing the local variability of the time series. Using scaling indices, one can study this time variability geometrically by comparing the density gradients in suitably constructed point sets on different distance scales. In physical contexts, scaling indices or “crowding indices” [1] originate from the dimensional analysis of single fractals and multifractals ([2–4], for an overview see [5]). For a mathematically oriented perspective see [6].

Scaling index analysis has primarily been used as a data-analytic tool so far, *i.e.*, as a numerical algorithm extracting potentially interesting features embedded in a random background (see, *e.g.*, [7–11]). The crucial concept in a scaling index analysis is a so-called H_α histogram (see below) whose properties can be investigated and discriminated against the known or expected null mode histogram. This is no problem as long as the

deviations from the null behavior are marked enough to be easily visible. If this is not the case, it may be difficult to distinguish systematic but small deviations from deviations due to random fluctuations produced by the null process. Then a statistical test with an appropriate discriminant statistics is required, in which the error of the first kind – to falsely classify a time series as “deviant” although it is in null mode – is well controlled. If deviations occur only sporadically, it may also be interesting to single out those time intervals in which “something deviant seems to go on”.

The goal of this paper is to propose methods for (i) the construction of statistical tests based on scaling indices, and (ii) the identification of conspicuous parts of the observed time series, selected by particular ranges of scaling indices. In this way, we focus on a precise statistical formulation of scaling index analyses which was not addressed in earlier publications listed above. Section 2.1 indicates the basic definitions and procedures for a scaling index analysis, Section 2.2 describes the test construction, and Section 2.3 gives a detailed example. Section 3 addresses how scaling index analyses can be used to enhance the signal-to-noise-ratio by selecting and investigating those parts of a time series that are responsible for deviations from a random null process. Section 4 discusses and summarizes the main results.

2. STATISTICAL SCALING INDEX ANALYSIS

2.1. Basic Notations and Framework

Consider a (univariate) time series $\{x(t): t = 1, \dots, T\}$ of finite length T . This notation implies a discrete notion of a parameter time t , which will subsequently be utilized to indicate the temporal ordering of observations $x(t)$. Using the technique of delay coordinates [12] one can construct an associated d – dimensional time series,

$$z(t) = (x(t), \dots, x(t + d - 1)), \\ t = 1, \dots, T - d + 1.$$

For every t and pre-defined radii $r_1, r_2 (0 < r_1 < r_2)$ let a *scaling index* $\alpha(t)$ be defined as

$$\alpha(t) = \frac{\log N_2(z(t)) - \log N_1(z(t))}{\log r_2 - \log r_1}, \quad (1)$$

where $N_{1,2}(z_{\text{ref}}) = \#\{t: |z(t) - z_{\text{ref}}| \leq r_{1,2}\}$ denotes the number of all points $z(t)$, $t = 1, \dots, T - d + 1$, whose Euclidean distance to the reference point z_{ref} does not exceed $r_{1,2}$.

Any scaling index analysis is based on the distribution of the scaling indices. This distribution can be represented by a histogram, the so-called H_α histogram. Its integrated version, the (*cumulative*) *empirical distribution function (edf) of the scaling indices (si)*

$$H(\alpha) = (T - d + 1)^{-1} \#\{t: \alpha(t) \leq \alpha\} \quad (\alpha \geq 0), \quad (2)$$

is more appropriate for our purposes, because of its improved statistical stability. $H(\alpha)$ will be referred to as the *si-edf associated with* $\{x(t)\}$.

The basic idea is to compare this si-edf $H(\alpha)$ with an appropriate reference distribution function H_{ref} characterizing the null process. In cases in which a continuous time process can be characterized by a multifractal measure, the limiting H_{ref} for $T \rightarrow \infty$, then $r_{1,2} \rightarrow 0$, increases in the range of α corresponding to the (set of) dimension(s) on which the measure is concentrated. (For single fractals, the increase of H_{ref} reduces to a Heaviside step function corresponding to a δ function type histogram.) If the null process is random, the range of α , over which H_{ref} increases, keeps growing toward larger α with increasing embedding dimension d . By contrast, for regular or chaotic deterministic processes the increase of the reference distribution stays at the same range of α for increasing d , provided that d is sufficiently greater than the “true” (set of) dimension(s) of the process.

In the case of a fixed finite observation segment and fixed radii, H_{ref} is generically broadened. As a consequence, sharp discriminations as in the limits $T \rightarrow \infty$, $r_{1,2} \rightarrow 0$ become impossible (even for single

fractals). Nevertheless, non-random contributions to a random null process tend to increase the frequency of small scaling indices also in the finite case, such that mass of the si-edf is shifted to smaller values of α . For this reason, discrepancies between the si-edf and H_{ref} in the left wing of the distributions are of particular interest.

Assume for example that the $x(t)$ are independent random variables distributed uniformly on the interval $[0, 1]$, and $d = 3$. Up to random fluctuations, $\{z(t)\}$ then covers the three-dimensional unit cube in a homogeneous manner. The corresponding scaling index distribution function increases over a small range of α centered around a value $d^* < d$ where $(d - d^*)$ depends on T and $r_{1,2}$. Other mechanisms (random, chaotic, or regular) typically produce clusters and voids, resulting in a scaling index distribution whose tails are more pronounced.

The distribution of the scaling indices also provides us with a clue to more refined analyses. For example, it allows to sort out, and study in itself, those parts of the time series corresponding to small scaling indices, which are of special interest as potential carriers of signals or patterns. Properly assessing the statistical significance of effects is particularly important in such a kind of refined analysis.

The main ingredients of the statistical framework adopted here are as follows.

- (a) It is assumed that the null mode behavior of the system under study can be described by a stochastic process $\{X_0(t): t = 1, \dots, T\}$ having a certain distribution \mathcal{D}_0 .
- (b) The given time series $\{x(t)\}$ is interpreted as the realization of another stochastic process $\{X(t): t = 1, \dots, T\}$ with distribution \mathcal{D} . The null hypothesis \mathcal{H}_0 is that $\{X(t)\}$ is in null mode, that is, has the same distribution as $X_0: \mathcal{D} = \mathcal{D}_0$.
- (c) In addition, there are n *calibration time series* $\{x_{\text{cal},i}(t): t = 1, \dots, T\}$ ($i = 1, \dots, n$). They are assumed to represent a sample of size n from the null process X_0 , that is, they are assumed to be independent and identically distributed copies of $\{X_0(t): t = 1, \dots, T\}$.

Some remarks: Subsequently, we drop the notational distinction between time series considered as a sequence of measurement data or considered as a stochastic process. Secondly, it should be noted that the null hypothesis does not impose any restriction on the internal structure of the null process. However, our proposals make sense only if the process is stationary (or nearly so at least). Thirdly, \mathcal{D} denotes the distribution of the whole stochastic process $\{x(t)\}$, not the (marginal) distribution of a single random variable $x(t)$ at fixed t (and likewise for \mathcal{D}_0 and $\{x_{\text{cal},i}(t)\}$). Finally, the “true” null distribution \mathcal{D}_0 need not be known, but it is assumed that a reasonable model is available for the null process which allows us to generate a sample of *auxiliary time series* $\{x_{\text{aux},j}(t): t = 1, \dots, T\}$ ($j = 1, \dots, m$) with a distribution close to \mathcal{D}_0 .

Both calibration data and auxiliary data can be considered as surrogate data [13], though of different type. Usually, surrogate data are generated to simulate the distribution of the test statistics under the null hypothesis. To serve this purpose, calibration data from a real physical system are available in our case. The auxiliary data are generated for a different purpose, namely to construct a test statistics adapted to the null process (see Section 2.2). Furthermore, in our approach the distribution \mathcal{D}_0 as a whole should be matched. This is in contrast to the typical use of surrogate data (*e.g.*, based on randomized Fourier phases [14]), where only some characteristics of \mathcal{D}_0 are taken into account.

2.2. Test Construction

As indicated above the test of the null hypothesis \mathcal{H}_0 proposed here will be based on the difference

$$\Delta(\alpha) = H(\alpha) - H_{\text{ref}}(\alpha)$$

of the si-edf $H(\alpha)$ associated with the time series $\{x(t)\}$ and the reference scaling index distribution function H_{ref} . As H_{ref} we use the mean distribution of the si-edf's $H_{\text{cal},i}$ as obtained from the

calibration time series $\{x_{\text{cal},i}(t)\}$:

$$H_{\text{ref}}(\alpha) = \overline{H_{\text{cal}}}(\alpha) = n^{-1} \sum_{i=1}^n H_{\text{cal},i}(\alpha). \quad (3)$$

This mean distribution may be regarded as an estimator of the expected value of $H(\alpha)$ (with respect to \mathcal{D}_0). The test depends on the partial maxima

$$M(\alpha) = \max_{\alpha_{\min} \leq \beta \leq \alpha} |\Delta(\beta)|$$

and a function $c(\alpha)$, henceforth referred to as the *deviation profile*. The null hypothesis \mathcal{H}_0 (*i.e.*, $\{X(t)\}$ is in null mode) is rejected if $M(\alpha)$ crosses the deviation profile $c(\alpha)$ somewhere in the interval $[\alpha_{\min}, \alpha_{\max}]$, *i.e.*, if

$$M(\alpha) > c(\alpha) \text{ for some } \alpha \in [\alpha_{\min}, \alpha_{\max}]. \quad (4)$$

The deviation profile $c(\alpha)$ and the interval $[\alpha_{\min}, \alpha_{\max}]$ have to be chosen such that the error of the first kind is (approximately) equal to the desired test level ε (typical choices are $\varepsilon = 0.05$ and $\varepsilon = 0.01$),

$$P_0(M(\alpha) > c(\alpha) \text{ for some } \alpha \in [\alpha_{\min}, \alpha_{\max}]) = \varepsilon. \quad (5)$$

The notation P_0 indicates that the probability of the event in brackets is computed with respect to the distribution \mathcal{D}_0 of the null process.

The deviation profile $c(\alpha)$ has to account for the strong α -dependence of the variance (under \mathcal{H}_0) of the random variables $\Delta(\alpha)$. Typically this variance is much larger for α -values in the central part of a scaling index distribution rather than in its tails. If c were chosen as a constant, possible weak systematic deviations in the left wing of $H(\alpha)$ would be concealed by large chance fluctuations in the center. For instance, this would be the case for a classical Kolmogorov–Smirnov type test, which rejects \mathcal{H}_0 if $\max_{\alpha} |\Delta(\alpha)| > c'$ – or, in terms of the partial maxima, if $M(\alpha) > c'$ for some α .

Using partial maxima $M(\alpha)$ instead of the (sometimes strongly fluctuating) function $\Delta(\alpha)$ has a certain regularizing effect, but the essential

features in the left tail of $H(\alpha)$ are retained because $|\Delta(\alpha)|$ tends to increase in that range. Another way to avoid strong fluctuations in $\Delta(\alpha)$ can be achieved by considering only subintervals $[\alpha_{\min}, \alpha_{\max}]$. These intervals have to be adjusted with respect to the parameters T and $r_{1,2}$.

The deviation profile $c(\alpha)$ will be constructed to reflect the local variability of $M(\alpha)$ as expected for the null process. This can be done in different ways. Two of them are described as Versions (A) and (B) below. Both versions use auxiliary time series $\{x_{\text{aux},j}(t)\}$ to obtain preliminary profiles. The final profile has to comply with condition (5), which is achieved by a simple multiplicative correction. For that purpose the calibration time series $\{x_{\text{cal},i}(t)\}$ are used. The null hypothesis is rejected according to (4).

Version (A):

$$c_A(\alpha) = \gamma_A \mu(\alpha),$$

where

$$\begin{aligned} \mu(\alpha) &= m^{-1} \sum_{j=1}^m M_{\text{aux},j}(\alpha) \\ &= m^{-1} \sum_{j=1}^m \max_{\alpha_{\min} \leq \beta \leq \alpha} |\Delta_{\text{aux},j}(\beta)|. \end{aligned} \quad (6)$$

The factor $\gamma_A > 0$ is determined (e.g., by trial and error) such that the (empirical) profile crossing probability

$$P_A = n^{-1} \#\{i \leq n: M_{\text{cal},i}(\alpha) > c_A(\alpha) \text{ for some } \alpha \in [\alpha_{\min}, \alpha_{\max}]\} \approx \varepsilon, \quad (7)$$

with

$$M_{\text{cal},i}(\alpha) = \max_{\alpha_{\min} \leq \beta \leq \alpha} |\Delta_{\text{cal},i}(\beta)| \quad (8)$$

calculated from the calibration data. The relative frequency P_A represents a bootstrap estimate of the corresponding error probability in (5). Of course, the number of calibration data and auxiliary data should be large enough to make this a stable estimate.

Version (B):

$$c_B(\alpha) = \mu(\alpha) + \gamma_B \sigma(\alpha),$$

where μ is the same as in (A), and $\sigma(\alpha)$ is computed as the root of the empirical variance

$$\sigma(\alpha)^2 = m^{-1} \sum_{j=1}^m M_{\text{aux},j}(\alpha)^2 - \mu(\alpha)^2 \quad (9)$$

from the auxiliary data. The factor γ_B is determined such that the empirical profile crossing probability P_B , defined analogous to P_A , satisfies (7).

Other variants of this approach in addition to these two versions may also be worth considering. For example, one could use the empirical standard deviations $s_{\Delta,\text{aux}}(\alpha) = \sqrt{(1/m) \sum_{j=1}^m \Delta_{\text{aux},j}(\alpha)^2}$ to construct the preliminary profile and then correct it by a factor γ_{Δ} such that one has (with $c_{\Delta} = \gamma_{\Delta} s_{\Delta}$)

$$P_{\Delta} = n^{-1} \#\{i \leq n: |\Delta_{\text{cal},i}(\alpha)| > c_{\Delta}(\alpha) \text{ for some } \alpha \in [\alpha_{\min}, \alpha_{\max}]\} \approx \varepsilon.$$

The null hypothesis \mathcal{H}_0 is then rejected if $|\Delta(\alpha)| > c_{\Delta}(\alpha)$ for some $\alpha \in [\alpha_{\min}, \alpha_{\max}]$.

2.3. An Example

Details of the proposed test construction for a specific example are illustrated in Figures 1–6. We had $n = 1320$ calibration time series (as described in [9]) at our disposal, each of length $T = 10000$. As a working hypothesis it is assumed that the $x_{\text{cal},i}(t)$ ($t = 1, \dots, T$; $i = 1, \dots, n$) are independent random variables, identically distributed according to a binomial distribution $B(n_b, \pi_b)$ with $n_b = 200$ trials and a probability of success $\pi_b = 1/2$. A number of $m = 2500$ auxiliary time series (roughly twice as much as the calibration data) were generated according to this “binomial model”. The embedding dimension is $d = 4$, and the radii are $r_1 = 4.6$, $r_2 = 12.7$.

Figure 1 shows H_{ref} together with an individual si-edf $H_{\text{cal},i}$. The $\alpha(t)$'s can only attain values in a certain finite set determined by T and $r_{1,2}$ and are

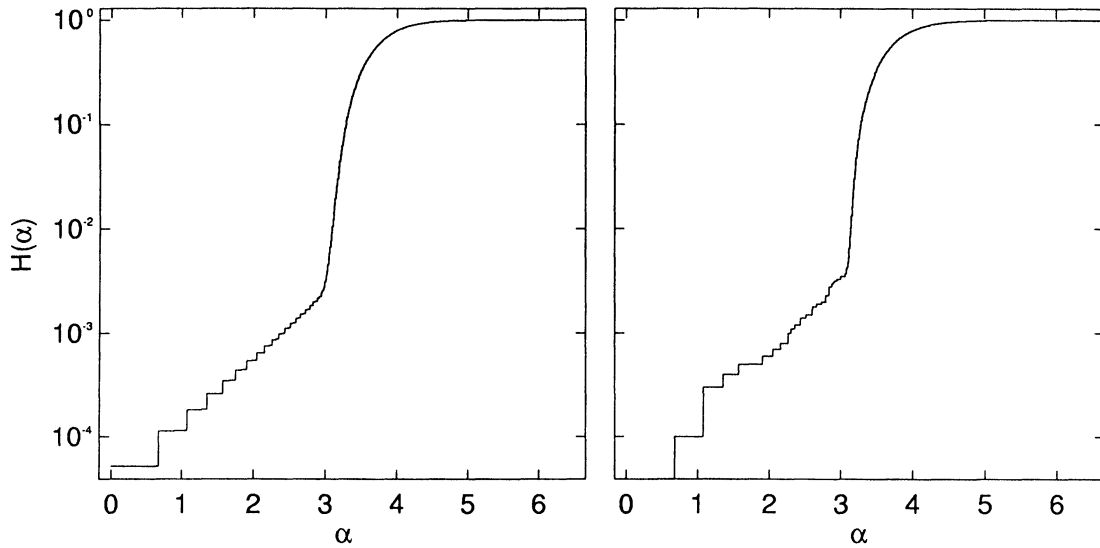


FIGURE 1 Distribution functions $H(\alpha)$ of scaling indices α . Left: Reference distribution function $H_{\text{ref}}(\alpha) = \overline{H_{\text{cal}}}(\alpha)$ as the empirical mean over 1320 calibration data sets of finite length $T = 10000$. Right: Distribution function $H_{\text{cal},i}$ of a single calibration data set. Notice the step-like behavior for $\alpha < 3$ in both distribution functions.

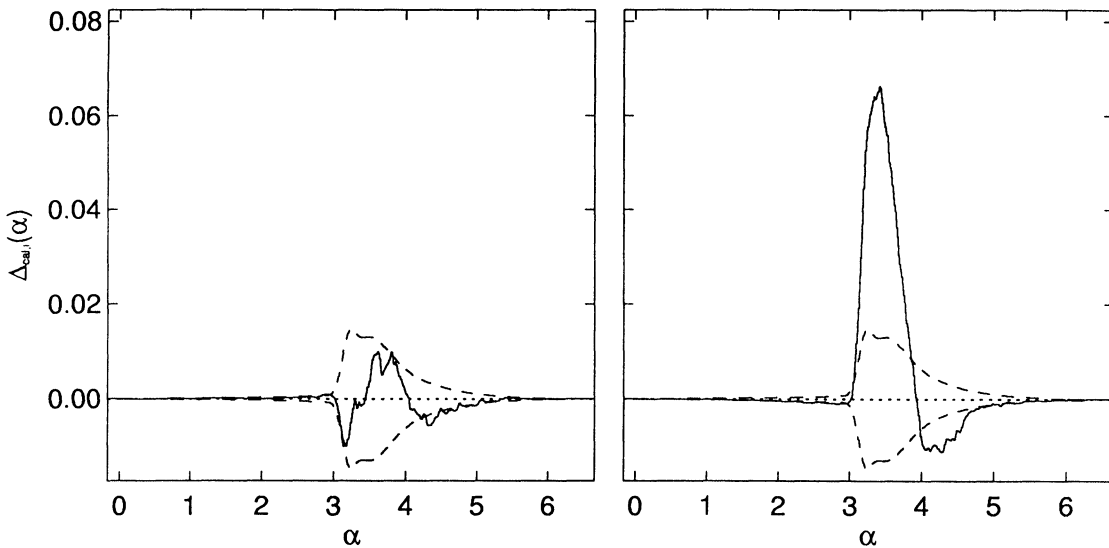


FIGURE 2 Deviation $\Delta_{\text{cal},i}(\alpha)$ of a single distribution function $H_{\text{cal},i}(\alpha)$ from the reference distribution function $H_{\text{ref}}(\alpha)$. Left: $\Delta_{\text{cal},i}(\alpha)$ for a single calibration time series. The dashed line indicates the 1σ contour of the calibration data sets used to calculate $H_{\text{ref}}(\alpha)$. Right: $\Delta_{\text{cal},i}(\alpha)$ for a time series generated with an i.i.d. random process of length $T = 10000$ with a binomial distribution $B(182, 0.55)$. The dashed line is the same 1σ contour as in the left panel.

therefore generically discrete (*cf.* (1)). In the left tail of H , points with small occupation numbers $N_{1,2}(z(t))$ produce discrete steps which are clearly visible in Figure 1. The central part of the distribution appears to be continuous, however,

since the corresponding occupation numbers are large.

The difference $\Delta(\alpha)$ of an individual si-edf H and H_{ref} is shown in Figure 2. In both cases H is computed from a simulated time series where the

$x(t)$ are independent and identically distributed random variables with a binomial distribution $B(n_b, \pi_b)$. The first case corresponds to the binomial model ($n_b = 200$, $\pi_b = 1/2$), whereas in the second case we set $n_b = 182$, $\pi_b = 0.55$. This leaves the expectation value of the $x(t)$'s (approximately) unchanged, but it reduces their variance from 50 to 45. The empirical standard error of the Δ 's calculated from the calibration data, $s_{\Delta, \text{cal}}(\alpha) = \sqrt{(1/n) \sum_{i=1}^n \Delta_{\text{cal}, i}(\alpha)^2}$, is also plotted in Figure 2 for comparison. For $B(182, 0.55)$ a prominent deviation from zero can be found within the range $3 < \alpha < 4$, whereas for $\alpha > 4$ Δ is of the same order as $s_{\Delta, \text{cal}}$. The partial maxima $M(0, \alpha)$, with lower limit $\alpha_{\min} = 0$, are shown in Figure 3 for the same data sets. The notation $M(\alpha_{\min}, \alpha)$ instead of $M(\alpha)$ indicates the interval over which the maximum is taken. The same notation will be adopted for the empirical error probabilities in (7), e.g., $P_A = P_A(\alpha_{\min}, \alpha_{\max})$.

Figure 4 exhibits the deviation profiles c_A and c_B (with the preliminary choice $\gamma_A = \gamma_B = 2$), and, in the lower plot, the (empirical) *pointwise* profile crossing probabilities $Q_A(\alpha)$ and $Q_B(\alpha)$. They are defined as the relative frequencies with which

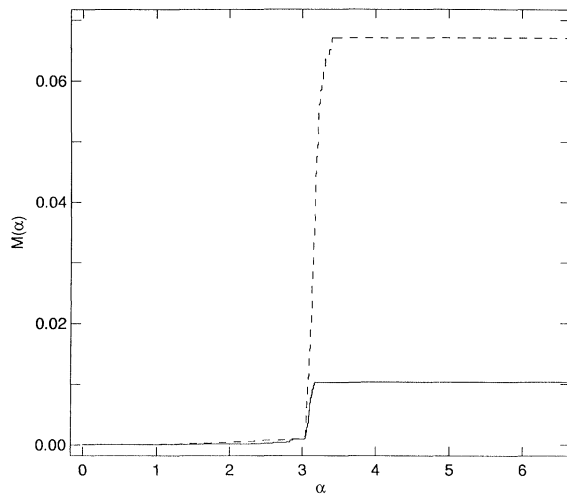


FIGURE 3 The solid line shows the partial maxima $M_{\text{cal}, A}(\alpha)$ (Eq. 8) for the case shown in the left panel of Figure 2. The dashed line represents the case shown in the right panel of Figure 2.

$M_{\text{cal}, i}(0, \alpha)$ exceeds the profile $c_{A, B}(\alpha)$ at a fixed value of α . The rather erratic behavior of these probabilities for small values of α is due to small occupation numbers $N_{1,2}(z(t))$. It prompts us to set $\alpha_{\min} = 3$ in Eq. (5), which approximates the “continuity onset” in Figure 1.

Figure 5 shows the (empirical) profile crossing probabilities $P_{A, B}(3, \alpha_{\max})$ as a function of α_{\max} , for $\gamma = 2$ and $\gamma = 3$. In all cases $P_{A, B}(3, \alpha_{\max})$ is constant for α_{\max} larger than ≈ 3.5 . Therefore α_{\max} is set to the (somewhat arbitrary) value 4. With this choice of α_{\min} and α_{\max} , and with an error level of $\varepsilon = 0.05$, the γ 's satisfying (7) are found to have the (approximate) values $\gamma_A = 3.42$, $\gamma_B = 2.94$. The deviation profiles $c_{A, B}$ corresponding to these specifications are shown in Figure 6.

Similar considerations will be necessary in any application of the proposed testing procedure. In particular, since the asymptotic distribution of the test statistics (under \mathcal{H}_0) is unknown, a sufficiently large pool of calibration time series is required in order to determine the γ -adjustment of the deviation profiles.

It might be interesting to mention that Ref. [9] discusses a concrete (real-world) example in which a similar, truncated form of Version A is used to detect non-random contributions in a random time series. In contrast to the detailed procedure proposed here, the procedure of [9] includes no adjustment of γ , and the profile $c_A(\alpha)$ is assumed as $\mu(\alpha)$ ($\gamma = 1$).

3. ENHANCING THE SIGNAL-TO-NOISE RATIO

Let us assume that a random null process is sporadically substituted or superimposed by some faint regular or deterministic signal. Is it possible to identify those subseries of the entire time series which contain the “deviant” non-random contributions? This section deals with the question whether those subseries can be found by selection procedures based on scaling indices. If this can be done, it is an attractive idea to construct tests of

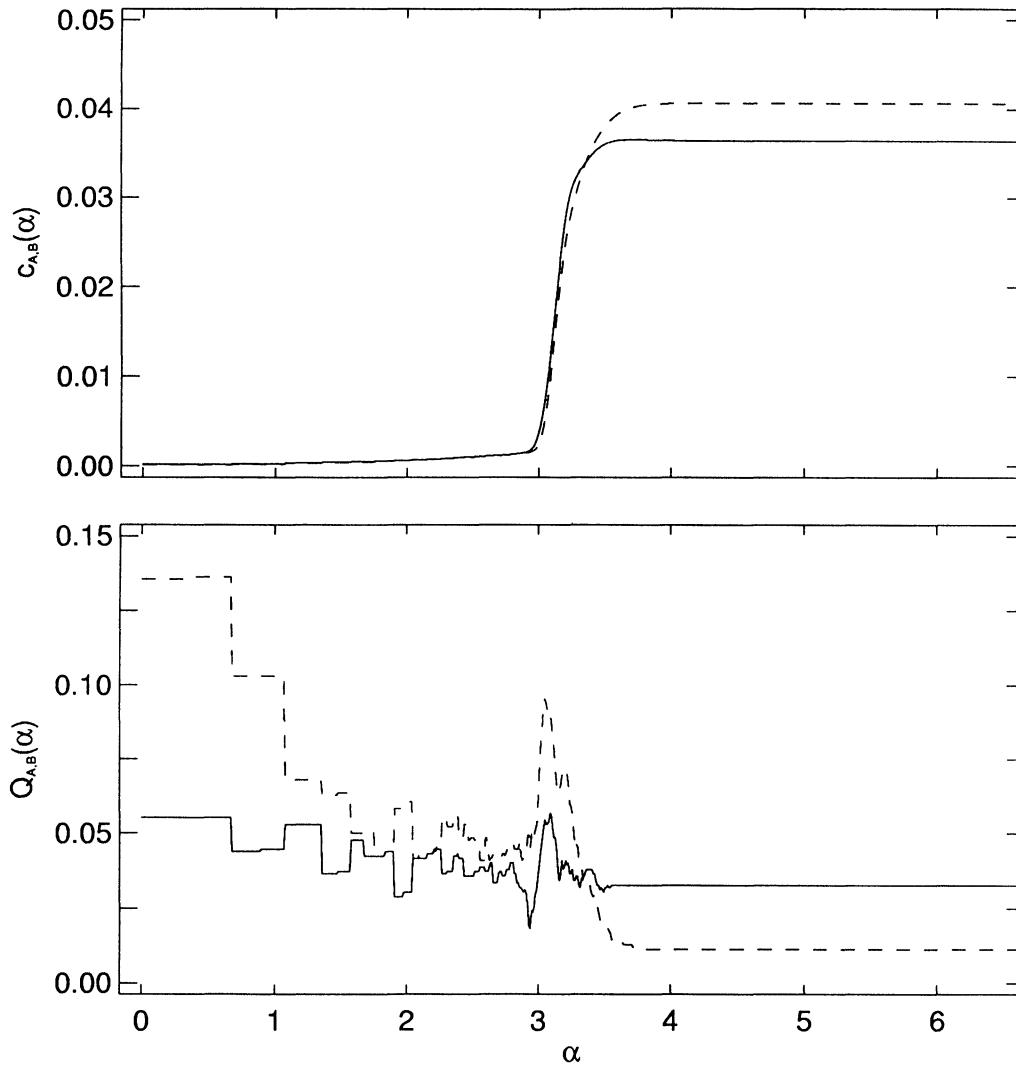


FIGURE 4 Upper panel: deviation profile $c(\alpha)$ for the two Versions (A) (dashed line) and (B) (solid line) of Section 2.2 with $\gamma=2$. Lower panel: the pointwise profile crossing probabilities Q_A (dashed line) and Q_B (solid line) are the relative frequencies with which $M_{\text{cal},f}(\alpha)$ exceeds the deviation profile of the upper panel for a single value α .

the null hypothesis based on the selected subseries and study their statistical characteristics. Some preliminary steps in this direction are included at the end of this section. Details in this respect will be addressed in future work.

Obviously there are other conceivable methods to detect deviations from randomness in a time series; *e.g.*, any type of autocorrelation analysis or, equivalently, Fourier analysis. If one focuses on the selection of local subseries containing those

deviations rather than a global characterization of them, wavelet techniques might be good candidates for alternative approaches. This, however, would be beyond the scope of the present paper.

Our main purpose in this paper is to make a concrete proposal for a selection method based on scaling index distributions and to investigate, by simulation, whether it yields a higher signal-to-noise ratio in the selected subseries than in the original time series. In this context, the signal-to-noise

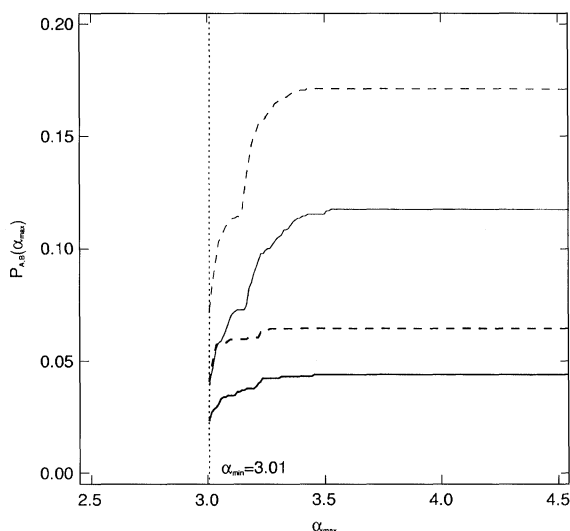


FIGURE 5 The profile crossing probabilities $P_{A,B}$ are the relative frequencies (as a function of α_{\max}) with which the function $M_{\text{cal},i}(\alpha)$ exceeds the deviation profile, $M_{\text{cal},i}(\alpha) > c(\alpha)$, somewhere in the range $3 < \alpha \leq \alpha_{\max}$. The dashed lines indicate P_A for the deviation profile of Version (A) of Section 2.2, the solid lines indicate the probability P_B for Version (B) of Section 2.2. Thin lines indicate $\gamma=2$, thick lines $\gamma=3$.

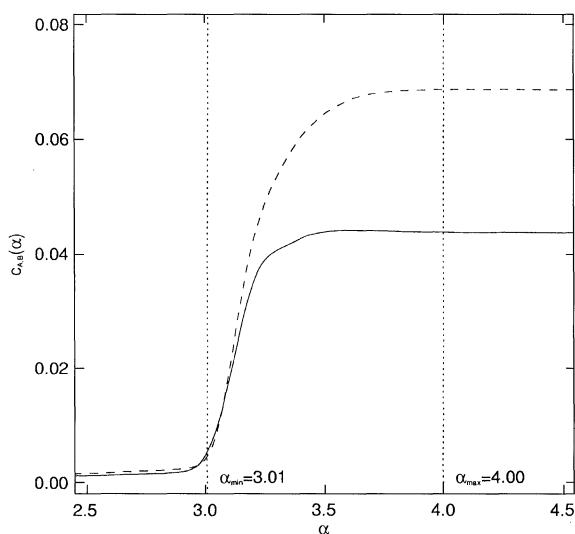


FIGURE 6 Deviation profiles $c_A(\alpha)$ (dashed line) and $c_B(\alpha)$ (solid line) according to Versions (A) and (B) of Section 2.2.

ratio is defined as the relative number of data points $x(t)$ due to the non-random signal.

Since non-random signals are basically expected to be associated with small scaling indices, the key

idea is to select all data points $x(t)$ whose scaling index $\alpha(t)$ falls below some threshold value A_1 . In order to avoid the discreteness problems discussed in Section 2.3, a lower threshold value A_0 is required in addition. Since $\alpha(t)$ depends on the d -tuple $z(t) = (x(t), \dots, x(t+d-1))$ rather than $x(t)$, selecting a data point at time t in principle implies selecting the subsequent $d-1$ data points as well. Therefore the selected subseries consists of data points $x(t)$ at instants t belonging to a set S which can be formally defined as

$$\begin{aligned} S &= \cup_{k=0}^{d-1} S_k, \text{ where} \\ S_0 &= \{t: A_0 \leq \alpha(t) \leq A_1\}, \\ S_k &= \{t+k: t \in S_0\} (1 \leq k \leq d-1). \end{aligned} \quad (10)$$

Depending on the nature of the non-random contribution to be detected, it may be advantageous to consider only subsets of S .

The choice of the threshold values is crucial and not obvious. For our purposes it will be conceived as depending both on H_{ref} and the si-edf H of the time series $\{x(t)\}$ under study. If the difference $\Delta(\alpha)$ attains its maximum at $\alpha = \alpha_+^*$ then it is expected that most of the α values corresponding to a non-random signal occur at $\alpha < \alpha_+^*$. The following strategy is proposed to fix this range: Define A_0 resp. A_1 as the minimal resp. maximal $\alpha < \alpha_+^*$ at which

$$\Delta(\alpha) = 0.05\Delta(\alpha_+^*) \quad \text{resp.} \quad \Delta(\alpha) = 0.75\Delta(\alpha_+^*).$$

Modifications of this strategy may be appropriate if A_0 falls below the continuity onset of H_{ref} or if A_1 is larger than the median of H_{ref} , say d^* . For the examples to be presented below, the fractions 0.05 and 0.75 have been found to produce the largest signal-to-noise enhancement.

To study the performance of the selection procedure, auxiliary time series are generated according to the binomial model of Section 2.3. These auxiliary time series are then perturbed by “signals” in such a way that at randomly selected times $\tau_1 < \dots < \tau_v$ the segment $x(\tau_j), x(\tau_j+1), \dots, x(\tau_j+L-1)$ of $\{x(t)\}$ is replaced by some L -tuple $\xi = (\xi_0, \dots, \xi_{L-1})$ fixed ad hoc.

(For $\tau_{j+1} - \tau_j < L$, each segment is overwritten to the extent to which it overlaps with the subsequent segment.) The perturbed data series are then used as the time series $\{x(t)\}$ under study. Within this setting, the signal-to-noise ratio is defined as $\#R/T$ for the full time series, and as $\#R \cap S / \#S$ for the subseries $\{x(t); t \in S\}$. Here R denotes the set of all instants for which the original x value has been replaced by a ξ value.

Table I summarizes the results for various tuples ξ of different length and form. Each entry is based on 20 time series generated as described above, with the same embedding dimension d and radii $r_{1,2}$ as in Section 2.3. The largest enhancements of the signal-to-noise ratio, up to a factor of 19, are found when $L \geq d (=4)$ and when the ξ values are close to the expectation value 100 of the $B(200, 1/2)$ distribution. This observation may be heuristically explained as follows.

Suppose that $L = d$, and that $\xi = (\xi_0, \dots, \xi_{d-1})$ gives rise to a scaling index of

$$\alpha = \frac{\log N_2 - \log N_1}{\log r_2 - \log r_1}, \text{ where}$$

$$N_{1,2} = \#\{t: |z(t) - \xi| \leq r_{1,2}\},$$

where the $z(t)$'s are constructed from the unperturbed time series. Let $\alpha' = \alpha + \delta\alpha$ and $N'_{1,2} = N_{1,2} + \delta N_{1,2}$ denote the corresponding quantities

for the perturbed series in which a (small) number $\nu = fT$ of tuples ξ replaces the original z tuples at randomly selected times. It is easy to see that $\delta\alpha \leq 0$, such that the scaling indices always decrease under such replacements. A rough quantitative estimate of that shift can be obtained on noting that a fraction f of the $(T - N_{1,2})$ tuples $z(t)$ outside the ball of radius $r_{1,2}$ around ξ is expected to be replaced by ξ -tuples. This gives the estimate $\delta N_{1,2} \sim f(T - N_{1,2})$, or

$$\delta\alpha \sim \log \frac{1-f+fT/N_2}{1-f+fT/N_1} \Big/ \log \frac{r_2}{r_1}$$

$$\sim -\frac{f}{1-f} \left(\frac{T}{N_1} - \frac{T}{N_2} \right) \Big/ \log \frac{r_2}{r_1}, \quad (11)$$

provided that $[f/(1-f)]T/N_{1,2}$ is sufficiently small. Therefore, the (negative) shift $\delta\alpha$ increases for decreasing N_1 . At first sight this seems to contradict the claim that the best enhancement is achieved for ξ values close to 100, where N_1 is large (and $\alpha \approx d^*$). However, the statistical variability of $\Delta(\alpha)$ decreases sharply with decreasing α in the left wing of $H(\alpha)$. On the other hand, in rarefied regions of the embedding space, where N_1 is relatively small, the (expected) α shift is large, so that α' tends to fall below the threshold A_0 .

The perturbed time series have a mean and standard deviation different from that of the

TABLE I Results of the selection procedure for random time series in which $\nu = 25$ sequences of x -values have been replaced by the ξ -tuples given in column 1. For each type of ξ -tuple $N_{\text{all}} = 20$ independent time series have been generated. N_{test} is the number of time series for which the test rejects \mathcal{H}_0 , hence indicating non-random contributions. N_{match} is the number of those elements in the selected subseries matching (in both time and amplitude) with elements of the inserted ξ -tuples. The medians and the standard deviations of N_{match} are listed. R_{sn} is the signal-to-noise ratio for the selected subseries and the total time series, respectively. For the selected subseries the medians and the standard deviations are listed as well

L -tuple $\{\xi_0, \xi_1, \dots, \xi_{L-1}\}$	N_{test} $N_{\text{all}} = 20$	N_{match} median \pm S.D.	R_{sn} (selected) median \pm S.D.	R_{sn} (total)
[100, 100, 100]	13	49 \pm 6.3	0.079 \pm 0.027	0.0075
[100, 100, 100, 100]	20	86 \pm 2.9	0.191 \pm 0.033	0.0100
[100, 100, 100, 100, 100]	20	111 \pm 2.3	0.240 \pm 0.033	0.0125
[107, 107, 107]	2	1 \pm 12.3	0.016 \pm 0.081	0.0075
[107, 107, 107, 107]	6	79 \pm 30.2	0.039 \pm 0.236	0.0100
[107, 107, 107, 107, 107]	20	105 \pm 5.0	0.135 \pm 0.291	0.0125
[101, 99, 98, 96]	16	86 \pm 7.6	0.128 \pm 0.046	0.0100
[101, 98, 101, 98]	17	83 \pm 3.6	0.137 \pm 0.030	0.0100
[107, 99, 101, 103]	5	41 \pm 27.6	0.032 \pm 0.022	0.0100

unperturbed series. However, the differences are too small to be easily detected by a t -test or a (variance ratio) F -test. What happens if these tests are applied to the selected subseries? Table II shows the number of subseries within the 20 time series of Table I for which the hypothesis is rejected that the mean and the variance are equal to those of the binomial model. These results can be compared to those for subseries of the same length, which are selected randomly (*i.e.*, not by the described procedure) from the respective perturbed time series. While the numbers for those randomly selected subseries for which the null hypothesis is rejected are consistent with the required type I error $\varepsilon = 0.05$, the numbers for the subseries selected by the described procedure significantly exceed this type I error except for the t -test and tuples with values

close to the mean of the unperturbed series. It has to be noted, however, that the null distributions of the test statistics for the selected subseries generally are different from the usual ones. To obtain a fair comparison further simulations are required to determine appropriate reference (null) distributions.

Tables III and IV show corresponding results for random time series according to the binomial model in which the elements have been reordered in such a way that the same ξ -tuples as used for Tables I and II are created, however without introducing new values of $x(t)$. In contrast to the replacement procedure discussed above, such a reordering keeps the state distribution invariant but may change the distribution of transition probabilities between states considerably. For each type of tuple, 20 independent time series have been

TABLE II Results of t -tests and F -tests applied to selected subseries. N_{test} is the same as in Table I. P_t is the number of subseries selected by scaling indices for which the hypothesis of no mean shift is rejected (t -test, two-sided, $\varepsilon = 0.05$). P_F is the number of subseries selected by scaling indices for which the hypothesis of no variance change is rejected (F -test, two-sided, $\varepsilon = 0.05$). P_t (random) and P_F (random) denote the corresponding numbers for randomly selected subseries

L -tuple $\{\xi_0, \xi_1, \dots, \xi_{L-1}\}$	N_{test} $N_{\text{all}} = 20$	P_t	P_F	P_t (random)	P_F (random)
[100, 100, 100]	13	4	20	1	1
[100, 100, 100, 100]	20	1	20	1	2
[100, 100, 100, 100, 100]	20	0	20	2	2
[107, 107, 107]	2	5	17	1	0
[107, 107, 107, 107]	6	13	20	1	0
[107, 107, 107, 107, 107]	20	15	19	0	0
[101, 99, 98, 96]	16	18	20	1	0
[101, 98, 101, 98]	17	3	20	0	2
[107, 99, 101, 103]	5	11	20	1	0

TABLE III Results of the selection procedure for binomial model time series which have been reordered to generate 25 ξ -tuples given in column 1. N_{test} , N_{all} , N_{match} , and R_{sn} as in Table I

L -tuple $\{\xi_0, \xi_1, \dots, \xi_{L-1}\}$	N_{test} $N_{\text{all}} = 20$	N_{match} median \pm S.D.	R_{sn} (selected) median \pm S.D.	R_{sn} (total)
[100, 100, 100]	6	48 \pm 13.3	0.086 \pm 0.032	0.0075
[100, 100, 100, 100]	19	84 \pm 3.9	0.216 \pm 0.028	0.0100
[100, 100, 100, 100, 100]	20	109 \pm 2.7	0.260 \pm 0.027	0.0125
[107, 107, 107]	2	1 \pm 16.9	0.004 \pm 0.009	0.0075
[107, 107, 107, 107]	6	81 \pm 23.3	0.070 \pm 0.101	0.0100
[107, 107, 107, 107, 107]	20	106 \pm 2.4	0.117 \pm 0.187	0.0125
[101, 99, 98, 96]	12	83 \pm 4.0	0.131 \pm 0.063	0.0100
[101, 98, 101, 98]	15	80 \pm 23.0	0.141 \pm 0.055	0.0100
[107, 99, 101, 103]	2	43 \pm 32.0	0.027 \pm 0.042	0.0100

TABLE IV Results of runs tests for the reordered time series of Table III. P_{total} (P_{selected}) is the number of total (selected) reordered time series for which the independence hypothesis (no temporal order) is rejected ($\varepsilon = 0.05$)

L -tuple $\{\xi_0, \xi_1, \dots, \xi_{L-1}\}$	N_{test} $N_{\text{all}} = 20$	P_{total}	P_{selected}
[100, 100, 100]	6	3	9
[100, 100, 100, 100]	19	1	6
[100, 100, 100, 100, 100]	20	1	7
[107, 107, 107]	2	1	9
[107, 107, 107, 107]	6	4	11
[107, 107, 107, 107, 107]	20	8	12
[101, 99, 98, 96]	12	0	9
[101, 98, 101, 98]	15	2	19
[107, 99, 101, 103]	2	1	12

generated. The same selection procedure as for Tables I and II has been applied. As in case of replacement, reordering shows that the selection of subseries with small α can lead to an enhancement of the signal-to-noise ratio up to a factor of 21.

The reordered time series has the same state distribution as the original time series but different transition probabilities. In particular, the elements $x(t)$ are statistically dependent. Therefore, sequential correlations (order) are introduced in the generated time series. These correlations are checked with a runs test (Tab. IV). Comparing P_{total} and P_{selected} in Table IV shows that the runs test is more sensitive when applied to the selected subseries of the reordered time series rather than to the whole reordered series. Of course, the same qualifying remarks as made above concerning t - and F -tests apply with respect to P_{selected} . Comparing the whole reordered time series to the original random time series, only tuples of the type [107,107,107, 107,107] lead to a significant difference. For all other types of tuples, P_{total} in Table IV shows no significant difference from the original random time series.

4. DISCUSSION

There are two ways in which the test proposed in Section 2.2 depends on the null process. (i) It is not

distribution free, in the sense that the distribution of the test statistics is in general different for different underlying null distributions \mathcal{D}_0 . (ii) Even the test statistics itself depends, *via* the deviation profiles $c(\alpha)$, on the assumed model (for the null process) used for the generation of the auxiliary time series. As a consequence of (i), the γ -adjustments have to be computed anew (on the basis of the given calibration data) if the test is to be applied to another null process. Further massive computation is needed for the construction of the deviation profiles.

The null processes considered in our examples are all of the i.i.d. type. It is expected that qualitatively similar results would be obtained for state (null) distributions different from the binomial distribution $B(200,1/2)$ (our main example), provided they are not too long-tailed. The situation must be more complex with null processes exhibiting statistical dependencies between states, *e.g.*, ARMA processes containing deterministic components.

The power of the test depends on the special type of deviation from the null hypothesis. If the process $\{x(t)\}$ under study is i.i.d. and only changes of the state distribution of $\{x(t)\}$ are considered, there are well-known tests which are more efficient than the one proposed here. Since $H(\alpha)$ is invariant under translations of the form $x(t) \rightarrow x(t) + \text{constant}$, the test has no power at all against mean shifts in the state distribution. It does have some power against scale dilations (variance changes), which cannot compete, however, with the power of the (variance ratio) F -test which is especially tailored to deviations of that kind. For instance, for time series generated according to the binomial model of Section 2.3 with $B(193, 0.482)$, corresponding to a reduction of the variance by 3.63%, our test rejects the null hypothesis in about 20% of the data series, whereas the F -test (one-sided, $\varepsilon = 0.05$) has a rejection rate of about 50%.

More interesting situations arise if the time series is not stationary and/or its elements are correlated, particularly when this is the case only sporadically within the observed segment. Corresponding types

of deviations can be studied with random time series (of the i.i.d. type) modified by various kinds of perturbations at randomly selected instants (or intervals). For example, tuples of L subsequent x -values may be replaced by some ad hoc fixed L -tuple ξ , or the time series may be reordered in such a way that the ξ 's appear at randomly selected instants (or intervals).

As can be seen from Tables I and III, the test proposed in Section 2.2 is the more sensitive the less “complex” the non-random process is. For instance, a deterministic process with a small number of different states around the median of the $x(t)$'s of the random (null) process with a simple temporal order of the tuples results in a high detection rate. The test works equally well when the empirical distribution of the perturbed $x(t)$'s is the same as for the unperturbed ones, *i.e.*, when the perturbation consists in a reordering of the time series. If the perturbation consists of a replacement by a fixed L -tuple ξ , the power tends to be the higher the closer the ξ -values are to the median of the $x(t)$'s. This may be contrasted with the behavior of standard tests which often react sensitively to outliers but draw little information from data points close to the center of the state distribution.

Summing up, the test proposed in this paper is well suited for time series in which (1) states generated by a deterministic process with low complexity are to be separated from random elements, and (2) the deterministic elements are close to the median of the random elements. Within the investigated time series, subseries can be selected for which the signal-to-noise ratio is enhanced compared to the total time series.

References

- [1] Grassberger, P., Badii, R. and Politi, A. (1988). Scaling laws for invariant measures on hyperbolic and nonhyperbolic attractors. *J. Stat. Phys.*, **51**, 135–178.
- [2] Mandelbrot, B. B. (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**, 331–358.
- [3] Frisch, U. and Parisi, G. (1985). On the singularity structure of fully developed turbulence. In: *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, Eds. Ghil, M., Benzi, R. and Parisi, G., North Holland, New York, pp. 84–88.
- [4] Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I. and Shraiman, B. I. (1986). Fractal measures and their singularities: the characterization of strange sets. *Phys. Rev. A*, **33**, 1141–1151.
- [5] Paladin, G. and Vulpiani, A. (1987). Anomalous scaling laws in multifractal objects. *Phys. Rep.*, **156**, 147–225.
- [6] Falconer, K. (1990). *Fractal Geometry, Mathematical Foundations and Applications*. Wiley, New York.
- [7] Atmanspacher, H., Scheingraber, H. and Wiedenmann, G. (1989). Determination of $f(\alpha)$ for a limited random point set. *Phys. Rev. A*, **40**, 3954–3963.
- [8] Atmanspacher, H., Wiedenmann, G. and Amann, A. (1995). Descartes revisited: the endo/exo distinction and its relevance for the study of complex systems. *Complexity*, **1**(3), 15–21.
- [9] Atmanspacher, H., Bösch, H., Boller, E., Nelson, R. D. and Scheingraber, H. (1999). Deviations from physical randomness due to human agent intention? *Chaos, Solitons, and Fractals*, **10**, 935–952.
- [10] R ath, C. and Morfill, G. (1997). Texture detection and texture discrimination with anisotropic scaling indices. *J. Opt. Soc. Am. A*, **14**, 3208–3215.
- [11] Wiedenmann, G., Scheingraber, H. and Voges, W. (1997). Source detection with the scaling index method. In: *Data Analysis in Astronomy V*, Eds. di Gesu, V. *et al.*, World Scientific, Singapore, pp. 203–211.
- [12] Packard, N. H., Crutchfield, J. P., Farmer, J. D. and Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, **45**, 712–716.
- [13] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, **58**, 77–94.
- [14] Theiler, J., Linsay, P. S. and Rubin, D. M. (1993). Detecting nonlinearity in data with long coherence times. In: *Time Series Prediction*, Eds. Weigend, A. S. and Gershenfeld, N. A. Addison-Wesley, Reading, pp. 429–455.