Detection of a nonlinear oscillator underlying experimental time series: the sunspot cycle

Milan Paluš

Abstract

After a brief review of a nonlinearity test based on information-theoretic functionals (redundancies) and surrogate data technique, we discuss problems of this and similar tests for nonlinearity. In particular, we stress that a formal rejection of a linear stochastic null hypothesis does not automatically mean an evidence for nonlinear dynamical origin of studied data. In an example we show how a variable variance could be mistaken for nonlinearity in a series of surface air pressure. Therefore we find a detection of nonlinearity in a series of sunspot numbers insufficient for an identification of a mechanism underlying the sunspot cycle. As a solution in this case we propose to test for a property of nonlinear oscillators – mutual dependence between their instantaneous amplitude and frequency. This behavior is detected in yearly and monthly records of the sunspot numbers using histogram-adjusted isospectral surrogate data and Barnes model as ARMA surrogates. The instantaneous amplitudes and frequencies are obtained by means of the analytic signal approach using the discrete Hilbert transform. In several tests the amplitude-frequency correlation has been found significant on levels ranging from $p < 0.03$ to $p < 0.07$, which supports the hypothesis of a driven nonlinear oscillator as a mechanism underlying the sunspot cycle.

1 Introduction

Let $\{y(t)\}$ be a time series, i.e., a series of measurements done on a system in consecutive instants of time $t = 1, 2, \ldots$. Can we identify a mechanism underlying temporal evolution of such a system?

The time series $\{y(t)\}$ can be considered as a realization of a stationary linear stochastic process $\{Y(t)\}$. Without loss of generality we can set its mean to zero. Then the linear stochastic process $Y(t)$ can be written as:

$$Y(t) = Y(0) + \sum_{i=1}^{\infty} a(i)Y(t-i) + \sum_{i=0}^{\infty} b(i)N(t-i), \qquad (1)$$

where $b(0) = 1$, $\sum_{i=1}^{\infty} |a(i)| < \infty$, $\sum_{i=0}^{\infty} |b(i)| < \infty$, $\{N(t)\}$ is an independent, identically distributed (iid), normally distributed process with zero mean and finite (constant) variance. (For more details see [Priestley, 1981].)

Alternatively, the time series $\{y(t)\}$ can be considered as a projected trajectory of a dynamical system, evolving in some measurable d-dimensional state space. To be more specific, let X_t denote a state vector in R^d. Then the measurements $y(t)$ are obtained as $y(t) = g(X_t)$, where $g(.)$ is a projection (measurement function), and

temporal evolution of X_t may be described by a discrete-time dynamical system (a difference equation):

$$X_t = F(X_{t-1}), \tag{2}$$

with $X_0 \in R^d$ and for $t \geq 1$.

Due to ubiquity of noise, it is more realistic to replace the above states by random variables and the dynamics by a Markovian model such as

$$X_t = F(X_{t-1}, e_t), \tag{3}$$

where $t \in Z_+$, $F : R^{2d} \to R^d, \{e_t\}$ is a sequence of independent and identically distributed d-dimensional random vectors and e_t is independent of X_s, $0 \leq s < t$. We call $\{e_t\}$ the *dynamic noise.* Following [Tong, 1995], we refer to equation (2) as the *skeleton* of model (3), considering $F(X) = F(X, 0)$. For convenience, it is frequently assumed that the dynamic noise is additive so that equation (3) reduces to the *model with additive noise*

$$X_t = F(X_{t-1}) + e_t. \tag{4}$$

Detection of nonlinearity in experimental time series, i.e., identification from experimental data of underlying mechanism such as the model (4) with a nonlinear function F is usually based on rejection of a linear null hypothesis by a statistical test. Typically, the considered null hypothesis is a linear Gaussian process such as (1) or a Gaussian process passed through a static nonlinear transformation, or a similar simple alternative. Rejection of such a null is frequently interpreted as a detection of a deterministic nonlinear relation (2 or 4) in data under study. This is, however, only one of possible alternatives. Other alternatives will be discussed in Sec. 3, in the following section we will briefly review a test for nonlinearity proposed by in [Paluš, 1995]. Then we will present an example of a spurious nonlinearity detection due to a variable variance in a case of air surface pressure data (Sec. 4). Nonlinearity in a series of sunspot numbers is tested in Sec. 5. Section 6 introduces the amplitude-frequency correlation in nonlinear oscillators. Detection of such a behavior in the sunspot cycles is presented in Sec. 7, and results are discussed in Sec. 8.

2 A test for nonlinearity based on redundancies and surrogate data

Consider n discrete random variables $X_1, \ldots, X_n$ with sets of values $\Xi_1, \ldots, \Xi_n$, respectively. The probability distribution for an individual X_i is $p(x_i) = \Pr\{X_i = x_i\}$, $x_i \in \Xi_i$. We denote the probability distribution function by $p(x_i)$, rather than $p_{X_i}(x_i)$, for convenience. Analogously, the joint distribution for the n variables $X_1, \ldots, X_n$ is $p(x_1, \ldots, x_n)$. The redundancy $R(X_1; \ldots; X_n)$, in the case of two variables also known as mutual information $I(X_1; X_2)$, quantifies average amount of common information, contained in the n variables $X_1, \ldots, X_n$:

$$R(X_1; \ldots; X_n) = \sum_{x_1 \in \Xi_1} \cdots \sum_{x_n \in \Xi_n} p(x_1, \ldots, x_n) \log \frac{p(x_1, \ldots, x_n)}{p(x_1) \ldots p(x_n)}. \tag{5}$$

When the discrete variables $X_1, \ldots, X_n$ are obtained from continuous variables on a continuous probability space, then the redundancies depend on a partition ξ

chosen to discretize the space. Various strategies have been proposed to define an optimal partition for estimating redundancies of continuous variables (see [Paluš, 1995, Weigend and Gershenfeld, 1993] and references therein). Here we use the "marginal equiquantization" method described in detail in [Paluš, 1993, Paluš, 1995].

Now, let the n variables $X_1,\ldots,\ X_n$ have zero means, unit variances and correlation matrix $\mathbf{C}$. Then, we define the *linear redundancy* $L(X_1;\ldots;X_n)$ of $X_1, X_2, \ldots, X_n$ as

$$L(X_1;\ldots;X_n) = -\frac{1}{2}\ \sum_{i=1}^{n} \log(\sigma_i), \tag{6}$$

where σ_i are the eigenvalues of the $n \times n$ correlation matrix $\mathbf{C}$.

If $X_1, \ldots, X_n$ have an n-dimensional Gaussian distribution, then $L(X_1;\ldots;X_n)$ and $R(X_1;\ldots;X_n)$ are theoretically equivalent.

In practical applications one deals with a time series $\{y(t)\}$, considered as a realization of a stochastic process $\{Y(t)\}$, which is stationary and ergodic. Then, due to ergodicity, all the subsequent information-theoretic functionals are estimated using time averages instead of ensemble averages, and the variables X_i are substituted as

$$X_i = y(t + (i-1)\tau). \tag{7}$$

Due to stationarity the redundancies

$$R^n(\tau) \equiv R(y(t); y(t+\tau); \ldots; y(t+(n-1)\tau)) \tag{8}$$

and

$$L^n(\tau) \equiv L(y(t); y(t+\tau); \ldots; y(t+(n-1)\tau)) \tag{9}$$

are functions of n and τ, independent of t.

The surrogate-data based nonlinearity tests [Theiler et al., 1992, Paluš, 1995] consist of computing a *nonlinear* statistic from data under study and from an ensemble of realizations of a linear stochastic process, which mimics "linear properties" of the studied data. If the computed statistic for the original data is significantly different from the values obtained for the surrogate set, one can infer that the data were not generated by a linear process; otherwise the null hypothesis, that a linear model fully explains the data, cannot be rejected and the data can be analyzed, characterized and predictions can be obtained by using well-developed linear methods. For the purpose of such test the surrogate data must preserve the spectrum and consequently, the autocorrelation function of the series under study. An isospectral linear stochastic process to a series can be constructed by computing the Fourier transform (FT) of the series, keeping unchanged the magnitudes of the Fourier coefficients, but randomizing their phases and computing the inverse FT into the time domain. Different realizations of the process are obtained using different sets of the random phases.

The FT surrogates tend to have a Gaussian distribution which is not always the case of the tested data. In order to avoid a possible influence of different histograms of the data and of the surrogates, the histogram adjusted FT (HAFT) surrogates are constructed. (In [Theiler et al., 1992] the term "amplitude-adjusted" – AAFT surrogates is used.) In this case, the raw data undergo a nonlinear transformation which leads to a Gaussian distribution of the transformed data ("gaussianization" — see [Paluš, 1995] and references within). The gaussianized data are used to generate the

FT surrogates as described above, and the obtained surrogate data are transformed in order to have the same histogram as the original raw data. An application of the HAFT surrogates effectively means a reformulation of the null hypothesis of a Gaussian linear stochastic process (1) into a hypothesis of a process (1), realizations of which are passed through a static nonlinear transformation.

To evaluate the test, usually, as well as in [Paluš, 1995], the test statistic is defined as a difference between the redundancy obtained for the original data and the mean redundancy of a set of surrogates, in the number of standard deviations (SD's) of the latter. Such an approach requires a relative small number of surrogate replications, however, a normal distribution of the test statistic is considered. The latter is not always the case, and therefore, as a more reliable approach, in Sec. 5 we generate a large number of surrogate realizations and estimate percentiles (e.g., 2.5th and 97.5th percentile) of the surrogate distribution. Then redundancy values are simply compared with related percentile values. It is also possible to directly estimate the "significance", i.e., the p value of the test by counting the percentage of the surrogate realizations yielding redundancy values equal to or greater than the related redundancy obtained from analyzed data.

Note, that both the redundancies and redundancy-based statistics can be evaluated as functions of lag τ and embedding dimension n. Evaluating the redundancies and related statistics for broad ranges of lags (and several dimensions), however, can bring the problem of simultaneous statistical inference (see [Paluš, 1995, Paluš and Novotná, 1994] and references within for details). This approach, however, can be more reliable than single-valued tests, as it was demonstrated in [Paluš, 1995].

The redundancy $R^n(\tau)$, based on probability distributions, measures general dependences among the series $\{y(t)\}$ and its lagged versions, whereas the linear redundancy $L^n(\tau)$, based on correlations, reflects only their linear relations. Comparing the plots of $R^n(\tau)$ and $L^n(\tau)$ can provide an informal test for important nonlinearities in the studied data [Paluš et al., 1993]. This approach, in [Paluš, 1995] referred to as *qualitative testing*, or, *qualitative comparison*, can bring some additional information to the results of the quantitative (surrogate data based) test. Moreover, one can not always construct good surrogate data. That is, despite theoretical expectations, in numerical practice linear properties of the surrogates may differ from those of the data under study. Changes in linear properties are reflected in nonlinear statistics as well, and thus they may result in spurious detection of nonlinearity in linear data [Paluš, 1995]. Therefore, we also evaluate a statistic based on the linear redundancy $L^n(\tau)$ (defined analogously to the statistic based on $R^n(\tau)$), which specifically reflects changes in linear properties. Then, only those significant differences in the nonlinear statistic can reliably count for nonlinearity, which are not detected in the linear statistic [Paluš, 1995].

3 The null hypothesis of nonlinearity tests and its negations

Consider that the above described test yielded a significant result, i.e., the null hypothesis was reliably rejected. The null hypothesis was equivalent[1] to a linear stochastic process such as that described by the ARMA model (1). It is very common in nonlinear dynamics literature to consider the rejection of the null (1) as an evidence for a process such as (4) with a nonlinear skeleton (2). This is, however, only one of possible negations of (1). A number of different processes should be considered, which possess a linear deterministic skeleton[2], i.e., a linear AR part – the first sum in (1), or no deterministic skeleton at all (MA processes), however, their innovations $\{N(t)\}$ do not fulfill the conditions given above. Generally, one or more of the following properties could reject the null (1):

1. The innovations $\{N(t)\}$ are not Gaussian.
2. The innovations $\{N(t)\}$ are not an iid process, where iid means that the innovations should be not only uncorrelated, but generally independent.
3. The variance of $\{N(t)\}$ is not constant.

4 An example of surface air pressure series

A series of daily recordings of surface air pressure {P(t)} (t = 1 – 65,536 days; i.e., 180 years) was described and analyzed in [Paluš and Novotná, 1994]. Here we analyze differences from long-term daily averages. This transformation of data (almost entirely) removes the oscillations with the period of one year (seasonality in mean). We ask the question about possible long-term nonlinear dependence and perform the linear redundancy – redundancy surrogate data test, described above.

The results are presented in Fig. 1. For lags larger than a few days there is only a weak linear dependence, as measured by the linear redundancy (Fig. 1a) and reflected in the surrogates, but the (nonlinear) redundancy detects a clear dependence as an oscillatory structure with a yearly periodicity. This difference is highly significant (10 – 30 SD's, Fig. 1d), while no significance in the linear statistic (Fig. 1c) confirm the quality of the surrogates, which reflect correctly "the linear properties" of the data (in the sense of the model (1)). Can this result be understood as an evidence for the model (4) with a nonlinear periodic skeleton F, which could provide predictability of the atmospheric pressure for several years in advance?

The seasonality in mean present in this data (Fig. 2, upper panel) was mostly removed by considering differences from the long term averages.[3] The problem is that also the variance of this data is not constant, but clearly seasonal (Fig. 2, lower panel, the standard deviation (SD) is the square root of variance). This property is "nonlinear" in the sense that the surrogate data and the model (1) possess a constant

[1] Principal equivalence and technical differences between the surrogate data constructed by using the Fourier transform and the ARMA modeling are discussed in [Theiler et al., 1992, Theiler and Prichard, 1996].

[2] Obviously, for a linear function F, the model (4) is a special case of (1).

[3] The seasonality in mean can be entirely removed by filtration in spectral domain, as done in [Paluš and Novotná, 1994].

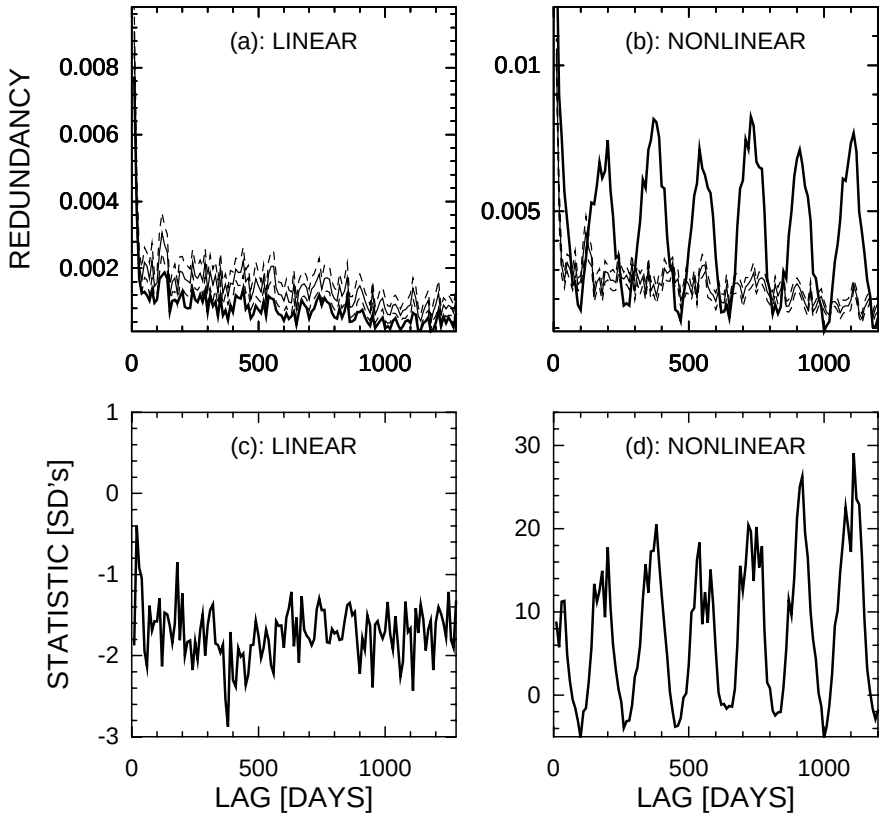

Figure 1: (a): Linear redundancy $L(y(t); y(t+\tau))$ (solid line), (b): nonlinear (general) redundancy $R(y(t); y(t+\tau))$ (solid line), for a series of differences from the long term averages of the surface air pressure (Prague-Klementinum station) and for its FT surrogates (thin solid and dashed lines present mean and mean$\pm$SD, respectively, of a set of 30 surrogate realizations); (c): linear (L-based), and d): nonlinear (R-based) statistics; as functions of the time lag τ, measured in days.

variance and cannot reproduce the seasonality in variance. After rescaling the data in order to obtain a constant variance, the effect of the false long-term nonlinear dependence is lost (Fig. 3). In the rescaled data there is no long-term dependence except of weak linear link due to not entirely removed seasonality in mean.

The above example of the surface air pressure clearly demonstrated the influence of variable variance on the redundancy – surrogate data nonlinearity test (introduced in Sec. 2 and in detail in [Paluš, 1995]). The effect of non-Gaussian innovations $\{N(t)\}$ was discussed in [Paluš, 1995], and a possible influence of non-iid $\{N(t)\}$ (i.e., innovations containing (nonlinear) temporal structures) is understandable. It is important to note that similar effects of "defective" innovations in a process under study would effect not only this particular test for nonlinearity, but all tests which use some type of FT/ARMA surrogates, and also any method which contain the process

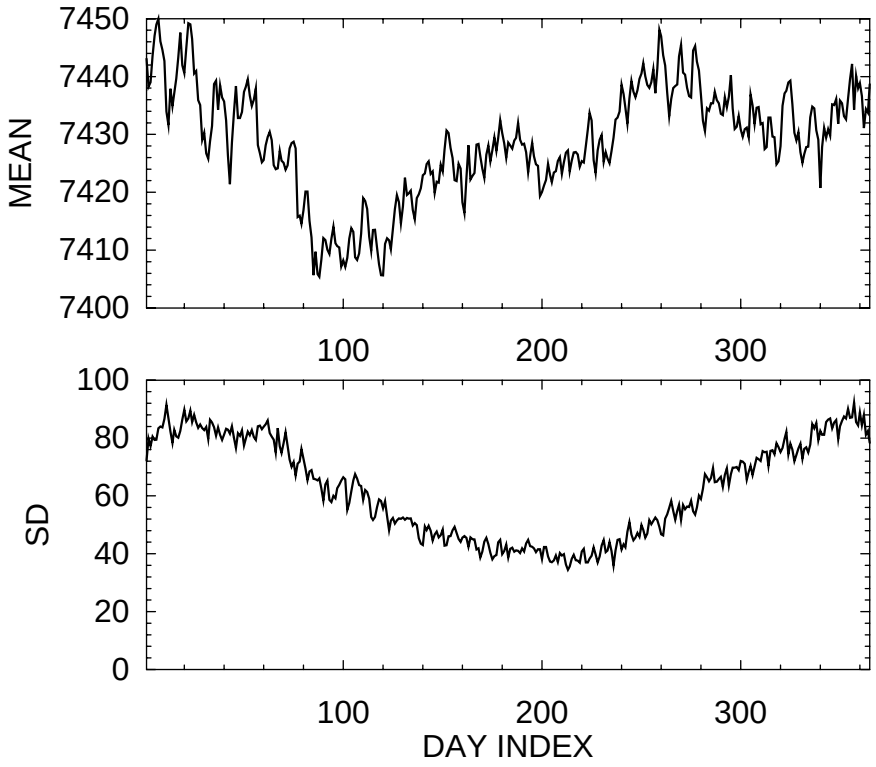

Figure 2: Seasonality in mean (upper panel) and in variance (lower panel) of the surface air pressure series – long term means (upper panel) and standard deviations (square root of variance, lower panel) for each day in a year. The days are consecutively numbered, January 1 has the index 1, January 2 has the index 2, ..., February 1 has the index 32, etc.

(1) at least implicitly in its construction. Also, all entropy-related statistics, that is, not only the above information-theoretic functionals, but also, for instance, statistics based on correlation integral [Prichard and Theiler, 1995], are extremely sensitive to variable variance and/or to (non)Gaussianity of data/innovations. Therefore one must very carefully assess results of nonlinearity tests in order to avoid confusing this kind of effects with actual nonlinear functional dependence in the data under study.

5 Nonlinearity in the sunspot cycle

An energy output of the Sun as the main basis of life on the Earth is nearly constant. However, the Sun is far from being uniform. The best observed solar inhomogeneities are spots on the solar surface in which the luminosity is diminished but magnetic fields appear which are stronger than usual magnetic fields on the rest of the solar surface. Appearance of the spots on the Sun is quantified by so-called sunspot numbers, or index, which have been collected from the beginning of the 18th century.

The historical data of the sunspot index have been attracting researchers for

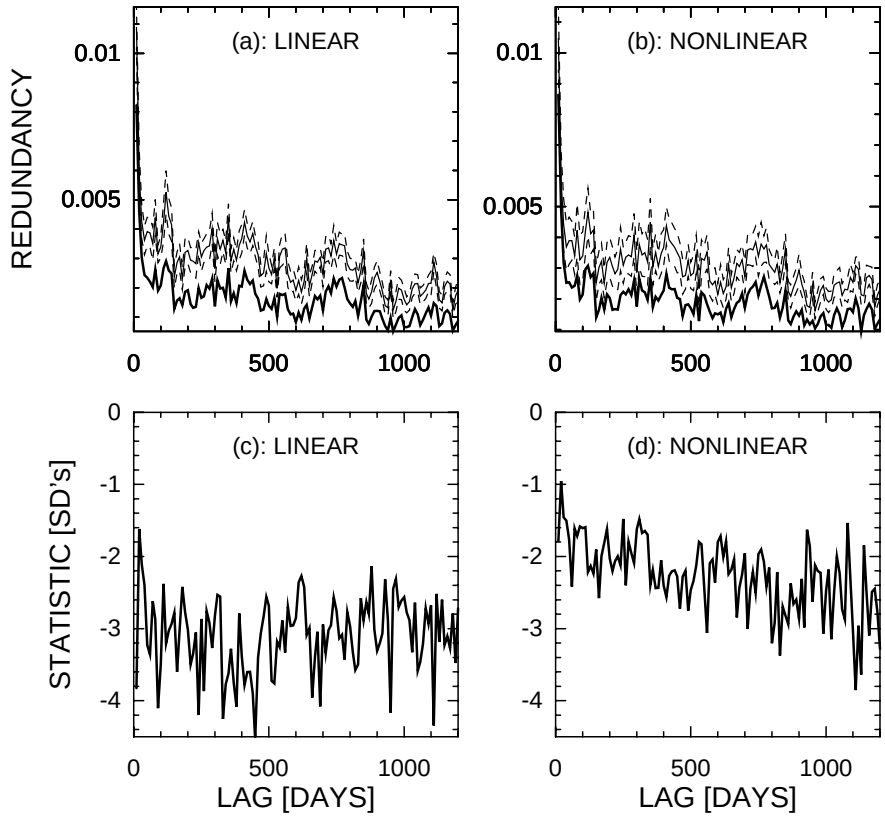

Figure 3: (a): Linear redundancy $L(y(t); y(t+\tau))$ (solid line), (b): nonlinear (general) redundancy $R(y(t); y(t+\tau))$ (solid line), for a series of differences from the long term averages of the surface air pressure, rescaled in order to have a constant variance, and for its FT surrogates (thin solid and dashed lines present mean and mean±SD, respectively, of a set of 30 surrogate realizations); (c): linear (L-based), and d): nonlinear (R-based) statistics; as functions of the time lag τ, measured in days.

more than a century. The now well-know 11-year cycle was reported already in [Wolf, 1852]. Of course, the sunspot cycle is not strictly periodic, but fluctuations in its amplitude as well as in its frequency (i.e., in the cycle duration) occur. Therefore researchers have turned towards stochastic models in order to make predictions of a future behavior of the sunspot cycle (see [Withbroe, 1989] and references therein). On the other hand, development in nonlinear dynamics and theory of deterministic chaos, namely methods and algorithms for analysis and prediction of (potentially) nonlinear and chaotic time series have naturally found their way into the analyses of the sunspot series. Several authors ([Mundt et al., 1991, Kremliovsky, 1994] and references therein) have claimed an evidence for a deterministic chaotic origin of the sunspot cycle, based on estimations of correlation dimension, Lyapunov exponents and an increase of a prediction error with a prediction horizon. The dimensional algorithms, however, have been found unreliable when applied to relatively short ex-

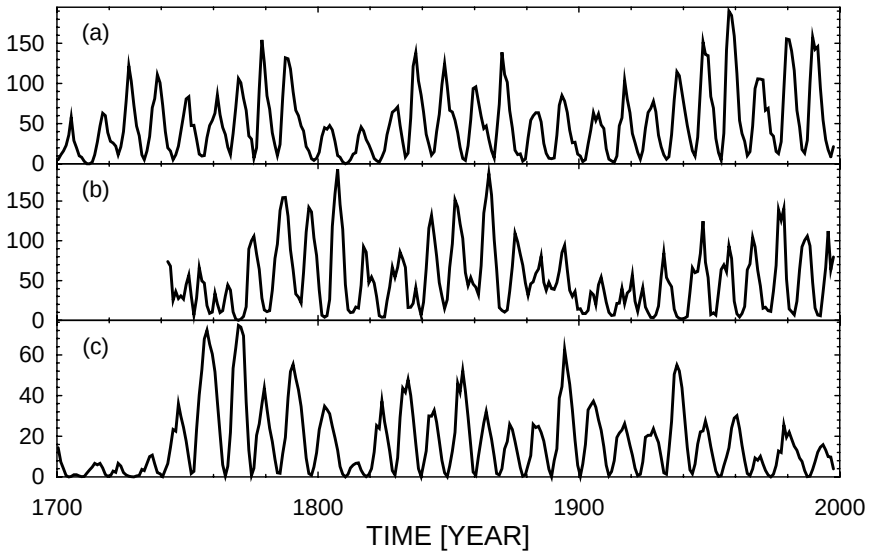

Figure 4: (a) The series of yearly sunspot numbers (1700 – 1997). (b) A realization of the HAFT surrogate data for the "last" 256 samples. (c) A 298-sample realization of the Barnes model.

perimental data, and properties consistent with stochastic processes (colored noises) such as autocorrelations can lead to a spurious convergence of dimensional estimates [Theiler, 1986]. Similar behavior has been observed also for Lyapunov exponent estimators [Dämmig and Mitschke, 1993, Paluš, 1998]. And the increase of a prediction error with an increasing prediction horizon is not a property exclusive for chaos, but it can also be observed in systems with a deterministic skeleton and an intrinsic stochastic component ("dynamical noise"). Therefore, such results cannot be considered as a convincing evidence for a nonlinear dynamical origin of the sunspot cycle.

In this section we present results of the test for nonlinearity (Sec. 2) applied to the sunspot data. The series of yearly sunspot numbers from the period 1700 – 1997 was obtained from the Sunspot Index Data Center[4], and is illustrated in Fig. 4a. As test results, here we do not present differences from surrogate mean in number of standard deviations as in the previous section, but using 15,000 surrogate replications we estimate the 2.5th and 97.5th percentiles of distributions of redundancies computed from surrogate data. If values of related redundancy obtained from the studied data lies outside the range given by these two percentile values, the null hypothesis is rejected since the test is significant with $p < 0.05$. It means that the probability of the null hypothesis (that the studied data can be explained by the surrogate model) is less than 5%.

The linear redundancy $L(y(t); y(t+\tau))$ (solid line in Fig. 5a) lies clearly inside the 2.5th and the 97.5th percentiles of the $L(y(t); y(t+\tau))$ HAFT surrogate distribution (dashed lines in Fig. 5a). This test is just a check of quality of the surrogate data

[4]Sunspot Index Data Center, Royal Observatory of Belgium, Av. Circulaire, 3, B-1180 Brussels, Internet address: `http://www.oma.be/KSB-ORB/SIDC`, file `yearssn.dat`.

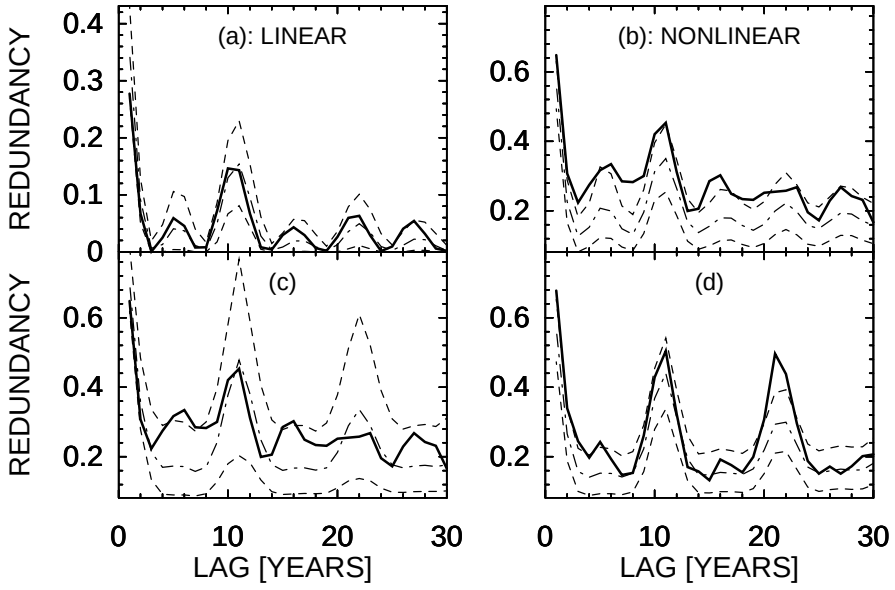

Figure 5: (a): Linear redundancy $L(y(t); y(t+\tau))$ (solid line) for the series of sunspot numbers as a function of the lag τ; linear redundancy $L(y(t); y(t+\tau))$ for the related HAFT surrogate set: dash-and-dotted line presents the mean, dashed lines illustrate the 2.5th and the 97.5th (bottom and top line, respectively) percentile of the surrogate $L(y(t); y(t+\tau))$ distribution. (b): Nonlinear (general) redundancy $R(y(t); y(t+\tau))$ (solid line), for the series of sunspot numbers as a function of the lag τ; nonlinear redundancy $R(y(t); y(t+\tau))$ for the related HAFT surrogate set, the same line codes as in (a). (c): Nonlinear redundancy $R(y(t); y(t+\tau))$ (solid line), for the series of sunspot numbers as a function of the lag τ; nonlinear redundancy $R(y(t); y(t+\tau))$ for the Barnes model surrogate set, the same line codes as in (a). (d): Nonlinear redundancy $R(y(t); y(t+\tau))$ (solid line), for a realization of the Barnes model as a function of the lag τ; nonlinear redundancy $R(y(t); y(t+\tau))$ for the related HAFT surrogate set, the same line codes as in (a).

and it says that the linear properties (dependence structures) in the sunspot data do not differ from those of the HAFT surrogates (a realization of which is presented in Fig. 4b), so that the surrogates should not be a source of a spurious detection of nonlinearity.

The nonlinearity test itself is presented in Fig. 5b, where the (nonlinear) redundancy $R(y(t); y(t+\tau))$ (solid line in Fig. 5b) is, for majority of studied lags, higher than the 97.5th percentile of the HAFT surrogate distribution (the upper dashed line in Fig. 5b). Thus the null hypothesis of a linear stochastic process (1), possibly passed through a static nonlinear transformation, is rejected.

Does this rejection really mean that a nonlinear dynamical system such as (2) or (4) underlies the sunspot cycle, or can this rejection be explained by any of the reasons listed in Sec. 3 ? This question is hard to answer. For instance, we cannot evaluate properties of innovations (model residuals) without a-priory knowledge of a valid model. And no direct connection of the sunspot data to any

physically inspired nonlinear dynamical model is known. Physical models trying to explain the variation of the solar activity come from the dynamo theory (cf. [Krause and Raedler, 1980, Proctor and Gilbert, 1994]). The principle of such a self-exciting dynamo is that the magnetic field is amplified and maintained by the interaction of mainly three types of hydrodynamic plasma motions, namely differential rotation, turbulent convection and helicity. It is interesting to mention that there are some rather simple conceptual dynamo models which show a rich dynamical behavior and can explain several facts known from observations [Feudel et al., 1993]. Such models, however, are not fitted directly to the experimental data, and are evaluated only in a qualitative way. Therefore, we should also consider concurrent linear stochastic models, such as the Barnes model [Barnes et al., 1980]:

$$z_i = \alpha_1 z_{i-1} + \alpha_2 z_{i-2} + a_i - \beta_1 a_{i-1} - \beta_2 a_{i-2}, \tag{10}$$

$$s_i = z_i^2 + \gamma(z_i^2 - z_{i-1}^2)^2, \tag{11}$$

where $\alpha_1 = 1.90693$, $\alpha_2 = -0.98751$, $\beta_1 = 0.78512$, $\beta_2 = -0.40662$, $\gamma = 0.03$ and a_i are IID Gaussian random variables with zero mean and standard deviation SD=0.4.

The Barnes model (10, 11) is a rather simple but efficient model to mimic essential properties of the sunspot numbers. It incorporates the structure of an autoregressive moving average ARMA(2,2) model (10) with a nonlinear transformation (11) which ensures that the generated series remains asymmetric and positive and tends to increase more rapidly than it decreases. Moreover, the stochastic Barnes model can mimic some seemingly nonlinear properties such as behavior of correlation integrals [Weiss, 1990] and phase portraits [Smith, 1997] obtained from the sunspot series.

We can evaluate, in the sense of the above nonlinearity test, whether the Barnes model can explain the sunspot data, by using realizations of the Barnes model as surrogate data. The result of such a test is presented in Fig. 5c.

The nonlinear redundancy $R(y(t); y(t+\tau))$ (solid line in Fig. 5c) only slightly exceeds the 97.5th percentile of the Barnes surrogate distribution in 3 from the 30 studied lags. The rejection of the Barnes model by this test is not very convincing.

On the other hand, when we test for nonlinearity in a realization of the Barnes model (Fig. 5d), the HAFT surrogates are rejected. (The rejection is clear only in the lags 21 and 22, however, $R(y(t); y(t+\tau))$ of the tested series there exceeds the whole range of the surrogate values, i.e., $p = 0$ and the test is significant even considering the simultaneous statistical inference (see [Paluš, 1995] and references therein). This result could be expected, since the nonlinear transformation (11) is not static, therefore the HAFT surrogates are rejected.

To summarize the last two tests, the realization of the Barnes model appeared in the nonlinearity test as nonlinear, and the rejection of the Barnes model as the null hypothesis is not very convincing. Can we find any solid argument for a nonlinear dynamical origin of the sunspot cycle, or should we accept a linear stochastic explanation, such as the Barnes model?

6 Amplitude-frequency correlation in nonlinear oscillators

In this section we demonstrate a typical property of nonlinear oscillators (a class of nonlinear dynamical systems), namely the correlation between instantaneous ampli-

tude and frequency of signals (solutions) generated by such systems.

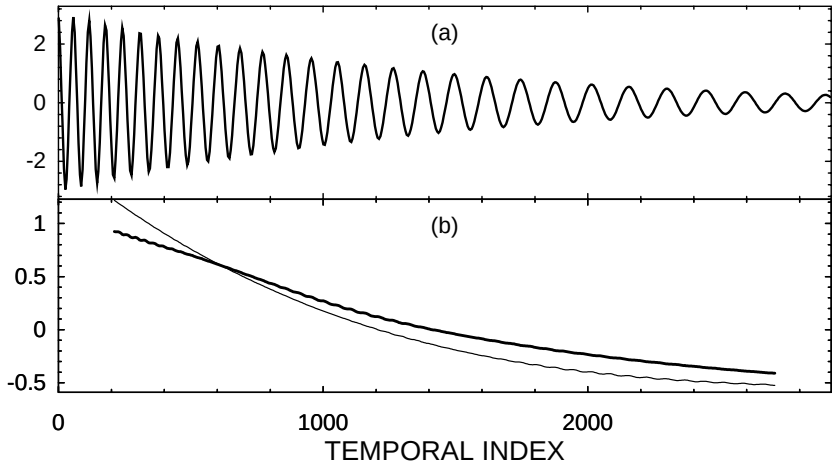

Figure 6: (a) A solution of the nonlinear Duffing oscillator without any external driving force, and (b) the related instantaneous amplitude (thick line) and frequency (thin line).

As a demonstrative example of a nonlinear oscillator (not a model for the sunspot cycle) we will consider the Duffing oscillator

$$\ddot{x} + 0.05\dot{x} + x + x^3 = F(t). \tag{12}$$

If $F(t) = 0$ and without the cubic member x^3, the equation (12) represents a damped linear oscillator with a constant frequency and an exponentially decreasing amplitude. The presence of the nonlinear (cubic) member x^3 in the equation (12) leads to a time dependent frequency, and considering again $F(t) = 0$, both the amplitude $A(t)$ and frequency $\omega(t)$ exponentially decrease and are correlated (Figs. 6a,b).

Now, consider that the nonlinear oscillator (12) is driven by a random driving force $F(t)$. In the numerical examples presented here we consider a simple random walk with a few jumps as the driving force $F(t)$.

The relation between $A(t)$ and $\omega(t)$ is a nonlinear function and may vary in time, however, the level of the correlation between $A(t)$ and $\omega(t)$ depends on the driving force: With a relatively weak driving (Fig. 7a), $A(t)$ and $\omega(t)$ are almost perfectly correlated (Fig. 7c), with a stronger driving force $F(t)$ (Fig. 8a) some differences between $A(t)$ and $\omega(t)$ emerge, however, $A(t)$ and $\omega(t)$ are still correlated (Fig. 8c).

The above presented amplitude-frequency correlation is a property which can be tested in experimental signals, even in scalar cases (single measured time series). The instantaneous amplitude and phase of a signal $s(t)$ can be determined by using the analytic signal concept of Gabor [Gabor, 1946], recently introduced into the field of nonlinear dynamics within the context of chaotic synchronization by [Rosenblum et al., 1996]. The analytic signal $\psi(t)$ is a complex function of time defined as

$$\psi(t) = s(t) + j\hat{s}(t) = A(t)e^{j\phi(t)}, \tag{13}$$

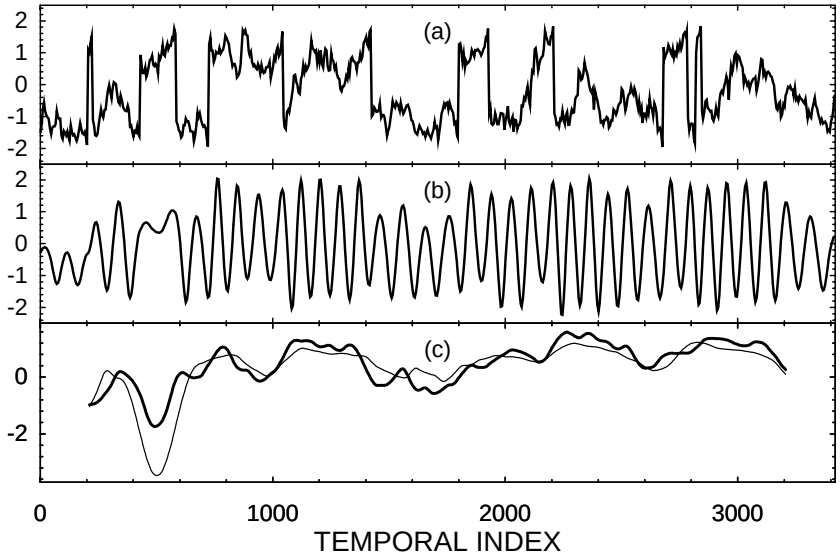

Figure 7: (a) A random driving force (a random walk with a few jumps). (b) A solution of the nonlinear Duffing oscillator with the random driving force $F(t)$ plotted in panel (a); (c) instantaneous amplitude (thick line) and frequency (thin line) extracted from the solution in panel (b).

where the function $\hat{s}(t)$ is the Hilbert transform of $s(t)$

$$\hat{s}(t) = \frac{1}{\pi} \text{ P.V.} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau. \tag{14}$$

(P.V. means that the integral is taken in the sense of the Cauchy principal value.) $A(t)$ is the instantaneous amplitude and the instantaneous phase $\phi(t)$ of the signal $s(t)$ is

$$\phi(t) = \arctan \frac{\hat{s}(t)}{s(t)}. \tag{15}$$

The instantaneous frequency $\omega(t)$ is the temporal derivative $\dot{\phi}(t)$ of the instantaneous phase $\phi(t)$.

7 Amplitude-frequency correlation in the sunspot cycle

A possible amplitude-frequency correlation (AFC thereafter) in the sunspot cycle, in particular, the importance of the amplitude in determining the length of the related cycle has already been noted in thirties by [Waldmeier, 1935] and recently discussed in [Hathaway et al., 1994]. In this section we demonstrate that the amplitude-frequency correlation found in the sunspot cycle is probably a non-random phenomenon and propose its explanation by an underlying nonlinear dynamical system.

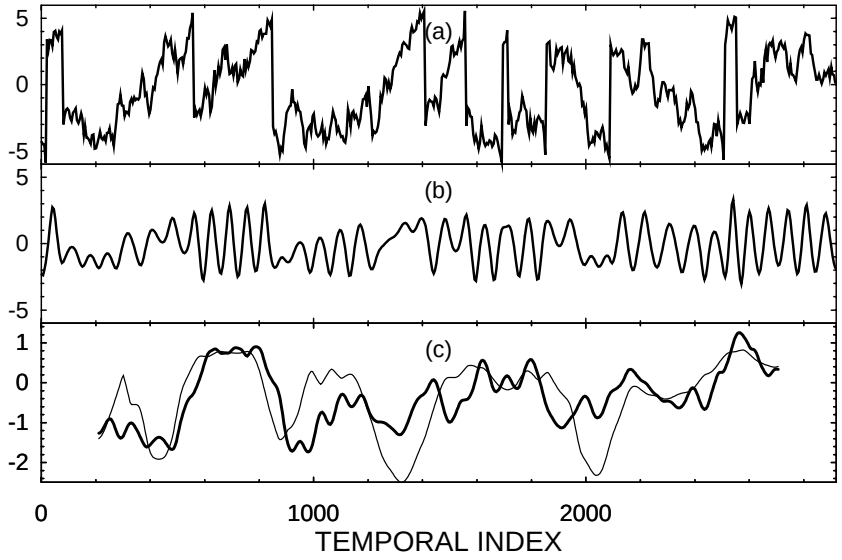

Figure 8: (a) Another example of a random driving force ("stronger", i.e., with higher amplitude than in the previous case in Fig. 7). (b) A solution of the nonlinear Duffing oscillator with the random driving force $F(t)$ plotted in panel (a); (c) instantaneous amplitude (thick line) and frequency (thin line) extracted from the solution in panel (b).

The series of yearly sunspot numbers from the period 1700 – 1997 (Fig. 4a) has been filtered by a simple moving average (MA) band-pass filter: First, the MA's from a 13-sample window have been subtracted from the data in order to remove slow processes and trends, and then a 3-sample MA smoothing has been used in order to remove high-frequency components and noise. Then the discrete version of the Hilbert transform (14) using the window length of 25 samples has been applied in order to obtain the instantaneous amplitude $A(t)$ and the instantaneous phase $\phi(t)$. For obtaining a more robust estimation of the instantaneous frequency $\omega(t)$ than the one yielded by a simple differencing the phase $\phi(t)$, the robust linear regression [Press et al., 1986] in a 7-sample moving window has been used. Finally, the series of $A(t)$ and $\omega(t)$ have been smoothed using a 13-sample MA window. The resulting series of the instantaneous amplitude and frequency of the yearly sunspot numbers, plotted in Fig. 9a, yield the crosscorrelation equal to 0.505. Does this value mean that the amplitude and frequency of the sunspot cycle are correlated as a consequence of an underlying dynamics, or could this correlation occur by chance? Searching for an answer, we test the statistical significance of this correlation using the surrogate data approach.

In the first step we use the FT and HAFT surrogates, defined in Sec. 2, where this kind of surrogates play the role of a linear stochastic process with the same spectrum and histogram as the studied data. Testing nonlinearity in general, it is stressed that the (HA)FT surrogates replicate the linear "properties" (more exactly, temporal dependences), while do not contain any nonlinear dependence structure.

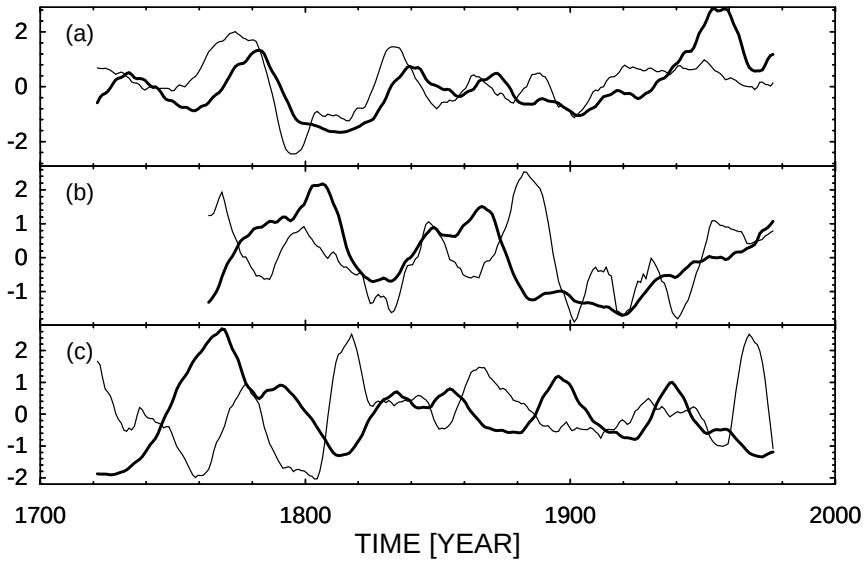

Figure 9: The instantaneous amplitude (thick line) and frequency (thin line) of (a) the yearly sunspot numbers series, (b) a realization of the related HAFT surrogate data, and (c) a realization of the Barnes model.

Here, testing the significance of AFC, we consider the (HA)FT surrogates as a data with cycles oscillating with the same frequencies as the sunspot cycles, however, not possessing any systematic amplitude-frequency correlation. Since for generating the (HA)FT surrogates we use the Fast Fourier Transform (FFT) [Press et al., 1986] which requires the number of samples equal to a power of two, we perform two tests, using the "first" and the "last" 256 samples, i.e., the subseries of the whole 298 sample series obtained by cutting away 42 samples at the end, or at the beginning, respectively, from the whole yearly sunspot numbers record. Thus, in each test, the surrogate data replicate the sample spectrum of the related 256-sample subseries.

The (HA)FT surrogates are generated from the raw (unfiltered) 256-sample segments of the sunspot data. Also, the 256-sample subseries are used for estimating the amplitude-frequency correlation related to the particular subseries, applying the procedures described above. Then, each realization of the (HA)FT surrogates, generated with respect to the raw data, undergoes the same processing as the raw data, i.e., the MA bandpass filtering, the Hilbert transform and the robust linear regression for the $\omega(t)$ estimation, and the final $A(t)$ and $\omega(t)$ smoothing are performed before computing the AFC for each surrogate realization. Then the *absolute* values of the AFC's for 150,000 surrogate realizations are evaluated in order to assess the significance of the related AFC value found in the sunspot data. The first 256-sample subseries of the sunspot yearly numbers yields the AFC equal to 0.605, while the mean value of the *absolute* AFC for the HAFT surrogate set is 0.26 with the standard deviation (SD) equal to 0.17.

In usual surrogate tests the significance is derived from the difference between the data value and the surrogate mean, divided by the surrogate SD, provided normal

distribution of the surrogate values. Having generated the large amount of the surrogate replications, here we directly estimate the p-value of the test, i.e., the probability that the assessed correlation occurred by chance (randomly) within the chosen null hypothesis (surrogate model), by simply counting the occurrences in the surrogate set of absolute AFC values greater or equal to the assessed raw data value, i.e., 0.605 in this case. The number obtained is 3637, which is equal to 2.43%. Statistically speaking, the test result is significant on $p < 0.03$, or, in other words, the probability that the amplitude-frequency correlation found in studied segment of the sunspot data occurred by chance (as a random event) is smaller than 3%.

Processing the "last" 256-sample segment of the yearly sunspot numbers, the obtained AFC is equal to 0.532, while the values from the HAFT surrogates are the same as above, however, the p-value in this case is 6.58%. Still, we can conclude that the test result is significant on $p < 0.07$. An example of the HAFT surrogate realization is plotted in Fig. 4b, its instantaneous amplitude and frequency in Fig. 9b.

The results from the tests using simple FT surrogates (i.e., without the histogram adjustment) are practically the same as those from the above HAFT surrogates. Testing the monthly sunspot numbers[5], the segments of ("first" and "last") 2048 samples were used. The same data processing has been applied as described above in the case of the yearly data with the windows lengths equivalent in real time, i.e., multiplied by 12 in number of samples. The obtained results are perfectly equivalent to those yielded by the yearly data, i.e., $p < 0.03$ and $p < 0.07$ for the "first" and the "last" 2048-sample segments, respectively.

In the second step of testing the significance of the sunspot cycle AFC, we use realizations of the Barnes model as surrogate data. As noted above, the Barnes model mimics some important properties of the sunspot data and it is hard to reject it by standard nonlinearity tests, however, realizations of the Barnes model do not possess any systematic amplitude-frequency correlation. A realization of the Barnes model is plotted in Fig. 4c, its instantaneous amplitude and frequency in Fig. 9c.

In the test, 150,000 298-sample realizations of the Barnes model have been generated and processed by the same way as the sunspot series. The mean absolute AFC is equal to 0.21, SD=0.15, comparison with the AFC obtained for the whole 298-sample yearly sunspot series (AFC=0.505) yields the p-value equal to 4.36%. Thus, considering the Barnes model, the probability that the whole yearly sunspot series AFC=0.505 occurred by chance is $p < 0.05$.

8 Discussion

Recent development in the art of identifying nonlinear dynamics underlying experimental time series has led several authors to find ways of preventing a spurious detection of nonlinearity due to flawed surrogate data. In [Paluš, 1995] we have proposed to test whether linear dependences contained in data under study were prefectly reproduced in the surrogates. Other authors have proposed sophisticated methods for constructing special constrained surrogate data. Nevertheless, we should

[5]Monthly sunspot numbers from the period 1749 – 1997 were also obtained from the Sunspot Index Data Center, `http://www.oma.be/KSB-ORB/SIDC`, file `monthssn.dat`.

be aware of the fact that a formal rejection of a linear stochastic null hypothesis does not automatically mean an evidence for nonlinear dynamical origin of studied data. Looking for a more concrete explanation of a process under study, we propose to test specific features of nonlinear dynamical systems. Such a feature, present in nonlinear oscillators, is their amplitude-frequency correlation. Using the analytic signal approach and the Hilbert transform we can estimate the instantaneous amplitude and frequency even from scalar signals and test significance of obtained amplitude-frequency correlation using a proper null model, as demonstrated here in the case of the sunspot index series.

Using two different types of stochastic models (scaled isospectral surrogates and the Barnes model) which replicate some properties of the sunspot cycle, we have obtained a statistical support for the hypothesis that the amplitude-frequency correlation observed in the sunspot cycle did not occur by chance (as a random event) but is probably a property of an underlying dynamical mechanism. Well-known systems, possessing this property, are nonlinear oscillators, in which a significant AFC can be observed also in cases of external, even random, driving force. Therefore the presented results can be considered as a statistical evidence for a nonlinear oscillator (with an external, possibly random driving force) underlying the dynamics of the sunspot cycle, unless the amplitude-frequency relation is explained by a different mechanism.

Although no particular model for the solar cycle has been proposed here, the presented statistical evidence for a nonlinear dynamical mechanism underlying the sunspot cycle can be understood as a first step in bridging the gap between reliable statistical analyses of the experimental sunspot data (dominated by linear stochastic methods) and physical models such as nonlinear dynamo models [Feudel et al., 1993, Kurths et al., 1997] (compared with data only on a qualitative level). In a next step, we can consider statistical comparison of the data with the dynamo models, with the aim to construct a realistic data-driven model for the solar cycle. Another interesting problem is identification of a force driving the sunspot cycle.

In conclusion we state that the proposed method for identification in experimental time series of amplitude-frequency correlation, a feature of nonlinear oscillators, have brought interesting information about the sunspot cycle. The method can also be applied in different fields where time series generated by possibly nonlinear processes are registered and studied.

Acknowledgements

The author would like to thank J. Kurths, D. Novotná and I. Charvátová for valuable discussions and cooperation in research of solar and climate variability. This research was supported by the Grant Agency of the Czech Republic (grant No. 205/97/0921).

References

[Barnes et al., 1980] Barnes, J., Sargent, H., and Tryon, P. (1980). Sunspot cycle simulation using random noise. In Pepin, R., Eddy, J., and Merrill, R., editors, *The Ancient Sun*, pages 159–163, New York. Pergamon Press.

[Cini Castagnoli and Provenzale, 1997] Cini Castagnoli, G. and Provenzale, A., editors (1997). *Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models*, Proc. Intl. School of Physics "Enrico Fermi", Amsterdam. IOS Press.

[Dämmig and Mitschke, 1993] Dämmig, M. and Mitschke, F. (1993). Estimation of lyapunov exponent from time series: the stochastic case. *Phys. Lett. A*, 178:385–394.

[Feudel et al., 1993] Feudel, U., Jansen, W., and Kurths, J. (1993). Tori and chaos in a nonlinear dynamo model for solar activity. *Int. J. Bif. & Chaos*, 3:131–138.

[Gabor, 1946] Gabor, D. (1946). *J. IEE London*, 93:429.

[Hathaway et al., 1994] Hathaway, D., Wilson, R., and Reichman, E. (1994). The shape of the sunspot cycle. *Solar Physics*, 151:177–190.

[Krause and Raedler, 1980] Krause, F. and Raedler, K.-H. (1980). *Mean-Field Magnetohydrodynamics and Dynamo Theory*. Berlin.

[Kremliovsky, 1994] Kremliovsky, M. (1994). Can we understand time scales in colar activity? *Solar Physics*, 151:351–370.

[Kurths et al., 1997] Kurths, J., Feudel, U., Jansen, W., Schwarz, U., and Voss, H. (1997). Solar variability: simple models and proxy data. In [Cini Castagnoli and Provenzale, 1997], pages 247–261.

[Mundt et al., 1991] Mundt, M., Maguire II, W., and Chase, R. (1991). Chaos in the sunspot cycle: analysis and prediction. *J. Geophys. Res.*, 96(A2):1705–1716.

[Paluš, 1993] Paluš, M. (1993). Identifying and quantifying chaos by using information-theoretic functionals. In [Weigend and Gershenfeld, 1993], pages 387–413.

[Paluš, 1995] Paluš, M. (1995). Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D*, 80:186–205.

[Paluš, 1998] Paluš, M. (1998). Chaotic measures and real-world systems. In Kantz, H., Kurths, J., and Mayer-Kress, G., editors, *Nonlinear Analysis of Physiological Data*, pages 49–66, Heidelberg. Springer.

[Paluš et al., 1993] Paluš, M., Albrecht, V., and Dvořák, I. (1993). Information-theoretic test for nonlinearity in time series. *Phys. Lett. A*, 175:203–209.

[Paluš and Novotná, 1994] Paluš, M. and Novotná, D. (1994). Testing for nonlinearity in weather records. *Phys.Lett. A*, 193:67–74.

[Press et al., 1986] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W. (1986). *Numerical Recipes: The Art of Scientific Computing*. Cambridge Univ. Press, Cambridge.

[Prichard and Theiler, 1995] Prichard, D. and Theiler, J. (1995). Generalized redundancies for time series analysis. *Physica D*, 84:476–493.

[Priestley, 1981] Priestley, M. (1981). *Spectral Analysis and Time Series*. Academic Press, New York.

[Proctor and Gilbert, 1994] Proctor, M. and Gilbert, A. (1994). *Lectures on Solar and Planetary Dynamos*. Cambridge Univ. Press, Cambridge.

[Rosenblum et al., 1996] Rosenblum, M., Pikovsky, A., and Kurths, J. (1996). Phase synchronization of chaotic oscillators. *Phys. Rev. Lett.*, 76:1804–1807.

[Smith, 1997] Smith, L. (1997). The maintenance of uncertainty. In [Cini Castagnoli and Provenzale, 1997], pages 177–246.

[Theiler, 1986] Theiler, J. (1986). Spurious dimension from correlation algorithm applied to time series data. *Phys. Rev. A*, 34:2427–2432.

[Theiler et al., 1992] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58:77–94.

[Theiler and Prichard, 1996] Theiler, J. and Prichard, D. (1996). Constrained-realization monte-carlo method for hypothesis testing. *Physica D*, 94:221.

[Tong, 1995] Tong, H. (1995). A personal overview of nonlinear time series analysis from a chaos perspective. *Scandinavian Journal of Statistics*, 22(2):399–445.

[Waldmeier, 1935] Waldmeier, M. (1935). *Astron. Mitt. Zürich*, 14(133):105.

[Weigend and Gershenfeld, 1993] Weigend, A. and Gershenfeld, N., editors (1993). *Time Series Prediction: Forecasting the Future and Understanding the Past*, volume XV of *Santa Fe Institute Studies in the Sciences of Complexity*, Reading, Mass. Addison–Wesley.

[Weiss, 1990] Weiss, N. (1990). *Philos. Trans. R. Soc. London, Ser. A*, 330:617.

[Withbroe, 1989] Withbroe, G. (1989). *Spacecraft*, 26:394.

[Wolf, 1852] Wolf, R. (1852). *Acad. Sci. Comp. Rend.*, 35:704.