PHYSICAL REVIEW E **66**, 025201(R) (2002)

Statically transformed autoregressive process and surrogate data test for nonlinearity

D. Kugiumtzis*

Department of Mathematical and Physical Sciences, Polytechnic School, Aristotle University of Thessaloniki, Thessaloniki 54006, Greece

(Received 15 October 2001; published 23 August 2002)

The key feature for the successful implementation of the surrogate data test for nonlinearity on a scalar time series is the generation of surrogate data that represent exactly the null hypothesis (statically transformed normal stochastic process), i.e., they possess the sample autocorrelation and amplitude distribution of the given data. A conceptual approach and algorithm for the generation of surrogate data is proposed, called the *statically transformed autoregressive process* (STAP). It identifies a normal autoregressive process and a monotonic static transform, so that the transformed realizations of this process fulfill exactly both conditions and do not suffer from bias in autocorrelation as the surrogate data generated by other algorithms. The appropriateness of STAP is demonstrated with simulated and real world data.

DOI: 10.1103/PhysRevE.66.025201 PACS number(s): 05.45.Tp, 05.10.Ln

The surrogate data test for nonlinearity has been widely used in real applications in order to establish statistically the existence of nonlinear dynamics and justify the use of nonlinear tools in the analysis of time series [1,2]. The most general null hypothesis H_0 of this test so far is that the examined time series $\mathbf{x}=[x_1,\ldots,x_n]'$ is a realization of a normal (linear) process $\mathbf{s}=[s_1,\ldots,s_n]'$ undergoing a possibly nonlinear static transform h, i.e., $x_i=h(s_i)$, $i=1,\ldots,n$. To test the null hypothesis, an estimate q from a nonlinear method applied to the original data, q_0, is compared to the estimates $q_1,\ldots,q_M$ on M surrogate time series representing the null hypothesis [3,4]. A properly designed surrogate time series $\mathbf{z}$ should possess the same autocorrelation as the original data, $r_z(\tau)=r_x(\tau)$, for a range of lags τ, and the same amplitude distribution, $F_z(z_i)=F_x(x_i)$ [$F_x(x_i)$ is the cumulative density function (CDF) of x_i], and be otherwise random.

It has been reported that erroneous results are likely to occur mainly due to the insufficiency of the algorithms to generate surrogate data that preserve the original linear correlations [5]. The prominent algorithm of the amplitude-adjusted Fourier transform (AAFT) [6], used in most applications so far, is built based on the assumption of monotonicity of h. When h is not monotonic, the AAFT algorithm is found to favor the rejection of H_0 due to the mismatch of the original linear correlations [7]. The iterated AAFT (IAAFT) algorithm improves the match of the autocorrelation of AAFT [8], but with about the same accuracy for all the surrogates, so that the small variance, combined with the small bias, may be another cause for false rejections in some cases [7]. Another algorithm making use of simulation annealing seems to perform similarly to IAAFT [9]. Recently, a correction of AAFT (CAAFT) that results in an unbiased match of the linear correlations was proposed in [10].

In this paper, we develop further the somehow profound rationale of the CAAFT algorithm and formulate a conceptual approach for the generation of surrogate data consistent with H_0, which we then solve analytically. The new algorithm, called STAP, generates the surrogate data as realizations of a suitable *statically transformed autoregressive process* (STAP), i.e., the process under H_0 is designed as a static transform of a suitable normal process.

The main idea behind the STAP algorithm is that for *any* stationary process $\mathbf{X}=[X_1,X_2,\ldots]'$, for which we have finite scalar measurements $\mathbf{x}$, there is a scalar *linear* stochastic process $\mathbf{Z}$ with the same autocorrelation ρ_X and marginal CDF Φ_X as for the observed process [$\rho_Z(\tau)=\rho_X(\tau)$ and $\Phi_Z(z_i)=\Phi_X(x_i)$], i.e., $\mathbf{Z}$ is a scalar "linear copy" of the observed $\mathbf{X}$. The objective is to derive $\mathbf{Z}$ through a static monotonic transform g on a scalar normal process $\mathbf{U}$ with a proper autocorrelation ρ_U, i.e., $Z_i=g(U_i)$. Thus g and ρ_U have to be properly selected, so that $\mathbf{Z}$ has the desired properties. In practice, the surrogate data set is a finite realization $\mathbf{z}=[z_1,\ldots,z_n]'$ of the process $\mathbf{Z}$ and g and ρ_U are estimated based solely on $\mathbf{x}$. Note that g and $\mathbf{U}$ are in general different from h and $\mathbf{S}$, respectively, of H_0 (they are the same if h is monotonic [10]). Thus with this approach H_0 can be formulated more generally, i.e., the time series is generated by a linear stochastic process.

Let Φ_0 be the marginal CDF of a standard normal process $\mathbf{U}$. A suitable choice for g, so that $\Phi_Z(z_i)=\Phi_X(x_i)$, is defined as [11]

$$Z_i=g(U_i)=\Phi_X^{-1}(\Phi_0(U_i)), \tag{1}$$

where g is monotonic by construction. Assuming that $\Phi_X(x_i)$ is continuous and strictly increasing and that $-1<\rho_U<1$, which are both true for all practical purposes, there is a function ϕ depending on g, such that $\rho_Z=\phi(\rho_U)$ for any lag τ [12,13]. If g has an analytic form, then it may be possible to find an analytic expression for ϕ as well. In that case, given that ρ_X is known and by setting $\rho_Z\equiv\rho_X$, one can invert ϕ to find $\rho_U=\phi^{-1}(\rho_X)$, if ϕ^{-1} exists.

In general, the function g, as defined in Eq. (1), does not have an analytic form because Φ_X is not known analytically, but it can be approximated by an analytic function, e.g., a polynomial p_m of degree m,

$$Z_i=g(U_i)\simeq p_m(U_i)=a_0+\sum_{j=1}^{m}a_jU_i^j. \tag{2}$$

---

*Electronic address: dkugiu@gen.auth.gr;
URL: http://users.auth.gr/dkugiu

1063-651X/2002/66(2)/025201(4)/$20.00

Low degree polynomials have been used to approximate such transforms [14,15]. Then using the definition for the autocorrelation, the approximate expression for ϕ reads

$$\rho_X=\phi(\rho_U)=\frac{\sum_{s=1}^{m}\sum_{t=1}^{m} a_s a_t(\mu_{s,t}-\mu_s\mu_t)}{\sum_{s=1}^{m}\sum_{t=1}^{m} a_s a_t(\mu_{s+t}-\mu_s\mu_t)}, \tag{3}$$

where an arbitrary lag τ is implied as argument for the autocorrelations, μ_s is the sth central moment of U_i being $\mu_{2k+1}=0$, $\mu_{2k}=1\times 3\times \ldots \times(2k-1)$, $k\geqslant 0$, and $\mu_{s,t}$ is the sth-tth central joint moment of the bivariate standard normal distribution of $(U_i,U_{i-\tau})$, defined as follows [16]:

$$\mu_{2k,2l}=\frac{(2k)!(2l)!}{2^{k+l}}\sum_{j=0}^{\nu}\frac{(2\rho_U)^{2j}}{(k-j)!(l-j)!(2j)!},$$

$$\mu_{2k+1,2l+1}=\frac{(2k+1)!(2l+1)!}{2^{k+l+1}}\times\sum_{j=0}^{\nu}\frac{(2\rho_U)^{2j+1}}{(k-j)!(l-j)!(2j+1)!},$$

$$\mu_{i,j}=0 \quad \text{if} \quad k+l=\text{odd},$$

where $\nu=\min(k,l)$. By substituting the expression for the moments in Eq. (3), the expression for ϕ can be brought to a polynomial form of the same order m,

$$\rho_X=\phi(\rho_U)=\sum_{j=1}^{m} c_j\rho_U^j, \tag{4}$$

where the vector of coefficients $\mathbf{c}=[c_1,\ldots,c_m]'$ is expressed only in terms of $\mathbf{a}=[a_1,\ldots,a_m]'$ (the expressions are rather involved and therefore not presented here). Simpler expressions can be derived using the Tchebycheff-Hermite polynomials [17]. Thus an analytic expression for ρ_U is possible if Eq. (4) can be solved with respect to ρ_U. From our simulations, we conjecture that if g is monotonic, then ϕ is also monotonic in $[-1,1]$. Then ϕ^{-1} exists and a unique solution for ρ_U can be found from Eq. (4). The proper standard normal process $\mathbf{U}$ is completely defined by ρ_U, and applying the transform g of Eq. (1) to the components of $\mathbf{U}$, the "linear copy" $\mathbf{Z}$ of the given process $\mathbf{X}$ is constructed.

Note that the solution for ρ_U is given analytically from the polynomial approximation of g and it requires only the knowledge of the coefficients $\mathbf{a}$ of the polynomial and the autocorrelation ρ_X.

In practice, we operate with a single time series $\mathbf{x}$ rather than a process $\mathbf{X}$ and with the sample estimates F_x and r_x for Φ_X and ρ_X, respectively. The steps of the algorithm are as follows:

(i) Estimate the vector of coefficients $\mathbf{a}=[a_1,\ldots,a_m]'$ of the polynomial p_m from the graph of $x_i=F_x^{-1}(F_0(w_i))$, i.e., the graph of $\mathbf{x}$ versus $\mathbf{w}$ after their ranks are matched, where $\mathbf{w}=[w_1,\ldots,w_n]'$ is standard white normal noise.

(ii) Compute $\mathbf{c}=[c_1,\ldots,c_m]'$ for the given $\mathbf{a}$ from Eqs. (3) and (4).

(iii) Find r_u from Eq. (4) for the given $\mathbf{c}$ and r_x using the sample estimates r instead of ρ. The common practice is that the solution exists and it is unique. If this is not the case, repeat the steps (i)–(iii) for a new $\mathbf{w}$ until a unique solution is obtained.

(iv) Generate a realization $\mathbf{u}$ of a standard normal process with autocorrelation r_u. We choose to do this simply by means of an autoregressive model of some order p, AR(p). The parameters $\mathbf{b}=[b_0,b_1,\ldots,b_p]'$ of AR(p) are found from r_u using the normal equations solved effectively by the Levinson algorithm [18]. The AR(p) model is run to generate $\mathbf{u}$,

$$u_{i+1}=b_0+\sum_{j=1}^{p} b_j u_{i-(j-1)}+e_i, \quad e_i\sim N(0,1).$$

(v) Transform $\mathbf{u}$ to $\mathbf{z}$ by reordering $\mathbf{x}$ to match the rank order of $\mathbf{u}$, i.e., $z_i=F_x^{-1}(F_0(u_i))$.

Note that $\mathbf{u}$ possesses the sample normal marginal CDF F_0 and the proper r_u, so that $\mathbf{z}$ possesses $F_z=F_x$, $r_z=r_x$, and is otherwise random, as desired. In practice, however, the equality $r_z=r_x$ is not exact and r_z may vary substantially around r_x. Two possible reasons for this are the insufficient approximation of g in step (i) and the inevitable variation of the sample autocorrelation of the generated $\mathbf{u}$ in step (iv), which decreases with the increase of data size. The former is due to the limited power of polynomials in approximating monotonic functions, and this shortcoming causes also occasional repetitions of the first steps of the algorithm as stated in step (iii) [20]. The latter constitutes an inherent property of the so-called "typical realisation" approach (i.e., a model is used to generate the surrogate data) and cannot be controlled. However, less variation in the autocorrelation is achieved when the AR(p) model is optimized making the following steps, in the same way as for the CAAFT algorithm [10].

(i) Apply the algorithm presented above K times and get $\mathbf{z}^1,\ldots,\mathbf{z}^K$ surrogate time series.

(ii) Compute $r_{z^1},\ldots,r_{z^K}$ and find the one r_{z^l} closest to r_x [21].

(iii) Use the parameters $\mathbf{b}$ of the l repetition to generate the M surrogate data [steps (iv) and (v) of the algorithm above].

The K repetitions above as well as the occasional repetitions of steps (i)–(iii) of the first part of the algorithm may slow down the algorithm if the time series is long, but they have no impact on the principal function of the algorithm. Simply, some realizations of white noise $\mathbf{w}$ are discarded in the search of the parameters $\mathbf{b}$ of the most suitable AR model that generates the surrogate data (through the g transform).

The free parameters of the STAP algorithm are the degree m of the polynomial approximation of g, the order p of the AR model, the number K of repetitions for the optimization of AR(p), and the maximum lag $\tau_{\max}$, used to compare $r_{z^1},\ldots,r_{z^K}$ to r_x. Usually, a small m ($m\leqslant 10$) is sufficient. For p, there is no optimal range of values but it may vary with the shape of r_x, e.g., a slowly decaying r_x may be better modeled by a larger p. In all our simulations, we set $K=M=40$ and $\tau_{\max}=p$.

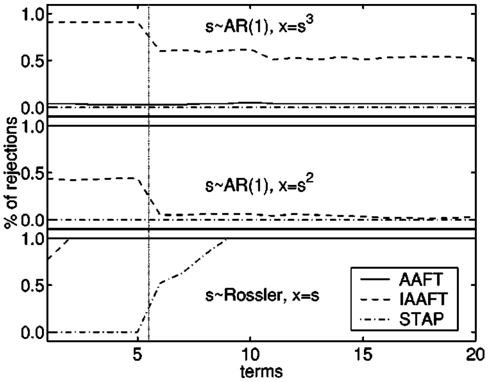

FIG. 1. Percentages of rejections of H_0 using as discriminating statistics the fit of Volterra polynomials from 100 realizations for each of the three cases in the three panels, as indicated. Three types of surrogates are used in each test as shown in the legend (for STAP $m=5$, $p=5$). The vertical gray line distinguishes the linear from the nonlinear statistics.

The proper performance and superiority of STAP over AAFT and IAAFT were confirmed from simulations on different toy models. CAAFT was found to perform very similarly to STAP. We show in Fig. 1 comparative results for AAFT, IAAFT, and STAP for three representative synthetic systems: the cube of an AR(1) process $[s_{i+1}=0.3+0.8s_i+e_i, e_i\sim N(0,1), x_i=s_i^3]$, the square of the same AR(1) ($x_i=s_i^2$), both being consistent with H_0, and the x variable of the Rössler system [19], not consistent with H_0. For each system, we generate 100 time series of 2048 samples each and for each realization we generate $M=40$ surrogate data of each type.

As discriminating statistics q^i we choose the correlation coefficient (CC) of the fit with the series of Volterra polynomials of degree 2 and order $v=5$. The polynomials for the first i terms, where $i=1,\ldots,5$, are linear and for terms $i=6,\ldots,20$ they are nonlinear (see also [10]). To quantify the discrimination we use the significance $S^i=|q_0^i-\langle q^i\rangle|/\sigma_q^i$ for each polynomial of i terms, where q_0^i is the statistic q^i on the original data, and $\langle q^i\rangle$ and σ_q^i are the mean and standard deviation of the statistic q^i on the M surrogate data. The null hypothesis H_0 is formally rejected at the 0.05 significance level when $S>1.96$, under the assumption of normality for the statistic q, which turns out to hold in general. The percentages of rejections for each of the three systems are shown in Fig. 1. Very similar results were found when deciding the rejection from the rank ordering of $q_0^i, q_1^i, \ldots, q_M^i$.

For the linear statistics, STAP gives consistently and correctly no rejections, i.e., an unbiased match of the linear correlations, whereas AAFT and IAAFT give very large percentages of rejections for all but the first case where AAFT gives about 5% rejections, as expected. For IAAFT, the rejections occur because σ_q^i is very small (10 to 20 times smaller than for AAFT and STAP), though the bias $q_0^i-\langle q^i\rangle$ is smaller than for STAP.

For the nonlinear statistics, the feature of the linear statistics persists for the first two systems (consistent with H_0) and for all three algorithms, i.e., correctly no rejections with STAP and erroneous rejections with AAFT and IAAFT. We cannot explain why the polynomial fit for the IAAFT surrogates improves with the addition of nonlinear terms (see Fig. 1). For the nonlinear system, STAP properly converges with the addition of a few nonlinear terms to the correct 100% rejection level, which AAFT and IAAFT possessed already with linear statistics.

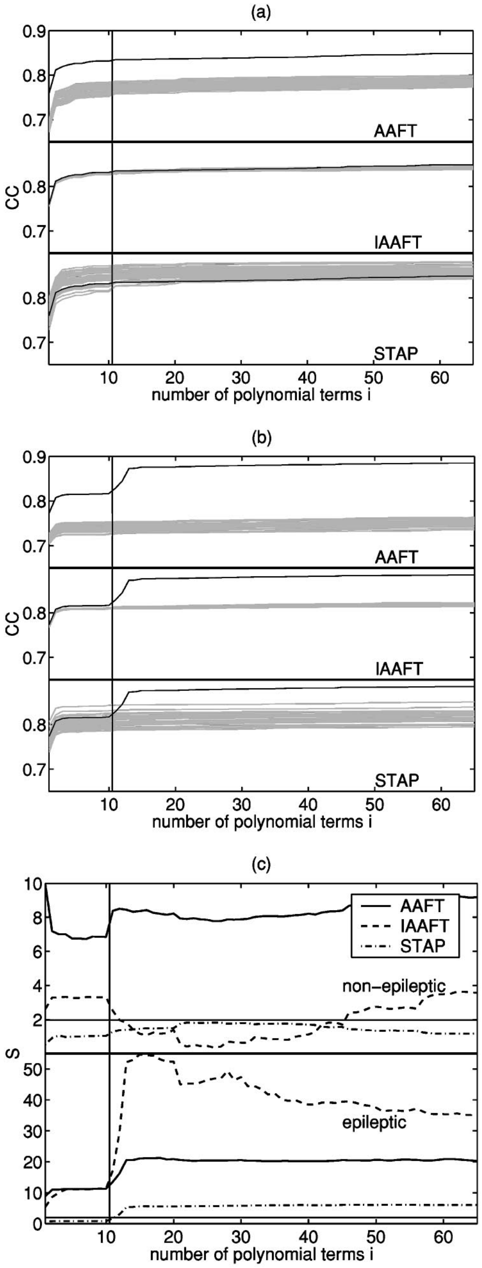

FIG. 2. (a) The correlation coefficient of the fit with Volterra polynomials on the original normal EEG data (black line) and 40 AAFT, IAAFT, and STAP surrogates (gray lines) in the three panels, as indicated (for STAP $m=10$, $p=10$). (b) The same as in (a) but for the epileptic EEG. (c) The significance for the fits in (a) (upper panel) and in (b) (lower panel) . The vertical lines in the plots distinguish the linear from the nonlinear polynomials.

Next, we verify the three algorithms on two human EEG data sets, one recorded many hours before an epileptic seizure accounting for normal brain activity and another recorded during an epileptic seizure (the sampling time is $\tau_s = 0.01$ s and the data size is $n=2048$ and 1700, respectively). The epileptic EEG seems to exhibit a pattern of oscillations indicating a more deterministic character than the nonepileptic EEG, and this constitutes a well-established result in physiology. This is demonstrated also with the fit of Volterra polynomials in Fig. 2. The fit does not improve with the inclusion of the first nonlinear terms for the nonepileptic EEG but it does for the epileptic EEG (which could as well be attributed to nonstationarity rather than nonlinear dynamics). These findings are confirmed statistically by the test with STAP surrogate data while the results of the test with AAFT and IAAFT are more or less confusing.

In particular, the H_0 on the normal EEG is erroneously rejected with AAFT because the difference in the fit between original and surrogate data is about the same for the linear and nonlinear polynomials [see Figs. 2(a) and 2(c)]. The same test result is obtained with IAAFT for large nonlinear polynomials, whereas again there is a significant difference in the linear fits between original and IAAFT surrogates (not easily discernible as both bias and variance are very small). Using STAP surrogates, the H_0 is not rejected for both linear and nonlinear fits.

For the epileptic EEG, there is again a clear difference in the linear fit between original data and AAFT surrogates and a smaller but equally significant difference between original data and IAAFT surrogates [see Figs. 2(b) and 2(c)]. The significance S for both AAFT and IAAFT increases with the addition of the first couple of nonlinear terms, much more for IAAFT due to the small variance of CC. However, the deviation in the linear fit does not support reliable rejection of H_0. On the other hand, using STAP surrogates, the H_0 is properly rejected only for the nonlinear statistics and with high confidence ($S<2$ for the linear fit and $S\simeq 5$ for the nonlinear fit).

In general, the test with STAP surrogates tends to be more conservative, i.e., small discriminations are found less significant, as the data size decreases. For example, the test on 296 sunspot samples (for which a small leap of the polynomial fit with the addition of nonlinear terms was observed) gave rejection of H_0 for AAFT and IAAFT but not for STAP (not shown here, see also [10]). However, this should not be considered as a drawback of the STAP algorithm, as one expects that the power of the test reduces with the decrease of data size.

An algorithm that generates surrogate data for the test for nonlinearity has been presented, called the statically transformed autoregressive process (STAP). The key feature of STAP is that it finds analytically the autocorrelation of an appropriate underlying normal process for the test. This is the main difference of the STAP algorithm from the corrected AAFT (CAAFT) algorithm, where the autocorrelation is estimated numerically. Both CAAFT and STAP algorithms do not suffer from the severe drawback of the AAFT algorithm, i.e., bias in the match of the original autocorrelation. The AAFT algorithm is essentially impractical for real applications because it favors the rejection of H_0 as a result of the bias in the autocorrelation. From the numerical simulations, it turns out that the IAAFT algorithm may also give small bias in the linear correlations, favoring also the rejection of H_0. On the other hand, the STAP algorithm performs properly and gives reliable rejections of H_0, only whenever this appears to be the case.

---

[1] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).

[2] C. Diks, *Nonlinear Time Series Analysis: Methods and Applications* (World Scientific, Singapore, 2000).

[3] T. Schreiber and A. Schmitz, Physica D **142**, 346 (2000).

[4] D. Kugiumtzis, in *Modelling and Forecasting Financial Data, Techniques of Nonlinear Dynamics*, edited by A. Soofi and C. Cao (Kluwer Academic Publishers, Boston, 2002), Chap. 12.

[5] D. Kugiumtzis, Int. J. Bifurcation Chaos Appl. Sci. Eng. **11**, 1881 (2001).

[6] J. Theiler, S. Eubank, A. Longtin, and B. Galdrikian, Physica D **58**, 77 (1992).

[7] D. Kugiumtzis, Phys. Rev. E **60**, 2808 (1999).

[8] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996).

[9] T. Schreiber, Phys. Rev. Lett. **80**, 2105 (1998).

[10] D. Kugiumtzis, Phys. Rev. E **62**, 25 (2000).

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (McGraw-Hill, New York, 1991), p. 113.

[12] N. Johnson and S. Kotz, *Distributions in Statistics, Continuous Multivariate Distributions* (Wiley, New York, 1972), Chap. 34, p. 15, Wiley Series in Probability and Mathematical Statistics.

[13] A. Der Kiureghian and P.-L. Liu, J. Eng. Mech. **112**, 85 (1986).

[14] C.D. Elphinstone, Commun. Stat.: Theory Meth. **12**, 161 (1983).

[15] D.M. Hawkins, Comput. Statist. **9**, 233 (1994).

[16] T.P. Hutchinson and C.D. Lai, *Continuous Bivariate Distributions Emphasising Applications* (Rumsby Scientific Publishing, Adelaide, 1990).

[17] N.M. Bhatt and P.H. Dave, J. Indian Statist. Assoc. **2**, 177 (1964).

[18] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, Springer Series in Statistics (Springer-Verlag, New York, 1991).

[19] O.E. Rössler, Phys. Lett. **57A**, 397 (1976).

[20] Other model classes may give better approximations of g and should be investigated, but polynomials were used here to derive a simple analytic form for the transform of the autocorrelations.

[21] We found that a robust way to achieve this is to compute the error $\Sigma_{i=1}^{\tau}[r_x(i)-r_{z^j}(i)]^2$ for $j=1,\ldots,K$ and $\tau = 1,\ldots,\tau_{\max}$ and then select the trial l that gives the minimum error most times, where $\tau_{\max}$ minima are totally computed, each time over the K trials.