

Measuring Nonstationarity by Analyzing the Loss of Recurrence in Dynamical Systems

Christoph Rieke,^{1,2} Karsten Sternickel,² Ralph G. Andrzejak,^{1,2} Christian E. Elger,¹ Peter David,² and Klaus Lehnertz¹

¹Department of Epileptology, Medical Center, University of Bonn, Sigmund-Freud-Strasse 25, 53105 Bonn, Germany

²Institut fuer Strahlen- und Kernphysik, University of Bonn, Nussallee 14-16, 53115 Bonn, Germany

(Received 9 January 2002; published 31 May 2002)

We propose a measure for nonstationarity which is based on the analysis of distributions of temporal distances of neighboring vectors in state space. As an extension of previous techniques our method does not require a partitioning of the time series. Moreover, the deviation of mean recurrence times from frequency distributions that would be expected under stationary conditions allows us to estimate the statistical significance of the method.

DOI: 10.1103/PhysRevLett.88.244102

PACS numbers: 05.45.Tp

Nonstationarity is a property of dynamical systems that is known from many fields of science including physics [1], engineering [2], physiology [3], and epidemics [4]. However, almost all methods of time series analysis, both linear and nonlinear, require some stationarity of the system under investigation. For a time series, the definition of *strong stationarity* often found in the literature (e.g., [5]) is constancy of all conditional probabilities in time. A deterministic system is usually regarded as stationary if the laws that govern the system remain constant in time. In most cases, however, these laws are unknown. Hence, the stationarity of the system can only be evaluated with respect to the observation time. The inference from stationarity estimated from the time series to stationarity of the dynamical system is not necessarily conclusive. However, this is the best that can be done.

Changes in dynamics during a measurement are often regarded as an undesired complication. Hence, nonstationary time series are often discarded as unsuitable for analysis. Nonstationarity, however, might actually represent an interesting aspect of the dynamics. In order to trace nonstationarity, a variety of techniques has been proposed in the framework of nonlinear dynamics. Most of these techniques are based on the relation between proximity in state space and time. The reconstruction of the state space and the identification of related states is a crucial point for these techniques. It has been shown that, for a D -dimensional deterministic system that is driven by P slowly time-dependent parameters, a time delay embedding of $m > 2(D + P)$ dimensions is sufficient to reconstruct essential aspects of determinism [6]. The identification of relationships in state space is usually carried out using graphical tools such as *recurrence plots* [7] or *space-time separation plots* [8]. Since these plots are difficult to read, efforts have been made to reduce their information content to a single number. Such methods are known as recurrence quantification techniques [9]. Other techniques include measuring the dissimilarity between density distributions [10] or the change of recurrence times in different segments of a partitioned time series [11]. Nonlinear cross prediction allows one to compare

the compatibility of nonlinear approximations to the dynamics found in different segments of the time series [12]. Almost all of these methods divide the time series into smaller segments. Often, the statistical power of these techniques is reduced due to small window sizes. In Ref. [13] a test has been proposed to detect nonstationarity in an unpartitioned time series. It is based on the deviation of the distribution of temporal distances between reference points and a fixed number of nearest neighbors in state space from a distribution that would be expected under stationary conditions. In this context, a system is regarded as stationary if the time index of a neighbor is independent from that of the reference.

In this Letter we propose, under similar assumptions as in Ref. [13], a measure of nonstationarity which allows us to determine a suitable neighborhood in a more general way. Furthermore, we consider the frequency distribution of distances in time under stationary conditions with respect to *each* reference point. For nonstationary systems we expect an increased deviation from these distributions due to the absence of distant time indices in the neighborhood of the reference. We refer to this phenomenon as *loss of recurrence*.

Let $\{x_i; i = 1, \dots, M\}$ denote an observed time series, where the physical time is related to the index i of x_i by $t = t_0 + i\Delta t$. Time delay embedding [14] in an m -dimensional state space leads to a set of vectors $V = \{\vec{x}_n; n = 1, \dots, N\}$. Let $\mathcal{U}_\varepsilon(\vec{x}_r) = \{\vec{x}_n : \|\vec{x}_n - \vec{x}_r\| \leq \varepsilon\}$ define a set of vectors in an ε neighborhood of \vec{x}_r for each reference vector $\vec{x}_r \in V$, with the number of neighbors $k = |\mathcal{U}_\varepsilon(\vec{x}_r)|$ and $\|\vec{x}_n - \vec{x}_r\| = \sqrt{\sum_{i=1}^m \frac{(x_n^i - x_r^i)^2}{m}}$. Alternatively, the neighborhood can be defined by a fixed number k of nearest neighbors. Let $l_r = \frac{1}{k} \sum_{\vec{x}_n \in \mathcal{U}} |n - r|$ be the mean time lag between \vec{x}_r and neighboring vectors. In the case of stationarity, the subspaces \mathcal{U} are revisited, which we refer to as recurrence. Assuming that all state space vectors have the same probability of recurrence, the expected value of the mean time lag is $E(l_r) = \frac{N}{2} - \frac{(r-1)(N-r)}{N-1}$. In contrast, in the case of nonstationarity the recurrence of related state space vectors is

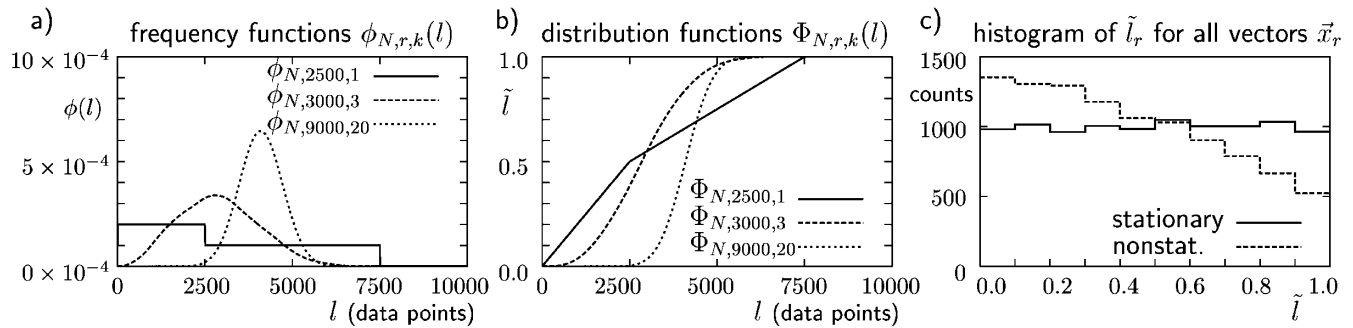


FIG. 1. Example of (a) frequency functions $\phi_{N,r,k}$ and (b) distribution functions $\Phi_{N,r,k}$ with $N = 10000$ for different parameters ($r = 2500, 3000, 9000$) with ($k = 1, 3, 20$). The histogram (c) of transformed mean time lags \tilde{l}_r collected over all \tilde{x}_r . Histograms are depicted for two time series of a stationary and a nonstationary system.

reduced. Thus, we expect that the observed mean time lag l_r is smaller than $E(l_r)$. We quantify this *loss of recurrence* as a deviation of the mean time lags l_r for all reference vectors \tilde{x}_r from the expected distributions.

Let $\phi_{N,r,k}(l)$ denote the *a priori* expected frequency distribution of the mean time lag under the assumption that for a stationary system each vector (except \tilde{x}_r itself) has the same probability to be found in the neighborhood of \tilde{x}_r . Since shorter time distances are more likely than longer ones, the frequency functions $\phi_{N,r,k}(l)$ are skewed left sided (cf. Figure 1a), and thus the probability of $l_r \leq E(l_r)$ is greater than 0.5, even for a time series under the assumption of $\phi_{N,r,k}(l)$. Furthermore, when averaged over all \tilde{x}_r , the difference of the observed mean time lag to the expected one contributes differently to the average. This complicates the quantification of the *loss of recurrence* and the specification of its significance. To solve this problem we propose the following transformation: The distribution function $\Phi_{N,r,k}(l) = \int_0^l \phi_{N,r,k}(l') dl'$ is the *a priori* probability that the observed mean time lag is less than or equal to l . For a time series under the above assumption, the transformed variable $\tilde{l}_r = \Phi_{N,r,k}(l)$ (cf. Fig. 1b) is uniformly distributed in $[0, 1]$ independent of N , r and k , and the probability of $\tilde{l}_r \leq 0.5$ is 0.5 by construction. To provide statistical validity we average over the mean time lag l_r for all reference vectors \tilde{x}_r . Figure 1c shows histograms of \tilde{l}_r for the time series of a stationary and a nonstationary system. The distribution of all l_r reflects the (non)stationarity of the system in the sense that stationarity leads to a uniform distribution, whereas for nonstationarity lower values of l_r will accumulate and therefore higher values are reduced.

Under the assumption of independent random variables \tilde{l}_r (each with a uniform frequency distribution), the significance of the median \bar{l} of the set $\{\tilde{l}_r\}$ can be obtained from the binomial distribution. We choose a one-sided test since we expect that the median \bar{l} is reduced for a nonstationary system and thus is less than 0.5. Then, the probability for a median μ to be less than or equal to \bar{l} is given by $P(\mu \leq \bar{l}) = \sum_{k=N/2}^N \binom{N}{k} \bar{l}^k (1 - \bar{l})^{N-k}$. Thus we can specify the significance level of the calculated median

\bar{l} . The transformed values, however, are not completely independent since a reference vector \tilde{x}_r is also a neighbor of its neighboring vectors \tilde{x}_n . Moreover, time lags of consecutive vectors can be correlated, in particular for oversampled sequences. Thus, the conditions of independent random variables \tilde{l}_r are not completely met but the estimated significance level is the best we can achieve.

For a numerical verification of our method we use stationary and nonstationary model systems. We apply a correction scheme by regarding $\tilde{x}_n \in \mathcal{U}(\tilde{x}_r)$ only if $|n - r| > \delta n$ (here $\delta n = \frac{N}{100}$) (cf. [15]). For oversampled sequences, it may be useful to apply additional correction schemes (cf. [16]). The *a priori* expected distribution functions $\Phi_{N,r,k}(l_r)$ are approximated numerically.

The first model is a generalization of the baker's map (cf. [12]): if $v_n \leq a: u_{n+1} = bu_n, v_{n+1} = v_n/a$; if $v_n \geq a: u_{n+1} = 0.5 + bu_n, v_{n+1} = (v_n - a)/(1 - a)$. With $a = 0.4$ we obtain a stationary system using $b = 0.5$ (B1) and two nonstationary systems by slowly varying $b = 0.4 + \frac{0.2}{M}n$ (B2) and $b = 0.2 + \frac{0.6}{M}n$ (B3), where M is the length of the time series (here $M = 100000$ data points). We record the sum $w_n = u_n + v_n$, subtract the running mean, and normalize to the running unit variance within an interval of 200 data points. As a second model we examine the Lorenz system as a dynamical flow [17]: $\frac{dx}{dt} = a(y - x), \frac{dy}{dt} = rx - y - xz, \frac{dz}{dt} = xy - bz$ with $a = 10, b = \frac{8}{3}$. For $25 \leq r \leq 90$ this system exhibits chaotic behavior (cf. [10]). We calculate $M = 100000$ data vectors (x, y, z) at fixed time intervals of $\Delta t = 0.01$ and focus on the z component $\{z_n\}$. The stationary time series is generated with $r = 25$ (L1) and the nonstationary time series by slowly varying the parameter $r = 25 + \frac{25}{M}n$ (L2) and $r = 25 + \frac{65}{M}n$ (L3).

The histograms of $\{\tilde{l}_r\}$ for the time series of the baker's map are depicted in Fig. 2. Histograms of $\{\tilde{l}_r\}$ for the time series of the Lorenz system look very similar. As expected, the distribution of $\{\tilde{l}_r\}$ for the time series B1 of the stationary system is almost uniform whereas those for the time series of the nonstationary systems show a gain

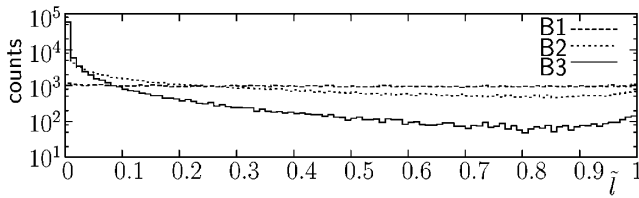


FIG. 2. Histogram for the time series of the stationary (B1) and nonstationary (B2, B3) baker's map (embedding dimension $m = 6$, hypersphere radius $\varepsilon = 0.1$).

of entries in lower sections and a loss of entries in higher sections which are caused by a *loss of recurrence* of vectors in state space.

The dependence of the median \bar{l} on ε for the time series of a stationary and a nonstationary baker's map is shown in Fig. 3. For the stationary system, the results agree with the expectation $\bar{l} = 0.5$. With increasing ε , the neighborhoods $\mathcal{U}_\varepsilon(\vec{x}_r)$ are filled up, the number of \vec{x}_r with a nonzero neighborhood increases, and the distribution of \bar{l} becomes more significant. For a sparsely filled state space, however, we also have to increase ε and thus allow more dissimilar vectors in the neighborhood. These vectors are only weakly or not at all related to the reference vector \vec{x}_r and, therefore, contribute randomly to l , which is also true for vectors of a time series under stationary conditions. Thus, $\{\bar{l}_r\}$ converges to a uniform distribution. The density of a state space depends on different properties, e.g., the length of the time series M , the investigated system, and the embedding dimension m . Thus the ε dependence of the median \bar{l} could contain relevant information for which we propose a measure l^* derived from the median \bar{l} as follows: For a time series with zero mean and unit variance we calculate the median \bar{l}_ε for 256 ε values within the range $[10^{-5}, 10]$ using a logarithmic partition. Finally we define $l^* := \{\bar{l}_\varepsilon : \max_\varepsilon |\{\vec{x}_r : |\mathcal{U}_\varepsilon(\vec{x}_r)| \leq 100\}|\}$, assuming that up to 100 neighboring vectors are sufficient to contain the relevant information. Other methods, such as weighted integration over a given ε range, are conceivable as well.

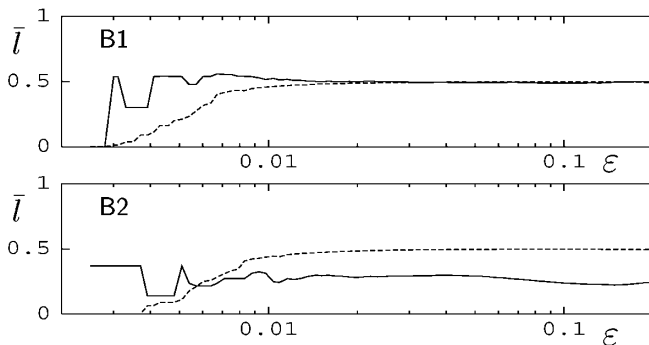


FIG. 3. Median \bar{l} (solid lines) for the time series from a stationary (B1) and a nonstationary (B2) baker's map. The dashed lines show the 0.001 quantile of the median.

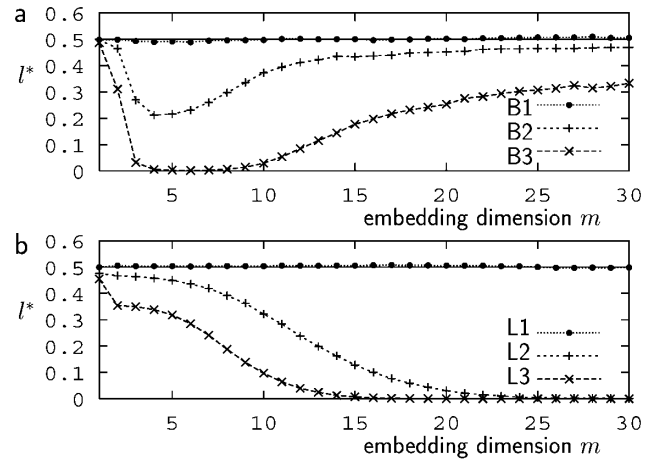


FIG. 4. l^* in dependence on the embedding dimension m for different time series of (a) the baker's map: stationary (B1) and nonstationary (B2, B3) and (b) the Lorenz system: stationary (L1) and nonstationary (L2, L3).

The dependence of l^* on the embedding dimension m is depicted in Fig. 4a for the baker's map and in Fig. 4b for the Lorenz system. An insufficient embedding dimension leads to false recurrences and thus distorts the set $\{\bar{l}_r\}$ to a uniform distribution. For stationary systems, l^* is almost independent of m . Nonstationary systems exhibit a nonmonotonous dependence of l^* on m . A local minimum of l^* can be observed at embedding dimensions coinciding with the lower bound, as suggested in [6].

A sufficient performance of the *loss of recurrence* for characterizing nonstationarity requires a sufficient number of nearby vectors in state space. This strongly depends on the observation time \mathcal{T} which should be long enough to capture all relevant time scales of a system.

In Fig. 5 we present the dependence of l^* on \mathcal{T} for the stationary and nonstationary systems. At an observation

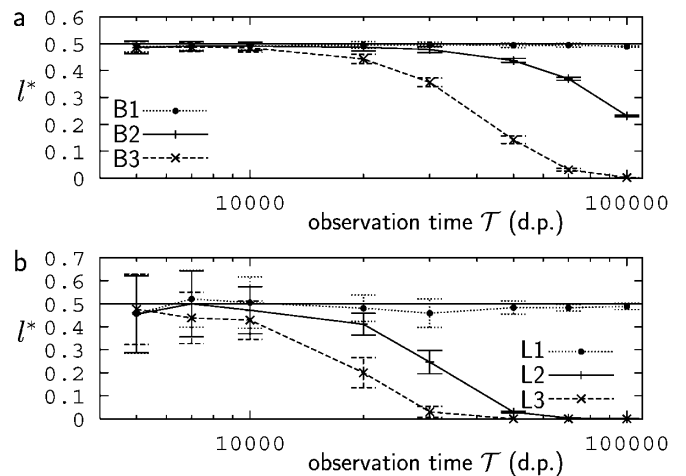


FIG. 5. Median l^* in dependence on the observation time \mathcal{T} for the time series of the baker's map (a) and the time series of the Lorenz system (b). Symbols denote the mean and bars denote the standard deviation of 100 realizations each.

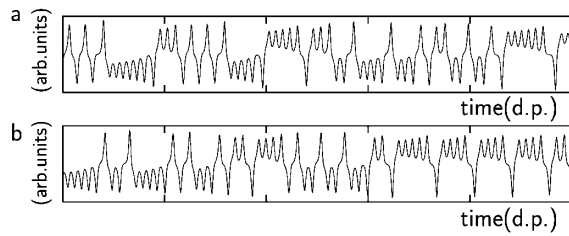


FIG. 6. Two segments [5000 data points (dp)] of the stationary Lorenz system. l^* indicates stationarity for the time series (a) and nonstationarity for (b).

time of $\mathcal{T} = 5000$, the distribution of l^* for nonstationary systems matches with the respective stationary system. Even stationary systems may lead to spurious detections of nonstationarity if the observation time is too short for a sufficient characterization of the system's dynamics (i.e., observation time is smaller than the system's characteristic time scale). When compared to the baker's map, the variability of l^* for the Lorenz systems shows a higher standard deviation due to a higher correlation of adjacent vectors. This does not match with the requirement of independent random variables of \tilde{l}_r , and thus reduces the significance of the results. Appropriate correction schemes might help to overcome this problem.

Values of $l^* > 0.5$ are due to a significant dominance of neighboring vectors with time distances larger than expected under stationary conditions. On longer time scales (e.g., $\mathcal{T} = 50\,000$ d.p.), however, the very property of the dynamics, i.e., the stationarity, is obvious.

Figure 6 shows two time series (length 5000 d.p.) of the stationary Lorenz system. Here l^* indicated stationarity for the time series depicted in Fig. 6a and nonstationarity for the time series depicted in Fig. 6b. Even visual inspection of the latter time series does not validate the stationarity of the underlying system but rather the nonstationarity as indicated by l^* . The false identification is due to an insufficient observation time. Thus the observation time is a parameter of crucial importance for the analysis of stationarity.

In conclusion we have proposed a technique which enables us to detect and quantify nonstationarity. The method can be regarded as a discriminative and independent test for stationarity and in particular can provide a powerful measure to quantify nonstationarity from an unpartitioned

time series. Application of our method to nonlinear model systems points to a crucial dependence on the observation time for the detection of nonstationarity. Thus our method provides a contribution to determine a suitable observation time, to trace characteristic time scales, and even to quantify nonstationarity in observed systems.

We thank Peter Grassberger, Thomas Kreuz, and Florian Mormann for useful comments. This work was supported by the SFB-TR3 of the Deutsche Forschungsgemeinschaft.

- [1] A. Davis, A. Marshak, W. Wiscombe, and R. Cahalan, *J. Geophys. Res. Atmos.* **99**, 8055 (1994); R. Manuca and R. Savit, *Physica (Amsterdam)* **99D**, 134 (1996); R. McGuire, *Bull. Seismol. Soc. Am.* **85**, 1275 (1995).
- [2] S. Haykin and D. J. Thomson, *Proc. IEEE* **86**, 2325 (1998).
- [3] P. Bernaola-Galvan, P. C. Ivanov, L. A. N. Amaral, and H. E. Stanley, *Phys. Rev. Lett.* **87**, 168105 (2001); J. Kaipio and P. Karjalainen, *Biol. Cybernet.* **76**, 349 (1997).
- [4] B. T. Grenfell, O. N. Bjornstad, and J. Kappey, *Nature (London)* **414**, 716 (2001).
- [5] M. Priestley *Non-linear and Non-stationary Time Series Analysis* (Academic Press, New York, 1988).
- [6] R. Hegger, H. Kantz, L. Matassini, and T. Schreiber, *Phys. Rev. Lett.* **84**, 4092 (2000).
- [7] J.-P. Eckmann, S. Ollifson Kamphorst, and D. Ruelle, *Europhys. Lett.* **4**, 973 (1987).
- [8] A. Provenzale, L. Smith, R. Vio, and G. Murante, *Physica (Amsterdam)* **58D**, 31 (1992).
- [9] M. Casdagli, *Physica (Amsterdam)* **108D**, 12 (1997); J. Zbilut, A. Giuliani, and C. Webber, *Phys. Lett. A* **246**, 122 (1998); J. B. Gao and H. Cai, *Phys. Lett. A* **270**, 75 (2000).
- [10] L. Hively, P. Gailey, and V. Protopopescu, *Phys. Lett. A* **258**, 103 (1999).
- [11] J. B. Gao, *Phys. Rev. Lett.* **83**, 3178 (1999).
- [12] T. Schreiber, *Phys. Rev. Lett.* **78**, 843 (1997).
- [13] M. B. Kennel, *Phys. Rev. E* **56**, 316 (1997).
- [14] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L.-S. Young (Springer-Verlag, Berlin, 1980), Vol. 898, pp. 366–381.
- [15] J. Theiler, *Phys. Rev. A* **34**, 2427 (1986).
- [16] J. D. Farmer and J. J. Sidorowich, in *Evolution, Learning and Cognition*, edited by Y.-C. Lee (World Scientific, Singapore, 1989), pp. 277–330.
- [17] E. Lorenz, *J. Atmos. Sci.* **20**, 130 (1963).