

Testing the null hypothesis of the non-existence of a pre-seizure state

Ralph G. Andrzejak,^{1,*} Florian Mormann,^{2,3} Thomas Kreuz,^{1,2} Christoph Rieke,^{2,3} Alexander Kraskov,¹ Christian E. Elger,² and Klaus Lehnertz²

¹*John-von-Neumann Institute for Computing, Forschungszentrum Jülich, 52425 Jülich, Germany*

²*Department of Epileptology, University of Bonn, Sigmund-Freud-Straße 25, 53105 Bonn, Germany*

³*Helmholtz Institut für Strahlen- und Kernphysik, University of Bonn, Nußallee 14-16, 53115 Bonn, Germany*

(Dated: November 14, 2002)

A rapidly growing number of studies deals with the prediction of epileptic seizures. For this purpose various techniques derived from linear and nonlinear time series analysis have been applied to the electroencephalogram (EEG) of epilepsy patients. In none of these works, however, the performance of the seizure prediction statistics is tested against a null hypothesis, an otherwise ubiquitous concept in science. In consequence, the evaluation of the reported performance values is problematic. Here we propose the technique of seizure time surrogates based on a Monte Carlo simulation to remedy this deficit.

PACS numbers: 05.45.Tp, 87.19.La, 89.20.-a

Epilepsies are characterized by recurrent and often severe malfunctions of the brain that manifest themselves as epileptic seizures. Most epilepsy patients experience the onset of a seizure as a sudden and unexpected event. Guided by both a priori and a posteriori considerations, however, it has been hypothesized that the transition to the seizure (*ictal*) state might not be an abrupt phenomenon but rather evolves via a temporally extended *pre-ictal state* (e.g. [1]). Provided that such a pre-ictal state detection could be achieved with a sufficient sensitivity and specificity, seizure anticipation and prevention technologies could be envisaged which would be of great benefit for epilepsy patients. In Refs. [2–4] it has been investigated whether information about an impending seizure can be extracted from the electroencephalogram (EEG) using different *characterizing measures* derived from linear or nonlinear time series analysis. Common to these studies is a two-step procedure: First, a characterizing measure is calculated for a multi-channel EEG using a moving-window technique. In a second step, the resulting spatiotemporal profile of the characterizing measure is analyzed by means of an often highly elaborated evaluation scheme aiming at an extraction of information specific for the pre-ictal state. As distinct and complementary as the different approaches are, in the context of the present study they will be termed *seizure prediction statistics* in their collectivity. Their output in terms of sensitivity and specificity will be denoted as *performance*.

Let us now consider the following null hypothesis: "The transition from the inter-ictal to the ictal state is an abrupt phenomenon. An intermediate pre-ictal state does not exist." Despite the fact that in this case no information predictive of impending seizures could be ex-

tracted from the EEG, many of the seizure prediction statistics would probably still render non-zero performance values. Moreover, an a priori estimation of these performance values is problematic. Hence, it is impossible to decide whether a given performance value obtained from real data indicates the existence of a pre-ictal state or whether it is consistent with the null hypothesis stated above.

A similar problem is known from the application of nonlinear time series analysis techniques to stochastic dynamics. The framework of nonlinear time series analysis comprises a number of measures that allow a characterization of nonlinear deterministic dynamics [5]. For most of these measures, however, the ranges of values obtained for nonlinear deterministic dynamics and for linear stochastic dynamics overlap substantially [6]. It is therefore impossible to decide whether a given value of a nonlinear measure calculated from some unknown time series reflects a property of an underlying nonlinear deterministic dynamics or whether it is consistent with a linear stochastic model. This ambiguity has been addressed by the method of surrogate data [7]. This method allows the testing of a specified null hypothesis about the dynamics underlying a given time series. For this technique, which can be regarded as a Monte Carlo simulation, an ensemble of surrogate time series is constructed from the original time series in such a way that they have all properties that are consistent with this null hypothesis in common with the original, but are otherwise random. A discriminating statistics, which has to be sensitive to at least one property that is not consistent with the null hypothesis, is calculated for both the original time series and the surrogates. In case the null hypothesis is the assumption of a linear stochastic process, a measure

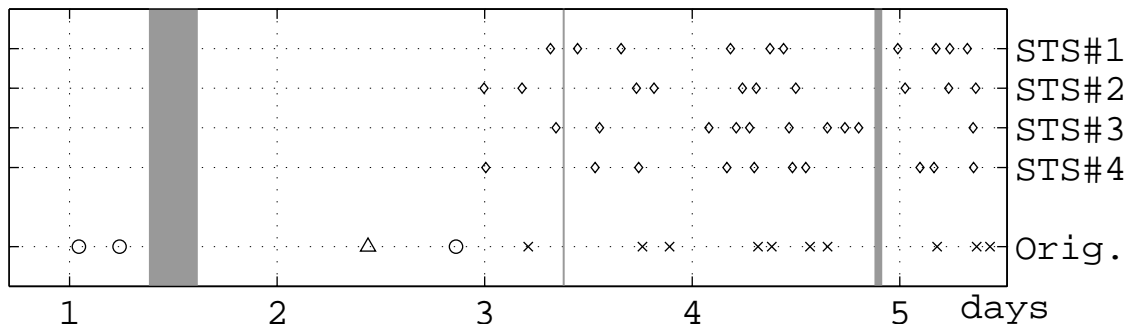


FIG. 1: Temporal distribution of relevant events that occurred during a six-day EEG recording from an epilepsy patient (lower row: \times seizures, \circ sub-clinical events, Δ hyperventilation, gray vertical bars: discontinuities). Four exemplary seizure time surrogates (STS) are displayed in the upper rows. Diamonds denote surrogate seizure onset times established by the constrained randomization procedure described in the text.

derived from nonlinear time series analysis can be used as a discriminating statistics. If the result for the original deviates from the distribution of the values obtained from the surrogates, the null hypothesis can be rejected at a level of significance determined by the number of surrogates used. In the beginning, the method of surrogates was mostly used to test the null hypothesis of a linear stochastic model and was regarded exclusively as a test for nonlinearity. Later it has been understood to be a more general and therefore also a more powerful concept. In Ref. [8] nonlinearity was even explicitly included in a null hypothesis. Furthermore, surrogate algorithms have been developed that allow the testing of almost arbitrary null hypotheses [9]. Problems associated with false positive rejections of null hypotheses have been discussed in Refs. [10]

In this Letter we propose a further generalization of the concept of surrogates by constructing *seizure time surrogates* that allow to validate the results of seizure prediction statistics. Given a continuous EEG recording, seizure time surrogates can be constructed by replacing the original seizure times with times randomly chosen from the inter-ictal intervals. Specified properties of the original sequence can be imposed as constraints on the surrogate seizure onset times. Subsequently, any given seizure prediction statistics can be carried out for both the original seizure times and the surrogates. Provided that a pre-ictal state exists and the prediction statistic is able to detect it, the statistics' performance should be highest for the original seizure times. A similar approach is used in seismology where null hypothesis tests are regarded as inevitable to evaluate the performance of earthquake prediction algorithms [11].

To illustrate this technique we analyzed the spatiotemporal distribution of a nonlinear measure that was calculated from a quasi-continuous EEG recorded over six days during the pre-surgical work-up [12] of an epilepsy patient independently from the design of our retrospective study. Using implanted electrodes equipped with

a total of 48 separate contacts the EEG was measured directly at the surface of the cortex and within deeper structures of the brain. EEG data was sampled at 200 Hz using a 16 bit analog-to-digital converter and filtered within a frequency band of 0.53–100 Hz.

Figure 1 shows a scheme of events that took place during the recording time and that have to be taken into account for the generation of surrogate seizure onset times. Twice the patient was briefly (13 and 54 min) disconnected from the EEG acquisition system. A longer discontinuity (340 min) was necessary to carry out a magnetic resonance imaging scan to determine the exact location of the implanted electrodes. All ten seizures occurred spontaneously within the second half of the recording. The latency of the first seizure can be explained by the remaining effect of anti-epileptic drugs that were withdrawn after implantation of the electrodes. During the first three days only three sub-clinical events took place, i.e. events during which seizure-like activity can be observed in the EEG while the patient does not show any clinical signs of an ongoing seizure. On the third day the patient was asked to perform a hyperventilation, a seizure provocation technique that may cause alterations of the EEG. For our study, four intervals of 20 minutes starting at the beginning of both the hyperventilation and the three sub-clinical events as well as ten intervals of one hour starting at the onset of the seizures were excluded from the analysis. The last step was carried out since the ictal and post-ictal EEG differs substantially from the EEG recorded during the inter-ictal state. Both these exclusions and the aforementioned discontinuities will be referred to as recording gaps. The remaining length of the analyzed EEG amounted to 101.1 hours.

Nineteen seizure time surrogates were generated by replacing original seizure times with times randomly chosen in the inter-ictal intervals (cf. Fig. 1). The following properties of the original seizure times were imposed as constraints on the seizure time surrogates: The total

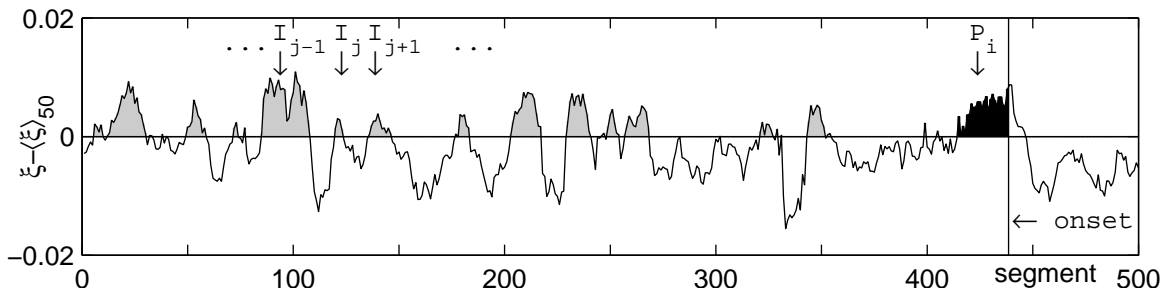


FIG. 2: Parametrization of an exemplary $\xi(t) - \langle \xi \rangle_{50}$ profile. Data from the ictal and post-ictal state were included for completeness. Gray shaded areas denote inter-ictal peaks (I_j). A pre-ictal peak (P_i) is shown in black.

number of seizures ($= 10$), the distribution of intervals between consecutive seizures, and the clustering of the seizures in the second half of the recording. The intervals between consecutive seizures and the interval from the first seizure back to an arbitrarily defined starting point T_0 at 12am on the third day are called D_1, \dots, D_{10} . For the generation of each of the seizure time surrogates the following steps were carried out: First, a new starting point was defined as $T_0^* = T_0 - \varepsilon \cdot 1\text{hour}$ with ε being a random number uniformly distributed in $[0, 4]$. Starting at T_0^* , surrogate seizure onset times T_1^*, \dots, T_{10}^* were generated from a random permutation of D_1, \dots, D_{10} . The sequence was discarded whenever a recording gap was located within the last hour prior to any of the T_1^*, \dots, T_{10}^* .

As a characterizing measure of the EEG we used the *degree of nonlinear determinism* ξ . Following Ref. [13], ξ was defined from a combination of the coarse grained flow average Λ [14] and iterative amplitude adjusted surrogates [15]. As a direct test for determinism, Λ quantifies the alignment of nearby trajectory segments in state space. Here, the use of surrogates is essential to correct for an alignment that is caused by autocorrelations rather than by deterministic dynamics. Using a moving window technique, the EEG was divided into non-overlapping segments of 20.48 s ($N = 4096$ data points). For each of these segments a set of four surrogates was generated. The dynamics were reconstructed using the method of delays [16] with a fixed embedding dimension ($m = 6$) and varying time delay τ . We defined $\xi \equiv \sum_{\tau=5}^{20} (\Lambda_{EEG} - \langle \Lambda_{SUR} \rangle)(\tau)$ with Λ_{EEG} denoting the value obtained for the EEG segment, and $\langle \Lambda_{SUR} \rangle$ denoting the mean value obtained for the surrogates. All parameters were adopted from a previous study in order to avoid any in-sample overtraining [13]. Only the number of surrogates was reduced from nine to four since the latter value was found to provide a sufficiently reliable estimate of $\langle \Lambda_{SUR} \rangle$. A $\xi(t)$ -profile was obtained for each of the 48 EEG channels for segments $t = 1, \dots, 17764$. In order to disregard short-term fluctuations and rather focus on long-term trends of $\xi(t)$, a moving-average filter of eleven consecutive segments was applied.

In Ref. [13] we have compared mean ξ values obtained

from the inter-ictal EEG recorded from within the epileptic focus and from other brain areas of epilepsy patients. A correct localization of the epileptic focus could be derived from increased values of ξ in all investigated cases. Following the basic concept of [2], we hypothesized that the pre-ictal state would be reflected in an increase of ξ (cf. [4]), and accordingly designed a simple evaluation for $\xi(t)$ (Fig. 2):

First, a reference level was defined by the median $\langle \xi \rangle_{50}$ of the distribution of all $\xi(t)$ values for each EEG channel. For every interval B between two crossings of $\xi(t)$ and $\langle \xi \rangle_{50}$ we quantify the area $A = \sum_{t \in B} (\xi(t) - \langle \xi \rangle_{50})$. The evaluation was restricted to positive areas, which we will refer to as *peaks*. Let s denote the number of seizures that were directly preceded by a peak instead of a drop of $\xi(t)$ below the reference level. For those seizures this peak is termed pre-ictal and its area is denoted by P_i for $i = 1, \dots, s$. All other peaks are termed inter-ictal and their areas are denoted by I_j for $j = 1, \dots, k$ with k being the total number of inter-ictal peaks. The P_i were only integrated up to the seizure onset times. In order to compare the distributions of P_i and I_j , we calculated $f \equiv \frac{\langle P_i \rangle_{50} - \langle I_j \rangle_{50}}{\langle P_i \rangle_{50} + \langle I_j \rangle_{50}}$ from the medians of the two distribution. Finally, we defined $F \equiv \langle f \rangle$ as the average over all channels. By construction f and F are restricted to $[-1, 1]$ and should tend to zero if the distributions of pre-ictal and inter-ictal peaks match.

Figure 3 shows the distribution function of I_j along with corresponding values of P_i determined for the original seizure onset times for one exemplary EEG channel. Among the $s = 7$ pre-ictal peaks, 5 peaks were found whose area exceeded the median area of the inter-ictal peaks. For this channel we obtained $f = 0.94$. After averaging the results over all channels, we obtained $F = 0.81$ for the original seizure times. At first glance, this value appears quite promising in the sense that it might indicate that the pre-ictal peaks were more pronounced than the inter-ictal peaks, confirming that the pre-ictal state is indeed reflected in a increase of ξ .

This interpretation, however, does not necessarily hold: Suppose we selected peaks randomly from a sequence like the one depicted in Fig. 1. If the probability

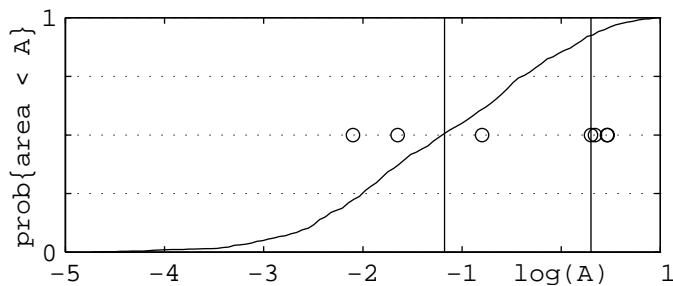


FIG. 3: Distribution function $\text{prob}\{\text{area} < A\}$ for inter-ictal peaks I_j (line) along with values P_1, \dots, P_7 (circles) obtained for the original seizure times for one exemplary channel. The two vertical lines indicate the medians $\langle I_j \rangle_{50}$ and $\langle P_i \rangle_{50}$, respectively

to be selected were the same for all peaks, scores with areas below and above the median of the distribution of areas of all peaks would be equiprobable. If we, in contrast, drew samples from such a sequence by randomly selecting points in time, we would be more likely to draw long peaks than to draw short peaks. Consequently, the area of our samples would tend to exceed the aforementioned median. Taking this into account, a value of $F > 0$ would be expected even under the assumption of our null hypothesis. Hence, interpreting the significance of the observed value of F appears quite difficult. One could consider a normalization or correction of F based on the distribution function of I_j but this might not be sufficient to eliminate any bias caused by further, unforeseen problems and pitfalls.

A more straightforward answer can be obtained from the application of seizure time surrogates. From Fig. 4 it becomes evident that the F value obtained for the original seizure times was within the distribution obtained for the seizure time surrogates. On the level of single EEG channels, i.e. based on f -values, the null hypothesis could be rejected for four of the 48 channels. However, if a test with a nominal size of $\alpha = 0.05$ is repeated 48 times there is a 9% chance to obtain up to four rejections. Hence, we could not reject the null hypothesis of the non-existence of a pre-ictal state by means of the applied seizure prediction statistics.

The fact that the null hypothesis could not be rejected does by no means prove its correctness. Rather, there exist numerous alternative explanations for this result. For several reasons the applied seizure prediction statistic might simply lack any discriminative power for the hypothesis test: Even though ξ did allow a characterization of the spatial distribution of the inter-ictal epileptic dynamics [13] it may still be incapable to detect any feature of the EEG specific for the pre-ictal state. An explanation for such a finding would be that the inter-ictal epileptic dynamics and the seizure-generating process are two distinct dynamical phenomena each imposing different features on the EEG. On the other hand, even if ξ

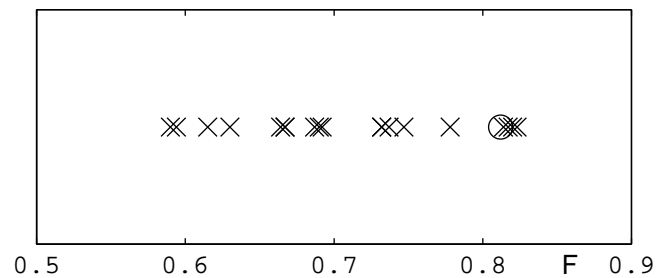


FIG. 4: F values for the original seizure times (\circ) and the distribution of nineteen seizure time surrogates (\times).

were capable to detect the pre-ictal state, the relevant information could be missed by our rather simple evaluation of $\xi(t)$. Furthermore, our study was based on the EEG recording of only one epilepsy patient. It would be highly speculative to draw any conclusions about the multi-faceted disease epilepsy from such a limited sample. It is far beyond the scope of the present study to prove or disprove the existence of a pre-ictal state. Rather, the aim was to propose a simple technique that allows to validate the performance of seizure prediction statistics. In some cases, e.g. if only a collection of very short recordings each containing one seizure is available, a randomization of seizure onset times might not be possible. In these cases, one could randomize the time course of the characterizing measure and keep the original seizure times fixed. For this purpose, the technique of constrained randomization [9] could readily be employed.

Further studies are underway that apply seizure time surrogates in combination with different seizure prediction statistics and to larger samples of EEG data to further elucidate the problem of pre-ictal state detection [17]. In this context, we expect seizure time surrogates to be a powerful tool to differentiate statistics unsuited for a detection of the pre-ictal state from more promising approaches.

We are grateful to Peter Grassberger for carefully reading this manuscript and valuable discussions. C.E.E., T.K., K.L., F.M., C.R. acknowledge support from the Deutsche Forschungsgemeinschaft.

* Electronic address: r.g.andrzejak@fz-juelich.de

- [1] S. Viglione and G. Walsh, *Electroencephalogr. Clin. Neurophysiol.* **39**, 435 (1975); Z. Rogowski, I. Gath, and E. Bental, *Biol. Cybern.* **42**, 9 (1981); J. Gotman et al., *Epilepsia* **23**, 432 (1982); A. Siegel, C. Grady, and A. Mirsky, *Epilepsia* **23**, 47 (1982); L. D. Iasemidis et al., *Brain Topogr.* **2**, 187 (1990); C. E. Elger and K. Lehnertz, in *Epileptic seizures and syndroms*, edited by P. Wolf (J. Libbey and Co, London, 1994), pp. 547–552;
- [2] K. Lehnertz and C. E. Elger, *Phys. Rev. Lett.* **80**, 5019 (1998).

- [3] J. Martinerie et al., *Nat. Med.* **4**, 1173 (1998). I. Osorio, F. G. Frei, and S. B. Wilkinson, *Epilepsia* **39**, 615 (1998). F. Mormann et al., *Physica D* **144**, 358 (2000). A. Petrosian et al., *Neurocomputing* **30**, 201 (2000). K. K. Jerger et al., *J. Clin. Neurophysiol.* **18**, 259 (2001). L. D. Iasemidis et al., *J. Combinatorial Optimization* **5**, 9 (2001). M. Le Van Quyen et al., *Lancet* **357**, 183 (2001). K. Schindler et al., *Clin. Neurophysiol.* **113**, 604 (2002). V. Navarro et al., *Brain* **125**, 640 (2002). B. Litt et al., *Neuron* **30**, 183 (2001).
- [4] R. G. Andrzejak et al., *Epilepsia (Suppl. 7)* **41**, 202 (2000).
- [5] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge Univ. Press, Cambridge, UK, 1997).
- [6] A. R. Osborne and A. Provenzale, *Phys. D* **35**, 357 (1989).
- [7] J. Theiler et al., *Physica D* **58**, 77 (1992).
- [8] J. Theiler, *Phys. Lett. A* **196**, 335 (1995).
- [9] T. Schreiber, *Phys. Rev. Lett.* **80**, 2105 (1998).
- [10] P. Rapp et al., *Phys. Lett. A* **192**, 27 (1994). J. Timmer, *Phys. Rev. E* **58**, 5153 (1998). D. Kugiumtzis, *Phys. Rev. E* **60**, 2808 (1999). P. Rapp et al., *Int. J. Bifurcation and Chaos* **11**, 983 (2001).
- [11] Y. Y. Kagan, *Geophys. J. Int.* **131**, 505 (1997).
- [12] J. Engel Jr, ed., *Surgical Treatment of the Epilepsies* (Raven Press, New York, 1993).
- [13] R. G. Andrzejak et al., *Epilepsy Res.* **44**, 129 (2001).
- [14] D. T. Kaplan and L. Glass, *Phys. Rev. Lett.* **68**, 427 (1992).
- [15] T. Schreiber and A. Schmitz, *Phys. Rev. Lett.* **77**, 635 (1996).
- [16] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L.-S. Young (Springer-Verlag, Berlin, 1980), vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381.
- [17] F. Mormann et al., Submitted (2002).