# Ordinal time series analysis

Christoph Bandt

*Institute of Mathematics, Arndt University, 17487 Greifswald, Germany*

**Abstract**

We discuss robust methods of time series analysis which use only comparisons of values and not their actual size. Local and global order structure are defined as matrices or by rank numbers. Local ranks, autocorrelation by Kendall's tau, and permutation entropy as complexity measure are introduced in such a way that they contain a scale parameter which allows to study time series on different scales.

*Key words:* time series, Kendall's tau, autocorrelation

## 1 Introduction

Many data sets in nature, in medicine and economy come in the form of time series. In ecology, these data will also become abundant with the development of monitoring equipment. Thus there is a need for methods which apply directly to real world data sets and work very fast. Classical time series methods, like Fourier spectrum and Gaussian ARMA processes, make assumptions which are rarely fulfilled by data sets from nature.

As an alternative, modifications of the classical models have been suggested: non-Gaussian ARMA, piecewise ARMA, GARCH processes, or models with fractional difference operators (FARIMA, FIGARCH) [15,14]. Fractal geometry enters time series analysis when we assume that the given time series is a noisy orbit of a chaotic dynamical system. Correlation dimensions, entropies and Lyapunov exponents are the main parameters for this approach [1], and fine-tuning of algorithms often needs expert knowledge. Fractal exponents are also determined in connection with long-range dependence.

*Email address:* bandt@uni-greifswald.de (Christoph Bandt).

The aim of this paper is to draw the reader's attention to still another alternative: simple, robust and fast methods which do not use all the structure of the time series

$$(x_1, ..., x_T) = (x_t)_{t=1...T} \ .$$

We shall not work with the actual values $x_t$ which are real numbers or integers. We shall only need to know whether $x_s < x_t$ or $x_s > x_t$ for every pair of time points $s, t$ between 1 and $T$.

It seems that such *ordinal* methods were used here and there, but, with exception of a long series of papers by Hallin and co-workers ([6,8,9], see section 4), have never been treated systematically. We have started to develop such methods together with B. Pompe, A. Groth and F. Shiha. Most of the methods are published here for the first time.

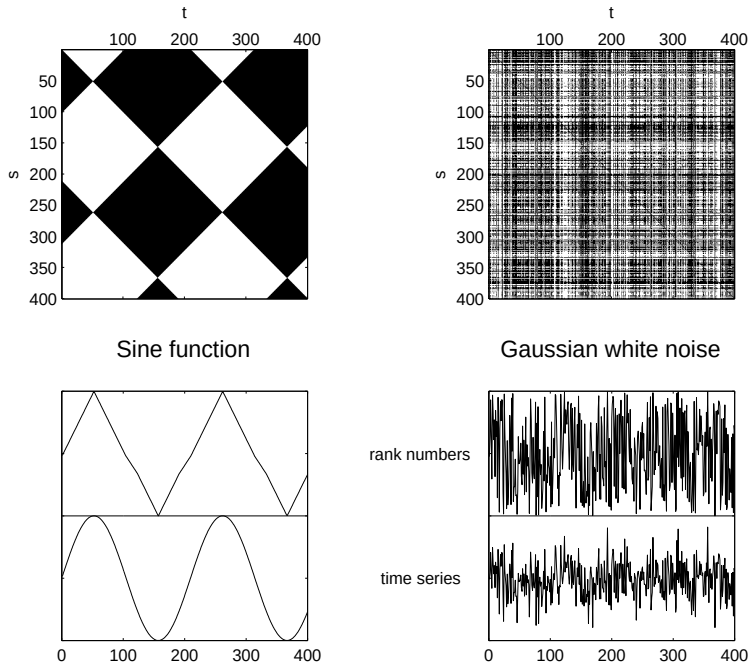


Figure 1. Order structure matrix (top), rank numbers (middle) and data (below) of a sine function and Gaussian white noise

To simplify our presentation, we shall assume that the range of values is so large that the case of ties $x_s = x_t$ is very rare and can be neglected. Our methods will not work when $x_t$ assumes only a few (say less then 10) different values. However, if $x_t$ is obtained from measurements, and there are a few ties, we can destroy the ties by adding a tiny random noise to the series. In practice, this works very well.

Note that in elementary statistics ordinal methods like the Wilcoxon test, the Kruskal-Wallis rank variance analysis, and rank correlation coefficients are very common. This is not the case for time series because without addition, we have no concept of superposition of two time series. For example, we cannot represent a given time series as a sum of signal and noise. On the other hand, we shall show now that ordinal methods have many advantages which justify their use.

Before going into details I would like to express my gratitude to M.-A. Martin who invited me to the stimulating PEDOFRACT workshop in El Barco, Spain, June 2002, where this work was presented as a lecture.

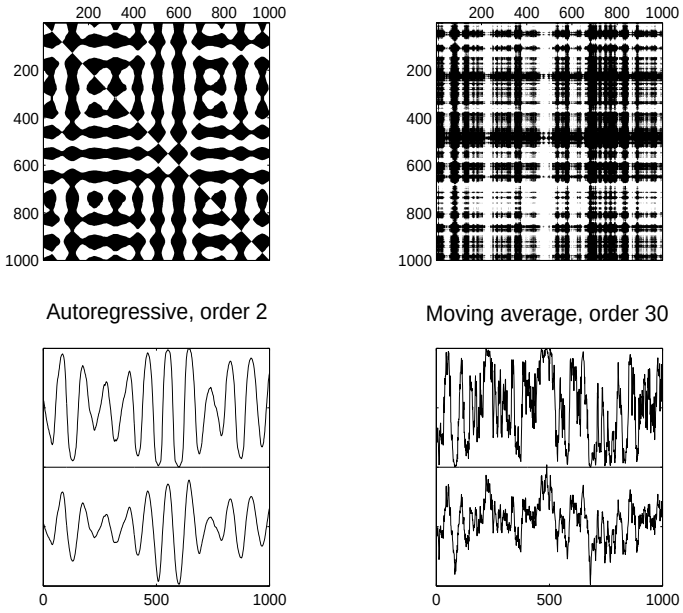


Figure 2. Left: AR(2) process $y_t = 1.97y_{t-1} - 0.98y_{t-2} + \epsilon_t$.
Right: MA(30) with equal weights. Gaussian noise.

## 2 Order structure matrix

**Definition.** First let us visualize the set of comparisons of all pairs $x_s, x_t$ with $1 \le s, t \le T$. We define the *order structure matrix*

$$B = (b_{st})_{s,t=1\ldots T} \quad \text{with} \;\; b_{st} = 1 \;\; \text{if} \;\; x_s \ge x_t \quad \text{and} \;\; b_{st} = 0 \;\; \text{if} \;\; x_s < x_t \,.$$

This is a square matrix. If we assign black color to all ones and white to all zeros, we obtain a black-and-white image as shown in figures 1-5. Similar matrices indicating (almost) equality instead of $\ge$ were introduced as *recurrence*

*plots* by Zbilut and Webber, cf. [7]. In unpublished work [13], Pompe compared different types of such structure matrices.

**Examples.** As seen in figure 1, a sine function is associated with a chessboard pattern turned around 45 degrees. White noise produces a matrix $B$ which shows no other patterns than the black and white lines which indicate very small and very large values. In figure 2 we show an AR(2) process which looks almost as smooth as the sine function, and a moving average process which is still rather noisy, in order to show how $B$ will change. Figure 3 shows on the left a spoken a: which among all speech signals is most similar to trajectories of chaotic dynamical systems [2].

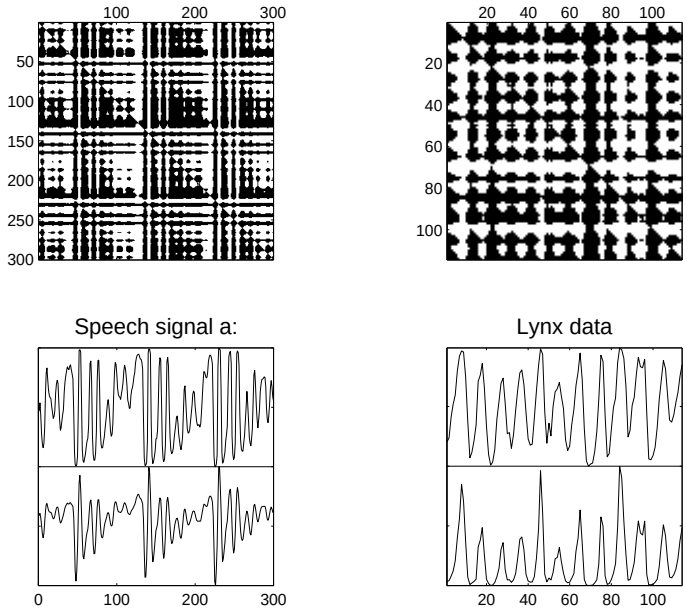


Figure 3. Real world data, from speech and ecology

On the right in figure 3 there is the main example of this paper. The classical Canadian lynx series is probably the best-known time series in Ecology. It has been studied by many authors. A comprehensive discussion with references is given by Tong [15], chapter 7.2. There are 114 values from the years 1821 to 1934. Compared to the other examples, this time series is very short. Nevertheless, our techniques will apply.

Figure 4 shows an example from the textbook of Shumway and Stoffer: 453 monthly values of the Southern Oscillation Index (SOI) which indicates the sea surface temperature of the Pacific and is related to the El Niño phenomenon, and 453 monthly values of estimated fish recruitment in the same time period 1950-1987. SOI looks erratic, although a yearly cycle does exist. The fish recruitment changes more slowly, but still in an irregular way.

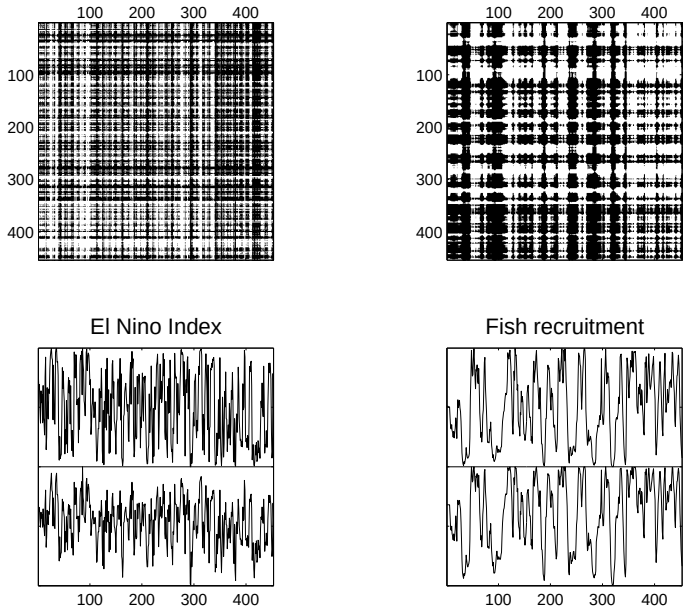


Figure 4. Left: Southern Oscillation Index. Right: Estimate of fish recruitment. Data taken from Shumway and Stoffer [15]

**Properties.** The matrix is skew-symmetric, that is, the pairs $(s, t)$ and $(t, s)$ are assigned different colors. The reflection at the main diagonal of the matrix turns black to white and white to black. The diagonal itself is black, by definition. It would be sufficient to draw only the upper or lower triangular matrix of $B$. Moreover, the smallest values $x_t$ correspond to black columns (and white rows) in $B$, and the largest values $x_t$ correspond to white columns (and black rows).

Periodicities in the time series correspond to periodicities in $B$ with respect to horizontal and vertical shift, and shift along the main diagonal. Even if there is no strict period, characteristic frequencies in the signal can be seen from the size of the patterns visible in $B$. The sine, the lynx series, and the autoregressive process have one characteristic frequency. White noise and the moving average process have no characteristic frequency at all. The speech signal a: shows two characteristic frequencies. For SOI and fish, the yearly cycle can be seen (the series contain values from almost 40 years), and there seem to be fluctuations with larger period.

## 3 Rank numbers

**Definition.** Ranking is everywhere in a competitive society: sailing contests, ranking of goods by price, ranking of universities ... The rank number $r_s$

indicates the position of $x_s$ in a rearranged series where all values are ordered by size. The smallest $x_s$ has rank 1, and the largest $x_s$ has rank $T$. In general, the rank $r_s$ corresponding to time $s$ is the number of values $x_t$ which are smaller than $x_s$,

$$r_s = \sum_{t=1}^{T} b_{st} \quad .$$

We have set $b_{ss} = 1$ in order to include $x_s$ in the count, and we should also include half of the $x_t$ which are equal to $x_s$, but as we said above we neglect ties in this paper.

It is obvious that $r_s < r_t$ if and only if $x_s < x_t$ so that original series and rank series have the same order structure matrix. Thus the rank numbers $(r_t)$ provide a complete description of the order structure of $(x_t)$, equivalent to the description by $B$.

In figures 1-4, the graphs of ranks $r_t$ are drawn above the time series $x_t$. The ranks of the sine form a symmetric sawtooth function. This rank function will be obtained whenever $x_t$ is the periodic extension of a symmetric function on an interval with one strictly increasing and one decreasing part, as for example $x_t = 1 - t^2$ on $[-1, 1]$. This means that from the ordinal viewpoint, all these time series are identified.

**Uniform distribution.** The smallest $x_t$ has rank 1, the second smallest has rank 2,..., and the largest has rank $T$. Each rank number appears once, so the $r_t$ are uniformly distributed on $\{1, ..., T\}$.

In this setting, the $r_t$ depend on $T$. It seems more appropriate to consider *relative rank numbers* $F(x_t) = r_t/T$ which are between 0 and 1. They are (almost) uniformly distributed on $[0, 1]$, and they will stabilize with increasing $T$ in the case of a stationary time series. In the following we shall need distinguish between $r_t$ and $F(x_t)$. Their uniform distribution is essential, however.

(Mathematical remark: $F(x_t)$ is the probability that an arbitrary value $x$ of the time series is smaller than the given $x_t$. Thus $F$ is the (cumulative) distribution function of the distribution of the values $x$. In the case of a continuous distribution, the values $F(x)$ are uniformly distributed on $[0, 1]$.)

**Graphical independence test.** As an application, we mention a graphical independence test which we introduced in [2], cf. [12]. Suppose we want to know whether $x_t$ and $x_{t+k}$ in the lynx series are independent. It is not enough to consider correlation and rank correlation since this will only check for linear and monotone dependence, respectively. A chi-square test can decide whether data deviate significantly from independence. However, our graphical method

will provide much more information than such a test – we can see the type of dependence.

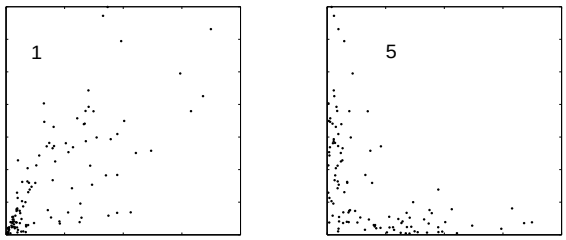


Figure 5a. Delay plots of the original data
of the lynx series are not meaningful.

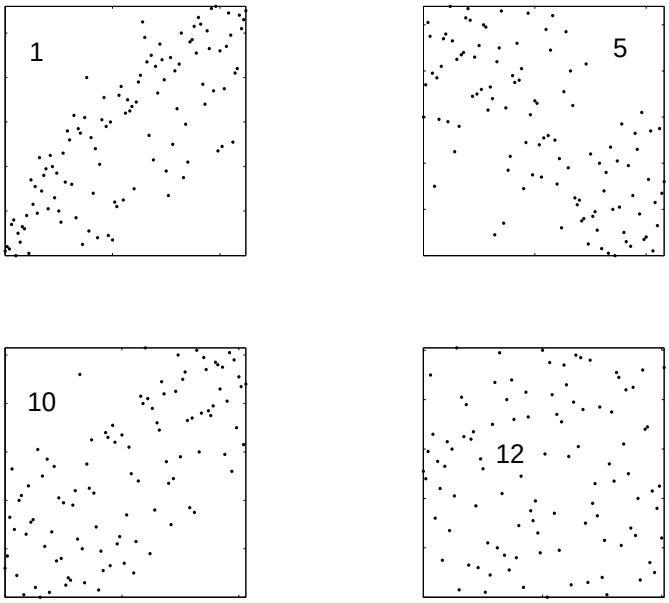


Figure 5b. The rank delay plots of the lynx series
give a clear picture of dependencies.

Figure 5a shows the points $(x_t, x_{t+k})$ for $t = 1, ..., T-k$ and for $k = 1, 5$. This so-called delay plot cannot verify independence: most of the 114 values of $x_t$ are small so that the points gather near the origin. In figure 5b, we have drawn the delay plot with $r_t$ and $r_{t+k}$. Both the horizontal and the vertical coordinate are uniformly distributed. So they are independent if and only if the points $(r_t, r_{t+k})$ are uniformly distributed in the square. Figure 5b shows that this is true for $k = 12$, as far as we can decide from 114 points. Thus it is impossible to give any information about the value after 12 years only knowing the present value.

Now let us study the dependence for $k = 10$. If $r_t$ is small, then $r_{t+10}$ cannot be very large, and conversely. This is true because the main period of the lynx series is 9 or 10 years. For $k = 5$, small values of $r_t$ correspond to larger values of $r_{t+k}$. We also see that $r_t$ and $r_{t+1}$ are not too much different. In most cases $r_{t+1}$ is a little larger than $r_t$ but sometimes it is much smaller.

## 4 Local order structure

**Motivation.** So far we have compared all values $x_t$ with each other. In some cases, it is more reasonable to compare only values which are not more than $d$ time steps apart. This will result in a fairly general but also very stable and robust structure.

Comparing values from quite different time points is not always appropriate. If we ask a wine-grower for the quality of this year's grapes, or a US citizen about his view on the president in power, then our opponent can in fact only compare with a few preceding values. Today's unemployment rate or Dow Jones index cannot be compared with the respective figure two years ago, because the whole environment as well as the rules of calculation have changed in the meantime. The same is true for time series from nature since technical equipment and methods of measurement change from time to time.

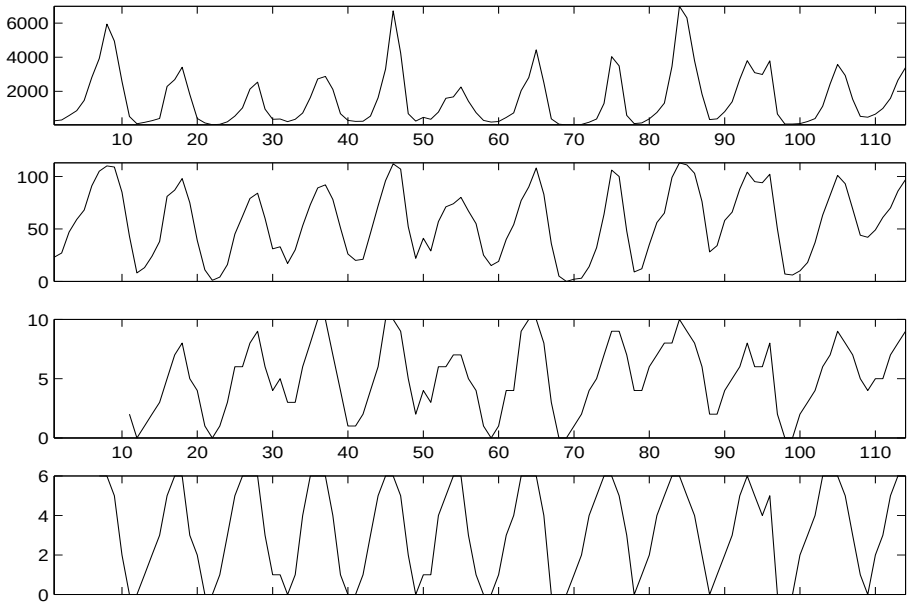


Figure 6. The lynx series, its global ranks, and its local ranks of level 10 and 6 (from above). The local structure of level 6 is very regular.

**Local rank numbers.** The local order structure of level $d$ is given by the

$d$ diagonal lines above and below the main diagonal in the matrix $B$. Alternatively, we can define local rank numbers by comparing each $x_s$ with the $d$ preceding values.

$$r_s^d = \sum_{k=1}^{d} b_{s,s-k} \qquad \text{for } s = d+1, ..., T.$$

Taking local rank numbers $r_s^d$ instead of the values $x_s$ is a kind of high-pass filter which removes a possible trend and very low frequencies from the time series. What remains is a very general structure which only involves local changes.

**Example.** The first row of figure 6 shows the lynx series, the second row the global rank numbers $r_t$ as in figure 3. The three extreme peaks of the $x_t$ are not so pronounced on the rank scale. The last two rows show the local rank numbers for $d = 10$ and $d = 6$ which start with $t = d + 1$. For $d = 10$ irregularities are caused by the fact that the basic period changes between 8 and 10. For $d = 6$, however, the local structure is surprisingly regular.

## 5 Kendall's tau for autocorrelation

**Definitions.** Autocorrelation is one of the basic methods in time series analysis which measures how much the given series $(x_t)$ coincides with the shifted series $(x_{t+k})$. Varying $k = 1, 2, ...$ we obtain the *ordinary autocorrelation function*

$$\rho(k) = \frac{1}{s^2} \cdot (\frac{1}{T} \sum_{t=1}^{T-k} x_t x_{t+k} \ - m^2)$$

where $m = \frac{1}{T} \sum x_t$ denotes the mean and $s^2 = \frac{1}{T} \sum x_t^2 \ - m^2$ the variance of the series [14]. If we replace $x_t$ by $r_t$ in this formula we obtain the *rank autocorrelation function* $\tilde{\rho}(k)$ studied extensively by Hallin et al [8,9]. It is based on the global order structure.

If we replace $x_t$ in the above formula by $b_{t,t+d}$ then we get another autocorrelation function $\tau^d(k)$ which we call *Kendall's tau with test delay d*. This parameter depends only on the local order structure. For the classical setting of correlation of two variables, such a parameter was suggested by Kendall [11]. For time series autocorrelation, a similar parameter, a kind of average of all $\tau^d(k)$ with $d = 1, ..., T-1$, was introduced by Hallin et al [6] for testing linear autoregression coefficients. We study $\tau^d(k)$ as a correlation function in

its own right, and we keep the test delay $d$ as a scale parameter which can be chosen appropriately.

The above formula standardizes all correlation coefficients to lie in $[-1, 1]$, where the value 1 indicates perfect coincidence of original and shifted time series. A non-standardized version of $\tau^d(k)$ can be calculated in a very simple way:

$$\tau_o^d(k) = \frac{1}{T} \sum_{t=1}^{T-k} b_{t,t+d} \diamond b_{t+k,t+k+d}.$$

Here $\diamond$ means logical equivalence: $b \diamond b' = 1$ if $b = b'$ and $b \diamond b' = 0$ otherwise.

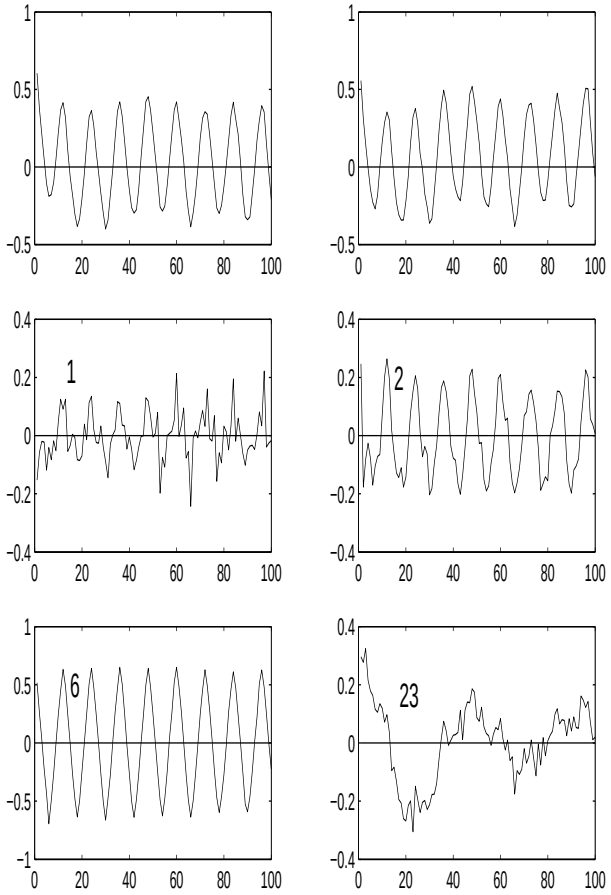


Figure 7. Autocorrelation functions for the SOI data. First row: Ordinary and rank correlation. Below: Kendall's tau with different delays.

**Interpretation.** The autocorrelation coefficient $\rho(k)$ measures the coincidence of values $x_t$ of the original series and $x_{t+k}$ of the series which was shifted $k$ steps. The rank autocorrelation coefficient $\tilde{\rho}(k)$ measures the coincidence of global rank numbers of original and shifted series.

The tau-correlation $\tau^1(k)$ measures the coincidence of the order relation between values and their right neighbors, in original and shifted series. $\tau^d(k)$ does the same for the order relation between values and their 'neighbors' $d$ steps apart.

For smooth time series, like those on the left in figures 1-3, the functions $\tau^1, \tau^2, \tau^3, ...$ will almost coincide, and they are similar to $\rho$ and $\tilde{\rho}$. For a noisy series, $\tau^1$ will not show much structure, and $d$ should be taken larger. On the other hand, if the test delay is chosen too large, the tau-correlation will usually be very small. Nevertheless, tau correlation can provide information not contained in ordinary and rank correlation, as we shall demonstrate now.

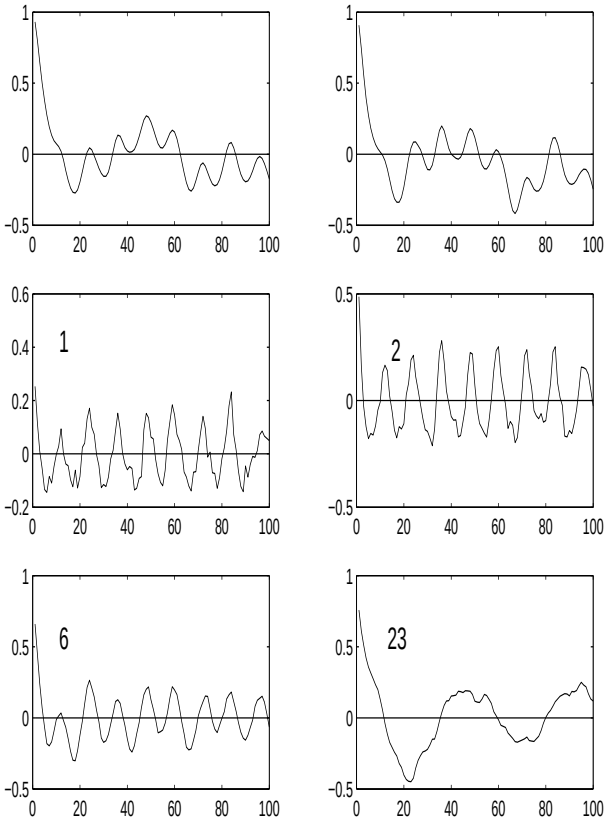


Figure 8. Autocorrelation functions for the fish recruitment series, as in fig. 7.

**Example.** Figures 7-8 show the autocorrelation of the SOI and fish recruitment series. Both ordinary and rank autocorrelation in the upper row show the yearly cycle, period 12 for the monthly series. This is also obtained by Kendall's tau for any $d$. However, the small $d = 1, 2$ leads to irregularities, especially for the noisy SOI series. $d = 6$ is much better. In fact a period $p$ is detected best with $d = p/2$.

To find the yearly cycle is neither difficult nor spectacular. We would like to know whether there are other, longer periods in the series. For this purpose,

we choose $d$ as a multiple of the main period $p = 12$. Then because $x_t$ and $x_{t+d}$ describe the same month of the year, their comparison will not be influenced by the yearly cycle. In our case, $d = 12$ is not very good, as $d = 1$ above. However, $d = 23, 24, 25$ give almost the same tau correlation functions, shown in the lower right part of figures 7-8. By clever choice of $d$ we have eliminated the influence of the main period, and found another period around 50. The smoothed spectrum with an appropriate window in Shumway and Stoffer ([14], p. 244) led to 46.5 months in both series, which is in good coincidence with our tau function. Spectral methods need a lot of fine-tuning, however, while tau for $d = 12, 24$ can be calculated directly.

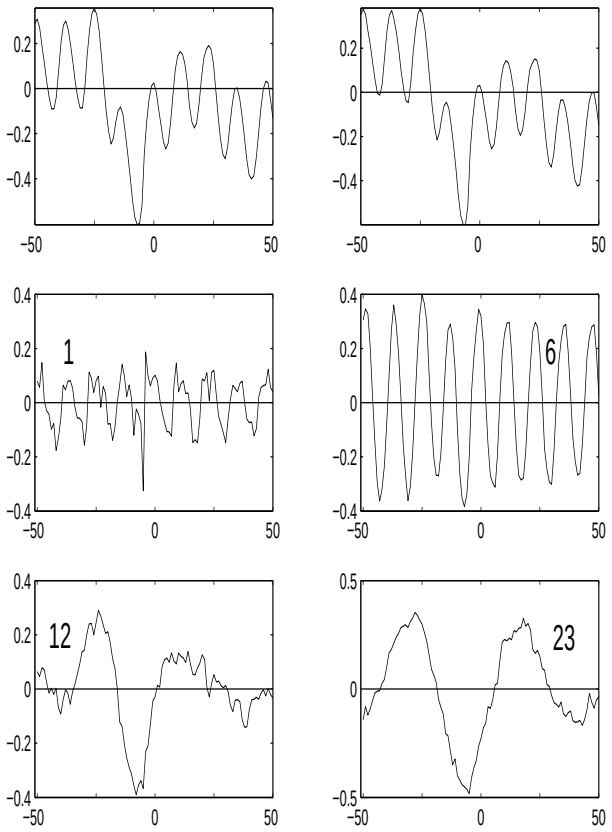


Figure 9. Crosscorrelation functions for SOI and fish recruitment. Details as in fig. 7. There is a minimum at -6 in all functions.

In figure 9 we have shown the cross-correlation between SOI and fish recruitment. There is a minimum at -6, which means that decrease of SOI today leads to increase of fish after 6 months. In the ordinary and rank crosscorrelation, this minimum is hard to detect since the function is dominated by the yearly cycle inherent in both series. This also holds for tau-crosscorrelation with delay 6, where the minimum almost disappears. For $d = 12$ or 23,24 the yearly cycle is eliminated, and the influence of SOI on fish recruitment becomes obvious. In the classical setting, partial correlation or spectral methods

have to be applied to get this result [14].

**Remark.** An important advantage of tau is that outliers and trend in the series will influence the result only slightly. We have done a number of experiments with EEG signals which contain a lot of so-called artefacts [5]. Series of 1 minute length (15000 values) could be treated by tau-correlation, and consecutive minutes have shown almost the same picture in spite of artefacts, while conventional correlation and spectral methods can only treat clean segments of 4 or at most 8 seconds.

## 6 Permutation entropy

**Order patterns.** The information contained in one comparison is one bit. A number with three decimals accuracy needs ten bits for its description, as much as 10 comparisons. Thus ordinal methods provide a huge potential for studying three and more time points at the same time, which cannot be done by usual correlation methods.

A first step in this direction was done in [3] where we studied the order patterns at $n$ equally spaced time points $t, t+d, t+2d, ..., t+(n-1)d$. Here $t$ runs from 1 to $T-(n-1)d$, and the scale parameter $d$ plays the same part as above in correlation.

For $n$ different numbers, there are $n! = 1 \cdot 2 \cdot ... \cdot n$ possible order patterns $\pi$ which are also called permutations. These permutations can be easily assigned to the numbers $1, ..., n!$ For example, for $d = 1$ we can set

$$\pi_t = r^1_{t+1} + 2!\, r^2_{t+2} + ... + (n-1)!\, r^{n-1}_{t+n-1} \ .$$

Some of these patterns will occur quite often in our time series. Some will only appear for a few $t$, and some will not appear. The idea now is to take the number of order patterns in the series as a measure for its complexity. To account for the different frequencies, we do not just take the number of $\pi$'s. We determine the entropy.

**Definition of permutation entropy.** For a permutation with number $\pi$, let $f(\pi)$ denote its frequency in the time series. In other words: $f(\pi)$ is the number of $t$ between 1 and $T-(n-1)d$ for which $\pi_t = \pi$. The relative frequency is $p(\pi) = f(\pi)/(T-(n-1)d)$. The *permutation entropy* of order $n$ of the time series is then defined as

$$H_n = -\sum_{\pi=1}^{n!} p(\pi) \cdot \log p(\pi) \ .$$

Clearly, permutation entropy refers to the local order structure of the time series. The smallest possible value of $H_n$ is zero. It will be attained for a monotonously increasing or decreasing series. The largest possible value of $H_n$ is $\log n!$ which is realized when all permutations have equal probability. This happens for white noise. (Here we refer to natural logarithms, but it does not matter which log we take.)

Thus large $H$ indicates noise and random influence while small $H$ means regular, deterministic behavior. $H_n$ should be compared for fixed $n$, but we may vary the test delay $d$ to study the complexity of a time series on different scales. This is done here for the first time. Instead of fixing $d = 1$, we let $d$ run from 1 to 20, and draw $H_4$ as a function of $d$.

Two interesting properties of these functions should be mentioned: if the series has main period $p$ then the entropy is small for $d = p/2$ and $3p/2$, and it is large for $d = p$ and $d = 2p$. The first property can be explained as follows. Since we have a tendency that $x_t = x_{t+p}$, the relation $x_t < x_{t+p/2}$ will imply that most likely $x_{t+p/2} > x_{t+p}$. So certain permutations for $d = p/2$ will not appear, or will appear with very small probability, resulting in small $H$. The second property comes from the fact that when there is a tendency that $x_t = x_{t+p}$, then for $d = p$ any kind of permutation is likely to appear, as for white noise. (Here we have to assume, that there are no other larger periods.) We shall see that the second effect is less pronounced than the first one.

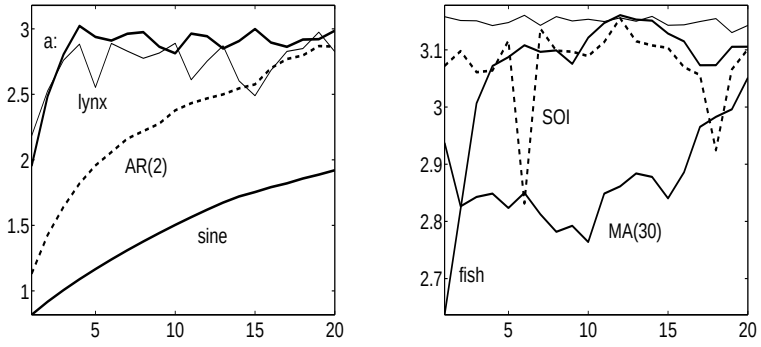


Figure 10. Permutation entropy as a function of test delay. Left: series with strong local dependencies. Right: noisy time series.

**Examples.** Figure 10 shows the permutation entropy of order 4 of the time series of figure 1-4, as a function of the test delay. The maximum value of $\log 4! \approx 3.178$ obtained by white noise is the end of the scale in both pictures. On the left we drew the fairly regular time series. Our sine function has the smallest entropy, and our AR(2) process has slightly larger entropy. The entropy increases with increasing $d$, and should approach its maximum value for $d \to \infty$.

For chaotic time series the approach should be very fast, since "sensitive dependence of the initial values" is a characteristic property of chaotic dynamical systems. In fact we see this behavior for the speech signal a: and for the lynx series. Some fluctuations of the permutation entropy could be caused from the fact that for 114 values, the statistics is not very accurate. However, given that the main period is about 10, it is a consequence that the entropy should be small for $d = 5$ and $d = 15$, and large for $d = 10$. This is what wee see!

On the right, time series with random influence were studied. Their entropy is so large that we took a different scale of values. The thin line at the top shows an evaluation of white noise entropy from 400 data points, indicating statistical fluctuations and the difference to the theoretical upper bound. The fat line below is the moving average process. For MA(30), the small values for $d = 10$ and $d = 15$ are probably not obtained by chance. The dashed line shows the SOI series, with two minima at $d = 6, 18$ and a maximum at $d = 12$ explained by the period 12. The fish recruitment series, however, does not show the minima, only a maximum at $d = 12$. It behaves rather like a trajectory of a chaotic dynamical system, which coincides with the impression of the time series itself (figure 4).

**Results of [3] and [4].** We studied $h_n = H_n/(n-1)$ numerically for time series generated by chaotic logistic maps $x_{t+1} = rx_t(1-x_t)$, for 5000 values of $r$ and $n = 2, ..., 16$. We found that $h_n$ is almost proportional to the so-called Lyapunov exponent of the map, which is taken as a complexity parameter of the time series. In [4] we proved that for arbitrary piecewise monotone one-dimensional maps, $\lim_{n\to\infty} h_n$ exists and equals the well-known Kolmogorov-Sinai entropy of the map. Moreover, complete spoken sentences were taken as time series, and $h_n$ was determined as a function of time, using a series of several thousands sliding windows. The result was that voiced and unvoiced sounds can be distinguished very well, even with $n = 3$ or 4, and with very short window length of 128 for which Fourier methods do not apply any more.

**Conclusion.** Some new methods of time series analysis were presented. Since they make only use of the order of the values, they are robust under non-linear distortion of the signal, and their calculation is very fast. These methods seem to be useful for many fields of application, and they deserve to be developed further.

## References


[1] H.D.I. Abarbanel, Analysis of observed chaotic data, Springer, New York 1995

[2] C. Bandt and B. Pompe, Entropy profiles of speech signals, Phys. Letters A 175 (1993), 305-313

[3] C. Bandt und B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (2002), 174102

[4] C. Bandt, G. Keller und B. Pompe, Entropy of interval maps via permutations, Nonlinearity 15 (2002), 1595-1602

[5] C. Bandt, S. Laude und H. Lauffer, Kendall's tau as autocorrelation, Preprint, Greifswald 2000

[6] S.T. Ferguson, C. Genest und M. Hallin, Kendall's tau for serial dependence, Canadian J. Statist. 28 (2000), 587-604

[7] A. Giuliani, M. Colafranceschi, C.L. Webber Jr. und J.P. Zbilut, A complexity score derived from principal components analysis of nonlinear order measures, Physica A 301 (2001), 567-588

[8] M. Hallin und J. Jurečkova, Optimal tests for autoregressive models based on autoregression rank scores, Annals Stat. 27 (1999), 1385-1414

[9] M. Hallin und B.J.M. Werker, Optimal testing for semi-parametric AR models – from Gaussian Lagrange multipliers to autoregression rank scores and adaptive tests, Asymptotics, Nonparametrics, and time series (ed. S. Gosh), Marcel Dekker, New York 1999, 295-350

[10] K. Keller und H. Lauffer, Symbolic analysis of high-dimensional time series, Preprint 7 des SPP, Greifswald 2002

[11] M. Kendall und J.D. Gibbons, Rank correlation methods, 5th ed., London, Edward Arnold, 1990

[12] B. Pompe, Ranking and entropy estimation in nonlinear time series analysis, Nonlinear analysis of physiological data (eds. H. Kantz, J. Kurths, G. Mayer-Kress), Springer 1998, 67-90

[13] B. Pompe, Ordinale Zeitreihenanalyse, Semester-Kolloquium, Angewandte Physik, Greifswald, February 2001

[14] R.H. Shumway, D.S. Stoffer, Time Series Analysis and its Applications, Springer 2000

[15] H. Tong, Non-linear time series, Oxford University Press 1990