

ON THE NONLINEAR ANALYSIS OF TIME SERIES

Tom Kuusela

Contents

1. Introduction	3
2. Approximate Entropy (ApEn)	4
3. Sample Entropy (SampEn)	7
4. Cross Entropies (Cross-ApEn, Cross-SampEn)	9
5. Correlation Dimension (CD)	11
6. Pointwise Correlation Dimension (CDi)	14
7. Fractal Dimension by Dispersion Analysis (FD-DA)	16
8. Fractal Dimension by Curve Length (FD-L)	18
9. Similarity of Distribution (SOD)	20
10. Return Map (RM)	22
11. Detrended Fluctuation Analysis (DFA)	25
12. Spectrum Power-Law (SPL)	28
13. Spectral Entropy (SEn)	30
14. Symbolic Dynamics (SymDyn)	32
15. Lempel-Ziv Entropy (LZEn)	35
16. Stationarity (StatAv)	37
17. Fractional Spectral Radius (FSR)	39

1. Introduction

There are numerous methods and approaches available for time series analysis. This paper presents selected methods, which are or at least could be useful especially when analysing biosignals. The methods described differ greatly and some of them are commonly used characterization methods. One must always select the method to use case by case.

The methods described in this paper are often called *nonlinear* methods. The name is in fact not quite fitting and is actually even misleading because the methods themselves are not more nonlinear than e.g. common spectrum analysis using FFT algorithm. The idea behind the naming convention used is probably related to the fact that the system producing the time series is nonlinear, which in the case of biological systems might well be true. Sometimes these methods are described with the term *chaos-theoretical approach*. This is true only with several of the methods in which the assumption is made that the systems trajectory in a phase space is actually a chaotic small dimensional attractor. There is very little proof of the existence of this in biological systems. In particular, if we are studying the cardiovascular blood pressure regulating system and its typical time series we can find in the literature a growing number of results, which indicate this system to be rather stochastic than chaotic.

Biosignal time series are in many ways very difficult to analyse. Many methods demand that the time series must have thousands of data points, which may be impossible or difficult to achieve depending on the measurement methods used. When the time series is long, notable changes in the biological system under measurement will occur almost inevitably and the time series will start to drift in the parameter space of the system either due to internal or external influences. Such instability has a negative effect on all analysis methods in which a specific characteristic of the whole time series is to be compressed into a single statistic. Calculations in themselves of course are successful but what is the meaning of the statistic? Other disturbing factors in these signals are normally notable noise levels, strong discreteness of the signal due to limited amplitude resolution of the digitalisation and often occurring strong periodic signals e.g. breathing modulation on top of the otherwise chaotic or stochastic signal. It should be remembered that all the methods described in this paper were originally developed in physics for analysing very different and often practically almost ideal systems.

Despite the warnings stated above the analysis of biosignals using alternative methods might still be beneficial. Even though the calculated statistic might not exactly describe the character for which it was originally designed to represent it could still play an important role when searching for correlation with other clinically interesting parameters. However, the basic assumptions and limitations related to each method should be clearly understood when interpreting results and different explanatory models.

2. Approximate Entropy (ApEn)

Approximate entropy ApEn measures the complexity of a time series, i.e. a smaller ApEn value indicates a more regular time series. ApEn may be used as an alternative method for describing the variability of a signal.

Definition

ApEn is defined as follows. First, so called pseudo phase space vectors are formed from the initial time series $\{x(i)\}$, in which $i = 1 \dots N$, N being the number of data points.

$$u(i) = [x(i), x(i+1), x(i+2), \dots, x(i+m-1)],$$

m being the so-called embedding dimension. Altogether $N-m+1$ vectors $u(i)$ may be constructed. The distance of these vectors d is defined as the maximum distance of the corresponding components

$$d[u(i), u(j)] = \max\{|u(i+k) - u(j+k)| : 0 \leq k \leq m-1\}.$$

Next, we calculate for each vector $u(i)$ the number of other vectors $u(j)$, which are at a distance r , and get

$$C_i^m(r) = \{\text{the number of index } j \text{ for which, } j \leq N - m + 1, d[u(i), u(j)] \leq r\} / (N - m + 1).$$

Due to normalization the maximum value of C parameters is 1. Next we define an auxiliary quantity

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r).$$

Finally we get approximative entropy from the equation

$$\text{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r),$$

In other words ApEn measures the (logarithmic) probability that at m points similar (nearby) vectors are also similar in the next point.

ApEn in practice

$\text{ApEn}(m, r, N)$ is then dependent on three parameters, i.e. on the length m of the compared vectors, on tolerance parameter r and on the number N of data points. For biosignals the value $m = 2$ for which the method functions best is normally used. When the number of

data points is increased ApEn approaches its final value asymptotically. In practice when $N > 800$ and $m = 2$ gives a reliable result. ApEn depends strongly on the tolerance parameter. If r is chosen so that it is a fraction of the standard deviation SD of the signal, the ApEn value of the signal is not dependent on absolute variability and understandably neither on the unit of the signal. Then it is possible to compare e.g. the entropies of RRI and SAP time series with each other. The most used choice is $r = 20\%$ of SD.

Benefits of ApEn

- Measures variability of the signal without the need to fix the nature of changes. Normal spectrum analysis for example assumes changes to be sinusoidal.
- Calculations are not based on specific assumptions regarding the inner structure or dynamics of the system.

Shortcomings of ApEn

- Value depends on three parameters, i.e. direct comparison always requires fixing of parameters.
- ApEn is sensitive to the smallest linear trends in the data because comparison is based on the absolute values of the signal. This tendency can be dampened by concentrating on the subtraction of consecutive data points.
- In principle ApEn is not sensitive to changes in single data values but if the tolerance parameter is bound to deviation the situation could well be different.

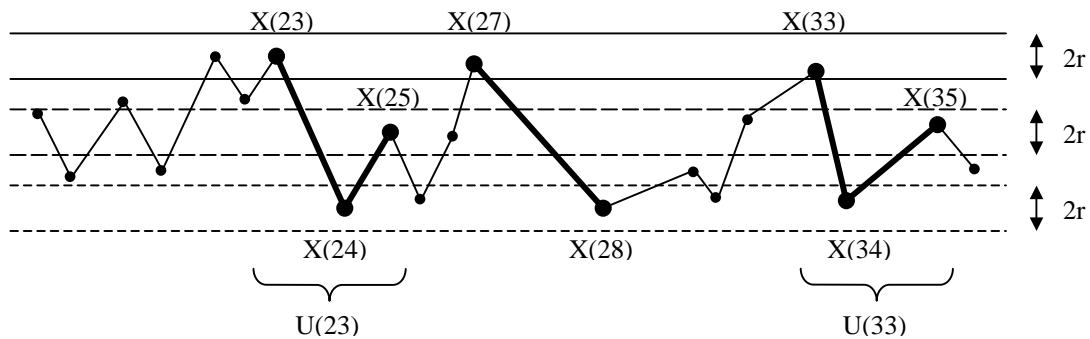


Figure 1. An example showing how with $m = 2$ a search is made for nearby vectors. When using vector $[x(23), x(24)]$ for comparison two nearby vectors, $[x(27), x(28)]$ and $[x(33), x(34)]$, are found meaning they increase parameter $C_{23}^2(r)$. When comparison is made also to the following component $x(25)$, only vector $[x(33), x(34), x(35)]$ is nearby, which means that only this vector is increasing factor $C_{23}^3(r)$.

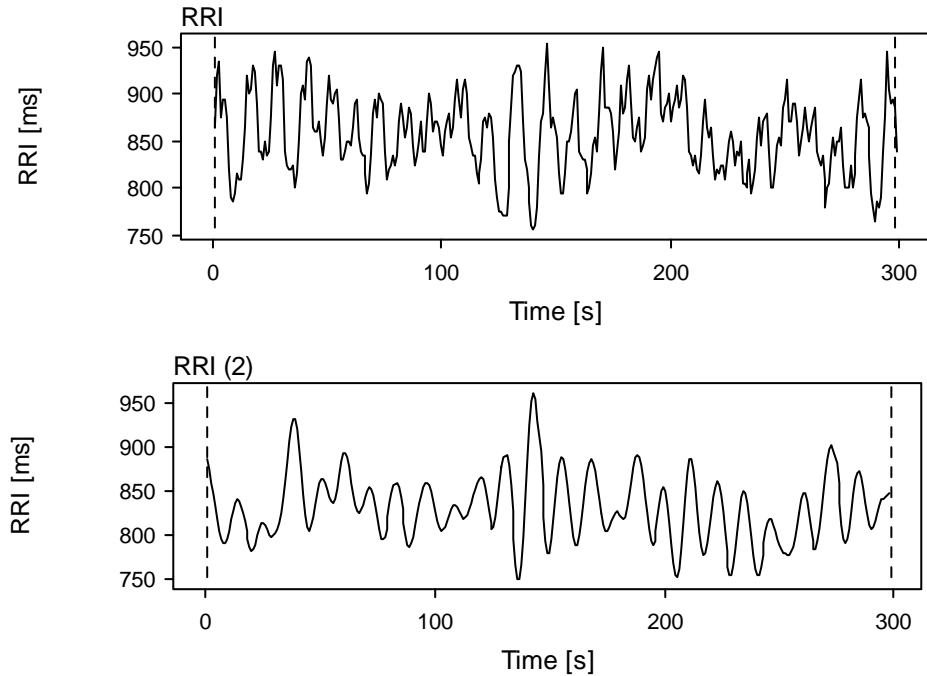


Figure 2. Two RRI signals with the same SD, but for the higher signal $ApEn = 1.149$ and for the lower signal $ApEn = 0.583$. Calculations have been made using $m = 2$ and $r = 20\%$ of SD.

References

Pincus S. M., Goldberger A. L. Physiological time-series analysis: what does regularity quantify? *Am. J. Physiol.* **266** (*Heart. Circ. Physiol.*): H1643 – H1656, 1995

Pincus S. Approximate entropy (ApEn) as a complexity measure. *CHAOS* **5**: 110 – 117, 1995

Bettermann H., Van Leeuwen P. Evidence of phase transitions in heart period dynamics. *Biol. Cybern.* **78**: 63 – 70, 1998

Cysarz D., Bettermann H., Van Leeuwen P. Entropies of short binary sequences in heart period dynamics. *Am. J. Physiol.* **278** (*Heart Circ. Physiol.*): H2163 – H2172, 2000

3. Sample Entropy (SampEn)

SampEn is a very similar type of statistic like ApEn, but there is a small yet fundamental difference in the way it is calculated. When calculating the number of similar vectors for each single vector in turn in the calculation of ApEn the vector itself is also taken into account. By this it is made certain that $C_i^m(r)$ parameters are always different from zero, which is essential since later on a logarithm is taken. This causes ApEn most of the time to give a result, which indicates greater regularity of the signal than regularity actually is.

SampEn is calculated in a way, which removes the above-described distortion from the results. When calculating the number of nearby vectors comparison to the vector itself is prevented:

$$C_i^m(r) = \{\text{the number of index } j \text{ for which, } j \neq i, j \leq N - m + 1, d[u(i), u(j)] \leq r\} / (N - m + 1)$$

Auxiliary factor Φ is also defined again without logarithms

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} C_i^m(r).$$

Now SampEn is defined as

$$\text{SampEn}(m, r, N) = \ln(\Phi^m(r) / \Phi^{m+1}(r)).$$

The interpretation as well as the use of SampEn remains exactly the same as for ApEn. Dependability on tolerance parameter r and the number of data points N is however different. ApEn reaches its maximum with a certain value of r , but SampEn decreases monotonically as r increases. SampEn is in principle also independent on the number of datapoint N , but with small values of N its statistical relevance is naturally poor. When r and N are large enough SampEn and ApEn give the same result.

Benefits of SampEn

- Provides in principle a more reliable estimate of the complexity of the signal compared to ApEn.
- May be used for considerably shorter time series than ApEn, even using under 200 data points.

Shortcomings of SampEn

- Comparison with ApEn may be difficult even using same parameter values, if the parameters have been selected badly.

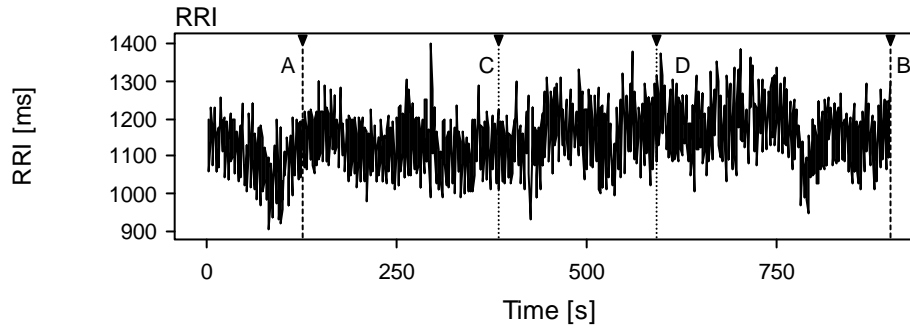


Figure 3. RRI signal with ApEn and SampEn for time periods of different lengths. Interval A – B contains 800 points, with ApEn = 1.331 and SampEn = 1.695; interval C – D contains 200 points, with ApEn = 0.808 and SampEn = 1.725. The results indicate that SampEn gives a good estimate already for considerably shorter time periods.

References

Richman J. S., Moorman J. R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* **278** (Heart. Circ. Physiol.): H2039 – H2049, 2000

4. Cross Entropies (Cross-ApEn, Cross-SampEn)

Both ApEn and SampEn present the possibility to calculate complexity between two signals. In such cases we try to evaluate the synchronicity between the signals by looking at typical patterns or shapes in the signals rather than looking at them in time periodical sense. If cross entropy is low the behaviour of the signals resemble each other.

Cross entropies are calculated as before, but the vectors to be compared are taken always in pairs from the two time series. If the time series of the two signals are $\{x(i)\}$ and $\{y(i)\}$, $i = 1 \dots N$, the vectors of the pseudo phase space are formed as before as follows:

$$u(i) = [x(i), x(i+1), x(i+2), \dots, x(i+m-1)],$$

$$v(i) = [y(i), y(i+1), y(i+2), \dots, y(i+m-1)].$$

In this case the distance of the vectors is

$$d[u(i), v(j)] = \max\{|u(i+k) - v(j+k)| : 0 \leq k \leq m-1\}.$$

After this calculations are carried out exactly as with one signal. ApEn and SampEn methods differ from each other in practice, since with ApEn it is not at all certain that for each vector taken from the other time series one can find a corresponding close enough vector from the other time series. If this indeed is the case we cannot calculate Cross-ApEn. Cross-SampEn allows much more in this sense because it suffices if we have even for one vector a corresponding vector close by from the other time series. Cross-SampEn is bipolar, i.e. the result does not depend on the order in which the signals are compared in contradiction to Cross-ApEn method.

Cross-ApEn and Cross-SampEn in practice

The same rules as for the entropies of a single signal apply for the selection of parameters. In order for the comparison of signals to make sense, it is usually necessary to normalize both signals and this is true especially in the case that the signals represent differing quantities. This is usually carried out as follows:

$$x^*(i) = (x(i) - \langle x \rangle) / SD(x),$$

where $\langle x \rangle$ is the mean of the signal and $SD(x)$ its standard deviation.

Benefits of Cross-ApEn and cross-SampEn

- The method provides a completely new way of comparing signals without limiting the analysis only on pure time synchronization.

Shortcomings of Cross-ApEn and Cross-SampEn

- Cross entropies are not very easy to grasp intuitively for which reason the interpretation of results might prove difficult.

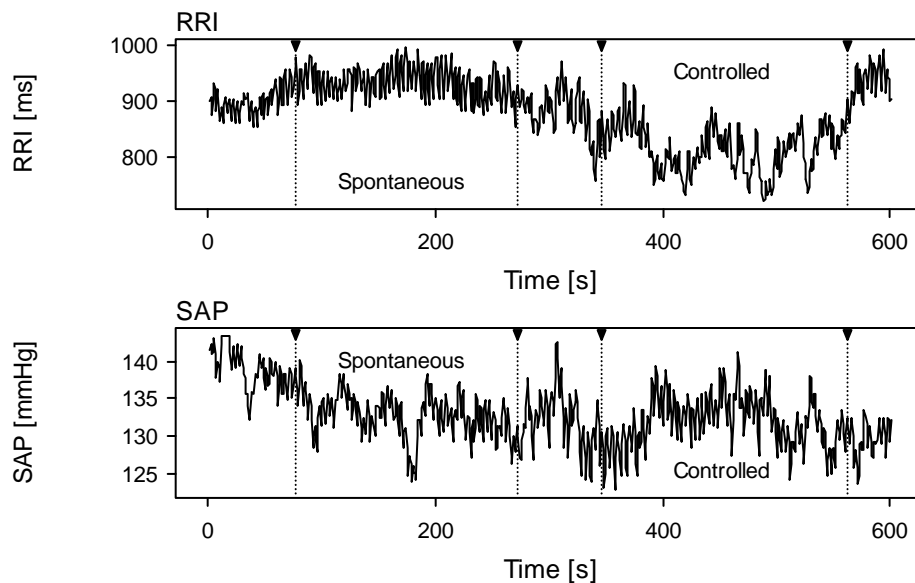


Figure 4. RRI and SAP time series in which the patient has breathed spontaneously in the beginning and then in a controlled way at a frequency of 0.25 Hz. RRI SampEn(spont) = 1.562, SAP SampEn (spont) = 1.612 and RRI-SAP Cross-SampEn(spont) = 2.028; RRI SampEn(contr) = 1.493, SAP SampEn (contr) = 1.632, RRI-SAP Cross-SampEn(contr) = 1.734. Entropy of RRI decreases slightly when breathing rhythm is changed to controlled mode but for SAP remains practically the same. Instead a clear change is in cross entropy can be seen.

References

Richman J. S., Moorman J. R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* **278** (Heart. Circ. Physiol.): H2039 – H2049, 2000

Pincus S., Singer B. H. Randomness and degrees of irregularity. *Proc. Natl. Acad. Sci.* **93**: 2083 – 2088, 1996

Pincus S. M., Mulligan T., Iranmanesh A., Gheorghiu S., Godshalk M., Veldhuis J. D. Older males secrete luteinizing hormone and testosterone more irregularly, and jointly asynchronously, than younger males. *Proc. Natl. Acad. Sci.* **93**: 14100 – 14105, 1996

5. Correlation Dimension (CD)

The functioning of the system can be described by measuring its attractor (the path towards which the system converges towards) dimension. Especially for chaotic systems the attractor can be fractal in which case its dimension is not an integer. Knowing the dimension of the attractor may also give useful information of the characteristics of the underlying system also when speaking of biological systems. Correlation dimension is sometimes referred to with the designation D2.

The basis for the calculations once again are in the reconstruction of the time series of the multidimensional phase space $\{x(i)\}$, $i = 1 .. N$, by using the vectors of the pseudo phase space

$$u(i) = [x(i), x(i+1), x(i+2), \dots, x(i+m-1)],$$

where m is the embedding dimension. Next we calculate for each vector $u(i)$ how many attractor points are at a distance r as measured from the observation point

$$C_i^m(r) = \{\text{the number of index } j \text{ for which, } j \leq N - m + 1, d[u(i), u(j)] \leq r\} / (N - m + 1),$$

where distance d is defined (i.e. differing from the ApEn method) as the normal euclidian distance

$$d[u(i), u(j)] = \left(\sum_{k=1}^m |u(i;k) - u(j;k)|^2 \right)^{1/2}.$$

Next we calculate the mean of the quantities $C_i^m(r)$ over all vectors, from which we then get the so-called correlation integral

$$C^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} C_i^m(r).$$

Correlation dimension CD is defined as a limit

$$CD(m) = \lim_{r \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\log C^m(r)}{\log r}.$$

In practice with limited data sets these limits cannot be calculated so CD is defined as the slope of the regression line calculated from a log-log representation and over an area with the required linearity.

CD in practice

When calculating CD, the embedding dimension m must be selected so that it is at least $2*D$, where D is the genuine dimension of the system under study i.e. the number of real dynamical variables. In order for the correlation integral to really describe the construction of the attractor correctly the number of data points should exceed 10^m . For example, when studying the blood pressure adjustment system we may assume that the number of dynamical variables >4 , which means that the time series must be very long. In addition, it is nearly impossible to find such a range of the distance r , in which $\log C^m(r)$ changes linearly as a function of $\log r$, the reason being noise contained in the data or unstationarity of the data. Due to these limitations calculation of the correct correlation dimension for biosignals is nearly impossible. Despite this fact this approach might still be useful. Promising results have been achieved by using $m = 20$ and by searching for the mean slope within $0.01 < C^m(r) < 0.1$. The quantity calculated in the described way without forgetting the above-mentioned limitations is called the *modified correlation dimension*. This quantity cannot accurately define the real dimension of the system but nevertheless does give a measure of the complexity of the system, i.e., when CD increases, the system becomes more complex.

Benefits of CD

- The method provides an alternative method to measure the complexity of a system.

Shortcomings of CD

- The method is based on the existence of a small dimensional (chaotic) attractor, the kind of which cannot necessarily be found for biological systems.
- The accurate calculation of CD for practical biosignal time series is nearly impossible.
- The selection of the parameters m and N have a fundamental impact on the end result.
- Calculation of CD for long time series is very time consuming.

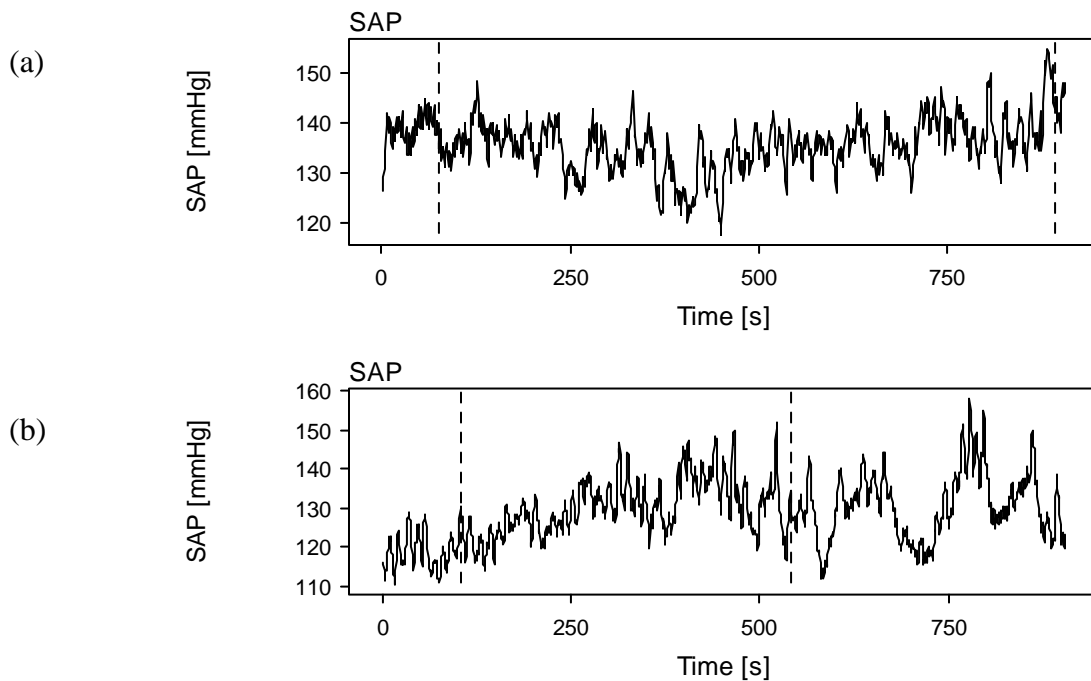


Figure 5. Two SAP time series sets measured with the patient at rest. (a) is normal situation, with $CD = 6.091$; in (b) the patient has been given terbutaline, and $CD = 3.545$. In each case the number of data points is 800 (area marked with dashed lines), but due to the different average heart beat the time periods differ from each other. According to dimension analysis the drug seems to strongly decrease the complexity of the system.

References

Grassberger P., Procaccia I. Characterization of strange attractors. *Phys. Rev. Lett.* **31**: 346 – 349, 1983

Kantz H., Schreiber T. Dimension estimates and physiological data. *CHAOS* **5**: 143 – 154, 1995

Fell J., Röscke J., Schäffner C. Surrogate data analysis of sleep electroencephalograms reveals evidence for nonlinearity. *Biol. Cybern.* **75**: 85 – 92, 1996

Yum M-K. Kim N-S. et al. Non-linear cardiac dynamics and morning dip: an unsound circadian rhythm. *Clin. Physiol.* **19**: 56 – 67, 1999

6. Pointwise Correlation Dimension (CDi)

Pointwise correlation dimension is defined in a very similar way compared to CD but instead of labelling the time series with a single value it is calculated as a function of time. This gives us the possibility to evaluate the changing of the system characteristics as a function of time, which is very important in unstationary cases. Pointwise correlation dimension is sometimes referred to with the designation D2i.

CDi in practice

When searching for the regression line we must once again select the area in which the relation is linear and to use a high enough embedding dimension value. When calculating the pointwise correlation dimension we can act as with the normal correlation dimension and fix $m = 20$ and by selecting the area $0.01 < CDi(r) < 0.1$. In addition we must note that even though CDi is calculated at each point and it can therefore in principle follow changes in data the calculation of CDi at each point requires the calculation of all other points. For this reason CDi is not applicable to clearly unstationary time series. Thus in practice it is advisable to use an additional condition which states that when calculating the regression line at each point in a log-log representation, the correlation factor of the thus achieved line must exceed a certain limit (e.g. 0.8) and if this criterion is not fulfilled the CDi value at the point in question is not reliable.

Benefits of CDi

- Correlation dimension as function of time gives some information of the internal changes of the time series, which in itself might be useful.

Shortcomings of CDi

- CDi does not necessarily function correctly for unstationary time series.

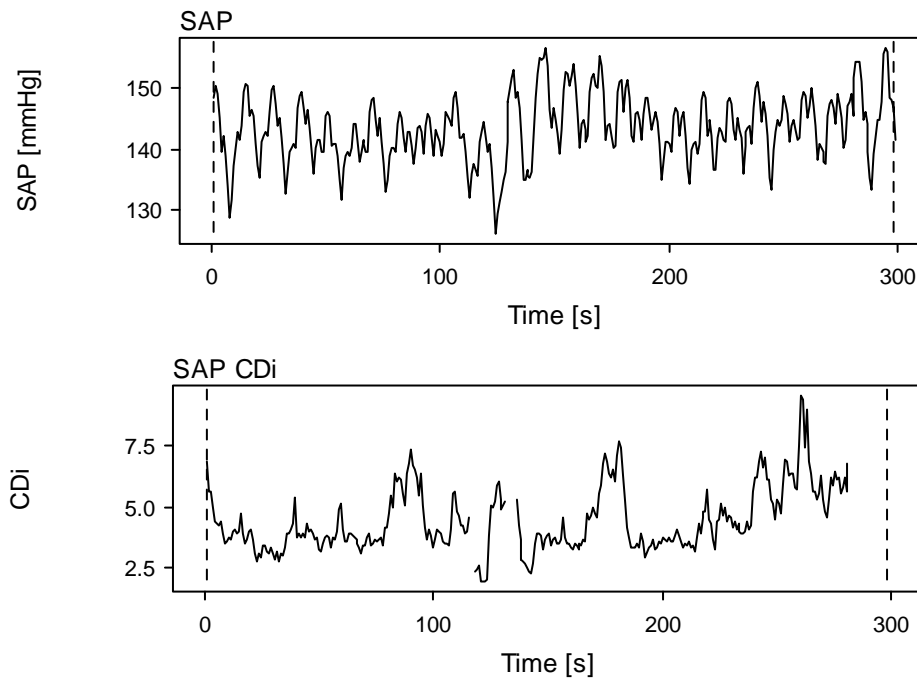


Figure 6. SAP time series and the corresponding point-by-point correlation dimension CDi. From the figure we can see that CDi increases slightly towards the end but occasionally has quite exceptional values. CDi curve has a discontinuity due to the fact that calculations could not be performed for each point this in turn rising from the requirement that the regression line correlation factor should be atleast 0.8.

References

Farmer J. D., Ott E., Yorke J. A. Dimension of chaotic attractors. *Physica D* **7**: 153 – 180, 1983

Mayer-Kress G., Yates F. E. et al. Dimensional analysis of nonlinear oscillations in brain, heart, and muscle. *Math. Biosci.* **90**: 155 – 182, 1988

7. Fractal Dimension by Dispersion Analysis (FD-DA)

Dimension analysis of time series may also be performed by studying the curve describing the time series itself rather than the dynamic system behind the signal. This kind of approach shows closest resemblance to image analysis. Because complex behaviour manifests itself in the time series as changes of at least apparently random changes, the study of the curves fractal structure will also give indirect information regarding the system itself.

In dispersion based analysis we first calculate from the time series the normal standard deviation

$$SD(1) = \frac{1}{N} \sqrt{N \sum_{i=1}^N x^2(i) - \left(\sum_{i=1}^N x(i) \right)^2},$$

where $\{x(i)\}$ is the time series under study with the time series having N data points. Next we calculate from the time series the mean of two consecutive data points for all data points and from this we get a new time series with $N/2$ datapoints. For this new time series we calculate standard deviation $SD(2)$. This is continued with group sizes 4, 16, 32 etc. until there are less than 4 data points left in the time series. By this way we have now a series of deviation values $SD(m)$. When we plot $\log SD(m)$ as a function of $\log m$, a line can be plotted through the points if the original time series curve was fractal. For this fractal object the dimension is $FD-DA = 1 - \text{slope of the line}$. Fractal dimension defined in this way can have values between 1 – 1.5, where 1 represents the case with a steady state signal and 1.5 represents maximal fractal characteristics.

FD-DA in practice

If the curve under study is not purely fractal, i.e. it contains a notable amount of noise, the $SD(m)$ values do not fit exactly a line in log-log representation. Especially in automated analysis in which case this line is not visually checked it is important to set some criteria for linearity. A well functioning choice again is the correlation factor of a well-fitted regression curve, from which we can demand a minimal value (e.g. 0.8).

This method is well suited for the study of all time series curves, because even though SD values in themselves are dependent on the unit of the quantity, scaling is apparent as standard deviation in log-log representation and this in turn has no effect on the dimension.

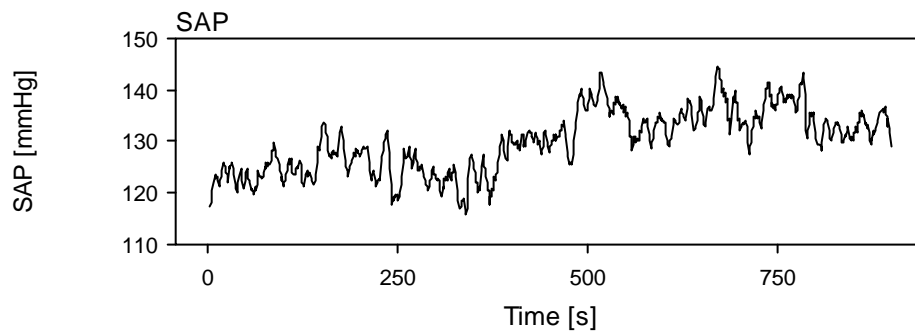
Benefits of FD-DA

- Method poses no special demands over the characteristics or dynamics of the system.
- Can be calculated quickly.

Shortcomings of FD-DA

- Because the method is used to study a directly measured signal curve, all measurement inaccuracies and disturbances affect the results directly.
- Sensitive to background trends of the signal.

(a)



(b)

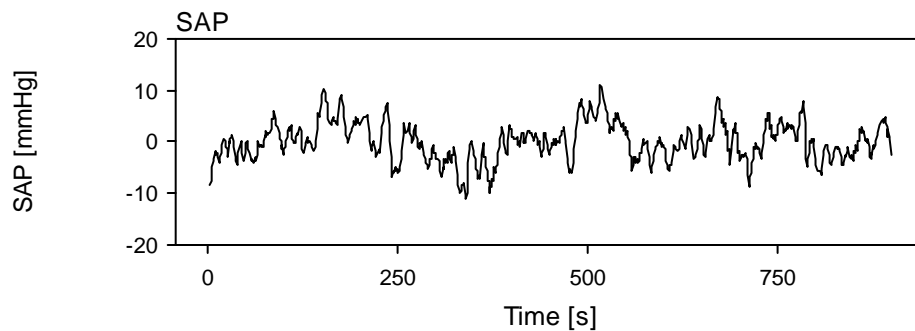


Figure 7. SAP time series. (a) original signal with FD-DA = 1.037 and (b) with background trend removed when FD-DA = 1.199. Removal of the trend has a strong impact on fractal dimension.

References

Bassingthwaite J. B., Raymond G. M. Evaluation of the dispersional analysis method for fractal time series. *Ann. Biomed. Eng.* **23**: 491 – 505, 1995

Yum M-K. Kim N-S. et al. Non-linear cardiac dynamics and morning dip: an unsound circadian rhythm. *Clin. Physiol.* **19**: 56 – 67, 1999

8. Fractal Dimension by Curve Length (FD-L)

This method also concerns observing the signal curve itself with the intention of defining a fractal dimension for the signal curve.

The calculation of FD-L is based on the measurement of the signal curve length using different measuring sticks, i.e., a study is made of how the length value of the curve changes with different measurement stick lengths. Figure 8 shows the results with two different measurement stick lengths used to evaluate the curve. The shorter the measurement stick length, the better one can achieve the real shape of the signal curve. Naturally enough one needs more sticks to achieve more accurate results. With a fractal curve the number of measurement sticks increases exponentially when the measurement stick length decreases. If $N(L)$ is the number of measurement sticks with the length L , fractal dimension may be defined as a limit:

$$FD = \lim_{L \rightarrow 0} \frac{\log N(L)}{\log(1/L)}.$$

In practice this limit cannot be calculated but when $\log N(L)$ is expressed as a function of $\log(1/L)$ and a regression line is plotted the slope of the line gives the fractal dimension value.

FD-L in practice

The number of measurement sticks increases exponentially as the length of the measurement stick decreases only up to a certain limit. When even the smallest details of the curve have been reached the exponential law no longer applies. That is why we must concentrate on this area in the log-log representation when fitting the regression line in

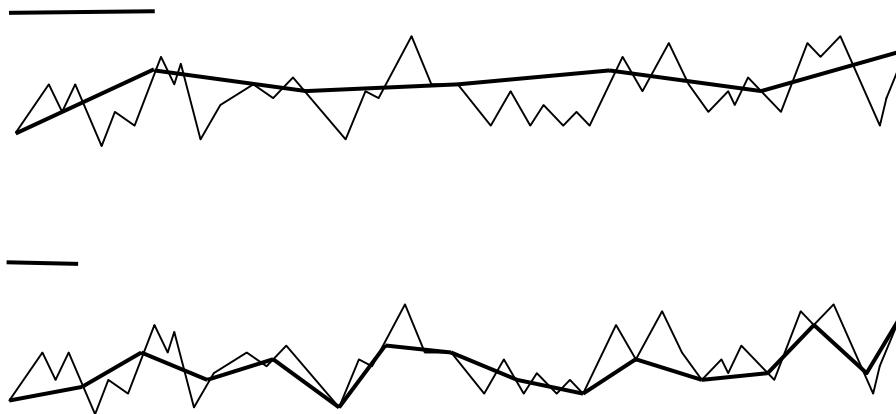


Figure 8. Signal curve (thin line) has been evaluated using a measure stick (thick line). When the measure stick is long (top curve) 6 sticks are required. If the length of the measure stick is halved, 15 sticks are required. The shorter the stick, the better the resemblance to the real curve.

the calculation of fractal dimension. For example with RRI interval this limit is usually the period between two beats but for more slowly changing signals and with tighter sampling (i.e. data points closer to each other) the limit might be much higher. FD-L values are >1 . In order for us to be able to compare fractal dimension values of different signals, we must first normalize the values. The most natural method is based on the following normalization

$$x^*(i) = (x(i) - \langle x \rangle) / SD(x),$$

where $\langle x \rangle$ is the mean of the signal and $SD(x)$ its standard deviation.

Benefits of FD-L

- The method poses no special characteristics for the dynamic behaviour of the system.
- Not particularly sensitive to slow trends of the signal.

Shortcomings of FD-L

- Sensitive to noise if the magnitude of noise is comparable to the magnitude of the smallest "real" changes of the signal.

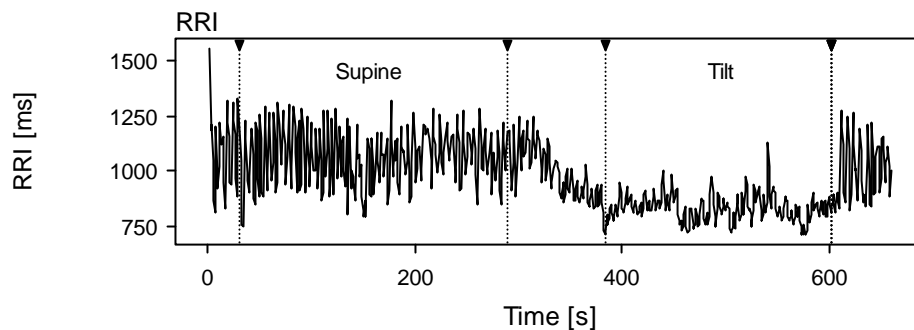


Figure 9. RRI time series measured in a tilt table test. Patient is positioned horizontally in the left side part of the diagram with $FD-L = 1.923$ and vertically in the right side of the diagram with $FD-L = 1.342$.

References

Chau N. P., Chanudet X., Bauduceau B., Gautier D., Larroque P. Fractal dimension of heart rate and blood pressure in healthy subjects and in diabetic subjects. *Blood Pressure* **2**: 101 – 107, 1993

Gough N. A. J. Fractal analysis of foetal heart rate variability. *Physiol. Meas.* **14**: 309 – 315, 1993

9. Similarity of Distribution (SOD)

This method is suited for measuring the stability of time series. If the time series is clearly unstable, i.e. it drifts in the systems parameter space (in which case the system changes considerably as a function of time), one must be very careful with the use of different linear and nonlinear analysis methods. SOD may also reveal whether a system is clearly stochastic.

SOD is calculated as follows:

- (i) a period starting from the moment t and with the length W is selected from the time series
- (ii) the amplitude distribution of this period is calculated using a suitable scale division
- (iii) a probability distribution $p_i(t)$ is made from the amplitude distribution by dividing it with the total integral of the distribution; $i = 1 \dots h$, h being the number of divisional steps
- (iv) a similar probability distribution for another period with the length W starting from the moment $t+\tau$, i.e. $p_i(t+\tau)$ is calculated
- (v) the auto correlation of these two probability distributions is calculated with

$$SOD(t) = \sum_{i=1}^h p_i(t)p_i(t+\tau)$$

- (vi) the above described calculation is performed for all data points of the time series. The result is SOD as a function of time.

SOD can have values between 0 .. 1. SOD is close to zero if the probability distributions are very wide as is the case typically for stochastic systems or if the distributions under comparison differ greatly as is the case for an unstable system. On the other hand if the distributions have maximums at approximately the same point, $SOD \approx 1$, in which case the system is stable and its behaviour is well predictable.

SOD in practice

When calculating distributions the selection of the shift parameter value τ has a great impact on how fast changes in the stability of the time series we are able to detect. Also window width W has an impact on the time resolution but if W is very short, the statistical relevance for singular SOD values begins to decrease strongly. In practise we must try to achieve the optimal compromise between these two parameters. In addition we must take into account that when calculating distributions the scale division is chosen so that the distributions contain enough details.

Benefits of SOD

- The method is a general test for stability and poses no special demands on the system.

Shortcomings of SOD

- Selection of parameters may strongly affect the results.
- Does not distinguish well very unstationary and stochastic time series from each other.

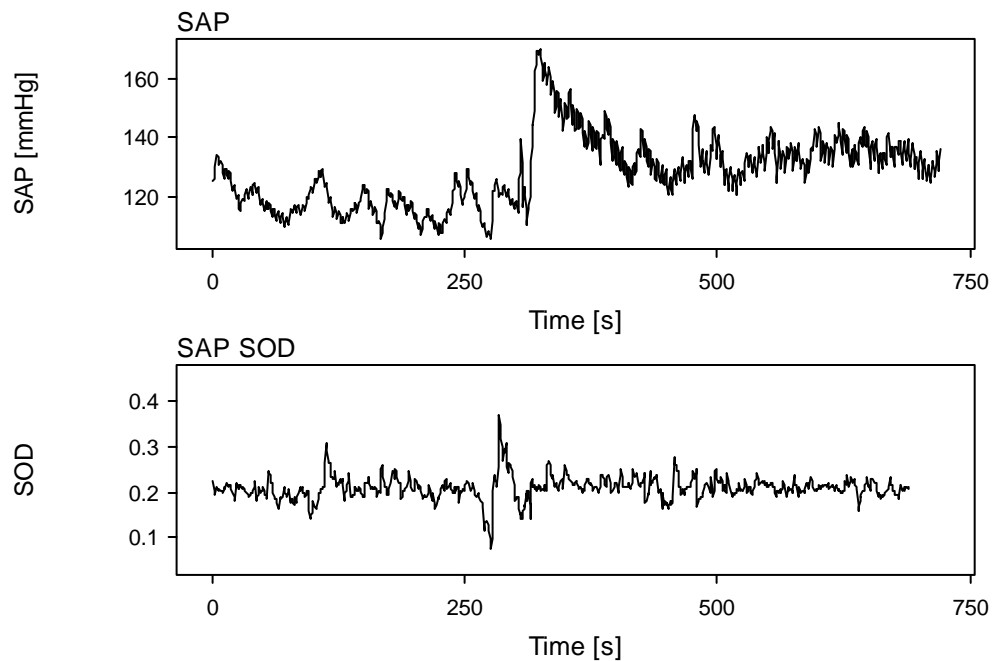


Figure 10. RRI and SOD time series made with the patient first laying down and then rising to a standing position. Even though SAP changes clearly, SOD does not change drastically. In standing position the shape of the probability distribution of SAP signal does not change. Calculations are based on a 30 second time window, which has been shifted forward in 15 second steps.

References

Zochowski M. Autocorrelations of R-R distributions as a measure of heart variability. *Phys. Rev. E* **56**: 3725 – 3727, 1997

10. Return Map (RM)

A dynamic system may be described mathematically in most cases by a differential equation group. If the variables get values only at specific discrete moments in time, as is the case with RRI time series, the differential equations may be replaced with discrete equations. The definition in general form is thus as follows:

$$\begin{aligned}x_{i+1} &= F(x_i, y_i, z_i, \dots) \\y_{i+1} &= G(x_i, y_i, z_i, \dots) \\z_{i+1} &= H(x_i, y_i, z_i, \dots) \\&\vdots\end{aligned}$$

where x, y, z etc. are dynamic variables of the system and F, G, H etc. are functions which define the dynamics. Usually these functions are not known but we may try to solve them by examining the measured time series. If there is only a single variable, the group can be expressed in a simpler form as

$$x_{i+1} = F(x_i).$$

Because this expression binds the new value x_{i+1} of the variable to its predecessor value x_i , we can solve function F in principle by plotting the points of the time series in a xy-graph with the previous value on horizontal axis and with the next value on vertical axis. This kind of a graph is called return map or Poincare mapping. If the description is exactly as defined above and there are enough data points, the method reveals the shape of the function F . Figure 11 shows a time series of a so called logistic map and its return map. Because the mapping is one-dimensional, i.e. it has only one dynamic variable, the return mapping reveals directly the function F , which in this particular case is a parabola.

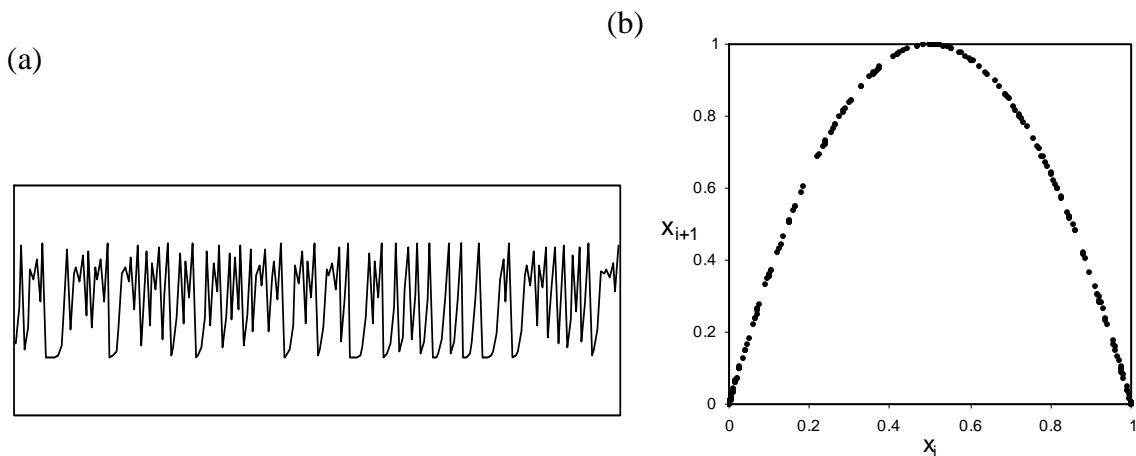


Figure 11. Time series (a) produced by the logistical description $x_{i+1} = 4x_i(1-x_i)$ and (b) return map made from (a).

RM in practice

If the dynamic system behind the time series is not one-dimensional, i.e. more than one variable have an impact on the system, a return map formed on the basis of a single measured variable naturally can not solve functions F , G etc. Even in such a case the single variable return map may prove to be useful although it is a some type of projection of the multidimensional system into a single dimension. Figure 12 shows the return map of RRI time series. The points are typically scattered to form an ellipsoid but the points could just as well form quite complex structures. When the points form an ellipsoid type area, its shape can be characterized by two distribution values: distribution in the direction of the diagonal, i.e. SD2 and parallel to this, i.e. SD1. These distribution values are by nature like variability quantities, since they measure the movement of the system in phase space.

Benefits of RM

- The method itself makes no assumptions regarding the system but its limitations with regard to multidimensional system description should be noted.

Shortcomings of RM

- Deviation parameters SD1 and SD2 do not necessarily reflect the real shape of the return map data set group very accurately.

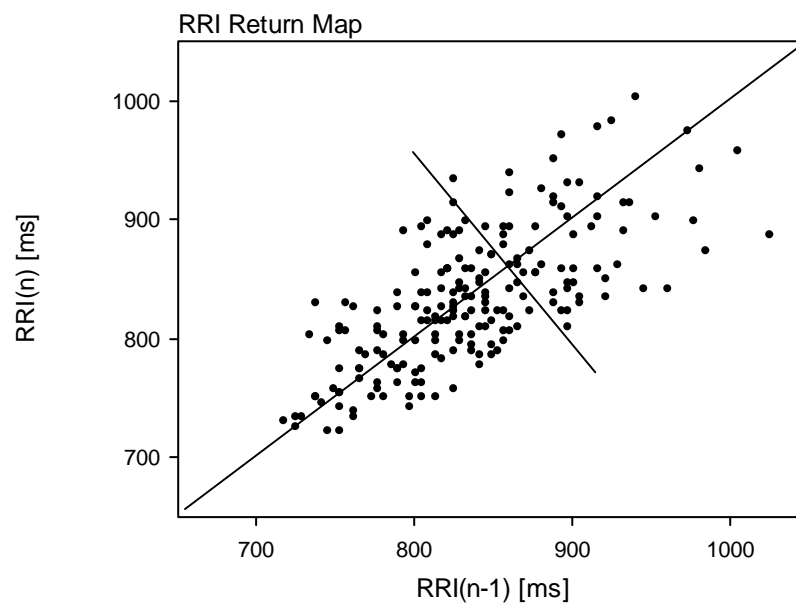


Figure 12. Return map made from a RRI time series. Deviation along the diagonal SD2 = 111 ms and perpendicular to this SD1 = 34 ms.

References

Huikuri H. V. et al. Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction. *Am. J. Cardiol.* **77**: 1073 – 1077, 1996

Woo M. A., Stevenson W. G., Moser D. K., Middlekauf H. R. Complex heart rate variability and serum norepinephrine levels in patients with advanced heart failure. *J. Am. Coll. Cardiol.* **23**: 565 – 569, 1994

11. Detrended Fluctuation Analysis (DFA)

DFA is used to measure the internal correlations of the time series. It is typical for biological systems that correlations are different for different time scales and DFA presents a possibility to characterize this as a function of correlation distance.

To calculate DFA we must first form an integrated version of the original time series $\{x(i)\}$, $i = 1 \dots N$, which gives us

$$y(k) = \sum_{i=1}^k (x(i) - \langle x \rangle),$$

where $\langle x \rangle$ is the mean of the original time series and $k = 1 \dots N$. Next we divide the time series $\{y(k)\}$ into equally spaced segments with the length n . For each segment we calculate separately the local trend by fitting a regression line $y_n(k)$ into the segment. The RMS (root-mean-square) fluctuation of the integrated time series is calculated by removing in each segment the linear trend of the segment, thus

$$DFA(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2}.$$

In the summing we must take into account that when the index k is stepped, $y_n(k)$ must be updated when moving into the next segment. DFA is calculated for several different segment length n values. Typically DFA increases when the segment length increases. If $\log DFA$ increases linearly as a function of $\log n$, the time series follows (fractal) scaling law and in this case the said slope α of the linear change defines the type of scaling. Different values of α correspond to the below listed basic time series types:

$0 < \alpha < 0.5$	small value followed most probably by larger value and vice versa
$\alpha = 0.5$	completely uncorrelated time series, i.e., white noise
$0.5 < \alpha < 1.0$	small value followed most likely by a small value and large value followed most likely by a large value
$\alpha = 1.0$	$1/f$ type noise
$1.0 < \alpha < 1.5$	noise of variable type
$\alpha = 1.5$	Brownian $1/f^2$ noise (integral of white noise)

When $\alpha > 1.0$, we are always looking at a signal with noise type characteristics, but which nevertheless contains correlations which are most clearly seen in the spectrum of the signal. The spectrum is not even but drops quickly as frequency increases.

DFA in practice

The most typical use for DFA is the analysis of long RRI time series. Heart beat has short range correlations which stem from e.g. the baroreflex mechanism, as well as long range correlations, which are related to the efforts to keep the variation of the beat cycle within certain limits. The measurement of the long range requires that the time series under investigation is very long, preferably at least a few hours so that statistical reliability would be at least reasonable.

Figure 13 shows a typical DFA for a long RRI time series. The limit for short and long range correlations is set typically to 10 or 11 beats (corresponds to 2.4 on the linear scale). The long range scaling exponent α_L for a healthy patient is ≈ 1 , which corresponds to $1/f$ behaviour. The short range scaling exponent α_S may vary but is usually between 0.5 – 1.5. Many factors affect it such as the functioning of the baroreflex mechanism, breathing modulation etc.

Benefits of DFA

- The method makes no special demands on the system.
- Enables distinct characterization of systems.

Shortcomings of DFA

- With long time series there exists always the possibility that the measured correlations are not only or are not at all a characteristic of the system but rather reflect instead environmental effects.

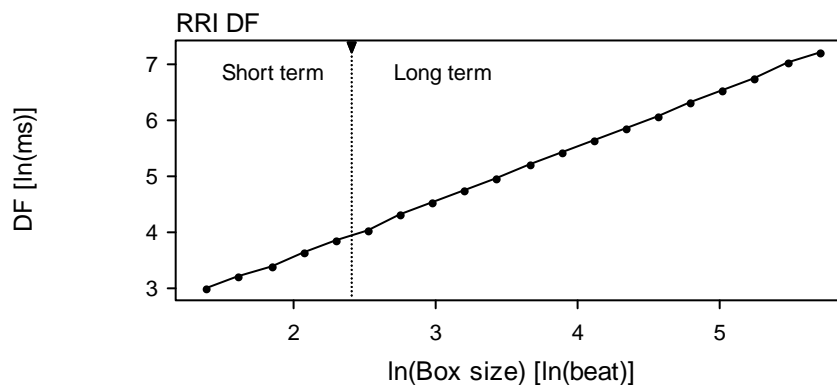


Figure 13. DFA analysis of an 8 hour long RRI time series. Scaling exponential ### corresponding to short range correlation $\alpha_S = 0.950$ and correspondingly for long range $\alpha_L = 0.998$.

References

Peng C-K., Mietus J., Hausdorff J. M., Havlin S., Stanley H. E., Goldberger A. L. Long-range anticorrelations and non-gaussian behaviour of the heartbeat. *Phys. Rev. Lett.* **70**: 1343 – 1346, 1993

Peng C-K., Havlin S., Stanley H. E., Goldberger A. L. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *CHAOS* **5**: 82 – 87, 1995

Iyengar N., Peng C-K., Morin R., Goldberger A. L., Lipsitz L. A. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *Am. J. Physiol.* **271** (Regulatory Integrative Comp. Physiol): R1078 – R1084, 1996

12. Spectrum Power-Law (SPL)

Long range correlations of the time series may also be analysed using the spectrum of the signal. In this case we study the most lowest frequency components of the spectrum and try to characterize its shape using simple exponential law.

If we presume that for a certain frequency spectrum we have $1/f^\beta$, the scaling exponent β can be calculated by presenting the spectrum on a log-log scale and by fitting a line over the desired frequency range. The slope of the line gives the exponent. The value of the exponent varies between 0 – 2. The border case 0 corresponds to a flat spectrum, i.e., white noise and the value 2 corresponds to Brownian noise.

SPL in practice

Usually the frequency range 0.0001 – 0.01 Hz of the spectrum is studied and this corresponds to an oscillation period of one minute to several hours. For the definition to make sense the time series must be several hours long. The spectrum is calculated almost without exception using FFT algorithm. Because the spectrum has a rather unsmooth form especially at the lowest frequencies, the use of some smoothing method is desirable. Replacing the regression line by a less sensitive fitting method may also improve the dependability of the end result.

Benefits of SPL

- The method is not based on the magnitude of the signal amplitude changes but on the special shape of the spectrum at low frequencies.

Shortcomings of SPL

- Often the spectrum cannot very well be described using exponential law, as the scaling exponential β no longer describes reality.

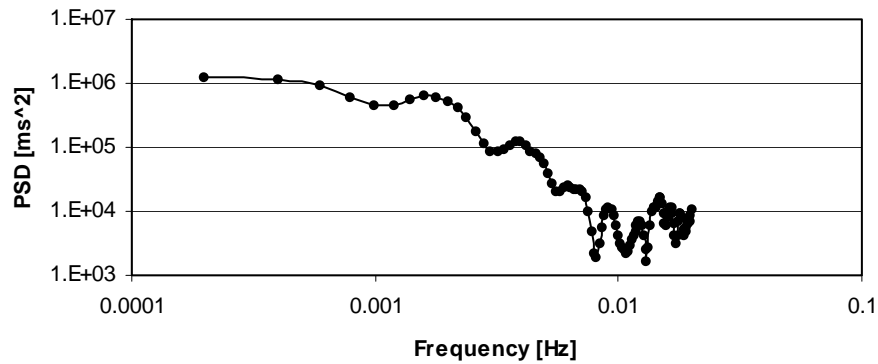


Figure 14. Spectrum calculated from a 4 hour long RRI time series in log-log scale. Typically the spectrum does not change linearly as a function of frequency but nevertheless a regression line can be fitted reasonably well between frequency range 0.0001 – 0.01 Hz, in which case the slope is 1.823. The result indicates that the time series can be described to consist mostly of Brownian noise.

References

Bigger J. T., Steinman R. C., Rolnitzky L. M., Fleiss J. L., Albrecht P., Cohen R. J. Power law behaviour of RR-interval variability on healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. *Circulation* **93**: 2142 – 2151, 1996

Iyengar N., Peng C-K., Morin R., Goldberger A. L., Lipsitz L. A. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *Am. J. Physiol.* **271** (Regulatory Integrative Comp. Physiol): R1078 – R1084, 1996

13. Spectral Entropy (SEn)

The nature of the spectrum may also be characterized with a single statistic by interpreting spectral density as a probability distribution. Then we may define the so called Shannon's entropy as follows

$$SEn = -\sum_f p_f \ln p_f ,$$

where the probabilities p_f are defined as normalized spectral densities (PSD), thus

$$p_f = \frac{PSD(f)}{\sum_{f=0}^{f_{\max}} PSD(f)} .$$

In practice spectral densities are calculated normally using the FFT routine but it is also possible to use the AR, in which case however the selection of the proper mode order may prove to be difficult. In order for the calculated entropy quantity to be independent of the number of the frequency components of spectral density (since spectrum is in practice always discrete), entropy is commonly normalized with the logarithm of the number of components, in which case SEn values have a range between 0 .. 1. SEn reflects the shape of the signal spectrum. If the signal contains only one or just several distinct separate frequency components, SEn will have low values, but on the other hand for wide band noise SEn has a high value.

SEn in practice

In practice SEn is easily and quickly calculated using the FFT routine and it does not require especially long time series but with very short time series its statistical reliability is naturally rather poor. Because SEn characterises the general shape of the spectrum, it cannot be used to distinguish between chaos and noise. If the signal is nonstationary by character the low frequency components may dominate the SEn quantity and its useability may be limited especially regarding RRI and SAP type signals, but it may be very useful when characterizing e.g. EEG time series.

Benefits of SEn

- Easy to calculate.
- Does not require long time series.

Shortcomings of SEn

- A rather unspecific statistic.

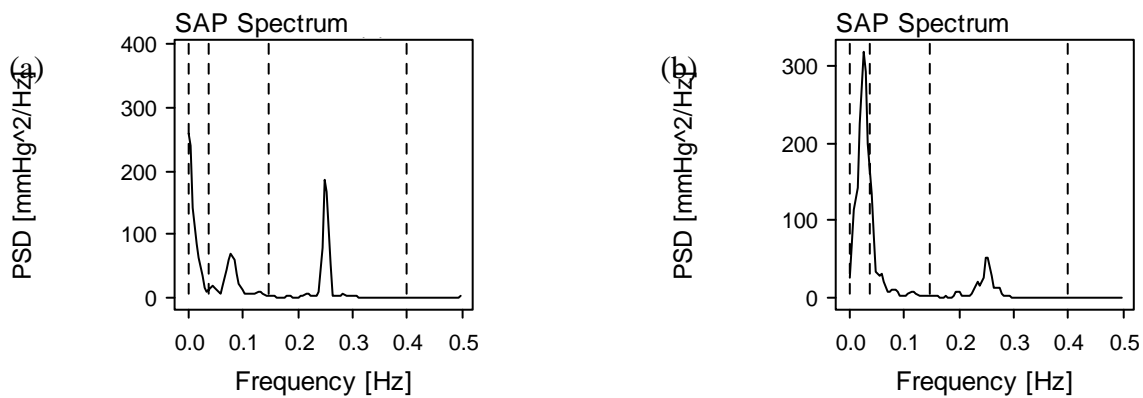


Figure 15. RRI signal spectrum measured from a test person who has been (a) breathing in phase with a metronome when $SEn = 0.647$, and spontaneously (b), when $SEn = 0.731$. Spectral entropy has been calculated over LF and HF bands in order to prevent the effects from the lowest frequency components which are caused probably by unstationarity.

References

Rezek I. A., Roberts S. J. Stochastic complexity measures for physiological signal analysis. *IEEE Trans. Biom. Eng.* **45**: 1186 – 1191, 1998

Steyn-Ross D. A., Steyn-Ross M. L., Wilcocks L. C. Toward a theory of the general-anesthetic-induced phase transition of the cerebral cortex. II. Numerical simulations, spectral entropy, and correlation times. *Phys. Rev E* **64**: 011918, 2001

14. Symbolic Dynamics (SymDyn)

The basic idea of this method is to replace the original time series with a much simpler and rougher replacement, which however retains the essential dynamic characteristics of the original time series. This is done by converting the time series into a string of symbols, of which there are only a limited number of. By this way we can reduce the study of dynamics into handling of the symbol string. Naturally enough we lose most of the information contained in the original time series but nevertheless we try to retain the most important characteristics intact.

Conversion of a time series into a symbol string may be done using several methods. One of them is described in figure 15. The signal is divided into two or more value ranges depending on how many symbols we wish to utilize. Value ranges may be absolute bands or based on signal averages or standard deviation (SD). If we for example have 4 different symbols, we may use the following bands:

- A** $\text{signal} \leq \text{average} - \text{SD}$
- B** $\text{average} - \text{SD} < \text{signal} \leq \text{average}$
- C** $\text{average} < \text{signal} \leq \text{average} + \text{SD}$
- D** $\text{signal} > \text{average} + \text{SD}$

After the bands have been selected the time series can be converted into a symbol string. The next step involves grouping the symbols in the string into *words*. A word is always formed by stepping forward one step in the symbol string. If we choose a word length of 3 and we have 4 different symbols, we get altogether $4 \cdot 4 \cdot 4 = 64$ different words. Each word corresponds to a specific graphical representation, which has at least a rough connection to the dynamics in the background.

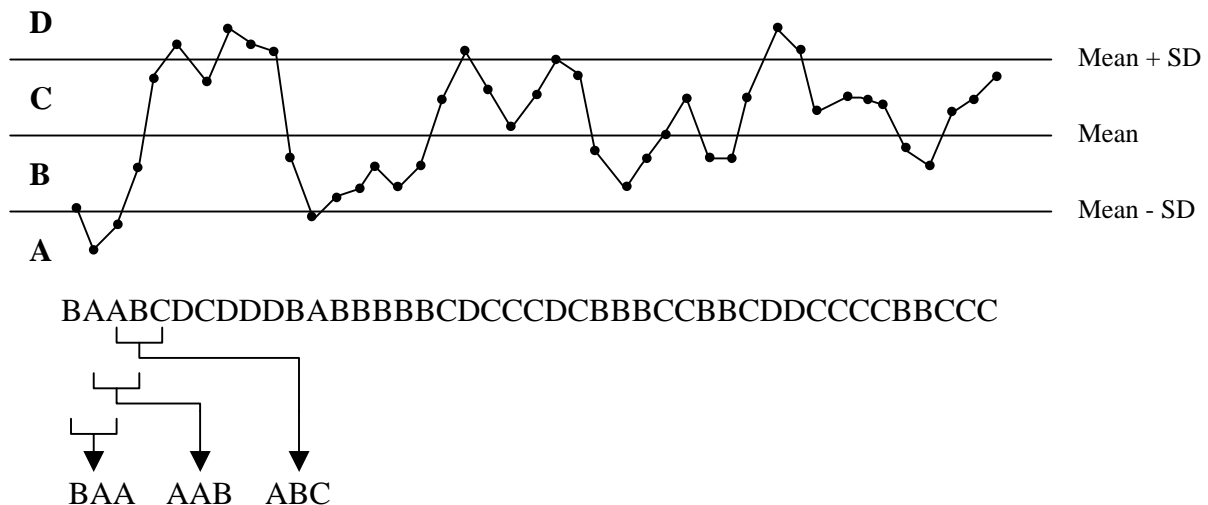


Figure 16. Conversion of the time series into a symbol row and grouping of the symbol row into words.

Different words do not have the same probability because dynamics favours certain words. In order to be able to study this effect more closely a distribution of the words is made. The words must be arranged in some particular way to make the representation of the distribution possible. If we have 4 symbols and the word length is 3, the words may be arranged in the following way:

- | | | | | | |
|----|-----|----|-----|-----|-----|
| 1. | AAA | 5. | ABA | 9. | ACA |
| 2. | BAA | 6. | BBA | ... | |
| 3. | CAA | 7. | CBA | 63. | CDD |
| 4. | DAA | 8. | DBA | 64. | DDD |

Figure 16 shows a typical RRI time series as a word probability distribution. Certain words have a clearly higher probability of existence than others. The shape of the distribution may itself act as a basis of further analysis but is also possible to measure the order related to the distribution. The simplest such measure is the so called Shannon's entropy

$$S = -\sum_{i=1}^n p_i \ln p_i ,$$

where p_i is the probability of each word and n the number of different words. Entropy may be normalized by displaying it as either bits/word or by dividing with the corresponding number of binary words, in which case $0 \leq S \leq 1$. The construction of the distribution is also described by the so-called forbidden words, i.e. the relative share of words with a probability < 0.001 compared to all words.

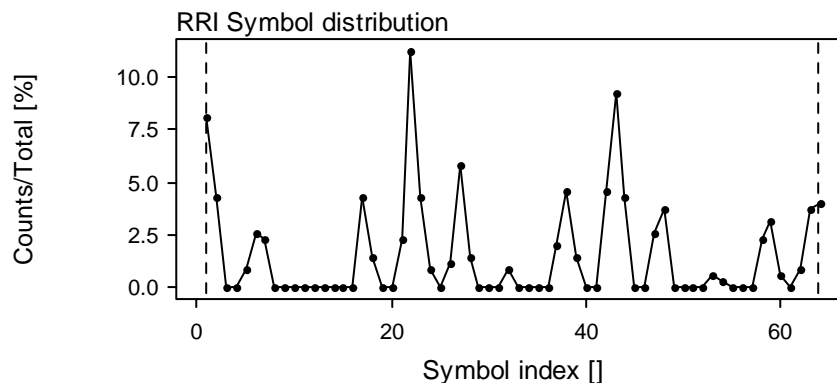


Figure 17. Probability distribution of a 5 minute long RRI time series. Distribution is based on the use of 4 different symbols, word length is 3. Shannons entropy = 4.543 bits/words, normalized 0.757, percentage of forbidden words is 52 %. The 6 most probable words are BBB, CCC, AAA, CCB, BBC and BCC. Channels are classified according to distribution.

SymDyn in practice

The number of different symbols is usually 2 - 6 and the length of the words 2 - 4. Using distribution in the selection of value bands of the signal is preferably since in this case the results are less dependent on absolute changes. When selecting parameters one must remember that the time series must be long enough for the distribution to contain enough words. For example with RRI time series it is considered normal that the distribution is relatively smooth, i.e. the share of forbidden words is low. In pathological cases the distribution is dominated by few singular words.

Benefits of SymDyn

- Independent of the dynamics of the system.
- Fast calculations.

Shortcomings of SymDyn

- Sensitive to trends in time series because channel definitions are based on absolute values.
- Interpretation of distribution shape is difficult.

References

Voss A., Kurths J., Kleiner H. J., Witt A., Wessel N. Improved analysis of heart rate variability by methods of nonlinear dynamics. *J. Electrocardiol.* **28**: 81 – 88, 1995

Voss A. et al. The application of methods on non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. *Cardiovascular Research* **31**: 419 – 433, 1996

Palazzolo J. A., Estafanous F. G., Murray P. A. Entropy measures of heart rate variation in conscious dogs. *Am. J. Physiol* **274** (Heart. Circ. Physiol.): H1099 –H1105, 1998

15. Lempel-Ziv Entropy (LZEn)

The complexity of a symbol string may also be measured directly without the use of distribution by defining the so called Kolmogorov entropy which states that complexity is equal to the shortest possible computer program (typically an algorithm) which can produce the string, with the length given in bits. If the program is short, complexity is low and vice versa. This definition makes sense intuitively but is impossible to calculate without special limitations.

When the symbol string described in the previous chapter is formed using only two symbols we actually get a binary string, for which we select the characters 0 and 1 as symbols. This kind of a string may be constructed using only two operations, which are the insertion of a singular symbol and the copying of a symbol substring. We can then calculate the Lempel-Ziv complexity of a string. For example, the following simple string

00000 ...

can be constructed with two operations: first we insert one zero and the rest are made by copying the zero. The complexity of the string is 2. In similar fashion the string

01 01 01 ...

can be constructed by inserting a 0 and a 1 and the rest of the string by copying the substring 01. The complexity of this string is 3. In similar fashion we can define the complexity of any binary string, although the required algorithm is a bit complex and will not be shown in more detail here.

It can be shown that the complexity of a random string containing K pcs of different symbols and with a length n is

$$b(n) = \frac{hn}{\log_K(n)},$$

where h is the so called normalized source entropy

$$h = \frac{-1}{\log(n)} \sum_{i=1}^n p_i \log p_i,$$

and the quantity p_i is the symbols probability in the string. When the string contains only two symbols and both have the same probability, $h = 1$ and $b(n) = n/\log_2(n)$. When the complexity of the string under study is normalized with the quantity $b(n)$, we get an entropy quantity LZEn, $0 \leq \text{LZEn} \leq 1$ which is independent of the length of the string.

LZEn in practice

The binary string of a time series is usually formed by giving all values below the average the value 0 and all values above the average the value 1. LZEn can be calculated for as short a string as desired but naturally for short strings the statistical relevance is poor. It should also be noted that normalized entropy is independent of the length of the string only when the length is >1000 .

Benefits of LZEn

- This method is a completely general characterisation method of time series.

Shortcomings of LZE

- Construction of a symbol string from a time series is really successful only when the signal average remains same.

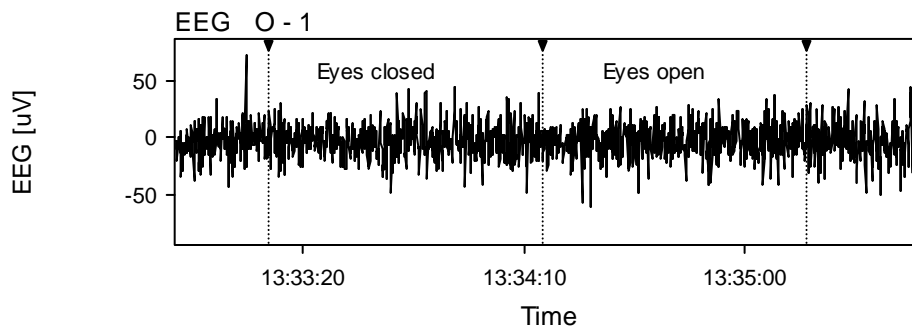


Figure 18. EEG signal, coupling O1-A2. When the eyes of the patient are closed, LZEn = 0.889 and when the eyes are open, LZEn = 0.940. Changes in the activity level of the brain appear as a more complex signal.

References

Kaspar F., Schuster H. G. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A* **36**: 842 – 848, 1987

Zhang X-S., Roy R. J. Predicting movement during anaesthesia by complexity analysis of electroencephalograms. *Med. Biol. Eng. Comput.* **37**: 327 – 334, 1999

16. Stationarity (StatAv)

There are many methods available for evaluating the stationarity of a signal depending on how we want to define stationarity. Stationarity can mean that there is no shifting in the base level of the signal or that the amplitude distribution, spectrum or autocorrelation function of the signal does not change as a function of time. More generally we can say that a signal is stationary if the parameters which define the working point of the system remain constant. In practice this evaluation is extremely difficult.

An example of a simple stationarity test which mainly tests changes in base level is as follows: The signal is divided into segments of suitable length and in each segment the signal average is calculated. When the standard deviation of these averages is divided by the standard deviation of the complete signal we get a measure of stationarity, StatAv. If StatAv is small, the signal is stationary.

StatAv in practice

The length of the segment is chosen so that it is not too long in order for the local changes to be detectable but it cannot be too short either in order to prevent the averages of the segments from varying too much. For RRI signal a good choice is to use 20 datapoints/segment. If $\text{StatAv} < 0.3$ the signal can be considered reasonably stationary.

Benefits of StatAv

- Calculations are very simple and fast.

Shortcomings of StatAv

- Sensitive mainly to baseline changes.

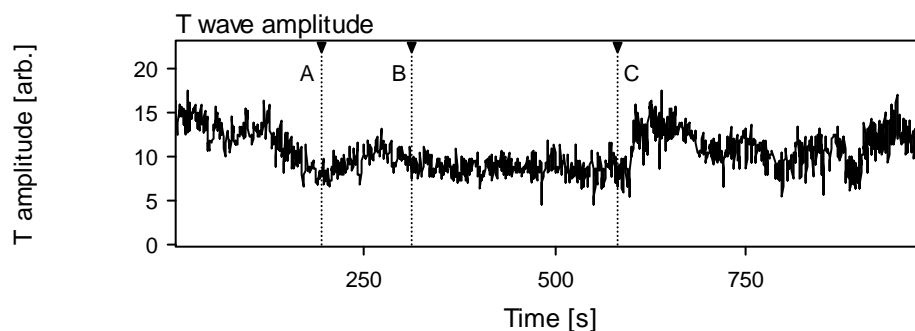


Figure 19. Amplitude of T-wave as a function of time. Between A-C $\text{StatAv} = 0.604$, but between B-C $\text{StatAv} = 0.293$. The method is sensitive to changes in baseline.

References

Palazzolo J. A., Estafanous F. G., Murray P. A. Entropy measures of heart rate variation in conscious dogs. *Am. J. Physiol* **274** (Heart. Circ. Physiol.): H1099 –H1105, 1998

17. Fractional Spectral Radius (FSR)

Characteristics of the system behind the signal may be studied in more detail by forming a local construction matrix by using pseudo space phase vectors which have been described previously. The vectors are formed from the measured time series using delayed coordinates. The construction matrix describes at each point of the path how the system is behaving. In practise it is easier to analyse the system by looking at a construction matrix which has been averaged over the path of the system (or by selecting a suitable time slice if it is desired to study changes as a function of time). We then get the so-called covariance matrix

$$\Xi = \begin{pmatrix} \sum_{i=1}^N x_i x_i & \sum_{i=1}^N x_i x_{i+1} & \cdots & \sum_{i=1}^N x_i x_{i+m-1} \\ \sum_{i=1}^N x_{i+1} x_i & \ddots & & \vdots \\ \vdots & & & \\ \sum_{i=1}^N x_{i+m-1} x_i & \cdots & & \sum_{i=1}^N x_{i+m-1} x_{i+m-1} \end{pmatrix},$$

where $\{x_i\}$ is the time series under study, N is the number of data points and m is the sinking dimension. Covariance matrix describes how the system behaves along the path as an average but because it contains m^2 numbers the use of the matrix as such for the analysis of time series is tedious. By calculating the eigenvectors and eigenvalues of the matrix we can find a coordinate system in which the description of the system becomes as simple as possible. Eigenvalues σ_i , $i = 1 \dots m$, on the other hand present a possibility to define how strong the movement in each coordinate direction is. Eigenvalues, which are zero or close to zero correspond to coordinates which are irrelevant from the viewpoint of the mapping. This means that the selected embedding dimension m is “too large”, the system can be described by fewer variables. By studying the distribution of the eigenvalues we may now also evaluate the real dimension of the system and through it also its complexity.

Because there are also m eigenvalues it is still necessary to simplify the approach. By arranging the eigenvalues in descending order (because covariance is symmetrical all of its eigenvalues are either positive or zero) we are able to define a new measure for the complexity of a system

$$FSR(j) = \frac{\sum_{i=1}^j \sigma_i}{\sum_{i=1}^m \sigma_i}.$$

this being the so called spectral radius of eigenvalues.

FSR in practice

If the system is stationary and does not contain noise at all, the number of nonzero eigenvalues is exactly the same as the dimension of the system. In reality the situation is far from being so clear, instead all of the eigenvalues differ from zero due to noise, which is always present in the signal. By selecting $j = 1$, i.e. by calculating the quantity $\text{FSR}(1)$, we obtain information regarding how much “power” there is in the other directions compared to the direction defined by the largest and at the same time the most important eigenvalue. If a system has numerous independent variables, $\text{FSR}(1)$ approaches zero and if only several variables have any meaning, $\text{FSR}(1)$ approaches one. So, in this way $\text{FSR}(1)$ is a measure of the systems complexity (note that the value 1 corresponds to a system with low dimensionality and low complexity, which is the reverse compared to entropy quantities).

Benefits of FSR

- No special requirements as to the characteristics of the system under investigation.
- Despite of complex calculations the method is quite fast.

Shortcomings of FSR

- The selection of embedding dimension and especially delay are somewhat critical.

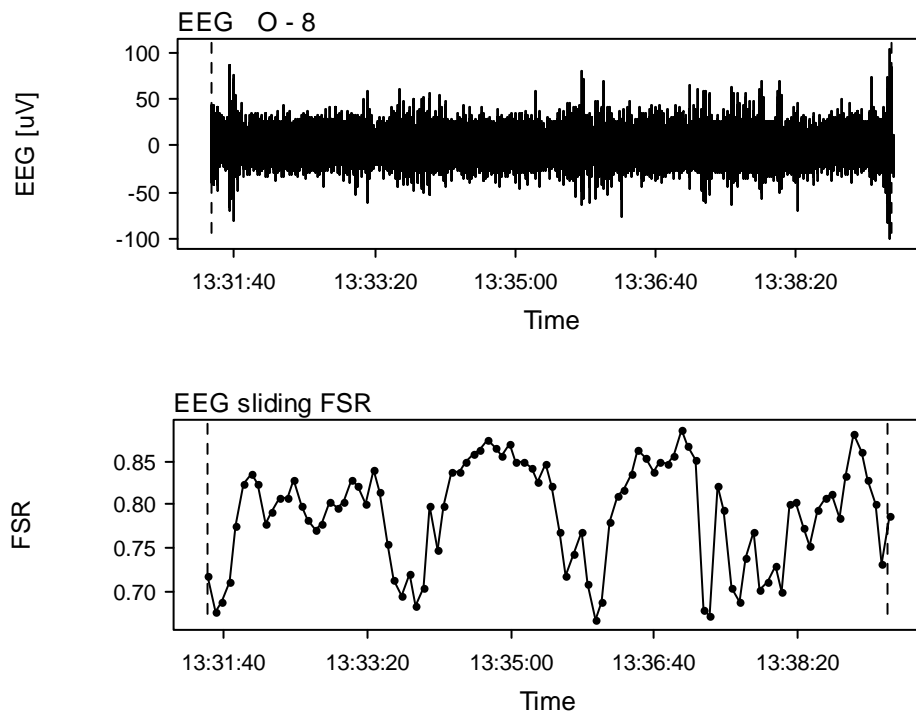


Figure 20. EEG signal and a floating FSR(1) calculated from the EEG signal. Calculations were based on a time window of 10 seconds and the window was shifted forward in 5 second steps, embedding dimension is 25 and time delay 40 ms.

References

Rezek I. A., Roberts S. J. Stochastic complexity measures for physiological signal analysis. *IEEE Trans. Biom. Eng.* **45**: 1186 – 1191, 1998

Broomhead D. S., King G. P. Extracting qualitative dynamics from experimental data. *Physica* **20D**: 217-236, 1986