Information Sciences 382–383 (2017) 359–373

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Data-driven detrending of nonstationary fractal time series with echo state networks

Enrico Maiorino [a], Filippo Maria Bianchi [b], Lorenzo Livi [c,*], Antonello Rizzi [a], Alireza Sadeghian [d]

[a] Dept. of Information Engineering, Electronics, and Telecommunications, SAPIENZA University of Rome, Italy
[b] Machine Learning Group, Department of Physics and Technology, University of Tromsø – The Arctic University of Norway, Norway
[c] Dept. of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK
[d] Dept. of Computer Science, Ryerson University, Toronto, Canada

ARTICLE INFO

Article history:
Received 24 October 2015
Revised 3 October 2016
Accepted 13 December 2016
Available online 14 December 2016

Keywords:
Fractal time series
Multiscaling
Fluctuation analysis
Detrending
Echo state network
Prediction

ABSTRACT

In this paper, we propose a novel data-driven approach for removing trends (detrending) from nonstationary, fractal and multifractal time series. We consider real-valued time series relative to measurements of an underlying dynamical system that evolves through time. We assume that such a dynamical process is predictable to a certain degree by means of a class of recurrent networks called Echo State Network (ESN), which are capable to model a generic dynamical process. In order to isolate the superimposed (multi)fractal component of interest, we define a data-driven filter by leveraging on the ESN prediction capability to identify the trend component of a given input time series. Specifically, the (estimated) trend is removed from the original time series and the residual signal is analyzed with the multifractal detrended fluctuation analysis procedure to verify the correctness of the detrending procedure. In order to demonstrate the effectiveness of the proposed technique, we consider several synthetic time series consisting of different types of trends and fractal noise components with known characteristics. We also process a real-world dataset, the sunspot time series, which is well-known for its multifractal features and has recently gained attention in the complex systems field. Results demonstrate the validity and generality of the proposed detrending method based on ESNs.

1. Introduction

Memory is one of the most interesting aspects of many processes in Nature and society [19]. In order to characterize and predict a system with memory, it is necessary to keep into account its past history. Memory can be quantified in different ways, depending on the particular features and effects of interest. One of the most common approaches in the study of real-valued time series is the analysis of the autocorrelation function. In such a linear setting, the extent of memory can be roughly quantified through the decay of the autocorrelation function, which indicates the characteristic time scales at which the series remains correlated. When the decay is exponential, the series is said to manifest *short-term memory* and the influence of the past to the current state is limited in time. Instead, if the decay follows a power-law, then there is no

* Corresponding author.
E-mail addresses: enrico.maiorino@uniroma1.it (E. Maiorino), filippomaria.bianchi@uniroma1.it (F.M. Bianchi), l.livi@exeter.ac.uk (L. Livi), antonello.rizzi@uniroma1.it (A. Rizzi), asadeghi@ryerson.ca (A. Sadeghian).

http://dx.doi.org/10.1016/j.ins.2016.12.015
0020-0255/

characteristic scale in the autocorrelation, i.e., the influences of the past have no cut-off. In this case, a time series is said to manifest *long-term memory* or *long-term correlation* (LTC) and the strength of this correlation is referred to as degree of *persistence* of the generating stochastic process. Persistence of a stochastic process [41] is quantified by the self-similarity coefficient of the process' fluctuations, called Hurst exponent $H \in [0, 1]$. A straightforward numeric approach to estimate the Hurst coefficient is the Fluctuation Analysis (FA), which evaluates the slope of the fluctuations scaling function $F(s)$. This function is in turn calculated by dividing the integrated time series in segments of equal sizes s and evaluating the root mean square difference between their extremal points. When the process corresponds to uncorrelated noise (e.g., white Gaussian noise), then the value of H is 0.5, whereas if the process is persistent (correlated) or antipersistent (anticorrelated) it will be respectively greater than or less than 0.5.

However, conventional methods employed to analyze the LTC properties of a time series (e.g., FA, spectral analysis, R/S analysis [3,43,44]) are misleading when such time series are non-stationary [8]. In fact, in many cases a process is driven by underlying trends [25], which operate at specific time scales, like seasons in the analysis of data related to a natural phenomenon and days in financial market analysis. Usually, when investigating memory properties of a process, one is interested in the fractal properties of the intrinsic fluctuations of such a process. Hence, to analyze the fluctuations of the stationary component of a time series, it is necessary to remove the non-stationary trend components. This can be done by employing one of the several methods proposed for this purpose, like detrended fluctuation analysis (DFA), detrended moving average, wavelet leaders [47], adaptive fractal analysis [40], and the so-called geometric-based approaches [17]. Notably, DFA has been shown to be successful in a broad range of applications [29,32,44]. The DFA has been generalized in the so-called Multifractal Detrended Fluctuation Analysis (MFDFA) [4,9,30,38], which accounts for the existence of multiple scaling exponents in the same data. DFA and the related variants remove trends from data by means of window-based local (polynomial) fittings. However, trends are often defined in terms of periodicities and/or fast-varying functions, resulting thus in a spurious detection of fractality [20]. For this reason, additional detrending methods are often used as a preprocessing step of the (MF-)DFA to single out these trends before the polynomial detrending takes place. In other research works, the local detrending step of DFA is modified or replaced with other ad-hoc methods [24,31,39].

The main problem with detrending lies in the difficulty of defining what exactly a trend is [48]. Local-fit based methods rely on the assumption that a trend is generally a slow-varying process, while the superimposed noise is a process characterized by higher frequencies. While this is often the case, it is still difficult to determine the right form and parameters of the fitting function without biasing the analysis. Moreover, window-based fitting algorithms are heavily influenced by the choice of the window sizes. In [48] a trend is defined as an intrinsically fitted monotonic function or a function in which there can be at most one extremum within a given data span. This method is not affected by border effects since it is not window-based. However, a problem with this definition is that it does not (fully) describe periodic trends in a consistent way. Chianca et al. [10] suggested to perform a detrending by applying a simple low-pass filter, in order to eliminate slow periodic trends from data. While this approach is suitable for systems with slow-varying trends, it is difficult to apply to more general cases, when the trends' frequencies span over a significant portion of the (power) spectrum. Another approach that has been demonstrated to be useful in the case of periodicities was proposed by Nagarajan [37]. As a first step, the signal is represented as a matrix, whose dimension has to be much larger than the number of frequency components of the periodic (or quasi-periodic) trends as shown by the power spectrum. The well-known singular value decomposition method is then applied to remove components related to large-magnitude eigenvalues, which correspond to the trend. Such a method, although interesting and mathematically well-founded, is very demanding in terms of computations and also assumes a deterministic form for trends.

In this work, we follow an approach similar to Wu et al. [48] and define a trend in a completely data-driven way. We consider the analyzed time series as a series of noisy measurements of an unknown dynamical process. We also assume that the dynamical process is predictable to a certain degree by means of a particular type of Recurrent Neural Network (RNN) called Echo State Network (ESN) [7,34]. RNNs have been shown to be able to predict the outcome of a number of dynamical processes [12]. In particular, a fundamental theorem formulated within the Neural Filtering framework, relates the number of neurons in a RNN hidden layer with the expected approximation accuracy of the estimated signal with respect to the true signal [33] of the process. Specifically, given a sufficiently large amount of processing units, a RNN that takes as input the measurement process is able to output an estimation that can be made as close as desired to the signal process, given its past input sequences. However, not all processes are predictable at the same level, as formally studied in [5,11], for instance. For example, chaotic processes are not predictable for long time-steps, while other deterministic systems, like a sinusoidal waveform, can be easily predicted. In a stochastic setting, instead, we note that white noise cannot be predicted at all, since the past observations do not convey any information about the future. On the other hand correlated noise signals, such as fractional Gaussian noise (fGn), are in theory partially predictable given the presence of memory in the process. To handle prediction problems of increasing difficulty, models characterized by a higher complexity or a larger amount of training data are required. In the case of ESNs, the complexity of the model is mainly determined by the properties and the size of its recurrent hidden layer. Here we propose to perform a data-driven detrending of nonstationary, fractal and multifractal time series by using ESNs acting as a filter. In this study, trends are the only form of nonstationarities that we consider. By means of ESNs, we predict the trend of a given input time series, which is always superimposed to the (multi)fractal component of interest. Such a trend is then removed from the original time series and the residual signal is analyzed with MFDFA in order to evaluate its scaling and (multi)fractal properties.