Journal of Hydrology 549 (2017) 715–730

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Temporal scaling phenomena in groundwater-floodplain systems using robust detrended fluctuation analysis

Abrar Habib [a,*], James P.R. Sorensen [b], John P. Bloomfield [b], Katie Muchan [c], Andrew J. Newell [b], Adrian P. Butler [a]

[a] Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK
[b] British Geological Survey, Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OX10 8BB, UK
[c] Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OXy 8BB, UK

ARTICLE INFO

Article history:
Received 23 January 2017
Received in revised form 12 April 2017
Accepted 14 April 2017
Available online 18 April 2017
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Stephen Worthington, Associate Editor

Keywords:
Robust detrended fluctuation analysis
Detrended fluctuation analysis
Fractal behaviour
Hurst Phenomenon
Time series analysis
High resolution hydrological data

ABSTRACT

In order to determine objectively the fractal behaviour of a time series, and to facilitate potential future attempts to assess model performance by incorporating fractal behaviour, a multi-order robust detrended fluctuation analysis (r-DFAn) procedure is developed herein. The r-DFAn procedure allows for robust and automated quantification of mono-fractal behaviour. The fractal behaviour is quantified with three parts: a global scaling exponent, crossovers, and local scaling exponents. The robustness of the r-DFAn procedure is established by the systematic use of robust regression, piecewise linear regression, Analysis of Covariance (ANCOVA) and Multiple Comparison Procedure to determine statistically significant scaling exponents and optimum crossover locations. The MATLAB code implementing the r-DFAn procedure has also been open sourced to enable reproducible results.

r-DFAn will be illustrated on a synthetic signal after which is used to analyse high-resolution hydrologic data; although the r-DFAn procedure is not limited to hydrological or geophysical time series. The hydrological data are 4 year-long datasets (January 2012 to January 2016) of 1-min groundwater level, river stage, groundwater and river temperature, and 15-min precipitation and air temperature, at Wallingford, UK. The datasets are analysed in both time and fractal domains. The study area is a shallow riparian aquifer in hydraulic connection to River Thames, which traverses the site. The unusually high resolution datasets, along with the responsive nature of the aquifer, enable detailed examination of the various data and their interconnections in both time- and fractal-domains.

1. Introduction

In the field of hydrology, the onset of the study of fractal behaviour of hydrological time series is marked with Hurst's investigation of the storage capacity of the Aswan High Dam in Egypt in 1951 (Hurst, 1951, 1956). This sparked further investigation of what later came to be known as the 'Hurst Phenomenon' (Hurst, 1951). The initial mathematical representation of the Hurst Phenomenon was described in terms of range, standard deviation and the number of samples considered. However, this relationship evolved into: $E\{X(T)\} \propto T^H$ with $H \neq 0.5$, where $X(T)$ is the aggregated series at scale T and H is the Hurst Exponent (Bras and Rodriguez-Iturbe, 1985). Of course, the relationship follows a power law and is linearly related to other measures of fractal behaviour such as the power-law exponent of the spectral density estimate and the scaling exponent α determined by detrended fluctuation analysis.

The mathematician Benoit Mandelbrot introduced a different concept to the Hurst Phenomenon that infuses the self-similarity property of fractals with that of Hurst (Mandelbrot, 1982). Mandelbrot introduced the term 'fractional noises' in 1968 to unify the different terms developed over time and across the different fields that describe series with long-term interdependence (Mandelbrot and Van Ness, 1968). Hence the term 'Fractal behaviour' will be used in this paper to refer to the 'Hurst Phenomenon' and 'long-term memory'; terms which are more common to hydrologists.

Evidently, fractal behaviour of time series has been investigated in various fields and a wide variety of techniques have been used to quantify it. Fractal behaviour has been studied in the fields of, amongst others, pharmacology: long-term correlations of DNA (Peng et al., 1994); cardiology: non-stationary heart beat time series (Peng et al., 1995); earth sciences: ocean wave height (Ozger,

* Corresponding author.
E-mail addresses: abrar.habib11@imperial.ac.uk, abr.habib@gmail.com (A. Habib).

http://dx.doi.org/10.1016/j.jhydrol.2017.04.034
0022-1694/

2011), temperature (Koscielny-Bunde et al., 1996) and seismicity (Alvarez-Ramirez et al., 2011); traffic control: traffic speeds time series (Shang et al., 2008), in marine transportation (Chen et al., 2016), solar physics: sunspot time series (Sadegh Movahed et al., 2006), finance: the economy and stock market (Reboredo et al., 2013) (Zunino et al., 2008; Caraiani, 2012) and even in music (Dagdug et al., 2007; Jafari et al., 2012; Hennig et al., 2011; Telesca and Lovallo, 2012). Finally, it has been widely used to investigate the fractal behaviour of hydrological systems, which is the focus of this investigation.

A variety of techniques have been used to study the fractal behaviour of time series. These include spectral analysis, wavelet analysis, rescaled-range (R\S), and detrended fluctuation analysis (DFA). Among these techniques, DFA and spectral analysis are the most commonly used, with DFA being the preferred technique by many researchers (Chen et al., 2002; Eichner et al., 2003; Zhang et al., 2011b; Hu et al., 2001, 2009; Matsoukas et al., 2000; Ozger, 2011) due to ease of detecting changes in scaling when compared to spectral analysis. Many hydrological time series are mono- and multi-fractal in nature with cut-offs in their scaling regime, i.e. they exhibit crossovers (Little and Bloomfield, 2010; Matsoukas et al., 2000; Li and Zhang, 2007; Tessier et al., 1996). Identifying these crossovers, or scaling breakpoints, is not generally done in a systematic or objective way, if it is acknowledged at all (Little and Bloomfield, 2010; Zhang et al., 2011b; Zhu et al., 2012; Williams and Pelletier, 2015; Yu et al., 2016; Li et al., 2015; Condon and Maxwell, 2014). In order to overcome this deficiency and to provide a means for quantifying reliable mono-fractal behaviour that can be used for further analysis – such as in conjunction with models or to infer causalities – this study presents a robust DFA procedure, named r-DFAn. The aim behind r-DFAn is to identify statistically different scaling regions in a signal along with the location of these changes, or crossovers, in a systematic way.

Even though fractal behaviour was found to be intrinsic to signals observed from diverse fields, a key stage in its development is Hurst's investigation of the storage capacity of the Aswan High Dam in Egypt in 1951. Analysing annual flows in the Nile, he noticed the clustering of high flows and low flows in the hydrological time series, and how these variations were scaled with the time over which they were considered. This effect came to be known as the Hurst Phenomenon (Hurst, 1951, 1956) and appears to be a fundamental property of many natural and anthropogenic systems, as the above examples show.

Hydrological and hydro-meteorological time series such as rainfall, river stage, river flow, temperature and more recently, groundwater levels have been characterised as being fractal (Eichner et al., 2003; Zhang and Schilling, 2004; Zhang and Yang, 2010; Fraedrich and Larnder, 1993; Gelhar, 1974; Kavasseri and Nagarajan, 2004; Li and Zhang, 2007; Little and Bloomfield, 2010; Zhu et al., 2012; Liang and Zhang, 2013), however, high resolution hydrological datasets are generally not available and this makes the study of the full range of fractal behaviour difficult. Among hydrological variables, groundwater levels, in particular, are not generally monitored at very short time intervals (such as one minute intervals), as for most purposes less frequent measurements are considered sufficient to capture any variations of interest. Indeed, in many aquifers the forcing processes are significantly damped such that there is very little value in monitoring at time intervals less than 1 day. However, this is not necessarily the case for shallow permeable aquifers, particularly if hydraulically connected with a river. In such cases, fluctuations in recharge due to variations in rainfall or changes in river stage during flood events can cause sub-daily groundwater level variations which can only be studied with high resolution data.

After presenting the r-DFAn procedure, a synthesized mono-fractal signal will be used to illustrate r-DFAn. In addition to this, high resolution, 1-min and 15-min, hydrological data from a study site in Wallingford will be presented in the time domain and their fractal behaviour will be analysed using r-DFAn. The datasets are: groundwater levels, river stage, groundwater temperature, river temperature, precipitation and air temperature.

The sections that follow include an explanation of the r-DFAn procedure followed by a detailed description of the study site and data collection and finally a presentation of the r-DFAn results along with a general discussion and some conclusions.

2. Methodology: r-DFAn procedure

Among numerous methods developed for studying fractal behaviour, detrended fluctuation analysis (DFA) is agreed to be a reliable method for non-stationary signals (Chen et al., 2002; Eichner et al., 2003; Zhang et al., 2011b; Hu et al., 2001; Matsoukas et al., 2000), among others). Nevertheless, in the case of mono-fractal signals that exhibit changes in there scaling regimes, determining crossovers is subjective and seriously affects the reliability of mono-fractal quantification.

Hence a procedure that includes DFA and statistical models was developed in order to overcome this shortcoming and to automate the entire quantification process. The procedure, which will be named r-DFAn, where n is the order of the detrending function, is explained below and illustrated on a synthetic signal.

2.1. Detrended fluctuation analysis

DFA of first order (i.e. DFA1) was first proposed by (Peng et al., 1994) when analysing correlations in DNA. DFA is presented in the following five steps:

1. Let $y(t_i)$ be a measurement of variable y observed at equally spaced time intervals, t_i, for N discrete times. Let $\bar{y}$ be the mean of $y(t_i)$. Compute $Y(t_i)$ by subtracting the mean from the time series and computing a cumulative sum:

$$Y(t_i) = \sum_{i=1}^{N}(y(t_i) - \bar{y}) \tag{1}$$

2. Divide $Y(t_i)$ into m non-overlapping segments each of length L so that $m = int(\frac{N}{L})$. Each segment will be notated as $Y_{j,k}(t_i)$ where $j = 1, 2, \ldots L$ and $k = 1, 2, \ldots m$, hence $i = (k-1)L + j$.
3. Determine the variance ($F_k^2(L)$) of the fluctuation in each segment (Y_k) after subtracting a best-fit polynomial of order n ($P_{j,k}^n(t_i)$) from each segment. DFAn refers to DFA detrending with polynomial of order n.

$$F_k^2(L) = \frac{1}{L}\sum_{j=1}^{L}(Y_{j,k} - P_{j,k}^n)^2 \quad for\ k = 1, 2, \ldots m \tag{2}$$

4. Determine an average variance measure for all segments of length L:

$$F(L) = \left[\frac{1}{m}\sum_{k=1}^{m}F_k^2(L)\right]^{1/2} \tag{3}$$

5. Repeat steps 1 to 4 for different values of L then plot $F(L)$ versus L on logarithmic axes to determine the scaling exponent (α) which is the slope of a best-fit line, as:

$$F(L) \approx L^{\alpha} \tag{4}$$