Physica A 392 (2013) 2195–2201

Contents lists available at SciVerse ScienceDirect

Physica A

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

# Using Detrended Cross-Correlation Analysis in geophysical data

E.B.S. Marinho, A.M.Y.R. Sousa, R.F.S. Andrade *

*Instituto de Física, Universidade Federal da Bahia, 40210-340 Salvador, Bahia, Brazil*

ARTICLE INFO

*Article history:*
Received 19 August 2012
Received in revised form 23 October 2012
Available online 17 January 2013

*Keywords:*
Detrended Cross-Correlation Analysis
Well-log data
Geological layers

ABSTRACT

Detrended cross correlation analysis (DCCA) is used to identify and characterize correlated data obtained in drilled oil wells. The investigation is focused on different petro-physical measurements within the same well, and of the same measurement from two wells in the same oil field. The evaluation of cross correlation exponents indicates if scaling properties in two measurements are alike. The work considers also the values of cross correlated coefficients, which provide an assessment on the local correlation between measurements. The existence of several highly correlated events provides information on the continuity of geological structures, including partial and global dislocations of deposited layers.


## 1. Introduction

The knowledge of the earth subsurface is central to human activities, either in the construction or exploration activities or to minimize the damage by natural phenomena. Such study, however, is hampered by the high cost of a direct access to the system or by its inherent complexity. Indeed, the different structures within the system were formed and modified in a wide variety of processes over geologic time scales [1–3]. To fully understand this system, it is necessary to unravel its preceding history using the fingerprints they have left in the geological structures. The ultimate goal is to identify the agents that have driven the system, the kind of intervening materials, and to present the sequence of processes and the conditions under which they occurred. Advances in the understanding of the geosphere, in particular of the lithosphere, are obtained by analyzing geological and geophysical data. Such knowledge is of utmost importance in the exploitation of natural resources, such as oil, reducing risk levels and making it more efficient. Well data profiles (well-logs) are particularly of great importance for the oil industry. They consist in the systematic recording of several physical and chemical properties of the subsurface materials, e.g., resistivity, radioactivity, temperature, as a function of depth. These data provide information about the sequence of rock layers in the drilled wells, layer interfaces, geological structures, being used to characterize hydrocarbon reservoirs and to assess the feasibility of oil exploitation [1].

Several mathematical tools that uncover scaling fluctuation properties have been used to analyze geophysical data [4–7]. In particular, well-log data sets have been subject of a comparative analysis by several nonlinear methods [8]. One contribution for the understanding of well log data is based on the analysis of local and global fluctuations existing in the series. Such a procedure allows to identify the presence of (auto) correlation in a single signal as a function of the distance between records, or between two different signals (cross-correlation). Due to short and large scale variability, respectively related to the local inhomogeneity in rock composition and the superposed layered structures, well-logs are clear examples of irregular, non-stationary data sets. There are various methods to identify and characterize both stationary and non-stationary complex system records. They are mainly targeted to evaluating the range of fluctuations existing in the series, the magnitude of which can vary over many time scales.

* Corresponding author. Tel.: +55 71 32836657.
*E-mail addresses:* randrade@ufba.br, randrade.ufba@gmail.com (R.F.S. Andrade).

0378-4371/$ – see front matter 
doi:10.1016/j.physa.2012.12.038

One key idea to characterize different fluctuation properties is to separate actual fluctuations from contributions due to local trends that persist over larger scales. The identification of patches within which the series can be assumed as stationary is also of great importance [9,10]. The so-called detrended fluctuation analysis (DFA) [11,12] leads to good estimates of usual Hurst exponent $H$ [13,14], provided the series are stationary over the considered patch. The identification of correlation within the same signal (autocorrelation), or between two different signals (cross-correlation) is another strategy in the record analysis. In the case of well-logs, for example, the strong correlation between two signals of different profiles may indicate the continuity of a geological layer. In this work, we show that correlations in well-logs can be detected and characterized by Detrended Cross-Correlation Analysis (DCCA) [15–17], which makes use of the same idea that justifies the DFA strategy: to separate the contribution of local trends from that of actual fluctuations. The DCCA framework has been recently enlarged to handle series where the presence of periodic component can be identified [18]. Although this phenomenon occurs very often in real world data, it was not the case of the well-log data we analyzed. Our results are based both on the scaling exponents of cross correlations between data fluctuations, as well as on the coefficients of the correlation function between the same data. It is worth commenting that recent works have been able to identify and quantify asymmetries between the upward and downward scaling behavior in financial data [19,20]. In the current work, we have not been able to identify clear asymmetric scaling properties in the data sets.

In the next sections of this paper we discuss the following issues. The DCCA procedure is reviewed in Section 2. It is divided into two subsections, where we briefly discuss the exponent and coefficient evaluation. Results obtained from these two different approaches are presented in Sections 3 and 4. Finally, Section 5 closes the paper with our concluding remarks and discussions.

## 2. Detrended cross correlation analysis

### 2.1. DCCA exponents

In general fluctuation analysis, we consider two data sets of increments $x_i$ and $x'_i$, both consisting of $N$ records, and build the corresponding integrated series

$$y_k = \sum_{i=1}^{k} x_i, \qquad y'_k = \sum_{i=1}^{k} x'_i, \tag{1}$$

with $k = 1, \ldots, N$. Each integrated series is divided into $M_\nu$ (possibly overlapping) boxes of width $\nu$. In each box labeled by $(m, \nu)$, with $1 \leq m \leq M_\nu$, we compute a measure of fluctuation that depends on the method. Throughout this work, we evaluate four fluctuation measures, respectively labeled by *DFA* (Detrended Fluctuation Analysis - Eq. (2)), *SCCA* (Standard Cross Correlation Analysis - Eq. (3)), *DCCA* (Detrended Cross Correlation Analysis - Eq. (4)), and $|DCCA|$ (Eq. (5)):

$$f^2_{DFA}(m, \nu) = \frac{1}{\nu} \sum_{k=I_{min}(m,\nu)}^{I_{max}(m,\nu)} [y_k - p_k(m, \nu)]^2, \tag{2}$$

$$f^2_{SCCA}(m, \nu) = \frac{1}{\nu} \sum_{k=I_{min}(m,\nu)}^{I_{max}(m,\nu)} [y_k - \bar{y}(m, \nu)][y'_k - \bar{y}'(m, \nu)], \tag{3}$$

$$f^2_{DCCA}(m, \nu) = \frac{1}{\nu} \sum_{k=I_{min}(m,\nu)}^{I_{max}(m,\nu)} [y_k - p_k(m, \nu)][y'_k - p'_k(m, \nu)], \tag{4}$$

$$f^2_{|DCCA|}(m, \nu) = \frac{1}{\nu} \sum_{k=I_{min}(m,\nu)}^{I_{max}(m,\nu)} |[y_k - p_k(m, \nu)][y'_k - p'_k(m, \nu)]|. \tag{5}$$

The label $|DCCA|$ in Eq. (5) emphasizes the use of the *absolute* value of the local fluctuation in each of the series. In Eqs. (2)–(5), $\bar{y}$ and $\bar{y}'$ are the mean values of $y_k$ and $y'_k$ in the box $(m, \nu)$ limited by $I_{min}(m, \nu)$ and $I_{max}(m, \nu)$; $p_k(m, \nu) = a(m, \nu)x_k + b(m, \nu)$ and $p'_k(m, \nu) = a'(m, \nu)x'_k + b'(m, \nu)$ are the first degree polynomials evaluated by the least squares method that express the local linear trends of the series in the box $(m, \nu)$. In particular, the $f^2_{DFA}(m, \nu)$ is the particular case of $f_{DCCA}$ or $f_{|DCCA|}$ with two equal series.

In the sequence, a fluctuation function is calculated for each width $\nu$ according to

$$F^2_X(\nu) = \frac{1}{M_\nu} \sum_{m=1}^{M_\nu} f^2_X(m, \nu), \quad X = DFA, SCCA, DCCA, |DCCA|. \tag{6}$$

If the series present scale properties related to cross-correlations, it is expected that a power law $F_X(\nu) \sim \nu^\lambda$ holds. The exponent $\lambda$ represents a measure of cross correlation between the two analyzed series. In the case of DFA, the exponent $\lambda$ becomes equivalent to the Hurst or roughness exponent, usually denoted by $H$ or $\alpha$.