Gradient pattern analysis of short nonstationary time series: an application to Lagrangian data from satellite tracked drifters

A.T. Assireu a,c,e, R.R. Rosa b, N.L. Vijaykumar b, J.A. Lorenzzetti c, E.L. Rempel b, F.M. Ramos b, L.D. Abreu Sá d, M.J.A. Bolzan e, A. Zanandrea f

a Oceanographic Institute, University of São Paulo, 05508-900 São Paulo, SP, Brazil
b Nucleus for Simulation and Analysis of Complex Systems (NUSASC-LAC), National Institute for Space Research (INPE), C.P. 515, São José dos Campos 12010-970, SP, Brazil
c Division of Remote Sensing, INPE, Brazil
d Centro de Previsão do Tempo e Estudos Climáticos (CPTEC), INPE, Brazil
e Instituto de Pesquisa e Desenvolvimento (IP&D), UNIVAP, Brazil
f Centro Técnico Aeroespacial (CTA), São José dos Campos 12228-900, SP, Brazil

Abstract

Based on the gradient pattern analysis (GPA) technique we introduce a new methodology for analyzing short nonstationary time series. Using the asymmetric amplitude fragmentation (AAF) operator from GPA we analyze Lagrangian data observed as velocity time series for ocean flow. The results show that quasi-periodic, chaotic and turbulent regimes can be well characterized by means of this new geometrical approach.

© 2002 Published by Elsevier Science B.V.

Keywords: Gradient pattern analysis; Nonstationary time series analysis; Turbulence; Nonlinear phenomena

1. Introduction

Time series analyses are typically statistical procedures in which the data series are regarded as subsets of a stochastic process. Additionally, the statistical properties of the time series can be used to determine the similarities and differences between distinct data sets. However, if the series is short (∼ 10^2 points) and nonstationary, these statistical measures (first and second moments, covariance and correlation functions, etc.) are less representative and useful [1]. As an example, Fig. 1a and b shows that the estimated mean value is not constant in time; different segments of these time series have different mean values and therefore these series are nonstationary. In these cases the ergodic assumption can be accepted only in a weak sense and the turbulent pattern coming from the underlying physical process, shown in Fig. 1a, cannot be characterized by means of pure statistical properties. In order to explore a more efficient way of characterizing variability patterns in short and nonstationary time series we used an operator from the gradient pattern analysis (GPA) formalism introduced by Rosa et al. [2–6]. The GPA formalism performs numerical investigation of spatio-temporal complex patterns by using computational operators such as AAF.
Fig. 1. (a) Time series from turbulent wind velocity measurement. (b) Time series from the solution of Kuramoto–Sivashinsky equation.
2. Methodology

Experimental data typically consist of discrete measurements of a single observable \( X(t) = [X(t_1), \ldots, X(t_i), \ldots, X(t_N)] \), whose time resolution is given by \( \tau = t_{i+1} - t_i \). For a more convenient notation, we can write the discrete time series as \( X(t) = [X_1, \ldots, X_i, \ldots, X_N] \) with \( X_j \in \mathbb{R} \).

In order to apply the GPA on \( X(t) \) it is necessary to convert it into a sequence of \( \ell \times \ell \) matrices, where \( \ell \geq 3 \). The length of temporal variability \( \tau \) is given by \( \tau = (\ell^2 - 1)\tau \). Note that, the highest temporal variability resolution is obtained by grouping the data into \( 3 \times 3 \) matrices row wise. This elementary case, we call a \((9-1)\tau\) variability pattern. Thus, each nine point sequence of a given time series produces the matrix \( M_{9\times 9} \):

\[
\begin{pmatrix}
X_{11} & X_{12}^+ & X_{13}^+ \\
X_{12}^- & X_{12} & X_{13}^-
\end{pmatrix}
\]

\[
\begin{pmatrix}
X_{21} & X_{22}^+ & X_{23}^+ \\
X_{22}^- & X_{22} & X_{23}^-
\end{pmatrix}
\]

\[
\begin{pmatrix}
X_{31} & X_{32}^+ & X_{33}^+ \\
X_{32}^- & X_{32} & X_{33}^-
\end{pmatrix}
\]

Note that, the distribution above (or its first transpose) is completely symmetric concerning temporal intervals: \([t_{i+8} - t_i], [t_{i+6} - t_{i+2}], [t_{i+3} - t_{i+1}], [t_{i+5} - t_{i+3}] \).

In order to characterize different maxima or minima, the gradient field \( \nabla M_{9\times 9} \) is used. The gradient field \( \nabla M_{9\times 9} \) is composed of nine vectors where each vector is associated to a grid point of the numerical lattice after a local numerical difference (see Appendix A). The gradient field \( \nabla M_{9\times 9} \) specifies quantitatively the variations at a given point and it is the proper quantitative indicator of the changes in the distribution of these maxima, i.e. an indicator of the variability pattern. In other words, in this approach, the relative values instead of the absolute values of the amplitude in the matrix are dynamically relevant.

The spatial distribution of vectors in the gradient field \( \nabla M_{9\times 9} \) will contain many vectors with the same magnitude, within a small tolerance, and these will form symmetric pairs if they have opposite orientations, and asymmetric pairs otherwise. After removing all symmetric pairs of vectors, the number of asymmetric vectors is denoted by \( L \). Any symmetry breaking can be quantified by connecting the middle points of the \( L \) vectors with \( I \) straight lines, generating a Delaunay triangulation field \( T_{\ell}(I, L) \). The triangulation field is a fractional field whose dimension is less than \( 2-\ell \) the lattice dimension. A measure of the asymmetry is the difference \( I - L \) normalized by \( L \), the so-called AAF parameter \( F_A \equiv (I - L)/L \) with \( I > L > 0 \) [4]. For symmetric patterns \( L = 0 \) and by definition, \( F_A = 0 \). The \( F_A \) parameter is a signature of the pattern \( \mathcal{P} \equiv \nabla M_{9\times 9} \), quantifying the degree of vectorial asymmetry and it is proportional to the symmetry breaking during the evolution of the gradient field pattern \( \mathcal{P} \). In the case of a random gradient pattern \( P_R \), when all vectors are distinct in modulus and orientation, \( F_A \) of \( P_R \) has
Fig. 2. (a) A typical $(9-1)\tau$ variability pattern extracted from a nonstationary time series of a generic variable $V(t)$, and its corresponding (b) $M_9$ matrix; (c) gradient (with nine asymmetric vectors) and (d) triangulation (with 20 connecting lines) patterns.

the highest value of 1.333. Indeed, the only possible values for the parameter $F_A$, calculated from $\nabla M_9$, are all restricted to the discrete domain $D_9$ composed of $\{0.778; 0.889; 1.000; 1.111; 1.222; 1.333\}$, and the uniqueness of $D_9$ can be proved by geometrical considerations [4].

As an example, one can see in Fig. 2 the operation on a nine point set that belongs to a typical nonstationary pattern. The gradient field is totally asymmetric and its value of $F_A$ is equal to $(I-L)/L = (20-9)/9 = 1.222$. Note that, when the variability of the nine point set is totally regular (periodic with a constant second moment) the corresponding gradient field is totally symmetric and therefore $F_A = 0$. It is worth noting that, for a given $X(t)$ composed by $N$ points it is possible to calculate the integer of $N/9$. This number (called here as $N_{9x}$) gives the amount of $F_A(\nabla M_9)$ and all the $N_{9x}$ values will belong to the six value domain $D_9$. It means that the most complex variability pattern can be detected by $6 \times 9 = 54$ points. In order to characterize the robustness of such a regime we need, by means of an average of $\langle F_A \rangle$, at least two intervals of 54 points, so that this technique starts to be useful for $N \geq 108$ points. Therefore, the choice of the $3 \times 3$ matrix dimension represents a good compromise between the robustness of the method and a possibility of obtaining an adequate resolution, which allows the application of the method to short time series. Taking bigger matrix sizes ($4 \times 4$, $5 \times 5$, ...) we have bigger sets of $F_A$ values and some of these values may be redundant, unlike $3 \times 3$ matrix where there are always a unique set of six distinct $F_A$ values.

As a preliminary validation of the AAF operator for time series analysis, we present two canonical time series together with the results of the mean $\langle F_A \rangle$ parameter as well as the first and the second moment statistics for long segments of the series separated by vertical lines as shown in Fig. 1a and b. The first example (Fig. 1a) represents a high frequency time series of purely turbulent velocity regime. These turbulent data (100,000 points) were obtained in the atmospheric surface layer during the large scale biosphere–atmosphere experiment in Amazonia (LBA). The fast response wind speed measurements, sampled at 60Hz, were performed by means of three-dimensional sonic anemometers [7]. The second example (Fig. 1b) represents a time series generated by varying the damp-
ing parameter \( \nu \) in a damped Kuramoto–Sivashinsky (DKS) model [8,9]. The time series in Fig. 1b corresponds to the temporal variability of one of the spatial coordinates of the system. The first regime \((t < 2000)\) is quasi-periodic \((\nu \geq 0.0299)\). The second regime \((t > 2000)\) is chaotic \((\nu < 0.0021)\).

For the first example, it is noted that despite the nonstationary characteristic of the series, the values of \( \langle F_A \rangle \) are quite stable. The results suggest that a value of \( F_A = 1.087 \pm 0.025 \) could be used as an alternative characterizing parameter of the turbulent pattern. The second example shows that the \( F_A \) parameter can be used to distinguish between a quasi-periodic \((\langle F_A \rangle = 1.001 \pm 0.002)\) and a chaotic regime \((\langle F_A \rangle = 1.040 \pm 0.020)\). Therefore, it seems that \( F_A \), for \( \gamma = (9 - 1)\tau \), can be useful to characterize quasi-periodic, chaotic and turbulent variability patterns that can be mixed in real nonstationary complex systems, mainly those in which both deterministic and stochastic processes are present. The robustness of this new methodology has been supported by further application on time series from chaotic maps and on Lagrangian experimental data. The results obtained from applications on chaotic maps (logistic and Hennon) will be reported later in a future paper.

3. Analysis of the Lagrangian oceanographic time series

Complex patterns are observed in velocity time series obtained from satellite tracked Lagrangian drifters. These drifters (buoys), when released in the ocean, are advected over large areas and consequently, are subjected to the different flow regimes. Thus, in a real ocean, the spectrum of oceanic turbulence contains a wide variety of eddies which cannot be easily separated. Therefore, nonlinear patterns may arise from the coupled dynamics of different eddies. Here, we apply the AAF operator to characterize different patterns associated with the velocity variability of drifters and consequently, of the ocean flow. The trajectory of a drifter is shown in Fig. 3a, indicating the different flow regimes for which the analysis was performed (segments 1 and 3—apparently quasi-linear flows and segment 2—nonlinear eddy flow). Fig. 3b shows the derived meridional velocity component time series, composed of 500 points, corresponding to the trajectory presented in Fig. 3a.

The data analyzed consists of the trajectory of a drifter deployed in Brazil Current waters. The drifter drogue depth (15 m) was chosen to minimize spurious wind and wave-drift effects on the buoy motions, and to allow measurement of the dynamics in the mixed surface layer. Positions of the drifters were determined several times a day by the ARGOS DCLS/NOAA satellite system [10]. Some location inaccuracies are caused by instrumental noise (mainly transmitter oscillator stability), buoy-satellite orbit geometry, number and distribution of messages received during a particular orbital pass, and inaccuracies in orbit and time coding. These errors were removed by application of Hansen and Poulain methodology [11]. Additional technical details on treating drifter data series can be found in [12].

4. Results and conclusions

The mean value, \( \langle \cdot \rangle \), and standard deviation, \( \sigma \), of the asymmetric fragmentation parameter, for each one of the three segments S1, S2 and S3 shown in Fig. 3, are presented in Table 1. The comparison of these results with those obtained in Section 2, indicates that the mean variability of the asymmetric fragmentation in segments S1 and S3 corresponds to a turbulent regime. Likewise, for segment S2, the values obtained for \( \langle F_A \rangle \) indicate a quasi-periodic regime as suggested by a visual analysis of the drifter trajectory and meridional velocity component.

In summary, we can say from these preliminary results, that the AAF operator from the GPA technique might be a robust and low cost computational tool to analyze the flow characteristics of nonstationary

<table>
<thead>
<tr>
<th>Segment</th>
<th>( \langle F_A \rangle \pm \sigma )</th>
<th>( \langle F \rangle \pm \sigma \text{ (cm s}^{-1}\text{)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1.11 \pm 0.02</td>
<td>-40.49 \pm 28.17</td>
</tr>
<tr>
<td>S2</td>
<td>1.00 \pm 0.02</td>
<td>1.38 \pm 32.24</td>
</tr>
<tr>
<td>S3</td>
<td>1.11 \pm 0.02</td>
<td>-15.50 \pm 24.38</td>
</tr>
</tbody>
</table>
Fig. 3. (a) Drifter trajectory divided into three different segments (S1, S2 and S3). (b) The corresponding meridional drifter velocity time series.
and short time series. An even better separation of flow regimes is possible by using a decibel scale for the fragmentation parameter, i.e., \( F_A (\text{dB}) = 10 \log_{10}(F_A) \). Using the decibel \( F_A \), the quasi-periodic regime would be characterized by \( 0 \leq F_A (\text{dB}) \leq 0.086 \) and the turbulent regime by \( 0.370 \leq F_A (\text{dB}) \leq 0.530 \). We are also applying this technique to investigate its utility for characterization of different kinds of intermittent and anomalous nonlinear diffusion phenomena, usually present in Lagrangian time series coming from velocity variability of drifters. A more complete characterization of chaotic regimes is also in progress and will be communicated later.

Acknowledgements

Thanks are due to R.P. Kane and V. Krishan for valuable discussion and text review. The authors are grateful to the Brazilian funding agency FAPESP under grants 97/13374-1, 98/15413-7 and 00/10043-9 and to the Brazilian National Buoy Program (PN-BOIA). We are also grateful to the referee for valuable suggestions.

Appendix A. Basic comments on AAF computational operator

The gradient pattern \( \nabla M^i_{ij} \) is obtained by the computation of numerical gradient \( d M/dx \) and \( d M/dy \) of the matrix \( M^i_{ij} \), where \( d M/dx \) and \( d M/dy \) correspond to the differences in the \( x \) (column) direction and the differences in the \( y \) (row) direction, respectively. The spacing between points in each direction is assumed to be one. A prototype of this kind of gradient routine is the one used in the MATLAB software \[13\]. The triangulation field is computed by means of a classical Delaunay triangulation routine. This convex triangulation maximizes the minimum internal angles of each triangle, i.e. each triangle is as equilateral as possible and obeys the empty circle criterion \[14\] which allows to construct the triangulation field directly from the sample set. The code for the AAF operator is written in C. There are versions for MATLAB and IDL environments. The IDL routine using the widget resource, permits an easy usage of it. Further details about the code and a public IDL version of AAF operator, can be supplied by the authors upon request.

References