Physica A 499 (2018) 233–240

Contents lists available at ScienceDirect

Physica A

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

Time series analysis of the Antarctic Circumpolar Wave via symbolic transfer entropy

Mingi Oh [a], Sehyun Kim [a], Kyuseong Lim [b], Soo Yong Kim [a,*]

[a] Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea
[b] Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon 16419, Republic of Korea

HIGHLIGHTS

- Antarctic Circumpolar Wave is examined through Symbolic transfer entropy.
- We calculate STEs for various embedding dimensions and test them for significance.
- STE contour maps are drawn to reveal the presence of ACW.
- Eastward information flow is observed only for oceanic variables, SST and SIE.
- Marked flow of information over the eastern Pacific Ocean is detected for all variables.

ARTICLE INFO

Article history:
Received 16 June 2017
Received in revised form 17 October 2017
Available online 12 January 2018

Keywords:
Antarctic Circumpolar Wave
Symbolic transfer entropy
Climate
ENSO
Information flow

ABSTRACT

An attempt to interpret a large-scale climate phenomenon in the Southern Ocean (SO), the Antarctic Circumpolar Wave (ACW), has been made using an information entropy method, symbolic transfer entropy (STE). Over the areas of 50–60°S latitude belt, information flow for four climate variables, sea surface temperature (SST), sea–ice edge (SIE), sea level pressure (SLP) and meridional wind speed (MWS) is examined. We found a tendency that eastward flow of information is preferred only for oceanic variables, which is a main characteristic of the ACW, an eastward wave making a circuit around the Antarctica. Since the ACW is the coherent pattern in both ocean and atmosphere it is reasonable to infer that the tendency reflects the Antarctic Circumpolar Current (ACC) encircling the Antarctica, rather than an evidence of the ACW. We observed one common feature for all four variables, a strong information flow over the area of the eastern Pacific Ocean, which suggest a signature of El Nino Southern Oscillation (ENSO).

1. Introduction

Due to the increasing availability of computing power and storage, accumulating large datasets has become a standard procedure in data science. In addition, the overall advancements in handling complex statistical models have allowed these disciplines to flourish in a wide spectrum of research fields, such as biophysics, brain signal processing, social media, econophysics and dynamical systems. More importantly, the data acquired from those studies are usually complex and nonlinear and therefore various statistical tools, such as complex networks, multifractal [1], detrended fluctuation analysis [2], and information theory [3] are most widely implemented to best understand these systems.

* Corresponding author.
E-mail address: sykim@kaist.ac.kr (S.Y. Kim).

https://doi.org/10.1016/j.physa.2017.12.019

Geophysical phenomena, one of the nonlinear and complex systems, have also been researched using computational methods in conjunction with analytical model approaches. In the southern hemisphere, which is observed relatively less than its northern counterpart, properties of inter-annual variability of the SO have been widely investigated. White and Peterson [4] (WP96 hereafter) pointed out the existence of a coherent and eastward propagating wave in both oceanic and atmospheric datasets, referred to as the ACW. The ACW is a spatial pattern of climate variables over the SO, which propagates eastward with the Antarctic Circumpolar Current with a periodicity of 4–5 years and makes an entire circuit around Antarctica in approximately 8–10 years. WP96 demonstrated and verified the existence of this circumpolar wave in SST, SLP, mean wind stress and sea–ice extent. Jacobs and Mitchell [5] showed a compatible result with WP96, observing the eastward propagation of the ACW in sea surface height, wind stress curl and SST. White et al. also suggested a mechanism of the ACW, in which eastward speeds of ACC model matched with those observed in SST anomalies in the presence of ocean–atmosphere coupling [6], and Christoph et al. [7] presented a coupled general circulation model in which a spatial structure of the ACW has zonal wave number 3 rather than 2. Connolley [8] analyzed the monthly sea level pressure for the period 1968–1999 and noted that a significant eastward propagation of the ACW with zonal wavenumber 2 was only shown during 1985–1995, the time period for which WP96 initially discovered the presence of the ACW. Before and after this time period, all the patterns of the ACW seemed to have a zonal wave number 3 or no clear sign of precession. In the same year, Venegas pointed out the simple picture of the ACW generation as a linear combination of propagating wavenumber-2 and wavenumber-3 patterns, each of which has different origins, namely a periodicity of 5 and 3 years, respectively [9]. However, Park et al. used Fourier harmonics of SST to reveal that the eastward propagating part accounted for only 25% of the total signal and most of the SO variability can be explained in conjunction with a standing wave pattern associated with ENSO [10].

Since late 1979, satellite observations have dramatically contributed to data mining in geoscience and enabled numerical analyses to be competitive against analytic methods, leading to active statistical approaches concentrated on time series data itself. As a statistical method, we used STE [11] to highlight the characteristics of the ACW.

The aim of the paper is by using the STE to investigate a coherent phenomenon related with four climate variables in two aspects. Firstly, a measure, STE, is introduced to explain the past findings of the ACW in terms of a biased direction of information flow. Secondly, incorporation of significant test of STE should be indispensable and invaluable draw a meaningful result. The paper is organized as follows. In Section 2, the data we treated are introduced in detail. Section 3 presents how to analyze data using transfer entropy method. Section 4 presents results and brief discussion. Finally, conclusion is made in Section 5.

2. Data

The dataset used in this work contains monthly means of SST, sea–ice concentration (SIC), SLP and MWS from the National Oceanic and Atmospheric Administration (NOAA). SSTs and SICs are included in NOAA optimum interpolation version 2 datasets [12] derived by averaging the appropriate daily values within a month to produce monthly averages spanning from Jan. 1982 through Dec. 2015. Both SST and SIC fields are on 1.0×1.0 degree regular latitude–longitude global grids. From SIC data, SIE is defined as a set of latitude values where SIC was at least 20% at each longitude. SIE is set to the continental boundary if such values cannot be found. SLP and MWS data covering the period from Jan. 1979 to Dec. 2015 at monthly intervals on T62 Gaussian grid are provided from 'NCEP/DOE Reanalysis project 2' generated by US National Centers for Environmental Prediction and the Department of Energy [13]. Four datasets can be downloaded from a web site at http://www.esrl.noaa.gov/psd/. Before calculating the entropy values, datasets as an input are preprocessed. Time series of monthly anomalies are produced in each variable by subtracting a monthly climatology, by month, from the time series of a raw data to remove the seasonal cycles in the series. Since the ACW appears in inter-annual time scale, time series of anomalies are band-passed with a filter of 2–7 years in most literatures, and filtering processes differ slightly between researchers. However, we use non-filtered anomalies to reduce arbitrariness and to preserve an original characteristic of signals as much as possible. The region of interest is defined as the latitude belt from 50°S to 60°S where the ACW is discovered. To improve computational efficiency, we compressed 3-dimensional data composed of time–latitude–longitude into 2-dimensional data by averaging latitudes.

3. Method

To identify a causal relation between two signals, the STE measure is used. STE was first suggested by Staniek and Lehnertz who adopted and developed permutation entropy introduced by Bandt and Pompe [14]. Similar to transfer entropy [15], STE can be obtained by counting how many cases are found in a time series, where cases refer to two symbols at successive times. The symbol mentioned above is defined as a permutation given by reordering the amplitudes values of time series in ascending order. Details for calculating STE follow below.

Given a sequence of observation X with length N, a vector is assigned for each element X_i, $i = 1, 2, 3, \dots, N$, in a following way: a small sequence(subset) of X, $\vec{X}_i = \{X_i, X_{i+\tau}, X_{i+2\tau}, \dots, X_{i+m-1\tau}\}$ is arranged in an increasing order $\{X_{i+\mathrm{k}_{i1}-1\tau} \le X_{i+\mathrm{k}_{i2}-1\tau} \le \dots \le X_{i+\mathrm{k}_{\mathrm{im}}-1\tau}\}$ where m and τ denote embedding dimension and time delay respectively and k's are ordinal numbers corresponding to an element of $\vec{X}_i$. Thus, a symbol corresponding to $\vec{X}_i$ is defined as $\hat{X}_i = k_{i1}, k_{i2}, \dots, k_{im}$. For example, if $k_{i1} = 3$, the third element $X_{i+2\tau}$ is the smallest value in $\vec{X}_i$. In this way, every symbol $\hat{X}_i$ is uniquely mapped