Journal of Atmospheric and Solar--Terrestrial Physics 154 (2017) 10–20

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar–Terrestrial Physics

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

A new tool for spatiotemporal pattern decomposition based on empirical mode decomposition: A case study of monthly mean precipitation in Taihu Lake Basin, China

Shen Chenhua[a,b,c,*], Yan Yani[a]

[a] College of Geographical Science, Nanjing Normal University, Nanjing, Jiangsu 210023, China
[b] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
[c] Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, Jiangsu 210023, China

ARTICLE INFO

Keywords:
Spatiotemporal pattern decomposition tool
Empirical mode decomposition
Empirical orthogonal function
Monthly mean precipitation
Taihu Lake Basin

ABSTRACT

We present a new tool for spatiotemporal pattern decomposition and utilize this new tool to decompose spatiotemporal patterns of monthly mean precipitation from January 1957 to May 2015 in Taihu Lake Basin, China. Our goal is to show that this new tool can mine more hidden information than empirical orthogonal function (EOF). First, based on EOF and empirical mode decomposition (EMD), the time series which is an average over the study region is decomposed into a variety of intrinsic mode functions (IMFs) and a residue by means of EMD. Then, these IMFs are supposed to be explanatory variables and a time series of precipitation in every station is considered as a dependent variable. Next, a linear multivariate regression equation is derived and corresponding coefficients are estimated. These estimated coefficients are physically interpreted as spatial coefficients and their physical meaning is an orthogonal projection between IMF and a precipitation time series in every station. Spatial patterns are presented depending on spatial coefficients. The spatiotemporal patterns include temporal patterns and spatial patterns at various timescales. Temporal pattern is obtained by means of EMD. Based on this temporal pattern, spatial patterns at various timescales will be gotten. The proposed tool has been applied in decomposition of spatiotemporal pattern of monthly mean precipitation in Taihu Lake Basin, China. Since spatial patterns are associated with intrinsic frequency, the new and individual spatial patterns are detected and explained physically. Our analysis shows that this new tool is reliable and applicable for geophysical data in the presence of nonstationarity and long-range correlation and can handle nonstationary spatiotemporal series and has the capacity to extract more hidden time-frequency information on spatiotemporal patterns.

1. Introduction

Decomposition of spatiotemporal pattern is a significant problem in geography, both in theoretical analysis and in practical application, playing an important role in analyzing and predicting environmental change (Li et al., 2013). The changes of geographical processes in spatiotemporal dimensions are often observed and recorded. Commonly, for a specified observational station, one variable should be required at least to record this change with time. For p observational stations, there are thus p spatiotemporal series. For convenience, a spatiotemporal series with series length of N is denoted as $\{x(i, t)\}$ ($1 \le i \le p$, $1 \le t \le N$).

For the purpose of decomposing spatiotemporal patterns that might be linked to physical mechanisms (Dommenget and Latif, 2002), an empirical orthogonal function (EOF) (Pearson, 1901; Kim et al., 1970) is commonly applied. EOF has been found to be one of the most important and effective methods for spatiotemporal pattern decomposition and has been widely applied in climatic fields (Hamlington et al., 2015; Jackson and Mound, 2010; Li et al., 2012; Lin and Wang, 2006; Pritchard and Somerville, 2009; Wei and Zhang, 2010). A number of useful findings have resulted from this method. To date, a variety of improved, extended, and adjusted approaches, such as an extended EOF (Weare and Nasstrom, 1982), rotated EOF (Cheng et al., 1995; Lian and Chen, 2012), complex EOF (Rasmusson et al., 1981; Barnett, 1985), wavelet EOF (Nayagam et al., 2009), and distinct EOF (Dommenget, 2007) (all based on EOF), have been established to improve understanding.

EOF might decompose spatiotemporal series into two parts through

* Corresponding author at: College of Geographical Science, Nanjing Normal University, Nanjing, Jiangsu 210023, China.
E-mail address: shenandchen01@163.com (S. Chenhua).

http://dx.doi.org/10.1016/j.jastp.2016.12.008
Received 24 July 2016; Received in revised form 10 December 2016; Accepted 20 December 2016
Available online 21 December 2016
1364-6826/

a reduction of data dimensionality: $x(i,\ t) = \sum_{m=1}^{m=p} V(i,\ m)Z(m,\ t)$ (Pearson, 1901). The first part, $V(i,\ m)$, is the spatial one, and the other, $Z(m,\ t)$, is the temporal one. For the series $\{x(i,\ t)\}$, there exists $xx^{\mathrm{T}} \equiv \sum_{t=1}^{t=N} x(i,\ t)^2 = (\mu^2(i)+\sigma^2(i))N$, where $\mu(i)=\sum_{t=1}^{t=N} x(i,\ t)/N$ and $\sigma^2(i) =\sum_{t=1}^{t=N} (x(i,\ t) - \mu(i))^2/N$ are designated as the mean of the series and the variance of the series, respectively. Although either stationary or nonstationary (Kantelhardt et al., 2002) spatiotemporal series can be decomposed into various spatiotemporal patterns, none of these patterns have time-frequency information. Furthermore, for stationary spatiotemporal series, since μ and σ remain unchanged statistically with an increasing series length (Shen et al., 2016), the decomposed spatiotemporal patterns can be directly utilized to predict a variable. In contrast, for nonstationary spatiotemporal series, since μ and σ are closely related to series length (Kantelhardt et al., 2002), the nonlinear change $xx^{\mathrm{T}} \equiv (\mu^2(i)+\sigma^2(i))N$ against series length appears. Thus, spatiotemporal patterns decomposed by EOF cannot be directly used to predict variables.

Spatiotemporal series in the presence of nonstationarity are in fact ubiquitous in nature and society (Malik et al., 2016). It is thus of great importance theoretically to discover an approach that can decompose temporal patterns of nonstationary series. Fortunately, empirical mode decomposition (EMD), proposed by Huang et al. (1998), is an effective method that can decompose nonlinear oscillatory patterns into a variety of intrinsic model function (IMF) components, based purely on the observed-data properties without relying on the concept of stationarity. This decomposition can directly extract the energy associated with various intrinsic timescales, which are the most important parameters of the series. A brief description of the procedure is as follows: first, all the local extrema of $\{x(t)\}$ are identified, and then all local maxima and minima are linked by a cubic spline to form the upper and lower envelopes. Their mean is denoted as m_1, and the difference between the data and m_1 under a specified requirement, defined as the first IMF component, is extracted. Repeating this process, we can achieve a decomposition of $\{x(t)\}$ into IMF_i ($1 \le i \le n$) and a residue, r_n, which can be either the mean trend or a constant. The maximum count of IMF is $\log_2 N-1$. Further analysis verifies that the cross-relation between IMF_i and IMF_j ($1 \le i,\ j \le n,\ i \ne j$) is not orthogonal (Huang et al., 1998) and that their cross-coefficient is about 10^{-2}–10^{-3} in magnitude. Under a low mathematical accuracy requirement, these IMFs are considered approximately orthogonal (Huang et al., 1998). For the purpose of obtaining fully orthogonal IMFs in mathematics to avoid energy leakage, the Gram-Schmidt method is applied.

EMD is well suited for one-dimensional series, either stationary or nonstationary. However, for p series with identical lengths of N in p geographical locations, the method that is used to directly decompose spatiotemporal patterns through EMD is important. Sun et al. (2008) calculated yearly-surface-mean temperature IMFs of 760 meteorological stations in China, obtained IMFs of temperature series for every station, and presented spatiotemporal patterns for yearly-surface-mean temperature through cluster analysis. Despite the fact that spatiotemporal patterns were identified, neither correlation between two or more series nor spatial distribution structure of the stations was considered.

In our study, we take the critical ideas of EMD and EOF to propose a new tool for decomposing spatiotemporal patterns. The new tool can address two challenges: to identify spatial patterns associated with the characterization of time and frequency and to handle nonstationary series. In order to test the capabilities of the approach, a theoretical analysis was carried out and some experiments were conducted to decompose spatiotemporal patterns of the monthly mean precipitation from January 1957 to May 2015 in Taihu Lake Basin, China. The results show that the approach is capable of identifying the spatiotemporal patterns of monthly mean precipitation series in the presence of nonstationarity through decomposition.

Our research differs from previous studies in the following key ways: 1) the spatial patterns decomposed by the new tool are associated with time frequency and 2) the new tool has the potential for application in a variety of research fields. Additionally, two new and individual spatiotemporal patterns are identified in the real dataset of monthly mean precipitation in Taihu Lake Basin.

Our objectives are to verify that the new tool is better at discovering the spatial patterns associated with time-frequency information and is capable of handling series in the presence of nonstationarity. Our findings will contribute to current knowledge of spatiotemporal pattern decomposition with nonstationary series and has the potential to predict variable change in the future.

2. Theoretical framework of the new tool

In this section, the new tool, a time-frequency analytical method to analyze spatial patterns at different timescales (TFSP), is defined and its theoretical framework is described. The decomposition of spatiotemporal pattern based on EOF is first briefly discussed, and then two hypotheses are given. Next, a theoretical analysis and derivation are provided and the new tool is presented.

The decomposition of spatiotemporal pattern using EOF (Pearson, 1901; Kim et al., 1970) occurs as follows: suppose that p variables X_1, X_2, ..., X_i, ..., X_p ($1 \le i \le p$), with an identical series-length of N, are used to describe the changes in monthly mean precipitation with time in p meteorological observation stations within a geographical region.

To achieve the decomposition of spatiotemporal pattern, a spatiotemporal series $X_i \equiv \{x(i,\ t)\}$ ($1 \le i \le p$, $1 \le t \le N$) is often assumed to be decomposed into two parts: a spatial one and a temporal one, i.e., $x(i,\ t) = \sum_{m=1}^{m=p} V(i,\ m)Z(m,\ t)$. The spatial part $V(i,\ m)$ designates the association degree between the ith observational station and the mth spatial field, and the temporal part $Z(m,\ t)$ represents the changes in monthly mean precipitation with time in the mth spatial field. Under the constraints of $VV^{\mathrm{T}}=1$ and $ZZ^{\mathrm{T}}=\Lambda$, the formulation of $xx^{\mathrm{T}}=VZZ^{\mathrm{T}}V^{\mathrm{T}}$ is given, where V is an eigenvector, the ith eigenvector is $(V(i,\ 1),\ V(i,\ 2),\ ...,\ V(i,\ m),\ ...,\ V(i,\ p))^{\mathrm{T}}$, and Λ is a diagonal matrix with λ_1, λ_2, ..., λ_p. Lagrange multiplier method is applied to estimate $V(i,\ m)$ and $Z(m,\ t)$ by means of an ordinary least square (OLS) and the optimal solutions of $V(i,\ m)$ and $Z(m,\ t)$ are obtained.

The key concepts of EOF are that 1) a spatiotemporal series can be decomposed into a sum over products of a spatial part and a temporal part, that 2) the spatial part can illustrate spatial differentiation for a given spatial field, that 3) the temporal part can depict the variable change with time for a fixed spatial field, that $Z(m,\ t)$ against $Z(m',\ t)$ is orthogonal, i.e., $\sum_{t=1}^{t=N} Z(m,\ t)Z(m',\ t)=\delta_{mm}\lambda_m$, $\delta_{mm'}$ is Kronecker denotation.

To arrive at a new approach, two hypotheses are given in our study. Absorbing the EOF's critical idea, naturally, the first supposition is that a spatiotemporal series can be decomposed into a sum over products of k spatial fields $V_e(i,\ m)$ and k temporal fields $Z_e(m,\ t)$. The corresponding formulation is shown in Eq. (1):

$$x(i,\ t) = \sum_{m=1}^{m=\infty} V_e(i,\ m)Z_e(m,\ t), \tag{1}$$

where $Z_e(m,\ t)$ is a complete orthogonal function. Clearly, Eq. (1) indicates that the spatial field $V_e(i,\ m)$ is independent of time, and that the temporal field $Z_e(m,\ t)$ is independent of space, and that $V_e(i,\ m)$ varies depending on spatial locations (observational stations) when m is fixed. When finite orthogonal functions are chosen, Eq. (1) can be written into Eq. (2):

$$\begin{aligned} x(i,\ t) &= \sum_{m=1}^{m=k} V_e(i,\ m)Z_e(m,\ t) + \sum_{m=k+1}^{m=\infty} V_e(i,\ m)Z_e(m,\ t) \\ &= \sum_{m=1}^{m=k} V_e(i,\ m)Z_e(m,\ t) + \xi(i,\ t), \end{aligned} \tag{2}$$

where $\xi(i,\ t)$ is a residual (truncated error) that is cause by $V_e(i,\ m)$ and $Z_e(m,\ t)$ ($m > k$).

We denote $\overline{X}(t) = \sum_{i=1}^{i=p} x(i,\ t)/p$ as the mean of precipitation over all observational stations. Hence,