LAND COVER DYNAMIC CHANGE MODES ANALYSIS BASED ON EMPIRICAL MODE DECOMPOSITION

Liang Tang[1], Haijun Yang[1*], Ping Tang[2], Zhuang Li[1], Yuhao Tang[2]

[1]Hainan Tropical Ocean University, Sanya, China - tyh9977@163.com,hyang@hntou.edu.cn, hitlz@163.com
[2]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China - tangping@radi.ac.cn, tangyh0707@163.com

KEY WORDS: Time-series, Land Cover, Dynamic Analysis, Empirical Mode Decomposition.

ABSTRACT:

The dynamic change mode of land cover not only reflects the change characteristics of the land cover itself, but also reflects the influence exerted by external factors. Empirical mode decomposition is an effective time-frequency analysis method for nonlinear and nonstationary signals. In this paper, empirical mode decomposition firstly uses to each pixel time series to get its intrinsic mode functions, each intrinsic mode function relates to change mode with some time scale, clustering those intrinsic mode functions over the image, and different land cover change modes are obtained. The experimental dataset used is MODIS subset from the Jan 2011 to Dec 2020, located in the Ningxia Hui Autonomous Region, China. The clustering results of half-annual cycle and one-year cycle of intrinsic mode functions show the influence of local climate and policy factors on land cover change.

1. INTRODUCTION

Land cover is one of the core topics in the study of global change. It not only affects the natural basis of human survival and development, but also affects the structure and function of the Earth's biochemical layer and the circulation of energy and material in the Earth system. On the global, national or regional scale, land use change always drives land cover change. The change information plays an important role in global climate change research, ecological environment monitoring and other fields. The Terra Moderate Resolution Imaging Spectrometer (MODIS) Normalized Difference Vegetation Index data products are employed for land cover change detection (Tucker, 1979). The vegetation in terrestrial ecosystem plays an important role in the global cycle of material and energy and is becoming more and more crucial to the sustainability in the context of global climate change and intensifying human activities. Vegetation changes can create feedback to climate system via the biogeophysical effects, i.e., increased leaf cover, tundra shrubification, and tree line advance can mediate land–atmosphere interactions (Zhang, 2013, 2018). To study non-stationary vegetation response, we selected Ningxia Hui Autonomous Region(NHAR) as a study area.

The normalized difference vegetation index (NDVI) is currently widely used to estimate vegetation cover, biomass and leaf area index and thus its changes can be used as a proxy for greening or browning (Alcaraz-Segura, 2010). In recent decades, several scholars have used NDVI data to conduct in-depth research on the vegetation change (Cihiar, 1991; Nicholson, 1994; Eklundh, 1998). Their results showed that the change of vegetation is closely related to not only climate factors but also human activities (Vicente-Serrano et al 2005).

NDVI time series are usually non-stationary, i.e., they present different frequency components, such as periodic or seasonal variations, long-term and short-term fluctuations, which may not be limited only to the mean of the series but also affects its overall variance structure. Typically, such series are characterized by patterns like periodicity, trends and localized abrupt changes or discontinuities resulting from disturbance events (de Beurs, 2005).

In this study, we selected Ningxia Hui Autonomous Region as the study area. We first decomposed the frequency components using the empirical mode decomposition (EMD) in the NDVI from Jan 2011 to Dec 2020. And then classified the half-annual cycle and one-year cycle of intrinsic mode functions using K-means clustering algorithm. The objective of this study is to develop a spatio-temporal coupling method for multi-scale change information extraction and analysis. These findings will promote the development and application of remote sensing time series image information mining technology.

2. DATA AND METHODS

2.1 Data and Preprocessing

The NDVI time series as the study area come from MOD13A2 product data of the Moderate-resolution imaging spectroradiometer (MODIS) during the period from January 2011 to December 2020. The product downloaded from website http://ladsweb.nascom.nasa.gov/data/search.html, 1km spatial resolution and 16day NDVI Maximum Value Composite (MVC) Product. The data has 23 phases per year and a total of 230 phases in 10 years. The MOD13A2 products are converted format from HDF to TIFF and re-projected by MODIS Reprojection Tools (MRT). Due to atmospheric variations, cloud contamination or multidirectional effects, MOD13A2 product still contain a lot of residual noises and outliers by MVC processed. These residual noises outliers hinder further analysis and have a risk to produce erroneous results, and be eliminated by further processing. The optimum-length Savitzky-Golay filter was selected as reprocessing to reconstruct high-quality NDVI time-series by approaching the upper NDVI envelope via an iteration process (Tang, 2019).

2.2 Empirical Mode Decomposition

EMD is proposed as a time-frequency signal analysis method by Huang et al. in 1998(N.E.Huang, 1998). It can decompose time-frequency signal into a finite number of intrinsic mode

---

function(IMF) according to the characteristics of the data signal itself. Each IMF represents the characteristic components of the signal on different time scales and does not interfere with each other. Any intrinsic modal function should satisfy the following two conditions: the difference between the number of extreme points and the number of zeros of the eigenmodal function is not more than 1; at any time, the mean value of the upper and lower envelope of the intrinsic modal function is zero. Unlike Fourier and wavelet transform, the decomposed IMF components is derived from the data itself, and contain local characteristic signals of different time scales of the original signal, which is especially suitable for the analysis of non-stationary and non-linear signals(Yu, 2005). Therefore, EMD method has been effectively applied in different engineering fields since it was proposed(Gloersen, 2003; Huang, 1998; Huang, 1999; N.E.Huang, 2003; Liang, 2000; Kaplan, 2013).

The essence of the EMD is to get the intrinsic mode pattern through data time scale features and perform data decomposition. The EMD of the one-dimensional signal $x(t)$ can be expressed as after equations:

$$\mathbf{x}(t) = \sum_{i=1}^{n} imf_i(t) + r(t) \tag{1}$$

where $imf_i(t) = i$th IMF
$r(t)$ = residual function
n = IMF component number

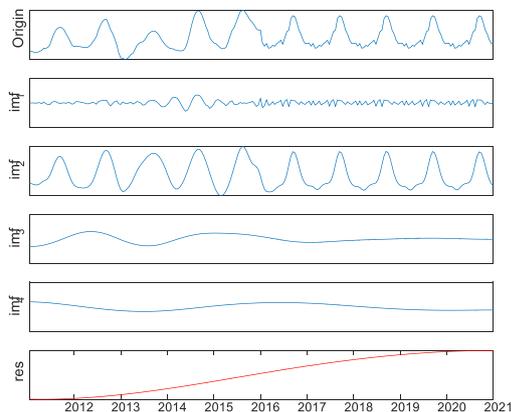Figure 1 shows the results of a l0-year NDVI time series decomposed by EMD at a certain point.

Figure 1. Decomposition results by EMD method

In figure 1, the frequency of signal from the component imf_1 to imf_4 decreased gradually, and the residual was denoted as res, which is a monotonic function and indicated an average trend. In res curve, NDVI value increased slowly from 2012 to 2013, and increased relatively rapidly between 2014 and 2018, and remains almost unchanged from 2011 to 2012, and until 2019, there was a slow increase.

2.3 K-means Clustering Algorithm

K-means clustering algorithm is an unsupervised clustering algorithm, which widely used in field of data mining across different disciplines in the past fifty years. The clustering results are not fixed, but have good convergence. In particular, the algorithm has low complexity and can quickly cluster chaos into several classes (Kumar, 2017). The basic idea of K-means algorithm is that for a given number of clusters k, an initial partition is created randomly, that is, some representative data points are randomly selected as the initial cluster centers. According to the distance between the remaining data points and each cluster center, it is divided into each class, and then the

new cluster center is determined again, and so on. The iterative method is used to move the cluster center continuously to try to further improve the division, until the cluster center is not Change again.

We begin by providing a formal description of the k-means problem and two serial algorithms for finding an approximate solution. Let $X = \{x_1, \ldots, x_n\}$ be a set of N data points, each with dimension d. The K-means algorithm aims to partition the observations into K sets: C={c_1,c_2,...,c_k}f g such that $K \leqq N$:The k-means problem seeks to find a set of k means M = $\{\mu_1, \ldots, \mu_k\}$ which minimizes the objective function

$$J(k) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \| X_i - \mu_k \|^2 \tag{2}$$

Where $d_{x_i} = \| x_i - \mu_k \|^2$ is the minimum distance measured between observation X_i and the cluster center μ_k and z_{ik} is an indicator variable belonging/not to the kth cluster.

3. EXPERIMENT AND DISCUSSIONS

3.1 Study Region

Study region is NHAR of China. It is situated in the upper and middle reaches of the Yellow River, north-western China, and covers an area of approximately 66,400 km², extending from 104°17′ E to 109°39′ E, and from 35°14′ N to 39°14′ N. It is a transitional zone between desert and the Loess Plateau and has a temperate continental arid and semi-arid climate. There are three major deserts surrounding it and desertification is a challenging threat there. Several decades of efforts have been made therefore to protect and restore vegetation for desertification combatting, including the implementation of the Grain for Green Program. That program is the largest ecological restoration project in central and western China initiated by the Chinese Central Government in 1999. It was reported that the desertification there has been significantly alleviated. Therefore, many scholars are of great interest to understand it from the perspective of vegetation cover change in this area(Tang, 2020). Figure 2 shows a true-color-composite image of Landsat TM.
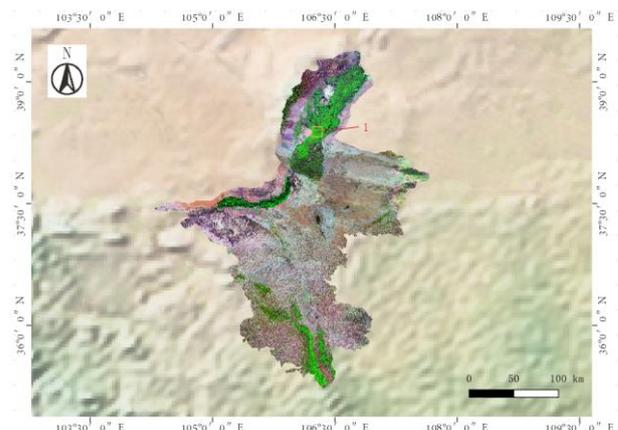

Figure 2. Location of the NHAR

3.2 Experiment procedure

We validated our decomposition model by a 16-day composition NDVI time series of forest type. Satellite image time series were collected from January 2011 to December 2020 (denoted as yellow block 1 in Figure 2), which was used to illustrate and discuss the time series decomposition procedure.

Figure 3 illustrates the time series of block 1, shown in Figure 2. The EMD decomposition result of this time series is further illustrated in Figure 4, and it was decomposed into five IMFs and one residue.

In figure 3, the NDVI value of this time series was at a relatively low value from 2011 to 2016, with little change. But from June 2016 to 2020, the value of NDVI had a relatively large increase, and then entered a period of equilibrium.
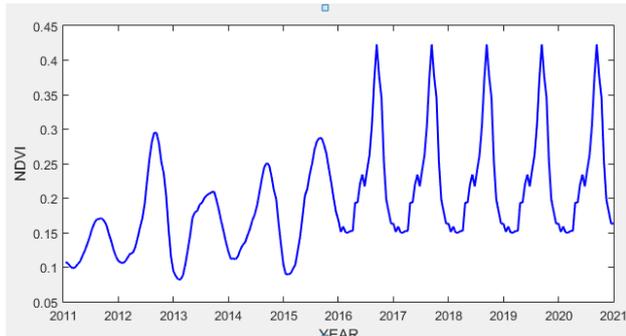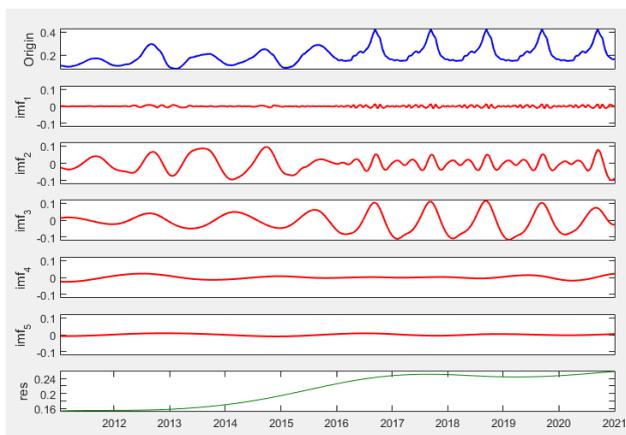

Figure 3. NDVI time series of block 1

Figure 4. NDVI time series decomposed results using EMD

In figure 4, the time series (blue line) is decomposed into 5 IMFs (red line) and a residue (green line). The imf_1, which was close to white noise, was subtracted from the original. Two components, imf_2 and imf_3, represent high-frequency component of the time series. which shown seasonal, weekly changes in vegetation. It shown period changes of vegetation. Imf_4 and imf_5 were low frequency components.

Components imf_2 and imf_3 represent the high-frequency component of the signal, which shows seasonal, periodic changes in vegetation. Imf_4 and imf_5 represent the low frequency components, which together with the res component constitute the trend term, representing the slow, trended change of vegetation. In res curve, NDVI value increased slowly from 2011 to 2013, and increased relatively rapidly between 2014 and 2017, and the trend of vegetation has a process of decreasing first and then increasing from 2017 to 2021.

To verify the validity of seasonal components, we analyzed the correlation between the biophysical characteristics of vegetation and the IMF2 and IMF3, taking into account that the NDVI is derived from the red and near-infrared reflectance ratio. It is easy to observe that the oscillation of the near infrared and red spectral reflectance time series is closely related to the IMF2 and IMF3 obtained by decomposing the NDVI time series. It is noted that the spectral reflectance time series in the near infrared (red) band have similar wave properties. In particular,

the semi-annual wave of IMF2 is caused by the semi-yearly wave of reflectivity in near-infrared band, while the annual wave in IMF3 is the joint result of near-infrared and red band(Kong, 2015).

The average IMF period after EMD decomposition is calculated, which is equal to the ratio of the number of samples of the IMF to its maximum value points (or minimum value points). We first obtain the IMF components corresponding to the semi-annual period and one-year period of each time series, and then carry out K-means clustering analysis on these two components(Kong, 2015). According to the characteristics of land cover types in the study area, the land features are divided into 10 categories. The color configuration table is shown in Table 1.

INDEX	TYPE	VOLUATION	COLOR
1	No data	0	
2	Cultivated land	10	
3	Forest	20	
4	Grassland	30	
5	Shrubland	40	
6	Wetland	50	
7	WaterBodies	60	
8	Artificial surface	80	
9	Bareland	90	
10	Permanent snow and ice	100	

Table 1. Figure Color Configuration Table

The K-means clustering results and the clustering centers of the semi-annual time series are shown in Figure 5 and 6.
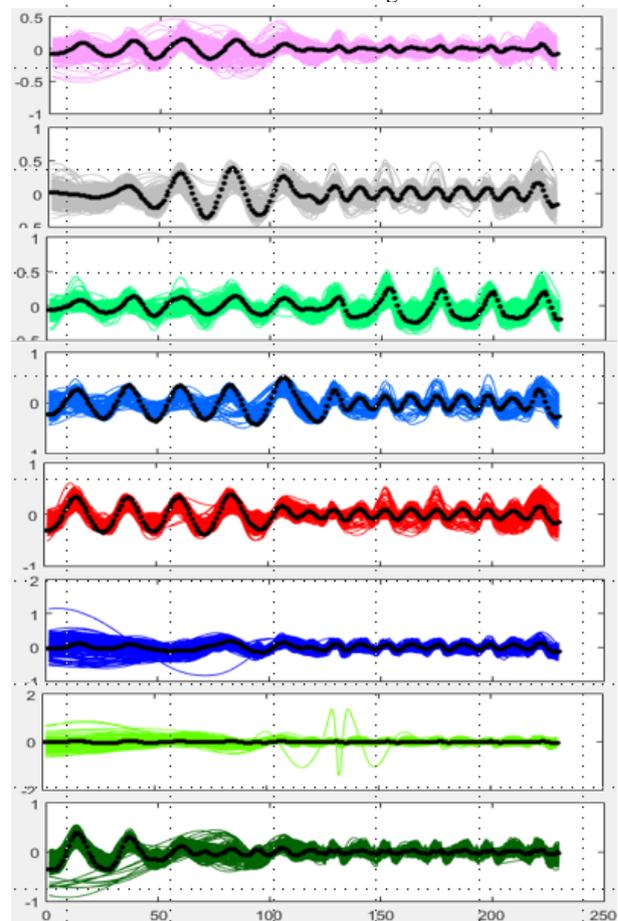
Figure 5. semi-annual time series k-means clustering results

In Figure 5 and 6. "*" represents the cluster center, and show the approximate proportion of each category. Because the isolated points and noise points will have a large impact on the k-means clustering results, the interpolation of missing values and the processing of noise such as cloud pollution are the key points. In this experiment, the results of time series pretreatment are not ideal, and need to be further strengthened.

The results of K-means clustering for the NDVI time series of the study area are shown in Figure 7 and Figure 8.
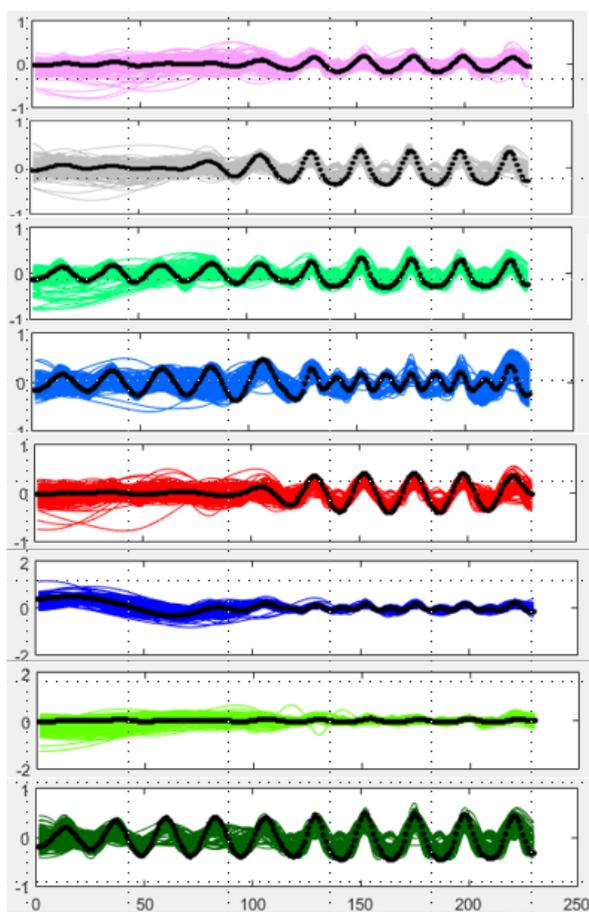

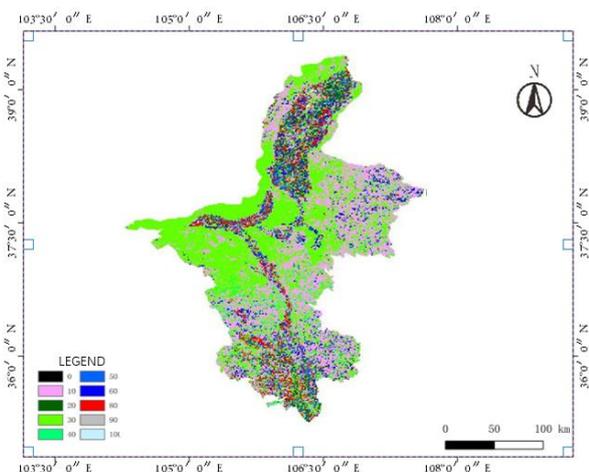Figure 6. annual time series k-means clustering results

Figure 7. k-means clustering results for semi-annual time series

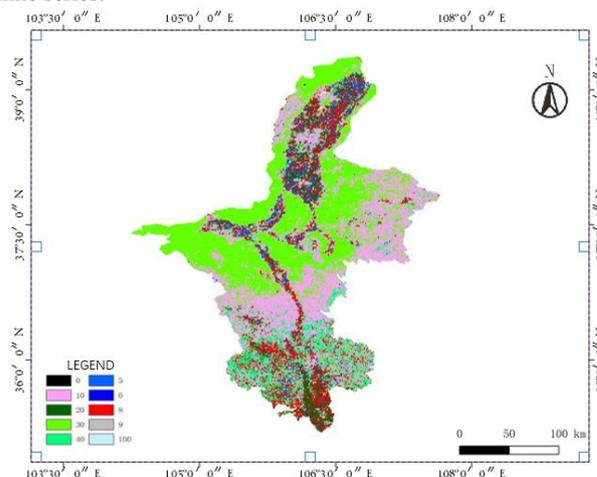Figure 7 shows the k-means clustering results for semi-annual time series.

Figure 8. k-means clustering results for a one-year period time series

Figure 8 shows the k-means clustering results for a one-year period time series. From the two clustering results, the results of the two clusters are very similar, and the change of the clusters reflects the influence of local climate and policy factors on land cover change.

From the results of Figure 7 and Figure 8, the fragmentation of classification results is serious. The reasons for this problem may be as follows: Firstly, there may be some problems in the temporal sequence preprocessing method, resulting in outliers and noise points affecting the EMD decomposition results; Secondly, K-means clustering algorithm is sensitive to mutation point, which makes the clustering results scattered. Third, there is no post-processing, and part of the clustering results should be classified and merged.

4. CONCLUSIONS

Due to the influence of vegetation phenology, the surface of a certain area usually presents seasonal changes with a period of half a year or a year, which causes periodic fluctuations in the satellite image sequence of the area. This fluctuation pattern is usually stable and is determined by the local land cover type. When the land cover changes, the fluctuation pattern of the satellite image series will change, from one kind of land to another.

In this paper, remote sensing time series has the characteristics of seasonality, instability, regionality, multiscale, spatiotemporal autocorrelation, The semi-annual and one-year cycle clustering results of K-means clustering based on empirical mode decomposition are introduced, and the results are analyzed, as well as the relationship with local climate and policy factors.

Based on the current relevant research status, the detection of dynamic change information in remote sensing time series images mainly reflects the following development characteristics:

(1) A large amount of accumulated historical data and the newly launched high temporal resolution remote sensing sensors provide basic data guarantee for the research in this field.

(2) Sustained change is a relatively slow change in land use and cover type.

Land surface change includes some specific change patterns, such as land use / cover of cultivated land / unused land change to land of construction, grassland, forest, bare land and sandy land. Using the time series data, we can get the characteristics

of the surface changes with time, which provides an effective data source for the information mining of the surface changes.
(3) Multi-feature temporal sequence decomposition of multidimensional EMD. Traditional EMD decomposition for a single feature, but the single feature is not enough to reflect the characteristics of land cover change, In order to give full play to the advantages of remote sensing data and help to extract spatio-temporal dynamic features of remote sensing time series, it is necessary to try multi-feature and multi-dimension time series EMD decomposition, which can not only simplify the analysis method, but also can verify each other and improve the accuracy of detection.

ACKNOWLEDGEMENTS

REFERENCES

Chan, K.L., Qin K., 2017: Biomass burning related pollution and their contributions to the local air quality in Hong Kong. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W7, 29-36. doi.org/10.5194/isprs-archives-XLII-2-W7-29-2017.

Alcaraz-Segura, D., Chuvieco, E., Epstein, H. E., Kasischke, E. S., Trishchenko, A., 2010. Debating the greening vs. browning of the north American boreal forest: differences between satellite datasets. Global Change Biology, 16(2), 760-770.

Cihiar, J., Laurent St, et al., 1991. Relation between the NDVI and ecological variables. *Remote Sensing of Environment*, 35(3), 279-298.

de Beurs, K. M., & Henebry, G. M., 2005. A statistical framework for the analysis of long image time series. International Journal of Remote Sensing, 26, 1551−1573.

Eklundh, L., 1998. Estimating relations between AVHRR NDV I and rainfall in East Africa at 102 day and monthly time scales. International Journal of Remote Sensing, 19(3), 563-568.

Gloersen, P., Huang, N.,2003. Comparison of interannual intrinsic modes in hemispheric sea ice covers and other geophysical parameters. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1062–1074. 19.

Huang, W., Shen, Z., N.E.Huang., et al, 1998. Use of intrinsic modes in biology: examples of indicial response of pulmonary blood pressure to step hypoxia. Proc. Natl Acad. Sci. USA, 1998, 95: 12766-12771.

Huang, W., Shen, Z., N.E.Huang, et al 1999. Nonlinear indicial response of complex nonstationary oscillations as pulmonary hypertension responding to step hypoxia. Proc. Natl Acad. Sci. USA, 1999, 96: 1834-1839.

Kaplan, N., Erer, I., 2013. An additive empirical mode decomposition based method for the fusion of remote sensing images. In Proceedings of the 6th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 12–14 June 2013; pp. 165–169.

Kong, Y. L., 2015. Anomaly Detection in Remote Sensed Imagery Time Series Based on EMD. Beijing: Institute of Remote Sensing and Digital Earth Chinese Academy of Sciences.

Kong, Y. L., Meng, Y., Li, W., Yue, A. Z., and Yuan, Y., 2015. Satellite image time series decomposition based on EEMD. Remote Sensing, 7(11): 15583–15604 [DOI: 10.3390/rs71115583]

Kumar, K., Reddy, A.R.M., 2017. An efficient k-means clustering filtering algorithm using density based initial cluster centers, Information Sciences, Volumes 418–419,2017, Pages 286-301, https://doi.org/10.1016/j.ins.2017.07.036.

Liang, H., Lin, Z., Mc, Callum RW, 2000. Artifact reduction in electrogastrogram based on the empirical mode decomposition method. Medical and Biological Computing, 2000, 38(1): 35-41.

N.E.Huang, Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., and Liu, H. H.,1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, in Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, 454(1971), The Royal Society, 903-995.

N.E.Huang, Wu, M. L., Qu, W. D., et al, 2003. Applications of Hilbert-Huang transform to non-stationary financial time series analysis. Applied Stochastic Models in Business and industry, 2003, 19: 245-268.

N.E.Huang, Shen, S. S., 2005. *Hilbert-Huang Transform and Its Applications*, World Scientific: Singapore, 2005; Volume 5. 20.

Nicholson, SE., Farrar, TJ., 1994. The influence of soil type on the relationship between NDVI, rainfall, and soil moisture in semiarid Botswana. *Remote Sensing of Environment*, 50(2), 107-120.

Tang, L., Zhao, Z., Tang, P., et al., 2019. SURE-based optimum-length SG filter to reconstruct NDVI time series iteratively with outliers removal[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2019,18(2): 20500010.

Tang, L., Yang, H., J., 2020. Monotonic Vegetation Trend Detection Based on Ensemble Empirical Mode Decomposition in Ningxia, China. Conference Series: Earth and Environmental Science 502 (2020) 012042.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150.

Yu, D., Cheng, J., and Yang Y., 2005. Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings. Mechanical systems and signal processing, 19(2), 259-270.

Zhang, W., Miller, P.A., Smith, B. et al., 2013. Tundra shrubification and tree-line advance amplify arctic climate warming: results from an individual-based dynamic vegetation model. Environ. Res. Lett., 8 (2013), Article 034023.

Zhang, W., Miller, P.A., Jansson, C., et al., 2018, Self-amplifying feedbacks accelerate greening and warming of the Arctic. Geophys. Res. Lett., 45 (2018), pp. 7102-7111.

Vicente-Serrano, SM., Lasanta, T., Romo, Alfredo., 2005. Analysis of spatial and temporal evolution of vegetation cover in the Spanish central Pyrenees: Role of human management. Environment Management, 34(6), 802-818。