

Earth Gravity Tide Signal Decomposition Model Based on PCA and Geophysical Information Extraction

ZHONG Dong-jun ^{a,*}, QUAN Hai-yan ^b

Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming650500, china

^a963240382@qq.com, ^bquanhaiyan@163.com

*Corresponding author

Keywords: Principal Component Analysis; the earth gravity tide signal; geophysical information; decomposition model

Abstract. Principal Component Analysis (PCA) is a main approach based on the second order statistics. It can remove the correlation between the signal components. PCA has been widely used in blind source separation and received attention because of its potential application in signal processing. PCA's principle algorithm, its simulation steps and its application in the earth gravity tide signal are introduced. The decomposition model can help us to analyze the geophysical information in signal. The result shows that PCA is a potential method in the earth gravity tide signal. So we can extract useful geophysical information from it to understand the interior structure of the earth and earthquake precursors.

Introduction

The earth gravity tide is a recurring phenomenon that is caused by the tide generating force of the moon and sun ^[1]. The earth is not completely rigid, but solid matter between elastic and plastic body. And the gravitational pull of the moon and sun not only cause ocean tide and atmospheric tide but the earth gravity tide as well. Although the mass of the sun is bigger than the moon's quality, but the moon's tidal forces is greater than the sun tidal forces. Because the distance of the moon is closer to earth than the sun is. As the other celestial body is far away from the earth, the earth's gravity influence can generally be neglected in the study of the earth gravity tide. Tidal force is the resultant force of the gravity of the sun and moon acting on the Earth's unit particle and inertial centrifugal force. With the role of different points and the change of position between the sun and the moon relative to the earth, the tidal force of the size and direction also changed.

In study of the earth gravity tide, there has been 100 years of history. The measurement of the earth gravity tide signal is an important method of studying earth crust deformation and movement, which means much to Geodynamics and the observation of earthquake prognostic ^[2]. Nearly 40 years, with the continuous improvement of instrument accuracy and stability, plus the wide application of electronic computers, the research of the earth gravity tide has developed rapidly ^[3-7].

Based on the research of the earth gravity tide signal, this paper proposes a new

decomposition model. It is the earth gravity tide signal decomposition model in Fig.1.

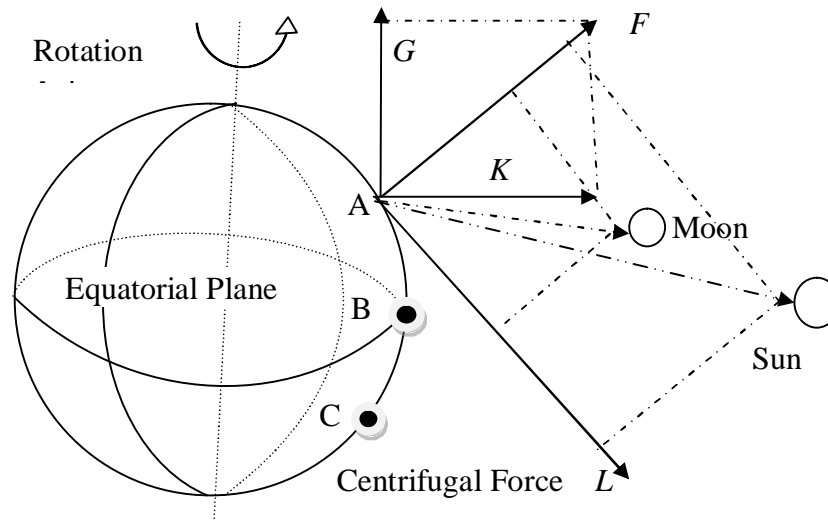


Fig.1. The earth gravity tide signal decomposition model.

The fig.1 shows a case that point A suffered tidal force in a particular place on Earth at a time. Tidal forces that are the interaction of the Sun and the Moon may be resolved into F and L . F 's direction is along the gravity direction. L 's direction is along the centrifugal force direction. F may be resolved into G and K . In fact, G and the axis of rotation are parallel to each other. K and equatorial plane are parallel to each other.

PCA

In practical problems, we often encounter the problem of multiple variables studied. But in most cases, there is often a certain correlation between multiple variables. Due to large number of variables and the correlation between variables, it is bound to increase the complexity of the analysis of the problem. How are we got a few representative variables out of those multiple variables and have them represent the vast majority of information of the original variables? We need to principal component analysis to make the new integrated variable unrelated.

Principal Component Analysis (PCA) is a linear method based on the decomposition of multi-channel signals [8-11]. The new global variables that are unrelated to each other and can represent as much as possible the amount of information of the original variables are a linear combination of the original variables. If choose the first one that is a linear combination of the variables recorded as the first comprehensive x_1 , we want it as much as possible to reflect the information of the original variables. The information is measured by the variance. So the bigger the variance is, the more information is. Therefore, x_1 is called the first principal component that should represent the largest variance in all linear combinations. If the first principal component is insufficient information on behalf of the p original variables, and then consider that the second linear combination should be selected. x_2 is the second principal component. In order to effectively reflect the original information, the available information does not need to re-emerge in

x_1 . And Mathematics language is a kind of special scientific language, which is to reflect and express the relationship between x_1 and x_2 :

$$\text{Cov}(x_1, x_2) = 0 \quad (1)$$

In this way, we can construct the third, fourth principal component and so on.

For Example: p variables: s_1, s_2, \dots, s_p , n array data, so the sample matrix:

$$S = \begin{pmatrix} s_{11} & s_{12} & \mathbf{L} & s_{1p} \\ s_{21} & s_{22} & \mathbf{L} & s_{2p} \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ s_{n1} & s_{n2} & \mathbf{L} & s_{np} \end{pmatrix} = (s_1, s_2, \mathbf{L} s_p) \quad (2)$$

$$\text{And : } s_j = \begin{pmatrix} s_{1j} \\ s_{2j} \\ \mathbf{M} \\ s_{nj} \end{pmatrix} \quad j = 1, 2, \mathbf{L} p$$

Principal Component Analysis is to regard comprehensive observation variables as the new variables.

$$\begin{cases} x_1 = a_{11}s_1 + a_{12}s_2 + \mathbf{L} + a_{1p}s_p \\ x_2 = a_{21}s_1 + a_{22}s_2 + \mathbf{L} + a_{2p}s_p \\ \mathbf{L} \\ x_p = a_{p1}s_1 + a_{p2}s_2 + \mathbf{L} + a_{pp}s_p \end{cases} \quad (3)$$

It can be expressed:

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \mathbf{L} + a_{jp}s_p \quad (4)$$

$$j = 1, 2, \mathbf{L}, p$$

In this way, x_1 is called the first principal component. x_2 is called the second principal component and so on. a_{ij} is the principal component coefficient.

Matrix form:

$$X = AS \quad (5)$$

And:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \mathbf{M} \\ x_p \end{pmatrix} \quad S = \begin{pmatrix} s_1 \\ s_2 \\ \mathbf{M} \\ s_p \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \mathbf{L} & a_{1p} \\ a_{21} & a_{22} & \mathbf{L} & a_{2p} \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ a_{p1} & a_{p2} & \mathbf{L} & a_{pp} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \mathbf{M} \\ a_p \end{pmatrix}$$

A is the principal component coefficient matrix.

According to research, the covariance of principal component analysis should be diagonal matrix. And diagonal elements are the characteristics value of original data correlation matrix. The principal component coefficient matrix is characteristics value with its eigenvector to form.

So x is a hybrid random variable with and mean is not zero. The PCA steps as follows:

The first step is to go to the sample mean of the random variable ^[12].

Hypothesis: $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ is the sample of random variable. And $t = 1, 2, \dots, N$,

the method as follows:

$$\bar{x}_i(t) = x_i(t) - \frac{1}{N} \sum_{i=1}^N x_i(t), \quad i = 1, 2, \dots, n \quad (6)$$

The second step is to find a linear transformation A .

after $\tilde{x} = Ax$, \tilde{x} conforms with the second order statistical independence between the components of the signal. And $R_x = E[\tilde{x}\tilde{x}^H] = I$, $E[\tilde{x}_i\tilde{x}_j] = d_{ij}$. d_{ij} is the Kronecker Delta function.

Assuming: R_x is the correlation matrix of mixed-signal vector. According to the nature of the correlation matrix:

$$R_x = Q\Sigma Q^T \quad (7)$$

The Σ is diagonal matrix. Its diagonal elements and R_x 's characteristics value:

I_1, I_2, \dots, I_N . The column vector of orthogonal matrix Q is corresponding to Ortho-normal eigenvector.

$$A = \Sigma^{-\frac{1}{2}} Q^T \quad (8)$$

Geophysical information extraction based on PCA

In the study of the earth gravity tide, there are two aspects of the data. On the one hand is actual observed data, on the other theoretical calculation data. On the theoretical calculation data, the moon and the sun often deal with separately. The formula for each celestial body is decomposed into many functions of different frequencies. So at any one time, tide force bits can be broken down into many components of different frequencies. The solid earth tidal research and time are closely related. Because Earth has a solid elastomeric lag, so the response of tidal force will have a hysteresis.

The earth gravity tidal wave is a periodic signal. And its periodicity is very complex, including

half-day period component, day period component, half-month period component, month period component, six months period component, one year period component and so on. Experiment data were chosen from the observation points of A, B and C. A's longitude is 110.0, latitude 40; B is on the equatorial plane, longitude 110.0, latitude 0; C (longitude 110.0, latitude -40) and A are symmetric around the equatorial plane. Selecting two symmetrical points is intended to eliminate the influence of the rotary shaft direction and leave the signals of the equatorial plane. The purpose of selecting the point C can easily distinguish the three PCA signals, which is the signal on the equatorial plane or on the axis of rotation (the similar two channel signals is equatorial plane signals, the other signal is the rotating axis signal).

Assuming: 24 hours a day, 30 days per month, 2 years data signal is generated as follows:

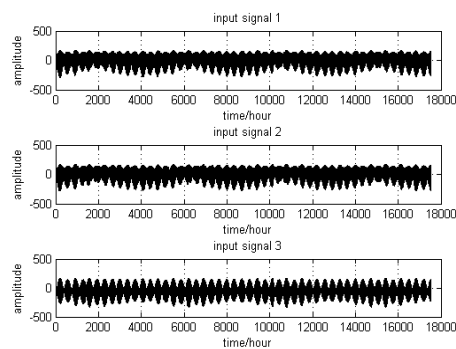


Fig.2.Input signals.

The input signals do PCA. We can get the independent signals as follows:

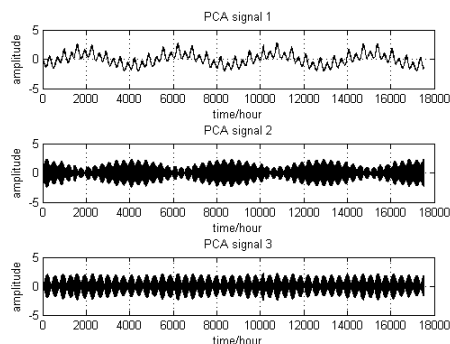


Fig.3.The earth gravity tide signals after PCA.

Form the three signals of Fig.3, we've come to the conclusion that first path is the signal of rotary shaft direction and the others are the signals of the equatorial plane. Through the analysis of the three signal's envelopes and frequency components, we can extract some useful geophysical information.

It reflects the relationship between the magnitude and time in Fig.4.

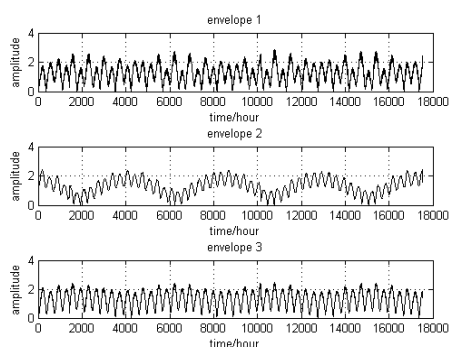


Fig.4.The signal envelopes.

They are the frequency component of the three signals by PCA in Fig.5, Fig.6 and Fig.7. Fig.5 reflects the relationship between amplitude and frequency. It reflects the relationship between frequency and time in Fig.6 and Fig.7. Because in Fig.4, first path signal envelope amplitude fluctuation is large. So in Fig. 5, it is made up of several frequency components and the frequency components are mixed together. We cannot distinguish frequency components in the time domain, so we will analyze it in the frequency domain. In the frequency domain, we will found that it is composed of several main components in Fig.5.

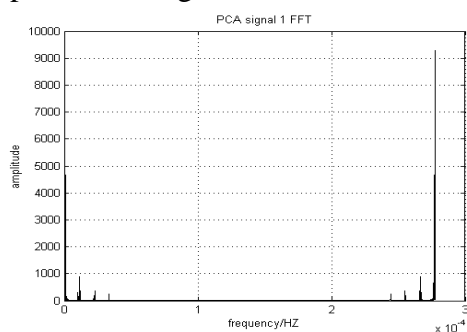


Fig.5.PCA signal 1 frequency domain.

From Fig.5, we can get the datum table.

Table 1 the contrast of between theoretical frequency and actual frequency.

components	half-day period	day period	half-month period	month period	six months period	one year period
Theoretical frequency/Hz	2.3148e-005	1.1574e-005	7.7160e-007	3.8580e-007	6.4300e-008	3.2150e-008
Actual frequency/Hz	2.315e-005	1.118e-005	8.403e-007	4.281e-007	6.342e-008	3.171e-008

In Fig.6, the majority of ordinate values are small-scale fluctuations in the vicinity of 1.158e-005. And the frequency of day period component is just in the vicinity of 1.158e-005. So Fig.6 reflects day period component. There is impulse in some points. Abscissa value of impulse points is about daily values and monthly values, indicating that the day period wave and month period wave are modulated together.

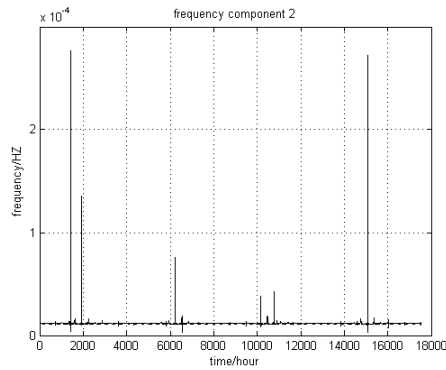


Fig.6.The day period wave.

In Fig.7 the majority of ordinate values are small-scale fluctuations in the vicinity of 2.3148×10^{-5} . And it is the frequency of half-day period wave. So Fig.7 reflects the half-day period components. There is impulse in some points. Abscissa value of impulse points is about half-monthly values, monthly values and six months values, indicating that the half-day period wave, half-month period wave, month period wave and six months period wave values are modulated together.

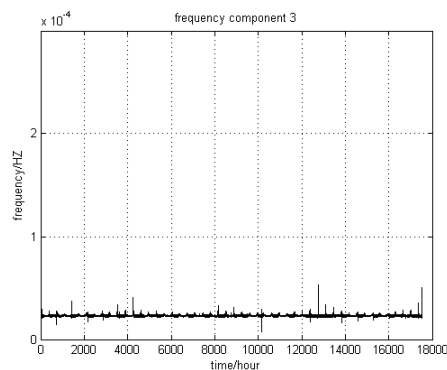


Fig.7.half-day period wave.

Conclusion

Principal component analysis can effectively remove the correlation. Because the earth gravity tide signal have correlation, the principal component analysis can apply to the earth gravity tide signal decomposition, which allows us to understand the basic information of the tidal signal and further knows that the tidal signal reflected geophysical information. Through the establishment of the tidal signal decomposition model, we will decompose this earth tide signal into the equatorial plane and the axis of rotation. Through computer simulation, the equatorial plane is generated half-day period wave and day wave. The axis of rotation direction is generated half-month period wave, month period wave, six months period wave and one year period wave. And the signals on the equatorial plane and the axis of rotation are mutually modulated together. These changes of parameters can great help us to understand the Earth's internal structure and explore earthquake precursors.

References

- [1] Eineder M, Balss U, Gisinger C, et al. TerraSAR-X pixel localization accuracy: Approaching the centimeter level[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. IEEE, 2014: 2669-2670.
- [2] Yague-Martinez N, Fielding E, Haghshenas-Haghighi M, et al. Ground displacement measurement of the 2013 M7. 7 and M6. 8 Balochistan Earthquake with TerraSAR-X ScanSAR data[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. IEEE, 2014: 950-953.
- [3]van Norden M F, Arroyo-Suarez E N, Najjar A S. Concept for military surveys to IHO standards without shore stations using the Real-Time GIPSY (RTG) Global Positioning System (GPS)[C]//OCEANS, 2005. Proceedings of MTS/IEEE. IEEE, 2005: 615-621.
- [4] Eineder M, Minet C, Steigenberger P, et al. Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X[J]. Geoscience and Remote Sensing, IEEE Transactions on, 2011, 49(2): 661-671.
- [5] Minet C, Eineder M, Yague-Martinez N. Haiti earthquake (12.01. 2010) surface shift estimation using TerraSAR-X data[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International. IEEE, 2011: 2488-2491.
- [6] Baek S H, Shum C K. Antarctic ocean tide signal restoration using differential InSAR technique[C]//Synthetic Aperture Radar (AP SAR), 2011 3rd International Asia-Pacific Conference on. IEEE, 2011: 1-4.
- [7] Yague-Martinez N, Eineder M, Cong X Y, et al. Ground displacement measurement by TerraSAR-X image correlation: The 2011 Tohoku-Oki earthquake[J]. Geoscience and Remote Sensing Letters, IEEE, 2012, 9(4): 539-543.
- [8]Nie B, Du J, Liu H, et al. Crowds' classification using hierarchical cluster, rough sets, principal component analysis and its combination[C]//Computer Science-Technology and Applications, 2009. IFCSTA'09. International Forum on. IEEE, 2009, 1: 287-290.
- [9] Lihong Z, Zikui G. Face Recognition Method Based on Adaptively Weighted Block-Two Dimensional Principal Component Analysis[C]//Computational Intelligence, Communication Systems and Networks (CICSyN), 2011 Third International Conference on. IEEE, 2011: 22-25.
- [10]Kottaimalai R, Rajasekaran M P, Selvam V, et al. EEG signal classification using principal component analysis with neural network in brain computer interface applications[C]//Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN), 2013 International Conference on. IEEE, 2013: 227-231.
- [11]Petrolis R, Krisciukaitis A. Multi stage principal component analysis based method for detection of fetal heart beats in abdominal ECGs[C]//Computing in Cardiology Conference (CinC), 2013. IEEE, 2013: 301-304.
- [12]Puyati W, Walairacht A. Efficiency improvement for unconstrained face recognition by weighting probability values of modular PCA and wavelet PCA[C]//Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on. IEEE, 2008, 2: 1449-1453.