Application of Empirical Mode Decomposition of cosmic ray in prediction of great geomagnetic storms

Cong Wang[1],Qian Ye[1],BingSen Xue[1]

[1]Key Laboratory for Space Weather, National Center for Space Weather, Beijing 100029, China

Key Points:

- Great geomagnetic storms
- Empirical Mode Decomposition
- Cosmic ray
- Prediction

Corresponding author: Cong Wang, `wangcong@cma.gov.cn`

Abstract

Influence of interplanetary perturbations on the CR intensity can be an indicator to predicting geomagnetic storm onsets. Case studies illustrating the complexity of the cosmic ray effects and related geomagnetic activity precursors are discussed. It is shown that some indices for cosmic ray activity are good tools for testing the reliability of cosmic ray characteristics for Space Weather forecasts. The use of cosmic ray data for Space Weather purposes is still in its infant stage, but suggestions for both case and statistical studies are made.

1 Introduction

Geomagnetic storms are extreme space weather events that are generally thought to be caused by the interaction between the southward component of the interplanetary magnetic field in the solar wind and the earth's magnetosphere(?, ?).Great geomagnetic storms(Kp$\geq$7) could affect satellites, aircrafts, VLF signal propagation and electric potential of power distribution network(?, ?, ?, ?). Meanwhile, great geomagnetic storms can also affect the ionosphere(?, ?, ?) , magnetosphere(?, ?) and even hurt passengers on airplanes. Hence, the prediction before the sudden commencement of the great geomagnetic storms is very important to prevent these negative effect.

Cosmic rays observed on Earth's surface were modulated by Earth's magnetic, the inhomogeneous magnetic field of the sun and the solar wind(?, ?, ?). Due to the influence of these factors, the features of the cosmic ray flux and anisotropy contain information about the disturbance of interplanetary space(?, ?, ?), which hidden in the recurrent and sporadic (mainly caused by coronal mass ejections(CMEs)) variation of cosmic ray. Through statistical analysis, ? (?) noted that most of Major Geomagnetic Storms (60%) , which occured between 1996 and 2005, were associated with a single CME at the sun, and 27 percent of Major Geomagnetic Storms were associated multiple CMEs. ? (?) investigated all moderate and strong geomagnetic storms between 2007 and 2012, and obtained similar result. Therefore, extracting the disturbance information caused by CMEs from cosmic ray intensity observed on the Earth ground could predict great geomagnetic storms in advance.

Many authers have studied to predicted great geomagnetic storms by analyzing cosmic ray data (?, ?, ?, ?, ?, ?, ?).? (?) presented that the great magnetic storms accompanied by cosmic ray Forbush-effects could be predicted by analyising cosmic ray data. Subsequent, ? (?)firstly systematically investigated cosmic ray precursors of geomagnetic storms. They statisticed and analyzed 14 major geomagnetic storm and 25 large geomagnetic storms observed from 1992 to 1998, and noted that cosmic ray precursors will appears 6 9 hours ahead to the large geomagnetic storms. Thought analysing online one-hour cosmic ray intensity, ? (?) suggested that the Forbush-decrease in cosmic ray could be used for predicting strong geomagnetic storms 10 to 15 hours in advance. In forecasting practice, ? (?) used the deviation between the cosmic ray flux in 8h and the average flux in this period to reflect the cosmic ray fluctuations, and tested this algorithm with data of whole year 2001. The final indicated that the accuracy rate of this algorithm was 80% and the error rate was 20%. ? (?) employed morlet wavelet to extract the abnormal fluctuations of cosmic ray before great geomagnetic storms, and advanced the forecast of great geomagnetic storms caused by CMEs to more than 12h.

To extracting the sporadic variation from cosmic ray flux, we uesd Empriical Mode Decomposes (EMD) to analyze cosmic ray intensity of oulu station. EMD is a key part of Hilbert-Huang transform(?, ?), which could decompose the nonstationary nonlinear signal into a finite set of intrinsic mod function (IMF) and a trend(?, ?). Nonstationary nonlinear signal could be cearly divided into quasi-periodic oscillatory signal and suerimposed random background signal by using EMD(?, ?). EMD has been widely in sapce weather because of its powerfull ability of decomposition which is based on the local characteristc timescal of the data(?, ?, ?, ?, ?, ?, ?, ?)

In the present paper, we use EMD to extract the random background signal of the cosmic ray intensity, which caused by CME. These random background signals contain the information about the disturbance of interplanetary space caused by CME, and will arrive the earth before CME.

Shocksdriven by energetic coronal mass ejections(CME's) and other interplanetary (IP) transients are mainly responsible for initiating large and intense geo- magnetic storms. Observational results indicate that ga- lactic cosmic rays (GR) coming from deep surface inte- ract with these abnormal solar and interplanetary condi- tions and suffer modulation effects.

Most of the events are associated with transient decreases in cosmic ray intensity. Intense storms are having their well defined solar origin as during solar maximum the occurrence rate is 55% while it is only 45% during solar minimum phase of solar cycle.

2 Methodology

2.1 Overview

Interplanetary perturbations is initiated in the solar atmosphere and affect cosmic ray(CR). In some cases their influence on the CR intensity results in data signatures that can possibly be treated as an indicator to predicting geomagnetic storm onsets(?, ?).But in addition, CRI also contains many other signals not related to geomagnetic storm, for instance, in the long run, CRs have inverse relationship with sunspot numbers, for short run, CRs have its own diurnal variations.

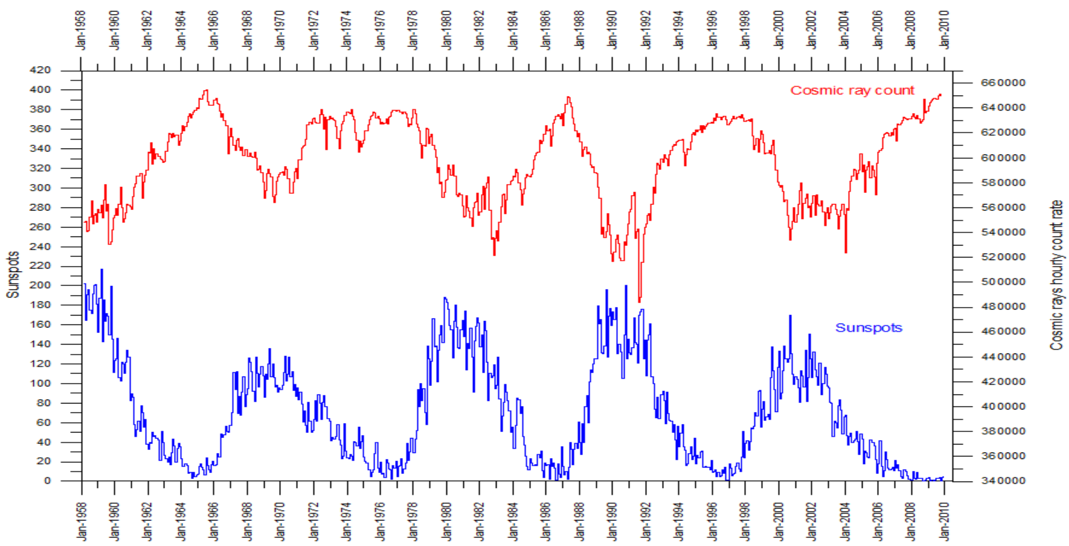

Figure 1. Inverse relationship with sunspot numbers

Signal processing can simplify this problem, traditional signal processing methods can be devided into time-domain methods and frequency-domain methods. Time-domain methods measure max, min, mean values of a signal along time which is easy to understand, but cannot reveal essence of the signal. Frequency-domain decomposes a siganl into a sum of functions according to basis function. Fast Fourier Tansform(FFT) Perhaps the most popular tool for signal processing whose basis function is sinusoidal function. But FFT can not obtain the signal variation of time. As a improvement of FFT, Short Time Fourier Transform(STFT) cotains time information, however, the width of the moving window adopted for the analysis has to be fixed as a function of the minimum frequency of interest, using the best compromise between resolution in both the time and frequency domains(?, ?).

As a time-frequency technique, Empirical Mode Decompisition(EMD) is self-adaptive to dealing with non-stationary and non-linear signals, but it still has "mode mixing" problem. To overcome this problem, we adopt Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) in our analysis, CEEMDAN shows good stability in our work. At last, we use the Hilbert transform on the result of CEEMDAN to monitoring intensity change of signal.

2.2 Empirical Mode Decompisition

The EMD method is a nonlinear signal transformation procedure introduced by Huang(?, ?), this method decomposes a signal into some intrinsic mode function(IMF) through the sfiting process, figure 4 shows the detailed process of sifting process. Each IMF can be obtained from an iterative process of finding the upper and lower envelopes. Cubic-spline is the most popular interpolation method.

Each IMF satisfies the following two conditions:

1. in the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one, and
2. at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

For a given signal $x(t)$, EMD ends up with a representation of the form:

$$x(t) = r(t) + \sum_{n=1}^{N} imf_n(t)$$

$imf_n$is N_{th} intrinsic mode function, and $r(t)$ represents residue corresponding to N intrinsic modes. But EMD has drawback : frequency mixing and un robust, is two figure under same method

2.3 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise

A recently demonstrated improved variant of the EMD method is Ensemble EMD (EEMD), inwhichEMD is performed on an ensemble of initial signals, each perturbed by low-amplitude white noise (Wu and Huang 2009). The noise helps the sifting process to avoid mode mixing and to provide more robust and physically meaningful IMFs. In the end the average of the results is designated as the true final result, and thus the direct effect of the noise is canceled out. Computing the EMD of a large ensemble of signals is computationally more intensive, but this difference in computation time can be reduced significantly since the separate ensemble members can be computed in parallel.

Because the added noise does not completely cancel out in the averaging process for any finite ensemble size, EEMD is no longer a strictly complete decomposition. This issue has been fixed in a EEMD variant called CEEMDAN (Torres et al. 2011). In CEEMDAN, the averaging over all ensemble members is carried separately for each IMF component. By changing the order of averaging over the ensemble and extracting the next IMF, at each point ofthe decomposition procedure the current residual together with the

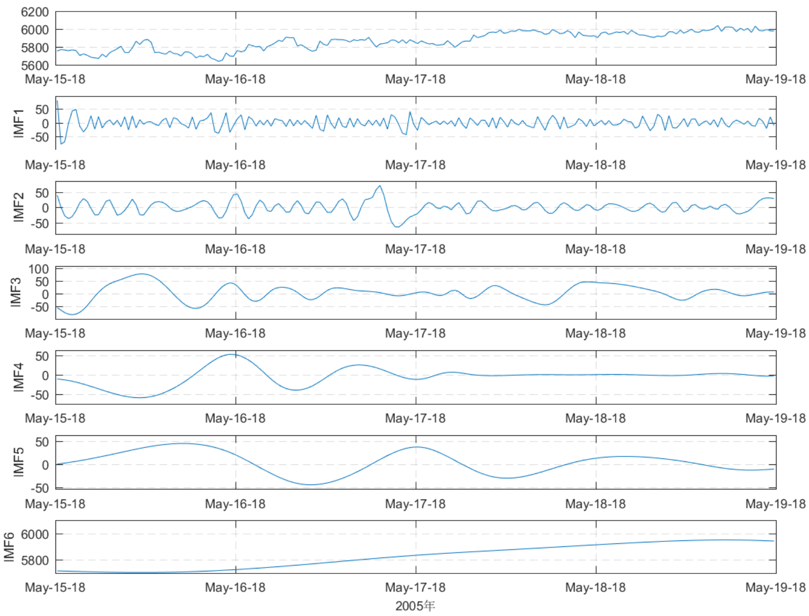

Figure 2. 2005-05-15-18 UTC

already extracted IMFs sums exactly (or up to numerical precision) to the original signal. This small change also seems to improve the algorithm's efficiency in recovering an underlying tone from an already noisy input signal (Colominas et al. 2012).

The EMD has great advantages in dealing with non-stationary and non-linear signals, but it still has "mode mixing" problem. Mode mixing refers to the presence of very similar oscillations in different modes or very disparate amplitude in a mode. By adding Gaussian white noise to the signal, the ensemble empirical mode decomposition (EEMD) algorithm largely eliminates the mode mixing in EMD algorithm[24]. However, the EEMD algorithm cannot completely eliminate Gaussian white noise after signal reconstruction, it cause reconstruction errors. To solve these problems, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was proposed as an improved version of EEMD[25]. It eliminate mode mixing more effectively, the reconstruction error is almost zero, and the cost of calculation is greatly reduced.

2.4 Hilbert Transform

2.5 Our Method

We

3 Result? Case Study

4 Discussion

Acknowledgments

Enter acknowledgments, including your data availability statement, here.

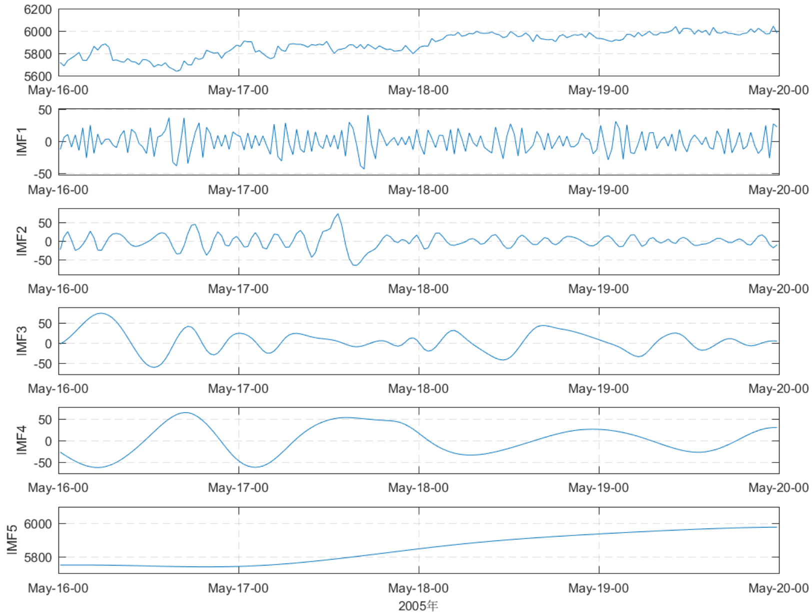

Figure 3. Process of EMD

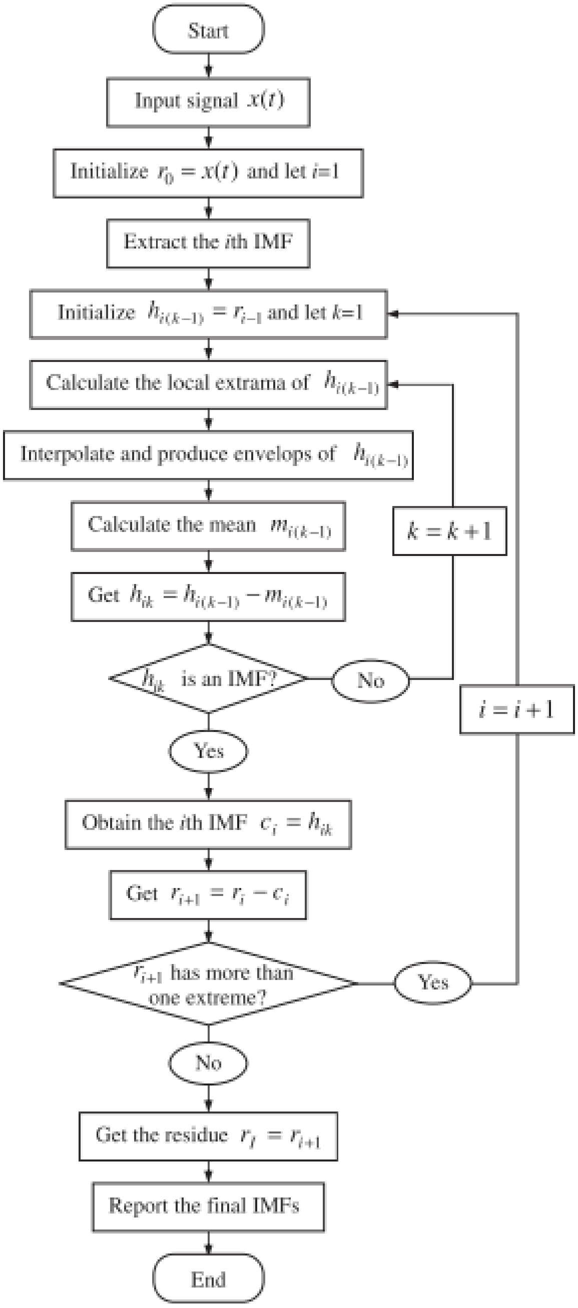

Figure 4. Process of EMD

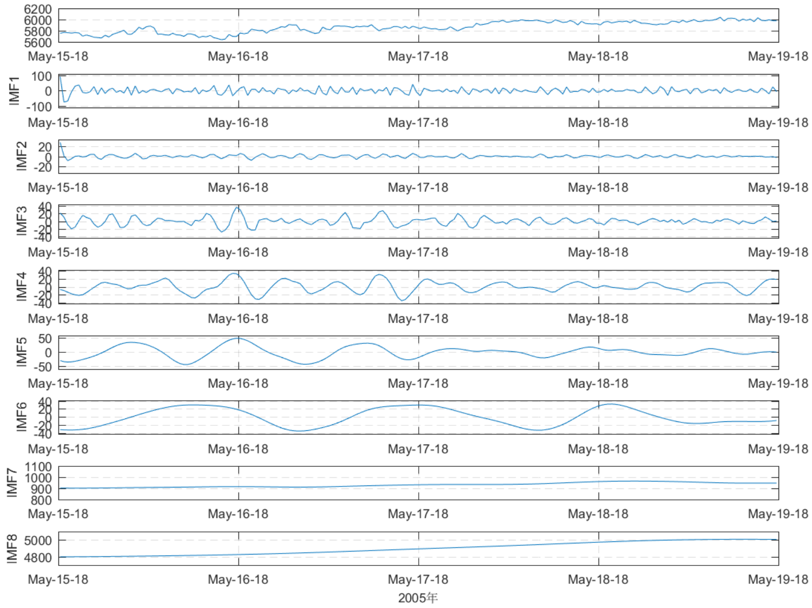

Figure 5. 2005-05-15-18 UTC

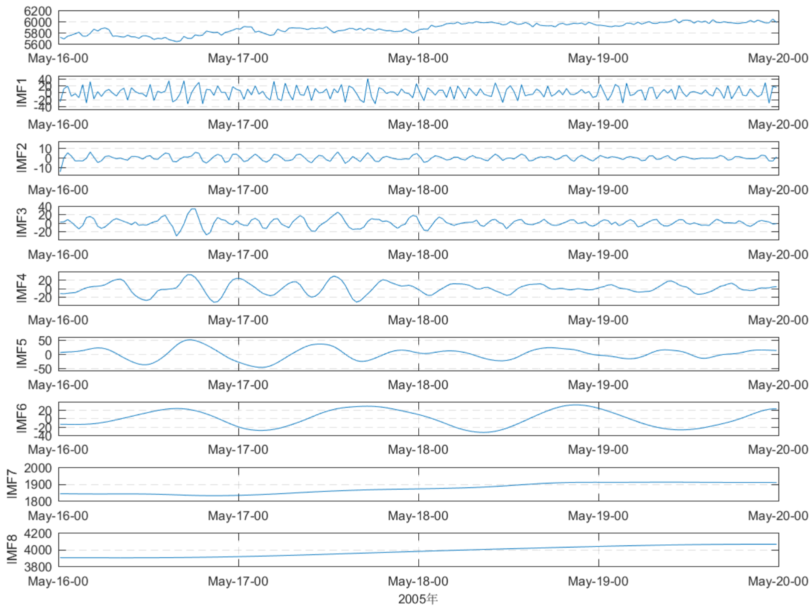

Figure 6. Process of EMD