Astronomy and Computing 43 (2023) 100695

Contents lists available at ScienceDirect

Astronomy and Computing

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

Full length article

An investigation of the relationship between the CME and the Geomagnetic Storm

A. Chattopadhyay *, M.H. Khondekar

Electronics and Communication Engineering, Dr. B.C. Roy Engineering College, Durgapur, India

ARTICLE INFO

Article history:
Received 21 June 2022
Accepted 16 February 2023
Available online 24 February 2023

Keywords:
Planetary K-index
Coronal mass ejection
Causal relationship
Phase relationship
Cross correlations

ABSTRACT

The primary objective of this work is to look for any conceivable link between the geomagnetic storm and coronal mass ejection (CME) events. The relevant data for both CME and geomagnetic storm occurrences have been obtained from the Solar and Heliospheric Observatory mission's LASCO and the NOAA Space Weather Prediction Center, respectively, for the same time span (February 1999 to December 2007). We performed Multivariate Singular Spectral Analysis (MSSA), Semblance Analysis, and Multifractal cross-correlation detrended fluctuation (MFXDFA) analysis to achieve our target. We also applied MSSA to investigate the causal link between the two data sets mentioned above. The analysis identifies the supportiveness factor between the data series and the possibility of a one-way statistical relationship between the data series is also verified. Then, Semblance and MFXDFA were used to detect the presence of any conceivable local phase relationship (time and frequency-dependent) and cross-correlation in the data sets. To make our analysis more persuasive, we must use all three methods. The Semblance analysis reveals positive and negative phase correlation between the signals under investigation at different time sub-intervals, whilst the MFXDFA validates the long-term power-law cross-correlations.

1. Introduction

The Earth's magnetic field can be considered our home planet's only defence against high-energy particles and harmful radiation emitted by the Sun. The scientific community is aware that significant disturbances in the geomagnetic field caused by various solar activities may jeopardise the functioning of critical infrastructures reliant on space-based assets and have terrestrial consequences. However, the underlying mechanisms are still poorly understood. This work has attempted to investigate the statistical association and nature of dependency between the Coronal mass ejection (CME) and the geomagnetic storm occurrence by analysing CME linear speed and Planetary K-index data.

Coronal mass ejection (CME) is a type of extreme solar event that happens when a giant cloud of magnetically charged plasma is ejected from the Sun's corona and travels at extremely high speeds (thousands of kilometres per second) into interplanetary space. When a CME collides with our planet's atmosphere, it injects large amounts of energy into the Earth's magnetosphere, which causes a massive disturbance in the geomagnetic field and triggers a geomagnetic storm. An intense geomagnetic storm poses a severe threat to our civilised society, as it can disturb almost every aspect of modern technology, like power grids, global communication navigation systems and complicate modern civilisation's operations (Hapgood, 2012; Kappenman, 1996, 2012; Thomson et al., 2010) The magnitude of the geomagnetic storms can be characterised and quantified by the value of the Planetary K-index (Kp-index) which is based on a global average of abnormal geomagnetic field variations (Menvielle and Berthelier, 1991). The Kp-index ranges from 0 to 9, with an intense geomagnetic storm defined as one with a value greater than or equal to 5. NOAA Space Weather Prediction Center (SWPC) derives the estimated 3-hour Planetary K-index using the data from ground-based magnetometers located in different countries.

The CME linear speed data were collected from February 1999 to December 2007 as obtained from the Solar and Heliospheric Observatory (SOHO) mission's Large Angle and Spectrometric Coronagraph (LASCO) as compiled in the CME catalogue[1] (Gopalswamy et al., 2009). Gopalswamy, et al. identified four primary attributes of a CME as 'linear speed', 'width', 'CPA', and 'acceleration' in their research work, whereas the 'linear speed' attribute has been taken here as the most appropriate parameter compared

* Corresponding author.
E-mail addresses: chatterjee.anirban7@gmail.com (A. Chattopadhyay), hossainkm_1976@yahoo.co.in (M.H. Khondekar).

[1] http://cdaw.gsfc.nasa.gov/CME_list/index.html.

https://doi.org/10.1016/j.ascom.2023.100695
2213-1337/

to the other three parameters for the investigation (Chattopadhyay et al., Stationarity and periodicities of linear speed of coronal mass ejection: a statistical signal processing approach, 2017) Similarly, the Kp-index data are chosen for the same period (February 1999 to December 2007) from the NOAA Space Weather Prediction Center by determining the average value of the eight 3-hour Kp indices per day for the analysis.[2] Here, an effort has been made to discover the statistical association between the geomagnetic storm and coronal mass ejection by analysing the Kp-index and CME speed signal for solar cycle 23.

Many geomagnetic storm occurrences have been reported during the solar cycle 23, but only a handful have influenced the Earth's environment. Intense geomagnetic storms were observed on October 4–7, 2000, and substantially influenced the ionosphere. Though the ion temperatures of the terrestrial magnetotail generally used remain consistent, it has been noticed that in October 2000, the measured ion temperature was 2–3 times more than the average value (Keesee et al., 2008). Another geomagnetic storm event was noticed on 21 October 2001, and its effect has been explained in Jordanova et al. research paper (Jordanova et al., 2008). They have observed the massive loss of electrons in the radiation belt into the atmosphere, and electron flux drops out due to the outward radial diffusion. In October 2003, various intense geomagnetic storm events were registered. As an effect of these geomagnetic storm events, different incidents have been recorded by multiple researchers like increased photoionisation effects in the dayside ionosphere (Villante and Regi, 2008), enormous ionospheric disturbances which had yielded a 'swirling' effect in a direction opposite to the Earth's rotation (Gopalswamy, 2009) and high-value GIC (geomagnetically induced currents) which is always a threat to our power grids (Ni, 2017) During the August–September 2005, various high-value Kp index peaks have were noticed while among them 24 August, 31 August, 11 September and 13 September most intense geomagnetic storm events were recorded with a peak K_p index value of ≈8 (Papaioannou et al., 2009). The geomagnetic storm events were recorded on 14–15 December 2006 when Earth's magnetosphere is affected by the CME-associated interplanetary shock. Ionospheric disturbances have been recorded due to these geomagnetic storm events, which produce increases and decreases in electron densities and total electron content (TEC) and also provide GIC (Sahai et al., 2012; Zhao et al., 2008; Suvorova et al., 2015). In space research, Earth's magnetosphere is always an exciting domain for researchers to explore (Horvath and Lovell, 2017; Holappa et al., 2014; Mansilla, 2013; Echer et al., 2006; Huttunen et al., 2008). Here, the causal relationship between CME and the geomagnetic storm has been investigated using Multivariate Singular Spectral Analysis (MSSA) technique (Hassani et al., Forecasting UK Industrial Production with Multivariate Singular Spectrum Analysis, 2013). The long-term power-law cross-correlations between these two signals have also been verified using the Multifractal cross-correlation detrended fluctuation analysis (MF-X-DFA) algorithm (Zhou, 2008). Also, to reveal the local phase relationship between two signals using continuous wavelet transform (CWT) based Semblance Analysis has been performed (Cooper and Cowan, 2008).

The rest of this work is planned as follows. In Section 2, the formal description of the algorithms for MSSA, CWT-based Semblance Analysis and MF-X-DFA methods are summarised. Section 3 includes different plots of our results along with a detailed discussion of the crucial observations, while the paper concludes in Section 4 with a summary of the results.

[2] https://www.swpc.noaa.gov/products/planetary-k-index.

2. Algorithms for analysis

2.1. Multivariate Singular Spectrum Analysis (MSSA)

MSSA technique evolved from the basic SSA algorithm, which is a nonparametric method and can be worked with any statistical process without making any assumptions about its stationarity and nonlinearity. Here, the causal relationship between two time series signals (CME and Kp-index) has been investigated by using the MSSA algorithm. This method can be summarised as follows (Hassani et al., Forecasting UK Industrial Production with Multivariate Singular Spectrum Analysis, 2013) (Khondekar et al., 2012).

2.1.1. Basic SSA

Consider the one-dimensional real-valued nonzero time series S_t of length $l(1 \le t \le l)$ is converted into the $L \times K$ trajectory matrix $\psi = \psi_1, \ldots, \psi_k = \left(s_{n,m}\right)_{n,m=1}^{L,k}$ with lagged vectors $S_n = [s_n, s_{n+1}, \ldots, s_{n+L-1}]^T \in \Re^L$ where, $k = l - L + 1$ is the transformation step, and L is the window length. This ψ is a Hankel matrix where diagonal elements are always equal i.e. $\left(\psi_{n(m-1)} = \psi_{(n-1)m}\right)$. Construction of the trajectory matrix is the first step of basic SSA. In the second step, singular value decomposition (SVD) is applied to a new matrix $Q = \psi\psi^T$ which provides L eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \lambda_L \ge 0$ and the corresponding eigenvectors $E_1 \cdots E_L$ where each eigenvector is normalised to the corresponding eigenvalue. The SVD of the trajectory matrix ψ can be expressed as

$$\psi = \sum_{i=1}^{d} \psi_i \tag{1}$$

where, $\psi_i = \sqrt{\lambda_i} E_i P^T$ and $d = \max\{i : \lambda_i > 0\}$.

$\sqrt{\lambda_i}$, E_i and P_i are called as ith 'eigentriple' of the SVD where $P_i = \psi^T E_i / \sqrt{\lambda_i}$; $(i = 1, \ldots, d)$. The third step is all about forming different groups by splitting the elementary matrices and summing the matrices within each group. Considering $I_1 = \{i_1, \ldots, i_p\}$ be the first subset of indices $i_1, \ldots, i_p$. Then the resulting matrix corresponding to the subset of I_1 is defined as $\psi_{I_1} = \psi_{i_1} + \cdots + \psi_{i_p}$. In the final step, the matrix $\tilde{\psi} = \sum_{k=1}^{p} E_{i_k} E_{i_k}^T \psi$ is determined as an approximation to the trajectory matrix ψ. Now initial time series S_t can be estimated by averaging over the diagonals of the matrix $\tilde{\psi}$.

2.1.2. Causality test

To find out the causality, the forecasted values are needed to determine first by using SSA and MSSA methods. Then, these forecasted values are compared with the actual values to evaluate the error values for both SSA and MSSA. If the error value obtained using MSSA is considerably smaller than that obtained using SSA, then it suggests a causal relationship between these time series (Hassani et al., 2010).

The causality test method can be summarised using the MSSA algorithm (Chattopadhyay et al., Multivariate Singular Spectral Analysis (MSSA) to explore geomagnetic storm and CME bond, 2019). Consider R and S are two trajectory matrices of two time series signals $R_N = r_1, \ldots, r_N$ and $S_N = s_1, \ldots, s_N$ respectively where L is the window length. The joint trajectory matrix U can be defined as

$$U = \begin{bmatrix} R \\ S \end{bmatrix} \tag{2}$$

Now, R_N is separated into two parts $R_X = r_1, \ldots, r_{N-f_m}$ and $R_Y = r_{N-f_m+1}, \ldots, r_N$ where $f_m \le N/10$ is the number of data

required to forecast. The sub-signal R_X is used to determine the noise-free signal $\tilde{R}_X^{SSA}$ using SSA considering the trajectory matrix is R^X.

$$\tilde{R}_X^{SSA} = r_1^{SSA}, \ldots, r_{N-f_m}^{SSA} \tag{3}$$

Like R_N, time series S_N also split into S_X and S_Y for the same window length. Let, S^X be the trajectory matrix of S_X then the joint trajectory matrix U^X for R_X and S_X can be expressed as

$$U^X = \begin{bmatrix} R^X \\ S^X \end{bmatrix} \tag{4}$$

With this trajectory matrix U^X the reconstructed noise free signal $\tilde{R}_X^{MSSA}$ is determined from R_X and S_X using MSSA algorithm, given as

$$\tilde{R}_X^{MSSA} = r_1^{MSSA}, \ldots, r_{N-f_m}^{MSSA} \tag{5}$$

Here, Iterated forecasting algorithm (Golyandina et al., 2001) is used to forecast $\tilde{R}_Y^{SSA}$ and $\tilde{R}_Y^{MSSA}$ from noise free signals $\tilde{R}_X^{SSA}$ and $\tilde{R}_X^{MSSA}$ respectively for next f_m data points (Chattopadhyay et al., Multivariate Singular Spectral Analysis (MSSA) to explore geomagnetic storm and CME bond, 2019). The supportiveness between two signals is judged by the criterion discussed in the paper (Hassani et al., 2010), and can be expressed as

$$\Im_{R|S} = \frac{\Delta_{MSSA}}{\Delta_{SSA}} \tag{6}$$

where Δ_{MSSA} and Δ_{SSA} are the root mean square error (MSE) between $\tilde{R}_Y^{MSSA}$ with R_Y and $\tilde{R}_Y^{SSA}$ with R_Y respectively. Similarly, another criterion $\Im_{S|R}$ can be estimated.

Based on this criterion, a conclusion can be made about the forecasting supportiveness of the signals. The existence of a relationship between R and S is validated if the value of $\Im_{R|S}$ is significantly small i.e. $\Im_{R|S} < 1$. Also $\Im_{R|S} < 1$ suggests that S is more supportive of forecasting R. If both $\Im_{R|S}$ and $\Im_{S|R}$ are very small (i.e. $\Im_{R|S} < 1$ and $\Im_{S|R} < 1$) then it verifies the presence of mutual supportiveness between R and S. If $\Im_{R|S} < \Im_{S|R}$, then it suggests that S is more supportive compared to R whereas if $\Im_{S|R} < \Im_{R|S}$, then an opposite conclusion can be made.

2.2. CWT-based semblance analysis

Continuous Wavelet Transform (CWT) technique is generally performed to decompose a signal into wavelets. The CWT of a signal $z(t)$ is described as the inner product of the family of wavelets $\psi_{xy}(t)$ with the signal.

$$CW(x,y) = \left\langle z(t)\, \psi_{xy}(t) \right\rangle = \frac{1}{\sqrt{y}} \int_{-\infty}^{\infty} z(t)\ \psi\left(\frac{t-x}{y}\right) dt \tag{7}$$

Here, x is the translation parameter while y defines various scales where the wavelet can be stretched and $\psi_{xy}(t) = \frac{1}{\sqrt{y}}\psi\left(\frac{t-x}{y}\right)$.

The CWT also can be defined as the convolution of the signal with a scaled version of the wavelet, given as (Rucka and Wilde, 2006)

$$\begin{aligned} CW(x,y) &= \frac{1}{\sqrt{y}} \int_{-\infty}^{\infty} z(t)\ \psi\left(\frac{-(x-t)}{y}\right) dt \\ &= \frac{1}{\sqrt{y}} z(t) * \psi\left(\frac{-x}{y}\right) = z(t) * \bar{\psi}_y(x) \end{aligned} \tag{8}$$

where $\bar{\psi}_y(x) = \frac{1}{\sqrt{y}}\psi\left(\frac{t}{y}\right)$.

The complex Morlet wavelet is used in this work to get better frequency resolution, which can be expressed as (Teolis and Benedetto, 1998)

$$\psi(t) = \frac{1}{\pi f_m} e^{2\pi j f_c t} e^{\frac{-t^2}{f_m}} \tag{9}$$

where f_c is the wavelet centre frequency and f_m tunes the wavelet bandwidth.

Assuming, CW_1 and CW_2 are the continuous wavelet transforms of two time series signals $r(t)$ and $s(t)$ respectively. Since the chosen wavelet is complex here, the CWT will also become complex, and its imaginary and real parts produce a Hilbert transform pair with orthogonality. The cross-wavelet transform can be expressed as (Torrence and Compo, 1998)

$$CW_{1,2} = CW_1 \times CW_2^* \tag{10}$$

The amplitude (A) and local phases (ϕ) of $CW_{1,2}$ can be defined as $A = \left|CW_{1,2}\right|$ and $\phi = \tan^{-1}\left[\frac{\Im(CW_{1,2})}{\Re(CW_{1,2})}\right]$ respectively, for $-\pi \leq \phi \leq \pi$.

Hence, the semblance S can be determined as

$$S = \cos^n(\phi) \tag{11}$$

where n are odd positive values.

The degree of correlation between two signals can be decided based on the value of S. If $S \to 0$, then it can be concluded that a poor correlation is present between the signals. Similarly, the value of S is close to 1 validates that the correlation between two signals is positive, whereas it is negative for $S \approx -1$.

Hence, the value of S is entirely dependent on phase value, not amplitude value, and it is less sensitive to noise. This problem is mitigated by determining the parameter D, which can be defined as (Cooper and Cowan, 2008)

$$D = S\left|CW_1 \times CW_2^*\right| \tag{12}$$

2.3. Multifractal cross-correlation detrended fluctuation analysis (MF-X-DFA)

Assuming $\{r_i\}$ and $\{s_i\}$ two time series of length L where $i = 1, \ldots, L$. MFXDFA algorithm is summarised as follows based on MFDFA algorithm (Chattopadhyay et al., Fractality and singularity in CME linear speed signal: Cycle 23, 2018) (Zhou, 2008):

- Step 1: Determine the signal profile of time series,

$$R(i) = \sum_{n=1}^{i} (r(n) - \bar{r}); \quad i = 1, \ldots, L \tag{13}$$

$$S(i) = \sum_{n=1}^{i} (s(n) - \bar{s}); \quad i = 1, \ldots, L \tag{14}$$

  where $\bar{r}$ and $\bar{s}$ are the sample averages.
- Step 2: Profile $R(i)$ and $S(i)$ are split into L_s number of non-overlapping segments of length s, where $L_s = int(L/S)$. Usually, a short tail remains at the signal profile because the data length L is not a multiple of time scale s. Hence the same process is repeated again but from the opposite end, this time to incorporate this profile tail. Therefore, $2L_s$ segments are obtained altogether.
- Step 3: The variance $F^2(s,m)$ has been computed by determining the local trends for each segment m using least square fit method.

$$\begin{aligned} F^2(s,m) = \frac{1}{s}\sum_{i=1}^{s} &\left|R((m-1)s+i) - \tilde{R}_m(i)\right| \\ &\times \left|S((m-1)s+i) - \tilde{S}_m(i)\right|; \quad m = 1, \ldots, L_s \end{aligned} \tag{15}$$

$$\begin{aligned} F^2(s,m) = \frac{1}{s}\sum_{i=1}^{s} &\left|R(L-(m-L_s)s+i) - \tilde{R}_m(i)\right| \\ &\times \left|S(L-(m-L_s)s+i) - \tilde{S}_m(i)\right|; \quad m = L_s+1, \ldots, 2L_s \end{aligned} \tag{16}$$

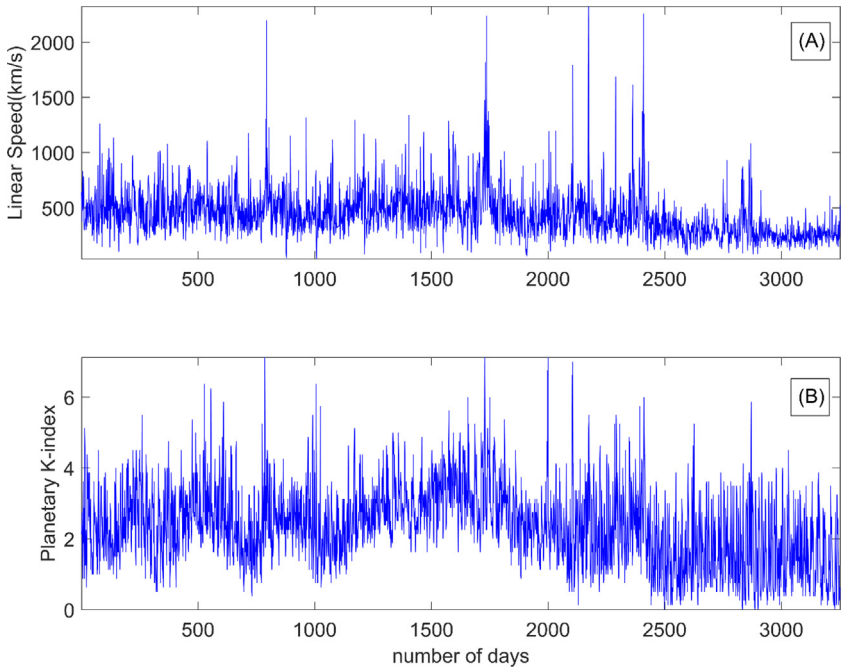

Fig. 1. (A) CME linear speed signal from SOHO; (B) Kp-index signal from the NOAA - SWPC.

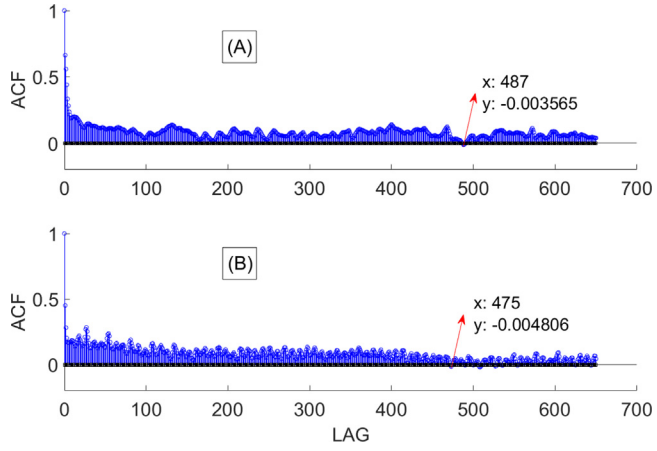

Fig. 2. ACF vs. Lag (A) CME and (B) Kp-index signal.

Here, $\tilde{R}_m(i)$ and $\tilde{S}_m(i)$ are the fitting polynomial in the segment m.

- Step 4: By averaging over all segments, the qth order fluctuation function $F_q(s)$ can be determined as follows,

$$F_q(s) = \left[\frac{1}{2L_s}\sum_{m=1}^{2L_s}\left(F^2(s,m)\right)^{\frac{q}{2}}\right]^{\frac{1}{q}}; \quad q \neq 0 \tag{17}$$

Whereas, if $q \to 0$ then L'Hospital rule is used to determine the fluctuation function as follows,

$$F_0(s) = exp\left[\frac{1}{4L_s}\sum_{m=1}^{2L_s} ln\left|F^2(s,m)\right|\right] \tag{18}$$

- Step 5: In this final step, we determine the slope of the log–log plot of $F_q(s)$ Vs s. If $F_q(s)$ is increasing continuously for large values of s as power-law $\left(F_q(s) \approx s^{h_{rs}(q)}\right)$ then it can be concluded that two time series $\{r_i\}$ and $\{s_i\}$ are long-range cross-correlated.

Here, $h_{rs}(q)$ is known as the cross-correlation Hurst exponent. The cross-correlation behaviour between two signals depends on the relationship between $h_{rs}(q)$ and q. If $h_{rs}(q)$ is dependent on q then the behaviour of cross-correlations between two signals is multifractal otherwise monofractal. The specific value of $h_{rs}(2)$ is known as bivariate cross-correlation Hurst exponent. If the value of $h_{rs}(2)$ is less than 0.5 then it can be concluded that signals are long-range anti-correlated or cross anti-persistent whereas if the value is greater than 0.5 then the signals are cross-persistent or long-range correlated. The cross-correlations between two signals are entirely missing when the value of $h_{rs}(2)$ is equal to 0, or it can be said that for $h_{rs}(2) = 0$, short-range cross-correlations present between signals.

The scaling exponents $\tau_{rs}(q)$ can be determined from the relation between Hurst exponents and scaling exponents, given as

$$\tau_{rs}(q) = qh_{rs}(q) - 1 \tag{19}$$

Eq. (13) can be modified as

$$\tau_{rs}(q) = qh_{rs}(q) - qh_{rs}(1) - 1 \tag{20}$$

If $\tau_{rs}(q)$ is varying non-linearly with q, then the behaviour of cross-correlation is multifractal. The Singularity spectrum can be obtained by using Legendre transformation, given as

$$f_{rs}(\alpha) = q\alpha_{rs} - \tau_{rs}(q) \tag{21}$$

Here, parameter α is known as the Lipschitz–Hölder exponent, which is computed from the slope of $\tau_{rs}(q)$ vs. q curve.

$$\alpha_{rs}(q) = \frac{d}{dq}\tau_{rs}(q) \tag{22}$$

3. Result and analysis

The original Kp-index and CME linear speed signals are represented in Fig. 1(A) and (B), respectively. Different analysis results for the MSSA-based causality test are shown in Figs. 2, 3 and Tables 1, 2. The performance of this method is mostly dependent upon the value of two parameters, (i) window length (L) and

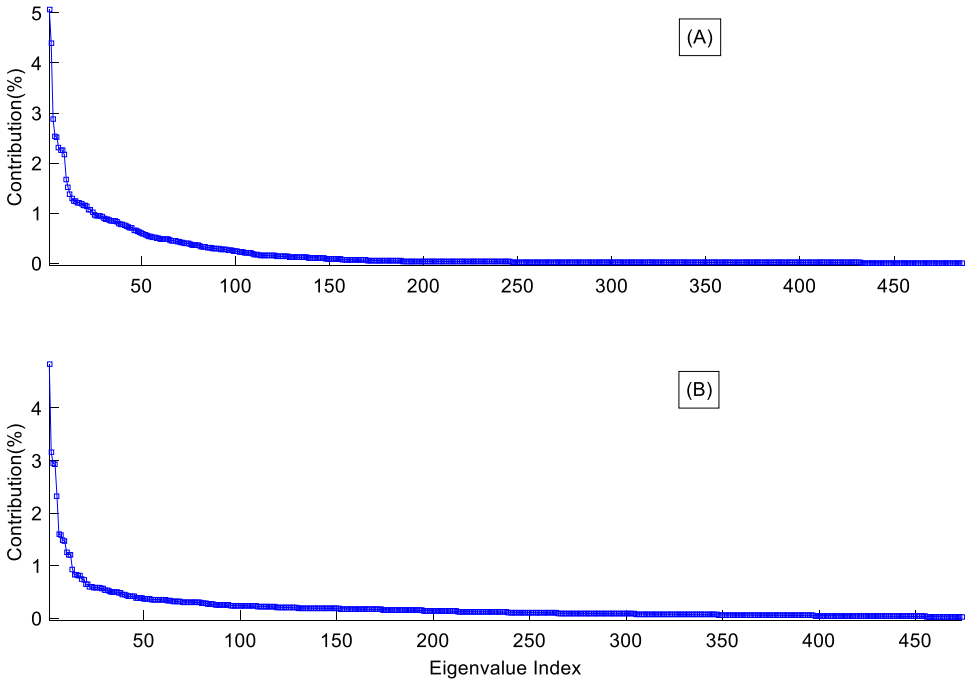

Fig. 3. Energy contribution for (A) CME (B) Kp-index time series.

Table 1
Window length and Eigen triples of Kp-index and CME time series.

Signal	Windows length (L)	Eigentriples (I)
Kp-index	475	1:441
CME	487	1:189

Table 2
Causality factors.

	$\Im_{R\|S}$	$\Im_{S\|R}$
R: CME linear speed & S: Kp-index	1.006203	0.900672

(ii) the number of significant eigentriples (I). Hence, to get an optimum analysis result, these two parameters needed to find out precisely. The window length is determined from the correlogram, where L is chosen as that particular lag value for which the Autocorrelation Function (ACF) intersects the standard Gaussian confidence interval (95% CI) for the first time.

In Fig. 2, it can be seen that 475 and 487 are the respective lag values where the ACFs of the Kp-index and CME speed signals intersect for the first time with 95% CI. Therefore, optimised window lengths have been computed here 475 and 487 for the signals under investigation. The 'Eigen triple grouping' criteria is based on ratio term (R_t), which approximates the energy contribution percentage for the tth principal component of the original signal. The ratio term (R_t) can be defined as,

$$R_t = \frac{\sigma_t}{\sum_{k=1}^{L} \sigma_k} * 100 \tag{23}$$

The group is formed by the number of elements (I_{max}); only those components have significant energy while ignoring others low energy components. The first I_{max} Eigen triples are chosen as the most significant Eigen triples (i.e. $I = 1 : I_{max}$). Fig. 3 reveals the energy contribution of the above signals, while the computed eigentriples (I) and window length (L) for the two signals are given in Table 1.

After determining I and L, the values of $\Im_{R|S}$ and $\Im_{S|R}$ are obtained to reveal the statistical relationship between the two signals, given in Table 2. Assuming here, R and S represent the CME linear speed and Kp-index signal, respectively. In Table 2, it is seen that the value of $\Im_{R|S} > 1$ implies that the information extracted from the CME signal may assist in forecasting the Planetary K-index signal. Also, it can be seen that the value of $\Im_{S|R}$ is less than 1, whereas the value of $\Im_{R|S}$ is greater than 1, which denies the possibility of any feedback system between these two signals. Now to judge which signal is more supportive compared to the other one, we compare between $\Im_{R|S}$ and $\Im_{S|R}$. Here, $\Im_{R|S} > \Im_{S|R}$ this clearly validates that R is more supportive of predicting S. Therefore, the final result of this analysis reveals that a one-way statistical association exists between CME and Kp-index signal and that the supportiveness of CME is more compared to the Kp-index signal.

It may not be enough to evaluate the supportiveness between two time series to investigate the statistical relationship. Therefore, correlation analysis has also been performed in addition to investigating the supportiveness of CME and Kp-index. To attain this purpose, the Semblance and MFXDFA analyses are practically applied. The semblance analysis between the CME and Kp-index time series has been depicted in Fig. 4. The positive phase correlations can be observed between CME and Kp-index at upper scales (wavelengths) between 200 and 500 during June 2003 to February 2004 and May 2005 to October 2005 from the Semblance analysis plot displayed in Fig. 4. There is a negative phase correlation on nearly every scale in between February 1999 to May 2000. From August 2000 to March 2001, a mixed-phase correlation has also been observed, with positive correlation at lower scales (100–300) and negative correlation at higher scales (200–500).

In Fig. 5, the dependency of the $F(q)$, $h(q)$, $\tau(q)$ on q, the behaviour of $\log(F(s))$ vs. $\log(s)$ plot and the nature of the singularity spectrum $f(\alpha)$ are presented for both CME linear speed and Kp-index signals using MFXDFA algorithm. Also, the variation of the $(h_{rr} + h_{ss})/2$ along with q has been shown in Fig. 5*c*. Here, the range of exponent q is taken as -10 to $+10$ with a fixed step interval of 0.5.

The nonlinear nature of $F(q)$ function can be noticed in Fig. 5*a* which advocates for the multifractal characteristics of both signals. It can be seen in Fig. 5*b* that the values of $\log(F(s))$ increase

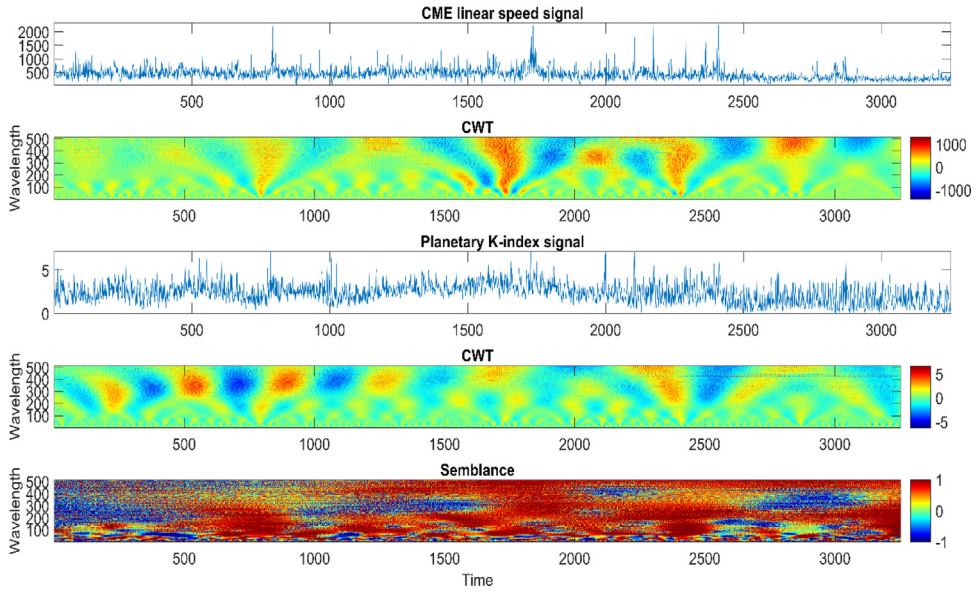

Fig. 4. Semblance analysis for the CME and Kp-index signal.

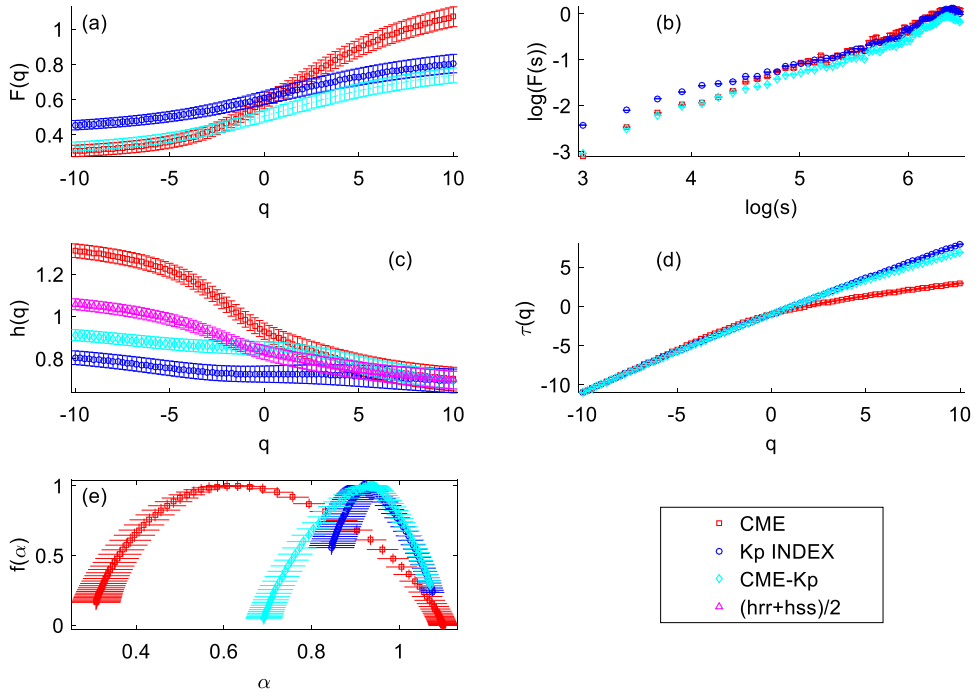

Fig. 5. MFXDFA plots for CME linear speed and Kp-index (a) $F(q)$ vs. q (b) $\log(F\ (s))$ vs. $\log(s)$ (c) $h(q)$ vs. q (d) $\tau(q)$ vs. q (e) $f(\alpha)$ vs. α.

as $\log(s)$ increases which confirms the presence of a power-law cross-correlation between CME and Kp-index. In Fig. 5*c*, the decreasing relationship between $h(q)$ and q has been observed. This nature of the $h(q)$ function reveals that the small fluctuation scaling properties are more dominant than the large fluctuation scaling properties for both signals. Also, the cross-correlation Hurst exponent has been determined from Fig. 5*c*, which is 0.82992 ± 0.036252. It is found that the value of $h_{rs}(2)$ is greater than 0.5, confirming the presence of positive persistence temperament in the cross-correlations. In Fig. 5*d*, the $\tau(q)$ function exhibits nonlinear nature with respect to q, which reveals the possibility of multifractal characteristics in corresponding cross-correlations. The singularity spectrum is one of the key functions used to quantify multifractal behaviour (Shimizu et al., 2002). The degree of multifractality of a signal can be measured by computing the width of the singularity spectrum. The singularity spectrum of MFXDFA has been portrayed in Fig. 5*f*. The maximum and minimum singularity strengths can be denoted as α_{min} and α_{max} respectively while these two parameters are the lowest and highest values of the Hölder exponent α of the spectrum for which $f(\alpha) = 0$. In this work, the computed values of α_{min} and α_{max} for cross correlations analysis using MFXDFA are 0.6892 and 1.068, respectively. The width of the spectrum can be defined as

Table 3
Cross-correlation parameters of CME and Kp-index using MFXDFA.

Cross correlation	$\alpha_{\min}$ $(f=0)$	$\alpha_{\max}$ $(f=0)$	$\Delta\alpha$ $(\alpha_{\max}-\alpha_{\min})$	$h_{rs}(q=2)$
$\text{MFXDFA}_{\text{CME-Kp}}$	0.6892	1.06892	0.37972	0.82992 ± 0.036252

the difference between the α_{max} and α_{min} (denoted as $\Delta\alpha$) which gives measures the length of the range of fractal exponents in the signal, i.e. the degree of multifractality. Here, the computed value of the $\Delta\alpha$ for MFXDFA is 0.37972 (see Table 3).

4. Conclusion

i. The CME Linear Speed information may assist in predicting the Planetary K-index signal.
ii. The calculation possibly indicates that the variation of the Kp-index may be expressed as a function of CME linear speed fluctuations but not vice versa. This, in turn, suggests that CME linear speed fluctuation possibly regulates the time variability of geomagnetic storms and hence possibly helps to determine its precision.
iii. The width of the singularity spectrum suggests that the range of the fractal exponents existing in the cross-correlation is quite significant.
iv. The actual Kp-index values support or agree with the outcome of the forecasting made from the CME linear speed signal to predict the Kp-index.
v. Positive phase correlations have been discovered between the Kp-index and CME.
vi. The CME and Kp, along with their cross-correlation, follow the power law.
vii. The structures in the CME linear speed and Kp-index signals are self-similar. Moreover, the structural patterns of CME linear speed are highly correlated with the Kp-index signal, and this correlation is long-range.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to my data in the article.

References

Chattopadhyay, A., Khondekar, M.H., Bhattacharjee, A.K., 2017. Stationarity and periodicities of linear speed of coronal mass ejection: a statistical signal processing approach. Astrophys. Space Sci. 362 (9), 179, (1-12).

Chattopadhyay, A., Khondekar, M.H., Bhattacharjee, A.K., 2018. Fractality and singularity in CME linear speed signal: Cycle 23. Chaos Solitons Fractals 114, 542–550.

Chattopadhyay, A., Khondekar, M.H., Bhattacharjee, A.K., 2019. Multivariate singular spectral analysis (MSSA) to explore geomagnetic storm and CME bond. In: IEEE-ASPCON. Jadavpur University, Kolkata, pp. 79–82.

Cooper, G.R., Cowan, D.R., 2008. Comparing time series using wavelet-based semblance analysis. Comput. Geosci. 34 (2), 95–102.

Echer, E., Gonzalez, W.D., Alves, M.V., 2006. On the geomagnetic effects of solar wind interplanetary magnetic structures. Space Weather 4 (6), S06001, (1-11).

Golyandina, N., Nekrutkin, V., Zhigljavsky, A., 2001. Analysis of Time Series Structure: SSA and Related Techniques. Chapman and Hall/CRC.

Gopalswamy, N., 2009. Introduction to special section on large geomagnetic storms. J. Geophys. Res. Space Phys. 114 (A3), 1–4.

Gopalswamy, N., Yashiro, S., Michalek, G., Stenborg, G., Vourlidas, A., Freeland, S., Howard, R., 2009. The SOHO/LASCO CME catalog. Earth Moon Planets 104 (1), 295–313.

Hapgood, M., 2012. Astrophysics: Prepare for the coming space weather storm. Nature 484 (7394), 311–313.

Hassani, H., Heravi, S., Zhigljavsky, A., 2013. Forecasting UK industrial production with multivariate singular spectrum analysis. J. Forecast. 32 (5), 395–408.

Hassani, H., Zhigljavsky, A., Patterson, K., Soofi, K., 2010. A comprehensive causality test based on the singular spectrum analysis. Causality Sci. 379–406.

Holappa, L., Mursula, K., Asikainen, T., Richardson, I.G., 2014. Annual fractions of high-speed streams from principal component analysis of local geomagnetic activity. J. Geophys. Res. Space Phys. 119 (6), 4544–4555.

Horvath, I., Lovell, B.C., 2017. Investigating the development of localized neutral density increases during the 24 2005 geomagnetic storm. Space Phys. 122 (11), 765–783.

Huttunen, K., Kilpua, S., Pulkkinen, A., Viljanen, A., Tanskanen, E., 2008. Solar wind drivers of large geomagnetically induced currents during the solar cycle 23. Space Weather 6 (10), 1–8. doi:10.1029/2007SW000374.

Jordanova, V.K., Albert, J., Miyoshi, Y., 2008. Relativistic electron precipitation by EMIC waves from self-consistent global simulations. J. Geophys. Res. Space Phys. 113 (A3), 1–11. doi:10.1029/2008JA013239.

Kappenman, J.G., 1996. Geomagnetic storms and their impact on power. IEEE Power Eng. Rev. 16 (5), 5–8.

Kappenman, J.G., 2012. A perfect storm of planetary proportions. IEEE Spectr. 49 (2), 26–31.

Keesee, A.M., Scime, E., Moldwin, M.B., 2008. Remote measurements of ion temperatures in the terrestrial magnetotail. J. Geophys. Res. Space Phys. 113 (A3), 1–9. doi:10.1029/2008JA013130.

Khondekar, M.H., Ghosh, D.N., Ghosh, K., Bhattacharya, A.K., 2012. An investigation on the relationship between solar irradiance signal from ERBS and 8B solar neutrino flux signals from SNO. Astrophys. Space Sci. 342 (2), 287–301.

Mansilla, G.A., 2013. Variation of total ozone during 24 2005 magnetic storm: a case study. Atmos. Clim. Sci. 3, 475–480.

Menvielle, M., Berthelier, A., 1991. The K-derived planetary indices: Description and availability. Rev. Geophys. 29 (3), 415–432.

Ni, Y.Y., 2017. The statistical analysis of the geomagnetically induced current events occurred in Guangdong, China during the declining phase of solar cycle 23 (2003–2006). IOP Conf. Ser.: Mater. Sci. Eng. 339 (1), 1–5. doi:10.1088/1757-899X/339/1/012013.

Papaioannou, A., Mavromichalaki, H., Eroshenko, E., Belov, A., Oleneva, V., 2009. The burst of solar and geomagnetic activity in august–2005. Ann. Geophys. 27, 1019–1026.

Rucka, M., Wilde, K., 2006. Application of continuous wavelet transform in vibration based damage detection method for beams and plates. J. Sound Vib. 297 (3–5), 536–550.

Sahai, Y.D., Jesus, R., Fagundes, P.R., Selhorst, C.L., Abreu, A.J., Ram …, S.T., Bittencourt, J.A., 2012. Effects observed in the equatorial and low latitude ionospheric F-region in the Brazilian sector during low solar activity geomagnetic storms and comparison with the COSMIC measurements. Adv. Space Res. 50 (10), 1344–1351.

Shimizu, Y., Thurner, S., Ehrenberger, K., 2002. Multifractal spectra as a measure of complexity in human posture. Fractals 10 (1), 103–116.

Suvorova, A.V., Huang, C.M., Tsai, L.C., Dmitriev, A.V., Ratovsky, K.G., 2015. Long-duration positive ionospheric storm during the 2006 geomagnetic storm: Ionizing effect of forbidden electrons. Adv. Space Res. 56 (9), 2001–2011.

Teolis, A., Benedetto, J.J., 1998. Computational Signal Processing with Wavelets (Vol. 182). Birkhäuser, Boston, MA, USA.

Thomson, A.W., Gaunt, C.T., Cilliers, P., Wild, J.A., Opperman, B., McKinnell ., L.A., Lotz, S.I., 2010. Present day challenges in understanding the geomagnetic hazard to national power grids. Adv. Space Res. 45 (9), 1182–1190.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78.

Villante, U., Regi, M., 2008. Solar flare effect preceding halloween storm (28 october 2003): Results of a worldwide analysis. J. Geophys. Res. Space Phys. 113 (A3), 1–10.

Zhao, B., Wan, W., Tschu, K., Igarashi, K., Kikuchi, T., Nozaki ., K., Ning, B.B., 2008. Ionosphere disturbances observed throughout southeast Asia of the superstorm of 20–22 november 2003. J. Geophys. Res. Space Phys. 113 (A3), 1–16.

Zhou, W.X., 2008. Multifractal detrended cross-correlation analysis for two nonstationary signals. Phys. Rev. E 77 (6), 066211.