Investigation of nonlinearity and chaos in solar flare index signal

Shankhachur Mukherjee[1], Swetadri Samadder[2] and Koushik Ghosh[3]
[1] Department of Physics, The University of Burdwan, Golapbag,Burdwan-713104, West Bengal, India
ahlechaos@gmail.com
[2] Department of Mathematics, Fakir Chand College, Diamond Harbour, South 24Parganas, Pin-743331, India
[3] Department of Mathematics, University Instituteof Technology, The University of Burdwan Golapbag (North), Burdwan 713104, West Bengal, India
ahlechaos@gmail.com
(Submitted on 5.11.2020; Accepted on 30.03.2021)

Abstract. In the present work daily Solar Flare Index developed at Mc-Math-Hulbert Solar Observatory during the period 1st January, 1966 to 31st December, 2008 has been taken into consideration. As study of nonlinearity in a signal is quite motivating to understand its intrinsic nature and thereby provides exciting scope to explore further knowledge of the observed phenomenon, so it is indispensable to examine the presence of nonlinearity in the solar flare index signal. Moreover, the presence of nonlinearity is regarded as a direct implication of chaos which serves to be an essential element of complexity, natural to the underlying system or phenomenon. In order to understand this nonlinearity in the present signal Delay Vector Variance (DVV) analysis has been employed here to address our goal, which reveals the presence of some form of nonlinearity in it. Further, to judge the presence of chaos in the said signal, Recurrence Plot (RP) analysis with Recurrence Quantification Analysis (RQA) has also been taken into consideration. The present study possibly enables to establish a deterministic nonlinear chaotic profile of the phenomena of rapid magnetic energy bursts observed in the solar atmosphere.
Key words: Solar Flare Index, nonlinearity, chaos, delay vector variance, recurrence plot, recurrence quantification analysis.

1.Introduction

A solar flare is a sudden brightening observed over the Sun's surface or the solar limb, which is interpreted as energy release of large amount, up to the order of 6×10^{25} Joules, of energy (about a sixth of the total energy output of the Sun each second). Solar flare is a sudden eruption of magnetic energy released on or near the surface of the Sun, usually associated with sunspots, accompanied by bursts of electromagnetic radiation and particles. Solar flares strongly influence the local space weather in the vicinity of the Earth. They can produce streams of highly energetic particles in the solar wind, known as a solar proton event or "coronal mass ejection" (CME). These can impact the Earth's magnetosphere and present radiation hazards to spacecraft, astronauts and cosmonauts. 'Solar Flare Index' is the product of the intensity scale of importance and the duration of the flare in minutes (Kleczek, 1952).

In everyday language, 'chaos' means the situation or state exactly opposite of order. It originates from the Greek word "khaox" meaning 'empty space', specifying the elemental emptiness before the origin of everything else, as described in Greek cosmology. The phenomenon of chaos exhibited by certain dynamical systems in real world possesses infinitely complex pattern of behaviour lying just beyond the edge of total order. A system is said to be chaotic if it is predictable in principle and yet is unpredictable in practice over long periods due to its behaviour of an elevated sensitive dependence on initial conditions. The sensitivity mentioned above leads to an exponential growth of errors generated in the initial condition resulting randomness in the behaviour of the chaotic systems. In other words, if in a phenomenon (a little perturbation) at initial level can give rise to significantly magnified changes in its future levels, then that phenomenon is termed as chaotic. The takeaway from this analysis is that for a chaotic system no long-term prediction should be performed; only short-term prediction can be made after a careful study. Chaos deals with nonlinearity and nonlinearity is a necessary condition for the presence of chaos in a dynamical system.

The present work focuses on the possible nonlinear and chaotic behavior of the solar flare index obtained from Mc-Math-Hulbert Solar Observatory, which is a signal containing daily record of solar flare index (NOAA, 2019).

Here, solar flare index data between Jan 01, 1966 and Dec 31, 2008 is taken under investigation which covers four solar cycles (solar cycles 20–23).

Five measures of solar flare importance are added to obtain this index: (NOAA, 2019):

1. Sudden Ionospheric Disturbance importance (scale 0–3);
2. H-α flare importance (scale 0–3);
3. 10.7 cm solar radio flux magnitude (characteristic of log of flux);
4. Solar radio spectral type (Type II = 1, Continuum = 2, and Type IV with duration greater than 10 minutes = 3) and
5. Magnitude of 200 MHz flux (characteristic of log of flux).

Double Exponential Smoothing (Brown, 1956) has been employed on the present solar flare index signal to de-trend it. Delay Vector Variance (DVV) analysis (Gautama, Mandic and Van Hulle, 2004; Ahmed, 2010) has been carried out over this immediately obtained smoothed signal for the detection of nonlinearity, followed by subsequent analysis of recurrence pattern evident from the Recurrence Plot (RP) (Eckmann, Kamphorst and Ruelle, 1987), with Recurrence Quantification Analysis (RQA) (Webber and Zbilut, 1994; Zbilut and Webber, 1992; Atay and Altntas, 1999; Marwan, Wessel, Meyerfeldt, Schirdewan and Kurths, 2002; Marwan, Carmen Romano, Thiel and Kurths, 2007; Marwan, 2011; Hossain, Ghosh, Ghosh and Bhattacharjee, 2015; Samadder, Ghosh and Basu, 2015) to examine the possibility of chaos occurring in it.

2.Aims and Objectives

The present work aims to explore the behaviour of the solar flare index from the perspective of nonlinearity and chaos. The specific objectives are:

1. to find out whether the profile of the signal of solar flare index is nonlinear
2. to explore if solar flare index is deterministic or stochastic
3. to investigate further whether chaos is present in the signal of solar flare index if the profile is nonlinear.

3.Research Questions

The following questions are investigated in the present work:

First, does the mechanics of reorganization of magnetic loops, resulting from rapid conversion of a large amount of magnetic energy previously stored in the solar corona and dissipated through magnetic reconnections, possess nonlinear profile?

Second, if nonlinear, is this process deterministic?

Third, is the process of sudden release of magnetic energy on, or near the surface of the Sun sensitive to the initial conditions? Alternatively, can we make a long term prediction for this physical process?

4.Theory

4.1. Double Exponential Smoothing

The Double Exponential Smoothing, or Second Order Exponential Smoothing, was first developed by Brown (1956), later modified by Holt (1957) and Winters (1960), to effectively remove trends from a given signal which basically distorts the relationship of one's interest. The method of double exponential smoothing (Brown, 1956; Holt, 1957; Winters, 1960) is governed by the following system of equations:

$$x_1^{(p)} = x_1 \quad b_1 = x_2 - x_1 \tag{1}$$

$$x_i^{(p)} = \alpha x_i + (1-\alpha)(x_{i-1}^{(p)} + b_{i-1}) \tag{2}$$

$$\text{and} \quad b_i = \beta(x_i^{(p)} - x_{i-1}^{(p)}) + (1-\beta)b_{i-1} \tag{3}$$

(i = 2,3,4,...,N)

where $\{x_i\}_{i=1}^N$ is the observed discrete signal, $\left\{x_i^{(p)}\right\}_{i=1}^N$ is the smoothed signal, and $\{b_i\}_{i=1}^N$ is the trend traced in the signal, α and β are the 'signal smoothing parameter' and 'trend smoothing parameter', respectively. We have $0 < \alpha < 1$ and $0 < \beta < 1$. In this current work, to maintain the positional importance, the values of α and β are taken as 0.68 and 0.74, respectively. This pre-processing, using double exponential smoothing, has been performed as the present signal contains some trend with possible existence of chaos, and double exponential smoothing can effectively remove the noise, as well as the trend as perceived from the above analysis. The original solar flare index signal and its corresponding smoothed one are represented by the graphical illustration shown in Figure 1 and Figure 2.

4.2. Delay Vector Variance (DVV) Analysis

A time-series manifested by a temporal signal x_i can be represented in "phase space" by the method of time delay embedding. When a time delay is embedded in it, it can be expressed by a set of delay vectors (DVs) $x(k) =$

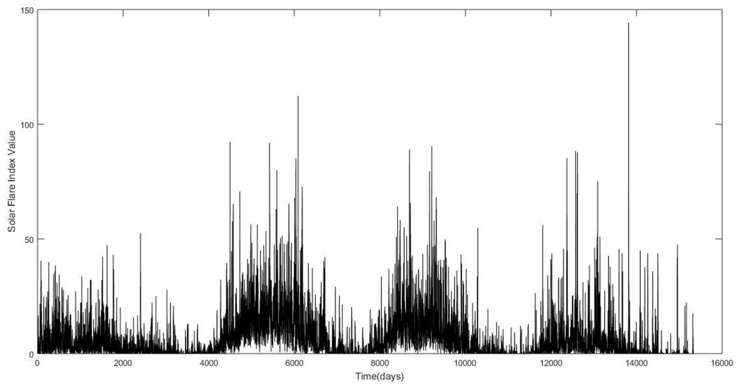

Fig. 1. Original Solar Flare Index Signal

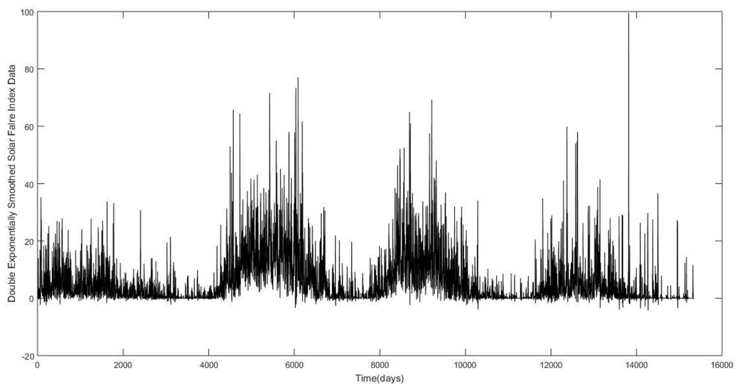

Fig. 2. Smoothed as well as de-trended Solar Flare Index Signal by Double Exponential Smoothing

$[x_{k-m\tau}, \cdots, x_{k-\tau}]$ (where $k = 1, 2, \cdots, N$), the embedding dimension is given by m and the embedded time delay lag is denoted by τ. Inside a certain Euclidian distance τ_d to DV $x(k)$, DVs are clustered which are denoted by $\lambda_k(\tau_d)$. The mean target variance σ^{*2} is calculated over all sets of $\lambda_k : k = 1, 2, \cdots, N$ to get optimal embedding dimension m. The embedding dimension which generates minimum σ^{*2} is the optimal one. The variation of the standardized distance facilitates the entire range of pair wise distances for the present examination (Gautama, Mandic and Van Hulle, 2004; Ahmed, 2010; Samadder, Ghosh and Basu, 2015; Hossain, Ghosh, Ghosh and Bhattacharjee, 2012). To standardize the distance axis, τ_d is supplanted by $(\tau_d - \mu_d)/\sigma_d$, where μ_d and σ_d are mean and standard deviation, respectively, simulated over all pair wise distances between DVs given by:

$$d(i,j) = \|x(i) - x(j)\|; \quad i \neq j \tag{4}$$

The DVV plots are generated by plotting target variance $\sigma^{*2}(\tau_d)$ vs. $(\tau_d - \mu_d)/\sigma_d$. The estimation of the noise present in the signal is given by minimum value of target variance $\sigma^{*2}_{\min} = \min_{\tau_d}\{\sigma^{*2}(\tau_d)\}$. The presence of noise is overriding in case of stochastic components. Hence, stochastic components should possess larger values of $\sigma^{*2}_{\min}$. On the other hand, smaller values of $\sigma^{*2}_{\min}$ indicate that the signal is deterministic. As all the DVs are interior to the same Universal set, and the variance of the targets is identical to the variance of the signal for maximum span, the DVV plots converge to unity at the extreme right.

Iterative Amplitude Adjusted Fourier Transform (IAAFT) (Kugiumtzis, 1999; Schreiber and Schmitz, 2000) has been used to obtain surrogate signal. The DVV plots of these surrogated signals are obtained using optimal embedding dimension of the original one. A DVV Scatter diagram can be composed by plotting target variance $\sigma^{*2}(\tau_d)$ of the original signal along horizontal axis and mean of $\sigma^{*2}(\tau_d)$ of surrogate signal along vertical axis. If the DVV plots of the surrogate and the original signal are analogous, then DVV Scatter diagram coincides with the bisector line and the signal is said to be linear. Else, if the two DVV plots are not similar, then DVV Scatter diagram deviates from the bisector line and the signal is said to be nonlinear. The nonlinearity can be understood as the root mean square error (RMSE) between the $\sigma^{*2}(\tau_d)$ of the original signal and mean of the $\sigma^{*2}(\tau_d)$ of the surrogate signal, shown below:

$$RMSE = \sqrt{\operatorname{mean}\left\{\sigma^{*2}(\tau_d) - \frac{\sum\limits_{k=1}^{N_S} \sigma^{*2}_{s,k}(\tau_d)}{N_S}\right\}^2} \tag{5}$$

where $\sigma^{*2}_{s,k}(\tau_d)$ is the target variance at span τ_d for the k-th surrogate, and the average is considered over all span of τ_d valid in all the surrogate and DVV plots (Ahmed, 2010; Samadder, Ghosh and Basu, 2015; Hossain, Ghosh, Ghosh and Bhattacharjee, 2012).

4.3. Recurrence Plot (RP) Analysis

A Recurrence Plot (RP) (Eckmann, Kamphorst and Ruelle, 1987) is a visual way to see the recurrence pattern of a dynamical system. A recurrence is defined as the return of the trajectory in its earlier state. A recurrence occurs when the system returns to the neighbourhood of an earlier point in the phase space. If a point $\overrightarrow{x_i} \in R^m$ is in a trajectory $\overrightarrow{x_1}, \overrightarrow{x_2}, \cdots, \overrightarrow{x_N}$, then the recurrence matrix $\boldsymbol{R}$ is expressed as:

$$R_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \|\overrightarrow{x_i} - \overrightarrow{x_j}\|\right); \quad i,j = 1,2,\cdots,N \tag{6}$$

where N is the number of points in the trajectory, ε is the appropriate threshold distance, $\Theta(\cdot)$ is the Heaviside function, i.e., $\Theta(a) = 0$ if $a < 0$, and $\Theta(a) = 1$ if $a \geq 0$ and $\| \cdot \|$ is a suitable norm. Hence, $\boldsymbol{R}$ is a matrix with elements either 0 or 1 only, and a Recurrence Plot is a graphical representation of $\boldsymbol{R}$, generated by staining a black dot for every 1 and white dot for every 0. Therefore, the RPs have very long diagonal lines for regular signals, very short diagonal lines for signals probably with sensitive dependence to the initial conditions and almost no diagonal line for homogeneous distribution of stochastic signals.

4.4. Recurrence Quantification Analysis (RQA)

The recurrence quantification analysis (RQA) (Webber and Zbilut, 1994; Zbilut and Webber, 1992; Atay and Altntas, 1999; Marwan, Wessel, Meyerfeldt, Schirdewan and Kurths, 2002; Marwan, Carmen Romano, Thiel and Kurths, 2007; Marwan, 2011; Hossain, Ghosh, Ghosh and Bhattacharjee, 2015; Samadder, Ghosh and Basu, 2015) is an efficient quantitative approach to study nonlinear data. It quantifies the number and duration of recurrences of a time series data occurring in its recurrence plot.

We estimate four recurrence variables. The first one is %REC which quantifies the percentage of recurrent points existing within predefined threshold. It estimates the probability that a specific state will recur by calculating black dots in the recurrence plot. %REC ranges between 0% to 100%. 0% indicates no recurrent point and 100% indicates all recurrent points in RP. More the %REC, more the chance of time series being regular. REC is expressed as

$$REC(\varepsilon_i) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon_i) \tag{7}$$

The second recurrence variable is %DET, or predictability, which quantifies the ratio of recurrent points forming diagonal lines to all recurrent points. The value of %DET is between 0% to 100%. 0% presents stochastic time series and 100% indicates deterministic time series. Any value between them indicates the possibility of chaotic time series. DET is given by

$$DET = \frac{\sum\limits_{i,j=1}^{N} D_{i,j}}{\sum\limits_{i,j=1}^{N} R_{i,j}} \tag{8}$$

where

$D_{i,j}$ =1,if (i, j), $(i + 1, j + 1)$ and $(i - 1, j - 1)$ are recurrent, and 0, otherwise.

The third recurrence variable is %LMAX, *i.e.*, linemax, which provides length of the longest diagonal line segment in the recurrence plot, except the main diagonal line. It is inversely proportional with Lyapunov exponent (Eckmann *et al.*, 1987), *i.e.*, shorter LMAX implies chaotic time series and longer LMAX implies non-chaotic time series.

If N_l is the number of diagonal lines and l_i is the length of i-th diagonal line, then

$$LMAX = \max(l_i) \text{ where } i = 1, 2, \cdots, N_l \tag{9}$$

The fourth recurrence variable is ENTR, *i.e.*, Shannon information entropy of all diagonal line lengths distributed over integer bins in a histogram. It estimates the probability to find a diagonal line of length l in RP. ENTR is given by

$$ENTR = -\sum_{i=1}^{N_l} p(l) \ln p(l) \tag{10}$$

where $p(l)$ is the probability distribution of lengths of diagonal lines. Entropy quantifies the complexity of the deterministic structure in the system. High entropy indicates non-chaotic nature and low entropy indicates chaotic nature of the data.

5. Results

In this work, our study is focused on the signal of solar flare index obtained from McMath-Hulbert Solar Observatory (NOAA, 2019) ranging from Jan 01, 1966 to Dec 31, 2008. First, the data is smoothed using double exponential smoothing and then DVV analysis is performed on the smoothed data to check if the data is nonlinear and if yes, if the data is deterministic in profiles or not. To have a proper analysis, it is essential to determine the proper embedding lag and embedding dimension separately (Kodba *et al*, 2005). Here the obtained embedding dimension (m) is 6 and embedding lag (τ) is 10. Minimum value of target variance $\sigma^{*2}_{\min}$ 0.2295, indicating possibly deterministic trend of the data, and corresponding root mean square error (RMSE) is 0.0323, clarifying nonlinearity of the data.

Figure 3 and Figure 4 depict the DVV analysis of the present smoothed Solar Flare Index signal.

From Figure 3, it can be observed that the DVV lines for the original and surrogate are somehow dissimilar. Figure 4 confirms that the scatter diagram deviates from the bisector line, and hence the present signal possesses some form of nonlinearity.

Smoothed Solar Flare Index data are normalized before producing RP plot to get a clear scenario of recurrence plot. Threshold value is taken as $\varepsilon = 0.1$, which is around 1% of maximum distance between two data points to get the clearer scenario of recurrence plot.

Figure 5 demonstrates the profile of RP analysis for the present signal.

RQA analysis is done to estimate complexity of time series in a quantified way. Table 1 summarizes the RQA variables.

The RPs of Solar Flare Index signal indicates the possibility of sensitive dependence on the initial conditions as some very short diagonal lines are visible in Figure 4, which in turn indicates a possible profile of sensitivity to the initial conditions. Table 1 explains the fact that the time series is chaotic. Low

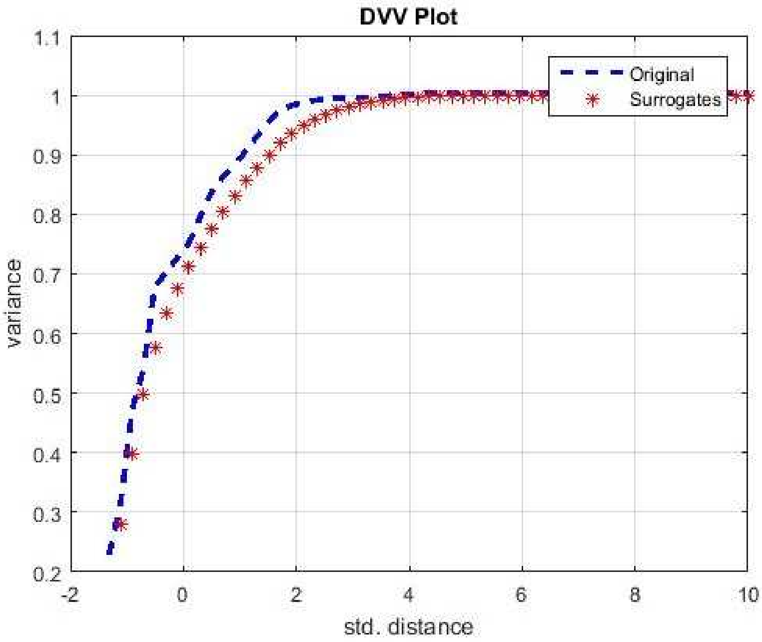

Fig. 3. DVV Plot of smoothed Solar Flare Index signal

Table 1. RQA Analysis for Solar Flare Index signal

REC	0.20
DET	0.68
LMAX	0.99
ENTR	3.86

value of %REC (20%) suggests that no reasonable amount of recurrence arises, which indicates the possibility that the solar flare index signal considered in our study is non-regular. The obtained value of %DET is 68%, which clearly shows that the time series is chaotic. Again, low values of LMAX(0.99) and ENTR(3.86) strengthen the possibility of chaotic nature of the Solar Flare Index signal.

6. Conclusion

In the present work, DVV analysis confirms the presence of nonlinearity in the Solar Flare Index signal, and the subsequent RP analysis with RQA confirms the existence of chaos in it. Hence, for this signal long term prediction cannot yield trust-worthy results; so one has to confine within short term predictions. Finally, the present study indicates that the mechanics of reorganization of magnetic loops, resulting rapid conversion of a large amount of

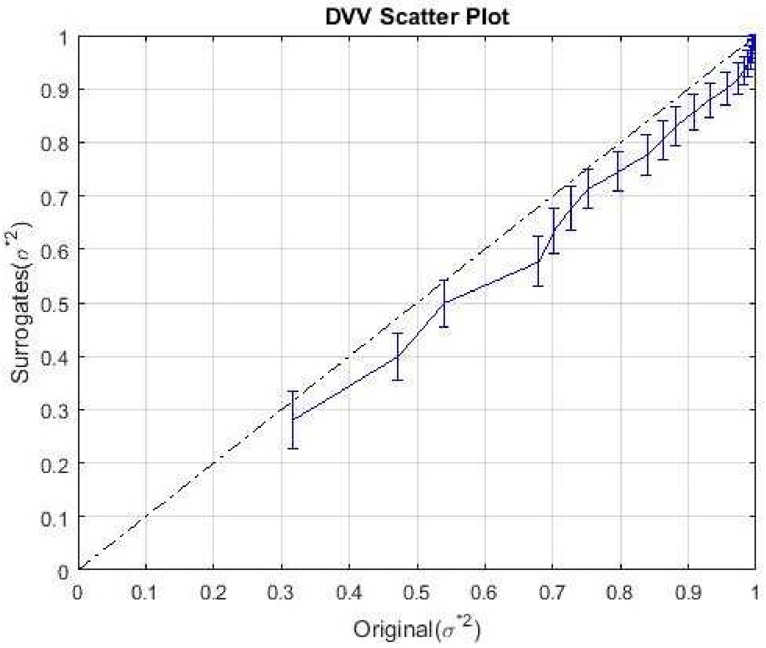

Fig. 4. DVV Scatter Plot of smoothed Solar Flare Index signal

magnetic energy previously stored in the solar corona and dissipated through magnetic reconnections, is a nonlinear and chaotic solar phenomenon. Earlier, by using *0–1* Test Gottwald and Melbourne, (2009); Mukherjee *et al*, (2017) showed the possible presence of chaos in the present signal. The present study re-establishes this fact with a more certainty. The study of solar flares is essential to understand the space weather in a better way as it directly impacts the ionosphere and radio communications in the vicinity of Earth. In other words, the present study not only puts a light over solar internal dynamics to a certain extent, but also helps us to realize the local space weather. The present study is very interesting and important, as it covers the range from the beginning of 1966 to the end of 2008, capturing almost four complete solar cycles (*viz.* solar cycles 20, 21, 22 and 23). The limitation of this work is that it has not take into account the current solar cycle 24 due to the unavailability of the data for the entire cycle. A future study, focused on solar cycle 24 should be more interesting as it will enable us to comprehend whether this cycle repeats the characteristics as prevailed in the old cycles, or it will show some novel features.

References

Ahmed, I. (2010) : Detection of nonlinearity and stochastic nature in time series by delay vector variance method, *International Journal of Engineering & Technology*, **10**(2), 22–27.

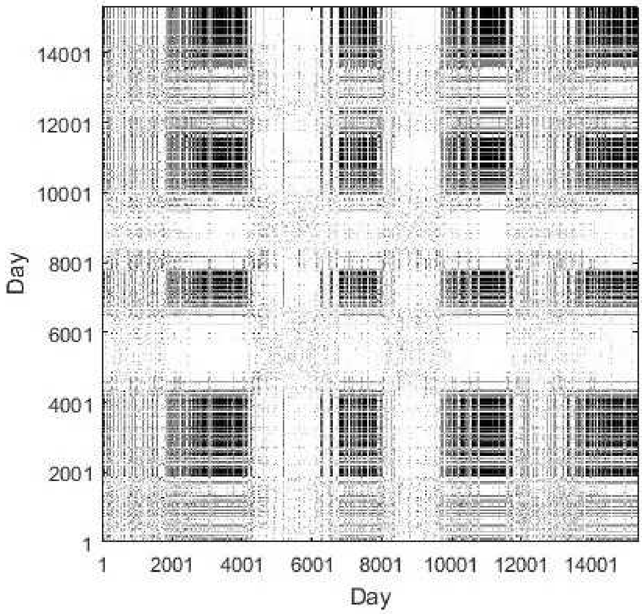

Fig. 5. Recurrence Plot Analysis of the present smoothed Solar Flare Index signal

Atay, F.M. and **Altýntas, Y.** (1999) : Recovering smooth dynamics from time series with the aid of recurrence plots, *Phys. Rev.* E, **59**(6), 6593–6598.

Brown, R.G. (1956), Exponential Smoothing for Predicting Demand, Cambridge, Massachusetts: Arthur D. Little Inc., 15.

Eckmann, J.P., Kamphorst, S.O. and **Ruelle, D.** (1987) : Recurrence plots of dynamical systems, *Europhys. Lett.*, **4**(9), 973–977.

Gautama, T., Mandic, D.P. and **Van Hulle, M.M.** (2004) : The delay vector variance method for detecting determinism and nonlinearity in time series. *Physica* D, **190**(3–4), 167–176.

Gottwald, G.A. and **Melbourne, I.** (2009) : "On the Implementation of the 0–1 Test for Chaos", *SIAM J. Appl. Dyn. Syst.*, **8**(1), 129–145.

Holt, C.C. (1957) : Forecasting Trends and Seasonal by Exponentially Weighted Averages, *Office of Naval Research Memorandum*, **52**, Carnegie Institute of Technology, Pittsburgh.

Hossain K.M., Ghosh D.N., Ghosh K. and **Bhattacharjee A.K.** (2012) : Nonlinearity and chaos in ^{8}B solar neutrino flux signals from Sudbury neutrino observatory, *Fractals*, **20**(1), 17–32.

Hossain, K.M., Ghosh, D.N., Ghosh, K. and **Bhattacharjee, A.K.** (2015) : Complexity in solar irradiance from the Earth radiation budget satellite, *IEEE System Journals*, **9**(2), 487–494.

Kleczek, J. (1952) : Publication of Astrophysical Observatory, No. **24**, Prague.

Kodba, S., Perc, M. and **Marhl, M.** (2005) : "Detecting chaos from a time seriess", *Eur. J. Phys.*, **26**, 205-215.

Kugiumtzis, D. (1999) : Test your surrogate data before you test for nonlinearity, *Physical Reiew* E, **60**(3), 2808–2816.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A. and **Kurths, J.** (2002) : Recurrence plot based measures of complexity and its application to heart rate variability data, *Phys. Rev.* E, **66**(2), 026702.1–026702.8.

Marwan, N., Carmen Romano, M., Thiel, M. and **Kurths, J.** (2007) : Recurrence plots for the analysis of complex systems, *Phys. Rep.*, **438**(5), 237–329.

Marwan, N. (2011) : How to avoid potential pitfalls in recurrence plot based data analysis, *Int. J. Bifurcation Chaos*, **21**(4), 1003–1017.

Mukherjee, S., Rakshit, K. and **Ghosh, K.** (2017) : *Bull. Cal. Math. Soc.*, **109**(3), 229–236.

NOAA(2019): ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FLARES/FLARES_/INDEX/DAILY.PLT.

] **Samadder, S., Ghosh, K.** and **Basu, T.** (2015) :, Investigation Of Nonlinearity And Chaos In Prime Indian And American Stock Exchange Indices, *Hyperion International Journal of Econophysics & New Economy*, **8**(1), 65–82.

Schreiber, T. and **Schmitz, A.** (2000) : Surrogate time series, *Physica* D, **142**(3), 346–382.

Webber, C.L. Jr. and **Zbilut, J.P.** (1994) : Dynamical assessment of physiological systems and states using recurrence plot strategies, *J. Appl. Physiol.*, **76**(2), 965–973.

Winters, P.R. (1960) : Forecasting Sales by Exponentially Weighted Moving Average, *Management Science*, **6**(3), 324–342.

Zbilut, J.P. and **Webber, C.L. Jr.** (1992) : Embeddings and delays as derived from quantification of recurrence plots, *Phys. Lett.* A, **171**(3), 199–203.