Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 55 (1): 27–33 (2018)

ISSN: 2518-4245 (print), 2518-4253 (online)

Pakistan Academy of Sciences

Research Article

Study of Coronal Index Time Series Solar Activity Data in the Perspective of Probability Distribution

Muhammad Fahim Akhter*, Shaheen Abbas, and Danish Hassan

Mathematical Sciences Research Centre, Federal Urdu University of Arts, Sciences and Technology, Karachi, Pakistan

Abstract: The different cycles of solar activity defines space weather variation and future space variability. Such as Solar Corona produces energy in micro and nano flares into the space climate. In this regards the time series analysis of monthly solar Coronal index data (1944 to 2008) is used, which contains six cycles of different length and peaks. In this study different probability distributions like Johnson SB, Beta, Gen. Pareto, Gen. Gamma, Triangular, Error, Dagum and Fatigue Life are fitted on Coronal cycles. The significance probability distributions are obtained using Kolmogrov- Smirnove, Anderson Darling and Chi square statistical tests. The Johnson SB distribution is found best fitted on all solar activity cycles along with the total time series data by Kolmogrov Smirnove test. The Coronal index monthly data is generated from 2008 to 2016 using a Monte Carlo simulation technique. While two other tests show variation in the fitted probability distributions for all cycles. With the help of significant probability distribution the expected length and peak of next Coronal index cycle data can be obtained.

Keywords: Solar activity, Coronal index cycle, probability distribution, statistical tests

1. INTRODUCTION

The sun produces amounts of solar energy that is caused by the different solar activities and terrestrial environment [1]. The lowest part of the sun's atmosphere is the photosphere gases that can be recognized by the disk of the sun. The Sun's atmosphere typically refers to all the regions above the photosphere [2]. The solar atmosphere separates into three layers, the photosphere, chromospheres, and corona on the basis of temperature, density, and composition [3]. The most above a layer of the sun's atmosphere is Corona. The sun's corona contains many spectrums of forbidden transition lines like, Green line Corona, Red line Corona, and yellow line Corona. The brightest emission corona is the Green line Corona that is FeXIV iron line of 530.03 nanometers [6]. The Corona is associated with sunspot activity and Coronal mass ejection occurs in close region of magnetic field that overlies the inversion line of magnetic field [10].

The first spectrum of the green line 530.3 nm is found by two astronomers, Harkness and Young in 1869 during full solar eclipse [8,9]. The beautiful Corona is shown in Fig.1 during full solar eclipse.

The Corona has a number of components K, E, F & T Corona. In E-Corona (Emission) it is composed of line emission of visible to EUV due to various atoms and ions in the Corona. In many forbidden transition lines, the strong lines are FeXIV 530.3 nm. The green line emissivity is due to forbidden transition of the FeXIV ion, peaks in a temperature of about 2x106 [6, 7 and 3]. This paper discusses the study of the green line coronal cycles in the perspective of adequate probability distributions.

2. MATERIAL AND METHODS

The study is based on the probability distributions of time series Coronal index cyclic data. For this

Received, June 2017; Accepted, March 2018
*Email: Muhammad Fahim Akhter: m_fahimakhter@yahoo.com

purpose, we observed 64 years monthly data (six cycles). Using Monte Carlo Simulation hundred sample points generated with the help of significant Probability distribution (Johnson SB). In this connection the smallest and longest cycle are found 10 years and 12 years long respectively. The fitted probability distributions of coronal cycles are evaluated by the following steps.

2.1 Step1: Fitting The Probability Distribution

The numerous probability distributions are applied on Coronal index Cycles. Different probability distributions Johnson SB, Beta, Gen. Pareto, Gen. Gamma, Dagum, Triangular, Eror and Fatigue [5] are examined and found best fitted probability distribution under three tests listed in Table 1.

Table 1. Proposed Probability Distributions along with Parameters (under three tests).

Probability Distribution Function	Parameters
Johnson SB $f(\mathrm{x}) = \frac{\delta}{\lambda\ \sqrt{2\pi}\ \ \mathrm{z}(1-\mathrm{z})} \exp\left(-\frac{1}{2}\left(\gamma + \delta \ln\left(\frac{\mathrm{z}}{1-\mathrm{z}}\right)\right)^2\right)$ $\zeta \le x \le +\lambda$	γ- continuous shape parameter δ- continuous shape parameter (δ>0) λ- continuous scale parameter (λ>0) ζ- continuous location parameter
Beta $f(\mathrm{x}) = \frac{1}{\beta(\alpha_1, \alpha_2)} \frac{(\mathrm{x}-\mathrm{a})^{\alpha_1 - 1} (\mathrm{b}-\mathrm{x})^{\alpha_2 - 1}}{(\mathrm{b}-\mathrm{a})^{\alpha_1+\alpha_2-1}}$ $a \le x \le b$	α_1 -Continuous shape parameter (α_1) α_2 - Continuous shape parameter (α_2) a, b - continuous boundary parameters (a<b)
Gen. Pareto $f(\mathrm{x}) = \begin{cases} \frac{1}{\sigma}\left(1 + \mathrm{k}\frac{(\mathrm{x}-\mu)}{\sigma}\right)^{-1-\frac{1}{\mathrm{k}}} & \text{for k} \ne 0 \\ 1 - \exp\left(-\frac{(\mathrm{x}-\mu)}{\sigma}\right) & \mathrm{K} = 0 \end{cases}$ $\mu \le x < +\infty, for\ K \ge 0\ \ \mu \le x \le \mu - \sigma/K\ \ for\ k < 0$	*K*- continuous shape parameter σ- continuous scale parameter (σ>0) μ- continuous location parameter
Gen. Gamma $f(\mathrm{x}) = \frac{\mathrm{k}(\mathrm{x}-\gamma)^{\mathrm{k}\alpha-1}}{\beta^{\mathrm{k}\alpha}\Gamma(\alpha)}$ $\gamma \le x < +\infty$	*k*- continuous shape parameter (*K*>0) α- continuous shape parameter (α>0) β- continuous scale parameter (β>0), γ-continuous location parameter (γ≡0)
Levy $f(\mathrm{x}) = \sqrt{\frac{\sigma}{2\pi}} \frac{\exp(-0.5\sigma/(\mathrm{x}-\gamma))}{(\mathrm{x}-\gamma)^{3/2}}$ $\gamma < x < +\infty$	σ- continuous scale parameter (σ>) γ- continuous location parameter (γ≡0)
Fatigue $f(x) = \frac{\sqrt{x/\beta} + \sqrt{\beta/x}}{2\alpha x} \emptyset \left(\frac{1}{\alpha}\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right)$ $\gamma < x < +\infty$	α- continuous shape parameter (α>0) β- continuous scale parameter (β>0) γ- continuous location parameter (γ≡0)
Error Function $\mathrm{f(x)} = \frac{\mathrm{h}}{\sqrt{\pi}} \exp(-(\mathrm{hx})^2)\ ,\ -\infty < x < +\infty$	h- continuous inverse scale parameter (h>0)
Dagum $f(\mathrm{x}) = \frac{\alpha\mathrm{k}\left(\frac{\mathrm{x}}{\beta}\right)^{\alpha\mathrm{k}-1}}{\beta\left(1+\left(\frac{\mathrm{x}}{\beta}\right)^{\alpha}\right)^{\mathrm{k}+1}}$ $\gamma \le x < +\infty$	k- continuous shape parameter (k>0) α- continuous shape parameter (α>0) β- continuous scale parameter (β>0) γ- continuous location parameter (γ≡0)
Triangular $f(\mathrm{x}) = \begin{cases} \frac{2(\mathrm{x}-\mathrm{a})}{(\mathrm{m}-\mathrm{a})(\mathrm{b}-\mathrm{a})} & \mathrm{a} \le \mathrm{x} \le \mathrm{m} \\ \frac{2(\mathrm{b}-\mathrm{x})}{(\mathrm{b}-\mathrm{m})(\mathrm{b}-\mathrm{a})} & \mathrm{m} < x \le b \end{cases}$ $a \le x \le b$	m (most likely value) - continuous mode parameter (a ≤ m ≤b) a, b- continuous boundary parameters (a <b)

2.2 Step 2: Goodness-of-Fit Tests

We adopted the probability distribution by using the three goodness-of-fit tests for the following null hypothesis:

H_0: the Coronal index cycle follows the significant distribution.

H_A: the Coronal index cycle does not follow the significant distribution.

2.2.1 Kolmogrov Smirnove Test

The Kolmogrov Smirnove statistic (D) is the largest vertical difference between the theoretical and the empirical cumulative distribution function (ECDF)

$$D = \max_{1 \le i \le n} \left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i) \right) \quad (1)$$

X_i = random sample. $i = 1,2,3 \ldots\ldots\ldots\ldots, n$.

$$CDF = F_n(x) = \frac{1}{n}[Number\ of\ observation \le x] \quad (2)$$

2.2.2 Anderson Darling Test

It is a test to compare to fit an observed cumulative distribution function to an expected cumulative distribution function. Generally, this test weight on tail than Kolmogrov test and denoted by A2 [4].

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left[lnF(X_i) + ln\left(1 - F(X_{n-i+1})\right) \right] \quad (3)$$

2.2.3 Chi-Squared Test

This test is applied to binned data, so value of test statistic χ^2 depends on how the value of data is binned. The number of bins is calculated by this formula.

$$k = 1 + \log_2 N \quad (4)$$

K is a number of bins and N is a sample size.

$$x^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

Where,

O_i = Observed frequency for bin i.

E_i = Expected frequency for bin i.

'*i*'= number of observations. (i=1, 2,………. K)

E_i is calculated by

$$E_i = F(x2) - F(x1) \quad (6)$$

F= The (CDF) of the probability distribution being tested, and x_1, x_2 are limits for bin data i [4].

2.3 Step 3: Significant Fit Of Probability Distribution

The three statistical tests are applied to segregate probability models for Coronal index cyclic data as the goodness-of-fit tests. Therefore, the smallest statistic value in all distributions with respect to each test was selected separately. The test statistic, tested critical value under significance (α=0.01 & 0.05), and checked no rejection of null hypothesis represents stimulated energy emission of Coronal cycle follows the particular distribution and found best fitted.

2.4 Monte Carlo Simulation by Best Fit Distribution

Monte Carlo is a computational technique. This technique based on constructing a random process. The random process carrying out a NUMERICAL EXPERIMENT by N-fold sampling from a random sequence of numbers with a PRESCRIBED probability distribution [11]. In This study data are simulated by the Johnson SB probability distribution parameter. The Coronal cyclic data are generated from 2008 to 2016. The simulated monthly data along with six cycles is furnished in Fig. 2.

3. RESULTS AND OUTLOOK

The study is utilized on Coronal index data (1944 to

Fig. 1. The beautiful Corona due to magnetic field and full solar eclipse.

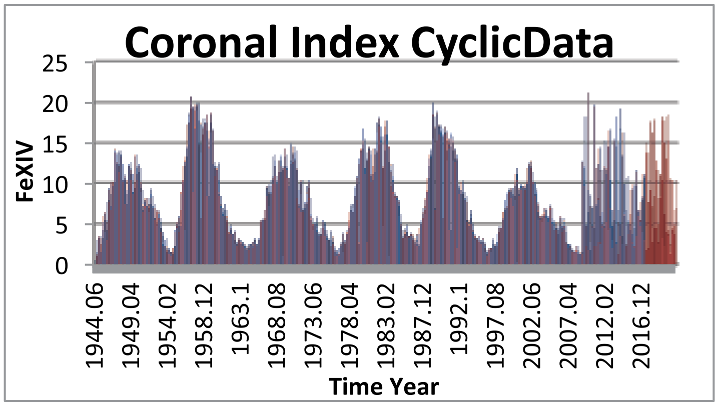

Fig. 2. Monte Carlo simulation from 2008 to 2016.

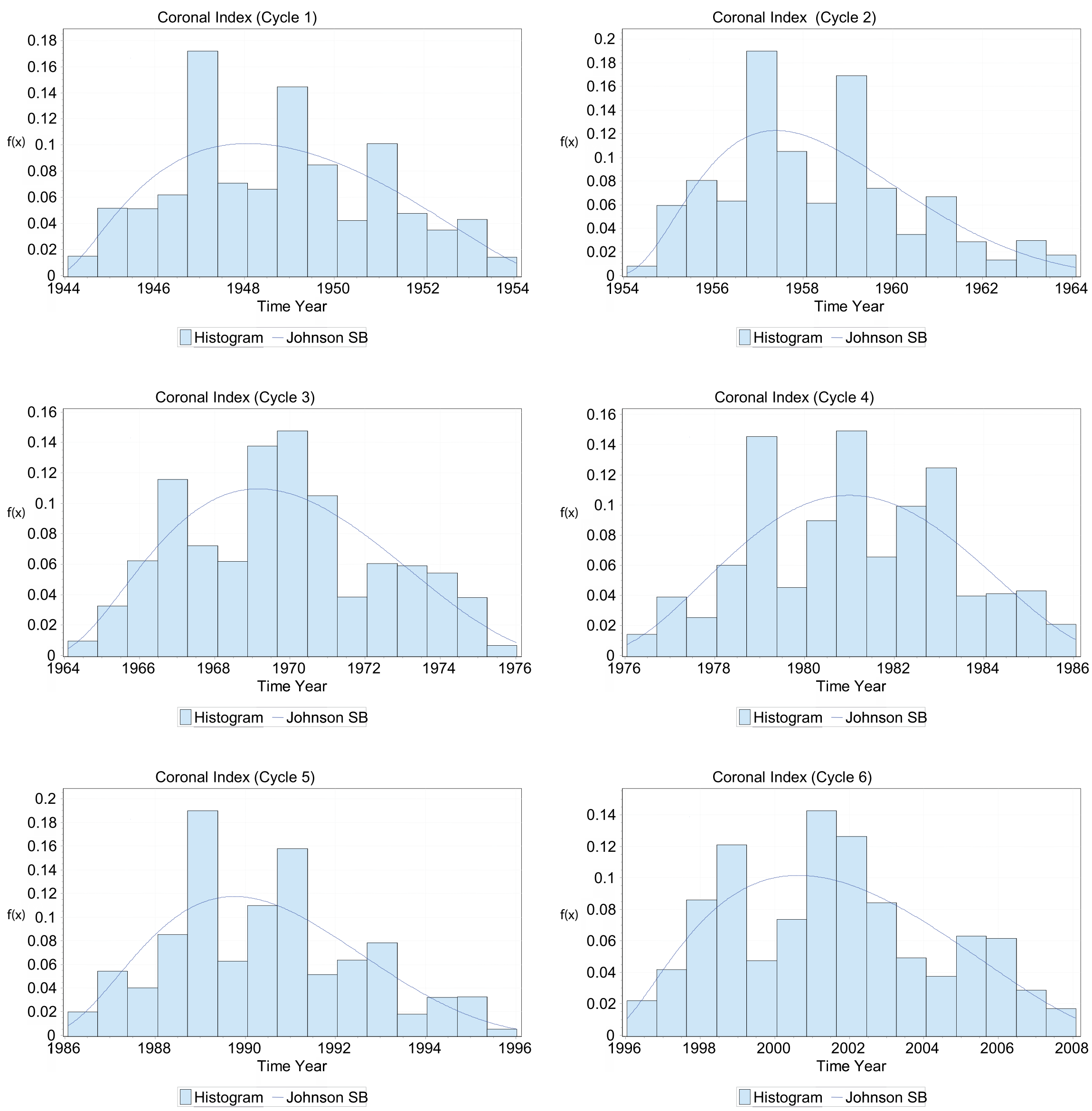

Fig. 3. Johnson SB probability distribution of all Coronal cycles.

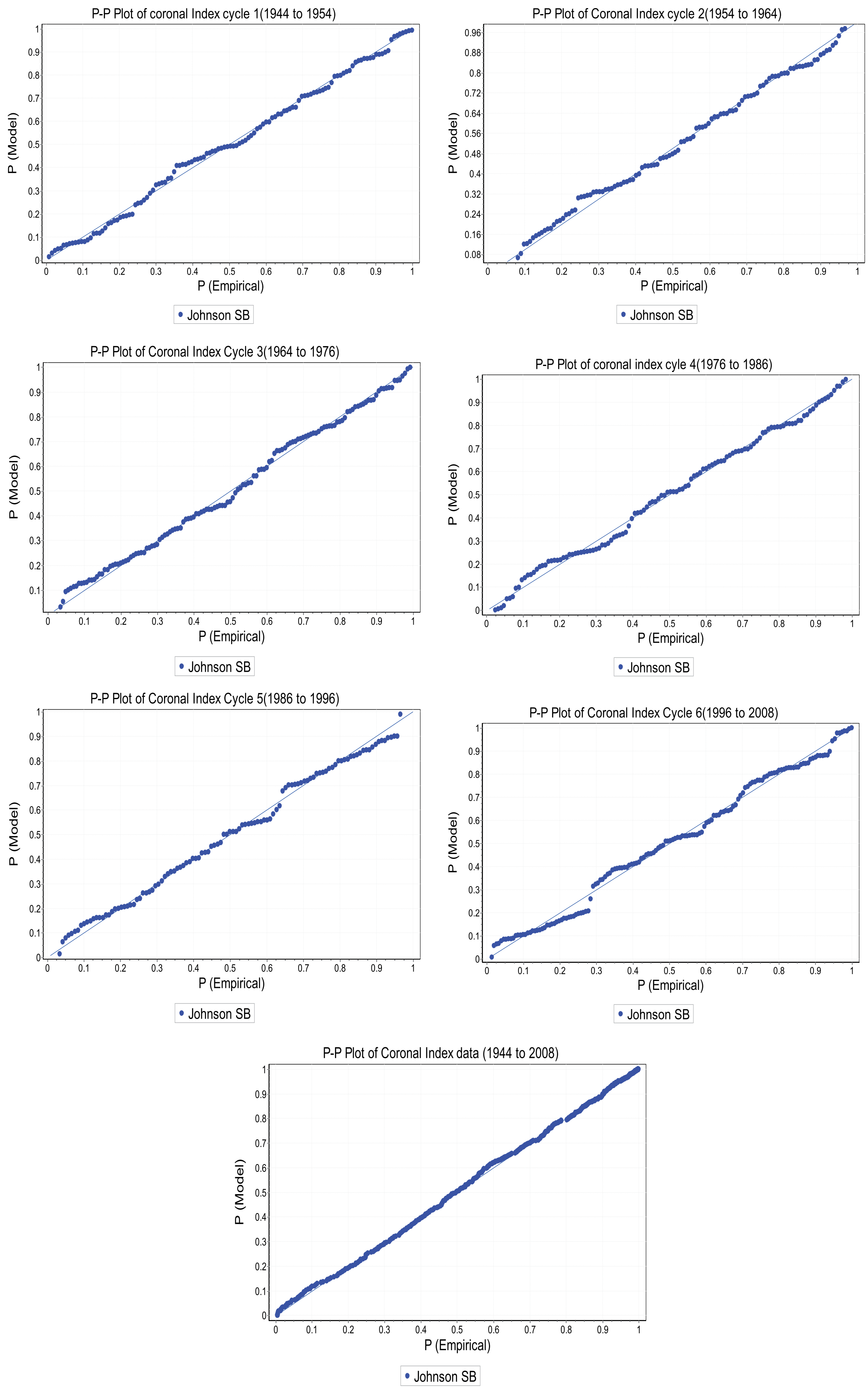

Fig. 4. P-P plot of coronal index cycles.

Table 2. Goodness-of-Fit Tests

Cycle	Kolmogrov Smirnove		Anderson Darling		Chi-Squared	
	Distribution	Statistics	Distribution	Statistics	Distribution	Statistics
1	Johnson SB	0.05737	Gen. Gamma (4p)	0.2958	Johnson SB	1.9379
2	Johnson SB	0.06597	Beta	1.9161	Levy (2p)	7.1667
3	Johnson SB	0.05023	Beta	1.1448	Fatigue Life	7.9527
4	Johnson SB	0.04989	Gen. Pareto	0.82832	Dagum	3.9108
5	Johnson SB	0.05658	Gen. Pareto	1.6817	Error	5.173
6	Gen. Pareto Johnson SB	0.06348 0.07012	Beta	0.97549	Gen. Gama (4p)	8.7881
1to 6	Johnson SB	0.02141	Triangular	1.0803	Triangular	14.164

2008) along with six cycles, including the minimum to maximum ranged from 0.02 to 20.79 nanometers. There is right skewed observed in Coronal index cycles, that is why mean of each cycle is greater than the median. Further negative excess Kurtosis of each cycle along with total data observed less peaked and has less extreme frequent value (less fat tails) than a normal distribution. Moreover the Coronal cycles have amount of variation relative to mean is around 50% to 66%. The list of fitted probability distribution under three tests is depicted in Table 1. The Johnson SB is observed best fitted distribution for all six Coronal cycles along with the total data. With the help of Kolmogrov test the Johnson SB shows minimum statistic value among numerous distributions. The variations for other distributions are observed in all six Corona index cycles under minimum statistic value regarding other two tests. All six cycle depicts in Fig. 2 along with the p-p plot and Johnson SB probability distribution is depicted in Fig. 3.

The long right tail of the Johnson probability distribution on each cycle indicates the heavy tail dynamics of Coronal time series data in the cycle (depicts in Fig. 3). All the solar activity, including Coronal index activity defines the space weather. Earth's climate change strongly depends on the space weather variability. The research work discusses the one of the space weather parameters that is Coronal index. All the Coronal index cycles show the same dynamical behavior in the perspective of Johnson SB probability distribution using Kolmogrove-Smirnove test. By the simulation of the significant distribution (Johnson SB) long term dynamics of Corona cycle can be observed. To understand the space weather and its forecast Coronal index simulated time series data (2008-2016) may be beneficial for space scientist.

Fig. 3 also proves the adequacy of Johnson SB probability distribution. The p-p plot of each Coronal index cycle represents the best fitted distribution model (Fig. 4).

While the fluctuation in Coronal cycle against Anderson Darling test gives more weight on tail than Kolmogrove Smirnove test. According to this test critical statistic value significant at alpha (0.01 and 0.05). Moreover Beta, Gen. Pareto, Gen. Gama (4p) and triangular distributions are best fitted. According to Chi-Square test, data grouped into intervals of equal probability or equal width due to this connection the best probability distribution settled on Coronal cyclic data are Johnson SB, Levy (2p), Fatigue Life, Dagum (4p), Gen. Gama (4p), Eror and Triangular Distribution. There is a null hypothesis acceptance level of significance at (0.2, 0.1, 0.05, 0.02, 0.01) except triangular, depicts the Table 2. Furthermore, 100 sample points are simulated by the best fitted Johnson SB distribution under the parameters of (γ=0. 84016 δ=0. 73135 λ=21. 184 ξ=1. 2422).

4. CONCLUSION

The stimulated energy emission of Green line

Coronal cycle performed under three goodness-of-fit tests. On the basis of Kolomogrov Smirnov test, Johnson SB probability distribution is identified in all six Coronal cycle data pattern. While Anderson Darling test repeated the Gen. Pareto and Beta because of the Coronal tail behavior. The Chi-square test shows the Johnson SB, Levy, Fatigue Life, Dagum, Error and Gen. Gama are best fitted probability distribution from cycle 1 to cycle 6 respectively. Using the Monte Carlo simulation technique the samples from probability distribution Johnson SB generated. The study may be helpful to other space variability parameters. The results obtained in this communication may also be useful to understand the space weather fluctuations in the future. The manuscript comprises the study of space weather by using Coronal index. The study can be extended by including the global climatic parameters to understand the effects of space weather on earth's climate.

5. ACKNOWLEDGEMENT

We are thankful to the NOAA for providing Coronal index monthly time series data. The results of this manuscript are the part of the first author's doctoral dissertation. The authors are also thankful to the anonymous referees and editor for their useful comments.

6. REFERENCES

1. Hassan, D., Abbas, S., Ansari, M. R. K., & Jan, B. Solar flares data analysis on application of probability distributions and fractal dimensions and a comparative analysis of North-south Hemispheric solar dlares data behavior. *Proceedings of the Pakistan Academy of Sciences* 51 (4): 345-353 (2014).
2. Babu, Arun. Coronal mass ejections from the sun-propagation and near earth fffects. *arXiv preprint arXiv:1407.4258* (2014).
3. Mariska, J. T. *The Solar Transition Region* (Vol. 23). Cambridge University Press (1992).
4. Sharma, M. A., & Singh, J. B. Use of probability distribution in rainfall analysis. *New York Science Journal 3* (9): 40-49 (2010).
5. Walk, C. *Handbook on Statistical Distributions for Experimentalists.* Internal Report SUF-PFY/96-01., University Stockholm (2007).
6. Wang, Y. M., Sheeley Jr, N. R., Hawley, S. H., Kraemer, J. R., Brueckner, G. E., Howard, R. A.,... & Schwenn, R. The green line corona and its to the photospheric magnetic z field. *The Astrophysical Journal* 485 (1): 419 (1997).
7. Podladchikova, O., Lefebvre, B., Krasnoselskikh, V., & Podladchikov, V. Classification of probability densities on the basis of Pearson's curves with application to coronal heating simulations. *Nonlinear Processes in Geophysics 10* (4/5): 323-333 (2003).
8. Minarovjech, M., Rušin, V., & Saniga, M. The green corona database and the coronal index of solar activity. *Contrib. Astron. Obs. Skalnaté Pleso 41:* 137-141 (2011).
9. Schwenn, R., Inhester, B., Plunkett, S. P., Epple, A., Podlipnik, B., Bedford, D. K.,... & Lamy, P. L. First view of the extended green-line emission corona at solar activity minimum using the LASCO-C1 coronagraph on SOHO. In: *The First Results from SOHO* (p. 667-684). Springer, Netherlands (1997).
10. Hundhausen, A.J. Sizes and locations of coronal mass ejections: SMM observations from 1980 and 1984□1989. *Journal of Geophysical Research: Space Physics* 98. A8: 13177-13200 (1993).
11. Vujic, J.L. *Monte Carlo Sampling Methods.* Nuclear Engineering Department, University of California, USA (2008).