Stochastic Modeling of Solar Flare Duration at Pakistan Atmospheric Region

Saifuddin Ahmed Jilani[1,*], M. Ayub Khan Yousuf Zai[2] and Afaq Ahmed Siddiqui[3]

[1]Department of Physics and Institute of Space & Planetary Astrophysics (ISPA), University of Karachi, Karachi-75270, Pakistan

[2]Department of Applied Physics and Institute of Space & Planetary Astrophysics (ISPA), University Karachi, Karachi-75270, Pakistan

[3]Department of Pharm. Chemistry, Faculty of Pharmacy, University of Karachi, Karachi, Pakistan

Abstract: Energy incident from the Sun is the chief deriving force responsible for all physical process existing in our terrestrial system. It is interesting to note that solar ultraviolet (UV) radiation created ozone in our stratosphere by the dissociation of O_2 molecules. On the other hand, the streams of solar particle flux deplete ozone by creating NO_x in our atmosphere. It is therefore, an important task to quantify the contribution of solar activity on OLD with the scientific assurance. In this communication the stochastic models of solar flare duration as solar activity have been investigated. Digisonde, at SUPARCO, HQ one of the ground based device provide us the record of solar flare duration by investigating the ionosphere disturbance. The behavior of solar activity have accomplished by the stochastic modeling in addition to their residual analysis. Since there are two major kinds of flares, it is necessary to establish what the different parametric configurations that causes their difference and their behavior in solar terrestrial relationship. Evidences suggest that gradual flares may become serious threat for our atmospheric and terrestrial disturbances. Their frequency most closely related with high activity periods. However sometimes this could be accomplished in low activity period as well. Hence, it is quite relevant to study theoretical and observational aspects of both high and low activity periods. The data recorded from March 1979 to March 2006 was consisting of mixed flares.

Keywords: Mixed Flares, ARIMA model, Solar Flare Duration (SFD), Solar Activity.

INTRODUCTION

The most powerful in an active region is a Solar Flare. The first flare ever detected was discovered by Carrington on 1 September 1859. The originally closed magnetic field in an active region, in which a filament (prominence) is embedded, suddenly opens. Reasons for it can be a newly emerging magnetic flux, a confined flare nearby, a wave disturbance coming along the solar surface from another source of activity, or some internal instability. As field lines open plasma begins to flow from the dense chromospheres upward to the corona so that gas pressure decrease and magnetic pressure begins to prevail. That leads to sequential reconnections of the open field lines. The reconnection process produce intense hating and it also accelerate particles. Flares are excellent indicators of coronal storms and indicate the strongest, fastest and most energetic disturbances coming from sun. Solar flares are one of two general types, Gradual and Impulsive flares. Gradual flares are large, occur high in the corona, have long duration soft and hard X-rays associated with coronal mass ejection (CME). Impulsive flares are more compact, occur lower in the corona, and produce short-duration radiation [1]. In 2002 NASA launched the Ramaty High Energy Solar Spectroscopic Imager (RHESSI), which has now captured views of certain solar flares. In doing so, RHESSI confirming that magnetic reconnection is responsible for both flares and coronal mass ejections [2]. These two most prominent classes of transient energy release from the sun have comparable maximum energies [3]. Roughly 40 % of coronal mass ejections are accompanied by solar flares that occur at about the same time and place [4].

*Address correspondence to this author at the Department of Physics and Institute of Space & Planetary Astrophysics (ISPA), University of Karachi, Karachi-75270, Pakistan; E-mail: sajilani@uok.edu.pk

METHODOLOGY

Stochastic Approach

A Stochastic process is a system expressing a phenomenon or experiment developed in some time with random variables. Most of the time series are stochastic in that the future is only partly determined by past values, so that exact predictions are impossible and must be replaced by the idea that future values have a probability distribution which is conditioned by knowledge of past values. Such a process is also probabilistic. Modeling these processes can be done using autoregressive (AR), autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) techniques. Among all of the above models of various orders the effort is being for selection

ISSN: 1814-8085 / E-ISSN: 1927-5129/15

of a model which is most adequate for the empirical study of solar flare activity. This quantitative study also includes forecasting of the solar flare duration using the selected models [5-7].

RESULTS AND DISCUSSION

Examine Stationary Condition

The first thing to note is that most of time series are non-stationary, and the AR and MA aspects of an ARIMA model refers only to a stationary time series. A time series is said to be stationary if there is no systematic change in mean (no trend) i.e. found for the data set of SFD but the variance is not uniform. The Table **1** shows non-stationary condition in variance, hence weak stationary condition may exist in the data series of SFD [8, 9].

Table 1: Comparison for Five Different Intervals of Mean and Variance; N = Total Counts = 65

Duration	Mean (SFD)	Variance (SFD)
Mar 79 – Jul 84	72.11	1093.05
Aug 84 – Dec 89	67.23	228.73
Jan 90 – May 95	37.72	375.74
Jun 95 – Oct 2000	68.03	2472.14
Nov 2000 – Mar2006	59.61	2512.62

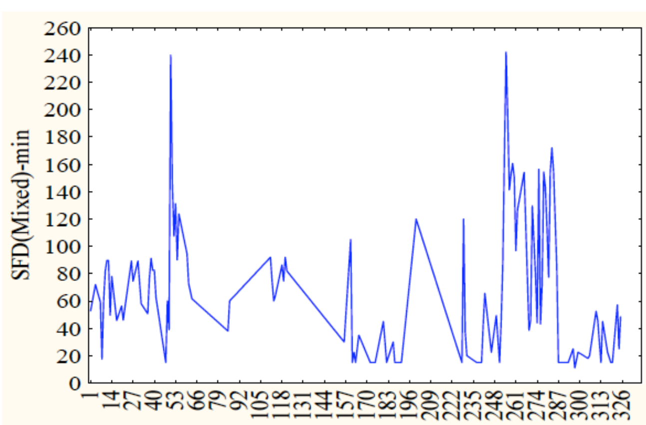

Figure 1: Average monthly observed Solar Flare Duration (SFD).

Stationary or non-stationary condition is the property of the process and not the data. Non-stationarity arises when the mechanism producing the data changes in time. However, a time series too short to capture the slowest variations of the measured quantity may produce the same effect.

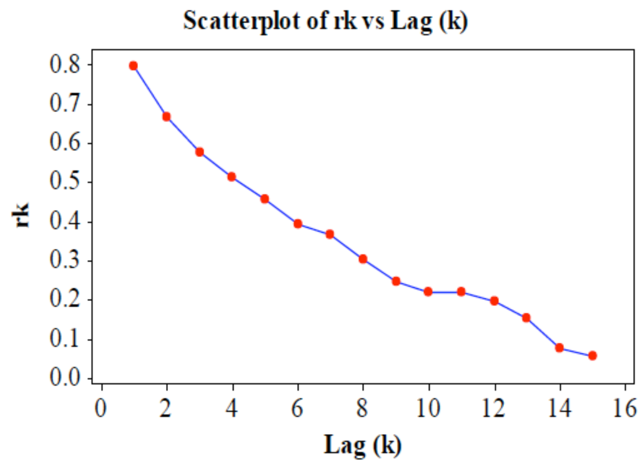

Figure 2: Non-stationary behavior revealed from data of SFD plotted against the correlation coefficient (r_k) and the time lag (k).

An autoregressive process will only be stable if the parameters are within a certain range: for example, if there is only one autoregressive parameter then is must fall within the interval of $-1 < x_t < 1$. In case of AR (1) the parameter estimated as 0.94243 which shows stationary series. However, as $\phi_1 \rightarrow 1$ there is indication of seasonal and / or polynomial trends in the series [10, 11].

For that purpose we get the first difference of the original series and calculate their autocorrelation. This indicates stationary condition as all of the autocorrelations are not significantly different from zero except lag1.

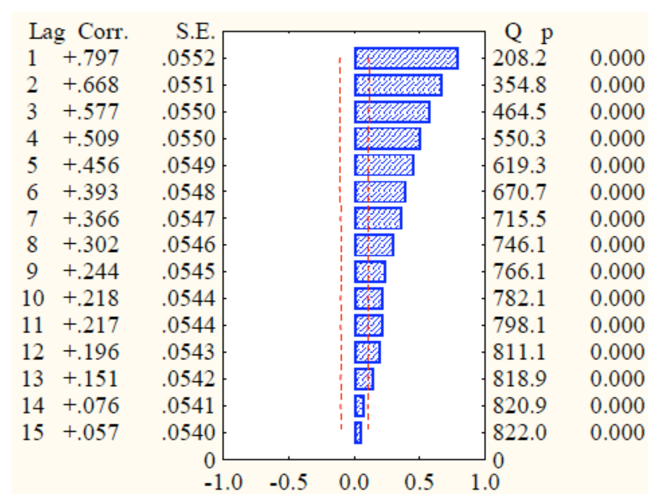

Figure 3: Autocorrelation Function for the series of SFD.

Model Identification

This has been identifying that the process of solar flare duration is autoregressive because its autocorrelation coefficient (of original series) decline to zero exponentially and its partial autocorrelation drop after lag 1 and has no exponential behavior suggesting that an MA model would be inappropriate [12, 13].

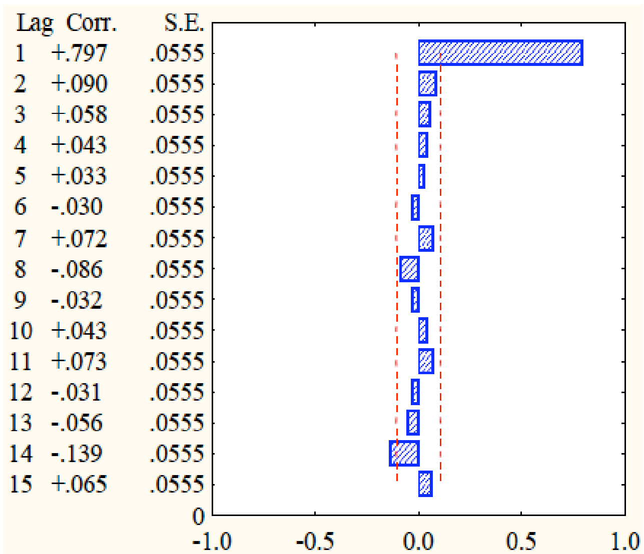

Figure 4: Partial Autocorrelation Function for the series of SFD.

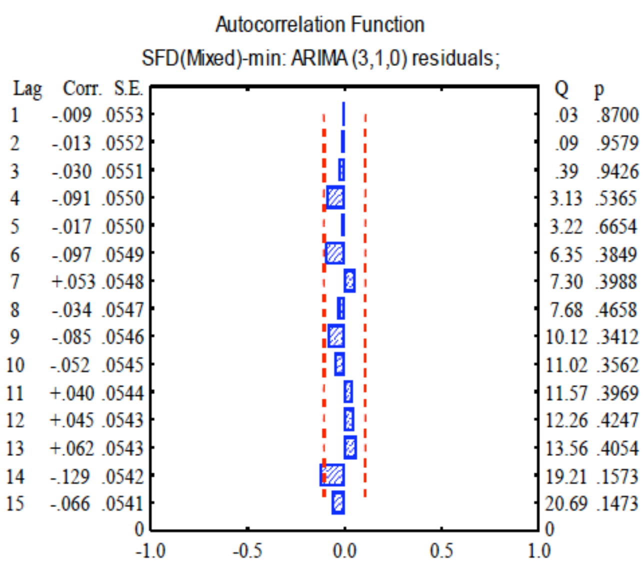

Figure 5: ARIMA for the mixed series of flares.

The model equation for ARIMA (3, 1, 0) is as follows.

$$X_t = \mu + X_{t-1} + \beta_1(X_{(t-1)} - X_{(t-2)}) + \beta_2(X_{(t-2)} - X_{(t-3)}) + \beta_3(X_{(t-3)} - X_{(t-4)}) \quad (1)$$

Diagnostic Checking

For an appropriate model residuals are expected to be random and close to zero. It is important to look at the few values of r_k, particularly at lags 1, 2 and see if any are significantly different from zero using the crude limit of $\pm\frac{2}{\sqrt{N}}$. If they are there is need to modify the model. However, if only one (or two) values of r_k are just significant at lags which have no obvious physical meaning (e.g. k = 5), then there would not be enough evidence to reject the model [14].

For the data series of SFD, ARIMA (3, 1, 0) has only one of their ACF at r_{14} is outside the confidence interval.

Coefficient of Determination R^2: If R^2 =1, then 100 per cent of the total variation in the dependent variable y has been explained by the model. The fit of the model is said to be 'better' the closer the value of R^2 = 1.

The coefficient of determination (R^2) can be obtained by the following relationship.

$$R^2 = 1 - \frac{SS_E}{SS_y} \quad (2)$$

where SS_E = Residual sum of squares

SS_y = Total sum of squares

The coefficient of determination found highest for ARIMA (3, 1, 0).

The object here is to find a model that minimizes the differences between the forecast values and the actual values. The quality of forecast has been obtained by the following methods.

Mean Absolute Forecast Error (MAFE)

$$\text{MAFE} = \frac{\sum|y_t - \hat{y}_t|}{m} \quad (3)$$

where y_t is the actual value of Y observed at time t and $\hat{y}_t$ is the forecast value of Y for time t.

Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{\sum[|y_t - \hat{y}_t|/y_t]}{m} \times 100 \quad (4)$$

Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum(y_t - \hat{y}_t)^2}{m}} \quad (5)$$

where m is the number of time periods for which forecasts have been made.

Table 2: Summary of Residual Analysis for Mixed Series of Flares

	R^2	MAFE	MAPE	RMSE
ARIMA (3, 1, 0)	0.580	12.48	56.67 %	13.73

Table 3: Forecast of Mixed Series of Flares from ARIMA (3, 1, 0)

Forecast	Lower 95.00 %	Upper 95.00 %	Std. Error
46.814	0.041	93.588	23.66
47.019	− 12.237	106.276	23.63
44.687	− 22.461	111.835	23.59
45.332	− 27.906	118.570	23.55
45.473	− 34.291	125.236	23.52

CONCLUSION

The final selected model for forecasting the mixed series of flares is ARIMA (3, 1, 0). The stochastic models are more reliable to forecast solar flare activity. The natural events are mostly non-stationary and they can be predicted better by stochastic process. However, this ARIMA model is being selected under the local condition and may vary with other spatial conditions.

REFERENCES

[1] Dwivedi BN, Parker EN. Dynamic Sun. Cambridge University Press 2003.

[2] Holman GD. The Mysterious Organize of Solar Flare. Scientific American, April 2006.

[3] Mullan DJ. Physics of the Sun. CRC Press Chapman & Hall / CRC Taylor & Francis Group 2010.

[4] Lang KR. The Cambridge Encyclopedia of the Sun. Cambridge University Press 2001.

[5] Forouzan BA. Data communication and networking. Second Edition, McGraw-Hill 2000.

[6] Lajos T. Stochastic Processes. Science Paperbacks and Methuen & Co. Ltd. 1966.

[7] Peter G. Stochastic Modeling of Scientific Data', First Edition, Chapman & Hall 1995.

[8] Wall JV, Jenkins CR. Practical statistics for Astronomers', Cambridge University Press 2003.

[9] Graham B. Statistics of Earth Science Data. Springer-Verlag Berlin Heidelberg N.Y. 2003.

[10] Chatfield C. Analysis of Time Series. Chapman & Hall, London 1989.

[11] Clinton SJ. Chaos and Time-Series Analysis. Oxford University Press 2003.

[12] Makridakis S, Wheelwright SC, Mcgee VE. Forecasting: Methods and applications. 2nd ed., John Wiley & Sons. 1983.

[13] Diebold FX. Elements of Forecasting. South-Western College Publishing, Cincinnati, Ohio 1998.

[14] Chatfield C. Analysis of Time Series. Chapman & Hall, London 1989.

Received on 18-02-2015 Accepted on 04-03-2015 Published on 06-03-2015

http://dx.doi.org/10.6000/1927-5129.2015.11.34