Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 56 (1): 59–68 (2019)

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

Pakistan Academy of Sciences

Research Article

Heavy Tail Analysis of Sunspot Cycles Based on Stochastic Modeling

Danish Hassan[1,2*]

[1]Mathematical Sciences Research Centre, Federal Urdu University of Arts, Sciences and Technology, Karachi, Pakistan
[2]Department of Applied Sciences, National Textile University, Karachi Campus

Abstract: Among other stochastic models, fractional auto regressive integrating moving average (FARIMA) is distinct because of its appropriateness for modeling stationary time series with long range dependence (long memory or persistence). Results obtained in this manuscript shows appropriateness of FARIMA model for the analysis of sunspot number. Analyzing for stationary, each cycle out of the 24 sunspot cycles were modeled. FARIMA can be used for modeling using different techniques. In view of the parameters obtained by maximum likelihood test two most appropriate techniques are adopted. These two are Direct Method and Whittle approximation Method. Results are obtained by applying these two types of FARIMA, the significance of these two methods were observed and compared using significance tests. For FARIMA models the fractional differencing parameter d is most decisive for the determination of persistency. In this regard four types of model (0, d,0), (1, d,0), (0, d,1) and (1, d,1) are used. The adequacy of each of the models is determined with the help of Akaike, Bayesian-Schwarz and Hannan-Quinn Information criterion. The investigations made using these models are reliable for both short and long sunspot cycles. Finally, tail analysis is performed in view of the parameter (α) it is observed that heavy tails exist for each sunspots cycle confirming long range dependence. The study is useful to examine the sunspot historical data using the FARIMA model to understand their long term behavior.

Keywords: Akaike Information criterion (AIC), Bayesian-Schwarz Information criterion (BIC), Hannan-Quinn Information criterion (HIC), Log-likelihood, FARIMA & Heavy tails.

1. INTRODUCTION

Sunspot cycles exhibit random variations in their lengths and occurrence of peaks. According to the past records [1] the maximum and minimum length of the cycles is found to be 14-years and 9-years respectively. The type of modeling needed to study the variation of solar activity cycles need to be capable of handling short and long term and long-term forecasts [2]. The different solar cycles are depicted in Fig. 1 indicating the various periods (Maunder minimum, Dalton minimum and Modern maximum) over the last 400 years.

FARIMA modeling is the stochastic process that deals with both short-term and long-term variations. This type of modeling is more feasible than the traditional models like Moving Average (MA), Autoregressive (AR), Autoregressive Moving Average (ARMA) and Autoregressive Integrating Moving Average ARIMA [3]. These traditional models do not provide significant information about the noises of long-term data. The fractional Gaussian noise (FGN) and fraction differencing noise (FDN) are the self-similarity models which are adequate for long-term data only [4]. Both FGN and FDN are the types of FARIMA. The FARIMA (p, d, q) models strongly depend on the fractional differencing parameter d. Data shows stationary behavior if d belongs to the interval (-0.5, 0.5). If d belongs to the interval (0, 0.5) then the data shows a persistent behavior, for $d < 0$ it will be anti-persistent and for $d = 0$ it will show Brownian behavior. The increasing of parameter d from 0 to

Received: June 2018; Accepted March 2019
* Corresponding Author: Danish Hassam <danish10ansari@gmail.com>

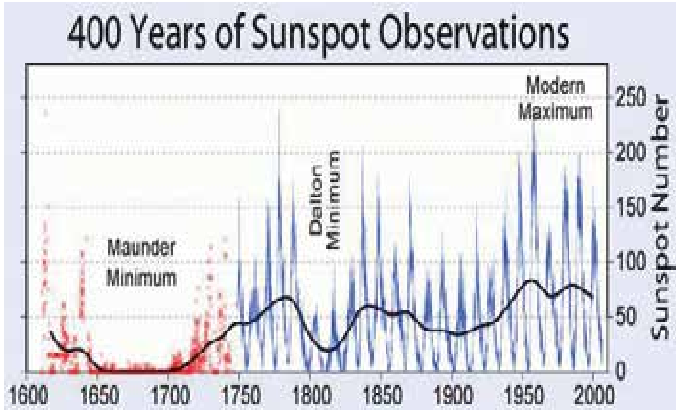

Fig. 1. The durations of sunspot cycles (1600 to 2000) for the last 400 years (Curtsy by Thomas and Weiss, 1992)

0.5 indicates the increasing of smoothness of any noise. The differencing parameter *d* depends on the self-similarity parameter *H*, these two are related by the equation $d = H - 0.5$ [5]. The significance and adequacy of the FARIMA models is determined by Akaike information criterion (AIC), Bayesian Schwarz information criterion (BIC) and Hannan Quinn information criterion (HIC). The range of Heavy Tail (HT) parameter α also depends on the differencing parameter and hence on the self-similarity parameter [5, 6], for $d > 0$ the HT parameter $\alpha > 1$. For any noise (sunspot number) the HT parameter is α belongs to the interval (0, 2) [7].

2. MATERIALS AND METHODS

This manuscript consists of the analysis of 24 sunspots cycles from 1754 to 2014. Several cycles have the same length approximately. For example, four cycles (2,3,5,16) are of 9-years duration, five cycles (8,18,19,21,22) are of 10-years duration, five cycles (7,11,12,14,17) are of 11-years duration, six cycles (1,9,10,13,15,20) are of 12-years duration, one cycle (23) of 13-years duration and two cycles (4,6) are of 14-years duration. FARIMA models (0, *d*,0), (1, *d*,0), (0, *d*,1) and (1, *d*,1) are investigated for appropriateness using AIC, BIC and HIC. For each sunspot cycle the maximum likelihood parameter ϕ and θ for standard error were estimated. Log likelihood parameters have also been estimated for each sunspot cycle. For the FARIMA models, each cycle has been further categorize and analyze using direct and the Whittle approximation method (Periodogram). The heavy tail parameter α is also calculated for each cycle to examine the strength of the cyclic data within the tails.

2.1. FARIMA Modeling of Sunspot Cycles

It is mentioned earlier that the duration of short sunspots cycle is 9 years and long sunspots cycle is 14 years so, for the analysis of sunspot cycles requires a modeling that can handle both short and long sample sizes. FARIMA modeling is such a modeling that can be used to analyze both the short and long range data. The FARIMA (*p*, *d*, *q*) models depend on parameters *p* (Autoregressive (AR)) and *q* (Moving Average (MA)) mainly on the differencing parameter *d*. The form of simple FARIMA (*p*, *d*, *q*) model is as the following.

$$\Phi(\mathrm{B})(1-\mathrm{B})^{\mathrm{d}}\mathrm{X}_{\mathrm{t}} = \theta(\mathrm{B})\epsilon_{\mathrm{t}}, \text{ for } \mathrm{d} \in (-0.5, 0.5) \quad (2.1)$$

Where $\{\epsilon_t\}$ is called a white noise and *B* is the back shift operator. White noise is the energy per frequency and representing the equal intensity of random signal at different frequencies. The condition $-0.5 < d < 0.5$ indicates the stationary behavior of any noise [8].

For the short range cycles the autocorrelation function (ACF (ρ)) test exhibits exponential decay which is defined as $\sum_{\tau=-\infty}^{\infty} \rho\mathrm{SRD}(\tau) = \mathrm{const} < \infty$ and also called short range correlation (Short Range Dependence). For the long range cycles the ACF diverges to a sum which is represented as $\sum_{\tau=-\infty}^{\infty} \rho\mathrm{LRD}(\tau) = \infty$ (Long Range Dependence).

$$\gamma(\tau) = \mathrm{L}(\tau)\tau^{-2(1-\mathrm{H})} \text{ with } \frac{1}{2} < H < 1. \quad (2.2)$$

The power spectral density or spectral density applies to larger signals over a time period, the time interval can be infinite. The long range cycles follow fractional process which is associated with the spectral density of equation (2.2) expressed as follows

$$f(\lambda) \sim \left|1 - e^{i\lambda}\right|^{-2H+1} L\left(\frac{1}{\lambda}\right),\ \lambda \to 0^{+} \quad (2.3)$$

Where the function $L(\lambda)$ varies regularly. Where '~' represents the left and right hand side ratio converges to one [2, 8]. The fractional Brownian motion (fBm) is the common form of Brownian motion and also known as fractal Brownian motion. The fBm from a stationary sequence is given by

$$Y_H(k) = B_H(k) - B_H(k-1),\ K \in \mathbb{Z} \quad (2.4)$$

The sequence $\{Y_H(k)\}_{k\in\mathbb{Z}}$ is known as the fractional Gaussian noise (FGN) [4]. The moving average components and autoregressive are expressed by the following polynomials in *q* and *p* respectively.

$$\phi(z) = 1 - \sum_{i=1}^{p} \emptyset_i z^i,\ \varphi(z) = 1 + \sum_{j=1}^{q} \varphi_j z^j, \quad (2.5)$$

The operator (1-*B*)*d* is known as the fractional differencing operator which expands as a power series. The FARIMA (*p*, *d*, *q*) process is stationary for $0 < d < 0.5$. $\emptyset_i$ and φ_j are the Maximum likelihood parameters $(\phi_1, \dots, \phi_p, d, \varphi_1, \dots, \varphi_q)$ are estimated by Whittle's approximation [2, 9].

2.2. Maximum Likelihood Parameters

The variation of parameters for each sunspot cycle depends on the number of spots in it. The stationary FARIMA (*p*, *d*, *q*) modeling works as a Gaussian process depending on the parameter

$$\theta = (d, a_1, \dots, a_p, b_1, \dots, b_p, \sigma_\eta) \quad (2.6)$$

The probability density of the estimates of the data is expressed as follows.

$$p(x|\theta) = (2\pi)^{-\frac{N}{2}} |\Sigma\, \theta|^{-\frac{1}{2}} e^{-\frac{1}{2}x^{+}\Sigma_{-1}(\theta)x}, \quad (2.7)$$

Where $\Sigma(\theta)$ is known as the autocovariance matrix and x^{+} represents the transpose of matrix *x* of the forecasts. The log-likelihood for the above condition is given by $l(\theta|x') = \log \mathcal{L}(\theta|x')$. The x' is the given realization on which θ is based [5, 9, 14]. The maximum log-likelihood in terms of argument can be describe as

$$\hat{\theta} = \arg \max_{\theta} l(\theta|x') \quad (2.8)$$

2.3. Whittle Approximation Method

This method approximates in terms of log-likelihood and depends on the parameter θ.

$$l(\theta|\, x') = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log|\Sigma(\theta)| - \frac{1}{2}x^{+}\Sigma_{-1}(\theta)x \quad (2.9)$$

The determinant $\Sigma|(\theta)|$ and $x^{+}\Sigma_{-1}$ strongly depend on θ. The approximation for both the terms range over the spectral density and involves an integral $S(\omega;\theta)$ This process is approximated by a Riemann sum for the successive step of spectral density. The discrete version for the whittle estimator for the minimization is expressed as follows.

$$Q(\theta^{*}) = \sum_{j=1}^{[(N-1)/2]} \frac{I(\omega_j)}{S(\omega_j;\theta^{*})} \quad (2.10)$$

Where $I(\omega_j)$ is the periodogram of the realization x' at frequencies of the Fourier $\omega_j = \frac{2\pi j}{N}$ and j = 1,... $\left[\frac{N-1}{2}\right]$. In view of [2] the discrete version with respect to θ^{*} of the Whittle estimator is given as follows.

$$\hat{\sigma}_\eta^2 = 2\pi\hat{\theta}_1 = \frac{4\pi}{N} Q(\theta^{*}). \quad (2.11)$$

2.4. Heavy Tail Parameters for Sunspot Cycles

For the sunspot cycle's heavy tails parameters strongly depends on the self-similarity parameter *H* also known as Hurst exponent. For the persistent noise the range of heavy tail parameters is $0 < \alpha < 2$. The long-range dependence parameter *d* has the range [0, 1 - 1/ α] whereas $H < 1$. The above condition is a necessary condition for a stationary series. For finite variance $H = d + ½$ and for infinite variance $H = d + 1/\alpha$. Since $d > 0$ so $1 < \alpha$. As suggested by [5, 10 & 11], the degree of long-range dependence is given by.

$$H = \frac{(3-\alpha)}{2} \quad (2.12)$$

For $d < 0.5$ it follows a power series expansion for all the sunspot cycles. For $\alpha < 2$, the tails are asymptotically equivalent to Pareto law [12].

2.5. Significance Tests for FARIMA Modeling

To develop suitable adequate models reliable significant tests are needed. For this purpose Akaike information criterion (AIC), Bayesian information criterion (BIC) and Hannan Quinn information criterion (HIC) are used to test the FARIMA (p, d, q) models. AIC is most easy to apply. It provides the distance between true probability density and probability estimated by the density function, $p(x|\theta)$. AIC is defined as follows:

$$\mathrm{AIC} = -2\mathrm{l}(\theta|x) + 2m, \tag{2.13}$$

Here is the number of parameters used in the model and $l(\theta|x)$ represents the log-likelihood. According to AIC the best model has the smallest distance from the true model attaining the smallest value with respect to AIC.Including the sample size N the difference of the AIC is known as Hannan-Quinn information criterion define as:

$$\mathrm{HIC} = -2l(\theta|x) + 2mc(\log(\log N)), \text{ here } c > 1 \tag{2.14}$$

The Bayesian information criterion (BIC) is formulated under the framework of the Bayesian modeling. It is one of the consistent criteria that provides the adequacy for the models in short and long sample sizes. It is define as:

$$\mathrm{BIC} = -2l(\theta|x) + m(\log N) \tag{2.15}$$

with N as the sample size of the used data. BIC is also known as Schwarz information criterion [1, 5 & 12].

3. RESULTS AND DISCUSSIONS

The variations of 24 sunspots cycles are discuss here in the using FARIMA models. Four FARIMA models (0, d, 0), (1, d, 0), (0, d, 1) and (1, d, 1) are utilized here. These models are tested according to AIC, BIC and HIC. The adequacy of these models is also tested using Log-likelihood. For each cycle the results obtained were then compared using direct method and Whittle approximation method. Out of the four models mentioned above FARIMA (1, d, 1) is found to be significant. The results obtained by Log-likelihood parameters using Whittle approximation are depicted in Tables 1, 2 and Fig 2. These results also confirmed the significance of (1, d, 1) model. The adequacy is further tested by using AIC, BIC and HIC which further asserted the reliability of FARIMA (1, d, 1) model. In view of the errors it is found that HIC is most appropriate test, see table 2. Comparison of both the methods shows the Whittle approximation method is more reliable then the direct method. The results are depicted in Table 1 & 2. Three main parameters used in the study are self-similarity parameter H, differencing parameter d and heavy-tails parameter. In view of the four models solar cycle 4 show the least and cycle 5 shows the greatest log-likelihood values. Whittle approximation method is found to be more adequate than the direct method for FARIMA (1, d, 1) because it is based on Periodogram. In case of direct method FARIMA (0, d, 1) appeared to be inappropriate for all the solar cycles as the parameter θ does not converge. Similarly, the standard error estimated by FARIMA (0, d, 1) for 24 solar cycles did not appear appropriate. The standard error for

Table 1. FARIMA (1,d,1) model along with the related parameters of sunspots cycles (1-24) by direct method

Direct method											
Cycle	**N**	**d**	**φ1**	**St. error**	**θ1**	**St. error**	**log LH**	**AICC**	**BIC**	**HIC**	**Model**
1	150	0.3147	0.945723	0.071618	-0.78263	0.183786	-648.9	1306.0	1308.7	1301.2	X(t) - 0.946X(t-1) = Z(t) - 0.783Z(t-1)
2	104	0.2216	0.745855	0.245248	-0.29118	0.401332	-492.7	993.6	995.5	988.4	X(t) - 0.746X(t-1) = Z(t) - 0.291Z(t-1)
3	111	0.2229	0.940371	0.062143	-0.6585	0.181504	-550.5	1109.2	1111.2	1104.1	X(t) - 0.940X(t-1) = Z(t) - 0.658Z(t-1)
4	168	0.2611	0.953839	0.051188	-0.72211	0.170792	-804.1	1616.3	1619.3	1611.6	X(t) - 0.954X(t-1) = Z(t) - 0.722Z(t-1)
5	115	0.3602	0.923946	0.110671	-0.73741	0.25703	-446.6	901.4	903.5	896.4	X(t) - 0.924X(t-1) = Z(t) - 0.737Z(t-1)

6	169	0.3516	0.957925	0.042614	-0.89021	0.06858	-671.4	1350.9	1353.9	1346.2	X(t) - 0.958X(t-1) = Z(t) - 0.890Z(t-1)
7	137	0.2949	0.956451	0.049627	-0.82554	0.118001	-597.3	1202.7	1205.2	1197.8	X(t) - 0.956X(t-1) = Z(t) - 0.826Z(t-1)
8	116	0.213	0.923182	0.105612	-0.59649	0.343552	-567.4	1143.0	1145.1	1137.9	X(t) - 0.923X(t-1) = Z(t) - 0.596Z(t-1)
9	151	0.2927	0.920345	0.10058	-0.68624	0.249373	-709.6	1427.3	1430.1	1422.5	X(t) - 0.920X(t-1) = Z(t) - 0.686Z(t-1)
10	136	0.2999	0.951628	0.053938	-0.7667	0.144204	-607.1	1222.3	1224.8	1217.5	X(t) - 0.952X(t-1) = Z(t) - 0.767Z(t-1)
11	139	0.2343	0.930396	0.080769	-0.61194	0.267388	-671.7	1351.7	1354.2	1346.8	X(t) - 0.930X(t-1) = Z(t) - 0.612Z(t-1)
12	135	0.3298	0.939194	0.077979	-0.78959	0.17407	-584.2	1176.6	1179.1	1171.7	X(t) - 0.939X(t-1) = Z(t) - 0.790Z(t-1)
13	137	0.3304	0.953495	0.050701	-0.81508	0.114329	-606.3	1220.7	1223.2	1215.8	X(t) - 0.953X(t-1) = Z(t) - 0.815Z(t-1)
14	130	0.3119	0.960515	0.043698	-0.86708	0.080856	-567.4	1142.9	1145.3	1138.0	X(t) - 0.961X(t-1) = Z(t) - 0.867Z(t-1)
15	143	0.2589	0.919588	0.093182	-0.64935	0.249788	-653.5	1315.1	1317.7	1310.3	X(t) - 0.920X(t-1) = Z(t) - 0.649Z(t-1)
16	115	0.3113	0.854201	0.204348	-0.54303	0.416137	-502.2	1012.6	1014.7	1007.5	X(t) - 0.854X(t-1) = Z(t) - 0.543Z(t-1)
17	128	0.2347	0.92585	0.100651	-0.62454	0.313174	-604.5	1217.2	1219.5	1212.2	X(t) - 0.926X(t-1) = Z(t) - 0.625Z(t-1)
18	117	0.2134	0.888777	0.129052	-0.4193	0.379693	-583.0	1174.2	1176.3	1169.1	X(t) - 0.889X(t-1) = Z(t) - 0.419Z(t-1)
19	126	0.2222	0.938647	0.071783	-0.60021	0.272444	-659.6	1327.4	1329.7	1322.4	X(t) - 0.939X(t-1) = Z(t) - 0.600Z(t-1)
20	144	0.2689	0.936897	0.074823	-0.67252	0.23057	-665.6	1339.4	1342.1	1334.6	X(t) - 0.937X(t-1) = Z(t) - 0.673Z(t-1)
21	119	0.2376	0.939469	0.076629	-0.66049	0.259154	-599.7	1207.6	1209.8	1202.6	X(t) - 0.939X(t-1) = Z(t) - 0.660Z(t-1)
22	124	0.2432	0.943264	0.062809	-0.67958	0.198364	-627.3	1262.9	1265.1	1257.9	X(t) - 0.943X(t-1) = Z(t) - 0.680Z(t-1)
23	154	0.2405	0.953301	0.055028	-0.72948	0.187148	-734.2	1476.5	1479.3	1471.7	X(t) - 0.953X(t-1) = Z(t) - 0.729Z(t-1)
24	56	0.199	0.894076	0.129387	-0.43066	0.351102	-237.8	484.0	484.3	478.0	X(t) - 0.894X(t-1) = Z(t) - 0.431Z(t-1)
1-24	3184	0.387	0.965347	0.014941	-0.8055	0.032628	-15198.7	30405.5	30415.0	30402.9	X(t) - 0.965X(t-1) = Z(t) - 0.806Z(t-1)

Table 2. FARIMA (1, d, 1) model along with the related parameters of sunspots cycles (1-24) by Whittle Approximation method

Whittle approximation method											
Cycle	**N**	**D**	**φ1**	**St. error**	**θ1**	**St. error**	**log LH**	**AICC**	**BIC**	**HIC**	**Model**
1	150	0.3147	0.9999	0.009444	-0.76754	0.139972	-642.8	1293.62	1296.5	1289.0	X(t) - 1.000X(t-1) = Z(t) - 0.768Z(t-1)
2	104	0.2216	0.9999	0.01185	-0.58608	0.237374	-488.1	984.127	986.2	979.2	X(t) - 1.000X(t-1) = Z(t) - 0.586Z(t-1)
3	111	0.2229	0.9999	0.010573	-0.61233	0.18434	-544.7	1097.33	1099.6	1092.4	X(t) - 1.000X(t-1) = Z(t) - 0.612Z(t-1)
4	168	0.2611	0.9999	0.00856	-0.6604	0.185069	-797.5	1603	1606.1	1598.5	X(t) - 1.000X(t-1) = Z(t) - 0.660Z(t-1)
5	115	0.3602	0.9999	0.010559	-0.69345	0.202215	-440.2	888.45	890.8	883.6	X(t) - 1.000X(t-1) = Z(t) - 0.693Z(t-1)

6	169	0.3516	0.9999	0.009329	-0.89019	0.052877	-665.9	1339.79	1342.9	1335.3	X(t) - 1.000X(t-1) = Z(t) - 0.890Z(t-1)
7	137	0.2949	0.9999	0.009832	-0.80237	0.110543	-591.5	1191.09	1193.8	1186.4	X(t) - 1.000X(t-1) = Z(t) - 0.802Z(t-1)
8	116	0.213	0.9999	0.010527	-0.54961	0.277432	-561.7	1131.32	1133.6	1126.5	X(t) - 1.000X(t-1) = Z(t) - 0.550Z(t-1)
9	151	0.2927	0.9999	0.009357	-0.69272	0.174031	-703.5	1414.99	1417.9	1410.4	X(t) - 1.000X(t-1) = Z(t) - 0.693Z(t-1)
10	136	0.2999	0.9999	0.009615	-0.72834	0.143866	-600.8	1209.63	1212.3	1204.9	X(t) - 1.000X(t-1) = Z(t) - 0.728Z(t-1)
11	139	0.2343	0.9999	0.009459	-0.54508	0.255775	-665.5	1338.98	1341.7	1334.3	X(t) - 1.000X(t-1) = Z(t) - 0.545Z(t-1)
12	135	0.3298	0.9999	0.009921	-0.77293	0.137139	-578.2	1164.47	1167.1	1159.7	X(t) - 1.000X(t-1) = Z(t) - 0.773Z(t-1)
13	137	0.3304	0.9999	0.009693	-0.78927	0.10942	-600.1	1208.22	1210.9	1203.5	X(t) - 1.000X(t-1) = Z(t) - 0.789Z(t-1)
14	130	0.3119	0.9999	0.010543	-0.86218	0.065816	-562.1	1132.25	1134.8	1127.5	X(t) - 1.000X(t-1) = Z(t) - 0.862Z(t-1)
15	143	0.2589	0.9999	0.009544	-0.64155	0.19023	-647.5	1303.02	1305.8	1298.4	X(t) - 1.000X(t-1) = Z(t) - 0.642Z(t-1)
16	115	0.3113	0.9999	0.010709	-0.61379	0.22638	-496.3	1000.56	1002.9	995.7	X(t) - 1.000X(t-1) = Z(t) - 0.614Z(t-1)
17	128	0.2347	0.9999	0.010009	-0.5919	0.237746	-598.6	1205.13	1207.7	1200.4	X(t) - 1.000X(t-1) = Z(t) - 0.592Z(t-1)
18	117	0.2134	0.9999	0.010347	-0.39348	0.293814	-576.9	1161.81	1164.1	1157.0	X(t) - 1.000X(t-1) = Z(t) - 0.393Z(t-1)
19	126	0.2222	0.9999	0.009772	-0.48266	0.290022	-653.2	1314.44	1316.9	1309.7	X(t) - 1.000X(t-1) = Z(t) - 0.483Z(t-1)
20	144	0.2689	0.9999	0.00932	-0.61933	0.208117	-659.2	1326.46	1329.3	1321.8	X(t) - 1.000X(t-1) = Z(t) - 0.619Z(t-1)
21	119	0.2376	0.9999	0.010209	-0.57609	0.267986	-593.7	1195.33	1197.7	1190.5	X(t) - 1.000X(t-1) = Z(t) - 0.576Z(t-1)
22	124	0.2432	0.9999	0.009951	-0.61694	0.204629	-621.3	1250.51	1253.0	1245.7	X(t) - 1.000X(t-1) = Z(t) - 0.617Z(t-1)
23	154	0.2405	0.9999	0.008994	-0.66821	0.194277	-728.0	1464.06	1467.0	1459.5	X(t) - 1.000X(t-1) = Z(t) - 0.668Z(t-1)
24	56	0.199	0.814928	0.201703	-0.20047	0.219182	-240.2	488.41	489.2	482.8	X(t) - 0.815X(t-1) = Z(t) - 0.200Z(t-1)
1-24	3184	0.387	0.999898	0.002947	-0.82922	0.02437	-15190.5	30388.9	30398.4	30386.4	X(t) - 1.000X(t-1) = Z(t) - 0.829Z(t-1)

Table 3. The long range dependence (LRD), fractional differencing (FD) and heavy tail (HT) parameter of sun-spot cycles (1-24)

Cycle	Duration	N	0.5<H<1 (LRD)	0 < d < 0.5 (FD)	1 < α < 2 (HT)
1	1754.01-1766.06	150	0.8147	0.3147	1.3706
2	1766.07-1775.02	104	0.7216	0.2216	1.5568
3	1775.03-1784.05	111	0.7229	0.2229	1.5542
4	1784.06-1798.05	168	0.7611	0.2611	1.4778
5	1798.06-1807.12	115	0.8602	0.3602	1.2796
6	1808.01-1822.01	169	0.8516	0.3516	1.2968
7	1822.02-1833.06	137	0.7949	0.2949	1.4102
8	1833.07-1843.02	116	0.713	0.213	1.574
9	1843.03-1855.09	151	0.7927	0.2927	1.4146
10	1855.1-1867.01	136	0.7999	0.2999	1.4002

11	1867.02-1878.08	139	0.7343	0.2343	1.5314
12	1878.09-1889.11	135	0.8298	0.3298	1.3404
13	1889.12-1901.04	137	0.8304	0.3304	1.3392
14	1901.05-1912.02	130	0.8119	0.3119	1.3762
15	1912.03-1924.01	143	0.7589	0.2589	1.4822
16	1924.02-1933.08	115	0.8113	0.3113	1.3774
17	1933.09-1944.04	128	0.7347	0.2347	1.5306
18	1944.05-1954.01	117	0.7134	0.2134	1.5732
19	1954.02-1964.07	126	0.7222	0.2222	1.5556
20	1964.08-1976.07	144	0.7689	0.2689	1.4622
21	1976.08-1986.06	119	0.7376	0.2376	1.5248
22	1986.07-1996.1	124	0.7432	0.2432	1.5136
23	1996.11-2009.08	154	0.7405	0.2405	1.519
24	2009.09-…….	56	0.699	0.199	1.602
1-24	1754-2014	3184	0.887	0.387	1.226

Table 4. Relative empirical mean squared error (REMSE) of FARIMA (p, d, q) models using Whittle approximation method for sunspots cycles

sunspot cycles	FARIMA(0,d,0)		FARIMA(1,d,0)		FARIMA(0,d,1)		FARIMA(1,d,1)	
	AICC/BIC	HIC/BIC	AICC/BIC	HIC/BIC	AICC/BIC	HIC/BIC	AICC/BIC	HIC/BIC
1	0.998053	0.996561	0.99793	0.995404	0.997933	0.995412	0.997778	0.994212
2	0.997917	0.995794	0.997907	0.994302	0.997915	0.994324	0.997886	0.992864
3	0.998106	0.996212	0.998023	0.994838	0.998047	0.9949	0.997975	0.993526
4	0.998389	0.997194	0.998199	0.9962	0.998214	0.996231	0.998053	0.995235
5	0.997581	0.995235	0.997497	0.993606	0.997506	0.993626	0.997414	0.991958
6	0.998012	0.996595	0.997845	0.995465	0.997846	0.995465	0.997662	0.994296
7	0.998006	0.996334	0.99788	0.995081	0.997884	0.995091	0.997752	0.993813
8	0.998147	0.996295	0.998012	0.994947	0.998036	0.995009	0.997951	0.993671
9	0.998248	0.996862	0.998095	0.995784	0.998101	0.995798	0.997958	0.994702
10	0.998018	0.996405	0.997925	0.995168	0.997934	0.995188	0.997799	0.993915
11	0.998215	0.996763	0.998084	0.995588	0.998102	0.995629	0.997976	0.994481
12	0.998007	0.996252	0.997853	0.994981	0.997857	0.99499	0.997728	0.993688
13	0.998065	0.996381	0.997915	0.995164	0.997919	0.995172	0.997784	0.9939
14	0.997917	0.99616	0.997846	0.994865	0.997849	0.994872	0.997736	0.993551
15	0.998175	0.996641	0.998	0.995458	0.99801	0.99548	0.997874	0.994301
16	0.997815	0.995774	0.997773	0.99431	0.997783	0.994335	0.997703	0.992857
17	0.998135	0.99645	0.997995	0.995179	0.998013	0.995223	0.9979	0.993956
18	0.998172	0.996392	0.998047	0.995062	0.998074	0.995128	0.997989	0.993828
19	0.998266	0.996765	0.998176	0.995578	0.998195	0.995623	0.9981	0.994473
20	0.998201	0.996698	0.998025	0.99553	0.998041	0.995566	0.9979	0.994395
21	0.998214	0.996467	0.99808	0.995192	0.998101	0.995244	0.998015	0.993983
22	0.998187	0.996597	0.998111	0.99538	0.998131	0.995428	0.998031	0.994207
23	0.998315	0.996978	0.998129	0.995898	0.998142	0.995926	0.997998	0.994862
24	0.99756	0.99257	0.997967	0.989654	0.998007	0.989856	0.998485	0.98709
1-24	0.999754	0.999763	0.999737	0.999683	0.999737	0.999683	0.999687	0.999603

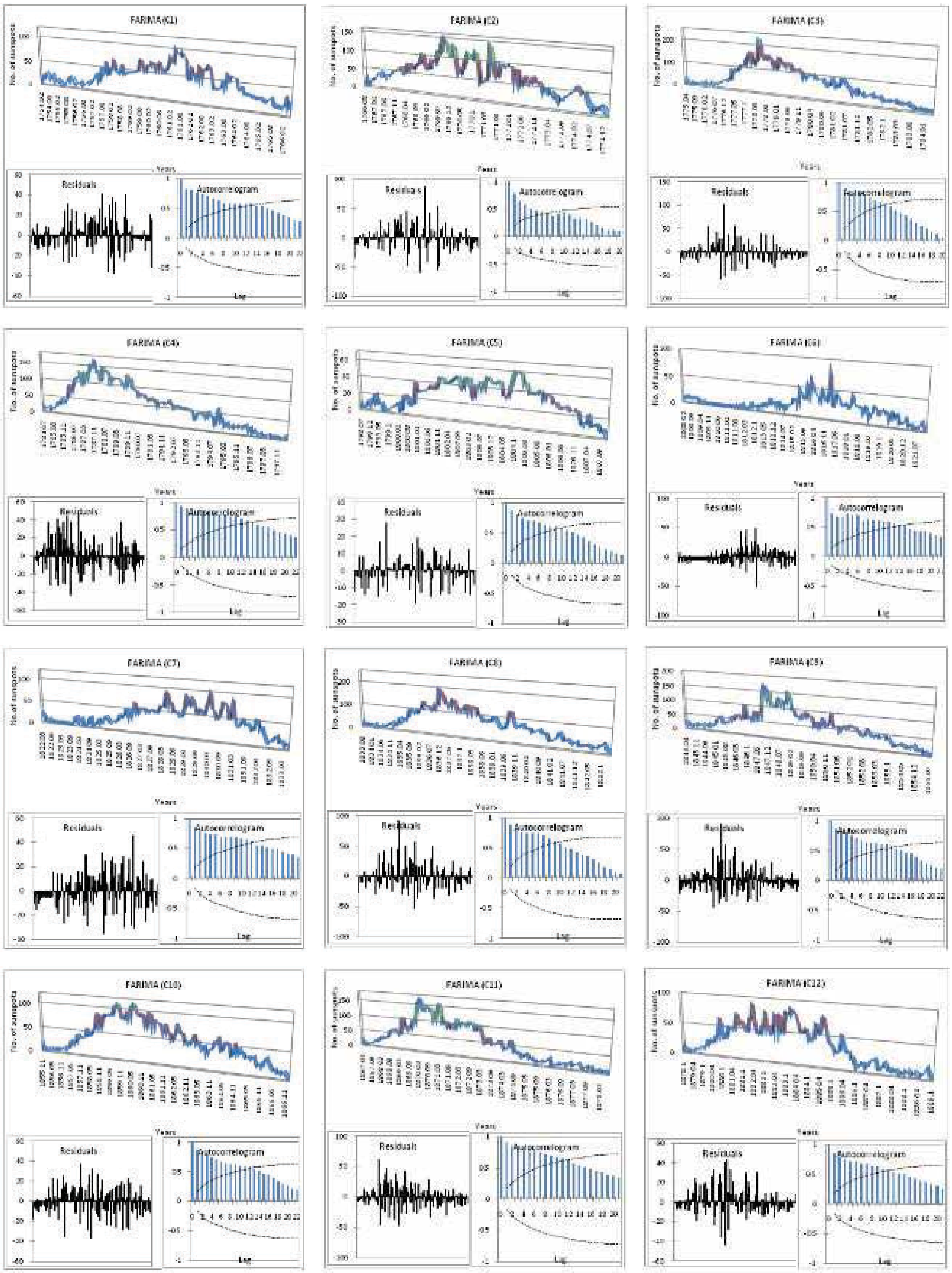

Fig. 2. Sunspot number's plot cycle wise (1-12) with the residuals and autocorrelogram plots. The autocorrelogram shows decreasing in correlation with increasing the lag.

all sunspot cycles follows the white noise process. Parameters obtained using Maximum likelihood estimator (MLE) for both the methods are depicted in Table1 & 2. The long-range dependence and differencing parameter are also estimated for each solar cycle. The results obtained confirmed that the solar cycles are stationary ($0 < d < 0.5$) as depicted in table 3. The sunspots data for each cycle are the positive random numbers and also persistent ($0.5 < H < 1$), furnished in table 3. The heavy tail parameter(α) was found significant in all the solar cycles (1 < < 2) and shows smallest value for solar cycle 5 and greatest for solar cycle 8 (see table 3). For $d < 0.5$, it shows a power series expansion for all the 24 sunspot cycles and for < 2, the tails are representing asymptotically equivalent to Pareto law. The relative empirical mean square prediction error (REMSPE) was estimated for each solar cycle to understand the model prediction reliability. This is obtained for each FARIMA (p, d, q) model in the perspective of BIC and found greater values for the ratio (AIC: BIC) in both methods (see Table 4). The prediction reliability of model FARIMA using Whittle approximation methods was confirmed.

4. CONCLUSION AND OUTLOOK

FARIMA models were developed and applied on each sunspots cycle. The significance of FARIMA models for each sunspot cycle was tested according to AIC, BIC and HIC. Results show that for each cycle HIC is more consistent than AIC and BIC. The adequacy of FARIMA (1, d, 1) have also been verified in the perspective of parameters obtained with the help of log-likelihood technique using Whittle approximation method. All the sunspot cycles exhibited stationary behavior as parameter d lies within the interval (0, 0.5). The heavy tail parameter (1 < < 2) was found in sunspot cycles, represents the smooth long-term behavior for solar activity and expected to continue in future. The REMSPE also confirmed the prediction reliability for the Whittle approximation method.

5. ACKNOWLEDGEMENT

The sunspots monthly data was taken from the NOAA. I am thankful to the editor and reviewers for their precious time and useful comments. The results of this paper are the part of my PhD thesis submitted to FUUAST Karachi in August, 2016 (http://prr.hec.gov.pk/jspui/handle/123456789/10946). The manuscript is dedicated to my beloved father Mr. Mujtaba Hassan Ansari (Late) and dearest friend Mr. Farhan (Late).

6. REFERENCES

1. Hassan, D., S. Abbas., M.R.K. Ansari, & B. Jan, The study of sunspots and K-index data in the perspective of probability distributions. *International Journal of Physical and Social Sciences*, *4*(1): 23-41(2014).
2. Ocher, D. Stationary & Nonstationary Farima Models-Model Choice, Forecasting, Aggregation & Intervention, Doctoral Dissertation, University of Konstanz, (1999), retrieved from: http://kops.uni-konstanz.de/handle/123456789/733.
3. Wang, M. J., G.H. Tzeng, & T.D. Jane, A fuzzy ARIMA model by using quadratic programming approach for time series data, *Int J Inf Syst Logist Manag*, *5*(1): 41-51 (2009).
4. Cappé, O., E. Moulines., J.C. Pesquet., A. Petropulu, & X. Yang, Long-range dependence and heavy-tail modeling for tele traffic data. *Signal Processing Magazine, IEEE*, *19*(3): 14-27 (2002).
5. Feldman, R, & M. Taqqudus, M, *A practical guide to heavy tails: statistical techniques and applications*. Springer Science & Business Media (1998).
6. Rust, H.W., M. Kallache., H.J. Schellnhuber, & J. Kropp, Confidence intervals for flood return level estimates using a bootstrap approach. *Springer*, 61-81(2011).
7. Meerschaert, M. M., P. Roy, & Q. Shao, Parameter estimation for exponentially tempered power law distributions. *Communications in Statistics-Theory and Methods*, *41*(10): 1839-1856 (2012).
8. Liu, J., Y. Shu., L. Zhang., F. Xue, & O.W. Yang, Traffic modeling based on FARIMA models, *Electrical and Computer Engineering, 1999 IEEE Canadian Conference(IEEE).*1:162-167 (1999).
9. Gourieroux, C., & J. Jasiaky, Truncated maximum likelihood, goodness of fit tests and tail analysis, *Quantification and Simulation of Economic Processes,* 36: (1998).
10. Barunik, J., & L. Kristoufek, On Hurst exponent estimation under heavy-tailed distributions. *Physica A: Statistical Mechanics and its Applications*, *389*(18): 3844-3855 (2010).
11. Katsev, S., & I. L'Heureux, Are Hurst exponents estimated from short or irregular time series meaningful? *Computers & Geosciences*, *29*(9): 1085-1089(2003).
12. Sun, W., S. Rachev., F. Fabozzi, & P. Kalev,

Long-range dependence and heavy tailedness in modelling trade duration. *Workig Paper, University of Karlsruhe*. (2005).

13. Reisen, V., B. Abraham, & S. Lopes, Estimation of parameters in ARFIMA processes: A simulation study. *Communications in Statistics-Simulation and Computation*. *30*(4): 787-803 (2001).
14. Stanislavsky, A. A., K. Burnecki., M. Magdziarz., A. Weron, & K. Weron, FARIMA modeling of solar flare activity from empirical time series of soft X-ray solar emission. *The Astrophysical Journal*, *693* (2): 1877 (2009).