A STUDY OF LONG-TERM SUNSPOTS AND K-INDEX GEOMETRIC CYCLES USING PROBABILISTIC MODELING

Danish Hassan¹, ², Hamza Khan³, ⁴, ⁵, Muhammad Fahim Akhter², Muhammad Danish Khan⁵, Shaheen Abbas²

1. Department of Applied Sciences, National Textile University Karachi Campus, Karachi Pakistan.
2. Mathematical Sciences Research Center Federal Urdu University of Arts, Sciences and Technology, Karachi, Pakistan.
3. Department of Textile and Clothing, National Textile University Karachi Campus, Karachi Pakistan 74900.
4. Doctoral School of Applied Informatics and Applied Mathematics Óbuda University, Budapest Hungary.
5. Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Karachi, Pakistan.

ABSTRACT

The research work done in this paper comprises the application of different well-known probability distribution models. This includes the understanding of the behavior and dynamics of 24 sunspot cycles with total data. The time-series data sets were selected from 1749 to 2014. To observe the solar activity effects on K-index activity the double cycles from 1932 to 2014 were also incorporated in the study. The comparative study is useful to observe the long-term solar-terrestrial connection. The magnetic field of the sun reverses its polarity after every 11 years of the cycle. So after every 22 years, the north pole becomes again north pole. By using the two well-known tests Kolmogorov-Smirnov (KST) and Anderson-Darling test (ADT) the probability distribution models were obtained for each sunspot cycles and compare. The significant probability models for all the sunspot cycles have been obtained. The fitted probability distribution models on selected data sets may be useful to understand the trend of solar and geomagnetic activity.

KEYWORDS: Sunspot cycles; geomagnetic activity; Kolmogrove-Smirnov test (KST); Anderson-Darling test (ADT).

*Corresponding author: Dr. Hamza Khan (Email: hamzakhan@ntu.edu.pk, Phone: 0092-333-3032982)

1. INTRODUCTION

The study of sunspots and its cycles has been unique importance in understanding the space climate, global climate and solar-terrestrial relationship. The solar cycles are depending on the sunspots and on its variation that is because of polar reversibility on sun. the quiet and active durations are also strongly connected with the number of sunspots [1, 15]. The other phenomena on sun like Plages, flares, solar wind and prominences are also depending on sunspots activity [14]. The astrophysics provides the information about the solar-terrestrial connection in the perspective of magnetic field of sun, electromagnetic radiation and emission of different particles. The composition of different gases on sun are H (60%), He (40%) and heavier particles (1%). The cumulative mass is $2 \times 10^{33}$ grams with $5 \times 10^9$ years of age. The earth is $1.5 \times 10^8$ Km away from the sun with a nearly colloidal magnetic field [1, 4, 7]. According to the Lorentz force equation perpendicular to both the direction of the velocity vector and the field (the cross product), the K-index of geomagnetic activity
in the perspective of the climate of the earth has been one of the important indices to experience the variation. It categorizes from 0-9 (very quiet to extremely disturb) scale on a scale [10]. The data from any observatory represents the measured at three hours of intervals. The index depicts the third of a unit (5− is 4 2/3, So is 5 0/3, and 5 + is 5 1/3) [4, 7]. The average of activity K-index is the Kr index that is determined and standardized by the data of all station. A storm is labeled a major storm if it is equivalent or surpasses forty [11]. In 1932 to 1989 more than thousands of such storms were happened. A distinct index is sustained by the number of storms which by Allen in 1978 [5,10] has been presented and recognized as Ap index. The independent classification has the C-figures arithmetic mean in every day magnetic activity by separate observatory on a scale of 0 to 22 (quiet-storm). The average of the dimension of two antipodean observatories is entitled aa index. Individuals two antipodal observatories are the Greenwich-Melbourne characterizes the magnetic activity on earth given in Gammas [5].

2. METHODOLOGIES
This manuscript analyzes the sunspots-K-index geomagnetic cyclic activity time series data in the perception of probability models and tail of sunspot cycles depending on the two tests KST and ADT. The tests are substantial to recognize the best fitted probability distribution. For the probability distribution models, the data of sunspot cycles 1749-2014 were incorporated while for assessment of the sunspots-geomagnetic relationship double-cyclic time series 82 years data 1932-2014 were selected. For these purposes a most suitable statistical software Easyfit (EF) was experienced.

3. DETAILS OF MATERIAL AND METHODS
The sun comprises numerous kinds of cycles, many features of the passing solar atmosphere and lots of assets of the quiet as well as the active sun are linked with sunspots. The number of sunspots and latitude foundation is considered almost periodic, again over an 11-year cycle [11, 12]. A specific sunspot remains at persistent latitude in its lifetime; also flourishing sunspots tend to format increasingly worse altitudes, moving from the poles towards the equator. As the end of sunspots in one cycle disappear near the sun’s equator, a new cycle starts near ± 40° (North and South) of the equator. The major number of sunspot maximum arises at transitional latitudes. The solar-terrestrial events are Wolf sunspot numbers well-defined as R = K(f + 10g) where f is the whole number of spots perceived, g is the number of concerned regions and K is a constant of observatory, associated to the sensitivity of the detecting instrument. There are additional important indices for occurrence the number of solar flares observed through one solar variation and the flux of 10 cm radio releases. Jointly these indices in turn narrate to the sunspot numbers [4, 5]. Solar singularities like flares, Plages and Prominences with connection to sunspot activity. In some cases, Sunspot-flares or dissimilarities are detected during a period of small solar activity. Consequently, it is understood that the quiet solar corona can be normally further extendable at the equator as compare to the poles [4]. In an 11-year cycle, the sunspot will be for all time that have the same deviation in one hemisphere. A north pole in the physical northern hemisphere, nevertheless the main sunspot in the further
hemisphere will have the reverse polarity [5]. When a large collection of sprawling spots exists then it creates a twisted magnetic field shape, a fundamentally bipolar field is existent. The polarizations will additionally be reserved in time of the forthcoming cycle of 11 years. In the north hemisphere the sunspot will prime with a magnetic polarity in south. In overall, the dipole park of the sun will variation with the intent that the north-pole of the sun will control from the geographic north-pole to the geographic south-pole. At minimum sunspot time, where the first sunspot is begins to form at higher latitudes. It is noticed that the change of polarity occurs always. For this purpose, the sun is known to have a cycle of 22 year [3, 5]. In this Manuscript, the behavior of 24 sunspots cycles have been investigated 1749-2014 with the 4 K-index cycles 1932-2014. The performance of the selected two-time series data was compared during eight last sunspot cycles 1932-2014. It is also known that 4 K-index cycles associated to these eight-sunspot double-cycles in the duration. Depending on two tests Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) the appropriate probability models were obtained for two data sets. The distribution log-gamma along with parameters are given as under.

\[
f(x) = \frac{\ln[x - y + 1]^{a-1} (x - y + 1)^{1+\beta}}{\beta a \gamma(a)}
\]

(1)

Here \(x > y, \ a > 0, \ \beta > 0\) and the mean of log-gamma distribution is \((1 - \beta) - a + y - 1\). The unique example of the gamma distribution is the chi-square distribution. There are two parameters contains in the family of gamma distribution. The random variable \(x\) of continuous type has the Probability Distribution Function (PDF) contains parameters \(a = \frac{\gamma}{2}\) and \(\beta = 2\), where \(\gamma\) is depicts the positive integer [6].

The obtained models may be the most applicable distribution for the sunspot and K-index cycles. Since the probability distribution helps to realize the potential future variation of the variables. The aim of this research work was to Test the behavior of all the sunspot cycles using probability distribution. It is obvious that 4 K-index cycles correspond to these 8 sunspot cycles in the duration [5]. The normal distribution of continuous variables is the most important distribution fitted on the selected time series. It can be demonstrated mathematically as follows.

\[
f(x) = \frac{1}{\rho \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x - \mu}{\rho})^2}, -\infty < x < \infty
\]

(3)

Where, \(\mu\) is recognizes as the mean of the sample and \(\rho\) is the standard deviation. Similarly, with parameters \(a\) and \(\beta\), if the random variable \(Y\) depicts the gamma distribution, then the likelihood of \(Y\) can be indicated in the form as.

\[
g(Y) = \frac{\beta^a}{\gamma(a)} y^{a-1} e^{-\beta y}, Y \geq 0, a > 0, \beta > 0
\]

(4)

Where \(a\) and \(\beta\) are the shape and scale parameter respectively. \(E(Y) = \frac{\alpha}{\beta}, \text{Var}(Y) = \frac{\alpha}{\beta^2}\). Moreover, it is normally have noticed that the normal distribution is shifted to gamma distribution and this gamma distribution can be further transformed to log-normal distribution [6]. For the log-normal distribution the mathematical model is expressed as,

\[
h(x) = \frac{1}{x \rho \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\ln x - \mu}{\rho})^2}, -\infty < x < +\infty
\]

(5)

\[
E[X] = e^{\mu + \frac{\sigma^2}{2}} \text{Var}(X), e^{2\mu + \sigma^2 (e^{\sigma^2} - 1)}
\]

(6)
Equation 6 signifies the mean (E[X]) and variance (Var (X)) of experiments respectively [6, 14]. Where parameters μ = δ > 0 and the variable x > 0 belongs to real numbers. The log-normal distribution, has been used as a first estimate sometimes known as Landau distribution, explains the loss of energy of a heavy charged particle by ionization [6, 9]. The normal distribution can be assessed the chi-square distribution and is given by

\[ f(x; n) = \frac{n}{2} \frac{x^2}{2} e^{-\frac{x^2}{2}} \]

The variable x ≥ 0 and positive integer n is the number of degrees of freedom in the chi-square distribution [9]. The hyper secant distribution is like a normal distribution, characterized by parameters mean and standard deviation. The hyper secant distribution is used for those data sets which are symmetric in shape like the normal distribution. The mathematical expression for hyper secant distribution is given by

\[ f(x) = \frac{\tanh(y)}{2\sigma}, \text{ where } y = \frac{\pi}{2\sigma}(x - \mu) \]  

(8)

4. GOODNESS-OF-FIT TESTS

The Kolmogorov-Smirnov test (KS) is one of the tests which assesses the empirical distribution function with the distribution function of the hypothesized distribution. The KS-test does not vary on the grouping of the data. This test reduces those problems which are linked to interval specifications. The main objective of KS-test on any empirical data is to verify the theoretical distribution with the help of expected parameters the data could have originated [6, 9]. However, as the KS-test does not involve any grouping of the data hence it has several benefits over the chi-square test. Furthermore, this test can be used for the exact sample size, which makes it powerful against alternative distributions, more meaning full. This test can be applied for any type of input data excluding the discrete distribution functions, and it does not distinguish tail discrepancies very well, which is one of the limitations of this test [8, 13]. Thus, we subject the constructed model to the KS-test as.

\[ D = \max|F(x) - G(x)| \]  

(9)

Where F(x) and G(x) are pre-determined cumulative and sample cumulative distribution respectively, which is corresponding to the given sample of size n [5]. Regarding the tails analysis of the distributions there is a difference between Anderson-Darling (AD) test and the KS-test. AD test was designed to detect discrepancies in the tails especially [6]. Although the AD test is similar to the KS-test, but it gives more insight information related to the tails of the distribution. The dependent behavior, in this case, can be considered the strength of the number of the periods. Moreover, the AD test can only be utilized for the input data sample which is considered the faults of the test [6, 15]. We apply the AD test [Anderson-Darling AP2P-test] [5, 6], where the hypothesis testing is based on comparing the parameters in the question with the larger values of the statistic.

\[ A^2 = -\frac{1}{n} \sum_i^n (2i - 1)[\ln u_i + \ln(1 - u_{n-i+1})] - n \]  

(10)

where u_i is the value of the theoretical cumulative distribution at the i^{th} largest observation, n being as usual the number of observations [2, 5].

5. RESULTS AND DISCUSSIONS

For the sunspots and K-index cycles the probability distributions were discovered in the first section. In the study data have been used from 1749 to 2014 for sunspot cycles and 4 K-index cycles from 1932 to 2014. The Last cycle of sunspots and K-index are yet not complete, that
is in the development. The probability distributions of time series are assessed in terms of two tests Kolmogorov-Smirnov and Anderson-Darling. Table 2 describes that the 24 sunspot cycles have various probability distributions which are according to the Kolmogrove-Smirnov test they are Normal Distribution (12.5%), Log Gamma Distribution (79.17%), Hyper-scat Distribution (4.17%) and ChiSquare Distribution (4.17%). The same 24 sunspot cycles following the Chi-square distribution (95.83%) and Log-gamma distribution (4.17%) all the probability distributions which are tested according to Anderson-Darling test (see figure 1). Since ADT gives better results for tail analysis as compared to the KS-test, Chi-square distribution among all, the sunspot cycles is 95.83% according to AD test. The results show that 23 out of 24 sunspot cycles represent the Chi-square distribution tested with the help of AD. While testing in KS-test it shows just 4.17% of Chi-square distribution. The comparison of the Log-gamma distribution for both tests KS-test and AD test is almost vice versa to the Chi-square distribution. The Log-gamma distribution is just 4.17% in AD test and 79.17% in KS-test. The evaluation of the two tests depicts that Chi-square distribution is increasing from KS-test to AD test while Log-gamma distribution is decreasing among the sunspot cycles. It represents that the dynamics of sunspot cycles are periodic and regular. The two tests for K-index cycles show that probability distributions vary from a Normal distribution to Chi-square distribution and chi-square distribution to Normal distribution in KS-test and AD test respectively except the last cycle which is in progress. Table 2 depicts the related variations (also see related figure 3). Table 1 and 2 shows that for any probability distributions in KS-test the chi-square distribution mostly occurs when tested according to the AD test. The Chi-square distribution has the greater right tail as compared to the other used probability distribution. Since AD test gives the significant results of the tail analysis. The Chi-square distribution containing the greater length of tail as compared to the others probability distributions. The comparative analysis of two tests KS and AD for 22-year’s sunspots and K-index cycles show that cycle 1 follows the Normal distribution and Chi-square distribution, respectively. The probability distribution for 22-year sunspot-double-cycle 2 represents the Log-gamma distribution and Log-normal distribution respectively for KS and AD tests. The 22-years, sunspot cycle 3 shows the Log-gamma distribution and Gamma distribution for both tests respectively, shown in figure 2. The last cycle of 22-years, sunspot data in progress confirms the Chi-square distribution for both the tests. The probability distribution of sunspot data (1749-2014) shows the Normal distribution and Chi-square distribution under the above both discussed tests respectively, while 22-years (approximately) sunspots and K-index 4 cycles total data (1932-2014) represents that Log-gamma distribution and Log-normal distribution are significant depicted in figure 4.

Table 1. Comparison of probability distributions of sunspots 24 cycles along with the total time series data in the perspective of KST and ADT, in which 24th cycle is in progress (1754-2014)

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Kolmogorov-Smirnov Test (KST)</th>
<th>Anderson-Darling (ADT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distribution</td>
<td>KST value</td>
</tr>
<tr>
<td>1</td>
<td>Normal</td>
<td>0.09932</td>
</tr>
</tbody>
</table>
Table 2. Comparison of probability distributions of sunspots and K-index 22-years cycles along with the total time series data using the KST and ADT (1932-2014)

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Duration</th>
<th>Mean(S C)</th>
<th>Mean(K-index)</th>
<th>KST(SC)</th>
<th>ADT(SC)</th>
<th>KST(K-index)</th>
<th>ADT(K-index)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1932.01-1953.12</td>
<td>62.4</td>
<td>14.583</td>
<td>Normal</td>
<td>Chi-square</td>
<td>Normal</td>
<td>Chi-square</td>
</tr>
<tr>
<td>2</td>
<td>1954.01-1976.06</td>
<td>73.811</td>
<td>14.344</td>
<td>Log-gamma</td>
<td>Log-normal</td>
<td>Chi-square</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>1976.07-1998.09</td>
<td>79.72</td>
<td>15.519</td>
<td>Log-gamma</td>
<td>Gamma</td>
<td>Normal</td>
<td>Chi-square</td>
</tr>
<tr>
<td>4</td>
<td>1996.10-20.00</td>
<td>48.834</td>
<td>11.006</td>
<td>Chi-square</td>
<td>Chi-square</td>
<td>Chi-square</td>
<td>Log-gamma</td>
</tr>
<tr>
<td>1-4</td>
<td>1932.01-20.00</td>
<td>67.446</td>
<td>14.103</td>
<td>Log-gamma</td>
<td>Log-normal</td>
<td>Log-gamma</td>
<td>Log-normal</td>
</tr>
</tbody>
</table>
sunspots cycle 3

KST-ADT Chi-Squared (2P) Log-Gamma

Time (years)

sunspots cycle 4

KST-ADT Log-Gamma

Time (years)

sunspots cycle 5

KST-ADT Normal Chi-Squared (2P)

Time (years)

sunspots cycle 6

KST-ADT Hypersecant Chi-Squared (2P)

Time (years)

sunspots cycle 7

KST-ADT Normal Chi-Squared (2P)

Time (years)

sunspots cycle 8

KST-ADT Chi-Squared (2P) Log-Gamma

Time (years)

sunspots cycle 9

KST-ADT Chi-Squared (2P) Log-Gamma

Time (years)

sunspots cycle 10

KST-ADT Chi-Squared (2P) Log-Gamma

Time (years)
Hassan et al., J. mt. area res. 07 (2022) 44-55

sunspots cycle 11

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1878 1876 1874 1872 1870 1868

No. of sunspots

0.44 0.4 0.36 0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0

sunspots cycle 12

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1888 1886 1884 1882 1880 1878

No. of sunspots

0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0

sunspots cycle 13

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1900 1898 1896 1894 1892 1890

No. of sunspots

0.26 0.24 0.22 0.2 0.18 0.16 0.14 0.12 0.1 0.08 0.06 0.04 0.02 0

sunspots cycle 14

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1912 1910 1908 1906 1904 1902 1900

No. of sunspots

0.3 0.28 0.26 0.24 0.22 0.2 0.18 0.16 0.14 0.12 0.1 0.08 0.06 0.04 0.02 0

sunspots cycle 15

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1924 1922 1920 1918 1916 1914 1912

No. of sunspots

0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0

sunspots cycle 16

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1933 1932 1931 1930 1929 1928 1927 1926 1925 1924

No. of sunspots

0.36 0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0

sunspots cycle 17

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1944 1942 1940 1938 1936 1934

No. of sunspots

0.36 0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0

sunspots cycle 18

KST- ADT Chi-Squared (2P) Log-Gamma

Time (years)

1954 1952 1950 1948 1946 1944

No. of sunspots

0.4 0.36 0.32 0.28 0.24 0.2 0.16 0.12 0.08 0.04 0
Figure 1. Comparison of probability distributions of sunspots 24 cycles in terms of KST and ADT in which 24\textsuperscript{th} cycle is incomplete. The long right tail is indicating for the Chi-square distribution (1754-2014).
Figure 2. Comparison of probability distributions of sunspots 4 cycles (approximately 22-years of each) in the perspective of KST and ADT, the last 4th cycle is in progress. The single line in cycle 2, 3 and 4 is representing the overlapping of two probability distributions. (1932-2014)

Figure 3. Comparison of probability distributions of K-index 4 cycles (approximately 22-years of each) in the perspective of KST and ADT. The last 4th cycle is incomplete. Chi-square distribution is representing the right long tail.
CONCLUSION
The study reveals that the 24 cycles of sunspots (1749-2014) and 4 cycles of sunspots and K-index (1932 to 2014) of 22 years show several distributions which are analyzed according to KS-test and AD test by comparison of each sunspot and K-index double-cycles. Distribution diagrams depict that each cycle is right tail, which confirms the periodicity and consistency of the dynamics involve as well. It is recommended: Since AD test gives the significant result for tail as compared to the KS-test, the results obtained for 24 cycles of sunspots are found to be Chi-square distributed except 4th cycle according to AD test. chi-square distribution has right long tail, that represents the long dynamics for the sunspots and K-index activity in the right tail corresponding to the Anderson-Darling test.

DECLARATIONS
Funding: We do not have any funding for this research.

Conflicts of interest/Competing interests: The authors declare no any conflict of interest.

Data availability: Not applicable.

Code availability: Not applicable.

Authors’ contributions: Danish Hassan and Shaheen Abbas: Conceptualization, Methodology Danish Hassan and Hamza Khan Visualization, Investigation Danish Hassan, Hamza Khan, Fahim Akhter and Muhammad Danish Khan Software, Validation Hamza Khan: Writing-the final version.

REFERENCES


Received: 12 Nov. 2021. Revised/Accepted: 20 April 2022.

This work is licensed under a Creative Commons Attribution 4.0 International License.