International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

Statistical Modeling and Forecasting of EEG Signal for BCI System Using ARIMA Model

Asst. Prof. Dr. Naufel B. Mohammed
Head of the Computer Department, KISSR
Naufel.bahjat@kissr.edu.krd

Abstract

This paper presents a description and building of a statistical model of EEG signals with an ARIMA forecasting process.
EEG measurement principles are explained for understanding EEG signals features and noise which is essential for building real-time Brain-Computer Interface (BCI).
Different statistical modeling and forecasting methods of EEG signals are discussed.

Keywords: Statistical modeling, EEG signals, BCI, Forecasting, Correlation, ARIMA.

Introduction

The human brain is always active. The brain controls the different activities of the body. Brain functions can be monitored by observing the electrical signals generated in the neurons. This signal is called an electroencephalogram (EEG) signal. The signal can be extracted using electrodes and can be viewed using a voltmeter, oscilloscope, or on a computer screen. The EEG signal can be recorded and the phenomenon is known as electroencephalography. The EEG signal can be used to investigate the condition of the human brain and the overall health of the person. These signals are roughly less than 100 µV and 100 Hz and can be measured with electrodes placed on the scalp, non-invasively (Sabbir) [1].

The human brain consists of millions of interconnected neurons.

The patterns of interaction between these neurons are represented as thoughts and emotional states. According to human thoughts, this pattern will be changing which in turn produces different electrical waves. A muscle contraction will also generate a unique electrical signal. All these electrical waves will be sensed by the brain wave sensor and it will convert the data into packets and transmit through Bluetooth medium. The basic idea of BCI is to translate user-produced patterns of brain activity into corresponding commands. A typical BCI is composed of signal acquisition and signal processing (including preprocessing, feature extraction, and classification) (Siliveru) [2].

A brain-computer interface (BCI) controlled lower extremity prosthesis may be one such novel approach. It can be envisioned that a combination of an invasive brain signal acquisition system and implantable functional electrical stimulation (FES) electrodes can potentially act as a

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

permanent BCI prosthesis. However, for safety reasons, the feasibility of brain-controlled ambulation must first be established using noninvasive systems (An. H. Do.) [3].

The use of brainwaves to control robotic devices has produced promising clinical results in terms of feasibility. Restoration of a certain degree of motor functions and high accuracy control of robotic prosthetic arms using invasive BCIs have already been demonstrated. Nevertheless, in order for such BCI-controlled robotic applications to achieve end-user maturity, the use of noninvasive, portable, and relatively low-cost systems is considered a required development (Alexander) [4].

Almost all bio signals are nonstationary too, meaning their statistics may change in time, hence time-dependent. In contrast to deterministic signals that are rigidly periodic, stochastic biosignals are difficult to be modeled precisely by mathematical functions due to uncertainty in the parameters of the computational equations.

Descriptive statistics can give an overall summary of the time series, and characterize a general shape, they may not be able to capture the precise trend movements aka the patterns of evolving lines. In particular, we are interested in distinguishing the abnormal time series from those that are normal. The difference in movements of the trends is estimated by Dynamic Time Wrap (DTW). DTW algorithm aims at finding the degree of matching between the alignments of two-time series. The normal group of time series is represented by the mean values of the points from all the time series. A particular time series under test is then compared to the mean sequence of points, and the deviation or extent of matching/similarity is then computed by the DTW algorithm (Simon) [5].

Understanding EEG noise is essential for EEG-based neurology and applications such as real-time brain-computer interfaces (BCIs), which must make accurate control decisions from very short data epochs.

Improving the performance of real-time neurological algorithms: Certain neurological signal processing tasks, such as extracting event-related potentials (ERPs), increase signal-to-noise ratios by averaging many epochs of data recorded over long experimental periods (Alan Paris) [6].

Two types of data were recorded: discrete emotional states and the responses to complex emotional scenarios. Typical trial times of 60 and 300 seconds were used for each type of data, respectively. A fifteen-minute break was taken between each trial so that the subject could return to her baseline, emotionally relaxed state (Mitchel) [7].

However, due to large subject variability in the studies, the time domains of these measurements are not the same. This random temporal variability, if not accounted for, can cloud the statistical analysis. Furthermore, many of the functions of interest are approximately periodic and sampled at a very high resolution. Direct analysis of these long biosignals is often computationally inefficient and prone to being affected by noise. Thus, a common approach is to summarize such biosignals using cross-sectional statistics such as the mean, median, and variance, and then use standard statistical tests as diagnostic tools (Sebastian) [8].

1. ARIMA Model:

The main part of the ARIMA model combines AR and MA polynomials into a complex polynomial. The ARIMA (p, d, q) model is applied to all the data points of the EEG data.

$$y_t = \mu + \sum_{i=1}^{p} (\sigma y_{t-1}) + \sum_{i-1}^{q} (\theta \varepsilon_{t-1}) + \varepsilon_t \qquad (1)$$

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

where the notation is as follows:
µ: the mean value of the time series data;
p: the number of autoregressive lags;
σ: autoregressive coefficients (AR);
q: the number of lags of the moving average process;
ϴ: moving average coefficients (MA);
Ɛ: the white noise of the time series data;
d: the number of differences calculated from (2)

$$Dy_t = y_i - y_{i-1} \quad (2)$$

The value of the ARIMA parameters (p, d, q) for AR and MA can be obtained from the behavior of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) [1]. These functions help to estimate the parameters that can be used to forecast data using the ARIMA model (Hussan) [9].

2. Autocorrelation:

One of the basic assumptions in the linear regression model is that the random error components or disturbances are identically and independently distributed. So, in the model y = X β + u, it is assumed that:

$$E\,(U_r, U_{i-s}) = \begin{cases} \overset{2}{\underset{u}{\sigma}} & if\ s = 0 \\ 0 & if\ s \neq 0 \end{cases} \quad (3)$$

i.e., the correlation between the successive disturbances is zero.

In this assumption, when $E\,(U_r, U_{i-s}) = \overset{2}{\underset{u}{\sigma}}, s = 0$ is violated, i.e., the variance of disturbance term does not remain constant, then the problem of heteroskedasticity arises. When $E\,(U_r, U_{i-s}) = 0,\ s \neq 0$ is violated, i.e., disturbance terms are correlated, then such a problem is termed the problem of autocorrelation (Shalabh) [10].

4. Data collection:

To measure and record EEG signals, A volunteer female was chosen to be under the test, A quiet lab environment condition, and a relaxing period before starting the test. The block diagram of the EEG measurement system is shown in fig. (1).

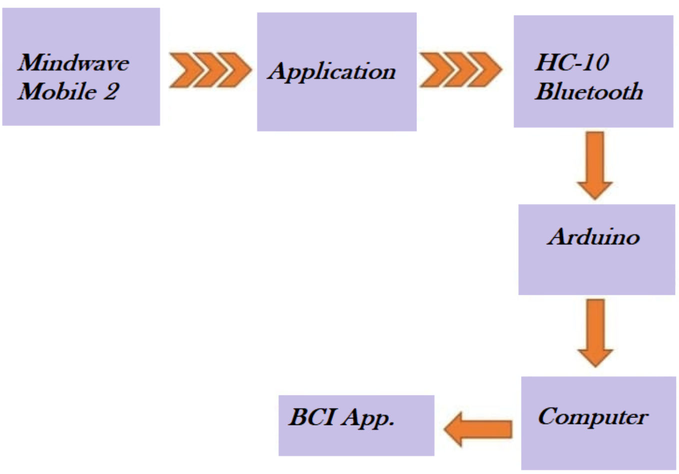

fig. (1) Block diagram for EEG measurement system

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

A Mindwave mobile2 shown in fig. (2) was used to measure the different EEG signals patterns and pass them to the Arduino ATmega via Bluetooth HC-06, the EEG signals were supplied to a computer to analyze the data, the time slots of measurements were done every 30 minutes for 5-minute data collection.

The statistical modeling and the forecasting process are done by using computer software (Wessa P.) [11].

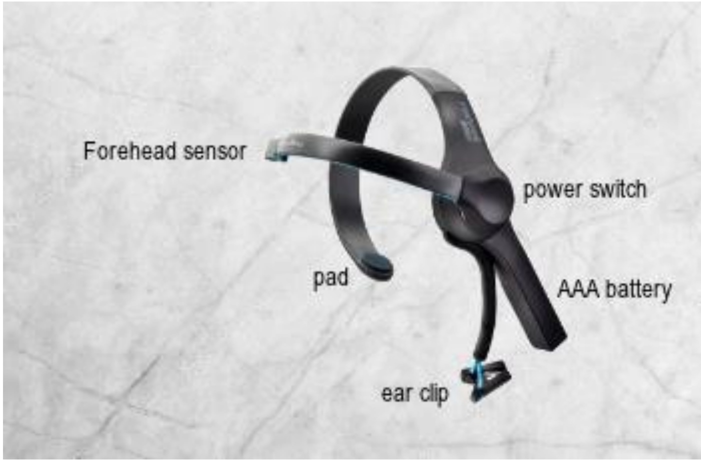

Fig. 2 Mindwave mobile 2 (Neurosky, Inc.)

5. EEG results:

The EEG signal measured is shown in fig. (3), this signal will be considered as time series data, and an ARIMA (p, d, q) model was calculated for the signal.

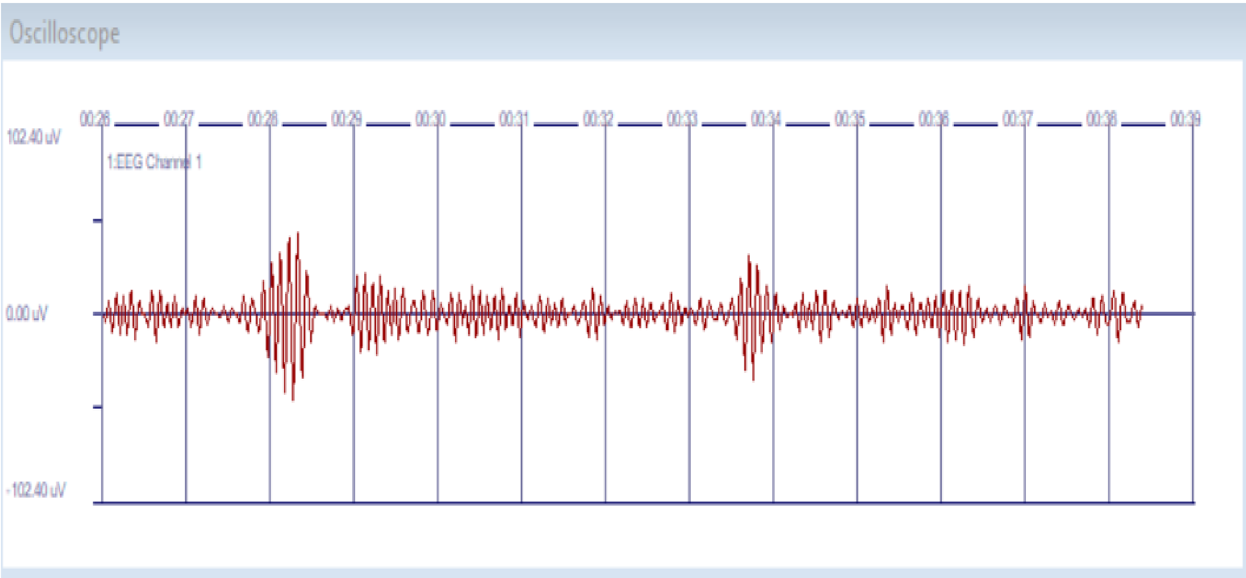

Fig. (3) measured EEG raw signal

raw EEG signal was sampled and the frequencies and their amplitudes were extracted using the FFT process as shown in fig. (4), the spectrum shows the most power amplitudes are of the lower frequencies.

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

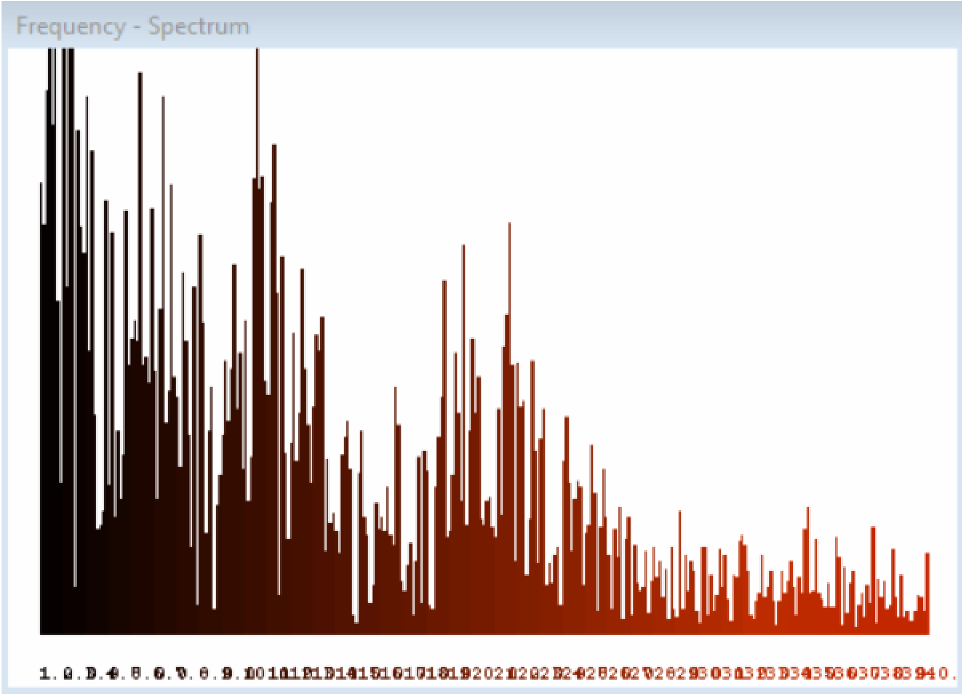

Fig. (4) The frequencies and amplitudes of the measured EEG signal

The 3D plot of the raw EEG signal can give a better understanding of the signal components and their variation and the type of noise that we had with the measurements; fig. (5) shows the details of the measured EEG signal.

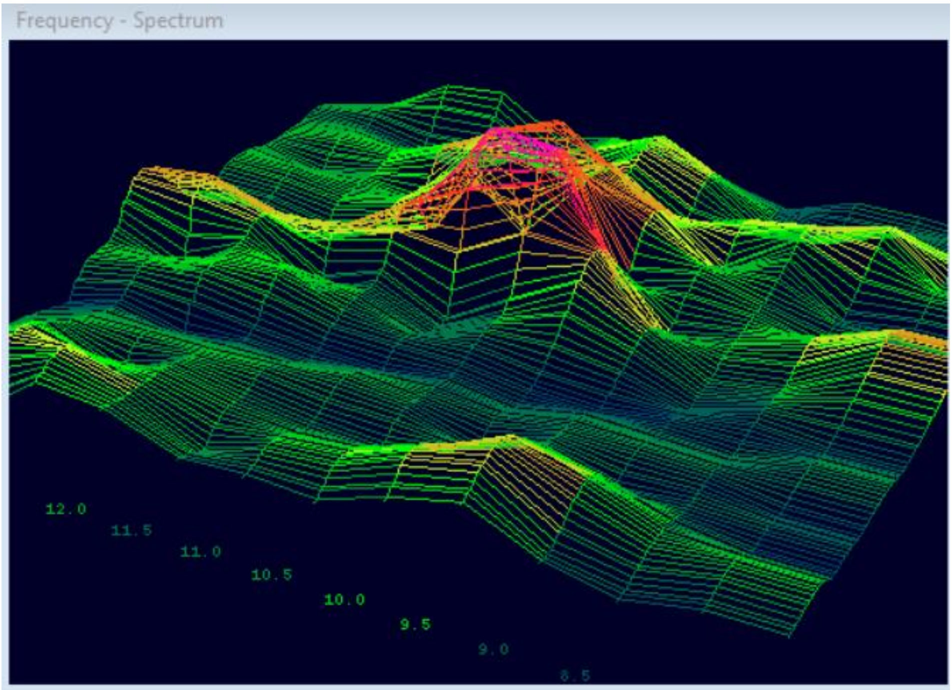

Fig. (5) 3D plot of the measured EEG signal

The extracted amplitudes were used to build a time series as shown in fig. (6).

The range of the measured amplitudes varied from 3-57 microvolts for this time slot, from this time series we can decide the seasonality and the forecasting parameters for the ARIMA forecasting model.

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

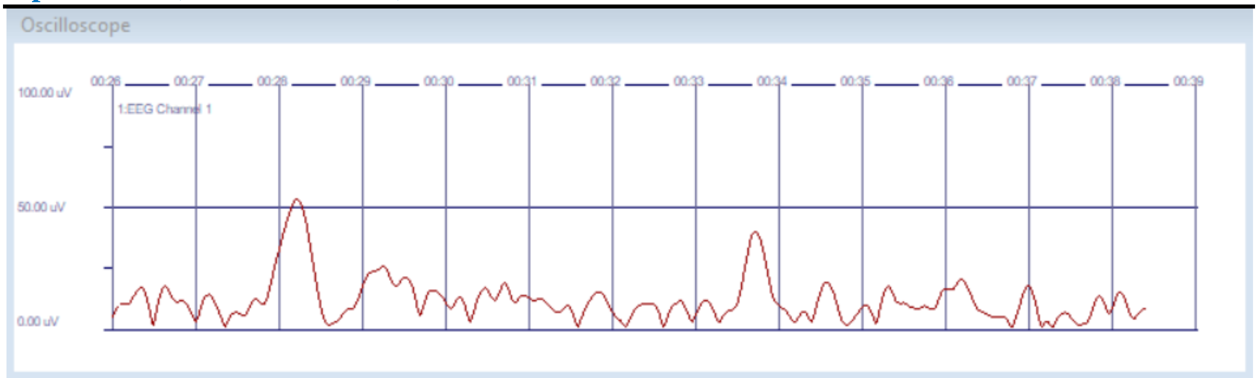

Fig. (6) EEG time series built from the Raw EEG signal.

6. Statistical model:

The measured EEG signals transformed into time series with (112) observations, then an ARIMA model was built, table (1) shows the calculated forecasted results, an accuracy of 95% reaches.

Table (1) ARIMA forecasting results

Univariate ARIMA Extrapolation Forecast

time	Y[t]	F[t]	95% LB	95% UB	p-value (H0: Y[t] = F[t])	P(F[t]>Y[t-1])	P(F[t]>Y[t-s])	P(F[t]>Y[107])
106	7	-	-	-	-	-	-	-
107	4	-	-	-	-	-	-	-
108	12	4.5444	0.9222	17.9759	0.1383	0.5317	0.5317	0.5317
109	10	4.1704	0.6353	20.3098	0.2395	0.1708	0.1708	0.5083
110	14	4.0785	0.4258	25.7271	0.1845	0.2959	0.2959	0.5028
111	7	3.869	0.2881	30.3578	0.4084	0.2267	0.2267	0.4961
112	9	3.7173	0.1945	36.1676	0.3748	0.4214	0.4214	0.4932

Often, measuring forecasting accuracy (or error) is difficult, for that only by experiments we can decide which can be used to evaluate the accuracy of the forecasting process, table (2) shows the (KPI) calculation.

Table (2) KPI calculation results

Univariate ARIMA Extrapolation Forecast Performance

time	% S.E.	PE	MAPE	sMAPE	Sq.E	MSE	RMSE	ScaledE	MASE
108	1.508	0.6213	0.6213	0.9013	55.5867	0	0	1.9882	1.9882
109	1.9745	0.583	0.6021	0.862	33.984	44.7853	6.6922	1.5546	1.7714
110	2.7082	0.7087	0.6376	0.9406	98.4367	62.6691	7.9164	2.6457	2.0628
111	3.4931	0.4473	0.5901	0.8495	9.8033	49.4527	7.0323	0.8349	1.7559
112	4.4538	0.587	0.5894	0.8457	27.9066	45.1434	6.7189	1.4087	1.6864

The forecasting plot of the EEG time series is shown in fig. (7), the resultant plot shows good similarity with the original EEG signals.

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

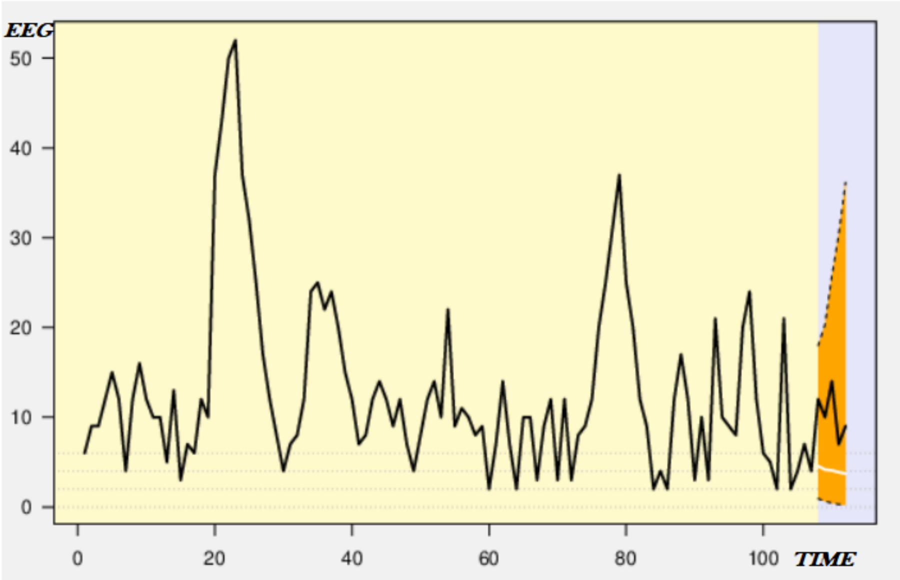

fig. (7) plot of the forecasted EEG signal.

The frequency table of the EEG time series is shown in fig. (8)

Table (3) Frequency distribution of the EEG time series.

Frequency Table (Histogram)

Bins	Midpoint	Abs. Frequency	Rel. Frequency	Cumul. Rel. Freq.	Density
[0,5[	2.5	20	0.178571	0.178571	0.035714
[5,10[	7.5	38	0.339286	0.517857	0.067857
[10,15[	12.5	28	0.25	0.767857	0.05
[15,20[	17.5	7	0.0625	0.830357	0.0125
[20,25[	22.5	11	0.098214	0.928571	0.019643
[25,30[	27.5	0	0	0.928571	0
[30,35[	32.5	2	0.017857	0.946429	0.003571
[35,40[	37.5	3	0.026786	0.973214	0.005357
[40,45[	42.5	1	0.008929	0.982143	0.001786
[45,50[	47.5	1	0.008929	0.991071	0.001786
[50,55]	52.5	1	0.008929	1	0.001786

The result of the frequency distribution is plot as a histogram as shown in fig. (9).

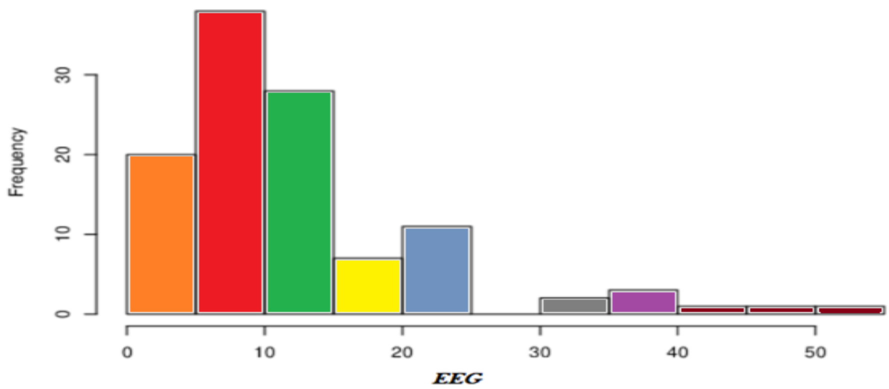

fig. (9) EEG signal histogram

International Journal of Scientific Trends- (IJST)
ISSN: 2980-4299
Volume 2, Issue 3, March, 2023
Website: https://scientifictrends.org/index.php/ijst
Open Access, Peer Reviewed, Scientific Journal

Extrapolation used to infer the unknown data of the EEG signal from previous known EEG data, or in another words the prediction data, fig (10) the ARIMA extrapolated EEG signals.

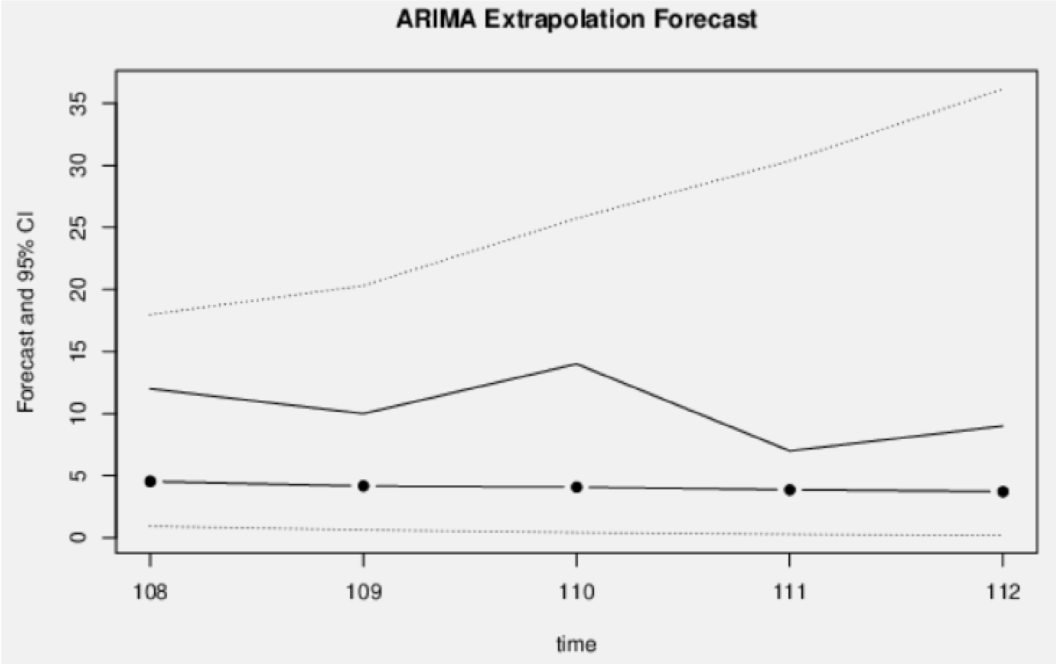

Fig. (10) ARIMA Extrapolation forecast.

To study the similarity between the original EEG signal and the forecasted one, cross correlation function (CCF) applied, the result shown in fig. (11),

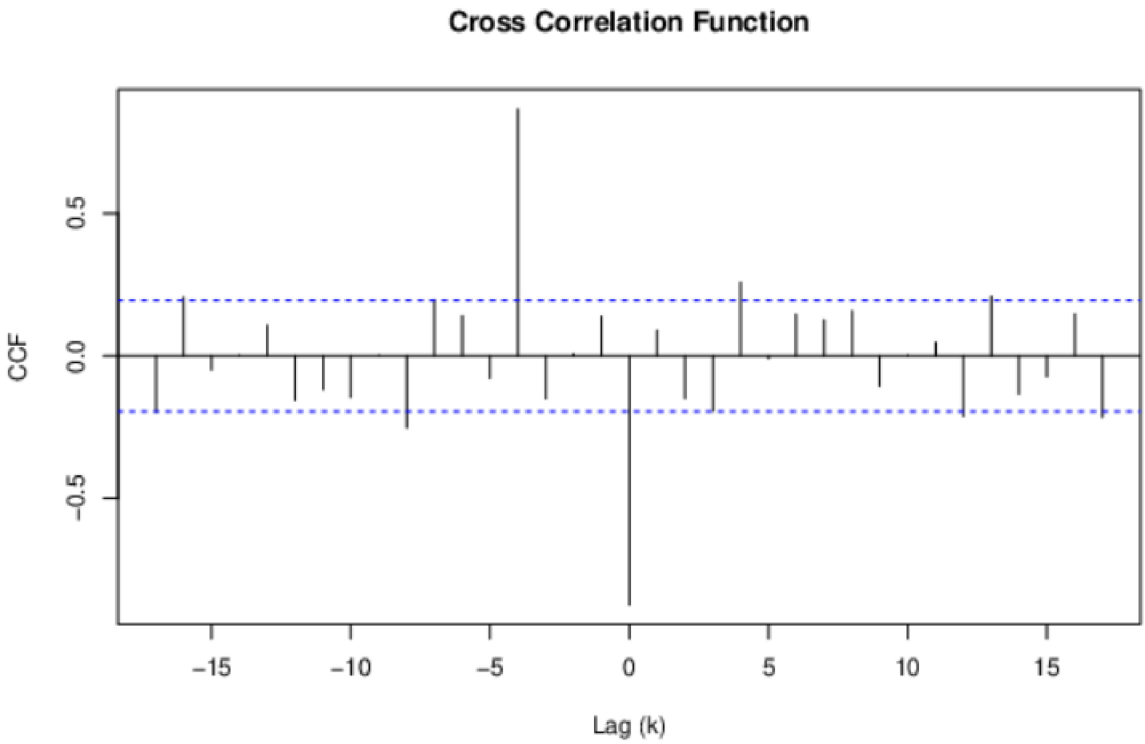

Fig. (11) plot of the Cross-Correlation Function applied to EEG signal

Conclusion:

This paper presents a statistical model of EEG signals using ARIMA for the BCI system.

A review of related research was discussed, then measurements of EEG signals were done in a suitable environment's lab conditions, the measured EEG signal was filtered and then transformed into a time series EEG data.

FFT analysis was carried out to collect data information required to build the ARIMA model, a forecasting process applied using this model.

The forecasting shows a good prediction result with an efficiency of 95%, to evaluate the accuracy of the prediction results a Key Performance Indicator (KPI)

Used.

A cross-correlation function was used to investigate the similarity between the original and the transformed EEG signals.

References:

1. Sabbir Ibn Arman," Cost- Effective EEG signal Acquisition and Recording System", International Journal of Bioscience, Biochemistry, and Bioinformatics, Vol,2, No. 5, Sep. 2012.
2. Siliveru, et.," Brainwave Controlled Robot Using Bluetooth", International Journal of Advanced Research in Electrical, Electronics, and Instrumentation Engineering, Vol.3, Issue 8, Aug. 2014.
3. An. H. Do.," Brain-Computer Interface Controlled Robotics Gait Orthosis", *Journal of Neuro Engineering and Rehabilitation* volume 10, Article number: 111 (2013)
4. Alexander Astaras, et. "Towards Brain-Computer Interface Control of a 6-Degree-of-Freedom Robotic Arm Using Dry EEG Electrodes", Hindawi Publishing Corporation, Advances in Human-Computer Interaction, Article ID 64107, 6 pages, 2013.
5. Simon Fong, et., "A TIME-SERIES PRE-PROCESSING METHODOLOGY FOR BIOSIGNAL CLASSIFICATION USING STATISTICAL FEATURE EXTRACTION", proceeding of the IASTED international conference, Austria, Feb. 2013.
6. Alan Paris, et., "A new statistical model of electroencephalogram noise spectra for real-time brain-computer interface", IEEE transaction bio-medical engineering, Jul. 2016.
7. Mitchel b., et., "BIOSIGNAL ANALYSIS AND ITS APPLICATION IN A PERFORMANCE SYSTEM", International Journal of Engineering Research & Technology, Vol.2, Issue 2, 2014.
8. Sebstian Kurtek, et.," Segmentation, Alignment and Statistical analysis of biosignal with application to disease classification", Journal of applied Statistics, Vol. 40, No. 6, 2013.
9. Hussan Al-chalabi, et.," Time series Forecasting Using ARIMA Model", ADVCOMP, 2015.
10. Shalabh," Linear Regration Analysis and Forecasting", IIT, Kanpur, Jan. 2017.
11. Wessa P., (2021), ARIMA Forecasting (v1.0.11) in Free Statistics Software (v1.2.1), Office for Research Development and Education, URL http://www.wessa.net/rwasp_arimaforecasting.wasp/