

Modeling and Forecasting of Sunspots Cycles: An Application of ARMA (p, q)-GARCH (1, 1) Model

Asma Zaffar (✉ asmazaffar105@gmail.com)

Sir Syed University of Engineering and Technology

Research Article

Keywords: ARMA (p, p) – GARCH (1, 1) process, stationary, Langrage multiplier, Root Mean Square Error, Skewness, Kurtosis

Posted Date: April 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-412946/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The influence of the earth climatic condition of oscillations of solar activity is measurable only in the long run duration. Modeling of the sunspots is an initial role for the mankind utilization of the benefits because solar activity has influenced on the earth's climates. Time series analysis and modeling have proved to stick out amidst with other statistical tools when estimating and predicting solar activities. This study emphasis on the appropriateness of the generalized autoregressive conditional heteroskedasticity (GARCH) models with specification autoregressive ARMA (p, q) process in terms of their performance for delivering volatility forecasts for Sunspot cycles. In this study, individual sunspots cycle's ranging from cycle 1st to 24th (1755–2019) are considered. To notice the appropriateness of Autoregressive Conditional Heteroscedastic (ARCH) effect on sunspot cycles data, Lagrange Multiplier test is used. ARMA (p, q)-GARCH (1, 1) process expresses leptokurtic that is fat and heavy tail (values are strongly correlated to each other). The sunspot cycles ARMA (p, q)-GRACH (1, 1) process expresses the positive skewness except cycles 4th and 19th. Most of the Sunspot cycles (1st, 4th, 12th, 13th, 14th, 15th, 16th, 19th, 20th, 23rd and 24th) follow Auto-Regressive and moving Average (ARMA (2, 2))-GARCH (1, 1). Sunspot cycles (5th, 6th, 7th and 15th) follow ARMA (3, 3)-GARCH model. Whereas the cycles (2nd and 11th) show appropriate model is ARMA (5, 1) - GARCH (1, 1) process. ARMA (5, 3) -GARCH (1, 1) process expresses cycles (18th and 19th). The ARMA (2, 2)-GARCH (1, 1) stationary volatility model expresses the finest forecasting model as compared with other models. Though, ARMA (2, 2)-GARCH (1, 1) is the adequate model for estimation and forecasting most of the sunspot cycles. The results that are obtained by this study are very beneficial for observing the influence of solar activity on the earth's climate.

1. Introduction

Different layers of the sun spin at different rates, creating a magnetic field for the solar sphere. Convection currents create local magnetic fields in hot gas bubbles. Larger local magnetic fields and bubbles rise to the surface. At the surface, north and south polarity is split into pairs of disturbances. Large pairs usually create sunspots. Large sunspot groups often create flares and mass coronal ejections. Solar activity is established via spots dark on the Sun surface which is called Sunspots. The counting of Sunspots changes from time to time. Approximately, sunspots have 11-year cycles (Muraközy and Ludmány 2012). The solar cycle effect on the activity changes in the sun, solar material ejection and the solar radiations level. The solar cycle appearance depends on the variations in sunspot numbers, flares, and other manifestations.

Time series are very essential for various solar physics disciplines. The study which belongs to the climate change study also goes to the area alike. After eliminating trend and periodicities from a time series, the components of stochastic endure there. The long range correlation recommends the positive autocorrelations presence that continue significantly high over large time lags, so as to the autocorrelation function of the series demonstrate a slow asymptotic decay. The persistence or strength of the long-range correlations constrained in experimental time series can be evaluated by various well-known methods (Box and Jenkins 1994). The involvement of time series in solar physics frequently reveals persistence, where sequential values are positively correlated. In statistical analysis, huge data is a way to associate the trends

Loading [MathJax]/jax/output/CommonHTML/jax.js study solar activities, we have certain the sunspot cycle (from

cycle 1 to cycle 24) individual data and total sunspot cycles data (1755 to 2019). Each sunspots cycle data has long term trend. The prediction and correlation of large time series data has long term trend behavior. Whereas small data has short term trend behavior.

In the conditional Heteroscedastic process, an autoregressive model is used. It can follow because of the presence of outliers (very small or very large). The GARCH model (Goh and Khor C 2016) is one of the most advanced statistical techniques which is applied in volatility. It is used to analyze forecast volatility. GARCH model is a variance model and used to forecast the variance of the forthcoming period as a weighted average of the long-term average variance. GARCH model is forecast just a single period, it turns out that absorbed based on one period forecast due to second period forecast can be made (Bollerslev and Engle 1994). GARCH model is mean reverting and conditionally heteroskedastic in which unconditional constant variance are involved (Engle 2001). ARMA model is strongly significant to volatility modeling. ARMA methods are frequently used and most popular in time series models compared to other models like Markov Chains, Artificial Neural Network Models, Fuzzy networks, etc (McKenzie 1984). The ARMA models have a flexible nature. Thus, it can be used in numerous types of time series with different orders. It compromises regular extensiveness at individual phases (identification, estimation, and diagnostic checks) for an appropriate model. In these models, one of the highest difficulties is the essential for large data (W. Ji and K. Chee 2011). A large amount of literature has been explored by using GARCH (1, 1) model (Engle 2001, Salisu and Fasanya 2013, Epaphra 2017, Pham and Yang 2010). All of these literatures reported that GARCH (1, 1) is more appropriate in analyzing time series data. It is the simplest and strongest among volatility models (Engle 1982) and fit various data series as well (Hill, Griffiths, and Lim 2011). GARCH (1, 1) is adequate to capture the volatility clustering in the data (Brooks 2014). Moreover, (Olson and Wu 2017) revealed that analysis can be sufficient with only one lag for each variable. Furthermore, GARCH (1, 1) is leptokurtic (a process having a kurtosis value greater than 3). The generalized autoregressive conditional heteroscedastic (GARCH) models can relate to ARMA models. Residual diagnostic checking like ARCH LM, normality test and correlogram squared residuals found the selection of the adequate model. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are confirmed by forecasting evolutions. Furthermore, the Akaike information criterion (AIC), Bayesian Schwarz information criterion (BIC) and Hannan Quinn information criterion (HIC) values are also calculated. The best-fitted model residuals are selected by diagnostic checking. Forecasting evolution of each sunspot cycle is calculated via the normality test, which is based on Skewness, Kurtosis and of Jurque-Bera statistic tests. ARMA (p, q)-GARCH (1, 1) model of sunspot cycles also verified the presence of leptokurtic except cycles 2, 7, 18 and 19 which are platykurtic flat tail (kurtosis value less than 3). The sunspot cycles GRACH (1, 1) follow positive skewness except cycles 4 and 19. These two cycles show negative skewness.

2. Data Description And Methodology

The data of sunspots cycles from 1755 to 2019 (1–24) is the mean monthly under deliberation. The data is collected from the World Data Centre (WDC). The main emphasis is on the Box-Jenkins method for the stationary process of ARMA-GRACH. The adequate models ARMA (p, q)-GRACH (1, 1) are selected by Akaike information criterion (AIC). Bavesian Schwarz information criterion (BIC) and Hannan Quinn

Loading [MathJax]/jax/output/CommonHTML/jax.js

information criterion (HIC). The forecasting ability of each model of sunspot cycles will be judged by diagnostic checking tests like Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Mean maximum likelihood estimation is used to evaluate ARMA (p, q)–GARCH (1, 1) model. The Statistical EViews version 9.0 software is used for calculation and analysis of ARMA (p, q) –GARCH (1, 1) model and respective graphs. For instance, time series plots and fitted, residual and forecasted plots for total sunspot cycles. This section consists of two subsections.

2.1: Basic equations of statistical analysis

This section consists of short statistical analysis.

2.1.1: Diagnostic Test

Lagrange multipliers (LM) are used to check the ARCH effect of the existing data. Correlogram squared residual test is also used to confirm the ARCH effect in the time series data. In addition, the usual normality test is also executed for the verification of the utilization of GARCH. Root mean square error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are calculated to verify the accuracy of the forecasts. The Gaussian quasi-maximum likelihood estimation (GQMLE) is executed to indicate the fitting models. The GQMLE is normally used in GARCH models for keeping the heavy-tailed returns (Bollerslev, Engle and Nelson 1994, Thomas. and Denial 2002). The Gaussian quasi-maximum Likelihood Estimator (GQMLE) is almost normally distributed with a variance which is at least as lesser as those of other asymptotically normally distributed estimators. GQMLE constantly produces consistent estimates of the parameters of appropriately specified conditional mean. The adequacy of selected models is verified by the Akaike information criterion (AIC), Bayesian Schwarz information criterion (BIC) and Hannan Quinn information criterion (HIC). Forecasts with the best-fitted model of sunspot cycles were tested for accuracy with the help of a Root mean square error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). A description of these terminologies is given in the following.

Akaike information criterion: The AIC test was introduced by Hirotogu Akaike in 1973. It is the extension of the maximum likelihood principle. The selection criterion is focused on the least value of AIC.

$$AIC = -2\text{Log (likelihood)} + 2S \quad (1)$$

Where S is the model parameter numbers. The likelihood is a measure of the fit model. Maximum values exhibit the best fit.

Schwarz criterion: The SIC test is used to select the most appropriate model among finite models. The appropriate model is based on the least value of SIC. Schwarz criterion (SIC) was developed by Gideon E. Schwarz. It is closely related to the AIC.

$$SIC = -2\ln (\text{Likelihood}) + (S + S \ln (N)) \quad (2)$$

Where S is the model parameter numbers. N exhibits the number of observations.

Hannan-Quinn criterion: The HQC is the criterion for model selection. This test is an alternative to AIC and SIC.

$$\text{HQC} = -2 \log(\text{Likelihood}) + 2(S + S \ln(N)) \quad (3)$$

Where S is the model parameter numbers. N exhibits the number of observations.

Durbin-Watson Test: The DW statistics is a test for measuring the linear association between the adjacent residual from a regression model. The hypothesis of Durbin-Watson statistics is $\rho = 0$ is the specification.

$$U_t = \tau U_{t-1} + \epsilon_t \quad (4)$$

Durbin-Watson (DW) is equal to 2 shows there is no serial correlation. If Durbin-Watson (DW) is less than 2 indicate that positive correlation and the range from 2 to 4 represents that negative correlation. The series is strongly correlated if the value nearly approaches to zero.

Mean Absolute Error: The mean absolute error is expressed as a mathematically formed.

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (5)$$

Where n is the number of observations. Mean Absolute Error (MAE) processes the absolute deviation of forecasted values from real ones. It is also called Mean Absolute Deviation (MAD). It expresses the magnitude of overall error caused by forecasting. MAE does not cancel out the effect of positive and negative errors. MAE does not definite the directions of errors. It should be as small as possible for good forecasting. MAE depends on the data transformations and the scale of measurement. Extreme forecast error does not exist by MAE (Adhikari R. 2013).

Mean Absolute Percentage Error (MAPE): The Mean Absolute Percentage Error (MAPE) is defined as

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{X_t} \right| \times 100 \quad (6)$$

Mean Absolute Percentage Error (MAPE) provides the percentage of the average absolute error. It is independent of the scale measurement. MAPE does not locate the direction of Error. The extreme deviation is not penalized by MAPE. In this measure, opposite signed errors do not offset each other in MAPE [1]. This means that due to the benefits of freedom and commentary on the absolute percentage error (MAPE) scale, one of the most extensively used measures of prediction accuracy. Whereas it is independent of the scale of measurement but affected by data transformations (Schwabe H. 1844).

Root Mean Squared Error (RMSE): The root mean squared error (RMSE) is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (7)$$

RMSE calculates the average squared deviation of forecasted values. The opposite signed errors do not offset one another. RMSE provides the complete idea of the error that happened during forecasting. By using the accuracy measures, errors that are small and are getting good, such as 0.1 RMSE and 1% MAPE, can often be achieved. In RMSE, the total forecast error is affected by the large individual error. For example, a large error is much more expensive than small errors. It does not reveal the direction of overall errors. RMSE is affected by the data transformation and the change of scale. RMSE is a good measure of overall forecast error (Adhikari R. 2013).

Theil's U-Statistics (U): Theil's U-Statistics is defined as

$$U = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n \epsilon_t^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n f_t^2} \sqrt{\frac{1}{n} \sum_{t=1}^n X_t^2}} \quad 0 \leq U \leq 1 \quad (8)$$

Where f_t represent the forecasted value and X_t shows that the actual value. U is the normalized measure of the total forecast error. U is equal to 0 exhibits the perfect fit.

2.1.2: Tests for Normality

The normality test is executed to test whether the data under consideration is normally distributed or not. These tests are based on the analysis of two numerical measures, the shape skewness and the excess kurtosis. The data sets are normally distributed if those measures are close to zero. The acceptance of Jurque-Bera test also focused on skewness and kurtosis. Hence, the test of normality consists of checking the skewness and kurtosis on which the Jurque-Bera test is based.

Skewness: The skewness determines the degree of asymmetry of the data.

$$\text{Skewness} = \frac{\sum_{i=1}^n (X_i - \bar{X})^3}{(n-1) S^3} \quad (9)$$

Where \bar{X} is the mean and S is the standard deviation and n is the number of values (Christian and Jean-Michel 2004). The skewness of the normal distribution. If the data is normally distributed, then the skewness shows that the following data is symmetry. If the data is normally distributed if the symmetric distribution (skewness value is equal to zero). The distribution is positively skewed, if it is greater than zero and negatively skewed if it is less than zero.

Kurtosis: The Kurtosis measures the degree of peakness of the data. Kurtosis has been estimated as

$$\text{Kurtosis} = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{(n-1) S^4} \quad (10)$$

Where \bar{X} is the mean, S is the standard deviation and n is the number of values of the time series data.

esokurtic if it is equal to 3. Whereas it is leptokurtic if the value

is greater than 3. It is Platykurtic if the value is less than 3.

Jarque-Bera Statistics Test (JBS): The JBS is accepted with the normality of the data with skewness is equal to zero and excess kurtosis is also equal to zero. Jarque-Bera test is defined as follows.

$$\text{Jarque-Bera test} = \frac{n(\text{Skewness})^2}{6} + \frac{n(\text{Kurtosis} - 3)^2}{24} \quad (11)$$

Jarque-Bera test statistics are estimated as Chi-squared distribution with two degrees of freedom. Null hypothesis (H_0) is a normal distribution with skewness zero and excess kurtosis zero (which is the same as a kurtosis is 3). Alternate hypothesis (H_A) of given data is not normally distributed.

2.2: Methodology of the model

This section is based on the description of ARMA-GARCH model.

2.2.1: ARMA MODEL

A statistical approach to forecasting involves stochastic models to predict the values of sunspot cycles by using previous once. In the linear time series, two methods are frequently used in literature, viz. Autoregressive AR (p) and Moving Average MA (q) (Jenkins et.al. 1970 and Hipal et. al. 1994). ARMA models are developed by (Jenkins et. al 1994). An ARMA model is the combination of an idea of Autoregressive AR (p) and Moving Average MA (q) process. The concept of ARMA process is strongly relevant in volatility modeling. ARMA model is widely used for forecasting the future values. Autoregressive process (AR) is developed by (yule, 1927). In stochastic process, Autoregressive process AR (p) can be expressed by a weighted sum of its previous value and a white noise. The generalized Autoregressive process AR (p) of lag p as follow

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + \epsilon_t \quad (12)$$

Here ϵ_t is white noise with mean $E(\epsilon_t) = 0$, variance $\text{Var}(\epsilon_t) = \sigma^2$ and $\text{Cov}(\epsilon_{t-s}, \epsilon_t) = 0$, if $s \neq 0$. For every t, suppose that ϵ_t is independent of the X_{t-1}, X_{t-2}, \dots . ϵ_t is uncorrelated with X_s for each $s < t$. AR (p) models regress is past values of the data set. Whereas MA (q) model relates with error terms as a descriptive variable (Hipal et. al. 1994). The generalized Moving Average process MA (q) of lag q as follows.

$$X_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (13)$$

The process X_t is defined by the ARMA model.

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (14)$$

With ϵ_t is an uncorrelated process with mean zero. The prediction of ARMA (p, q) process shows the

2.2.2: GARCH MODEL

The generalized autoregressive conditional heteroskedasticity (GARCH) model is used to evaluate the volatility of an asset. It expresses that the volatility presence depends on the past observations and volatilities (Christian and Jean-Michel 2010). The time series τ_t can be modeled by

$$\tau_t = \sigma_t \epsilon_t \text{ with } \epsilon_t \sim \text{IID}(0, 1) \quad (15)$$

GARCH model is used to estimate the variance σ_t

$$\sigma_t^2 = \delta + \sum_{i=1}^q \alpha_i x_{t-i}^2 + \sum_{j=1}^p \beta_j \tau_{t-j}^2 \quad (16)$$

The GARCH (p, q) model is strictly stationary with finite variance, when the conditions $\delta > 0$, and $\sum_{i=1}^q \alpha_i x_{t-i}^2 + \sum_{j=1}^p \beta_j \tau_{t-j}^2 < 1$ are essential. The GARCH model has similar form with the ARMA model. Moreover, the GARCH process can be derived by using a similar theory and method with ARMA.

2.2.3: ARMA (p, q) – GARCH (1, 1) METHODS OF SUNSPOT CYCLES

The concept of ARMA models is strongly relevant in volatility modeling. The generalized autoregressive conditional heteroscedastic (GARCH) models can be linked as ARMA models. GARCH Models satisfy an ARMA equation with white noise. In time series, GARCH model supposition that conditional mean is zero. Generally, conditional mean of ARMA model can be structured. Identification of GARCH process focused on the square of residuals from the appropriate ARMA models. Moreover, in the ARAM process the quasi-maximum likelihood estimation is nearly independent of their GARCH process. ARMA estimation and GARCH estimation are strongly correlated if the ARMA – GARCH process has a skewed distribution (Csyer et. al 2008). The ARAM process and GARCH process have similar behavior in forecasting. ARMA – GARCH process provides a good estimation in time series data.

GARCH (1, 1) process specification with ARMA (p, q) is defined as follows ($t = 0 \pm 1, \pm 2, \dots$).

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (17)$$

$$\tau_t = \sigma_t \epsilon_t \text{ with } \epsilon_t \sim \text{IID}(0, 1) \quad (18)$$

$$\sigma_t^2 = \delta + \beta_1 \tau_{t-1}^2 + \beta_2 \tau_{t-2}^2 + \dots + \beta_p \tau_{t-p}^2 + \gamma \sigma_{t-1}^2 \quad (19)$$

Where $E(\tau_t) = 0$, variance $\text{Var}(\tau_t | \tau_{t-1}^2, \tau_{t-2}^2, \dots) = \sigma^2$ and $\text{Cov}(\tau_{t-s}, \tau_t) = 0$, if $s \neq 0$.

Moreover, The Box-Jenkins methodology with GARCH approach is used to develop models, to estimate the models and to forecast the sunspot cycle's data.

3. Result And Discussion

This study focused to estimate and forecast the future sunspots with Box-Jenkins ARMA (p, q) GARCH models. Using lags with second differences for making data series stationary. The Independent Generalized Autoregressive Centralized Heteroskedastic Model (GARCH) is mostly the only three parameters that allow an unlimited number of square roots to influence the present infinite variables. As the ARCH integrated the consultation independent feature in the absence of a return of sunspots cycles, parameters in GARCH (p, q) is frequently used for modeling, this model is insufficient parameters. They develop good estimates. The conditional variance estimated through GARCH is a weighted average of past residuals. Weight is low, but never zero. Essential for GARCH, it is the fact that it allows a vertical variable that it depends on the previous screen itself [6]. The novelty of this study to analyze the ARMA (p, q) -GARCH (1, 1) process of sunspot cycles. The ARMA (p, q) -GARCH (1, 1) model based on the least value of Darbin - Waston statistics test (DW). Least DW value (< 2) shows that each value of cycles is strongly correlated and persistence to each other. AIC, SIC, HQC and Log likelihood also estimate to each cycle. In Tables 1, 2 and 3 are depicted the GARCH (1, 1) model equations to specification ARMA (p, q) model of sunspot cycles by diagnostic checking test, forecast evolution and normality test. The Gaussian quasi maximum likelihood estimation is used to analysis ARMA (p, q) -GARCH (1, 1) model. Lagrange multiplier is used to verify the ARCH effect on following time series data. Ljung-Box test is used for serial correlation of each sunspot cycle. The novelty of this research to analysis the conditional mean and conditional variance effect on each sunspot cycle. The appropriate model which is frequently verified in sunspot cycles is GARCH (1, 1) with specification ARMA (2, 2) cycles (1st, 4th, 12th, 13th, 14th, 15th, 16th, 19th, 20th, 23rd and 24th). ARMA (3, 3) -GARCH (1, 1) process follows with means and standard deviation 0.153 ± 0.988 , 0.190 ± 0.983 , 0.097 ± 1.009 and 0.238 ± 0.973 of most appropriate sunspot cycles (5th, 6th, 7th and 15th) models respectively. ARMA (5, 1) -GARCH (1, 1) model follows cycles (2nd and 11th) with 0.167 ± 0.991 and 0.249 ± 0.968 respectively. Cycle (3rd) represents ARMA (1, 1) -GARCH (1, 1) model with 0.197 ± 0.983 . ARMA (5, 3) -GARCH (1, 1) process follows cycles (18th and 19th) with 0.258 ± 0.965 and 0.258 ± 0.966 respectively. Sunspot cycle (3rd) represents ARMA (1, 1) -GARCH (1, 1) model with 0.197 ± 0.983 . Sunspot cycle (8th) explore ARMA (3, 2) -GARCH (1, 1) process with 0.305 ± 0.954 . Sunspot cycle (9th) explore ARMA (4, 2) -GARCH (1, 1) model with 0.181 ± 0.984 . Sunspot cycle (10th) explore ARMA (4, 4) -GARCH (1, 1) model with 0.145 ± 0.992 . Sunspot cycle (22nd) explore ARMA (6, 1) -GARCH (1, 1) model with 0.206 ± 0.977 . Sunspot cycle (5th) appropriate model ARMA (3, 3) -GARCH (1, 1) based on least value of Akaike information criterion (AIC), Bayesian Schwarz information criterion (BIC) and Hannan Quinn information criterion (HIC) which has least value 7.075953, 7.197462 and 7.125322 respectively. According to log likelihood sunspot cycle (24th) best model ARMA (2, 2) -GARCH (1, 1) with maximum log likelihood value - 12808.63. Test of normality demonstrates that each sunspot cycles have positive skewed expect sunspot cycles (4th and 19th) have negative skewed with - 0.283 and - 0.1425 respectively. Each sunspot cycle has kurtosis value greater than 3 (Leptokurtic) heavy tail expect (2nd, 7th, 10th, 11th, 18th and 19th ,) cycles follows platokurtic flat tail. Jurque-Bera test failed in each cycle, which is shown that sunspot cycles are not normally distributed. Diagnostic Checking Test is chosen with compression of these techniques with the help of Root mean square error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Figure 1 shows the GARCH plots of sunspot cycles (1-24 cycles complete time series data

from 1855 to 2019) with conditional variance. In the Fig. 2 displayed that the forecasting evolution of each sunspot cycle (1–24) analysis in the view of diagnostic check, for each cycle Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) as well as in the Table 2 is determined the forecasting evolution of each sunspot cycle analysis in the view of diagnostic check, for each cycle Mean Absolute Error (MAE) has least value. 6th cycle of sunspot has the smallest value of RMSE, MAE and MAPE are 25.17780, 17.81173 and 79.46162 respectively.

Table 1
Diagnostic test of ARMA (p, q)- GARCH for Sunspot Cycles (1st to 24)

Cycles	ARMA(p, q)- GARCH (1,1)	R ²	ADJ R ²	SE Reg	Log Likelihood	AIC	SIC	HQC	DWS
1	ARMA(2, 2)- G(1,1)	0.658	0.65288	14.3671	-516.1398	8.158	8.292	8.213	1.450
2	ARMA(5, 1)- G(1,1)	0.557	0.549487	26.3177	-522.1865	9.266	9.410	9.325	1.516
3	ARMA(1, 1)- G(1,1)	0.827	0.824172	21.8562	-454.6855	8.611	8.7608	8.672	1.431
4	ARMA(2, 2)- G(1,1)	0.871	0.869758	16.3983	-673.2805	8.039	8.150	8.084	1.297
5	ARMA(3, 3)- G(1,1)	0.768	0.764758	9.07455	-517.6205	7.076	7.197	7.125	1.008
6	ARMA(3, 3)- G(1,1)	0.557	0.551562	12.0414	-583.7966	7.372	7.4878	7.419	1.471
7	ARMA(3, 3)- G(1,1)	0.571	0.567804	17.0997	-503.4886	8.546	8.663	8.593	1.198
8	ARMA(3, 2)- G(1,1)	0.624	0.617639	29.8739	-555.3826	9.356	9.4957	9.413	0.975
9	ARMA(4, 2)- G(1,1)	0.661	0.655573	22.5427	-603.2278	8.766	8.893	8.817	1.434
10	ARMA(4, 4)- G(1,1)	0.766	0.762319	15.8800	-584.8237	8.206	8.3296	8.256	1.103
11	ARMA(5, 1)- G(1,1)	0.786	0.783275	21.7951	-592.5847	8.551	8.6773	8.602	1.271
12	ARMA(2, 2)- G(1,1)	0.722	0.717787	14.3390	-549.7031	7.827	7.9517	7.878	1.330
13	ARMA(2, 2)- G(1,1)	0.801	0.798477	13.7487	-571.5942	7.805	7.9268	7.855	1.326

Cycles	ARMA(p, q)- GARCH (1,1)	R ²	ADJ R ²	SE Reg	Log Likelihood	AIC	SIC	HQC	DWS
14	ARMA(2, 2)- G(1,1)	0.590	0.583965	16.9820	-539.1395	8.017	8.1452	8.070	1.288
15	ARMA(3, 3)- G(1,1)	0.579	0.571190	22.2452	-492.8471	8.752	8.8957	8.810	0.802
16	ARMA(2, 2)- G(1,1)	0.727	0.722219	15.1424	-498.5507	8.073	8.2086	8.128	1.003
17	ARMA(2, 2)- G(1,1)	0.755	0.751202	20.2128	-535.5396	8.665	8.8004	8.720	1.177
18	ARMA(3, 5)- G(1,1)	0.658	0.652470	32.2541	-568.5037	9.418	9.5560	9.474	0.962
19	ARMA(5, 3)- G(1,1)	0.845	0.842241	28.4069	-578.0802	9.056	9.1885	9.110	0.794
20	ARMA(2, 2)- G(1,1)	0.817	0.814264	16.0709	-581.2417	8.389	8.5152	8.440	1.325
21	ARMA(2, 2)- G(1,1)	0.815	0.811452	24.8450	-538.7563	9.216	9.2957	9.212	1.138
22	ARMA(6, 1)- G(1,1)	0.792	0.788657	26.8954	-562.5152	9.170	9.3061	9.225	1.463
23	ARMA(2, 2)- G(1,1)	0.803	0.800593	19.0536	-593.3379	8.562	8.6880	8.613	1.280
24	ARMA(2, 2)- G(1,1)	0.797	0.794662	20.2114	-581.2489	8.459	8.4704	8.463	1.474

Table 2
Forecast Evolution of Sunspot Cycles by AMRA (p, q)-GARCH model (1st to 24)

Cycles	ARMA(p, q)- GARCH (1,1)	RMSE	MAE	MAPE	GARCH =A+ B f_Res^2 + C f_GARCH(-1)
1	ARMA(2, 2) - G(1,1)	32.67503	25.67541	56.50119	13.1418 + 0.1512*Resid(-1)^2 + 0.7964*GARCH(-1)
2	ARMA(5, 1) - G(1,1)	43.75567	33.31524	124.7012	13.5351 + 0.1668*Resid(-1)^2 + 0.8192*GARCH(-1)
3	ARMA(1, 1) - G(1,1)	82.97762	64.99252	88.91661	GARCH = 4.6145 + 0.2983*Resid(-1)^2 + 0.7611*GARCH(-1)
4	ARMA(2, 2) - G(1,1)	69.78911	53.91451	82.29495	13.0234 + 0.3321*Resid(-1)^2 + 0.6463*GARCH(-1)
5	ARMA(3, 3) - G(1,1)	28.93611	22.58619	85.55745	4.1899 + 0.2784*Resid(-1)^2 + 0.6817*GARCH(-1)
6	ARMA(3, 3) - G(1,1)	25.17780	17.81173	79.03446	1.5456 + 0.1084*Resid(-1)^2 + 0.8929*GARCH(-1)
7	ARMA(3, 3) - G(1,1)	35.06270	27.96897	113.0161	97.4526 + 0.3116*Resid(-1)^2 + 0.3843*GARCH(-1)
8	ARMA(3, 2) - G(1,1)	70.49102	53.86759	79.46162	18.1207 + 0.4027*Resid(-1)^2 + 0.6657*GARCH(-1)
9	ARMA(4, 2) - G(1,1)	55.68955	42.46238	61.37514	8.0428 + 0.1739*Resid(-1)^2 + 0.8282*GARCH(-1)
10	ARMA(4, 4) - G(1,1)	50.33423	40.31416	150.4605	5.2998 + 0.1693*Resid(-1)^2 + 0.8236*GARCH(-1)
11	ARMA(5, 1) - G(1,1)	68.21752	50.46720	163.9832	5.5674 + 0.2023*Resid(-1)^2 + 0.8006*GARCH(-1)
12	ARMA(2, 2) - G(1,1)	40.16201	30.27203	105.1265	2.0310 + 0.3175*Resid(-1)^2 + 0.7260*GARCH(-1)
13	ARMA(2, 2) - G(1,1)	44.99864	33.86295	108.7426	3.4242 + 0.2769*Resid(-1)^2 + 0.7417*GARCH(-1)
14	ARMA(2, 2) - G(1,1)	36.86236	27.61363	117.5639	0.3897 + 0.2509*Resid(-1)^2 + 0.7845*GARCH(-1)
15	ARMA(3, 3) - G(1,1)	53.34212	41.91097	79.70615	20.9228 + 0.5071*Resid(-1)^2 + 0.5522*GARCH(-1)
16	ARMA(2, 2) - G(1,1)	44.927819	35.31193	116.2730	9.9187 + 0.3086*Resid(-1)^2 + 0.6774*GARCH(-1)
17	ARMA(2, 2) - G(1,1)	67.16046	54.07226	96.16188	7.7621 + 0.1524*Resid(-1)^2 + 0.8407*GARCH(-1)
18	ARMA(3, 5) -	89.57113	72.19864	122.4058	9.4407 + 0.2034*Resid(-1)^2 + 0.8127*GARCH(-1)

Loading [MathJax]/jax/output/CommonHTML/jax.js

Cycles	ARMA(p, q)- GARCH (1,1)	RMSE	MAE	MAPE	GARCH =A+ B f_Res^2 + C f_GARCH(-1)
19	ARMA(5, 3) - G(1,1)	113.2087	88.62173	87.83657	$3.7640 + 0.3486 * \text{Resid}(-1)^2 + 0.6959 * \text{GARCH}(-1)$
20	ARMA(2, 2) - G(1,1)	63.56454	51.95468	74.98162	$618.4769 + 0.1576 * \text{Resid}(-1)^2 - 1.0532 * \text{GARCH}(-1)$
21	ARMA(2, 2) - G(1,1)	91.74100	72.87236	80.66910	$14.7176 + 0.1767 * \text{Resid}(-1)^2 + 0.8190 * \text{GARCH}(-1)$
22	ARMA(6, 1) - G(1,1)	87.08569	66.45792	75.91525	$11.4294 + 0.1612 * \text{Resid}(-1)^2 + 0.8380 * \text{GARCH}(-1)$
23	ARMA(2, 2) - G(1,1)	65.31769	50.61469	93.96324	$10.1606 + 0.1681 * \text{Resid}(-1)^2 + 0.8193 * \text{GARCH}(-1)$
24	ARMA(2, 2) - G(1,1)	54.37098	38.55961	205.0083	$4.2183 + 0.2292 * \text{Resid}(-1)^2 + 0.7863 * \text{GARCH}(-1)$

Table 3
Test of Normality ARMA (p, q)-GARCH (1, 1) process of Sunspot Cycles (1st to 24)

cycles	ARMA-GARCH	Mean	Median	Std.D	Skewness	Kurtosis	Jur-Bera
1	ARMA(2,2)-G(1,1)	0.143	-0.202	0.991	0.3875	3.4860	4.4639
2	ARMA(5,1)-G(1,1)	0.167	0.107	0.991	0.256	2.6522	1.913
3	ARMA(1,1)-G(1,1)	0.197	0.045	0.983	0.9322	6.1950	61.007
4	ARMA(2,2)-G(1,1)	0.140	0.121	0.991	-0.283	4.226	12.839
5	ARMA(3,3)-G(1,1)	0.153	0.061	0.988	0.1280	3.501	1.9499
6	ARMA(3,3)-G(1,1)	0.190	-0.027	0.983	1.1371	4.882	58.097
7	ARMA(3,3)-G(1,1)	0.097	0.018	1.009	0.4005	2.8365	3.3141
8	ARMA(3,2)-G(1,1)	0.305	0.178	0.954	0.6251	3.3824	8.5454
9	ARMA(4,2)-G(1,1)	0.181	0.093	0.984	0.8470	4.5254	30.096
10	ARMA(4,4)-G(1,1)	0.145	0.045	0.992	0.089	2.8840	0.2716
11	ARMA(5,1)-G(1,1)	0.249	0.322	0.968	0.046	2.9809	0.2716
12	ARMA(2,2)-G(1,1)	0.130	-0.003	0.994	0.6397	3.0275	9.6901
13	ARMA(2,2)-G(1,1)	0.140	0.0615	0.992	0.4077	3.8820	808934
14	ARMA(2,2)-G(1,1)	0.030	-0.112	0.996	0.9258	4.1931	27.4940
15	ARMA(3,3)-G(1,1)	0.238	0.098	0.973	0.4546	3.1964	4.1101
16	ARMA(2,2)-G(1,1)	0.165	0.148	0.987	0.1321	2.6787	0.9011
17	ARMA(2,2)-G(1,1)	0.167	0.095	0.985	0.2786	3.2299	1.8924
18	ARMA(5,3)-G(1,1)	0.258	0.100	0.965	0.2909	2.6303	2.4153
19	ARMA(5,3)-G(1,1)	0.258	0.297	0.966	-0.1425	2.4166	2.2661
20	ARMA(2,2)-G(1,1)	0.125	0.017	0.996	0.4562	3.1727	5.0298
21	ARMA(2,2)-G(1,1)	0.142	0.129	0.989	0.0006	3.7090	2.4922
22	ARMA(6,1)-G(1,1)	0.206	0.108	0.977	0.1528	3.6409	2.6051
23	ARMA(2,2)-G(1,1)	0.118	-0.050	0.993	0.6868	4.5192	24.4696
24	ARMA(2,2)-G(1,1)	0.142	0.318	0.990	0.5121	3.8258	218.54

The presentation of ARCH model and its modification namely generalized autoregressive conditional heteroskedasticity GARCH has been studied using sunspot cycles (1st – 24th). The sunspot cycles (1st – 24th) have been modelled and forecasted using GARCH volatility model with specification autoregressive ARMA (p, q) process are used to estimate and forecast evolution the sunspot cycles (1st -24th). GARCH stationary volatility model expresses the finest forecasting model as compared with other models. The Gaussian quasi maximum likelihood estimation is used to analysis ARMA (p, q)-GARCH (1, 1) process. The appropriate model is selected by residuals diagnostic checking (Lagrange multiplier LM test for knowing ARCH effect, Ljung-Box test for checking autocorrelation or ARCH effect in given data, and last normality test). ARMA (p, q)-GARCH (1, 1) process follows leptokurtic that is fat and heavy tail (values are strongly correlated to each other). The sunspot cycles ARMA (p, q)-GRACH (1, 1) process expresses the positive skewness except cycles 4 and 19. In this study, the sunspot cycles follow GARCH specification with ARMA (2, 2) model for cycles (1st, 4th, 12th, 13th, 14th, 15th, 16th, 19th, 20th, 23rd and 24th). Sunspot cycles (5th, 6th, 7th and 15th) follow ARMA (3, 3)-GARCH model. Whereas the cycles (2nd and 11th) show appropriate model is ARMA (5, 1) -GARCH (1, 1) process. ARMA (5, 3) -GARCH (1, 1) process expresses cycles (18th and 19th). Durbin-Waston (DW) statistics test value of each sunspot cycles are less than 2 which indicate that sunspot observations are correlated to each other. Akaike information criterion (AIC), Bayesian Schwarz information criterion (BIC) and Hannan Quinn information criterion (HIC) explored that the most appropriate model is a 5th sunspot cycle. ARMA (p, q)-GARCH (1, 1) process of sunspot cycles rejected Jurque-Bera test for normality test. Forecasting of each sunspot cycle was analyzed based on Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). For each cycle Mean Absolute Error (MAE) has the least value. Sunspot cycle 6th has the smallest value of RMSE, MAE, and MAPE which are 25.17780, 17.81173 and 79.03446 respectively. On the bases of ARMA-GARCH process results for the cycles 1st to 24th is stationary and linear. On the behalf of the study, we can predict that the cycle 25 will be stationary and linear.

The results exposed that ARMA (p, q)-GARCH (1, 1) process is the finest volatility modeling for solar activities. Based on the implications of the results, the scope of the future research directions will be expanded.

Declarations

Acknowledgments:

The authors are also thankful to the World Data Centre (WDC) and the National Oceanic and Atmospheric Administration (NOAA) for providing the Sunspots data. In addition, the work done by the referees in reviewing this manuscript is greatly appreciated.

References

Muraközy J., Ludmány A. (2012) Phase lags of solar hemispheric cycle. Monthly Notices of the Royal Astronomical Society 419 (4): 3624-3630 doi:10.1111/j.1365-2966.2011.20011.x, ADS:

Box G. E. P., Jenkins G. M. and Reinsel G. C (1994) Time Series Analysis: Forecasting and Control New Jersey: Prentice Hall, Inc.

Goh H H, Tan K L, Khor C Y and Ng S L (2016) Volatility and Market Risk of Rubber Price in Malaysia: Pre- and Post- Global Financial Crisis Journal of Quantitative Economics **14**(2): p 323–344

Bollerslev T., Engle, R. F. and Nelson, D.B. (1994) ARCH Models, Handbook of Econometrics, Amsterdam: Elsevier Science B. V

Engle Robert F. (2001) GARCH101: An Introduction to the use of ARCH/GARCH models in Applied Econometrics, NYU working paper no. FIN-01-030. Available at SSRN: <https://ssrn.com/abstract=1294571>.

E. McKenzie (1984) General exponential smoothing and the equivalent ARMA process, Journal of Forecasting **3**, 333 – 344.

W. Ji and K. Chee (2011) Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN, Solar Energy **85**, 808 – 817.

Salisu A A and Fasanya I O (2013) Modelling oil price volatility with structural breaks. Energy Policy p 554–562

Epaphra M (2017) Modeling Exchange Rate Volatility: Application of the GARCH and EGARCH Models Journal of Mathematical Finance **7**(1): p 121–143

Pham H T and Yang B S (2010) Estimation and forecasting of machine health condition using ARMA/GARCH model Mechanical Systems and Signal Processing **24**(2): p 546–558

Engle R F (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica **50** p 987–1007

Hill R C, Griffiths W E and Lim G C (2011) Principles of Econometrics ed Fourth (United States of America: John Wiley & Sons, Inc) chapter 14 pp 526

Brooks C (2014) Introductory Econometrics for Finance ed Third (New York: Cambridge University Press) chapter 9 pp 430

Olson D L and Wu D (2017) Predictive Data Mining Models (Singapore: Springer) chapter 6 pp 63

Thomas M. and Denial M. (2002) Whittle Estimate in a Heavy Tail GARCH (1, 1) Model, Stochastic Process and their Applications volume 100, p. 187-222

Jenkins, J.M., R.D. Milholland, J.P. Lilly and M.K. Beute. 1970. Commercial gladiolus production in North Carolina. N.C. Agric. Ext. Circ. 44: 1-34.

Hipel, K.W. and McLeod, A.I. (1994) Time Series Modelling of Water Resources and Environmental Systems. Elsevier, Amsterdam.

G. udny Yule (1927) On the Method of Investigating Periodicities in Disturbed Series with Special reference to Wolfer's Sunspot numbers, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character Vol. 226, pp. 267-298 (32 pages) published By: Royal Society

Figures

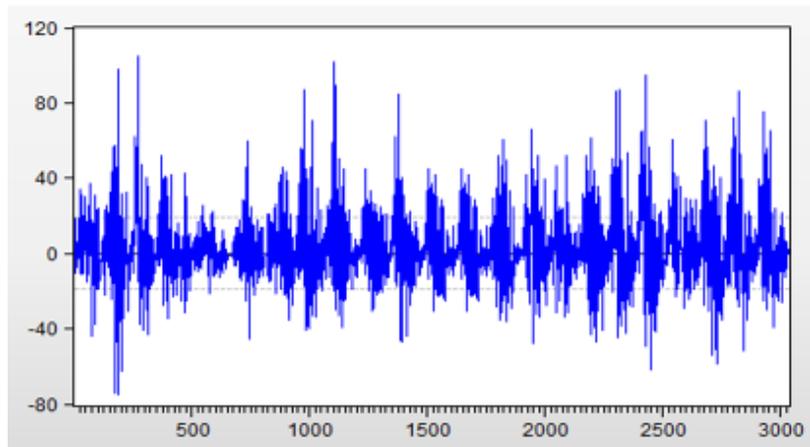


Figure 1

Residuals ARMA – GARCH models of Sunspot Cycle (1-24)

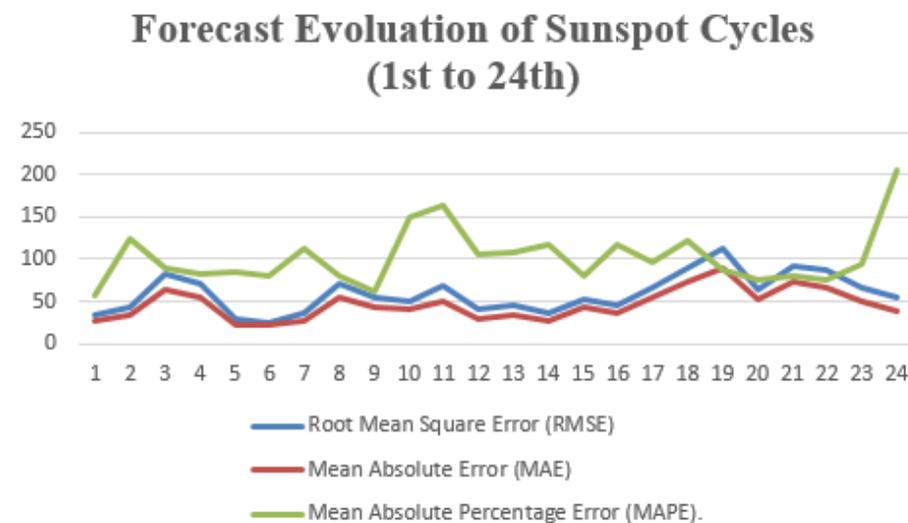


Figure 2

Forecasting Evolution of Sunspot Cycle (1-24)