MATEMATIKA, 2015, Volume 31, Number 2, 159–176

Time Series Modeling of Monthly Temperature Data of Jerusalem/Palestine

Bisher M. Iqelan
Department of Mathematics, The Islamic University of Gaza
PO BOX 108, Gaza-AlRimal, Palestine
e-mail: biqelan@iugaza.edu.ps

Abstract The objective of the research is to estimate the month-ahead temperature records in Jerusalem, Palestine. In this study, the modeling mechanism of analytic for forecasting is considered. This paper explores implementation of ARIMA and GARCH modeling techniques to fit a historical data set and estimate the coefficients of the suitable models for fitting the average monthly temperature data of Jerusalem in Palestine for the period from January 1964 to December 2013. The analysis of this study are carried out with the assist of R software. Eventually, using different statistical measures, comparison efficiency between ARIMA$(2, 0, 1)(2, 1, 1)_{12}$ and AR(1)-GARCH(1,1) models are produced. AR(1)-GARCH(1,1) is detected to be a better than ARIMA$(2, 0, 1)(2, 1, 1)_{12}$ model.

Keywords Jerusalem/Palestine; R software package; Time series; ARIMA; GARCH; Modeling; Estimation; Forecasting; Average monthly data of temperature; Volatility.

2010 Mathematics Subject Classification 62M10

1 Introduction

One goal of time series analysis is to forecast the future values of the time series data. Many methods and approaches for formulating forecasting models are available in the literature. The most widely used technique for analysis of the time series data is, undoubtedly, the Box Jenkins Autoregressive integrated moving average (ARIMA) methodology. One crucial assumption of the ARIMA model is the presence of linear dependence in the observations of the series. However, many financial time series show periods of stability, followed by unstable periods with high volatility. Volatility is a condition where the conditional variance changes between extremely high and low values, see Chatfield [1].

In order to account for volatility, we require "non-linear time series models" that allow for conditional changes in the variance. Prediction of the variability of the future values based on the past and current values is the main concern for the development of conditional variance process model [2, 3].

The most suitable parametric nonlinear time series model captures volatility is the Autoregressive conditional heteroscedastic (ARCH) model, which was pioneered by Engle [4]. ARCH model allows the conditional variance to change over time and its main purpose is to predict the future conditional variance. ARCH model can also be called the error term model [5]. However, forecasting of future conditional variances by using ARCH model involves the past squared returns only. Therefore, Bollerslev [6] extended the ARCH model which includes past conditional variances instead of past squared returns. This model is known as Generalized ARCH (GARCH) model. If the lag of past conditional variance is zero, ARCH and GARCH models are equal. For more details see [7].

In this paper, we consider Box Jenkins ARIMA as criterion model. The current study is focuses on forecasting of a monthly average temperature series that is from 1964 to 2012 in Jerusalem/Palestine using GARCH model. To fit the models, R-programming [8] is used with a step by step procedure. To estimate the coefficients of the fitted GARCH model, the `R` package '`fGarch`' [9] is particularly used. The GARCH model is used to provide a volatility measure of the temperature series. The goodness of fit and the forecasting models are measured by AIC and BIC information criterion. For comparison, we use the statistical measures: the mean absolute error (MAE), the root mean square error (RMSE), and the mean absolute percentage error (MAPE) respectively.

The remaining parts of this paper will be organized as follows. Section 2 reviews previous studies related to forecasting techniques and particularly to temperature time series. In Section 3, we introduce the ARIMA and GARCH models which are used to describe the distinct components and their processes for estimating and forecasting our monthly average temperature series. Section 4 describes data, analyzes different forecasting techniques, and illustrates the methodology followed in this study. Finally, the conclusions are summarized in Section 5.

2 Previous Study for Temperature Time Series

Tol [10] suggests using a generalized AR conditional heteroscedastic (GARCH) model for daily temperature observations. Tol's suggestion for selecting the GARCH model depends on the detection that the predictability of meteorological variables is not constant but shows regular variations. He estimates the proposed model on 30 years of daily temperature observations in De Bilti, The Netherlands (Beneth [11]). Franses, Neste and Dijk [12] suggest to use a so-called quadratic GARCH (QGARCH) model allowing for asymmetry in the impact of innovations on the conditional variance. They estimate the model on the same time series records as in Tol [10].

Campbell and Diebold [13] model the daily average temperature in a number of US cities by an AR time series with a seasonal AR conditional heteroscedastic (GARCH) type dynamics for the residuals.

3 Time Series Models

In this section we introduce the models which are used to describe the distinct components of our daily temperature series.

3.1 The ARIMA model

The process $\{X_t\}$ is called an autoregressive moving average (ARMA) process, denoted by ARMA(p,q) is given by

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \qquad (1)$$

or simply by

$$\phi_p(\mathbf{B}) X_t = \theta_q(\mathbf{B}) \varepsilon_t \qquad (2)$$

where,

$$\phi_p(\mathbf{B}) = 1 - \phi_1 \mathbf{B} - \phi_2 \mathbf{B}^2 - \cdots - \phi_p \mathbf{B}^p$$

is an autoregressive polynomial of $\mathbf{B}$ for order p.

Also,

$$\theta_q(\mathbf{B}) = 1 - \theta_1\mathbf{B} - \theta_2\mathbf{B}^2 - \cdots - \theta_p\mathbf{B}^p$$

is a moving average polynomial of $\mathbf{B}$ for order q.

In the above, $\mathbf{B}$ is the backshift operator, used to simplify the representation of lag values, by $\mathbf{B}X_t = X_{t-1}$. A generalization of ARMA model, to cover a wide class of non-stationary time series, is achieved by proposing "differencing" in the model. A non-stationary time series $\{X_t\}$ is said to follow a non-stationary autoregressive integrated moving average (ARIMA) denoted by ARIMA(p, d, q) if it is expressed as:

$$\phi_p(\mathbf{B}) \bigtriangledown^d X_t = \mu + \theta_q(\mathbf{B})\varepsilon_t \tag{3}$$

where ε_t are identically and independently distributed as $\mathrm{N}(0, \sigma^2)$, $t = 1, 2, \ldots, N$ and N is the number of samples, d is the order of non-seasonal differences and $\bigtriangledown$ is the non-seasonal differencing operator, $\bigtriangledown = 1 - \mathbf{B}$. μ is the mean of a series assuming that after differencing is stationary. One can notice that when $d = 0$, the ARIMA(p, d, q) model becomes an ARMA(p, q) model. More details can be found in [2, 3, 14, 15].

The ARIMA model (3) is for non-stationary non-seasonal observations. Box and Jenkins in [14] popularized this model to deal with seasonality. For time series possessing a seasonal component that repeats every s observations, their proposed model is known as the seasonal ARIMA model. For monthly time series $s = 12$ and for quarterly time series $s = 4$.

The situation is entirely analogous in dealing with Equation 3. The seasonal autoregressive integrated moving average model is given by

$$\phi_P(\mathbf{B}^s)\phi_p(\mathbf{B}) \bigtriangledown_s^D \bigtriangledown^d X_t = \mu + \theta_Q(\mathbf{B}^s)\theta_q(\mathbf{B})\varepsilon_t \tag{4}$$

and is denoted as an ARIMA$(p, d, q) \times (P, D, Q)_s$, where D is the order of seasonal differences and $\bigtriangledown_s$ is the seasonal differencing operator, $\bigtriangledown_s = 1 - \mathbf{B}^s$. It is worth noting that ARIMA$(p, d, q) \times (P, D, Q)_s$ processes with $d \geq 1$ and/or $D \geq 1$ are non-stationary.

3.2 The GARCH model

A process $\{\varepsilon_t\}$ is called an ARCH(q) process if its evolution is described by the following two equations

$$\begin{aligned} \varepsilon_t &= \sigma_t\eta_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i\varepsilon_{t-i}^2 \end{aligned} \tag{5}$$

where the constraints on the model parameters in Equation 5 are $\omega > 0$, $\alpha_i \geq 0$ for all i, $i = 1, 2, \ldots, q$ and $\sum_{i=1}^{q} \alpha_i < 1$. This is to ensure that the conditional variance, σ_t^2, is always nonnegative. It is further assumed that $\{\eta_t\}$ is a sequence of independent and identically distributed random variables with mean 0 and variance 1 and is independent of the past process ε_t for all t. For more details, see [7, p. 34], and see also [10, 16].

As previously mentioned, Bollerslev [6] proposed the Generalized ARCH (GARCH) model in which conditional variance is also a linear function of its own lags and its evolution is described by the following pair of equations:

$$\begin{aligned}\varepsilon_t &= \sigma_t \eta_t,\\ \sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2\end{aligned} \tag{6}$$

To ensure that the conditional variance, σ_t^2, is nonnegative, the constraint on the parameters of Equation 6 are $\omega > 0,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, q$ and $\beta_j \geq 0,\ j = 1, 2, \ldots, p$. It is further assumed that $\{\eta_t\}$ is a sequence of independent and identically distributed random variables with $\mathrm{E}(\eta_t) = 0$ and $\mathrm{Var}(\eta_t) = 1$ and η_t is independent of $\{\varepsilon_{t-k},\quad k \geq 1\}$ for all t. When $p = 0$, Equation 6 is reduced to an autoregressive conditional heteroscedastic, ARCH(q), model. The necessary and sufficient condition for Equation 6 to define a weakly stationary process $\{\varepsilon_t,\ t = 0, \pm 1, \pm 2, \ldots.\}$ with $\mathrm{E}(\varepsilon_t^2) < \infty$ is that

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$$

(for more details see [17, Theorem 2.5, p. 37] and see also [18–20]). It is worthy to mention that in applications, GARCH$(1,1)$ is the most famous GARCH model. Moreover, one can derive the ARMA representation of the GARCH(p,q) model. Let

$$\zeta_t = \varepsilon_t^2 - \sigma_t^2,$$

the process $\{\zeta_t\}$ is said to be the innovation of the squared returns process,$\{\varepsilon_t^2\}$. Then

$$\varepsilon_t^2 = \omega + \sum_{i=1}^{r} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \zeta_t + \sum_{j=1}^{p} \beta_j \zeta_{t-j}, \tag{7}$$

where $r = \max(p, q),\ \alpha_i = 0$ if $i > q$, and $\beta_j = 0$ if $j > p$. The last equation shows that $\{\varepsilon_t^2\}$ is an ARMA(r,p) process. For more fine points, see [7, p. 35-39] and [17, p.19-21].

3.3 Parameters Estimation

The estimation of parameters for ARIMA model is a nonlinear problem that demands some special procedures, like the maximum likelihood method or nonlinear least-squares estimation. At this stage of model building, the estimated parameter values should minimize the sum of squared residuals. For this purpose, many software packages are applicable for fitting ARIMA models. In this current study, R software package will be used.

To choose the best ARIMA model based on observation data, we use the corrected Akaike Information Criterion (AIC) and can be calculated by the formula:

$$\mathrm{AICc} = N \log(\frac{SS}{N}) + 2(p + q + 1)\frac{N}{(N - p - q - 2)}$$

where,
N: the number of items after difference ($N = n - d$),
SS: sum of squares of differences
p & q: the order of autoregressive and moving average model, respectively.

In literature, method of Maximum Likelihood, ML, has been used to estimate the parameters of GARCH model. Consider a GARCH(p, q) model of Equation 6. Suppose that $\varepsilon_t^2 \sim \mathrm{N}(0, \sigma_t^2)$. The probability density function of ε_t is given by

$$f(\varepsilon_t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{\frac{-\varepsilon_t^2}{2\sigma_t^2}\right\}$$

The log likelihood function is obtained by taking log of the joint probability density function of ε_t,

$$\begin{aligned} \ell &= \log L = \log \prod_{i=1}^{N} \left(\frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{\frac{-\varepsilon_t^2}{2\sigma_t^2}\right\} \right), \\ &= -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\left\{\log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}\right\}. \end{aligned} \tag{8}$$

where,

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

Then, maximize the log-likelihood function ℓ of Equation 8 in order to estimate the values $\hat{\omega}$, $\hat{\alpha_i}$, $i = 1, 2, \ldots, q$ and $\hat{\beta_j}$ $j = 1, 2, \ldots, p$.

3.4 Testing for ARCH Effects

Testing for ARCH effects or disorders can also be defined as testing the presence of heteroscedasticity in time series model. Engle in [4] introduced the Lagrange Multiple (LM) test to check for ARCH disorders.

Let $\varepsilon_t = y_t - \mu_t$ be the residual series. The squared series $\{\varepsilon_t^2\}$ is used to implement the LM test to check for conditional heteroscedasticity. Under the null hypothesis, we have

$$H_0 : \alpha_i = 0, \quad i = 1, 2, \ldots, q$$

versus,

$$H_1 : \alpha_i \neq 0, \quad \text{for at least one } i$$

in the linear regression

$$\varepsilon_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2, \quad t = q+1, \ldots, N, \tag{9}$$

where q is the length of ARCH lags and N is the number of observations used in that regression. Test statistic for LM-test is defined by:

$$\mathrm{LM} = NR^2 \tag{10}$$

where R^2 is the R-squared from the regression of ε_t^2 in Equation 9 and defined by

$$R^2 = \frac{\text{regression sum of squares}}{\text{total sum of squares}}$$

Under the null hypothesis, the test statistics NR^2 is distributed as a Chi-squared distribution with q degrees of freedom. Rejecting H_0 when $LM > \chi^2_\alpha(q)$ implies that ARCH effect exists.

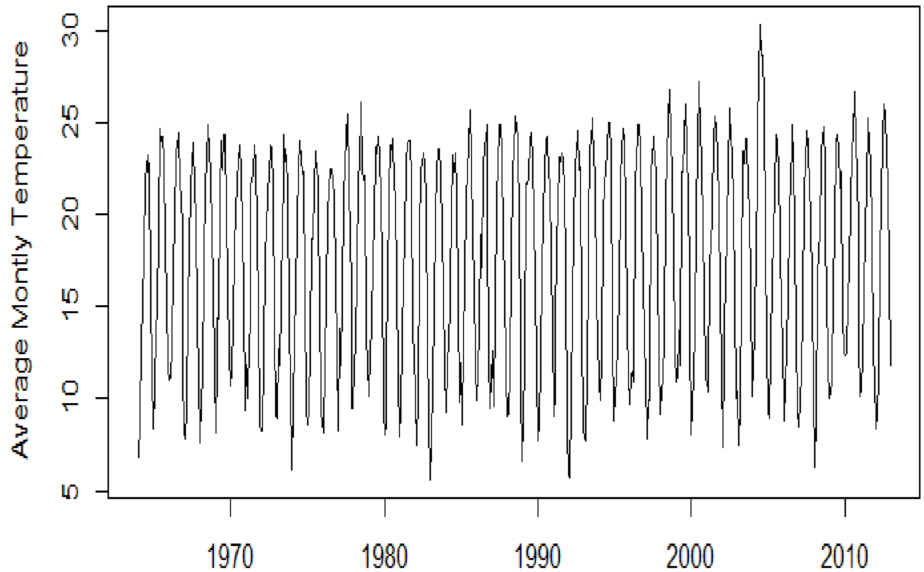

Figure 1: Data of Monthly Mean Temperature in Jerusalem/Palestine

4 Modeling of Jerusalem's Monthly Temperature Data

For the objective of research, the data set consists of 600 observations from average monthly temperature of Jerusalem, Palestine between 1964 and 2013. Considered holy to three major religions: Islam, Christianity, and Judaism, Jerusalem is one of the oldest cities in the world. During its ancient times, the city of Jerusalem has been demolished at least twice, besieged 23 times, attacked 52 times, captured 44 times, but remained a world heritage site. Jerusalem is 60 kilometers east of the Mediterranean Sea, and 35 kilometers west of the Dead Sea, the lowest body of water on Earth as seen in Figure 2. As a result, the city has a hot-summer Mediterranean climate with hot summers and wet winters. The geography and climate of this old city add more to its beauty and status. Therefore, it is worthy to hold statistical analysis and time series models for monthly temperature data of the old city of Jerusalem.

To identify and fit the model, 588 observations up to year 2012 are used for sampling and forecasting 12 months-ahead temperature records. The observations of year 2013 are considered for comparing the predicted average monthly temperature.

Figure 2: Map of Palestine Showing the Location of Jerusalem.

Data from January 1964 to December 2012 are plotted in Figure 1. The regular pattern of ups and downs is an indicator of seasonality. Moreover, to detect seasonality, we need to plot means for the 12 months of the year. Figure 3 shows that there are monthly differences (seasonality). Consequently, we will take a first seasonal difference of the data. The seasonally different data are shown in Figure 4. It seems that there is no seasonality for this time series and the data has mean 0. Therefore, we will not consider further differences.

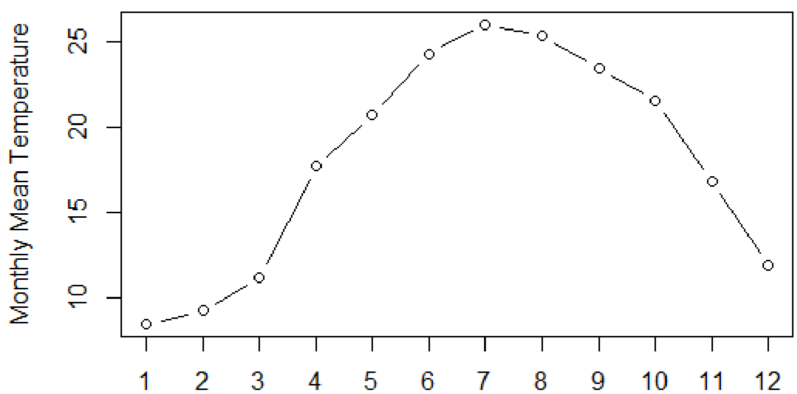

Figure 3: Data of Means of Jerusalem's Temperature for the 12 Months of the Year

4.1 Fitting of ARIMA Model

Our aim now is to find an appropriate ARIMA model based in the ACF and PACF shown in Figure 4. Looking at just the first 2 or 3 lags, it seems possible that an AR(2) might work based on the two spikes in the PACF. On the other hand, a tapering pattern in the beginning lags of ACF. That is, the pattern in the ACF is not indicative of any simple model.

For the seasonal behavior, we check the situation on around lags 12, 24 and so on. In the ACF, there is a collection of negative spikes around lag 12. In the PACF there is a clear spike at lag 12, 24 and also at lag 36.

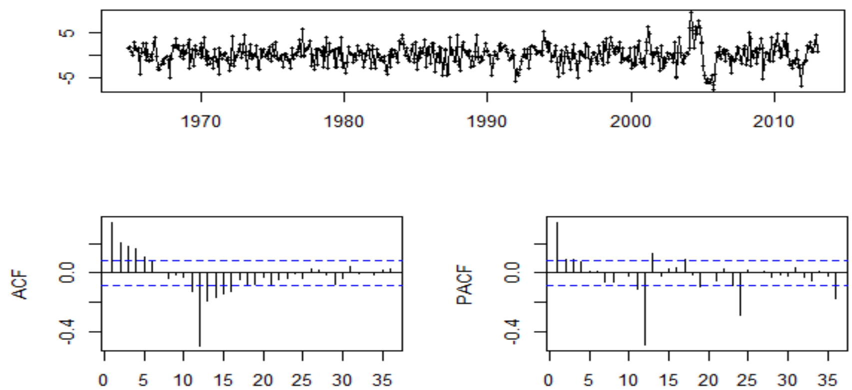

Figure 4: Seasonally Differenced Monthly Mean Temperature in Jerusalem

Consequently, this primary analysis proposes that the advisable model for these data is an ARIMA$(2,0,0)(3,1,1)_{12}$. We fit this model, along with some variations on it, and compute their AICc values which are shown in Table 1.

Table 1: AICc Values of Suggested ARIMA Models

ARIMA Model	AICc
ARIMA$(3,0,0)(3,1,1)_{12}$	2108.33
ARIMA$(3,0,1)(3,1,1)_{12}$	2106.62
ARIMA$(2,0,0)(3,1,1)_{12}$	2109.65
ARIMA$(2,0,1)(3,1,1)_{12}$	2105.78
ARIMA$(3,0,0)(2,1,1)_{12}$	2106.71
ARIMA$(3,0,1)(2,1,1)_{12}$	2104.97
ARIMA$(2,0,0)(2,1,1)_{12}$	2107.97
ARIMA$(2,0,1)(2,1,1)_{12}$	2103.82

All information criteria prefer the ARIMA$(2,0,1)(2,1,1)_{12}$ model, which is the model displayed in Equation 4. The fitted model in this case is

$$(1-0.0351\mathbf{B}^{12}-0.0632\mathbf{B}^{24})(1-0.9412\mathbf{B}+0.1119\mathbf{B}^{2})(1-\mathbf{B}^{12})\widehat{X}_t = 0.0017+(1-\mathbf{B}^{12})(1-0.6585\mathbf{B})\widehat{\varepsilon}_t \quad (11)$$

with estimated variance, $\widehat{\sigma}^2_{\varepsilon}=2.242$.

The diagnostics for the fit are displayed in Figure 5. There are one or two extreme standardized residuals. The Q-Q plot show some non-normality on the tails of the distribution which is accepted and the model seems to fit well.

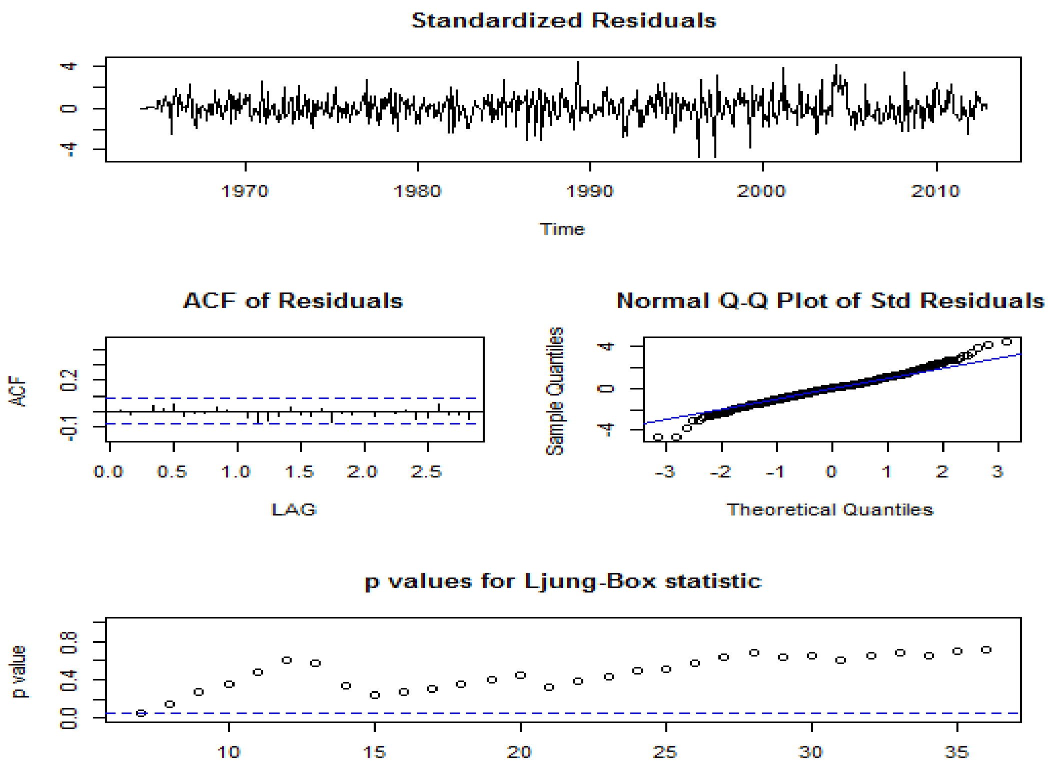

Figure 5: Residual Analysis for the ARIMA$(2, 0, 1)(2, 1, 1)_{12}$ fit to Jerusalem's Temperature Data Set.

The performance of the best model in Box-Jenkins modeling will later be compared with the best model in GARCH modeling.

4.2 Fitting of GARCH Model

Model identification of the GARCH Model is based on the ACF and PACF plots. On evaluating autocorrelations of squared residuals of the fitted ARIMA$(2, 0, 1)(2, 1, 1)_{12}$ model, recorded in Table 2, it is found that the autocorrelation is high at lag 12, which is 0.164. The ARCH-LM test statistic at lag 12 computed using Equations 9 and 10 is 41.5903, which is significant at 5% level. But it is not acceptable to apply ARCH model of order 12 because of the extraordinarily large number of parameters. Therefore, the parsimonious GARCH model is applied.

Table 2: Sample Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) of the Squared Residuals of the ARIMA$(2,0,1)(2,1,1)_{12}$ Series.

Lag	ACF of the squared residual series	PACF of the squared residual series
1	0.157	0.157
2	0.078	0.054
3	0.021	0.001
4	−0.009	−0.016
5	−0.066	−0.065
6	0.004	0.025
7	−0.029	−0.026
8	−0.039	−0.032
9	−0.036	−0.024
10	0.045	0.056
11	0.073	0.067
12	0.164	0.140
13	0.030	−0.028
14	−0.007	−0.034
15	−0.001	0.006
16	−0.022	−0.014
17	−0.009	0.014
18	−0.034	−0.036
19	−0.084	−0.068
20	−0.055	−0.017
21	−0.054	−0.032
22	−0.028	−0.026
23	0.071	0.065
24	0.129	0.092

A quick comparison of the results obtained from R to fit GARCH models, using the residuals of the ARIMA$(2,0,1)(2,1,1)_{12}$ series, is shown in Table 3. There are two models (the 2nd and the 4th) with low AIC within 2 decimals and both can be considered as the best. However, BIC is lower for the second model. Therefore GARCH(1,1) model with Mean and Variance Equation: AR(1)-GARCH(1,1) is preferred on the basis of minimum AIC and BIC values. Moreover, the LM test for this model has $p-$value $= 0.016199$ is less than 5% level of significance and the null hypothesis stated “no ARCH effect or no conditional heteroskedasticity in the residuals” is rejected. Thus the AR(1)-GARCH(1,1) model best fit our data set.

Table 3: Information Criterion Statistics with LM-ARCH p-value of Fitted GARCH Models

	Model	AIC	BIC	LM ARCH p-value
1	GARCH$(1,1)$	5.485320	5.515094	0.0173654
2	AR(1)-GARCH(1,1)	5.473084	5.510301	0.01619934
3	MA(1)-GARCH(1,1)	5.475204	5.512421	0.01470771
4	ARMA(1,1)-GARCH(1,1)	5.472166	5.516827	0.02477267

The $p-$values of residuals and squared residuals of Box-Ljung Q statistics of GARCH(1,1) are reported in Table 4. The $p-$values of residuals are less than 5% level of significance which shows no attendance of autocorrelations; however, the squared residuals show time dependency in the series.

Table 4: Standardised Residuals Tests of Residuals Using AR(1)-GARCH(1,1) Model.

			Statistic	p-value
Jarque-Bera Test	R	Chi^2	18621.04	0
Shapiro-Wilk Test	R	W	0.5995114	0
Ljung-Box Test	R	Q(10)	7.063854	0.7194033
Ljung-Box Test	R	Q(15)	19.31244	0.1999232
Ljung-Box Test	R	Q(20)	23.88994	0.2472329
Ljung-Box Test	R^2	Q(10)	1.541173	0.9988013
Ljung-Box Test	R^2	Q(15)	26.63124	0.03188855
Ljung-Box Test	R^2	Q(20)	27.81042	0.1139637
LM Arch Test	R	TR^2	24.72146	0.01619934

The residuals plot of the AR(1)-GARCH(1,1) model in Figure 6 attains a few peaks away from the boundaries and it displays volatility clustering.

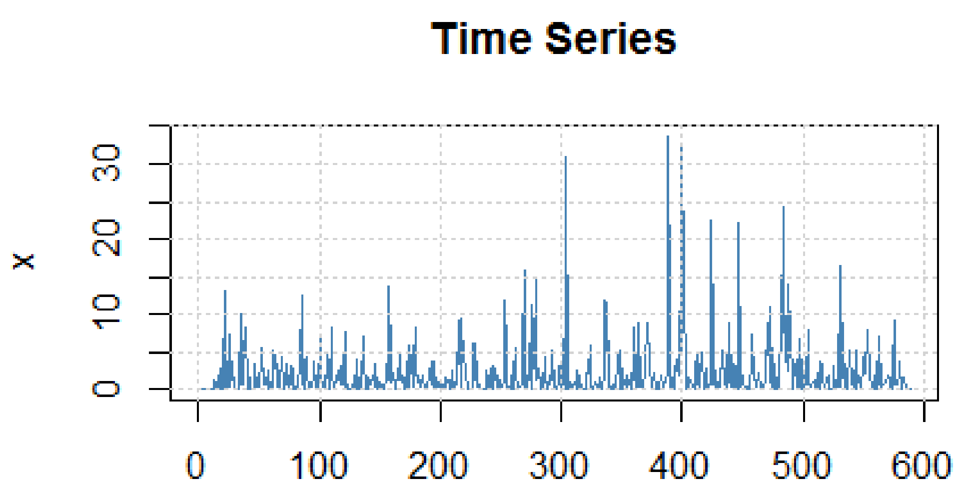

Figure 6: Time Series Plot of Residuals of AR(1)-GARCH(1,1) Process

The peaks of the model residuals coincide with the peaks of the standard deviation shown in Figure 7. Moreover, the standard deviation of AR(1)-GARCH(1,1) process shows that there is a high volatility in the middle and at the end of the year.

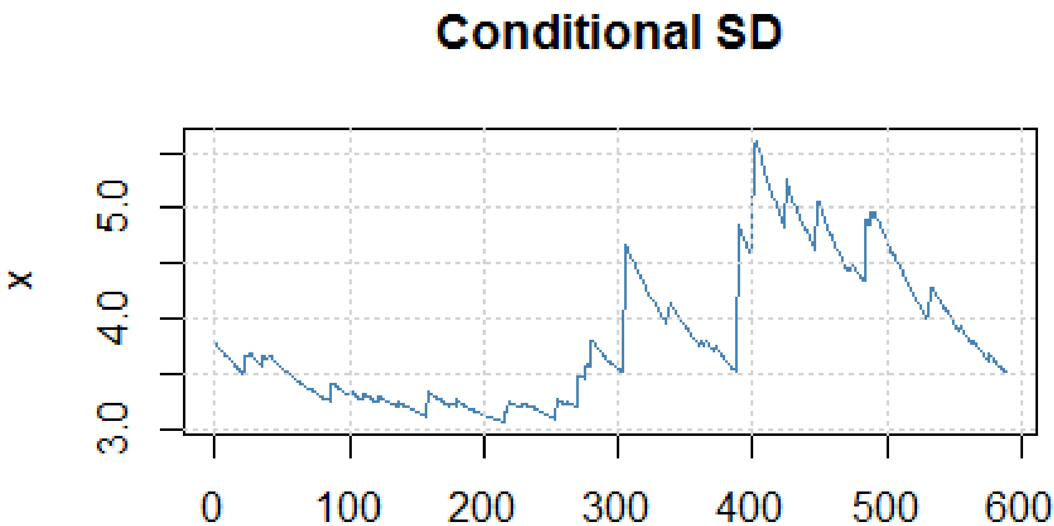

Figure 7: Time Series Plot of Standard Deviation of AR(1)-GARCH(1,1) Process.

The ACF of standardized residuals is shown in Figure 8. The ACF of standardized residuals in Figure 8 shows there are peaks i.e., autocorrelations but within ACF boundary. Notice the slow decay of the lag plots which means there is correlation between the magnitude of change in the residuals. In other words, there is serial dependence in the variance of the data. The ACF plot of squared standardized residuals is shown in Figure 9. As shown in Figure 9, some of the peaks of squared standardized residuals are reduced. The QQ plot is shown in Figure 10.

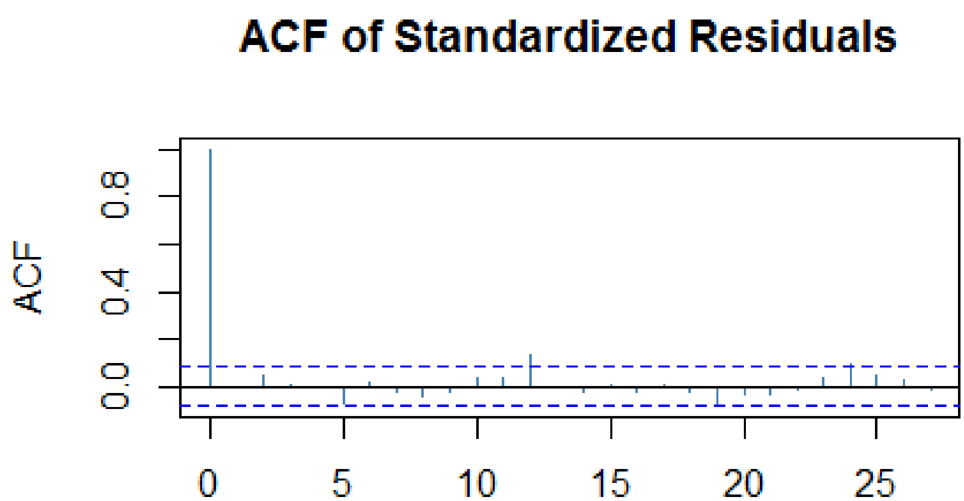

Figure 8: ACF Plot of Standardized Residuals.

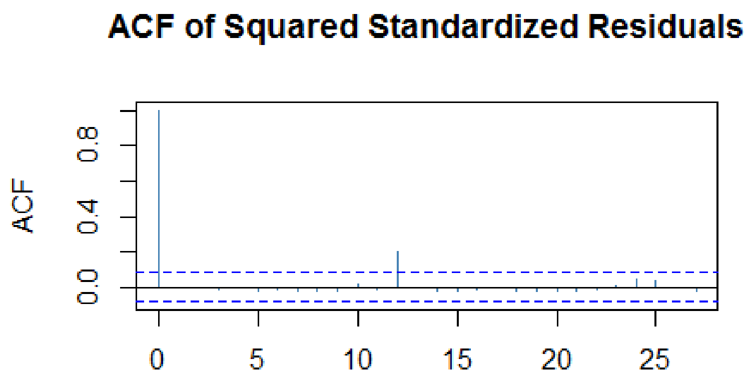

Figure 9: ACF Plot of Squared Standardized Residuals

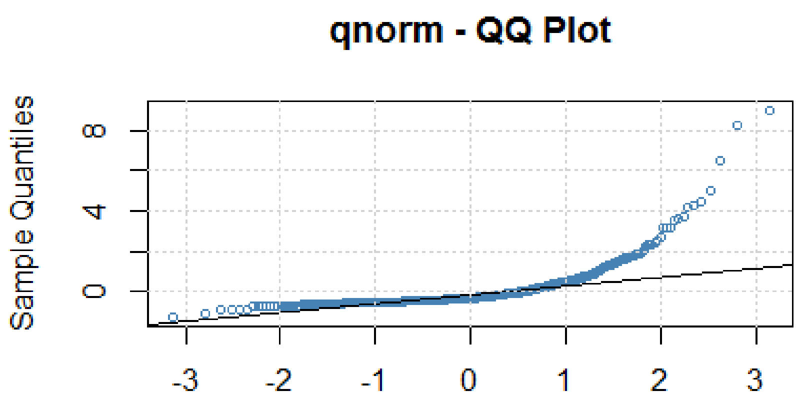

Figure 10: QQ Plot of AR(1)-GARCH(1,1) Model.

The '`garchFit`' function of the `R` package '`fGarch`' (see [9]) is used to estimate the coefficients of the AR(1)-GARCH(1,1) model. These coefficients are listed in Table 5:

Table 5: Coefficients of AR(1)-GARCH(1,1) Model with Error Analysis

	Estimate	Std. Error	t value	Pr($>\|t\|$)
mu	1.804746	0.173794	10.384	$< 2e-16$ ***
ar1	0.135693	0.045256	2.998	0.002715 **
omega	0.144262	0.049703	2.902	0.003703 **
alpha1	0.011225	0.003324	3.377	0.000732 ***
beta1	0.979603	0.004374	223.951	$< 2e-16$***

Full AR(1)-GARCH(1,1) model with the estimated coefficients is represented as:

$$\begin{aligned} X_t &= 1.804746 + 0.135693 X_{t-1} + \varepsilon_t, \\ \sigma_t^2 &= 0.144262 + 0.011225 \varepsilon_{t-1}^2 + 0.979603 \sigma_{t-1}^2 \end{aligned} \tag{12}$$

The $p-$values for all parameters are less than 0.05, showing that they are highly statistically significant. The value of $\alpha_1+\beta_1$ is less than but close to unity and $\beta_1 > \alpha_1$. This agrees that the volatility shocks are totally continual. The coefficient of the squared residuals is positive and statistically significant showing that strong GARCH effects are evident for our data. Higher value of β_1 shows a long memory in the variance. Thus the AR(1)-GARCH(1,1) model displayed in equation (12) adequately represents the residuals.

4.3 Forecasting of Jerusalem's Temperature Data

The forecast models are predestined in terms of their capability to predict the future values. In order to compare forecasting performance of different models, many statistical measures can be used for this purpose. The most widely reported error measure is RMSE. The other frequently used measures are MAE and MAPE.

The RMSE and MAE are used for comparison of model accuracy. Smaller values indicates better model performance. Both of MAE and RMSE together can be used to diagnose the variation in the errors of predicted values. Whenever value of MAE is less than RMSE, there is a variation in the errors. The three prediction error estimators are defined as follows (see [21, 22]):

- Mean Absolute Error (MAE):
  The MAE is used to measure how close forecasts or predictions are to the actual data and is given by:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|X_t - \widehat{X}_t\right| \tag{13}$$

- Root Mean Square Error (RMSE):
  The RMSE is a quadratic formula which measures the differences between values predicted by hypothetical model and observed values and is given by:

$$\text{RMSE} = \sqrt{\frac{1}{n}\left(X_t - \widehat{X}_t\right)^2} \tag{14}$$

- Mean absolute percentage error (MAPE):
  The MAPE measures the relative dispersion of forecast errors and is defined by the following formula:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|X_t - \widehat{X}_t\right|}{X_t} * 100 \tag{15}$$

Table 6 shows the 12-month-ahead prediction results of Jerusalem's temperature along with their corresponding standard errors inside the brackets ( ) for the year 2013 in respect of the above fitted models.

Table 6: The 12-Month-Ahead Prediction Results of Jerusalem's Temperature for the Fitted Models.

Months	**Observed Values**	**Forecasts By**	
		SARIMA	**AR(1)-GARCH(1,1)**
Jan 2013	10.019355	9.763649 (1.512443)	9.843522 (3.509865)
Feb 2013	11.903571	10.585964 (1.571673)	10.964575 (3.514317)
Mar 2013	15.219355	13.089481 (1.588858)	14.006754 (3.518723)
Apr 2013	16.393333	16.786828 (1.598088)	16.699214 (3.523083)
May 2013	21.829032	20.805272 (1.603809)	21.284091 (3.527397)
Jun 2013	22.983333	23.305218 (1.607454)	23.181534 (3.531667)
Jul 2013	23.245161	24.808717 (1.609788)	24.192735 (3.535893)
Aug 2013	24.374194	24.793023 (1.611285)	24.653073 (3.540075)
Sep 2013	22.270000	23.491077 (1.612245)	23.071035 (3.544213)
Oct 2013	18.958065	21.038515 (1.612859)	20.148405 (3.548309)
Nov 2013	17.410000	15.933020 (1.613248)	16.897626 (3.552363)
Dec 2013	9.706452	11.602616 (1.613479)	10.624668 (3.556375)

Figure 11 displays the forecast value for the monthly mean temperature of year 2013 using AR(1)-GARCH(1,1) model. In Figure 11 the solid line shows the forecasted value whereas the dashed lines are forecast temperature with ±2 standard deviations.

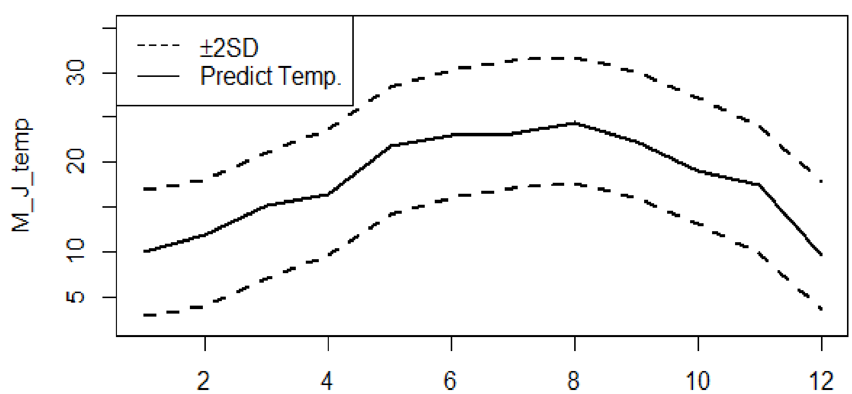

Figure 11: Forecast of Monthly Mean Temperature of Year 2013 by AR(1)-GARCH(1,1)

The actual and predicted Monthly mean temperature records by AR(1)-GARCH(1,1) model are plotted in Figure 12. From Figure 12 it can be shown that the trend of prediction temperature records follows the actual for the 12 months closely.

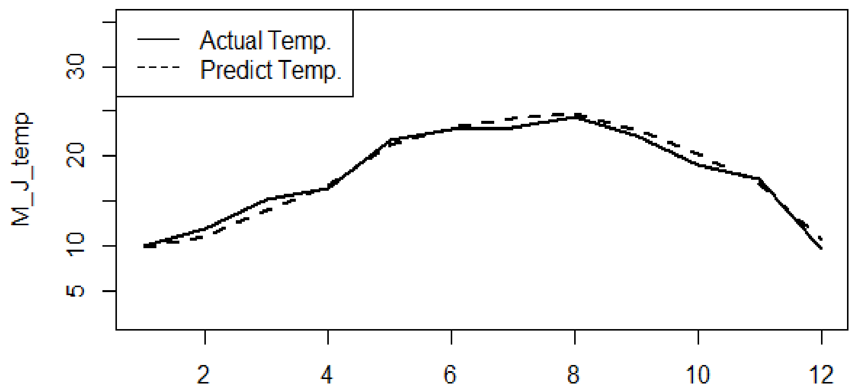

Figure 12: The Plot of Actual Temperature versus Forecast Temperature

After obtaining the 12-month-ahead prediction results reported in Table 6 above, we need to compute the three statistical estimators of equations (13), (14) and (15). The values of these measures of forecast errors are reported in Table 7. Based on lower values of MAE, RMSE and MAPE, AR(1)-GARCH(1,1) is a more convenient forecast model for Jerusalem's temperature values since it performs better than $\text{ARIMA}(2,0,1)(2,1,1)_{12}$.

Table 7: The Computed MAE, RMSE and MAPE for the Fitted Models.

Model	MAE	RMSE	MAPE
SARIMA	1.821726	1.349713	6.578911
AR(1)-GARCH(1,1)	0.5793938	0.7611792	3.744483

5 Conclusion

In this paper, the model that has been selected for forecasting Jerusalem's temperature values is AR(1)-GARCH(1,1). The model performs better than $\text{ARIMA}(2,0,1)(2,1,1)_{12}$ because of its capability to captivate the volatility or the time-varying conditional variances or errors. In the current study, AR(1)-GARCH(1,1) was achieved to be a better model than $\text{ARIMA}(2,0,1)(2,1,1)_{12}$ in forecasting Jerusalem's temperature because the values for the statistical estimators MAE, RMSE and MAPE using this model were smaller than those computed using $\text{ARIMA}(2,0,1)(2,1,1)_{12}$ model.

Future studies in this area can also use other types of GARCH models namely, Integrated GARCH (IGARCH) and Exponential GARCH (EGARCH).

R-programming is well appropriate for modeling and predicting of temperature data in this case.

References

[1] Chatfield, C. *Time Series Forecasting*. New York: Chapman and Hall. 2001.

[2] Cryer, J. D. and Chan, K. *Time Series Analysis With Applications in R.* 2nd Edition. New York: Springer. 2008.

[3] Shumway, R. H. and Stoffer, D. S. *Time Series Analysis and Its Applications.* New York: Springer. 2011.

[4] Engle, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica.* 1982. 50: 987-1007.

[5] William, W. S. W. *Time Series Analysis : Univariate and Multivariate Methods.* New York: Pearson Education. 2006.

[6] Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics.* 1986. 31: 307-327.

[7] Gouriéroux, C. *ARCH Models and Financial Applications.* New York: Springer-Verlag. 1997.

[8] R Core Team (2014). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.http://www.R-project.org/.

[9] Wuertz, D. and Chalabi, Y. with contribution from Michal Miklovic, Chris Boudt, Pierre Chausse and others . fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling. R package version 3010.82. http://CRAN.R-project.org/package=fGarch. 2013.

[10] TOL, R. S. J. Autoregressive Conditional Heteroscedasticity in Daily Temperature Measurements. *Environmetrics.* 1996. 7: 67-75.

[11] Benth, F. E. and Benth, J. S. *Modeling and Pricing in Financial Markets for Weather Derivatives.* New York: World Scientific. 2013.

[12] Franses, P. H., Neele, J. and Dijk, D. Modeling asymmetric volatility in weekly Dutch temperature data. *Environmental Modeling & Software.* 2001. 16: 131-137.

[13] Campbell, S. D. and Diebold, F. X. Weather Forecasting for Weather Derivatives. *Journal of the American Statistical Association* 2005. 469: 100.

[14] Box, G. E. and Jenkins, G. M. *Time Series Analysis Forcasting and Control.* 2nd Edition. New York: Holden-Day. 1976.

[15] Brockwell, P. J. and Davis, R. A. *Time Series: Theory and Methods.* 2nd Edition. New York: Springer Series in Statistics. 1991.

[16] Tan, Z., Zhang, J., Wang, J. and Xu, J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. *Applied Energy.* 2010. 87: 3606-3610.

[17] Francq, C. and Zakoïan, J. *GARCH Models Structure, Statistical Inference and Financial Applications.* London: John Wiley & Sons Ltd. 2010.

[18] Giraitis, L., Kokoszka, P. and Leipus, R. Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory.* 2000. 16: 3-22.

[19] Hall, P. and Yao, Q. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica.* 2003. 71: 285-317.

[20] Huang, D. Wang, H. AND YAO, Q. Estimating GARCH models: when to use what?. *Econometrics Journal.* 2008. 11: 27-38.

[21] Zambrano-Bigiarini, M. hydroGOF. R Package version 0.3-8 (2014-02-04). http://cran.r-project.org/web/packages/hydroGOF/ 2014.

[22] Ravindran, A. R. and Warsing Jr., D. P. *Supply Chain Engineering: Models and Applications.* London: CRC Press Taylor & Francis Group. 2013.