Dhaka Univ. J. Sci. 60(2): 159-162, 2012 (July)

An Assessment of Renewable Energy in Bangladesh through ARIMA, Holt's, ARCH-GARCH Models

Fatema Tuz Jhohura and Md. Israt Rayhan

Institute of Statistical Research and Training (ISRT), Dhaka University, Dhaka-1000, Bangladesh

Received on 07. 01. 2011. Accepted for Publication on 01. 12. 2011.

Abstract

Forecasting of the Renewable Energy plays a major role in optimal decision formula for government and industrial sector in Bangladesh. This research is based on time series modeling with special application to solar energy data for Dhaka city. Three families of time series models namely, the autoregressive integrated moving average models, Holt's linear exponential smoothing, and the autoregressive conditional heteroscedastic (with their extensions to generalized autoregressive conditional heteroscedastic) models were fitted to the data. The goodness of fit is performed via the Akaike information criteria, Schwartz Bayesian criteria. It was established that the generalized autoregressive conditional heteroscedastic model was superior to the autoregressive integrated moving average model and Holt's linear exponential smoothing because the data was characterized by changing mean and variance.

Key words: Autoregressive integrated moving average, Holt's linear exponential smoothing, Autoregressive conditional heteroscedastic, Akaike information criteria, Schwartz Bayesian criteria.

I. Introduction

Renewable energy comes from natural resources such as sunlight, wind, rain, tides, and geothermal heat, which are renewable (naturally replenished). The main forms of renewable energy are solar energy, wind power, hydropower, geothermal energy, and biofuel. Among these, solar energy is being seriously considered for satisfying significant part of energy demand in Bangladesh, as in the world.

Solar radiation data are a fundamental input for solar energy applications such as photovoltaic systems for electricity generation, solar collectors for heating, solar air conditioning climate control in buildings and passive solar devices. The solar radiation data should be measured continuously and accurately over the long term. Time series analysis of solar radiation data is useful in predicting long-term average performance of solar energy systems. Thus, many studies have been carried out on this subject.

A comparative study on Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroscedastic (GARCH) models in forecasting natural resource (crude oil) prices was proposed by Yaziz et al. (2011)[1], where GARCH is found to be a better model than ARIMA model. Akincilar et al. (2011)[2] found that GARCH model exhibited superior forecasting efficiency to ARIMA and Holt's smoothing models. Guo et. al. (2010)[3] used both ARMA and GARCH models and its extension for forecasting wind speed. Bulut and Büyükalaca (2007)[4] have developed a simple model for estimating the daily global radiation.

In the study of Sulaiman et al. (1997)[5], the Box-Jenkins approach was applied to daily solar radiation data from four different locations in Malaysia. A time series is a series or sequence of data points measured typically at successive times. These data points are commonly equally spaced in time (Chatfield, 2004)[6].

ARIMA model has been studied extensively to time series analysis and forecasting. They were popularized by George E. P. Box and Gwilym M. Jenkins in the 1976s[7], and their names have frequently been used synonymously with general ARIMA models. Exponential Smoothing is a very popular scheme to produce a smoothed time series. Whereas in moving averages the past observations are weighted equally, exponential smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weight in forecasting than the older observations. The related literature on exponential smoothing can be found in Brown (1962)[8], Holt et al. (1960)[9], Gardner (1985)[10], Makridakis (1982)[11].

The autoregressive conditional heteroscedastic (ARCH) models, with its extension to generalized ARCH, (GARCH) models as introduced by Engle (1982) [12-13] and Bollerslev (1986)[14] respectively, accommodates the dynamics of conditional heteroscedasticity. Campbell et al. (1997)[15] argued that it is both logically inconsistent and statistically inefficient to use and model volatility measures that are based on the assumption of constant variance over some period when the resulting series moves or progress through time. The ARCH-GARCH modeling considers the conditional error variance as a function of the past realization of the series.

II. Data and Methodology

This study was conducted for searching a univariate forecasting model for the solar radiation (in kWh per square meter) of Dhaka city. The data possess 36 observations during the time period 2003-2005. This secondary data were collected from Renewable Energy Research Centre (RERC), Dhaka University (DU) that covered the period January 2003 to December 2005 available from the source. The dataset provide monthly solar radiation rate of Dhaka city.

*Author for correspondence

Box- Jenkins Methodology for ARIMA Model

An autoregressive model of order *p is* conventionally classified as AR (*p*). A moving average model with *q* terms is classified as MA (*q*). If the object series is differenced *d* times to achieve stationarity, the model is classified as ARIMA (*p*, *d*, *q*), where the symbol ''I" signifies "integrated".

The equation for the ARIMA (*p, d, q*) model is as follows:

Y_t=C+$f_1 Y_{t-1}$+$f_2 Y_{t-2}$+…+$f_p Y_{t-p}$+e_t-$\theta_1 e_{t-1}$-$\theta_2 e_{t-2}$-…-$\theta_q e_{t-q}$

Or in backshift notation

(1-f_1B-$f_2 B^2$-…-$f_p B^p$) Y_t= c + (1-θ_1B-$\theta_2 B^2$-…-$\theta_q B^q$) e_t

where c denotes constant, f denotes i^{th} autoregressive parameter, θ_j denotes j^{th} moving average parameter, e_t denotes the error term at time t, B^k denotes the k^{th} order backward shift operator.

Box and Jenkins proposed a Methodology that consists three phases are known as Box-Jenkins Methodology. Phases are identification, Estimation of Diagnostic checking, and Application. Identification consists four steps (Step1: Stability in variance, Step2: Checking the Stationarity, Step3: Obtaining Stationarity, Step4: Model Selection) and estimation of Diagnostic checking consists three steps (Step1: Estimating the Parameters, Step2: selection of the Best Model, Step3: Diagnostic Checking)

Holt's Linear Exponential Smoothing Methodology

In exponential smoothing procedures the weights assigned to observation are exponentially decreased, as the observations get older. The forecast for Holt's linear exponential smoothing is found using two parameter smoothing constants, α and β with values between 0 and 1, and three equation:

L_t = α+(1-α) $(L_{t-1} + b_{t-1})$

b_t= $\beta(L_t$-$L_{t-1})$ + (1-β)b_{t-1}

F_{t+m} = L_t +$b_t m$

Here L_t denotes an estimate of the level of the series at time *t* and b_t denotes an estimate of the slope of the series at time *t* and F_{t+m} is forecasted value at period $t+m$.

To perform a Holt's linear exponential smoothing method we have to go through the three steps, namely Step 1: Initialization, Step 2: optimization, Step 3: Forecasting.

The initialization process requires two estimates-one to get the first smoothing valued for L_1 and the other to get the trend b_1.

One alternative is to set L_1= Y_1 and b_1= Y_2- Y_1 or b_1= (Y_4- Y_1) /3

Another alternative is to use least squared regression on the first few values of the series for finding L_1 and b_1 .

Methodology for GARCH Model

The GARCH (p, q) is an generalization of GARCH (1,1) with p as the autoregressive lag and q is the moving average lag. Formally a process { y_t} is GARCH (p, q) if

y_t = $\sigma_t \varepsilon_t$

σ_t^2 = α_0+$\sum_{i=1}^{q} \alpha_i\, y_{t-i}^2$+$\sum_{j=1}^{p} \beta_j\, \sigma_{t-j}^2$

= α_0+α (B) y_t^2+β (B) σ_t^2

where ε_t is Gaussian white noise while α (B) and β (B) are polynomials in the backshift operator.

Assuming the GARCH (p, q) process is second order stationary, that is

Var (y_t) = $E(y_t^2)$ = $\frac{\alpha_0}{1-\sum_{i=1}^{q}\alpha_i-\sum_{j=1}^{p}\beta_j}$

The autocovariance of a GARCH (p, q) model for k $\geq$ 1 where k is the lag is

$E(y_t, y_{t-k})$ = 0

Considering writing y_t^2 in terms of $v_t = y_t^2$-σ_t^2 yields

$y_t^2 = \sigma_t^2$+v_t

= α_0+$\sum_{i=1}^{q}\alpha_i\, y_{t-i}^2$

+$\sum_{j=1}^{p}\beta_j\, y_{t-i}^2 - \sum_{j=1}^{p}\beta_j\, v_{t-j}$+ v_t

Now let $m = max\,(p, q)$, then

y_t^2= α_0+$\sum_{i=1}^{m}(\alpha_i + \beta_i)\, y_{t-i}^2 - \sum_{j=1}^{p}\beta_j\, v_{t-j}$+ v_t

where $\alpha_i = 0$ for $i > q$ and $\beta_j = 0$ for $i > p$. Thus the equation of y_t^2 has an ARMA (m, p) representation.

Therefore assuming $y_1, y_2,$ …, y_q and σ_1^2, σ_2^2,…, σ_p^2 are known, the conditional maximum likelihood estimates can be obtained by maximizing the conditional log-likelihood given by

l = ln $f(y_{q+1}, \dots, y_t, \sigma_{p+1}^2, \dots, \sigma_t^2 \,|\, \theta, y_1, y_2,$ …, $y_q, \sigma_1^2, \sigma_2^2$,…, σ_p^2)

=$-\frac{1}{2}\sum_{t=m+1}^{T}$ ln$(2\pi\sigma_t^2)$ $-\frac{1}{2}\sum_{t=m+1}^{T}\frac{y_t^2}{\sigma_t^2}$

With θ= $(\alpha_0, \dots, \alpha_q, \beta_1, \dots, \beta_p)$ and $m = max(p, q)$.

Then the 1-step ahead volatility forecast is given by

y_t^2(1)= $E(y_{t+1}^2|y_t)$

= $\alpha_0 + \sum_{i=1}^{m}(\alpha_i + \beta_i)E(y_{t+1-i}^2|y_t)$

$-\sum_{i=1}^{p}\beta_i(v_{t+1-i}|y_t)$

Furthermore, the l-step ahead forecast of the conditional variance in a GARCH (p, q) model is given by

$$\sigma_t^2(l) = E(y_{t+l}^2 | y_t)$$
$$= \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \beta_i) E(y_{t+l-i}^2 | y_t)$$
$$- \sum_{i=1}^{p} \beta_i (v_{t+l-i} | y_t)$$

III. Analysis and Results

Estimation of an ARIMA Model

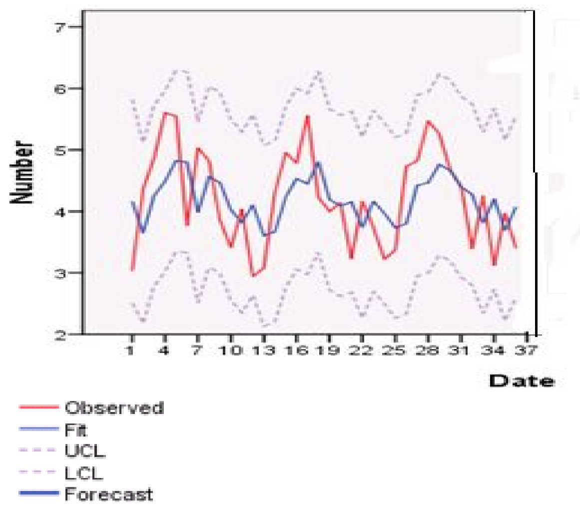

Fig. 1. The Time-Plot of Solar energy

From the observed time plot we found that over the time period of study solar energy is not stationary in the mean and variance. The data are not stationary and there exist seasonality which we observed from the autocorrelation function (ACF) of solar and the corresponding correlogram. Than the data set does not follow a white noise series since the values of Ljung-Box test are significant different from a null set. Now we have to take the differencing method.

This study uses the Akaike information Criterion (AIC) to choose the best model among the class of plausible models. The model which has the minimum AIC value is our model of interest.

Table. 1. The AIC values for ARIMA (p, d, q) model

	q = 0	q = 1	q = 2	q = 3
p = 0	115.65	91.54	91.87	93.73
p = 1	102.19	91.80	93.80	95.73
p = 2	97.08	93.79	93.28	95.36
p = 3	98.38	95.59	97.52	96.95

From Table-1, we found that the AIC value for the model ARIMA (0, 2, 1) is minimum. This model includes no AR coefficients and one MA coefficient and takes the form:

$\Delta^2 y_t = e_t - \theta_1 e_{t-1}$

Or, $y_t - 2y_{t-1} + y_{t-2} = e_t - \theta_1 e_{t-1}$

Or, $y_t = 2y_{t-1} - y_{t-2} + e_t - \theta_1 e_{t-1}$

Now we have to test the signification of the parameter.

Table. 2. The significance test of the parameters of ARIMA (0, 2, 1)

Coefficient	Parameters	Standard error	Z value	Decision
θ_1	-1.00	0.09	-10.88	Significant

From the above table we see that p values corresponding to the coefficient θ_j less than 0.05, which leads to the conclusion that this parameter is significant.

Estimation of Holt's Linear Smoothing Model

Through computer programming, we obtain the set of value α and β as smoothing parameters which gives the minimum Mean Square Error (MSE). Following this procedure we obtain the value of α=0.800 and β=0.000 which minimizes the MSE.

Thus, our Holt's linear model becomes,

L_t = 0.800Y_t+(1-0.800) $(L_{t-1} + b_{t-1})$

= 0.800Y_t+0.2$(L_{t-1} + b_{t-1})$

b_t = 0.000$(L_t - L_{t-1})$ + (1-0.000) b_{t-1}

$= b_{t-1}$

And $F_{t+m} = L_t + b_t m$ where m=1,2,3,…

Using m=1. We get

$F_{t+1} = L_t + b_t$

This model is used to forecast the future values.

Estimation of the ARCH-GARCH Model

From model selection criteria, the model which has minimum Akaike Information Criteria (AIC), and Schwartz Bayesian Criteria (SBC) value, is the best model.

Table. 3. Comparison of ARCH-GARCH models

Model	AIC	SBC
GARCH(1, 1)	212.42	217.18
GARCH(1, 2)	**212.41**	**217.17**
GARCH(2, 1)	212.88	217.63
GARCH(2, 2)	214.88	221.21

Therefore the GARCH (1, 2) in table 3 is chosen to be the most appropriate; hence we proceed to estimate the associated model parameters.

Table. 4. Comparisons of the ARIMA, HOLT'S and ARCH-GARCH Models

Measures of error	**ARIMA model**	**Holt's linear model with α=0.80 and β=0.00**	**GARCH (1,2) model**
Mean Error(ME)	-0.17	0.03	4.44×10^{-6}
Mean Absolute Error(MAE)	0.66	0.67	0.66
Sum Square Error(SSE)	23.54	22.31	22.29
Mean Percentage Error(MPE)	-5.64	-1.61	-3.66
Mean Absolute Percentage Error(MAPE)	16.50	16.42	16.41

From the above table we see that for all measures of errors, GARCH (1,2) model gives the better results over ARIMA (0,2,1) model, and Holt's linear smoothing model on solar energy data.

IV. Conclusion

Generation of typical solar radiation is very important for the calculations concerning many solar applications. It is one of the most important economic sectors of a country, like Bangladesh. Day by day, the infrastructural development is improving and that's why the solar radiation of this sector also shows increasing trend. This study focuses the extensive understanding of the theory of time series analysis and its application to a real life situation. Here although the ARIMA model captured these variations, the need to transform the data to stationary makes the model rely on some restricted assumptions resulting the GARCH model being superior. The GARCH model fits the data well. To formulate future development plan for this sector, it is essential to know the previous condition and also see the future trend. In this study, forecasting is done by using some sophisticated statistical tools so that the government and policy makers can easily realize about the future contribution of the solar radiation of Dhaka city to the overall solar radiation and could take initiatives to how to improve this sector.

....

1. Yaziz, S.R., M. H. Ahmad, L. C. Nian and N. Muhammad, (2011). A Comparative Study on Box-Jenkins and Garch Models in Forecasting Crude Oil Prices. *Journal of Applied Sciences,* **11**: 1129-1135.
2. Akincilar, A., I. Temiz, and E. Sahin, (2011). An Application Of Exchange Rate Forecasting In Turkey. *Gazi University Journal of Science,* **24(4)**:817-828.
3. Guo, Z, Y. Dong, J. Wang, and H. Lu, (2010). The Forecasting Procedure for Long-Term Wind Speed in the Zhangye Area. *Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2010, Article ID 684742, 17 pages, doi:10.1155/2010/684742.*
4. Bulut, H. and O. Büyükalaca, (2007). Simple model for the generation of daily global solar radiation data in Turkey. *Applied Energy* **84**: 477-491.
5. Sulaiman, M.Y., O. Hlaing, M. Wahab, (1997). Analysis of residuals in daily solar radiation time series. *Renewable Energy* **11**: 97-105.
6. Chatfield, C., (2004). The analysis of time series: An introduction. 3rd edition, Chapman and Hall text in statistical science, London.
7. Box, G, G.M. Jenkins and C.R. Gregory, (1976). Time series analysis: Forecasting and control. 4th Edition. Wiley series, probability and statistics, Prentice Hall Inc., London.
8. Brown, R.G., (1962). Smoothing, forecasting and prediction of discrete time series. Prentice Hall Inc., London.
9. Holt, C., F. Modigliani, J. Muth, H. Simon (1960) Planning Production, Inventories, and Workforce. Englewood Cliffs, NJ: Prentice-Hall.
10. Gardner, E. S., (1985). Exponential smoothing: The state of the art. *Journal of Forecasting,* **4**:1-28.
11. Makridakis, S. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *Journal of Forecasting,* **1**: 111-153.
12. Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of variance of the United Kingdom inflation. *Econometrica,* **50(4)**: 987-1006.
13. Engle,R.,(1982) The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspective* , **15(4)**: 157-168.
14. Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journal of econometrics ,* **51** : 307-327.
15. Campbell, J.Y., A.W. Lo and A.C. Mackinlay, (1997). Modelling daily value-at-risk using realized volatility and ARCH type models, *Journal of empirical finance* , **11(3)**: 379-398.