Application of Analytics using R Programming for Forecasting Day-ahead Electricity Demand

ChandraBhushan Kesavabhotla[1], Professor V. V. **HaraGopal**[2], Professor A. **Vinay Babu**[3]

[1]Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India

[2]Department of Statistics, Osmania University, Hyderabad, India

[3]Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Hyderabad, India

Abstract

Analytics technologies are becoming increasingly important for the organizations to mine and get insights into large amount of structured and unstructured data, and also to enable effective decision making for the management. One of the challenges for power industry is to estimate the Day-ahead power demand specifically to predict the peak and low demand periods. In this paper, the modeling techniques of analytics for forecasting are discussed. This paper investigates application of ARMA and GARCH modeling techniques to fit the historical data and estimate the coefficients to predict the Day-ahead electricity demand. R-Programming is used to fit the models.

***Keywords*:** ARMA, ARCH, Electricity Demand, Estimation, Forecasting, GARCH, Modeling, R-Code.

1. Introduction

The demand forecasting in electricity has become one of the most critical economy factors for many countries especially for those with scarce energy resources or high costs of generation of electricity. The power industry requires forecasts for both short term such as hours or days ahead and long term like 5,10, 15 years ahead. The short-term forecasts are becoming increasingly important to manage the peak and low demand days efficiently and also for planning periodic monthly, quarterly, and yearly electricity demand. A unique feature of electricity is that unlike the manufactured goods which can be stored and utilized, electricity is not storable which makes the supply and demand management very complex.

There are various factors that influence daily electricity demand on a power system such as seasons, weather conditions, and other local variables like public holidays, economy situation (economic boom or recession period) in any country. Many of these factors are different for each country depending on the geographical location, culture, and the country and/or state dependency on the resources of power generation. The electricity demand planning also needs to accommodate the additional demand due to the growth in new industries, residential and agriculture needs.

The demand prediction methods and techniques are chosen based on the scenario, "One model fit for all" may not work for multiple scenarios across the countries as the conditions may not be same. Sigauke and Chikobvu [1] predicted the daily peak electricity demand in South Africa using volatility forecasting models. They designed a hybrid regressive SARIMA (Seasonal Autoregressive Integrated Moving-Average) and GARCH (Generalized Autoregressive Heteroskedasticity) models to predict the daily demand with a Mean Absolute Percent Error (MAPE) of 1.42%. Ching-Lai and Simon J Watson [2] worked on ARIMA (Autoregressive Integrated Moving Average) and GARCH models to forecast the daily load of electricity in UK with Mean Absolute Percent Error of 1-3% for each month and analyzed the influence of the climate change in the power demand. Nowicka-Zagrajck and Weron [3] studied electricity load patterns of California state and proposed ARMA (Autoregressive Moving Average) time series with hyperbolic noise. Reinaldo and Javier[4] have studied electricity prices in Spain and California and proposed GARCH model to predict the volatility of Day-ahead electricity prices. Chengjun Li, Ming Zhang[5] compared the GARCH and ARMA models to predict the hourly electricity prices in California market. Ramakrishna and et al [6] proposed a Neural Network model to forecast monthly electricity load. They compared the ANN(Artificial Neural Networks) model with SARIMA. Some of the researchers indicated [7, 8, 9] that ANN may not always outperform as compared to other forecast models due to the challenges in validations, and accurate systematic testing.

The objective of this paper is to study the historical daily electricity load of Andhra Pradesh state, India and forecast the Day-ahead electricity demand using ARMA and GARCH models. R-Programming is used with step by step process to fit the models. The data considered in this paper is from Andhra Pradesh Transmission Company, India. The R-Project is an Open Source Software. The R Notes [11] from R-Project is referred for R Programming syntax.

2. Data Preparation

For the purpose of study, the data population of 375 observations from Daily Electricity Load of Andhra Pradesh State, India between 2005 and 2006 is considered. To identify and fit the model, 365 daily electricity observations of 2005 year are used for sampling and forecasting 10 days-ahead electricity demand. 10 observations of 2006 are considered for comparing the predicted electricity demand. The fitted model can be applied to all the other years to validate the model. The preliminary data analysis is indicated in the Table 1, consists of the average daily load in each year,

year on year change in average daily load, minimum daily load in each year, and maximum daily load in each year.

Table 1: Data Analysis for 5 years (Power Load in Giga Watts)

	Average Daily Load	%Change YoY	Min Daily Load	Max Daily Load
Year 1	137.78		103.40	170.63
Year 2	152.99	11%	114.33	169.80
Year 3	161.79	6%	116.29	197.23
Year 4	178.18	10%	128.03	230.12
Year 5	210.49	18%	164.30	260.09

The trend in the Table 1 shows that the average day load, minimum, and maximum day load increased across the years.

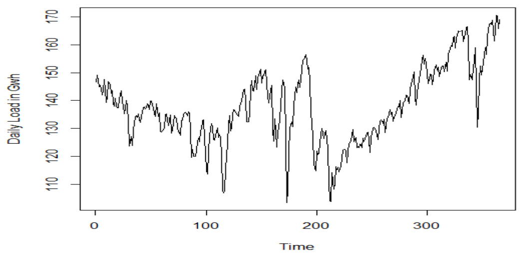

Figure 1: Time Series plot of the Daily Electricity Load pattern.

The time series of daily electricity load of 2005 is plotted in Figure 1. The data exhibits peaks and lows with increasing and decreasing trend of variances. There is no consistent trend (upward or downward) over the entire time span considered. The series appears to move up and down. In the case of annual data, the seasonal factors may not have significant influence on the variances. There are no outliers. It may be difficult to conclude from the above whether the variance is constant or not and needs to be tested with a suitable model. The data in Giga Watt is captured in a CSV file. The data is loaded into R-Project from a flat file and converted to time series as follows:

```
>read csv(dailyload,"dailydata.csv")
>dailyload <-ts(dailyload)
```

The step-wise methodology used in this paper is to identify a model, estimate the coefficients, fit the sample data and forecast the predicted load.

3. Time Series Model -ARMA

3.1 Identification of the model

The general ARMA statistical model is used to describe a time series that evolves over time.

The ARMA(p, q) process for $\{x_t\}$ time series is represented as

$$x_t = a_0 + a_1 x_{t-1} + \ldots + a_p x_{t-p} + b_1 e_{t-1} + \ldots + b_q e_{t-q} + e_t \quad \ldots(1)$$

where a_0 is set to zero if no intercept is included, p is the autoregressive process order and q is the moving average process order.

The 'Best-Fit' p and q order values of ARMA model can be determined by using the smallest AIC (Akaike Information Criterion) through the following R-Code:

```
>r.order <- c(0,0)
>smaller.aic <- Inf
>k <- 0
>l <- 0
>for (k in 0:2) for (l in 0:2) {
r.arma <- arima(dailyload, order = c(k,l))
smallest.aic <- AIC(r.arma)
if (smallest.aic < smaller.aic) {
r.order <- c(k,l)
r.arma <- arima(dailyload,order = r.order)
smallest.aic <- AIC(r.arma)
smaller.aic <- smallest.aic
}
}
>r.order
>acf(resid(r.arma))
```

The above code compares the AIC of each iteration with the previous iteration of ARMA function with changed p and q values. The output of the above code is 1, 1 with smallest AIC. ARMA (1, 1) is the identified model to estimate the coefficients and fit the data to the model.
The Autocorrelation Function plot of the model in Figure 2 depicts that the Autocorrelation is decayed rapidly to zero after initial lags.

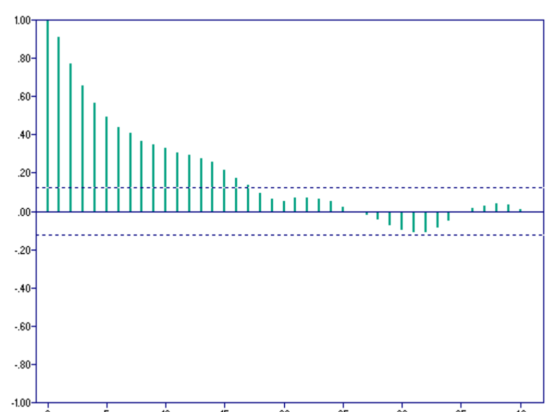

Figure 2: ACF plot of the observations.

3.2 Estimation

The Estimation of coefficients for the identified model ARMA (1, 1) is done as follows:

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 1, November 2012
ISSN (Online): 1694-0814
www.IJCSI.org

```
>armafit <- arma(dailyload,order=c(1,1))
>summary(armafit)
>acf(armafit$residuals)
```

The output of ARMA fit is shown in the Table 2, it consists of coefficients, p-values. ACF and PACF of residuals are plotted in the Figure 3 and Figure 4 respectively.

Table 2: ARMA (1, 1) Model Estimation

```
arma(x = dailyload, order = c(1, 1))
Model: ARMA(1,1)
Residuals:
    Min      1Q  Median      3Q     Max
-18.2413 -1.9954  0.4046  2.3832  9.0822
Coefficient(s):
          Estimate  Std. Error  t value    Pr(>|t|)
ar1        0.9285    0.0213    43.594    < 2e-16 ***
ma1        0.3161    0.0517     6.115    9.68e-10 ***
intercept  9.9109    2.9474     3.363    0.000772 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Fit:
sigma^2 estimated as 16.64,  Conditional Sum-of-Squares = 6039.01,  AIC = 2068.06
```

Both AR and MA p-values are significant at 5% level. This implies the first order coefficients of AR and MA are sufficient to fit the model to the data.

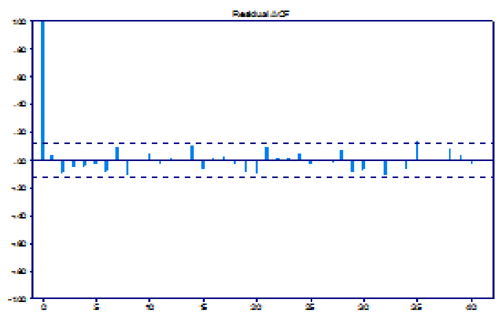

Figure 3: ACF plot of residuals of Model ARMA (1, 1)

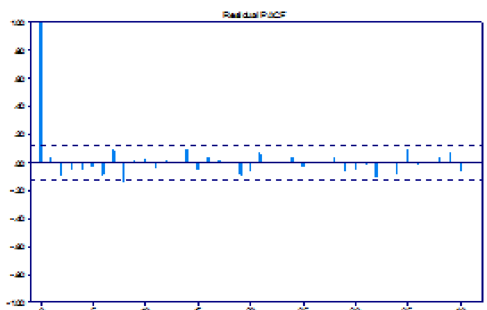

Figure 4: PACF plot of residuals of Model ARMA (1, 1)

There are no significant peaks in ACF and PACF plots, implies that there are no autocorrelations. The ACF plot of squared residuals is shown Figure 5.

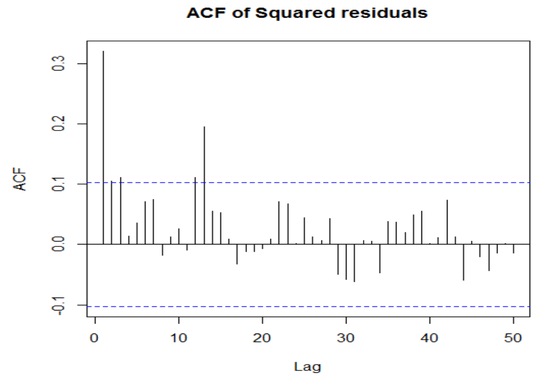

Figure 5: ACF plot of Squared Residuals

Except at the initial lags, all sample autocorrelations are within the 95 % confidence bounds. The residuals appear to be random after initial spikes. Ljung-Box Test is used to test the randomness on residuals and test is carried out for the lags till p-value 5% significance level. This can be achieved through the following R-Code:

```
>j <- 1
>pvalues <- array(1:10)
>lj_statistic <- array(1:10)
>lj_df <- array(1:10)
>for(k in 1:10) {
>ljung <- Box.test(armafit$residuals,lag=k)
>lj_chisquare[k] <- ljung$statistic
>lj_df[k]  <- ljung$parameter
>pvalues[k] <- ljung$p.value
>if (pvalues[k] > 0.05)
     j <- j + 1
  }
>lj_df
>lj_chisquare
>pvalues
```

The output of Ljung-Box test is shown in the Table 3.

Table 3: Ljung-Box Test on the Residuals series

Lag	Chi-Square	p-value
1	0.01237632	0.9114
2	1.22016414	0.5433
3	4.36113619	0.2250
4	4.64129359	0.3261
5	4.72277910	0.4506
6	7.54552729	0.2733
7	21.32018241	0.00332*
8	23.24596515	0.00306*
9	23.32674959	0.00550*
10	23.34491952	0.00954*

Ljung-Box test has given p-values significant at 5% from lag 7 onwards with a statistic value of 21.79 and degrees freedom of 22. This implies that there are some autocorrelations in the initial lags and these vanish after lag 6. The ARMA(1,1) is adequate to fit this data. The number of predictions ahead for short duration is more accurate than that of long duration ahead in this model which is good enough for forecasting the Day-ahead electricity load for weekly planning.

3.3 Forecasting of Day-ahead Electricity Demand

The ARMA(1,1) model equation with the estimated coefficients is

$$x_t = 9.91 + 0.9285x_{t-1} + 0.3161\, e_{t-1} + e_t \qquad \ldots(2)$$

The R-Code to forecast 10 Day-ahead demand is as follows:

```
> f_values <- forecast(armafit, 10)
> plot(f_values, main="Forecasting of Day-ahead Electricity Demand")
```

The predicted values with 80-95% confidence bounds are captured in Table 4 along with the observed values of 2006 for comparison.

Table 4: ARMA(1,1) predicted values

10 Day Forecast						
Day	Forecast	Lo 80	Hi 80	Lo 95	Hi 95	Observed
366	168.8871	163.6738	174.1005	160.9140	176.8603	166.977
367	166.6858	158.3663	175.0052	153.9623	179.4092	164.806
368	164.6468	154.3865	174.9071	148.9551	180.3385	164.351
369	162.7583	151.0874	174.4291	144.9092	180.6073	167.591
370	161.0091	148.2518	173.7664	141.4985	180.5196	166.415
371	159.3889	145.7685	173.0093	138.5583	180.2196	167.114
372	157.8883	143.5688	172.2078	135.9886	179.7881	169.797
373	156.4984	141.6053	171.3915	133.7214	179.2754	168.941
374	155.2111	139.8430	170.5792	131.7076	178.7146	166.530
375	154.0187	138.2545	169.7829	129.9094	178.1280	167.732

The predicted values of ARMA model are plotted in the time series in the Figure 6.

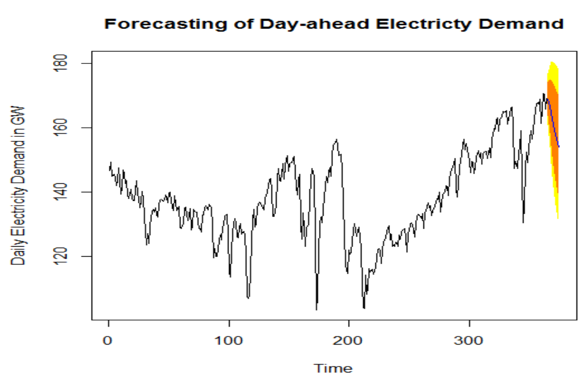

Figure 6: Predicted Day-ahead Electricity Demand

The above results show that ARMA(1,1) can be used for forecasting Day-ahead electricity load. However to predict for longer periods and to improve the accuracy within the confidence bounds, it is necessary to test non-linearity in the errors. The fitted model assumes constant variances i.e. homoskedasticity and the predictions become inaccurate with increase in the time lags , in reality the variances may not be constant. For example, electricity usage at nighttime may have more steady pattern compare to that of daytime usage i.e. the variances in nighttime may be lower than that during at daytime. Thus, the daytime usage plays a significant factor in the daily load variances. This behavior is called ARCH effect i.e. Autoregressive model with conditional heteroskedasticity (conditional on time). McLeod test can be used to analyze if there is any conditional heteroskedasticity in the residuals i.e. ARCH (Autoregressive Conditional Heteroskedasticity) effect.

3.4 Test for Heteroskedasticity in the residuals

McLeod-Li test checks for the presence of conditional heteroskedasticity by computing Ljung-Box test with the squared data or with the squared residuals from an ARMA model. The R-Code for the test is

```
> McLeod.Li.test(armafit, y=dailyload),main="McLeod Test")
```

The test with 25 lags has given a statistic value of 61.13, degrees of freedom 22 with a p-value of 0.00002 at 5% level of significance. These results are depicted in the Figure 7. Thus the null hypothesis stated "linear or constant variances" is rejected. There is an ARCH effect in the daily electricity load variances across the days.

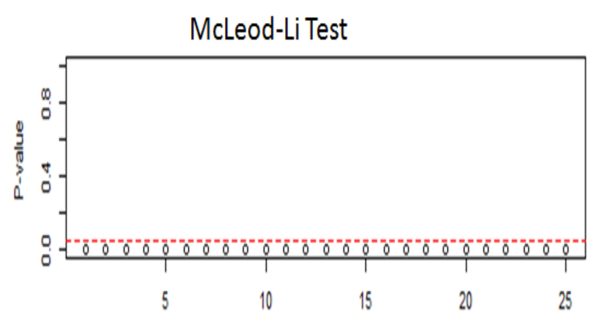

Figure 7: McLeod Test results on the residuals

Jarque-Bera Test is another popular test for testing simultaneously the normality and homoskedasticity of residuals in a time series data.

```
> jarque.bera.test(arimafit$residuals)
```

This test has given a statistics value = 137.4599, degrees of freedom = 2, and a p-value < 2.2e-16 at 5% level of significance. Thus, the null hypothesis stated "the residuals are normal and homoskedastic" is rejected.

Both the tests implied that there is a volatility in the daily load and the variances may have stochastic pattern. Thus, the heteroskedasticity of errors needs further analysis using a non-linear model such as GARCH where the variances themselves are modeled as AR(p) (Autoregressive process of order p) model. Since both positive and negative changes are observed in the daily electricity load this is another factor for considering a nonlinear model to analyze the errors.

4. GARCH Model for the variances

4.1 GARCH Model

GARCH, a non-linear model treats heteroskedasticity as a variance to be modeled[11], for each error term a prediction is estimated for the variance apart from correcting the deficiencies of least squares. The GARCH model measures the volatility like a standard deviation which is used in the decision making where the differential values are critical such as Electricity demand Day-ahead forecast, risk analysis, Electricity prices in the deregulated market, portfolio returns in the stock market, etc.

If a time series is defined as

$$y_t = \mu_t + \varepsilon_t \qquad \ldots\ldots (3)$$

where μ_t is the conditional mean, and ε_t is a residual process with mean zero.

The residuals are generated as $\varepsilon_t = \sigma_t z_t$

where z_t is an identically distributed process and independent with mean 0 and variance 1. Thus, $E(\varepsilon_t \varepsilon_{t+h}) = 0$ for all lags, h is not equal to zero and the residuals are uncorrelated.

If H_t denotes history of the process available at time t then the conditional variance of y_t is

$$Var(y_t \mid H_{t-1}) = Var(\varepsilon_t \mid H_{t-1}) = E(\varepsilon^2_t \mid H_{t-1}) = \sigma^2_t \quad ..(4)$$

Thus in the squared residual process , conditional heteroskedasticity is equivalent to autocorrelation.

The residual series is defined as

$$e_t = y_t - \hat{\mu}_t \qquad \ldots.(5)$$

The residuals are uncorrelated with mean zero when the autocorrelations in the series y_t are included in the conditional mean model. The residuals can still be serially dependent.

For Engle's ARCH test of autocorrelation in the squared residuals , the alternative hypothesis is given by the regression

$$H_\alpha : e^2_t = \alpha_0 + \alpha_1 e^2_{t-1} + \ldots. + \alpha_m e^2_{t-m} + u_t \qquad \ldots.(6)$$

where u_t is a noise error process. The null hypothesis is

$$H_0 : \alpha_0 = \alpha_1 = \ldots. = \alpha_m = 0 \qquad \ldots. (7)$$

Ljung-Box Q-Test can be conducted on the first m lags of the squared residual series to check for serial dependence, this is an alternative to Engle's ARCH effect.

4.2 Estimation of Coefficients

The GARCH(p, q) model is applied on the daily difference data to determine the p and q terms where p is the order of GARCH term σ^2 , standard deviation and q is the order of ARCH term ε^2, mean error. The R-Code to identify GARCH model & fit the model to the residuals data and predict the variances is given below:

```
>xlog <- diff(log(dailyload)) # generated for the series
> rgarchfit <- garchFit(~garch(1,1), xlog)  # Conditional Variance equation
>summary(rgarchfit)
>plot(rgarchfit)
```

The p and q values are changed iteratively in the above code. AIC, and LM (Lagrange Multiplier) ARCH Test p-values are captured for each iteration in the Table 5.

Table 5: GARCH model with p, q values

(p,q)	AIC	BIC	LM ARCH p-value
(1,1)	-4.212	-4.169	0.00614
(1,0)	-4.162	-4.13	3.6E-05
(2,0)	-4.19	-4.147	0.00231
(2,1)	-4.206	-4.153	0.00592
(2,2)	-4.213	-4.149	0.00438

There are two cases of p , and q : (1,1) and (2,2) with low AIC within 2 decimals in the Table 5. Either of these can be considered as the best but since BIC is lower for (1,1), the best values of p and q are taken as 1 and 1. The LM test for GARCH(1,1) has p-value <5% level of significance. Thus, the null hypothesis stated "no ARCH effect or no conditional heteroskedasticity in the residuals" is rejected.

The p-values of residuals and squared residuals of Ljung-Box Q statistics of GARCH(1,1) are given in the Table 6. The p-values of residuals are < 5% level of significance which indicates no presence of autocorrelations, however the squared residuals indicate some time dependency in the series and except at lag 15 where there are no autocorrelations.

Table 6: LB test of residuals using GARCH(1,1)

Standardised Residuals Tests:				
			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	222.5201	0
Ljung-Box Test	R	Q(10)	37.28246	5.05662e-05
Ljung-Box Test	R	Q(15)	48.04868	2.497476e-05
Ljung-Box Test	R	Q(20)	57.17543	1.930553e-05
Ljung-Box Test	R^2	Q(10)	5.613833	**0.8465973**
Ljung-Box Test	R^2	Q(15)	29.04761	0.01585906
Ljung-Box Test	R^2	Q(20)	30.60992	**0.06055534**
LM Arch Test	R	TR^2	27.68772	0.006144083

The time series plot of residuals in Figure 8 has a few peaks beyond the boundaries and it is exhibiting volatility clustering. The peaks of residuals match with the peaks of standard deviation plotted in Figure 9

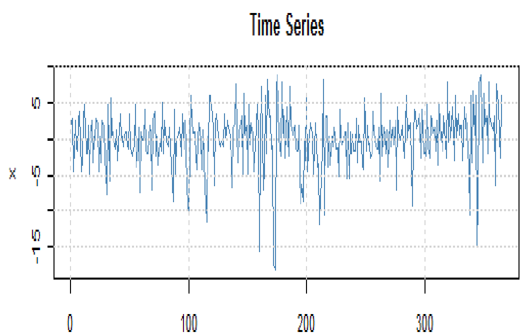

Figure 8 : Residuals plot of GARCH process.

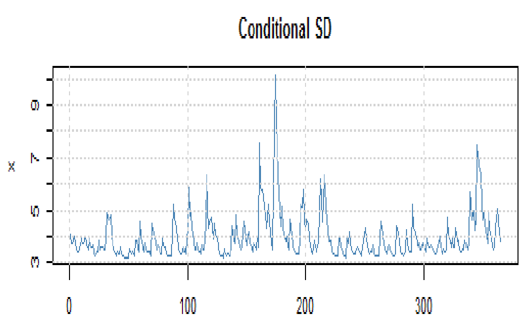

Figure 9: Standard deviation plot

The standard deviation of GARCH process indicates that there is a high volatility in the middle and at the end of the year. The QQ plot is shown Figure 10. The ACF of standardized residuals is shown in Figure 11. The ACF plot of squared residuals is shown in Figure 12.

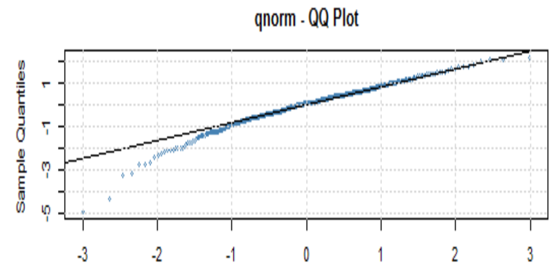

Figure 10: QQ plot of GARCH model

The QQ plot of GARCH indicates heavy tail.

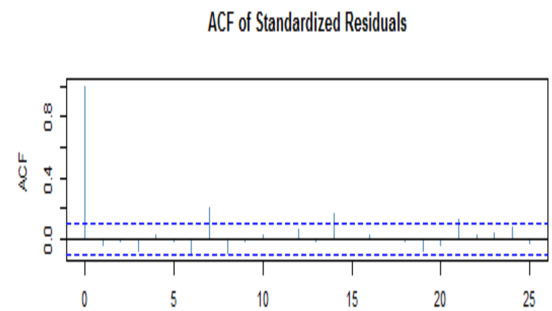

Figure 11: ACF plot of standardized residuals.

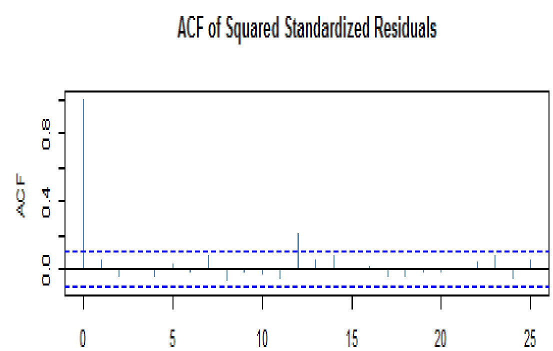

Figure 12: ACF plot of Squared Standardized Residuals

The ACF of residuals in the Figure 11 shows there are peaks i.e. autocorrelations but within ACF boundary. In the case of Squared residuals plot shown in Figure 12, some of these peaks are reduced.

The coefficients of GARCH(1,1) model are estimated using the 'garchFit' function. The coefficients estimated for GARCH(1,1) are shown in the Table 7.

Table 7 Coefficients of GARCH(1,1) model

	Estimate	Std. Error	t value	Pr(>|t|)
mu	1.110e-03	1.337e-03	0.830	0.406378
omega	2.651e-04	7.429e-05	3.569	0.000359 ***
alpha1	3.786e-01	9.531e-02	3.972	7.11e-05 ***
beta1	3.966e-01	1.045e-01	3.796	0.000147 ***

4.3 Forecasting of Variances

The GARCH(1,1) with the estimated coefficients is represented as :

$$\sigma^2_t = 0.00265 + 0.3786\ x^2_{t-1} + 0.3966\ \sigma^2_{t-1} \qquad \ldots(8)$$

The following R-Code uses the above coefficients for GARCH predictions.

```
>prd_garch <- predict(rgarchfit, 20)
> plot(prd_garch$standardDeviation, main="SD Forecast")
```

The output of R-Code is the Predicted variances, shown in the Table 8.

Table 8: Predicted values using GARCH(1,1)

	meanForecast	meanError	standardDeviation
1	0.0011	0.0271	0.0271
2	0.0011	0.0289	0.0289
3	0.0011	0.0302	0.0302
4	0.0011	0.0312	0.0312
5	0.0011	0.0319	0.0319
6	0.0011	0.0325	0.0325
7	0.0011	0.0329	0.0329
8	0.0011	0.0332	0.0332
9	0.0011	0.0335	0.0335
10	0.0011	0.0337	0.0337
11	0.0011	0.0338	0.0338
12	0.0011	0.0339	0.0339
13	0.0011	0.0340	0.0340
14	0.0011	0.0341	0.0341
15	0.0011	0.0341	0.0341
16	0.0011	0.0342	0.0342
17	0.0011	0.0342	0.0342
18	0.0011	0.0342	0.0342
19	0.0011	0.0342	0.0342
20	0.0011	0.0343	0.0343

The forecasted standard deviation(SD) is plotted in Figure 13. The variance in SD is high in the initial lags, it is flat i.e. constant after 20 lags.

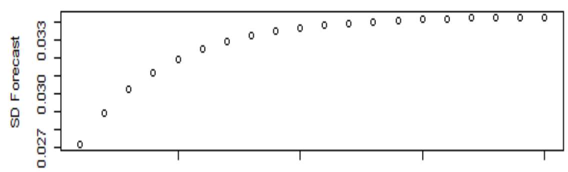

Figure 13: SD Forecast Plot from GARCH predictions.

The Mean Error (ME) plot of forecast in Figure 14 shows that it is in the range of 1-3.5%.

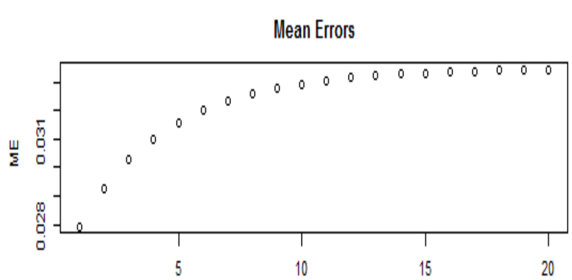

Figure 14: Mean Error of forecast

It is observed that the standard deviation and mean error values are equal in the predictions. Let us combine ARMA and GARCH for conditional mean and conditional variances and predicted the values as follows:

```
>rgarchfit <- garchFit(xlog~arma(1,1)+garch(1,1), xlog)
>prd_garch <- predict(rgarchfit, 20)
>plot(prd_garch$standardDeviation, ylab="SD Forecast"
>plot(prd_garch$meanerror, ylab="ME Forecast"
>plot(prd_garch$meanforcast,ylab="Mean Forecast",main="Mean Forecast – ARMA + GARCH
```

The predicted values are shown in Table 9. The Mean forecast plot is shown in Figure 15. The standard deviation of forecast errors is shown in Figure 16. The Mean Error Forecast is shown in Figure 17.

Table 9: Predicted values using ARMA+GARCH(1,1)

	meanForecast	meanError	standardDeviation
1	0.0054	0.028	0.0279
2	-0.0022	0.0292	0.0292
3	0.0032	0.0303	0.0301
4	-0.0007	0.0310	0.0308
5	0.0021	0.0315	0.0313
6	0.0001	0.0320	0.0317
7	0.0016	0.0323	0.0321
8	0.0005	0.0326	0.0323
9	0.0013	0.0328	0.0325
10	0.0007	0.0329	0.0326
11	0.0011	0.0330	0.0328
12	0.0008	0.0331	0.0328
13	0.0010	0.0332	0.0329
14	0.0009	0.0333	0.0330
15	0.0010	0.0333	0.0330
16	0.0009	0.0333	0.0331
17	0.0010	0.0334	0.0331
18	0.0009	0.0334	0.0331
19	0.0010	0.0334	0.0331
20	0.0009	0.0334	0.0331

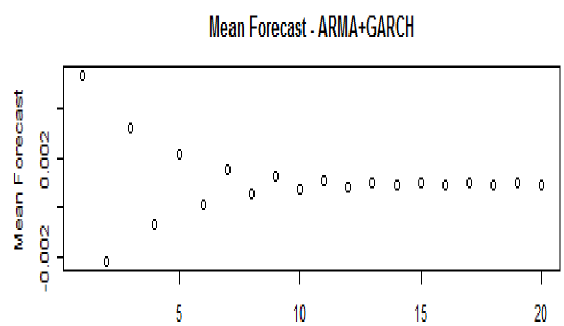

Figure 15: Mean forecast of ARMA+GARCH

The standard deviation and mean error have different values at lower decimal i.e. 4th decimal digit onwards.

The mean forecast values are varied up to lag 10 and became steady after that. In this also, at lower decimals the values are changing steadily. As substantiated in the section 3.2 where the predictions for short term is more accurate than that for long term, the conditional variances are also more accurate in the initial lags.

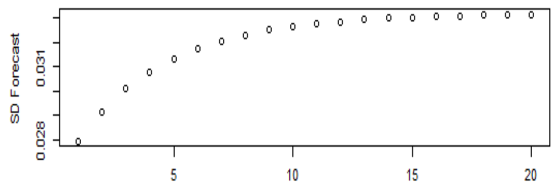

Figure 16: Standard Deviation of Forecast

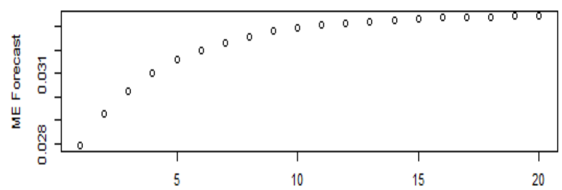

Figure 17: Mean Error Forecast

Thus, the predicted Electricity Demand based on the ARMA(1,1) along with the non-linear behavior of variances improved the accuracy of forecast.

5. Conclusion

In this paper , the ARMA and GARCH models are used to forecast the Day-ahead electricity demand. The ARMA(1,1) model has given predictions of Day-ahead electricity demand with 80-95% confidence bounds for short durations. However the assumption of constant variance of residuals is not true in reality for various reasons as: there are always variances in any annual data it could be daily electricity load data , daily electricity prices, returns and risk profile in stock market etc. The residuals are tested for ARCH effects i.e. conditional heteroskedasticity using McLeod and Ljung-Box tests. The GARCH(1,1) model is identified and estimated the coefficients to fit the model to the residuals and predicted the conditional variances. From the results it is concluded that it is always a good practice to test the volatility of variances or errors and standard deviations after fitting the linear models to improve the accuracy of predictions. The nonlinear issues of variances/errors can be handled appropriately through GARCH model which provide flexibility to coexist with the other models. The combination of ARMA and GARCH models is giving accurate forecasting in high volatility scenarios. R-programming is well suited for modeling and forecasting of electricity demand in this case. There is further scope to extend the model by considering larger sample with 2 to 3 years observations and compare the accuracy of the predicted electricity demand.

6. References

[1] C. Sigauke, D. Chikobvu, *Prediction of daily peak electricity demand in South Africa using volatility forecasting models, Energy Economic, Elsevier, 33(5), September 2011 :882:888*

[2] Ching-Lai Hor, Member, IEEE, Simon J. Watson Member, IEEE and Shanti Majithia , *Daily Load Forecasting and Maximum Demand Estimation using ARIMA and GARCH ; 9th International Conference on Probabilistic Methods Applied to Power Systems KTH, Stockholm, Sweden – June 11-15, 2006;pp-1-6*

[3] J.Nowicka-Zagrajek, R Weron; *Modeling electricity loads in California: ARMA models with hyperbolic noise, Elsevier , Singal Processing 82 (2002), pp-1903-1915*

[4] Reinaldo C. Garcia, Javier Contreras, Marco van Akkeren, and Joao Batista C. Garcia; *A GARCh Forecasting Model to Predict Day-ahead Electricity Prices, IEEE Transactions on Power Systems, Vol. 20, No. 2, May 2005, pp: 867-874*

[5] Chengjun Li, Ming Zhang ; *Application of GARCH Model in the Forecasting of Day-ahead Electricity Prices; IEEE Computer Society ,Third International Conference on Natural Computation (ICNC 2007)*

[6] R Ramakrishna, Naveen Kumar Boiroju and M Krishna Reddy, *Neural Networks Forecasting Model for Monthly Electricity Load in Andhra Pradesh, International Journal of Engineering Research and Applications , Vol. 2, Issue 1, Jan-Feb 2012, pp.1108-1115*

[7] S. Makridakis, S.C. Wheelwright, R.J. Hyndman, *Forecasting—Methods and Applications, 3rd Edition, Wiley, New York, 1998.*

[8] Curia, C. Chatfield , *Forecasting in the 1990s, in: Proceedings of the V Annual Conference of Portuguese Society of Statistics, ,1997 pp. 57–63*

[9] H.S. Hippert, C.E. Pedreira, R.C. Souza, *Neural networks for short-term load forecasting: a review and evaluation IEEE Trans. Power Systems, 16 (1) (2001) 44–55*

[10] W.N. Venables, D. M. Smith, and the R Core Team, *Introduction to R Notes on R: A Programming Environment for Data Analysis and Graphics Version 2.1 (2012-06-22)*

[11]Robert Engle , *GARCH101: The use of ARCH/GARCH Models in Applied Econometrics,2001*

K Chandrabhushan received his M. Tech degree from Osmania University and B.Tech from Indian Institute of Science. He has over 2 decades of experience in IT Industry in the areas - IT Consultancy, Business Analytics, Enterprise Application Packages Implementation and Custom Development. He is doing his Ph. D from JNTU, Hyderabad, India, in Computer Science in **Analytics** area.

Dr. V V HaraGopal received his Ph.D in Statistics from Osmania University. He has been a full professor in the Department of Statistics and Head. He has published over 40 papers to his credit so far in the areas- Multivariate Data Analysis, Operations Research, Pattern Recognition, Image Processing and Group testing. He has conducted national and international level conferences.

Dr. A. Vinay Babu, Professor in Computer Science, Principal, JNTU, Hyderabad, India, received his Ph.D in Computer Science. He guided 9 PhDs and many more PhD Scholars are working under his Guidance. He has published more than 82 Research publications to his credit so far in his specialization such as algorithms, data mining, parallel algorithms, networks and related areas. He received various awards for his contribution such as "Jewel of India 08", "Best Computer Science Teacher Award", "Distinguished Academician", and others