Short-term Electricity Load Forecasting with Time Series Analysis

Hung Nguyen
Department of Mathematics
Eastern Washington University
nguyenluonghung@eagles.ewu.edu

Christian K. Hansen
Professor and Associate Dean
College of Science, Technology, Engineering and Mathematics
Eastern Washington University
chansen@ewu.edu

***Abstract*—Load forecasting plays a fundamental role throughout all segments of system health management for utility companies, including, but not limited to, financial planning, rate design, power system operation, and electrical grid maintenance. Recently, due to the deployment of Smart Grid technologies, utility companies' ability to create accurate forecasts is of even greater importance, especially in consideration of demand response programs, charging of plug-in electric vehicles, and use of distributed energy resources. In this paper, several time series Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) models will be introduced for the purpose of generating forecasts of short-term load demand, at an hourly interval, based on data made available by the Electric Reliability Council of Texas (ERCOT). The case study which expands on the short-term data analyzed in [1] includes over 100,000 data points representing electricity load in Texas recorded over the past 14 years.**

***Keywords*—Time Series Analysis, ARIMA, SARIMA, Prognostics, System Health Management**

I. Introduction

Statistical forecasting is commonly used in virtually any industry and the energy/utility industry is not an exception. In fact, it has become more crucial than ever for utility companies to provide improved load forecasts due to increased deregulation of the energy market. Beginning in the late 1980s, many utilities became conservative about infrastructure upgrades, which pushed the load/capacity ratio higher than usual in order to accommodate the increase in electricity demand [2]. As result of electrical grids operating near and over their limit capabilities, utility companies are increasingly required to develop formal load forecasting models to support their decisions about operation, planning, and maintenance.

Unlike many other products and commodities, electricity has the very unique characteristic that in principle, it must be consumed at the instant it is produced. In any case, storage of electrical energy is either costly, inefficient, or impractical. Therefore, the demand and the supply of electricity must be balanced at all time. Meanwhile, electricity price is directly influenced by factors that affect the demand and supply equilibrium. On the demand side, factors affecting the price include economic activities, meteorological conditions, and general efficiency of consumption whereas on the supply side, resource availability is the main driver of the price. Ideally, to avoid extra costs and losses from storing and transmitting electricity, or from purchasing additional capacity, it is important that a utility company ensure their power systems generate enough capacity to meet demands at all time, especially demand during peak hours, often referred to as peak load. Failing to provide enough energy to meet peak load would result in financial losses in the market and/or even worse, system-wide blackouts. Notably, a number of blackouts resulted from supply shortage and equipment failure was due to the deregulation of the energy market, inadequate transmission investment in particular [3], and possibly their failures to predict peak loads. For this reason, more and more utility companies have directed their attentions toward the quality of their peak load forecasting models.

A. Types of Load Forecasting:

A single load forecasting model generally cannot satisfy all the business needs and planning purposes. Consequently, load forecasting problems are divided into the following categories:

- Very short-term load forecast: ranging from a few minutes to a few hours ahead.
- Short-term load forecast: ranging from one day to two weeks ahead.
- Medium-term load forecast: ranging from two weeks to three years ahead.
- Long-term load forecast: ranging from three to fifty years ahead.

Often, load forecasting problems are mainly categorized as short-term or long-term, using "two weeks" as the cut-off point

978-1-5090-0382-2/16/$31.00 ©2017 IEEE

978-1-5090-5710-8/17/$31.00 ©2017 IEEE

due to the limitation of weather forecasts beyond a two-week period [2]. Different models are employed to generate these different types of forecasts, and the models also serve different purposes. It is critical to understand the limitations of each of these different models. For example, it might be possible to predict the next day load with the mean absolute percentage error (MAPE) approximately 5% or lower; however, to predict the peak load for next year with the same MAPE level is nearly impossible due to the low accuracy in longer term of the ARIMA models [2].

B. The Smart Grid Era:

In recent years, the energy industry has gone through dramatic structural changes, from analog to digital, due to the deployment of the Smart Grid. Meanwhile, Smart Grid technologies have enabled other useful applications, such as the demand response program and distributed energy resources. Together with the great potential for greener and more efficient energy consumption, the load forecasting problems become even more challenging due to the influence of "negative load" resulting from alternative energy resources being connected to the grid.

A variety of mathematical techniques has been employed for load forecasting, such as Artificial Neural Network (ANN), Multiple Linear Regression (MLR) and Time Series Analysis with pros and cons provided in [2]. However, the broad class of time series Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models (discussed below) have proven to outperform other techniques in modeling the short-term load demand [2].

II. Time Series Modeling

The statistical literature, e.g. [4] and [5], provide different methods that can be used for model fitting, i.e. estimating the unknown parameters, and constructing forecasts for finite future values Y_t. In this study, several approaches will be used to eliminate non-stationarity of a time series and several stochastic models are used to produce load forecasts using different lengths of data. The forecast results are then compared directly with the actual hourly load data provided by ERCOT in January 2016.

A. Stationary ARMA Time Series:

Let Y_t denote the hourly load (in MW) at discrete time t. Consider the classical additive discrete time series model:

$$Y_t = m_t + s_t + c_t + X_t \tag{1}$$

where $m_t$is a trend component modeled by a polynomial of order k, s_t is a seasonal component with known period P_1, c_t is a cyclical component (deterministic factor) with a known period P_2 and X_t is an ARMA(p,q) time series. Both the seasonal and cyclic components are represented by periodic functions, but are separated in order to distinguish short-term (hourly) from long-term (daily) variation patterns. Both components may be decomposed using Fourier series as described in [1]. The individual components in (1) are thus parametrized as:

$$m_t = \alpha_0 + \alpha_1 t + \cdots + \alpha_k t^k \tag{2}$$

$$s_t = \sum_{n=0}^{\infty} a_n cos\left(\frac{2\pi nt}{P_1}\right) + \sum_{n=1}^{\infty} b_n sin\left(\frac{2\pi n}{P_1}\right) \tag{3}$$

$$c_t = \sum_{n=0}^{\infty} a_n cos\left(\frac{2\pi nt}{P_2}\right) + \sum_{n=1}^{\infty} b_n sin\left(\frac{2\pi nt}{P_2}\right) \tag{4}$$

$$X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \sum_{j=0}^{q} \theta_j Z_{t-j} \tag{5}$$

In this ARMA model, $\alpha_0, \alpha_1, \ldots \alpha_k, \phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q$ are unknown parameters, P_1 is the seasonal period, P_2 is the cyclical period and Z_t is assumed to be white noise, $Z_t \sim WN(0, \sigma^2)$.

The equation (5) can be rewritten in the more compact form:

$$\phi(B)X_t = \theta(B)Z_t \tag{6}$$

where ϕ and θ are the p^{th} and the q^{th} degree polynomials

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p \tag{7}$$

and

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q \tag{8}$$

and B is the backward shift operator

$$B^j X_t = X_{t-j} \tag{9}$$

B. ARIMA Models for Non-Stationary Time Series:

Consider the ARIMA(p,d,q) process which is expressed as follows:

$$\phi(B)(1-B)^d X_t = \theta(B)Z_t \tag{10}$$

where ϕ and θ are the p^{th} and the q^{th} degree polynomials; d^{th} is a non-negative differencing operation.

It is often the case that while the stochastic processes may not have a constant level, they inhere certain homogeneous behaviors over time. Accordingly, the ARIMA processes, which is a variation of the ARMA models, can be used to model a wide range of non-stationary time series, see [4]. In particular, if d is a non-negative integer, X_t is said to be an ARIMA(p,d,q) process if $(1-B)^d X_t$ is an ARMA(q,p) process.

C. SARIMA Models for Non-stationary Time Series:

In order to estimate the parameters describing the (random) time series component X_t, the deterministic components, i.e. the trend component, the seasonal component and the cyclical component must be identified and removed. In practice, however, it may not be reasonable to assume that the seasonal and cyclical components repeat themselves precisely period after period. Therefore, a SARIMA model may provide a better explanation and better fit a time series since it allows randomness in the seasonal pattern [4].

A SARIMA(p,d,q) × (P,D,Q) can be expressed as follows:

$$\phi(B)\Phi(B^s)(1-\mathrm{B})^{\mathrm{d}}(1-\mathrm{B}^{\mathrm{s}})^D Y_t = \theta(B)\Theta(B^s)Z_t \tag{11}$$

where $Z_t \sim WN(0, \sigma^2)$, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, $\Phi(B) = 1 - \Phi_1 B - \cdots - \Phi_P B^P$, $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$, and $\Theta(B) = 1 + \Theta_1 B + \cdots + \Theta_Q B^Q$.
In this SARIMA model, $\phi_1, \phi_2, \ldots, \phi_p$, $\Phi_1, \Phi_2, \ldots, \Phi_P$, $\theta_1, \theta_2, \ldots, \theta_q$, and $\Theta_1, \Theta_2, \ldots, \Theta_Q$ are unknown parameters, d and D are finite non-seasonal and seasonal differencing, respectively.

D. Accuracy Measurements and Errors Calculation:

The Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), and Mean Squared Deviation (MSD) statistics are used as the main criteria, as in [6], to assess performances and to make comparisons between similar models.

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|}{n} * 100 \quad (12)$$

$$MAD = \frac{\sum_{t=1}^{n} |Y_t - \hat{Y}_t|}{n} \quad (13)$$

$$MSD = \frac{\sum_{t=1}^{n} |Y_t - \hat{Y}_t|^2}{n} \quad (14)$$

where Y_t are the actual values, $\hat{Y}_t$ are forecast values, and n is the number of observations.

III. A Case Study

In this case study, time series forecasts are constructed using hourly load data for various regions in Texas made available by the Electric Reliability Council of Texas (ERCOT) [7]. Historical data is available from 1995 to 2015, however, due to the unavailability of data in 2001, only observations between 2002 and 2015 are used to produce 6 days ahead load forecasts for January 2016.

In addition, forecasts produced from different lengths of data ranging from 1 week to 6 years are evaluated in order to study the effect that varying the length of data has on the accuracy of forecasting models.

To effectively handle missing and extra observations at the beginning and the end of daylight savings from each year, additional hour, i.e. 2am, is created by averaging the 2-nearby observations. For instance, at the beginning of daylight savings, the 2am data is created by averaging the 1am and 3am observations. Similarly, at the end of daylight savings, the extra 2am data can simply be replaced by the average of the 2am observations. For modelling convenience, the additional days in February from leap years 2004, 2008, 2012, and 2016 shall be removed.

A. Model Fitting and Parameter Estimation:

First, a time series plot is constructed to determine whether the time series exhibits characteristics of trend, seasonality and/or cyclicality. Figure 1 shows a plot of the time series Y_t observed for 5110 days (n=122640 observations).

A close examination of the graph in Figure 2 shows that all three of the aforementioned characteristics are likely features of the data and should be included in the final model. In this case, the trend component can be interpreted as the increase of electricity usage over time due to population growth and more electronic devices being deployed. A linear least square regression, see [8], is used to identify the trend component. Now, to remove the trend component, the fitted value in the trend is subtracted from each observation, and refer this as the de-trended data, see Figure 3.

$$Y_t - \hat{m}_t = s_t + c_t + X_t \quad (15)$$

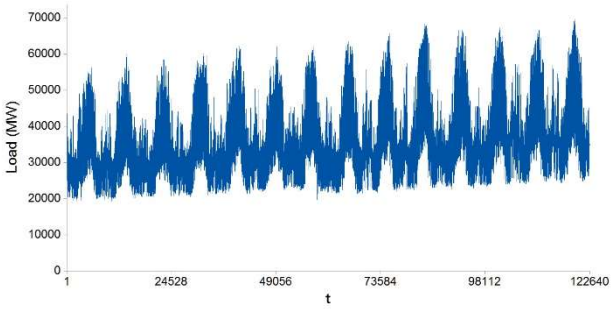

Figure 1: Hourly load on the ERCOT electrical grid from January 1, 2002 to December 31, 2015

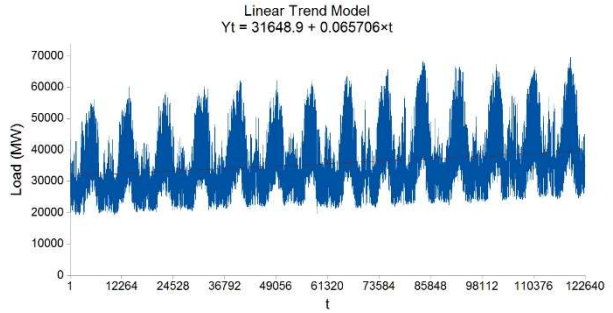

Figure 2: *Trend analysis for hourly load data*

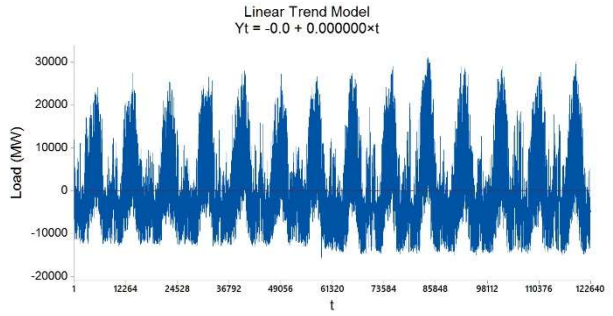

Figure 3: De-trended Data Plot

Next, to identify and remove the cyclical component, which represents the daily variation of electricity usage within a year, the daily average is derived from the de-trended data by averaging 24 observations per day and same day across 14 years, see Figure 4.

$$\overline{c_d} = \frac{1}{14}\left[\sum_{j=0}^{14} \frac{1}{24}\left(\sum_{i=24(d-1)+1+36}^{24d+365j} Y_i - \hat{m}_i\right)\right] \quad (16)$$

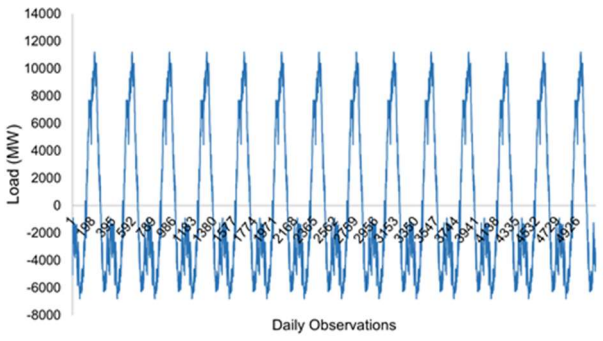

Figure 4: Daily average (Cyclical components)

Now, in order to model the seasonal component, which in this context represent the hourly variation in energy usage throughout the day, the daily average is temporarily removed from the de-trended data, referred to as the detrended and cyclically adjusted data, see Figure 5. To this end, fitting the seasonal component parametrically requires averaging observations recorded at the same hour across the 5110 days.

$$\widetilde{s_t} = Y_t - \widehat{m_t} - \overline{c_d} \tag{17}$$

$$\widehat{s_t} = \frac{1}{N}\sum_{d=1}^{N} \tilde{s}_{24(d-1)+t} \tag{18}$$

where $N = 5110$ and $t = 1, \dots 24$

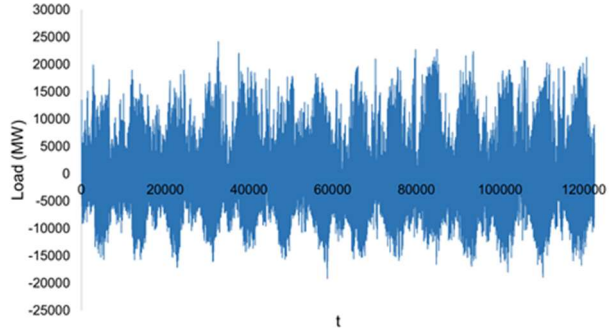

Figure 5: De-trended and cyclical adjusted data

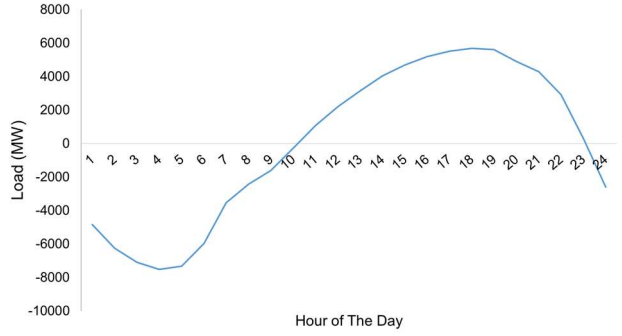

Figure 6: Hourly average of de-trended and cyclical adjusted data

Although fitting a model non-parametrically does not restrict the function of the detrended data, or $Y_t - \widehat{m}_t$, to any specific shape or mathematical form, a parametric model has the advantage of providing a simpler description with fewer characteristic parameters [1].

Recall equation (3), in this case, $f(t) = s_t$, and $P_1 = 24$, $a_0 = 0$. Here, the parametric model is limited to include no more than the first four harmonic components as in [1], and the unknown parameters are estimated using the least squares linear regression method [8] and the Minitab statistical software [9], see Figure 8.

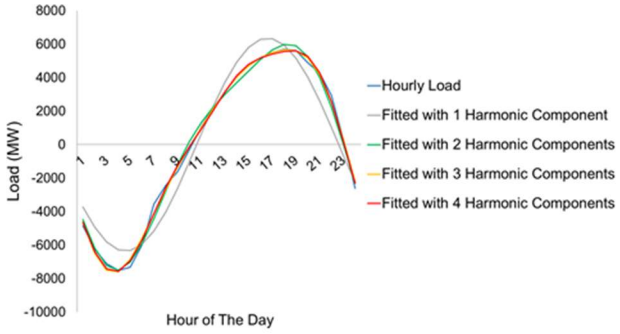

Figure 7: Comparison of seasonal component (nonparametric) and the fitted model (parametric) with the individual harmonic components

Regression Analysis: Hourly Avg versus cos1, cos2, cos3, cos4, sin1, sin2, sin3, sin4

Model Summary

S	R-sq	R-sq(adj)
240.570	99.73%	99.73%

Coefficients:

Term	Coef	SE Coef	T-Value	P-Value
cos1	-2264.51	0.97	-2330.96	0.000
cos2	-38.268	0.971	-39.39	0.000
cos3	96.206	0.971	99.03	0.000
cos4	-106.699	0.971	-109.83	0.000
sin1	-5950.45	0.97	-6125.04	0.000
sin2	-1412.00	0.97	-1453.43	0.000
sin3	-326.847	0.971	-336.44	0.000
sin4	52.361	0.971	53.90	0.000

Regression Equation:

Hourly Avg = -2264.51 cos1 - 38.268 cos2 + 96.206 cos3 - 106.699 cos4 - 5950.45 sin1 - 1412.00 sin2 - 326.847 sin3 + 52.361 sin4

Figure 8: Minitab output for the fitted harmonic model

Figure 7, and 8 show the fitted results and comparison between non-parametric fitted and parametric fitted models. Noticeably, the parametric fitted model provides an excellent fit with $R^2 = 99.73\%$. However, for a parsimonious model, the optimal fitting result shall be traded off at the expense of reducing the degrees of freedom [10,11]. Therefore, a model based on the first two harmonics is used resulting in $R^2 = 99.43\%$.

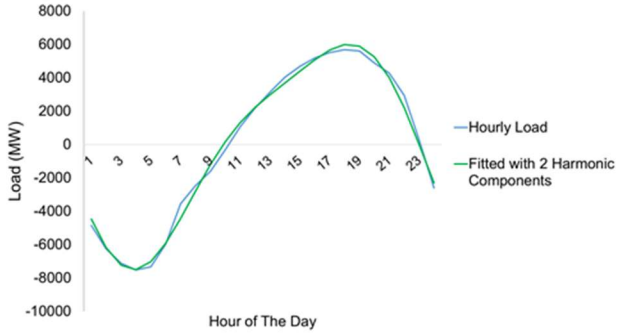

Figure 9: Hourly Variation fitted with 2 harmonic components

Regression Analysis: Hourly Avg versus cos1, cos2, sin1, sin2

Model Summary:

S	R-sq	R-sq(adj)
350.685	99.43%	99.43%

Coefficients:

Term	Coef	SE Coef	T-Value	P-Value
cos1	-2264.51	1.42	-1599.04	0.000
cos2	-38.27	1.42	-27.02	0.000
sin1	-5950.45	1.42	-4201.79	0.000
sin2	-1412.00	1.42	-997.05	0.000

Regression Equation:

Hourly Avg = -2264.51 cos1 - 38.27 cos2 - 5950.45 sin1 - 1412.00 sin2

Figure 10: Minitab output for the fitted model with 2 harmonic components

Once the parametric model for seasonality is established, the seasonal component from the de-trended data can be removed to parametrically estimate the cyclical component, referred to as de-trended and de-seasonalized data.

$$\tilde{c}_t = Y_t - \hat{m}_t - \hat{s}_t \tag{19}$$

$$\bar{\tilde{c}}_d = \frac{1}{14}\left[\sum_{j=0}^{14} \frac{1}{24}\left(\sum_{i=24(d-1)+1+36}^{24d+365j} \tilde{c}_i\right)\right] \tag{20}$$

Again, a similar technique as in equation (16), as shown in (20), is used to calculate the daily average, then equation (4) is applied to estimate the cyclical components parametrically. In this case, $g(t) = c_t$, $P_2 = 24 * 365$ and $a_0 = 0$. Figure 11 provides a visualization of the cyclicality and the comparisons of alternative parametrically fitted models.

To meet the model selection criteria, i.e. parsimony, the parametric model with 3 harmonic components is chosen over the model with 4 harmonic components. Although all 8 estimated parameters (coefficients) are statistically significant, the fitted models with 1 harmonic components and 2 harmonic components does not fully describe the cyclicality of the cyclical curve. Thus, the third order harmonic model is adopted.

Lastly, through similar procedures, both the seasonality and the cyclicality in the raw data are removed to obtain the de-seasonalized and de-cyclicalized data, then the linear least square regression is applied to refine the estimate of the trend component, as shown in Figure 13.

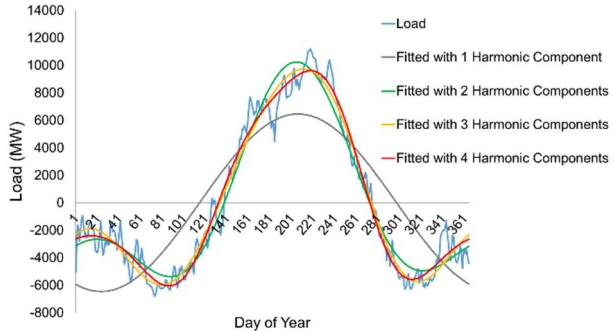

Figure 11: Comparison of cyclical component (non-parametric) and the fitted model (parametric) with the individual harmonic components

Regression Analysis: Daily Avg of versus cos1_1, cos2_1, cos3_1, cos4_1, sin1_1, sin2_1, sin3_1, sin4_1

Model Summary:

S	R-sq	R-sq(adj)
915.870	97.15%	97.15%

Coefficients:

Term	Coef	SE Coef	T Value	P Value
cos1_1	-5923.15	3.70	-1601.47	0.000
cos2_1	2781.34	3.70	752.01	0.000
cos3_1	830.21	3.70	224.47	0.000
cos4_1	-329.01	3.70	-88.96	0.000
sin1_1	-2573.53	3.70	-695.82	0.000
sin2_1	2575.19	3.70	696.27	0.000
sin3_1	303.28	3.70	82.00	0.000
sin4_1	-400.37	3.70	-108.25	0.000

Regression Equation:

Daily Avg of De-Hourly Adjusted = -5923.15 cos1_1 + 2781.34 cos2_1 + 830.21 cos3_1 - 329.01 cos4_1 - 2573.53 sin1_1 + 2575.19 sin2_1 + 303.28 sin3_1 - 400.37 sin4_1

Figure 12: Minitab output for the fitted harmonic model

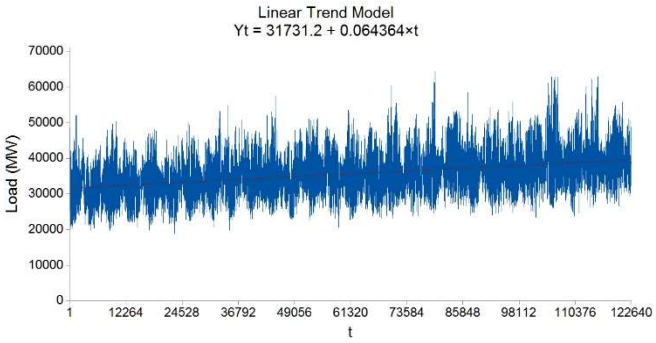

Figure 13: Trend analysis for the de-seasonalized and de-cyclicalized data

The processes of eliminating components by components to preliminarily fit the trend, seasonality and cyclicality may be repeated until all the refined estimated parameters (coefficients) remain unchanged.

$$\hat{x}_t = y_t - \hat{m}_t - \hat{s}_t - \hat{c}_t \tag{21}$$

After having estimated all the trend, seasonal and cyclical components, the residuals $\hat{x}_t$ can be obtained through equation (21). Provided that the deterministic components have been successfully isolated, a time series plot should now exhibit the characteristics of a stationary time series x_t, as shown in Figure 14. It is now appropriate to fit an ARMA(p,q) model (5, 6) or an ARIMA(p,d,q) model (10) to the residuals.

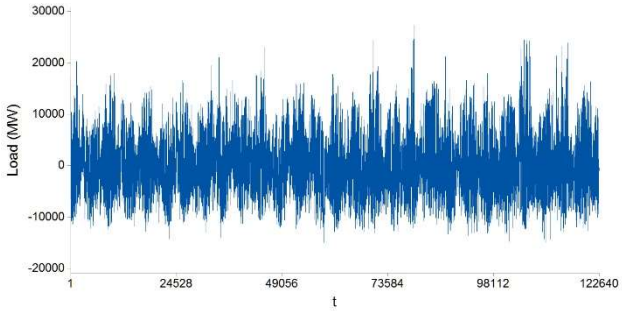

Figure 14: Time series plot x_t

As in [1], an ARMA(2,0) model is attempted to fit the time series and to produce forecast for January 1st-6th, then the

forecast result is compared with the actual hourly load provided by ERCOT. Figure 15 displays the output for the fitted ARMA(2,0) model from Minitab.

ARMA Model: Time Series (Final Iteration)

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1.7068	0.0019	913.19	0.000
AR	2	-0.7560	0.0019	-404.48	0.000

Number of observations: 122640
Residuals: SS = 65909869387 (backforecasts excluded)
MS = 537434 DF = 122638

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	20886.8	124959.7	140601.3	215142.0
DF	10	22	34	46
P-Value	0.000	0.000	0.000	0.000

Figure 15: Minitab output for the fitted ARMA(2,0)

IV. Comparisons of Load Forecasting Results

Without mentioning other useful applications of time series analysis, such as studying past behavior and assessing past performance, the most popular and most important application of time series analysis is to produce forecasts for a variable of interest. As discussed in the previous section, load forecasting is an integral part of the planning purposes in a utility company. The objective of this analysis is to decompose or to separate the time series (random components) from deterministic components, i.e. trend, seasonality, and cyclicality, and apply the ARMA processes to construct a load forecasting model.

A. Load Forecast with ARMA Model:

A forecast using an ARMA (2,0) model is provided and compared against the actual hourly load, as shown in Figure 16.

Although the predicted values for Y_t in Figure 16 did not fully capture the hourly variation of the load demand curve, they closely represent the daily variation (cyclical component). This suggests further investigation in the assumption that hourly variation is similar on daily basis. In other words, energy consumption throughout a day could gradually change over the years as a result deploying more efficient appliances and more smart monitoring devices.

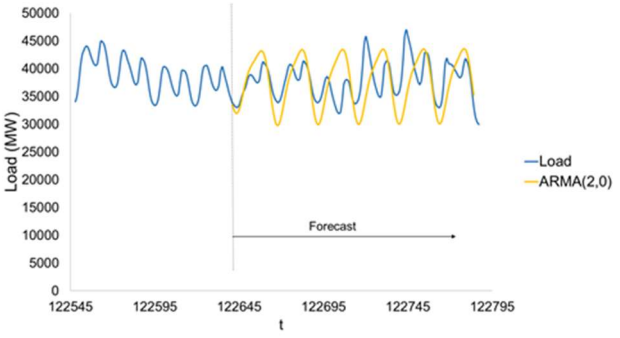

Figure 16: A comparison of the actual and forecasting model of the hourly load in Texas from January 1st to January 6th, 2016

B. Forecast from SARIMA Models and Comparisons:

Additionally, forecasts produced from SARIMA models using different length of data are also included, with accuracy measures provided in Table I and graphs in Appendix A.*

Although with 1 week of data, the SARIMA model poorly capture the weekly variation component, with 1 additional week of observations, it outperforms the ARIMA model produced from 14 years of hourly load observations. SARIMA models continue to improve when adding addition weeks of historical data to produce forecasts; however, somewhere between 3 to 8 weeks length, the MAPE suggests an optimal length of observations which the SARIMA model would best produce forecast with the lowest percentage error.

TABLE I. SARIMA Forecast Comparisons

Date Length	Forecasts			
	Model	***MAPE***	***MAD***	***MSD***
14-years	ARMA(2,0)	9.13%	3451	1.98E7
1 week	SARIMA(1,1,1) x (0,1,1)	12.41%	4605	2.81E7
2 weeks	SARIMA(1,1,1) x (0,1,2)	7.90%	2922	6.63E6
3 weeks	SARIMA(1,1,1) x (0,1,2)	4.40%	1658	5.94E6
4 weeks	SARIMA(1,1,1) x (0,1,2)	4.36%	1638	5.70E6
8 weeks	SARIMA(1,1,1) x (0,1,2)	4.45%	1669	5.61E6
6 months	SARIMA(1,1,1) x (2,1,2)	9.15%	33922	1.69E6
1 year	SARIMA(1,1,1) x (2,1,2)	10.49%	3892	2.20E7

V. Discussion

Several ARIMA models have been used to model the load demand curve after the trend, the seasonality and the cyclicality were isolated; however, considering the autocorrelation function (ACF) and the partial autocorrelation function (PACF), see [11], suggest that a correlation between observations recorded at the same hour that is still not accounted for†. These are indications that models such as ARIMA and SARIMA have limitations when it comes to producing accurate forecasts. Other issues must also be addressed in future study are whether a 14-year data set would be more desirable and/or how much data an utility actually need to construct better forecasts due to the rapid change in technology, more efficient energy consumptions, but at the same time higher energy demand?

References

[1] Hansen, Christian K. "A prognostic model for managing consumer electricity demand and smart grid reliability." Prognostics and Health Management (PHM), 2012 IEEE Conference on. IEEE, 2012.

[2] Hong, T., and M. Shahidehpour. "Load forecasting case study." EISPC, US Department of Energy (2015).

[3] Hines, Paul, Jay Apt, and Sarosh Talukdar. "Trends in the history of large blackouts in the United States." *Power and Energy*

* The coefficients, produced by Minitab 17, from the SARIMA models in table I are all statistically significant

† ACF and PACF are diagnostic tools to evaluate the stationarity of a time series. They both indicate a strong correlation at lag 24 and non-stationarity in the time series.

Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE. IEEE, 2008.

[4] Brockwell, Peter J., and Richard A. Davis. Time Series: Theory and Methods. New York: Springer-Verlag, 1991. Print.

[5] Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time series analysis: forecasting and control*. Hoboken, NJ: John Wiley & Sons, Inc., 2016. Print.

[6] Hyndman, Rob J., and George Athanasopoulos. *Forecasting: Principles and Practice*. Heathmont, Vic.: OTexts, 2014.

[7] http://www.ercot.com

[8] Wackerly, Dennis D., William Mendenhall, and Richard L. Scheaffer. *Mathematical Statistics with Applications*. Belmont, CA: Thomson Brooks/Cole, 2008. Print.

[9] http://www.minitab.com

[10] Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci. Introduction to Time Series Analysis and Forecasting. Hoboken, NJ: Wiley-Interscience, 2008. Print.

[11] Enders, Walter. *Applied Econometric Time Series*. Hoboken, NJ: Wiley, 2015. Print.

APPENDIX A:

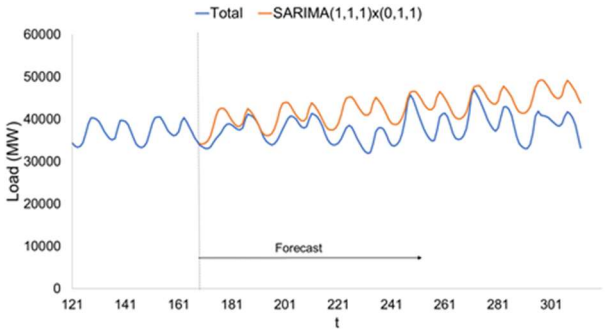

Figure 17: SARIMA(1,1,1) x (0,1,1) using 1-week length data

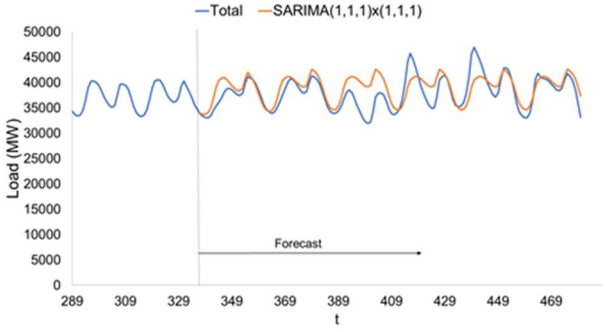

Figure 18: SARIMA(1,1,1) x (0,1,2) using 2-weeks length data

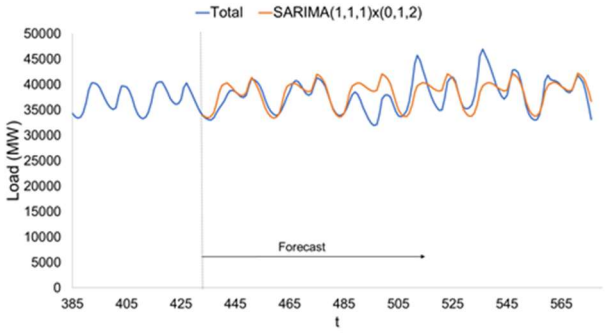

Figure 19: SARIMA(1,1,1) x (0,1,2) using 3-weeks length data

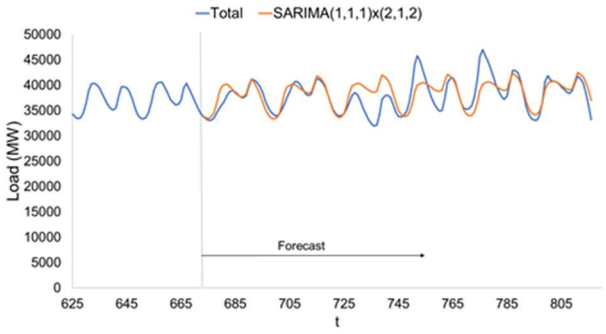

Figure 20: SARIMA(1,1,1) x (0,1,2) using 4-week data

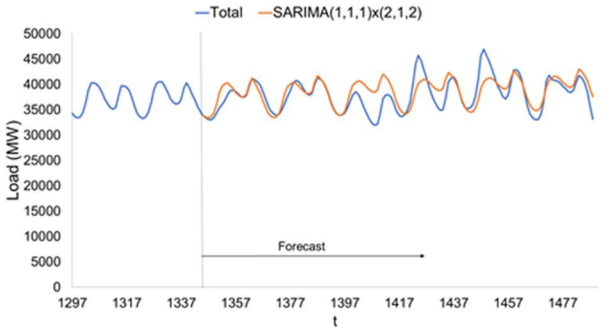

Figure 21: SARIMA(1,1,1) x (0,1,2) using 8-weeks length data

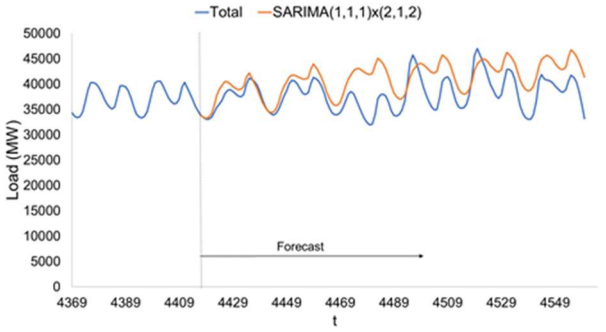

Figure 22: SARIMA(1,1,1) x (2,1,2) using 6-months length data

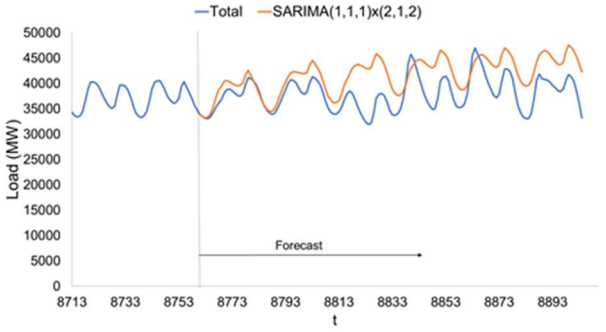

Figure 23: SARIMA(1,1,1) x (2,1,2) using 1-year length data