ERCOT South Central Weather Zone Energy Load Forecasting & Analysis

By: Ziam Ghaznavi

Time Series Analysis, Modeling and Control:
Spring 2020 Course Project - Final Report

May 19th 2020

Abstract

Accurate prediction of electrical load requirement is crucial to maintaining integrity of any electrical grid. Reliable energy prediction is especially important in the case of ERCOT or the Electric Reliability Council of Texas since it operates on a competitive wholesale bulk-power market or energy-only business model which requires frequent power dispatch scheduling based on short term 1 hour, daily and mid-term weekly forecasts. To ensure proper operation, ERCOT currently aims for a 12.5% power supply margin over maxim load utilizing complex hybrid models to predict energy load which rely on expensive proprietary data services to achieve necessary levels of accuracy. This work explores implementation of statistics-based time series modelling and analysis techniques for predicting energy load for ERCOT's South-Central weather zone. 43,824 load and local air temperature measurements each starting on Jan. 1st, 2015 01:00 through Dec. 31st, 2019 24:00 serve as the training set for model fitting purposes and model testing spans the first week or 168 data points of 2020. Time series models for load were trained on linear residuals of standardized data in order to address heteroscedasticity and remove non-stationarity whereas the temperature data model was fitted directly to raw values. Model selection consisted of F-test criterion and minimum AIC for scalar and vectoral ARMA models, respectively. Prediction error is defined in terms of Green's Function coefficients and variance of residuals by orthogonal Wold's decomposition and application of the conditional expectation of squared error. A scalar ARMA(36,35) model is found to be adequate for linear residuals of standardized load data and had a RSS of 262.34. Similarly, ARMAV(35,35,34) was found to be the most appropriate model for linear residuals of load data with air temperate as input and zero lag and had an RSS of 263.28. In both cases, actual load values were within 6% of 1-hour ahead load predictions meeting ERCOT's 12.5% margin requirement while actual load did exceed this margin for 24-hour ahead predictions. Likewise, actual load exceeded 12.5% of ARMA generated 1-week ahead forecasts over the testing data time span. However, ARMAV generated 1-week ahead forecasts were found to be acceptable for capacity planning requirements. While statistically inconclusive, this result does suggest ARMAV is better suited for predicting actual load and more appropriate for use in real-world power dispatch scheduling. Overall, time series modeling alone is able to accurately and reliably predict load values 1-hour ahead and even 1-week ahead utilitzing vectoral ARMA modelling with local air temperature input.

1 Introduction

Accurate prediction of electrical load requirement is crucial to maintaining grid integrity since energy storage technologies have not reached levels of scalability required for economic implementation in grid infrastructure to a significant measure. As a result, excessive power generation transmitted to the grid will impose undue thermal stress which can lead to subsystem damage if left unchecked. Conversely, proper energy capacity planning is necessary to ensure supply meets demand to prevent occurrence of black outs. Further complicating the problem, different generation sources possess inherit energy transmission limitations and require varying ramp up times to reach system specific acceptable efficiency ranges. For example, natural gas turbines may have a different transient power delivery profile verses a solar energy plant.

The Western, Eastern and ERCOT interconnections constitute the three main energy grid networks in the continental United States. Excluding the city of El Paso and some minor peripheral geographical regions of Texas, ERCOT or the Electric Reliability Council of Texas manages roughly 90% of the state's energy load serving more than 26 million customers [1]. ERCOT's primary function is scheduling power onto the grid and facilitating exchange between energy generators and distributors ensuring consistent power availability to end customers. The South-Central Weather Zone of ERCOT is particularly interesting as it includes the cities of Austin and San Antonio which are two of the nation's fastest growing in terms of population according to recent estimates by the U.S Census Bureau [2]. As a result, the present study is solely focused to the South-Central region of Texas where energy demand is projected to steadily increase at a rate 2% annually according to Warren Lasher, Senior Director of System Planning at ERCOT.

Reliable energy prediction is especially important in the case of ERCOT since it operates on a competitive wholesale bulk-power market or energy-only business model. In contrast to the other major interconnections, producers in ERCOT only receive compensation for energy delivered onto the grid whereas the capacity component factored into pricing is minor. ERCOT's pricing scheme promotes innovation in the energy sector and allows for competitive energy pricing lowering costs for end users, but the tradeoff is the need for frequent power dispatch scheduling based on short term 1 hour, daily and mid-term weekly forecasts. To ensure proper operation, ERCOT currently aims for a 12.5% power supply margin over maxim load.

Understandably, energy load prediction is a notoriously challenging problem. Researchers have undertaken numerous modelling approaches to address this need ranging from simple least squares regression fitting to more sophisticated methodologies including state vector machines and neural networks [3], [4]. Based on conversations with employees, ERCOT itself utilizes complex hybrid models to predict energy load which rely on expensive proprietary data services to achieve necessary levels of accuracy. The objective of this work is to implement statistics-based time series modelling and analysis techniques originally formalized by Box and Jenkins et al. to energy load predicting, and assess if this approach alone is acceptable per ERCOT-defined required power load margins [5].

2 Methods

2.1 Data Sets

Figure 1 shows hourly sampled adjusted metered load data for the ERCOT South-Central weather zone in units of megawatt hours (MWh). The term adjusted here simply means transmission losses are accounted for and data is representative of power delivered to customers.

43,824 load measurements starting on Jan. 1st, 2015 01:00 through Dec. 31st, 2019 24:00 serve as the training set for model fitting purposes. Model predictions span over the first week (168 data points) of 2020 i.e. testing data ranges from Jan. 1st, 2020 01:00 to Jan. 7th, 2020 24:00

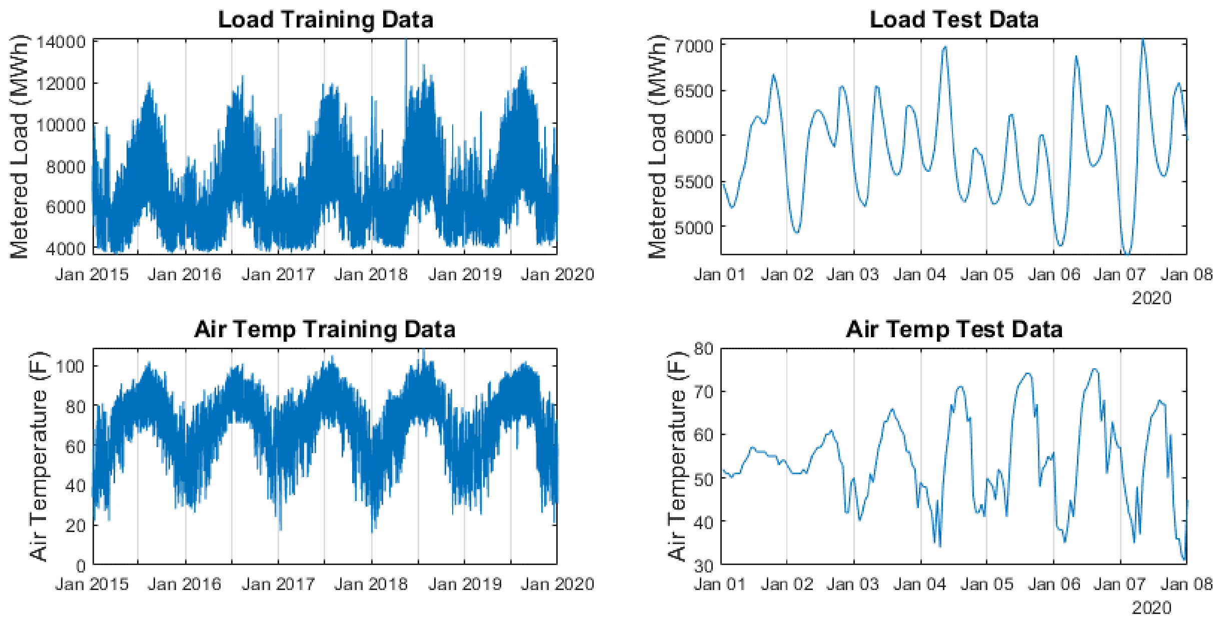


Figure 1: Training and testing data sets for both hourly energy load and air temperature

Similarly, Figure 1 also contains hourly sampled air temperature data collected locally at Austin Bergstrom Airport weather station. Visual comparison of temperature and metered load suggests the two time series share similar underlying dynamics as minimum and maximum temperatures match peaks in energy usage. Therefore, the vectoral model discussed in Section 3.2 utilizes air temperature (T_t) as an input to improve prediction accuracy of load data. Air temperature data training and testing sets span identical timeframes as corresponding load data sets.

2.2 Data Processing

Observation of Figure 1 reveals metered load date exhibits varying mean and variance with time. For instance, variance of load is greater in Summer months compared to variance in Spring and Autumn. In addition to heteroscedasticity, a positive linear deterministic trend exists in metered load data which is expected due to the rapid population growth in the region as mentioned earlier. As a result, preprocessing is required to transform load data into a wide-stance stationary time series appropriate for auto-regressive moving average (ARMA) modeling.

First, in order to address heteroscedasticity, metered load underwent standardization on an hourly basis according to the following procedure.

$$Z_t = Z_{h+24k} = \frac{L_{h+24k} - \mu_h}{\sigma_h}$$
$$\sigma_h = \sqrt{\sum_{k=0}^{d} \frac{(L_{h+24k} - \mu_h)^2}{d-1}} \tag{1}$$
$$\mu_h = \frac{1}{d} \sum_{k=0}^{d} L_{h+24k}$$

where,

$$h = 1,2,\ldots,24$$
$$k = 0,1,2,\ldots,d$$

Time series L_t and Z_t here are the original and corresponding standardized load data, respectively. Indices h and d represent hours in a day and total number of days in the entire data set, respectively.

Next, a linear trend was fitted to standardized load data via least squares regression following the form below.

$$Z_t = \beta_0 + \beta_1 t + Z_t^* \tag{2}$$

All subsequent models for load were trained on these linear residuals of standardized data (Z_t^*). Thus, model predictions are originally in terms of residuals of normalized load values, and conversion to natural units of MWh is achieved by inverting operations of equations (1) and (2) in reverse order. The air temperature model discussed in Section 3.2 is trained directly on raw temperature data, and predictions were conveniently in natural units.

2.3 Modelling Procedure

The approach for finding an adequate scalar ARMA models consisted of fitting increasingly higher order ARMA(2n,2n-1) models to data and assessing the significance of the reduction in residual sum of squares of errors (RSS) between the unrestricted and restricted models using F-test criterion at each iteration. This procedure is summarized below.

$$F = \frac{(A_1 - A_0)/s}{A_0/(N-r)} < F_{s,N-r}^{0.95} \tag{3}$$

where F is the F score of the two embedded models under consideration, A_1 is the minimized RSS of the restricted model, A_0 is the minimized RSS of the unrestricted model, s is the number of restricted parameters between the two models, N is the number of data points and r is the number of estimated parameters in the unrestricted model. The restricted ARMA(2n,2n-1) model is assumed adequate when equation (3) yields true. Additionally, auto-correlations of residuals of the adequate model were checked for 100 lags and considered acceptable the majority of values remained within the 95% confidence interval around mean zero and none were greater in magnitude than 0.1.

The vectoral or ARMAV model of linear residuals of standardized load data with air temperate as input was of the following structure.

$$\begin{bmatrix} Z_t^* \\ T_t \end{bmatrix} = \begin{bmatrix} AR_{11}(B) & AR_{12}(B) \\ 0 & AR_{22}(B) \end{bmatrix} \begin{bmatrix} Z_t^* \\ T_t \end{bmatrix} + \begin{bmatrix} MR_{11}(B) & 0 \\ 0 & MR_{22}(B) \end{bmatrix} \begin{bmatrix} a_{Z_t^*} \\ a_{T_t} \end{bmatrix} \tag{4}$$

where a_{i_t} are past errors, $AR_{ij}(B)$ are linear functions of the backshift operator with estimated AR coefficients and $MR_{ij}(B)$ are linear functions of the backshift operator with estimated MR coefficients. Per equation (4), load and temperature models were fitted independent of each other. The ARMA model for temperature data or the second row of equation (4) was fitted using the F-test criterion described.

Selection of the appropriate ARMAV model for load data employed Akaike Information Criteria (AIC). ARMAV(n,n,n-1) models of order n from 1 up to AR order of the corresponding adequate scaler ARMA model for load were fitted to the data. Limiting the max ARMAV model order in this manner provided for direct comparison between the univariant and multivariant approaches. Next, AIC values assuming Gaussian residuals were calculated for each model according to the equation below.

$$AIC = N \ln \frac{RSS}{N} + 2 \tag{5}$$

where k is the number of estimated model parameters. The ARMAV model with the minimum AIC value was selected as the most adequate model for analysis and prediction. Similar to the scalar approach, auto-correlations of residuals of the selected ARMAV load model were checked for 100 lags as well as cross-correlations of residuals between load and temperature models.

2.4 Characterization of Prediction Uncertainty

Assessment of model capabilities entailed the comparison of 1-hour and 24-hours ahead predictions as well as 1-week ahead forecasts spanning the testing set time span against recorded data per ERCOT requirements. Characterization of prediction error in terms of Green's Function coefficients and variance of residuals involved orthogonal Wold's decomposition of each model and application of the conditional expectation of squared error [6].This procedure is straightforward in the scaler ARMA model case, however requires partial fraction expansion when dealing with vectoral models. Based on equation (4), prediction error of the ARMAV load model in this study is explicitly expressed below.

$$\begin{gathered} Z_L^* = \frac{MA_{11}(B)}{1 - AR_{11}(B)} a_{Z_L^*} + \frac{AR_{12}(B) MA_{22}(B)}{\big(1 - AR_{11}(B)\big)\big(1 - AR_{22}(B)\big)} \, a_{T_t} \\ Var\big[\widehat{e_{Z_L^*}}(l)\big] = (\sum_{i=0}^{l-1} G_{1_i}^2) \sigma_{Z_L^*}^2 + (\sum_{i=0}^{l-1} G_{2_i}^2) \sigma_{T_t}^2 \end{gathered} \tag{6}$$

where $\widehat{e_{Z_L^*}}(l)$ is estimated prediction error, G_{X_i} are Green's Function coefficients for lag i and $\sigma_{\mathrm{X_t}}^2$ are variance of model residuals.

2.5 Checks for Stochastic Seasonality

Stochastic seasonality of periodicity P is suspected to exist if an adequate ARMA(n,n-1) model contains pairs of complex conjugate AR roots on or close to the unit circle i.e. $\lambda_{1,2} \cong e^{\pm\frac{2\pi}{P}j}$. To determine if there is statistical evidence for seasonality with period P to exist, the original time series(X_t) is subjected to the below transformation resulting in a parsimonious time series (Y_t) where the AR roots under investigation are deliberately set equal to one in magnitude.

$$Y_t = X_t \prod_{i=0}^{c} (1 - \lambda_i B) \tag{7}$$

where c is the number of complex roots with magnitude within 98% of unity. Next, an embedded parsimonious ARMA(n-c,n-1) is fitted to Y_t and the minimized RSS is compared to the RSS of the adequate ARMA(n,n-1) utilizing the F-test criterion similar to equation (3) [6].

3 Results & Discussion

3.1 Scaler ARMA Model

Iterative F-testing using equation (3) found ARMA(36,35) to be an adequate scaler model for linear residuals of standardized load data and had a RSS of 262.34. Figure 2 below shows the AR roots and autocorrelations of residuals for this model. Autocorrelations appear to be uncorrelated confirming F-testing results. However, close examination of Figure 2 reveals minor spikes at every 24-hour interval suggesting the existence of possible daily stochastic seasonality. Additionally, the system exhibits marginal stability, and 12 pairs of complex conjugate AR roots corresponding to periodicity of 24 hours and its harmonics lay within 98% of the unit circle.

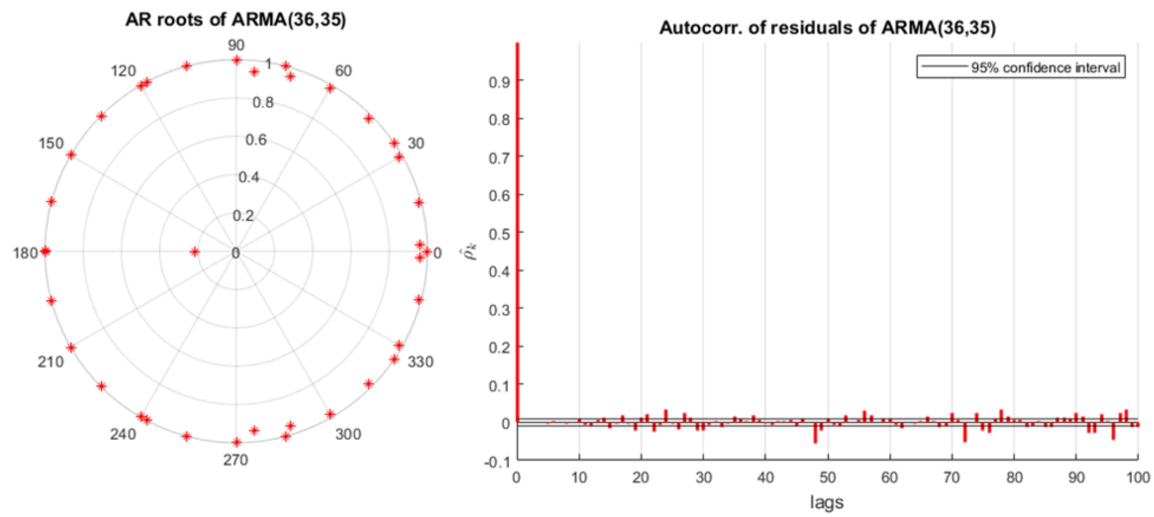


Figure 2: (left) Auto regressive roots of ARMA(36,35) model fitted to linear residuals of standardized load data, and (right) autocorrelations of the model out to lag 100 with 95% confidence bounds.

These results justified testing for existence of stochastic seasonality using the parsimonious modeling approached described in Section 2.5. Testing AR roots corresponding to periodicity of 24 hours and all of its harmonics simultaneously failed to provide statistical evidence for their existence. Similarly, testing the pair of roots corresponding to period of 24 hours alone also resulted in an inadequate parsimonious model. While not statistically detected, deterministic daily seasonality in energy load usage does exist based on prior knowledge, and the plot of Green's function coefficients of the ARMA(36,35) in Appendix A is indicative of this behavior.

Figure 3 below shows 1-hour and 24-hour ahead predictions of load data using the ARMA(36,35) model along with 95% confidence intervals over the first week of Jan. 2020. In both cases, nearly all metered load measurements are within the 95% confidence interval of predicted values with the exception of the 24-hour period of Jan. 1st, 2020. The discrepancy between actual and predicted load values is likely explained by the anomaly of New Year's Eve. In terms of ERCOT's required 12.5% margin over maximum energy load, actual load does not exceed 6% of 1-hour ahead predictions, and therefore, ARMA(36,35) 1-hour ahead predictions are valid for real world operational purposes. However, actual energy usage does exceed 12.5% of 24-hour ahead predictions which would result in insufficient energy capacity planning if used for dispatch scheduling. Appendix A provides plots for percent difference between predicted and actual load values for both cases.

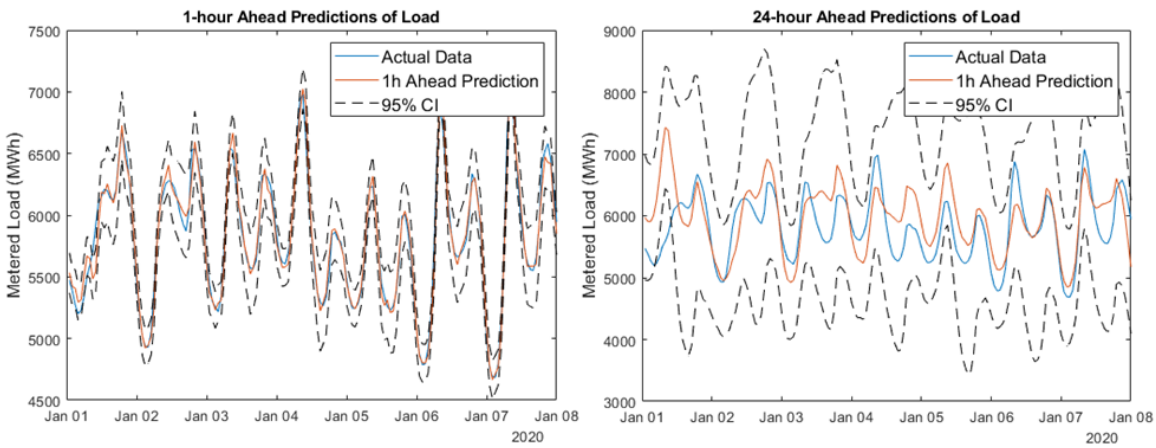


Figure 3: ARMA(36,35) (left) 1-hour ahead and (right) 24-hour ahead predictions of load data over the first week of 2020

Figure 4 shows forecasted load values and corresponding 95% confidence interval bounds over the first week of 2020. Note, forecasted load values are all generated at start of the year indicated by the green vertical dashed line and are increasing in lag moving forward i.e. forecasts are not being updated with time. As expected, the 95% confidence interval asymptotically increases in width for forecasts projected further out into the future due to decaying oscillatory Green's Functions coefficients (See Appendix A). Excluding the first 24 hours of forecasted values due to the New Year's Eve anomaly, all actual load values fall within the confidence interval bounds of forecasted values. However, actual load exceeds 12.5% of forecasted values rendering the scalar time series model unacceptable for 1-week ahead dispatch scheduling.

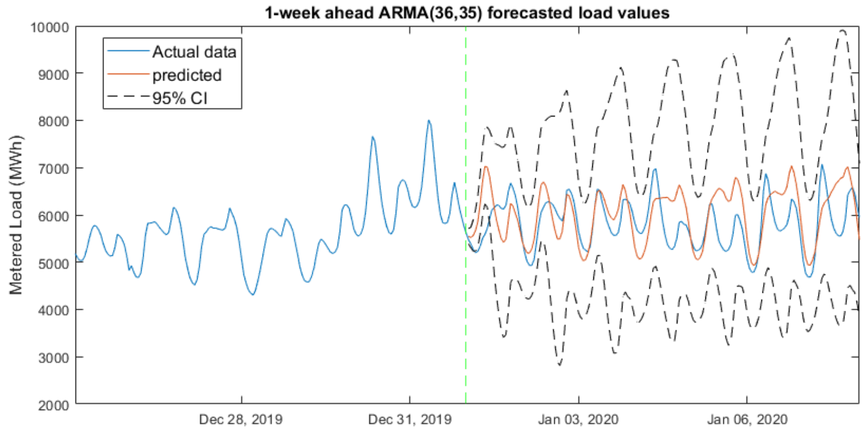


Figure 4: 1-week out forecasts of ARMA(36,35) model for load data showing 1 week worth of load data leading up to test time span. The vertical green dashed line indicates start of forecasting.

3.2 Vectoral ARMA Model

Vectoral ARMA model fitting according to minimum AIC procedure per Section 2.3 resulted in an ARMAV(35,35,34) as the most appropriate model for linear residuals of load data with air temperate as input with zero lag. Iterative F-testing found ARMA(12,11) to be adequate for modeling air temperature, however, autocorrelations of residuals of this model were not acceptable. Therefore, an ARMA(30,29) model was force fitted to temperature data which resulted in uncorrelated residuals. Figure 5 shows the autocorrelations of residuals for the vectoral ARMA load model and cross-correlations of residuals between the ARMA(30,29) temperature model. Similar to the scalar case, both autocorrelations and cross-correlations mostly fall within the 95% confidence interval for whiteness with slight peaks in correlation outside the confidence bounds at intervals of 24 hours none of which exceed 0.1 in magnitude. RSS of the ARMAV load model was found to be 263.28 which is slightly greater than the RSS of the scalar ARMA model suggesting no added benefit of using temperature data in predicting energy load.

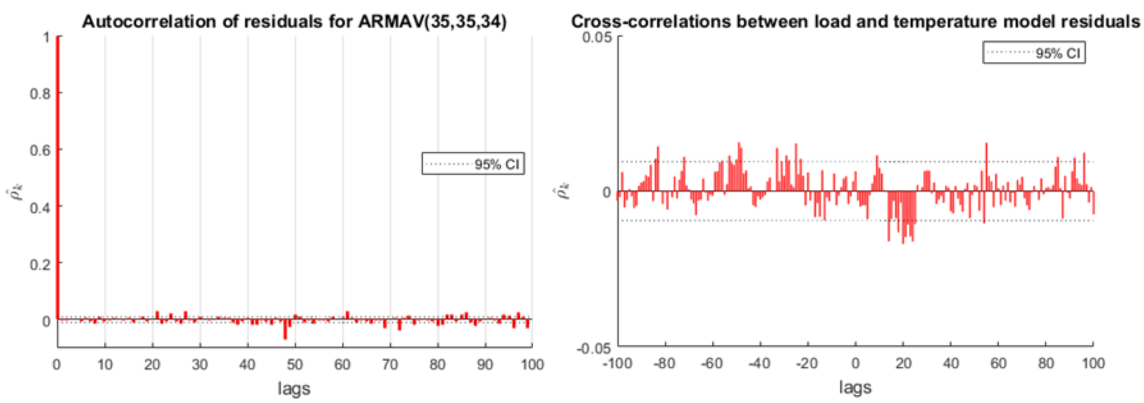


Figure 5: (left) Autocorrelations of residuals of vectoral ARMA model and (right) cross-correlations of residuals between load and temperature models

Figure 6 below shows ARMAV 1-hour ahead and 24-hours ahead load predictions and 95% confidence intervals defined by equation (6). Appendix B contains plots of Green's function coefficients and variance of estimated prediction error with respect to lag for the ARMAV model. Similar to the scaler case, nearly all 1-hour and 24-hour predictions fall within the 95% confidence intervals with the exception of only load values during Jan. 1st, 2020. Actual load consistently remains withing 6% or less of 1-hour ahead prediction making the ARMAV model sufficient for power dispatch scheduling. However, percent difference between actual load and 24-hour ahead predictions does exceed 12.5% with prediction accuracy on par with the scalar ARMA model.

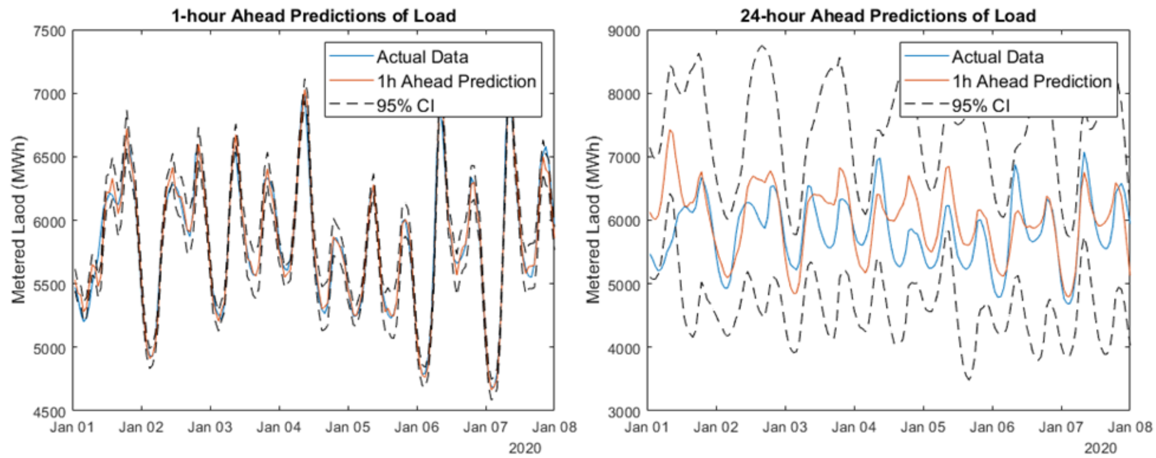


Figure 6: ARMAV(35,35,34) (left) 1-hour ahead and (right) 24-hour ahead predictions of load data over the first week of 2020

Figure 7 shows ARMAV generated forecasted load values and corresponding 95% confidence interval over the first week of 2020 with forecasted temperature values as future inputs to the model. Excluding Jan. 1st, 2020, all actual load values are within the confidence bounds of forecasted load. Surprisingly, all actual load values are within 12.5% of ARMAV model generated forecasts over the testing time span unlike its univariant counterpart. While statistically inconclusive, this result does suggest ARMAV modeling is better suited for predicting actual load and more appropriate for use in real-world power dispatch scheduling. Per ERCOT requirements, the ARMAV(35,35,34) forecasted load values out to one week ahead are, in fact, acceptable for capacity planning purposes.

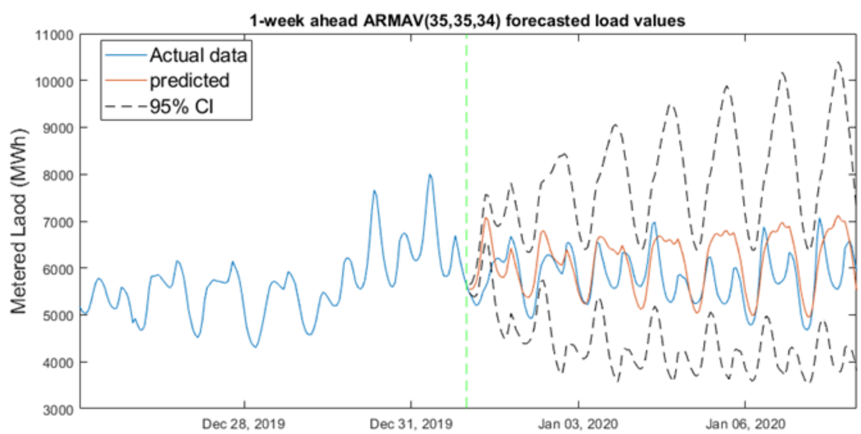


Figure 7: 1-week out forecasts of ARMAV(35,35,34) model for load data showing 1 week worth of load data leading up to test time span. The vertical green dashed line indicates start of forecasting

4 Conclusions & Future Work

Scalar and Vectoral ARMA models were fitting to hourly adjusted load values for ERCOT's South Central weather zone following preprocessing to address heteroscedasticity and non-stationarity present in the data. In the ARMAV case, local hourly air temperatures from AUS airport weather station served as inputs to the model with zero lag. 1-hour and 24-hour ahead load predictions were compared to actual load values over test data time span based on ERCOT's required 12.5% power capacity margin over maximum load. Both the univariant and multi-dimensional model generated 1-hour ahead predictions were within 6% of actual load values, and therefore, acceptable for power dispatch planning purposes. However, neither modeling strategy resulted in 24-hour predictions which meet ERCOT's power margin specification. Additionally, 1-week ahead forecasts revealed ARMA modeling alone is not robust enough to be used in operational capacity planning, whereas, the ARMAV load model with forecasted temperature data as future inputs did produce adequate forecasted load one values out to one week. This result suggests the ARMAV model meets ERCOT specifications and is superior in real-world application compared to scalar ARMA models. Future work includes expanding the ARMAV model introduced in this study to include additional input parameters such as influx of population, more comprehensive weather data and possible consideration for anomaly detection.

Bibliography

[1] "About ERCOT." http://www.ercot.com/about (accessed May 13, 2020).

[2] U. C. Bureau, "Fastest-Growing Cities Primarily in the South and West," *The United States Census Bureau*. https://www.census.gov/newsroom/press-releases/2019/subcounty-population-estimates.html (accessed May 13, 2020).

[3] G. K. F. Tso and K. K. W. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761–1768, Sep. 2007, doi: 10.1016/j.energy.2006.11.010.

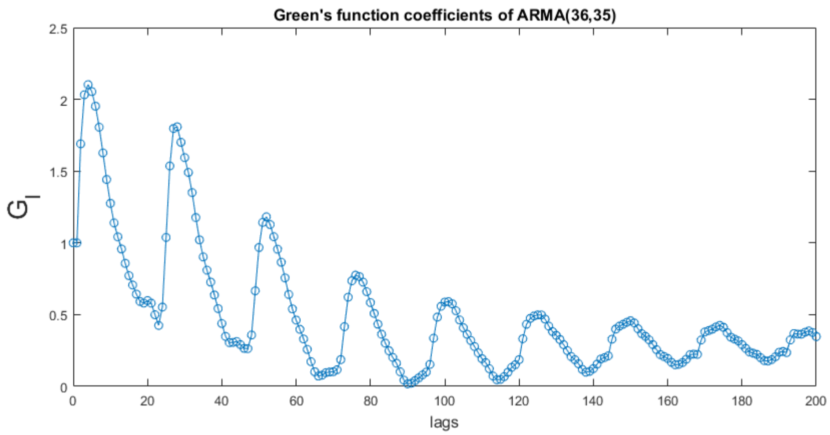
[4] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, Jul. 2017, doi: 10.1016/j.rser.2017.02.085.

[5] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

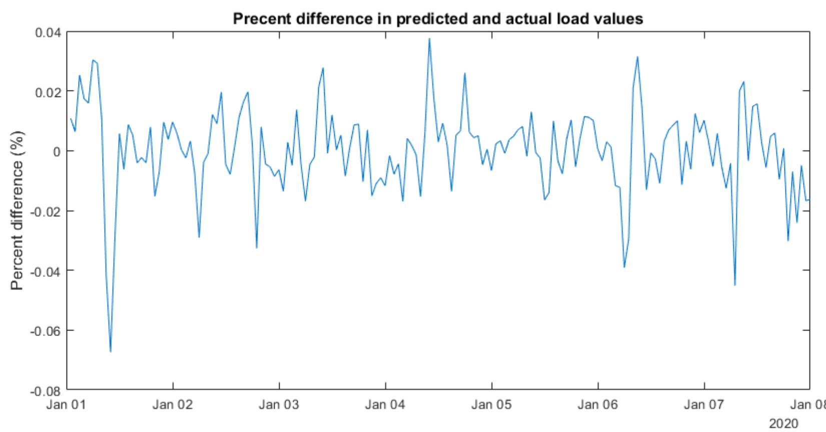
[6] S. M. Pandit, S. Wu, and T. I. Šmits, "Time Series and System Analysis with Applications by Sudhakar Madhavrao Pandit and Shien-Ming Wu," *The Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1924–1925, Jun. 1984, doi: 10.1121/1.390924.

Appendix A:

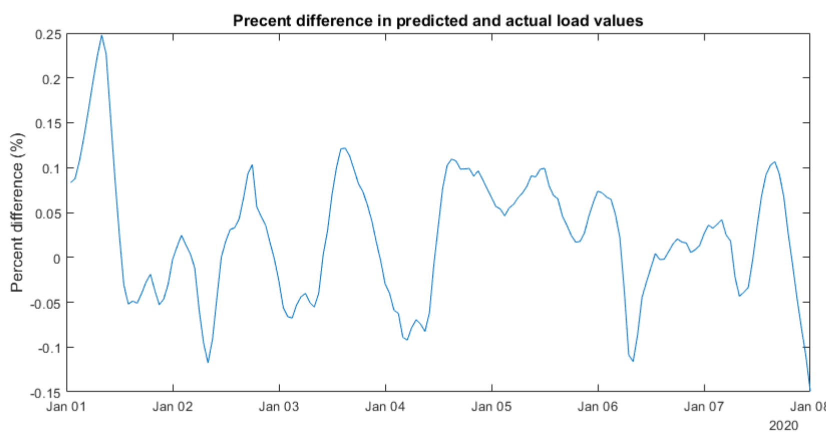
- ARMA model Green's Function Coefficients



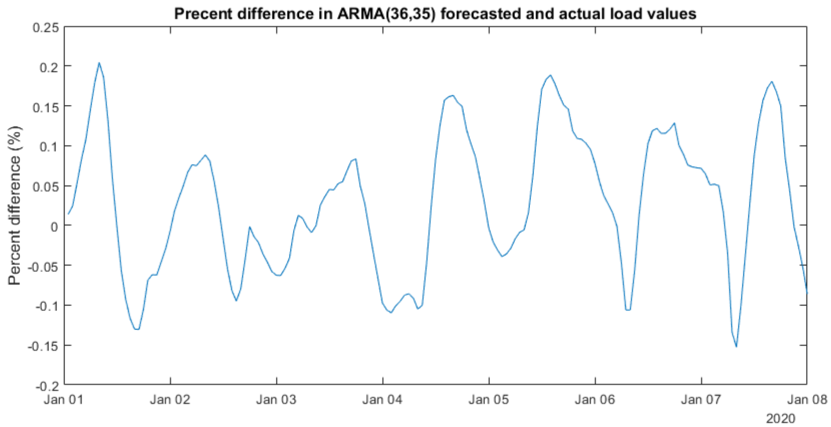
- ARMA 1-hour ahead prediction percent difference from actual load



- ARMA 24-hour ahead prediction percent difference from actual load

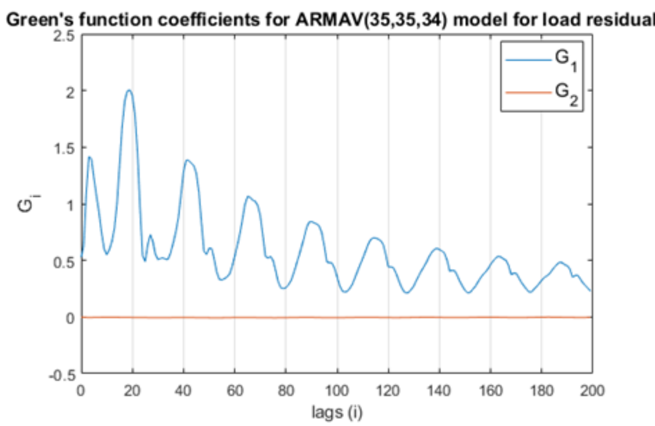


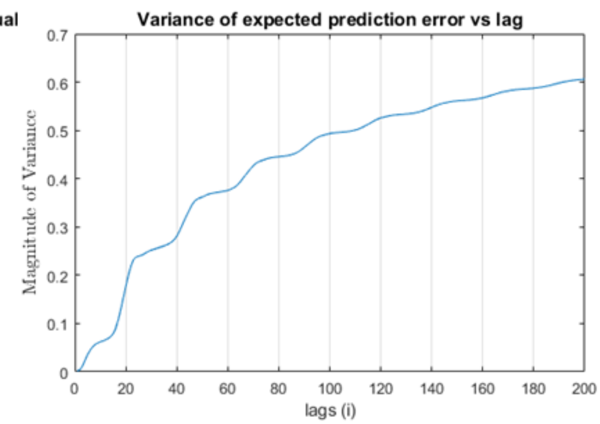
- ARMA 1-week forecast percent difference from actual load



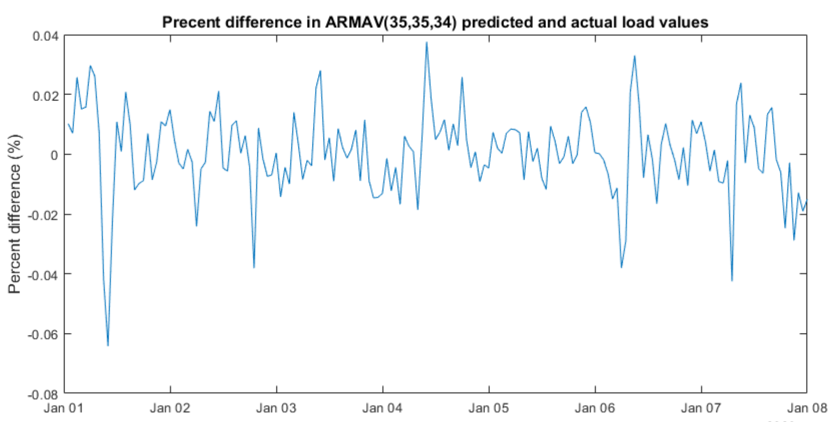
Appendix B

- ARMAV Greens function coefficients and variance of prediction error as a function of lag according to equation (6)

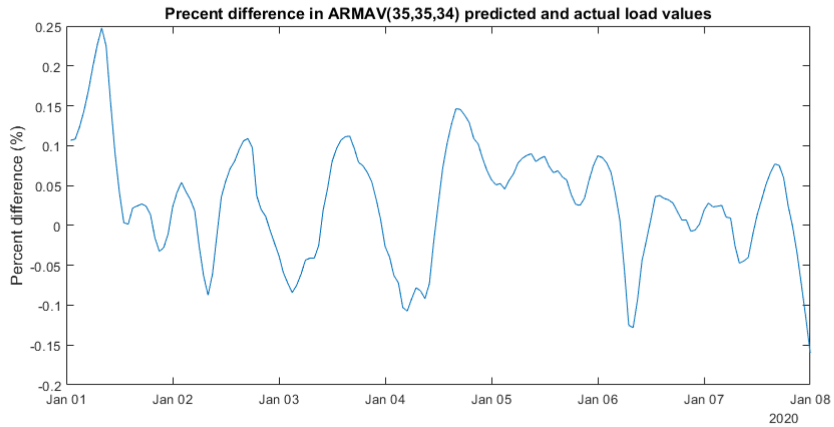




- ARMAV 1-hour ahead prediction percent difference from actual load



- ARMAV 24-hour ahead prediction percent difference from actual load



- ARMAV 1-week forecast percent difference from actual load