SMU Data Science Review

Volume 5 | Number 3 Article 5

2021

Predicting Power Using Time Series Analysis of Power Generation and Consumption in Texas

Joshua Eysenbach
Southern Methodist University, jeysenbach@mail.smu.edu

Bodie Franklin
Southern Methodist University, bodief@mail.smu.edu

Andrew J. Larsen
Southern Methodist University, larsend@mail.smu.edu

Joel Lindsey
Southern Methodist University, jdlindsey@mail.smu.edu

Predicting Power Using Time Series Analysis of Power Generation and Consumption in Texas

Andrew Larsen[1i], Bodie Franklin[1ii], Joshua Eysenbach[1iii], Joel Lindsey[1iv]
[1] Master of Science in Data Science, Southern Methodist University,
Dallas, TX 75275 USA
larsend@smu.edu[i], bodief@mail.smu.edu[ii], jeysenbach@mail.smu.edu[iii], jlindsey@smu.edu[iv]

Abstract. Due to the recent power events in Texas, power forecasting has been brought national attention. Accurate demand forecasting is necessary to be sure that there is adequate power supply to meet consumer's needs. While Texas has a forecasting model created by the Electricity Reliability Council of Texas (ERCOT), constant efforts are required to ensure that the model stays at the state-of-the-art and is producing the most reliable forecasts possible. This research seeks to provide improved short- and medium-term forecasting models, bringing in state-of-the-art deep learning models to compare to ERCOT's forecasts. A model that is more accurate than ERCOT's own models during certain time periods was found. To have the most accurate energy forecasts in Texas, it is recommended that ERCOT investigate using different models in their Coast, Far West, North, North Central and West zones specifically. No models produced in this analysis accurately predicted the actual load that the state experienced during Winter Storm Uri due to the predicted load exceeding the practical grid capacity given the extreme weather. Synthetic data that simulates these types of extreme weather events could aid in training models for prediction in the future.

1 Introduction

Texas electricity generation and consumption models are created by the Electricity Reliability Council of Texas (ERCOT). These models project the highest expected energy demand in the state in the summer and the annual total energy generation for the following decade. In 2021 for example, the forecasted peak demand is 77,144 MW [1]. ERCOT claims that "the grid may still experience temporary tight conditions if a high demand day occurs when one or a combination of the following occurs: there are significant maintenance outages, renewable energy production is low and/or there is extreme weather" [1]. With potentially unexpected factors such as maintenance outages, low renewable energy production and/or extreme weather, it becomes extremely important to be able to accurately forecast supply and demand under different, potentially extreme, conditions [1].

The effects of these extreme conditions can be seen in the February 2021 Texas Energy Crisis. This event took the lives of hundreds of people [2] and left millions without power for days [3]. The crisis was caused not only by record-breaking cold temperatures and snowfall causing extreme demand, but also a severe reduction in power generation across most of Texas' several types of power generators [4].

Transcendently severe weather events can be difficult to predict far in advance. Improved forecasting methods that can consider the effects of extreme conditions on power generation and the potential burden on demand could provide a means for better preparedness to reduce the damage of such events.

This research strives to get a handle on what kind of demand the Texas electrical power supply will need to accommodate in the short- and medium-term. This study analyzes the past power availability and consumption data to see how they have changed over time and how they fluctuate seasonally. This study then analyzes weather and population growth in the state, which are the most obvious variables in determining future electrical burden. Finally, this study attempts to forecast how much supply Texas can expect to need in the future. This study aims to build models for power supply and power demand that perform independently of the current models being used by ERCOT. The independent models will help provide the Texas power grid with a better understanding of how to deal with future demand and mitigate potential future problems. It is understood that the biggest factors affecting power consumption are expected to be population change and temperature. The ERCOT models will be the primary baseline for comparison to the models created in this study.

ERCOT currently has a model that is being used to forecast power generation and load across Texas. Their current model is a set of linear regression models that combine factors such as weather, economic factors and other miscellaneous variables that were determined to affect power grid use [5, 6, 7, 8]. Previous ERCOT forecasts used more advanced Neural Network models despite the current reliance on less advanced models. It was determined that weather (specifically temperature) is the most important factor to short-term energy use, but also the most difficult to forecast [8]. ERCOT has split their forecasts into eight separate weather zones to allow for more accurate temperature forecasts, as Texas is a large state and temperature can vary significantly from location to location (Figure 1).

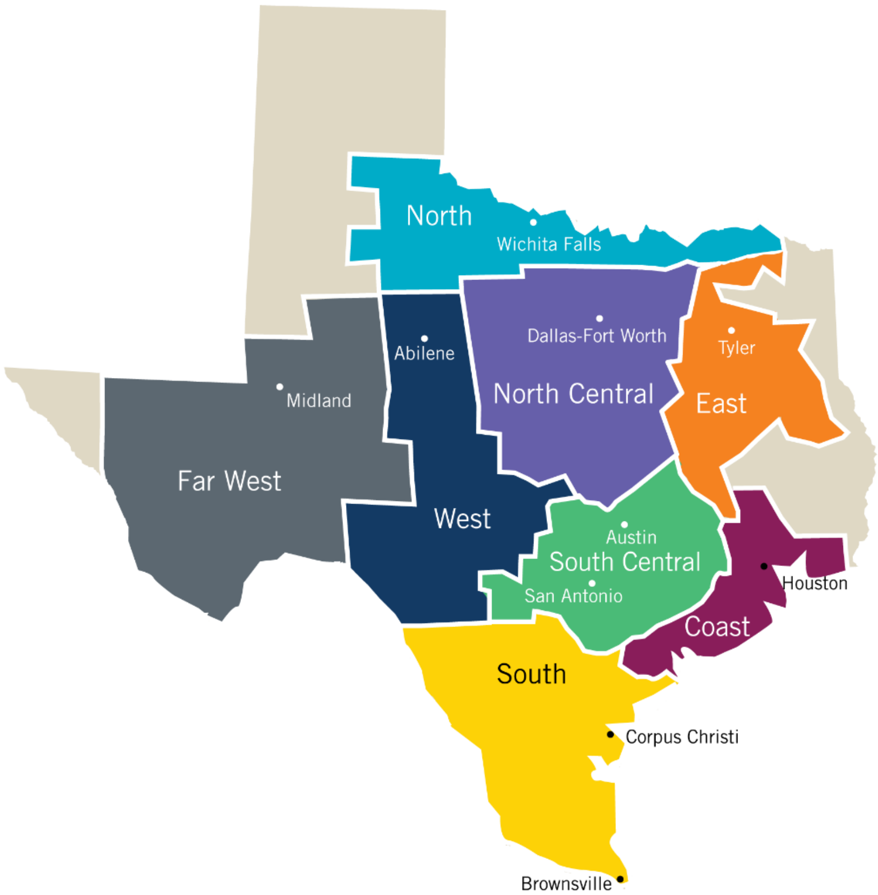


Figure 1: ERCOT Weather Zones [9]

Outside of Texas, most energy models that have been created can be separated into two distinct types: traditional statistical models and deep learning models. Traditional statistical models typically are less complex than their deep learning counterparts and allow for easier model interpretation. Deep learning models have the power to determine more complex relationships between variables. Unfortunately, it can also be difficult to interpret how a single variable can affect the forecast, or how the variables interact with each other to produce the forecast. Deep learning models also take significant amounts of time and resources to train. The choice of model is determined by balancing the importance of training time, interpretability, and accuracy.

Traditional statistical models that were found to be used for forecasting include Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models [10, 11], Vector Autoregressive (VAR) models [12], and regression and trend extrapolation models [13]. For studies that use more advanced deep learning models, traditional statistical models are often used as a baseline to compare to the deep learning model, are included in the model itself in ensembles or are used in the model to forecast a variable [12, 14, 15, 16, 17].

Deep learning models that were found to be used for forecasting include Convolutional Neural Networks (CNN) [16, 18], Recurrent Neural Networks (RNN) [15], Multilayer Perceptrons (MLP) [19], Radial Basis Function Networks (RBF) [19], Support Vector Machines with Gaussian Kernels (SVM) [19] and Support Vector Regression (SVR) [17]. Again, these models can be compared to traditional statistical models or can have the outputs from traditional statistical models used as inputs [12, 14, 15, 16, 17].

This study compares the efficacy of three (3) different time series models to predict future power consumption and generation in order to determine where there could be potential drops in consumption and spikes in generation. ARUMA, VAR and RNN models are built and tested. ARUMA and VAR are statistical models that use more traditional time series modeling techniques, while RNN models are machine learning models. The ARUMA models are univariate, while VAR and RNN are multivariate models. This means that ARUMA only utilizes past data for the target variable being forecasted, while VAR and RNN can include additional factors like population changes and weather. Univariate models are assumed to implicitly contain the effects of other factors in their past data without explicitly modelling using the factors in question. Univariate models do not explicitly consider additional factors and only consider trends of a single variable for forecasting and can be useful in demonstrating how inclusion of other variables in multivariate models can alter forecasts. VAR models are useful for multivariate forecasting with the ability to interpret how the added variables affect the model. RNNs can use machine learning to generate the most sophisticated models but lack the interpretability of simpler methods.

The historical grid data used in this study comes from the ERCOT website's grid data collection from 2010 – 2021 [9]. The data includes generation by fuel type and load by ERCOT weather zone. The weather data is obtained from the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information Division webpage for Climate Data Online [20]. The weather data for each ERCOT weather zone is obtained using the weather station at the airport of the most populous city in each zone. Conditions include various metrics for temperatures, windspeeds, and precipitation. The population data used in this study comes from the Texas Population Estimates Program created by the Texas Demographics Center [21]. They provide population estimates for every county in Texas at six-month intervals based on the previous census and known current trends for each county.

While this study looks at how changes in weather can affect load and generation, weather remains incredibly difficult to predict in the long-term. Predicting weather is outside of the scope of this study.

There are significantly fewer variables included in the models built in this study than those considered and included by ERCOT. This more simplistic methodology provides several important aspects for consideration in comparison to ERCOT's strategy. The first is to investigate the necessity and reasonability of the complexity of ERCOT's models in contrast to those that are simpler. If models that include fewer and more easily obtainable variables produce similar results, the time and effort expended in capturing non-essential information could be drastically reduced.

The second is a potential increase in interpretability of the models and the included variables' effects on forecasting. This can be useful for better understanding how a change in population or an extreme weather event can affect power generation and consumption, without having to worry about other less crucial factors.

The downside to generating models that use less variables is the possibility of less accurate models. There could also be variables that were not considered in this study that incidentally have greater impact than they were assumed to have.

The models created in this paper are compared to ERCOT's forecasts as of 12/31/2020 for 1/1/2021 - 1/7/2021 and as of 2/12/2021 for 12/13/2021 - 12/19/2021. These periods are chosen because one was a "normal" winter period, and one was during Winter Storm Uri where widespread power outages plagued the state. The models created in this paper outperform ERCOT's models for the Coast, Far West, North, North Central and West regions during the "normal" winter period. While this is a single data point and is not convincing evidence of issues with ERCOT's models, it is recommended that ERCOT investigate their models for the Coast, Far West, North, North Central and West regions to make sure they have the most accurate models possible.

No models performed well when compared to the actual load that was experienced during Winter Storm Uri. This is due to the predicted load being higher than the grid could handle, given the conditions. Because of this, it is recommended that ERCOT create synthetic data that simulates these types of events in order to gain a better understanding of what load can look like in extreme conditions.

2 Literature Review

This section is focused on research that has been done to create the most accurate demand generation models. While Texas has their own forecasting models [5], there have been numerous papers published on power generation in other locations. As this study's focus is to create a model independent of ERCOT's current models, the methods used will be guided by models used outside of Texas.

There are numerous factors that affect power generation and demand, increasing the complexity of the model needed to create accurate forecasts. Wind generation alone is affected by high energy cost, increasing desire for renewable energies and the development of modern technologies [22]. Due to the complexity of power forecasting, numerous models and variables were considered.

2.1 ERCOT Forecast Methods

Forecasting within Texas has mostly been performed by ERCOT and at times has been evaluated and revised by third-party companies such as Itron, Inc.

ERCOT forecasts for predicting future energy load are largely based on weather patterns, but the accompanying variables included in ERCOT's models that account for growth have changed over the last decade. It has been well documented in ERCOT's reports that the weather variable (most notably temperature) is both the most important to demand fluctuations and the most difficult to predict. Still, forecasters have experimented with the factors that affect the growth of energy use in long-term projections [8].

ERCOT recognized in 2013 that forecasts needed to account for a changing relationship between energy use and economic variables like Gross Domestic Product and employment which could be attributed to factors like energy efficiency and behavior changes [6]. This led to changing some of the key variables included in predicting the growth of future loads as well as modeling with a more modern technique in Neural Networks with help from Itron, Inc. Forecasting and Load Research Solutions (Itron) [7, 8].

The growth indices for the load forecasts were modified from using economic growth based on employment and general population increases to instead creating premise forecasts (forecasts of residential, industrial, and business locations growth) which were combined as a weighted average into a growth index [7]. This growth index was then included with the weather forecasts in each model as base parameters [7]. This method of using premise forecasts was found to be more stable and reliable than former methods using population growth. The same system of residential, business, and industrial premise forecasts is still in use in the most current load forecasts which aggregate a set of autoregressive (AR1) models [5].

ERCOT introduced forecasts for each of its eight weather zones based on Neural Networks beginning in 2013 with help from Itron [23]. The neural net models were validated by using regression models to formulate a baseline for comparison as well as to continue to track the performance progression of the neural net models. It was noted in these early models that the use of a neural network "decoupled the growth in demand and energy" as well as allowed for calculation of forecast sensitivities using multiple variations of models, but that interpretability of model parameters and causes for changes in forecasts was far more difficult [7]. These early neural nets were reduced in complexity by reduction to a single node to isolate the growth index for interpretation and improve model stability [7].

Despite the inclusion of neural nets in past models, the current long-term forecasts are produced by ERCOT using a set of linear regression models that combine the premise forecasts with weather and other factors [5]. The overall aggregate forecast for energy demand predicts gross summer peak demand, which is the highest expected load each year [5].

The weather component of the current ERCOT forecast uses the past fifteen years of weather data, and the grid is divided into eight weather zones that are forecasted individually [5]. The summation of these eight weather zone forecasts produces the overall ERCOT forecast. Annual forecasts using each individual year from 2005 to 2019 combined with the current premise forecast are calculated and the results are averaged [5]. These forecasts are calculated for every hour in a calendar year to identify and rank when the peak load is predicted to occur in each month and year [5].

The main forecast presented is based on peak summer demand, but peak winter demand is also calculated [5]. Various sensitivity cases are run on these models based on different weather scenarios as well as variations of several of the other included factors [5].

The newest factors added to the ERCOT forecast in 2021 that influence the growth index include predictions based on changes in on-site power generation, home solar panel use, and electric vehicles [5]. These are included in a list of variables ERCOT notes as extremely uncertain growth factors along with energy efficiency improvements, loads responsive to price changes, and new large industrial establishments. The most recent ERCOT reports have also noted that Lubbock will be integrated on the grid in 2021 and this currently presents challenges in determining what additional demand will be put on the system [5, 6].

ERCOT also makes use of “mid-term” (7-day) and “short-term” (next hour) load models and forecasts [6]. The “mid-term" 7-day forecasts (which this paper will later refer to as “short-term” forecasts when comparing the models created in this study to ERCOT’s own models) are currently developed daily based on seven different available models [6]. Two of the models are legacy and five are internally developed that blend regularly tuned autoregressive and non-autoregressive models [6]. The model used to publish the load forecast is decided on by ERCOT operators each day [6]. These models can include many unique measured weather variables and several calendar variables [6]. Each day, ERCOT also assesses the potential for forecast variability for the next three days by assigning a classification of low, medium, or high chance of forecast variability based on the uncertainty of current and near-future conditions [24].

2.2 Other Forecasting Methods

Most research done on energy demand forecasting has focused on areas outside of Texas. While these results are not directly applicable to Texas due to differences in geography and consumer behavior, they do motivate the methodology used in this research. The previous research that was done can be split into two focus areas: research that focuses on traditional statistical models and research that focuses on applying state-of-the-art machine learning models.

2.2.1 Traditional Statistical Models

Traditional statistical models have been developed and used for many years to forecast power consumption and generation. These methods include regression and time series approaches. These models are still used in research today and provide a valuable baseline for the more advanced machine learning and deep learning models. Unlike deep learning models, traditional statistical models are more interpretable and thus are sometimes favored over more complex models for their ease of interpretation. These models have been used for short-, medium- and long-term power forecasting.

Traditional time series methods have typically been the favored statistical models for forecasting power demand. Since power demand is typically represented as the amount of energy used over time, time series analysis methods have more obvious applications than other statistical methods. Boroojeni et al. uses seasonal and non-seasonal ARMA (Autoregressive Moving Average) methods [11] while Hwang et al. uses ARIMA (Autoregressive Integrated Moving Average) methods [10] to forecast. ARMA and ARIMA are univariate methods and thus are only able to consider a single variable. These models are only able to use previous demand and time to forecast future demand. Thus, they are not able to consider extra variables such as temperature or economic conditions to forecast future demand. Univariate models assume that these extra variables are latent variables, with their information already encoded in time and power demand. While they are not the most advanced models, ARIMA is used as a valuable baseline to compare to more complex models in previous research [15, 16, 17].

Multivariate statistical models were also considered in previous research. Vector Autoregression (VAR) was used by García-Ascanio & Maté in an ensemble model with an interval Multi-Layered Perceptron (iMLP) and outperformed the more complex iMLP in certain circumstances [12]. Wang et al. ensembled a weighted combination of a regression model and a trend extrapolation model to bring macroeconomic situations and development trends into their model [13].

These simple models can also be used in conjunction with more complex models. Craig et al. uses a piecewise function of the temperature multiplied by the slope found through linear regression to create a single piece of their more complex model [14]. García-Ascanio & Maté found that their VAR model outperformed their Multi-Layered Perceptron (MLP) model in specific cases and use a hybrid model where they use their MLP when certain factors are true and use their VAR model when other factors are true to increase their accuracy [12].

Traditional statistical models have been used for many years. The rapidly changing landscape of power generation, the availability of more data and the increasing use of machine learning models have pushed forecasting research toward newer state-of-the-art machine learning and deep learning methods.

2.2.2 Deep Learning Models

In addition to statistical methods, researchers have turned to machine learning to forecast demand on short-, medium- and long-term horizons. Most of the research has been focused on building variations of Neural Networks and Support Vector Regression (SVR) models. The main drawback to Neural Networks is the training time and computational requirements needed to implement. Each paper reviewed provided a framework of how to address these shortcomings to forecast power demand.

One of the simplest approaches to address the large amount of data was found by researchers del Real et al. and Kang et al. [16, 18]. Both research teams used a Convolutional Neural Network (CNN) to forecast power demand. CNN automates feature learning from the data input and requires no additional effort from the researcher. CNN "have sparse interactions" therefore "they use a fewer number of parameters (weights) with respect to fully connected networks" [16]. CNN performed very well in relation to other traditional machine learning models [16, 18]. The benefit of CNN is that feature learning can be automated, and CNN will provide a strong model [16, 18].

Despite the impressive prediction scoring of Neural Networks, there are drawbacks. For example, Recurrent Neural Network (RNN) has the drawback that as the number of layers increase within the network, “RNN will appear as gradient disappearance and gradient explosion” [15]. In simpler terms, if more data is used as input into the model, training time and computational requirements will increase significantly. Often electricity data has multiple dimensions and dimensionality will need to be reduced. The researchers Xu Et al. provides a framework of modeling techniques that are well known to forecast power demand [15]. It is a frequent practice to reduce data dimensionality through PCA (Principal Component Analysis). PCA works by calculating new variables as linear combinations of the raw variables. The construction is performed so that the new variables are not correlated, and the information of the old variables is compressed into the new components. This reduces the dimensionality of the data, therefore training time and computing requirements are significantly reduced. After PCA, the researchers inputted the PCA data into an LTSM (Long Short-Term Memory network) [15]. LTSM addresses the shortcomings of RNN because LTSM can draw more information from longer input [15]. Although, LTSM will suffer from longer computing time if PCA is not introduced. The model scored an impressive RMSE (Root Mean Squared Error) of .48. PCA-LTSM score significantly better than the other models within the study.

One of the more interesting approaches to forecasting demand was used by researchers Ciechulski & Osowski [19]. Before the data was inputted into the model, the researchers applied a deep autoencoder to the data. The purpose of the deep autoencoder was to act as a diagnostic feature [19]. The researchers describe autoencoding as “multilayer nonlinear generalization of PCA” [19]. Autoencoding is the usage of a neural network to reduce the vector size of the data. The deep autoencoder data was fed into three neural networks: Multilayer Perceptron (MLP), Radial Basis Function network (RBF) and Support Vector Machine with a Gaussian Kernel (SVM). The goal of the models was to forecast short-term demand of 24 hours by using MAPE (Mean Absolute Percentage Error) as scoring metric of the model. Their model results suggest that RBF performed the best at short-term forecasting with an MAPE of 1.47%, although MLP and SVM were comparable with scores of 1.96% and 2.44% respectively [19]. The significance of this research is that it provides another framework of dimensionality reduction. The application of this could aid in future demand forecasting.

Another popular modeling technique for power demand is Support Vector Regression (SVR). Researchers Acakpovi et al. outlined the benefits and drawbacks to utilizing SVR [17]. SVR is effective with “high-dimensional spaces and memory efficiency” [17]. The main drawback to SVR is that they do not work well “for large datasets” [17]. Therefore, SVR can provide a comparable for Neural Network performance.

2.2.3 Climate Change Effects

It must be noted that climate change will have a major impact on energy consumption and generation in the long-term. Craig et al. used the RCP 8.5 model to forecast weather conditions. They created an electricity demand model using a piecewise function as it was found that power consumption increases non-linearly as temperature changes [14]. This paper also noted that extreme temperatures decrease power plant output due to cooling issues but increase solar generation [14]. While climate change will severely affect power generation and consumption in the long-term, this study is only forecasting out to a maximum of 151 days (about 5 months). It is not expected that climate change will have a significant impact in such a brief period.

This study aims to build models for power supply and power demand that perform independently of the current models being used by ERCOT. The independent models will help provide the Texas power grid with a better understanding of how to deal with future demand and mitigate potential future problems. It is understood that the biggest factors affecting power consumption are expected to be population change and temperature. The ERCOT models will be the primary baseline for comparison to the models created in this study.

3 Methods

This analysis focuses on creating multiple competing models in order to create the most accurate forecasts possible. Many prior results used Autoregressive Integrating Moving Average (ARIMA), Vector Autoregressive (VAR) and Neural Network models in order to create their forecasts or as baseline models [5, 7, 10, 11, 12, 14, 15, 16, 17, 18, 19]. Likewise in this study, Seasonal ARIMA (ARUMA), VAR and Neural Network models are built in order to compare multiple models to be sure that the model found is as accurate as possible.

ARIMA models are univariate statistical models, so the only variable being used in the model is the variable that is being predicted. ARIMA models take the variable at previous time steps as inputs into the model and are trained to give these previous time steps a certain weight in determining the forecasted value. ARIMA models differ from ARMA models in that they can model non-stationary data. Non-stationary data is data without a constant mean, a constant variance, a constant autoregressive structure or a combination of the three. The load data does not have a constant mean, so ARIMA models must be used.

ARUMA models take ARIMA models and add a seasonal component to them. Data that repeats weekly, monthly or yearly typically takes advantage of these seasonal models by subtracting the trend from each data point. If the data is daily and the trend is yearly, each data point 365 days in the past is subtracted from each data point. In subtracting the trend, the data is made stationary and can be analyzed using ARMA models.

VAR models are multivariate statistical models, so these models can include weather and other factors that were determined to be important. VAR models are like ARIMA models where they take input from a certain number of time steps previously, but they differ in being able to use time series of multiple variables. A VAR model is a linear function of the past lags of both itself and of past lags of the other included variables. This can be used to create a forecast that is based on more information than just the past data of the variable of interest. Forecasting one time step forward will forecast all variables included in the VAR model one time step forward. So, in this case the VAR models can include weather and other pertinent variables. Both the ARUMA and VAR models were developed in R using the *TSWGE* package, and the VAR models additionally used the *vars* package. The VAR models are created by testing with one weather zone to determine which of the available variables could improve the forecast over using past load data alone. The variables are explored within various test comparisons to historical load data over different time periods to determine which are consistently suitable and significant contributors.

The neural networks used in this study are Recurrent Neural Networks (RNN), specifically Long Short-Term Memory Networks (LSTM). Since the data is in a time series, it was determined that time series analysis focused models must be used. RNNs are Neural Networks that take the result of the previous time step as input. LSTM models are a subset of RNNs. A simple RNN model can have issues in "remembering" things that happened a few time steps back. On the other hand, LSTM models are trained in a way that allows the model to "remember" important things that happened a significant amount of time steps in the past. Because of this, these models require a significant amount of data to train. LSTM models were created using *TensorFlow* in Python.

To create the inputs and outputs for the neural network, a windowing procedure was used. For each day in the training and testing data, a predetermined number of days are used as the inputs and outputs for the model. For example, the model would take in data from 6/23/2019 - 6/30/2019 and output predictions from 7/1/2019 - 7/14/2019 if the model were run on 6/30/2019. If the model had last year's data as its input as well, the input would also include 6/23/2018 - 6/30/2018. The next data point would then be "as of 7/1/2019", and its inputs would be 6/24/2019 - 7/1/2019, and its output would be predictions for 7/2/2019 - 7/15/2019. This windowing procedure allows for almost every day in the data set to be used as a training point but does require a choice to be made for how many days will be used as inputs and outputs. The input was determined through a grid search procedure, and the outputs were determined to be 14 and 151 days (about five months) as short-term and medium-term predictions.

The neural network parameters for short- and medium-term predictions in each of the eight weather zones were found using a grid search procedure. The parameters being searched for are the number of layers in the neural network, the number of nodes in each layer of the network, the number of days to use as inputs into the model and whether to use last year's data as inputs in the model. Each set of parameters was run five times and the mean of the Average Squared Error (ASE) of these five runs were taken. The top models were chosen based on a combination of their train and test ASEs. These top models were then used to predict the holdout data set in order to directly compare the neural network models to the VAR and ARUMA models. The results of this grid search can be seen in Appendix Table 1, Appendix Table 2 and Appendix Table 3.

The historical grid data used in this study comes from the ERCOT website's grid data collection from 2010 – 2021 [9]. The data includes generation by fuel type and load by ERCOT weather zone. It is provided as individual comma separated files in various formats and various time intervals; all files are combined, and the data is aggregated to represent one-day measurements of each variable in a single comma separated file.

The weather data is obtained from the NOAA National Centers for Environmental Information Division webpage for Climate Data Online [20]. The weather data for each ERCOT weather zone is obtained using the weather station at the airport of the most populous city in each zone. Based on examination of the available weather station data, it was found that airport weather stations collect the most comprehensive set of daily weather conditions. These conditions include various metrics for temperatures, windspeeds, and precipitation. The NOAA weather data is also aggregated to daily observations and compiled into the same comma separated file that contains load and generation data.

The population data used in this study comes from Texas Population Estimates Program under the Texas Demographics Center [21]. They provide population estimates for every county in Texas at six-month intervals (January 1 and July 1 of each year) based on the previous census and known current trends for each county. For use in this study, the population estimates are aggregated from the county level to the weather zone level. To match this data with the daily level of granularity that was provided by ERCOT, population increases between January 1 and July 1 are assumed to be linear. This is not likely to be an exact representation, but population changes for each weather zone did not increase enough to suggest that this assumption would be problematic for creating models. The rate of change between each population data point is calculated, and the data is interpolated between each six-month estimate and linearly extrapolated from the last available estimates to achieve a daily level of granularity.

Since weather is being used as an input into the multivariate models, future weather forecasts need to be created to forecast load. VAR models automatically create predictions for all variables included in the model based on past data. Due to the data windowing that is being used to fit the neural network, only historical weather data is being used to predict future load and generation. Because of this, the neural network models produced in this paper do not consider weather forecasts, only historical weather conditions. Any forecast more than a few days out created using the neural network model will likely rely on more stable metrics, like population.

Although ten years of data are available, not all the final models examined use the entirety of the ten years of available data. ARUMA and VAR, for example, may not have any functional need for data more than one or two years in the past depending on the parameters.

The periods that are being used to compare the models are 1/1/2021 - 1/7/2021, 2/13/2021 - 2/19/2021 and 1/1/2021 - 5/31/2021. These periods will allow comparisons of models during a historically "normal" 7-day forecast, a 7-day forecast with a significant weather event, in this case the winter storm of February 2021 (Winter Storm *Uri*), and a medium-term comparison. The metric used to compare the models is Average Squared Error (ASE). The 7-day forecasts are also compared to ERCOT's 7-day forecasts ("mid-term" forecasts as defined by ERCOT but referred to as "short-term" or "7-day" in this paper) for the two weeklong test periods. ERCOT's current 7-day forecasts are only available online for a short time, so the forecasts for 1/1/2021 - 1/7/2021 and 2/13/2021 - 2/19/2021 are obtained by contacting ERCOT directly and requesting the data. Figures 2 and 3 plot the forecasts published by ERCOT as well as the actual measured total system daily loads during and prior to these two periods.

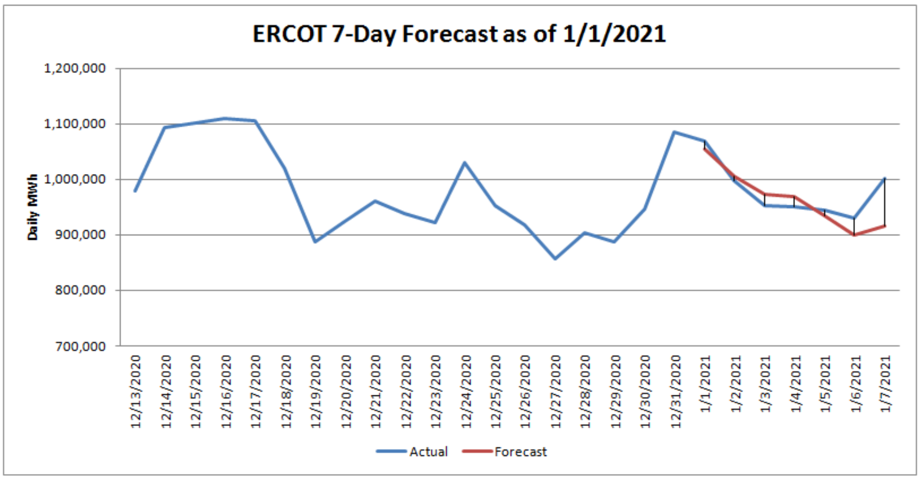


Figure 2: ERCOT 7-Day Load Forecast for 1/1/2021 - 1/7/2021

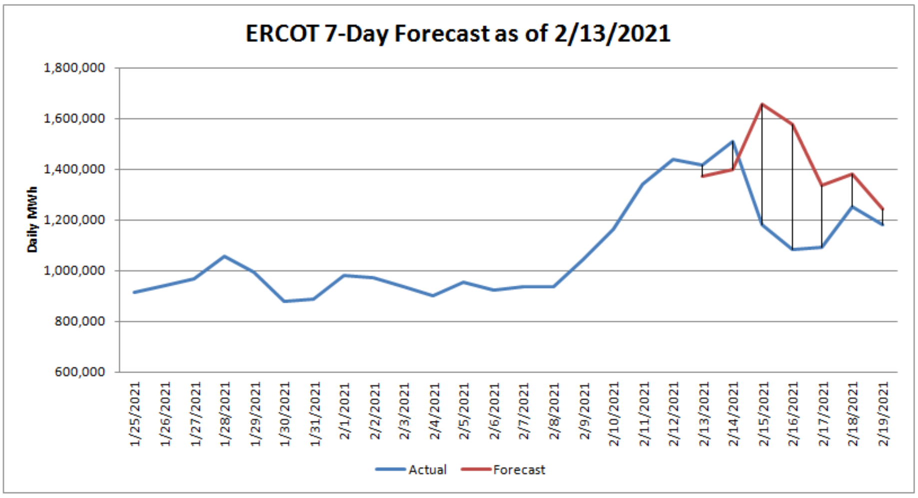


Figure 3: ERCOT 7-Day Load Forecast for 2/13/2021 - 2/19/2021

The figures below from ERCOT's 2021 Long-Term Demand and Energy Forecast Report [5] display the results of their Long-Term Demand Forecast model for Peak Summer System Demand (Figure 4) and Annual Energy Forecast (Figure 5). The report notes that the demand model forecasts an average annual increase of 1.2% in peak system demand for 2021-2030 despite a historical average increase of .9% per year from the previous decade. The Energy Forecast also forecasts a higher average annual rate of increase of 2% to 2030 versus a ten-year historical average increase of 1.5% [5].

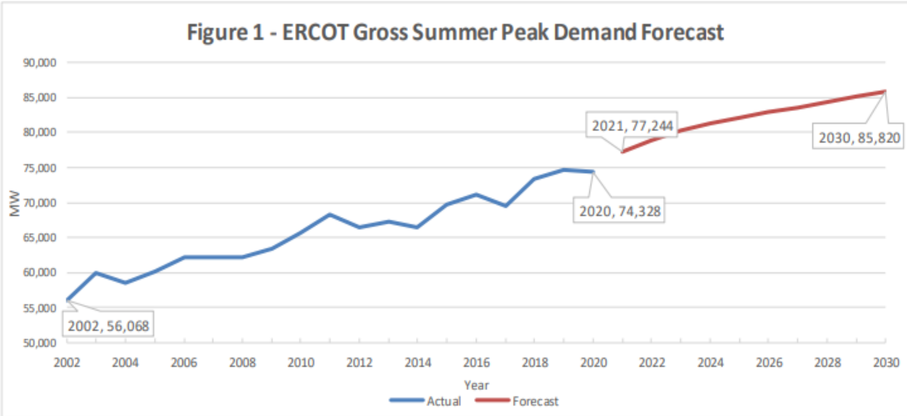


Figure 4: ERCOT Gross Summer Peak Demand Forecast [5]

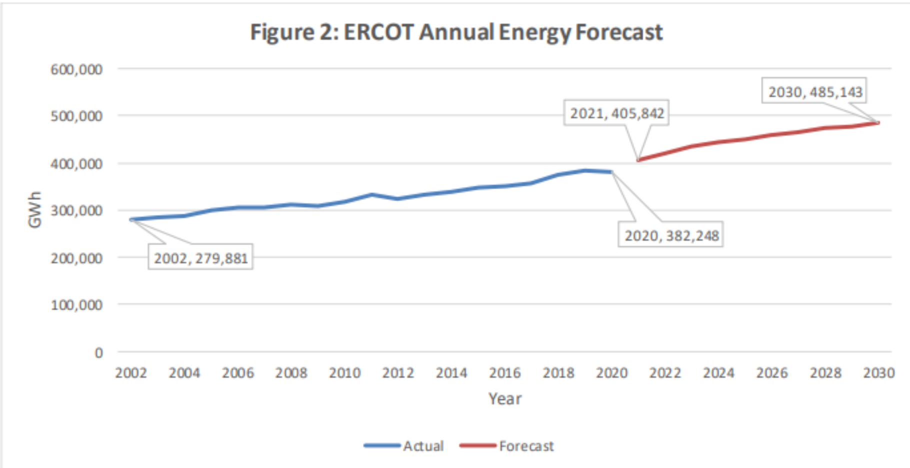


Figure 5: ERCOT Annual Energy Forecast [5]

When making longer-term validations of the models in this analysis, their trends should also be compared to the long-term forecasted increases from ERCOT. However, the models conceived in this analysis are only run during the period of January through May of 2021 for long-term comparison to actual measured values. The increases predicted by ERCOT in the 10-year forecast may not be as apparent in a much shorter span.

4 Results

There are a few things about the structure of the data that affects how data modelling is done. As discussed in the methods section, the data is a time series so it must be analyzed using time series methods. Appendix Figure 1 and Appendix Figure 2 display the historical load for each of the weather zones. Seasonal tendencies are clearly visible for all of the weather zones. There is also a distinct overall increasing trend over the ten years of available data for several of the zones. These indicate non-stationary behavior. Techniques that model stationary time series, such as ARMA, are not appropriate to analyze this data, especially in longer-term situations. This suggests that ARUMA and VAR, described in the methods section, must be used to analyze this data correctly. Overall, it looks like most zones have a similar long-term trend if the magnitude of load is ignored. This suggests that most zones can use a similar model with adjustments only to certain parameters.

Appendix Figure 3 has the correlations of the different load and generation types. Note that load in each location is highly correlated with load in all other locations, excluding far west load. Despite the size of the state, it seems that high loads in one location correlates with high loads in another location. This is likely due to the seasonality of load. Summer load is higher than winter load, regardless of location.

The data for generation in Appendix Figure 4 and Appendix Figure 5 display similar seasonal characteristics to load, which is expected as power generation is designed to accommodate changing seasonal power demand. However, the seasonal trends appear to behave differently for different power source types and not all power sources have a clear seasonal trend on their own. The plots also show how overall trends over the past ten years vary between sources. Renewables such as solar and wind are making up a larger portion of total generation while the contribution of coal is decreasing. This result suggests that significantly different models may be needed to accurately predict generation for each power supply and will be heavily dependent on future expectations for how power sources will change.

4.1 Model Results

4.1.1 ARUMA

The ARUMA model was intended to be the simplest model used in this study. The only input into the ARUMA model was the load data for each respective weather zone. Due to the simplicity of the model, historical data from 2010 to 2020 was used to train the ARUMA model. The 10-year history for each respective weather zone was log transformed to reduce skewness. During the model building process, the r package *TSWGE* function 'aic5.wge' was utilized to determine the values for *phi* and *theta* for each respective weather zone. It is important in the ARUMA building process to difference the data and remove seasonality. The effect of taking the first difference of the data and removing seasonality is to "stationarize" it. In other words, these methods enable the data to meet the assumptions of stationarity.

Since the data that was used as the input into the ARUMA model was log transformed, the forecast of the ARUMA model was log transformed as well. The ARUMA forecast was inverse log transformed to accurately compare against the ERCOT Model's forecast. The forecasting periods used for comparison were data as of 12/31/2020 to predict 1/1/2021 - 1/7/2021 and data as of 2/12/2021 to predict 2/13/2021 - 2/19/2021. The first week of January 2021 was used to provide an example of how a model would perform during a normal week or a week without extreme weather fluctuations. The February 2021 dates represent how a model would perform with extreme weather fluctuations.

The ARUMA model performed better than the ERCOT models for 1/1/2021 - 1/7/2021 for the Coast, Far West, North and West weather zones (Table 1). The ARUMA model performed better during the 2/13/2021 - 2/19/2021 period for Coast, Far West, North, and West regions than the ERCOT model (Table 2). The ARUMA model is still severely inadequate for capturing extreme deviations.

Weather Zone	ERCOT Model January 1 - 7, 2021	ARUMA Models January 1 - 7, 2021
Coast	324,735,334	97,971,661
East	7,519,166	11,007,423
Far West	9,795,965	7,316,770
North	11,571,249	1,116,671
North Central	382,483,068	566,309,208
South Central	39,659,348	208,869,299
Southern	6,503,418	16,366,323
West	55,786,132	6,675,611

Table 1: ARUMA Models: Average Squared Error for Load from 1/1/2021 - 1/7/2021

Weather Zone	ERCOT Model February 13 - 19, 2021	ARUMA Models February 13 - 19, 2021
Coast	6,499,966,149	2,714,164,886
East	88,332,434	406,926,062
Far West	2,000,900,363	1,775,402,632
North	86,290,196	72,984,110
North Central	8,789,099,199	28,127,438,465
South Central	1,183,362,088	9,949,946,871
Southern	575,913,544	1,777,955,558
West	226,078,908	105,533,741

Table 2: ARUMA Models: Average Squared Error for Load from 2/13/2021 - 2/19/2021

4.1.2 VAR

Predicted load using VAR is intended to improve on ARUMA models that use only past load data by including additional variables. Much of the transient weather data such as precipitation and wind are sparse by nature and thus not suitable for vectorized models, so they are excluded. Based on analysis of past data and ERCOT's own admissions, temperature is by far the most influential factor in determining load, so the primary focus of the VAR models is on inclusion of the daily maximum and minimum temperatures. Long-term VAR models include an annual seasonal factor to capture the expected fluctuations that occur regularly from year to year. In the model used for long-term forecasting, the population data is also included to account for any trends due to population changes.

Most of the testing for determining the best use for the VAR model is performed on the Coastal Weather Zone but using the same model-building methodology on other zones produces very similar results, as they followed similar trends in the past year. The VAR model for each weather zone is fit by first using the *vars* package to estimate the best parameter for p, which is the number of lagged components to include for each variable in the function. The best value for p is based on the AIC (Akaike Information Criterion). For the long-term models, this is determined using multiple years (the number depending on the length of the desired forecast) of data for load, minimum temperature, maximum temperature, population, and a 365-day seasonal component.

The length of historical data to use when determining p and running the model for longer-term forecasting affects the behavior of the VAR forecasts produced. When a shorter training history is used, such as one year or less, the forecast has more day-to-day variability. As the historical input increases up to the total ten years of available data, the day-to-day variability for the same forecast period decreases substantially and is dominated by the central seasonal trend. This also affects the significance of the population variable, which is only marginally significant when nearly all the available ten-year history is used. The final VAR models for the 151-day forecast use two or three years of history (based on differing performance results for different weather zones) to balance the overall seasonal trend with daily fluctuations, but this model is unable to capture any severe deviations from historical trends.

Notably, the total load during the winter storm did not behave the same for each weather zone. It appears some areas, like the Eastern Zone, did not have significant curtailments and outages that caused a large drop after the sudden extreme rise in energy use. The Eastern Zone reached what appears to be a maximum load and remained flat at this constant load for the duration of the week. This also is a very atypical behavior that inhibits the applicability of forecast models, especially those that rely heavily on autoregressive functions such as VAR.

For the short-term models, the last 365 days of data are used, and both the population and the seasonal component variables are removed. Separate values for p are determined for every forecast period as they are observed to change based on the nature of the historical data used to estimate the best value based on AIC. In most cases, the estimated value for parameter p is lower for the February 13^{th} – 19^{th} 7-day model than for the model for January 1^{st} – 7^{th}. The summary of each produced VAR model shows that the most significant lagged value for each variable can change dramatically based on the time frame examined, but the last lagged variable (furthest back in time) of those included is almost always significant indicating the estimation method for parameter p is acceptable. The values for parameter p are shown in Appendix Table 1.

The results of the short-term VAR models when compared to ERCOT's 7-day forecasts for 1/1/2021 - 1/7/2021 and 2/13/2021 - 2/19/2021 are shown in Table 3 and 4, respectively. While VAR models perform similarly or better to ERCOT's forecasts for the first week of 2021, they are severely inadequate for capturing extreme deviations from normal historical load behaviors like those exhibited in Winter Storm Uri.

Weather Zone	ERCOT Model January 1 - 7, 2021	VAR Models January 1 - 7, 2021
Coast	324,735,334	69,826,423
East	7,519,166	8,380,808
Far West	9,795,965	1,933,635
North	11,571,249	355,672
North Central	382,483,068	198,002,134
South Central	39,659,348	51,999,299
Southern	6,503,418	9,983,529
West	55,786,132	6,530,437

Table 3: VAR Models: Average Squared Error for Load from 1/1/2021 - 1/7/2021

Weather Zone	ERCOT Model February 13 - 19, 2021	VAR Models February 13 - 19, 2021
Coast	6,499,966,149	1,715,801,979
East	88,332,434	59,821,650,292
Far West	2,000,900,363	59,449,577,798
North	86,290,196	74,864,877,566
North Central	8,789,099,199	26,397,498,496
South Central	1,183,362,088	4,154,665,293
Southern	575,913,544	37,990,488,481
West	226,078,908	69,673,388,481

Table 4: VAR Models: Average Squared Error for Load from 2/13/2021 - 2/19/2021

4.1.3 Neural Network

A Neural Network model is intended to improve upon both the ARUMA models and the VAR models. Not only are Neural Network models able to take weather and population into account, but these types of models are also advanced enough to find complex interactions between variables without the need of feature engineering. Because of Neural Networks' ability to find these interactions, all variables were used as inputs in the model, and the load at each of the eight weather locations was used as output.

The models' parameters were chosen using a grid search procedure. The parameters were chosen based on a combination of a train and test ASE, after the data was split using a 70/30 train/test split. (See Appendix Table 2 and Appendix Table 3 for final neural network parameters.) The grid search procedure that was run to tune the models takes a couple of days to run, but once the parameters were tuned, the models themselves only take a few minutes to fit.

The predictions made by the neural networks created in this study and ERCOT's own models were compared using data as of 12/31/2020 to predict 1/1/2021 - 1/7/2021 and data as of 2/12/2021 to predict 2/13/2021 - 2/19/2021. As stated previously, these two weeklong windows give an example of how the models perform during a normal week and a week that exhibit extreme fluctuations. During the week of 1/1/2021 - 1/7/2021, the neural networks performed better in the Coast, North and West regions (Table 5). During the week of 2/13/2021 - 2/19/2021, the neural networks performed better in the Coast, Far West, North, and West regions (Table 6). These results may not hold if comparing these models across all time periods.

No models performed well during the February storm. Considering this was a completely unprecedented event, models trained using past data had no way of predicting an event that had never before been seen.

Predictions for 1/1/2021 - 5/31/2021 were created for each model and are compared in Table 7. Results were mixed, and no model consistently scored better than the other models in every weather zone. The neural network models scored the best in the Coast and Southern regions and VAR scored the best in the East, Far West, North, North Central, South Central and West Regions. It seems like multivariable models capture medium-term load trends better than the univariate models, but more complex univariate models score better in certain zones, while less complex models score better in other zones. It is recommended that ERCOT use different types of models in each of the different zones in order to have the most accurate models.

Weather Zone	ERCOT Model January 1 - 7, 2021	Neural Network January 1 - 7, 2021
Coast	324,735,334	41,046,612
East	7,519,166	45,160,852
Far West	9,795,965	670,875,584
North	11,571,249	1,705,800
North Central	382,483,068	887,822,656
South Central	39,659,348	169,509,088

Southern	6,503,418	109,817,704
West	55,786,132	6,206,557

Table 5: Neural Network: Average Squared Error for Load from 1/1/2021 - 1/7/2021

Weather Zone	ERCOT Model February 13 - 19, 2021	Neural Network February 13 - 19, 2021
Coast	6,499,966,149	1,505,596,032
East	88,332,434	616,197,312
Far West	2,000,900,363	483,185,504
North	86,290,196	40,824,056
North Central	8,789,099,199	16,216,733,696
South Central	1,183,362,088	7,420,276,736
Southern	575,913,544	831,805,824
West	226,078,908	45,413,172

Table 6: Neural Network Average Squared Error for Load from 2/13/2021 - 2/19/2021

4.1.5 Medium-Term Forecasts

Predictions for 1/1/2021 - 5/31/2021 were created for each model and are compared in Table 7. Results were mixed, and no model consistently scored better than the other models in every weather zone. The neural network models scored the best in the Coast and Southern regions and VAR scored the best in the East, Far West, North, North Central, South Central and West Regions. It seems like multivariable models capture medium-term load trends better than the univariate models, but more complex models score better in certain zones, while less complex models score better in other zones. It is recommended that ERCOT use different types of models in each of the different zones in order to have the most accurate models.

Weather Zone	Neural Network	VAR	ARUMA
Coast	824,580,352	1,014,296,406	5,465,597,591
East	82,909,064	42,621,641	65,452,560
Far West	555,289,024	117,976,894	123,389,604
North	15,963,795	7,888,691	18,489,127
North Central	5,226,702,336	3,190,195,369	7,760,485,934
South Central	1,206,753,792	847,013,252	2,251,980,146
Southern	166,896,704	173,131,033	646,439,896
West	49,811,392	24,225,634	42,587,819

Table 7: Average Squared Error for Load from 1/1/2021 - 5/31/2021

4.1.5 Generation

Basic models were created to predict short- and medium-term total power generation in Texas. Since ERCOT controls generation, it correlates significantly with total load across the state. In general, not much power was wasted. It is recommended that ERCOT runs analyses of their own power generation by type and location in order to fully understand how future demand will be met. The internal analyses of power generation can help to determine where and what type of generation should be considered for future investment.

5 Discussion

5.1 ARUMA

The results of the ARUMA model can be explained by the outages during the February storm. Based on Figure 3, the demand for power was there but the demand was unable to be met due to the blackouts. Therefore, the ARUMA model likely underpredicted the load output and since the demand could not be met, the ARUMA model scored well. In other words, the load output was lower during this period due to issues with power sources.

Log transformations of the weather zone data improved the ARUMA model significantly. The effect of log transformations is that the impact of extreme values is reduced. Since the winter storm has such extreme values, the log transformations reduced the effect of extreme values.

The last component that aided the ARUMA model was the differencing of data and removing seasonality. These methods enable the data to meet the assumptions of stationarity. Despite these enhancements to the ARUMA model, the ARUMA model would not be useful in production use.

5.2 VAR

Inclusion of temperature data in the vectorized models marginally improved the ability of the predictions to adjust to temperature changes that diverged from past seasonal tendencies. The 365-day seasonal component dominates the behavior of the model. The minimum and maximum temperature components only have small effects on the predicted load when included. This is most notable in extreme cases where the temperature varies wildly in a short span of time or from its historical seasonal past. The winter months in general have much more sporadic day-to-day loads due to the highly varied nature of Texas winter temperatures. Cold days that exhibit spikes in load are often flanked by mild days that require very low energy use. Even with lagged values of temperature included, these winter spikes are very difficult to capture for the VAR model that is still largely based on lagged and seasonal load. The VAR model will usually not adequately capture the peak of load representing the needed capacity for these sporadic cold days.

The VAR model performs adequately during the summer, but this is likely due to more constant expected temperatures in Texas. However due to this more predictable behavior, VAR models do not provide a significant advantage over basic ARUMA models that are based solely on past load data. A VAR model that only includes the minimum temperature might be slightly more valuable for use during the coldest winter months where the minimum temperature will be inversely related to the maximum expected load and these cold temperatures are more isolated, but a VAR model for use during the summer months is not recommended.

5.3 Neural Network

Neural Networks can find complex interactions between variables in order to achieve higher accuracy when making predictions. Unfortunately, this higher accuracy comes at the price of model explainability. While the neural networks outperformed the ERCOT models in some locations and were outperformed by ERCOT models in other locations, it is difficult to pin down why.

The Coast region has consistent high loads on each day. A 5% margin of error when predicting the Coast region results in a much higher ASE than a 5% margin of error in a region with less load. While ERCOT’s model was not off by a significant percentage, the large load in this region causes the ASE to be extremely high when compared to a region like the Far West.

The neural network performed significantly worse than ERCOT’s models in the Far West region, despite the relatively low load in that region compared to the Coast region. The Far West has had a major increase in power usage since 2010, going from about 30,000 MWH to 95,000 MWH in 10 years. The neural network predicted around 60,000 MWH in 2021, which was around the average power usage in the Far West from 2010 – 2020, and this resulted in large errors. It seems that the neural network performs well when there are not major changes in load over time, like in the Coast region.

5.4 Model Considerations

The primary goal of this study is to build models for power load that perform independently of the current models being used by ERCOT and examine if and how these relate to power generation. This study does not have the same data complexity as ERCOT’s full data set, which includes far more variables and blends a variety of autoregressive and non-autoregressive models [5]. This study only uses weather, population, location, and time in its models and forecasts, all of which are autoregressive components of historical load data. Having fewer factors involved improves the explainability of the model. In a field where being correct is more important than explainable, having more factors for marginal improvements is of vital importance. Occasionally reverting to simplified models may illuminate trends and issues that are not readily apparent in models geared toward maximizing predictive accuracy.

The most glaring issue with any load models, whether it be those produced by ERCOT or in this analysis, is the unpredictability of weather. There is value in the long-term predictions made by this study in terms of climate, but weather is the biggest factor in predicting load. Extreme weather events will still have a major effect on ERCOT's ability to generate power and the amount of power consumed in Texas. Predicting extreme weather events is outside of the scope of this study, but the models in this study can still be used to extrapolate how much power will be used.

When examining loads during Winter Storm Uri, both ERCOT's models and the models in this study fail to capture the extreme peaks and valleys exhibited in actual measured load of the storm. The reason for this likely differs for the competing models. The ERCOT 7-day forecast appears to represent a realistic expectation for what the grid demand would be if those loads could be maintained for that period, as the forecast estimates several daily loads that are higher than any point in the last 10 years. The models in this study are largely autoregressive with some influence of weather variables, so they are hindered by a lack of any extreme events in the past to provide proper training.

The actual measured loads during the winter storm elucidate issues that could be addressed to the benefit of prediction models. One of those issues is that ERCOT's model predicted loads that were higher than the grid capacity given the situation. In addition to universally extreme increased power demand, record-breaking freezes caused many power sources to shut down completely. The steep drops in measured load represent the outages and curtailments that ensued as a result. There is little doubt that ERCOT operators understood the looming implications of their forecast predicting loads higher than capacity. However, they were likely limited in foreseeing what loads would actually be achieved during that time due to the relationship between extreme demands and unpredictable grid limitations. This highlights the need for data that simulates situations that manifest this behavior. Conditions that cause grid capacity to be reduced could be introduced as factors in predictive load models if the limitations of power sources are better understood. This could include metrics such as weather thresholds for power sources that indicate when they cease to function under duress. As these events are rare, synthetic data that simulates them could greatly aid in training models for prediction in the future.

5.5 Ethics

There are a few possible ethical concerns in this study. This study does not explicitly take economic factors into account. It is possible that there are latent economic factors in the data that is being used, meaning that certain economic factors could be secretly encoded somewhere in the data. Care may need to be taken to ensure that lower income areas are given the amount of power that is needed to avoid blackouts. There are areas with major power growth in the last 10 years as well. There was extreme growth in power usage in the Far West region, specifically. Care must be taken to ensure that these areas experiencing high levels of growth are as well powered as established areas with consistent levels of power usage.

As seen during the 2021 February storms, these models are being fit on data that affects human life. Poor predictions and a lack of preparedness can cause economic hardship, human suffering and even loss of human life. Extreme care must be taken to avoid these types of catastrophes. Because these models are balancing economic hardship and human suffering, it is better to overpredict and overproduce power rather than under generate power. While having the most accurate model possible is desired, care must be taken whenever the models underpredict load. Understanding what factors go into underpredicting could go a long way to preventing human suffering.

Although it is better to overpredict load, the potential environmental impact of using non-renewable energies to supplement this generation must also be considered. The long-term impact of non-renewable energies has been well documented. In long-term load and generation projections, climate change projections must also be considered as it has been determined that weather is a major factor in power usage. A few degrees of temperature increase could result in large consumption increases, especially in the summer months. A few degrees of increase could also result in larger grid load and potential loss of life due to an increase in chaotic weather patterns, which is a cause of major power events such as the February 2021 storm. The nature of climate change makes these extreme weather events extremely difficult to predict in terms of time, severity and type. ERCOT can potentially cut down on climate change by having more accurate load models that result in reduced waste. Again, overproducing is better than underproducing in the short-term, but overproduction with non-renewable power sources could result in major problems down the road.

The environmental impact of increasing generation is not only important to keep in mind in short-term power generation, but also long-term investment. ERCOT's 2021 Long-Term Demand and Energy Forecast Report [5] (see Figure 4 and Figure 5) predicts long-term increases in load in the state of Texas, and the state will be forced to keep up with this demand by increasing their generation. In order to reduce ERCOT's impact on the environmental issues discussed in the previous paragraph, it is highly recommended that ERCOT invest in green, renewable energy as much as possible.

The data used in this study was obtained at the weather zone level. While it would be nice to be able to drill down all the way to the household level, this could result in some ethical concerns. The most obvious of these ethical concerns is privacy. Being able to drill down to the household level to see how much power a specific household uses is a major privacy concern, especially with publicly available data. Being able to control power generation down to the household level could also be a concern. If drilling down to the household level, care must be taken to ensure that all households have sufficient power, and that this data is kept private and secure.

6 Conclusion

This study took a few factors deemed to be important to generation and load of Texas' power grid and created three distinct types of models using this data. The data used included population, weather, load, and generation. Weather and population were chosen to supplement load and generation as it was determined that these two factors affect load and generation more than any other factor. The models created using this data were ARUMA models, VAR models and LSTM models for load and generation, as these models are all time series models that provide various levels of complexity. Models were compared across three time periods: 1/1/2021 - 1/7/2021 as of 12/31/2020, 2/13/2021 - 2/19/2021 as of 2/12/2021 and 1/1/2021 - 5/31/2021 as of 12/31/2020. These periods were chosen as they provide short-term comparisons during a "normal" winter period, comparisons during a major anomaly, and medium-term comparisons.

Some of the short-term models shown in this paper did better than ERCOT's own models between 1/1/2021 - 1/7/2021 in the Coast, North and West regions. None of the models created in this analysis were able to accurately predict the load between 2/13/2021 - 2/19/2021 due to extreme previously unexperienced conditions including predicted loads exceeding the practical grid capacity during the storm. It is recommended that ERCOT investigate their Coast, Far West, North, North Central and West region 7-day load prediction models to determine if simpler, and thus less time-consuming, models could be similarly accurate. It is also recommended that ERCOT create synthetic data for training prediction models that simulate extreme events in order to gain a better understanding of how models will respond to them.

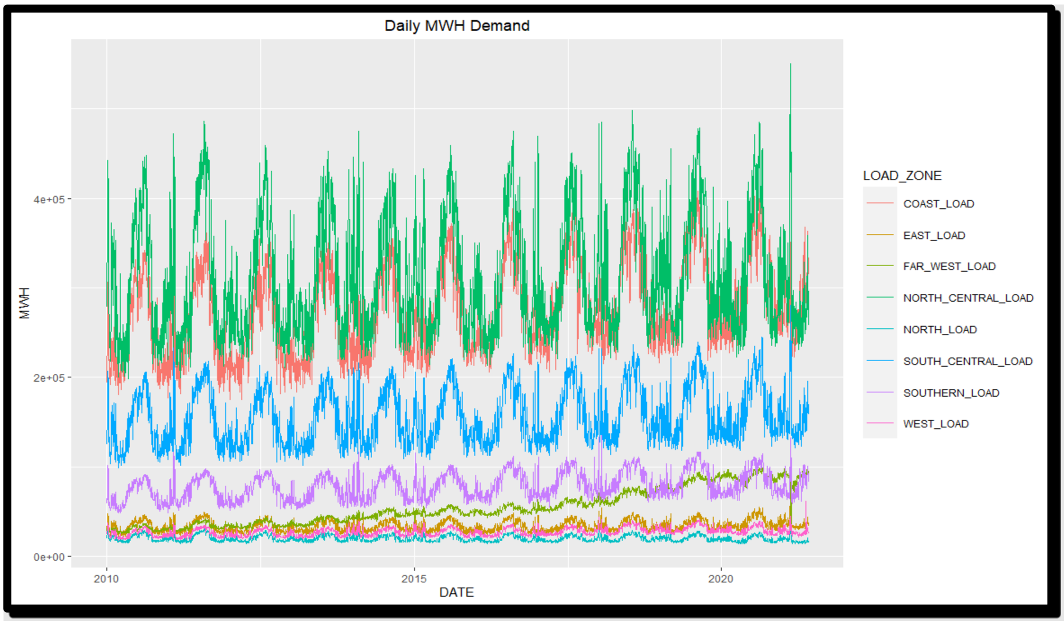
References

1. Electric Reliability Council of Texas, Inc. (ERCOT). (2021, May 6). *Record Breaking Demand Expected This Summer*. http://www.ercot.com/news/releases/show/230649
2. Texas Health and Human Services (2021). *Winter Storm-Related Deaths*. [Data Set]. Texas Department of State Health Services. https://dshs.texas.gov/news/updates.shtm#wn
3. Stelloh, T., Suliman, A., Chirbas, K., Sheeley, C. (2021, Feb 15). *Millions in Texas without Power as Deatly Storm Brings Snow, Freezing Weather*. NBC News. https://www.nbcnews.com/news/weather/knocked-out-texas-millions-face-record-lows-without-power-new-n1257964
4. Penney, V. (2021, Feb 19). How Texas' Power Generation Failed During the Storm, in Charts. *The New York Times*. https://www.nytimes.com/interactive/2021/02/19/climate/texas-storm-power-generation-charts.html
5. Electric Reliability Council of Texas, Inc. (ERCOT). (2021, January 8). *Long-Peak Demand and Energy Forecast*. http://www.ercot.com/content/wcm/lists/219761/2021_LTLF_Report.pdf
6. Opheim, C. (2019, February 12). Item 10.1: *Load Forecasting Overview* [Presentation]. ERCOT Board of Directors Meeting. http://www.ercot.com/content/wcm/key_documents_lists/161457/10.1_Load_Forecasting_Overview.pdf
7. Opheim, C., Quan, M. (2014, January 27). *Review of Preliminary Load Forecast* [Presentation]. http://www.ercot.com/content/meetings/other/keydocs/2014/0127-CDRWorkshop/2013%20ERCOT%20Load%20Forecast%20Review%20with%20Itron%20012714.pdf
8. Electric Reliability Council of Texas, Inc. (ERCOT). (2012, July 13). *2011 Long-Term Load Forecast Model Review* [Presentation].

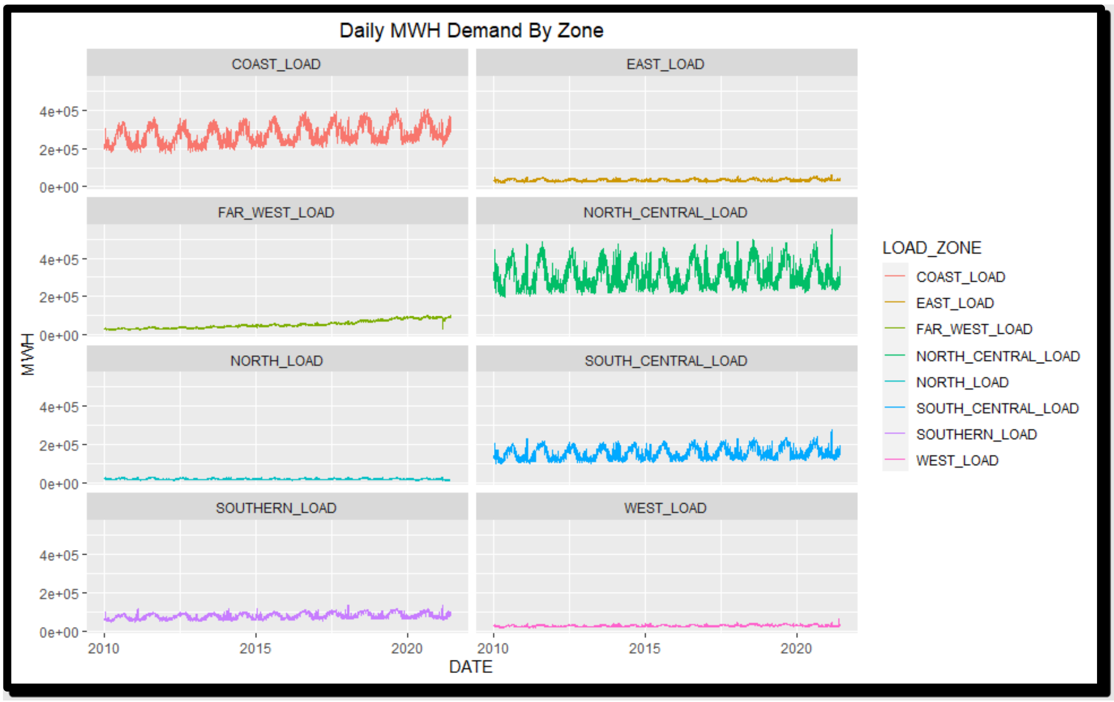
http://www.ercot.com/content/meetings/rpg/keydocs/2012/0713/2011_Long-Term_Load_Forecast_Model_Description_for_RPG_-_Jul.pdf
9. Electric Reliability Council of Texas, Inc. (ERCOT). (2021). *Grid Information*. http://www.ercot.com/gridinfo
10. Hwang, M. Y., Jin, C. H., Lee, Y. K., Kim, K. D., Shin, J. H., & Ryu, K. H. (2011). Prediction of wind power generation and power ramp rate with time series analysis. *2011 3rd International Conference on Awareness Science and Technology (ICAST)*. https://doi.org/10.1109/icawst.2011.6163182
11. Boroojeni, K. G., Amini, M. H., Bahrami, S., Iyengar, S. S., Sarwat, A. I., & Karabasoglu, O. (2016, September 21). *A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon*. Electric Power Systems Research. https://www.sciencedirect.com/science/article/pii/S0378779616303352.
12. del Real, A. J., Dorado, F., & Durán, J. (2020, May 3). *Energy Demand Forecasting Using Deep Learning: Applications for the French Grid*. MDPI. https://www.mdpi.com/1996-1073/13/9/2242/htm.
13. Wang, Z. G., Zhou, Q. J., & Zhou, X. H. (2013, September 3). *A Combined Power Demand Forecasting Model with Variable Weight*. Applied Mechanics and Materials. https://www.scientific.net/AMM.392.618.
14. Craig, M. T., Jaramillo, P., Hodge, B.-M., Nijssen, B., & Brancucci, C. (2020). Compounding climate change impacts during high stress periods for a high wind and solar power system in Texas. *Environmental Research Letters*, *15*(2), 024002. https://doi.org/10.1088/1748-9326/ab6615
15. Xu, K., Hou, R., Ding, X., Tao, Y., & Xu, Z. (2019). Short-term Time Series Data Prediction of Power Consumption Based on Deep Neural Network. *IOP Conference Series: Materials Science and Engineering*, *646*, 012027. https://doi.org/10.1088/1757-899x/646/1/012027
16. García-Ascanio, C., & Maté, C. (2009, November 22). *Electric power demand forecasting using interval time series: A comparison between VAR and iMLP*. Energy Policy. https://www.sciencedirect.com/science/article/pii/S0301421509007344.
17. Acakpovi, A., Ternor, A. T., Asabere, N. Y., Adjei, P., & Iddrisu, A.-S. (2020). Time Series Prediction of Electricity Demand Using Adaptive Neuro-Fuzzy Inference Systems. *Mathematical Problems in Engineering*, *2020*, 1–14.
18. Kang, T., Lim, D., Tayara, H., & Chong, K. (2020). *Forecasting of Power Demands Using Deep Learning*. Applied Sciences, 10(20), 7241–. https://doi.org/10.3390/app10207241 https://doi.org/10.1155/2020/4181045
19. Ciechulski, T., & Osowski, S. (2020, November 23). *Deep Learning Approach to Power Demand Forecasting in Polish Power System*. MDPI. https://www.mdpi.com/1996-1073/13/22/6154.
20. National Oceanic and Atmospheric Administration (NOAA) National Centers For Environmental Information. (2021). *Climate Data Online*. https://www.ncdc.noaa.gov/cdo-web/
21. Texas Demographics Center. (2021). *Texas Population Estimates Program.* (2021). https://demographics.texas.gov/Data/TPEPP/Estimates
22. Zhao, Z.-yu, Yan, H., Zuo, J., Tian, Y.-xi, & Zillante, G. (2012, December 20). *A critical review of factors affecting the wind power generation industry in China*. Renewable and Sustainable Energy Reviews. https://www.sciencedirect.com/science/article/pii/S1364032112006806.
23. Itron, Inc. Forecasting and Load Research Solutions (2014, January 10). *Review of ERCOT Premise Forecast.* http://www.ercot.com/content/gridinfo/load/forecast/Docs/PremiseFindings_January2014.pdf

24. Electric Reliability Council of Texas, Inc. (ERCOT). Market Information System. *Forecast Variability.* (2021). https://mis.ercot.com/public/static-assets/weather-files/forecast_variability.pdf

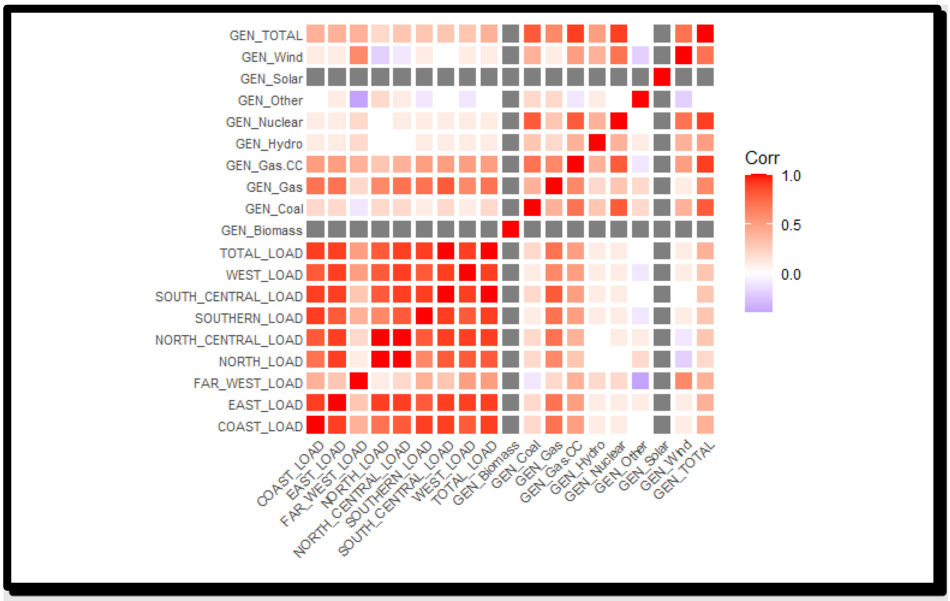
Appendix:



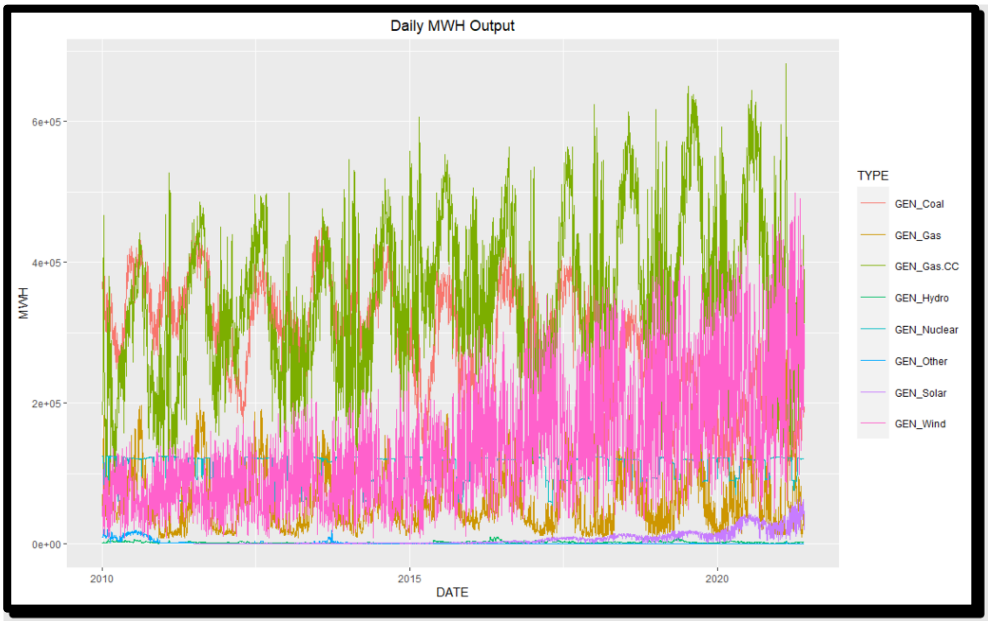
Appendix Figure 1



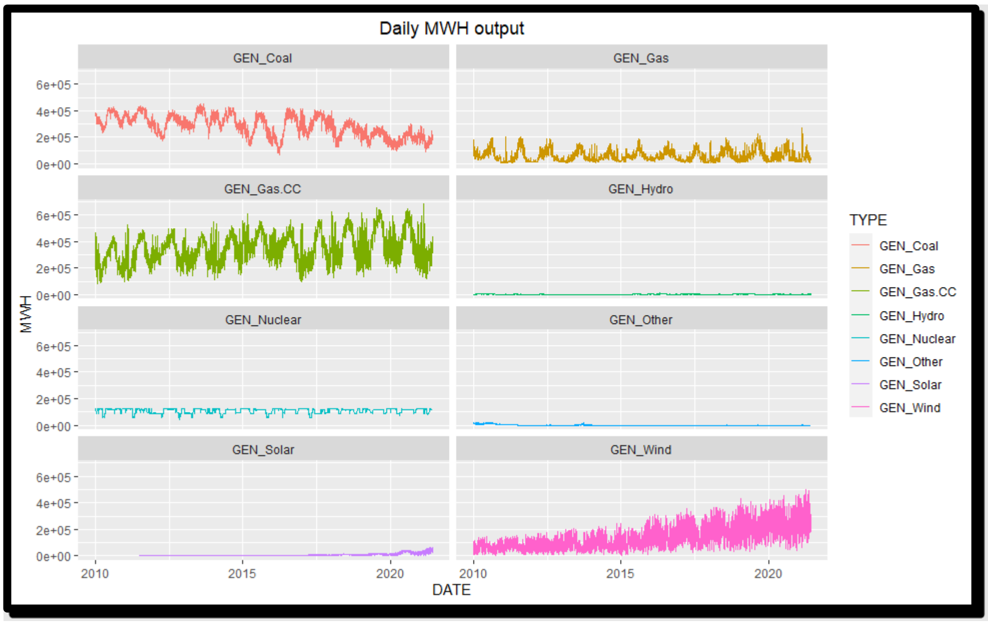
Appendix Figure 2



Appendix Figure 3



Appendix Figure 4



Appendix Figure 5

Target	Best Value for p based on AIC		
	January 1 - 7, 2021	February 13 - 19, 2021	January 1 - May 31, 2021
Coast	8	6	4
East	7	6	18
Far West	5	3	9
North	5	3	18
North Central	7	3	3
South Central	6	3	3
Southern	6	4	28
West	5	3	20

Appendix Table 1: VAR model Parameters for Forecast Periods

Target	Days forward	Days back	Year Back	Nodes 1	Nodes 2	Dense Layer
Coast	7	14	False	64	32	True
East	7	7	True	256	128	False
Far West	7	7	False	128	64	False
North Central	7	7	True	128	64	False
North	7	7	True	128	64	False
Southern	7	14	False	256	128	True
South Central	7	7	True	64	32	True
West	7	7	False	128	64	False

Appendix Table 2: Neural Network Short-Term Parameters

Target	Days forward	Days Back	Year Back	Nodes 1	Nodes 2	Dense Layer
Coast	151	7	True	64	32	True
East	151	7	True	256	128	False
Far West	151	7	False	128	64	True
North Central	151	14	True	256	128	False
North	151	7	True	256	128	True
Southern	151	7	True	64	32	True
South Central	151	7	True	128	64	False
West	151	14	True	128	64	True

Appendix Table 3: Neural Network Long-Term Parameters