This report is part one of the project forecasting electricity demand in zone **WFEC,** operated by the Southwest Power Pool (SPP). The report contains four sections: a brief introduction of WFEC, exploratory data analysis, evaluation of naïve methods, and description of possible explanatory variables.

1. **Introduction**

Western Farmers Electric Cooperative (WFEC) is a not-for-profit generation and transmission (G&T) cooperative headquartered in Anadarko, Oklahoma. It provides wholesale electric service to its members. The cooperative is Oklahoma’s largest locally owned power supply system.

* 1. **Brief History of WFEC**

WFEC was organized in 1941 with purpose of securing an adequate power supply at affordable rates for the farmers and rural industrial developers. It has experienced substantial growth since its launch. By 2001, WFEC owned and operated four power plants fueled by gas steam and coal. Entering the 21st century, while continuing in building new power plants, WFEC sought for opportunities to diversify its fuel resources into renewable energy. It began with signing a purchase power agreement (PPA) with a wind farm developer in January 2003. WFEC entered into a PPA with a solar power facility in early 2015. Later in September 2015, it decided to invest in multiple solar generating projects. Currently, WFEC’s power generation fleet consists of six steam and gas turbine power generation sites, five utility-scaled solar farms, and 13 community solar farms. The total combined power capacity for owned and contracted facilities is approximately 2,500 megawatts (MW).

* 1. **Member Profile**

WFEC’s member-owners consist of 21 distribution cooperatives, 17 are located in Oklahoma and four in New Mexico. Its membership is diverse with only two members accounting for over 10% of its sales and no member accounting for more than 12%. WFEC also serves Altus Air Force Base, which represents less than 1% of sales. WFEC supplies the electrical needs of more than two-thirds of the geographical region of Oklahoma, part of New Mexico, and small portions of Texas and Kansas. The members of WFEC serve estimated populations of 494,000 in Oklahoma and 52,000 in New Mexico. Population growth throughout the largely rural service area is minimal. WFEC’s members serve residential, industrial and commercial sectors. The revenue source appears to concentrate in commercial or industrial sectors, since some members have residential revenues as a low 9% to 10% of total revenues.

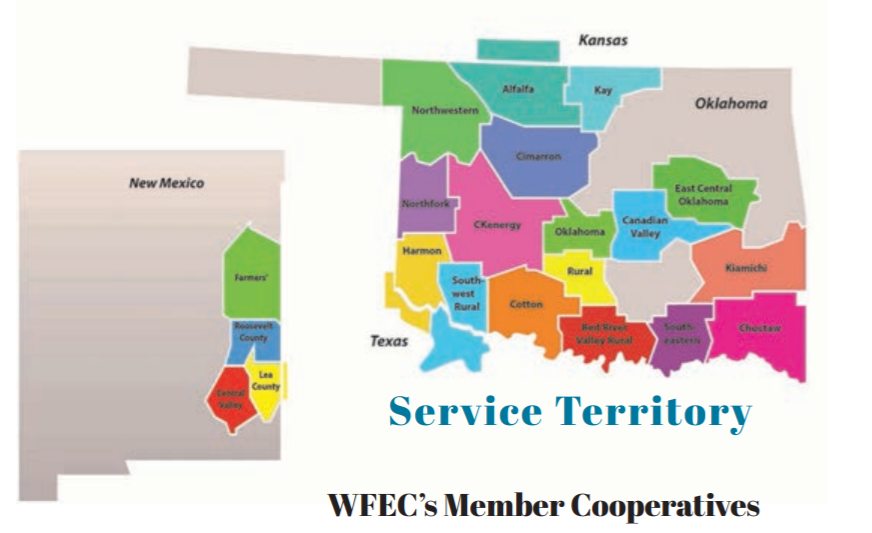
* 1. **WFEC Energy Demand**

The two major industries in rural Oklahoma are mining and agriculture. Oil and gas sector accounts for approximately 11% of Oklahoma’s gross domestic product (GDP). The state has a total of 34.4 million acres of land in farms. The energy demand growth in rural Oklahoma has typically been flat, although stronger oil and gas prices have periodically driven growth in some service territories. Future demand growth is expected to be low and varies with activities in the oil and gas sector.

* 1. **WFEC Energy Supply**

The energy mix of WFEC is made up of power purchase agreement, hydropower, wind energy, and solar. It maintains a diversified portfolio of generation resources that includes owned facilities and capacity. Several technologies and fuel types are part of that mix and ensure that WFEC can flexibly adjust the demand to its supply. As we can also see on WFEC website: “The generation and transmission (G&T) cooperative's energy resources include a coal-based power plant, as well as numerous natural gas units. Wind resources, solar energy and hydropower allocations also make up a part of the overall energy mix, as well.”

* 1. **Geographical and Climate Context**

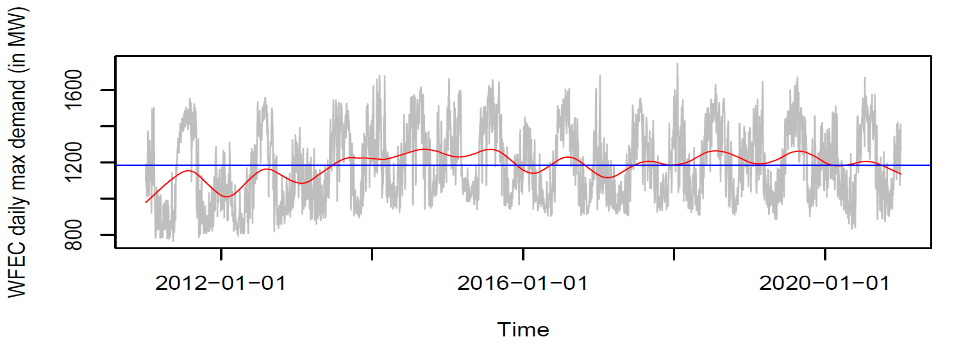
The territory covered by the Western Farmers Electric cooperative comprises the near entirety of Oklahoma State (OK), a small Kansas (KS) district to the north of Cherokee, small Texas (TX) district to the south of Tipton, and three non-negligible New-Mexico (NM) around Clovis, NM, and Portales, NM. The geography of this territory is dominated by the ecological features characterizing Oklahoma and its great diversity of landscape and climate. It is situated in the Great Plains and borders four primary mountain ranges that are the Arbuckle Mountains, the Wichita Mountains, the Ozark Mountains and the Ouachita Mountains. Given its mountain ranges to the West and the great plains to the East, the WFEC territory slopes downward from its western to its eastern boundaries. WFEC presents semi-arid high plains in the northwestern corner. The southwestern part is dotted of partial plains and small mountain ranges like Antelope Hills and the Wichita Mountains. The central portion of WFEC territory is made up of prairies and woodlands. Given its geography between the Great Plains stretching from Canada to the Gulf of Mexico and the mountain ranges bordering the WFEC territory, WFEC’s climate is representative of a diverse climate. It sits at the crossroads of three different air masses: humid air from the Gulf of Mexico, warm and dry hair from Mexico and the Southwestern U.S, and cold dry air from Canada. WFEC’s region can see frequent air mass changes, that produce dramatic change in both temperature and humidity, during the same season and sometimes, even, the same day. The territory is subjected to extreme meteorological conditions. WFEC areas are also characterized by frequent interaction between cold and warm air masses, producing tornadoes between April and June. Due to its location as an intersection between zones of differing temperature and winds, the weather patterns in the territory can change significantly between relatively short distances. Precipitation levels are lowest in winter months and then rise dramatically from March to a peak in May (owing to frequent thunderstorm activity) before decreasing by mid-summer when long periods of dry weather occur in July and August. Precipitation then reach a second yearly maximum from early to mid-fall before decreasing from October to December.

1. **Exploratory data Analysis**
   1. **Missing Data**

There are 7 missing data, which occurred at GMT 2017-09-28 00:00:00 and from GMT 2018-12-12 08:00:00 to 13:00:00. We searched online for the particular time intervals and had found no information that could explain why the data are missing. We observed that in the “sppdata” dataset, 10 out of 11 columns have missing data for the same period of time. That is, 10 out 11 services zones have no data at the same time. If there is a blackout of such a great scale, we must have found some information about the outage. The only information we found that might be related is that there was a winter storm with a mix of heavy rain, snows and ice during December 7 to 10, 2018, right before the date of the missing data. The storm damaged severely the Southeastern states. Oklahoma was on its path. There is a possibility that the missing data in the region could be attributed to the severe weather conditions. In any case, SPP mechanism for collecting or recording the electricity demand data seemed to be malfunctioned during the time periods and did not record the data successfully. Since the proportion of missing data is small (7 out of 87,669), we imputed them using the average of the three previous years for each hour. That is, use the average of the electricity demand from 8AM to 1PM on Dec. 12 in 2015, 2016 and 2017 to impute the missing demand from 8AM to 1PM on Dec. 12 in 2018. Same method applied on the missing data of Sep 28, 2017.

* 1. **Trends**

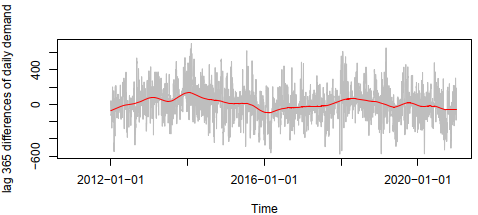
Below is a plot of the WFEC daily max demand over time:



There was an upward trend from 2011 to 2014. The trend then decreased till 2016, stayed flat between 2017 to 2019, and decreased again in 2020.

* 1. **Seasonality**

The daily max demand from 2011 to 2020 shows consistent seasonality. We used differencing to reduce the seasonal effect. Below is the difference plot with a lag of 365. It seems natural to choose the lag to be 365 days since the demand of electricity is likely to be at similar level on the same day of the year. The plot shows more clear trends than the ones in the last plot.

****

1. **Evaluation of naïve methods**
   1. **Data split**

It is observed that the max daily electricity experiences similar trend between the time period 2011-2016 and period 2017 onwards. We decided to split the data in the fashion that the model is trained on the entire first cycle, and validated and tested on the second cycle. We used equal length of time for the validation and the testing set. The split of data is as follows:

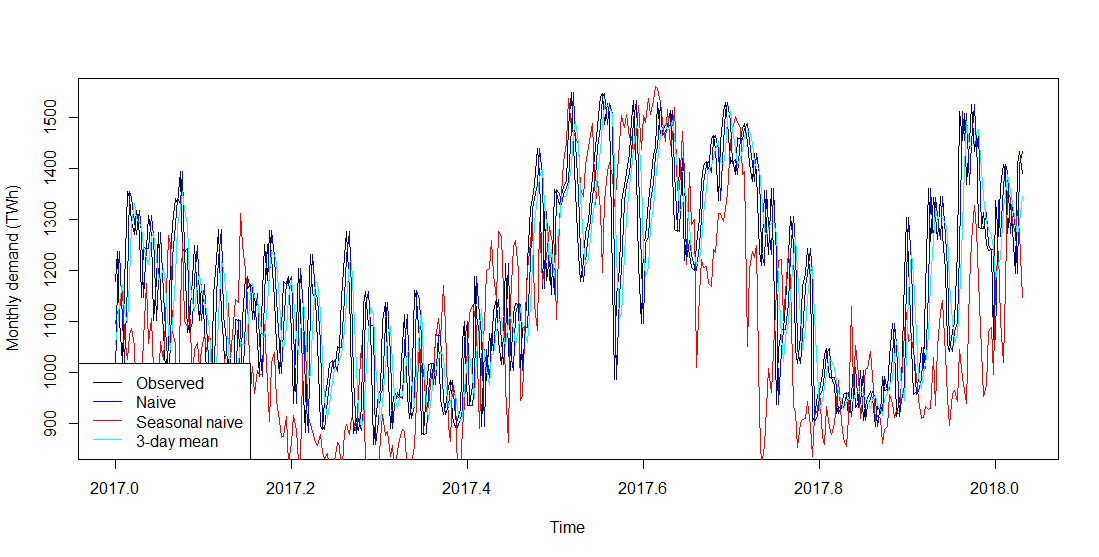
* Training data set starts from January 1, 2011 and ends December 31, 2016.
* Validation data set starts from January 1, 2017 and ends December 31, 2018.
* Testing data set starts from January 1, 2019 and ends December 31, 2020.
  1. **Naïve forecasting model**

We used three models. All three models are predicting on a daily basis.

1. Naïve model forecasts based on one day before.
2. Seasonal naïve model predicts the value based on the value of same date one year before.
3. Moving average predicts based on the latest 3-day average.
   1. **Performance**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Data Set | Bias | Pbias | Mape |
| Naïve | -0.27 | 0.32 | 6.22 |
| Seasonal naïve | -38.90 | -2.74 | 11.9 |
| Moving average | -0.25 | 0.08 | 3.34 |

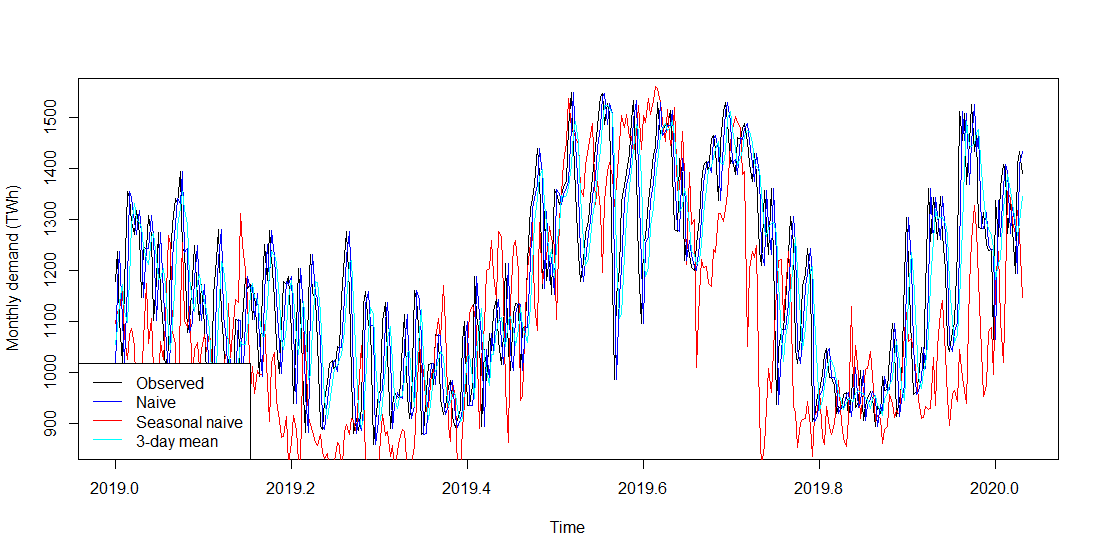
*Table 1 Performance metrics of validation data set*



*Graph 1 Performance metrics of validation data set*

|  |  |  |  |
| --- | --- | --- | --- |
| Test Data Set | Bias | Pbias | Mape |
| Naïve | -0.006 | 0.32 | 5.99 |
| Seasonal naïve | -18.60 | -0.61 | 12 |
| Moving average | -0.02 | 0.09 | 3.23 |

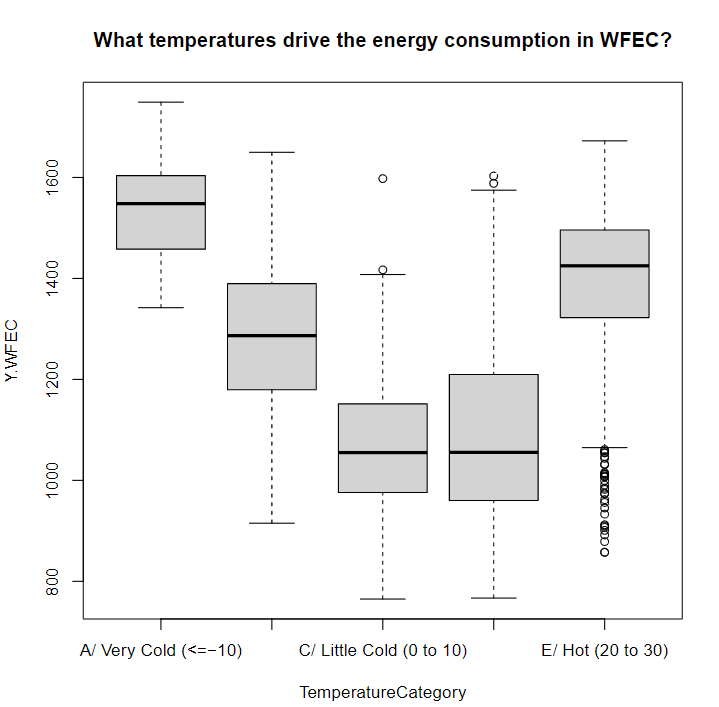
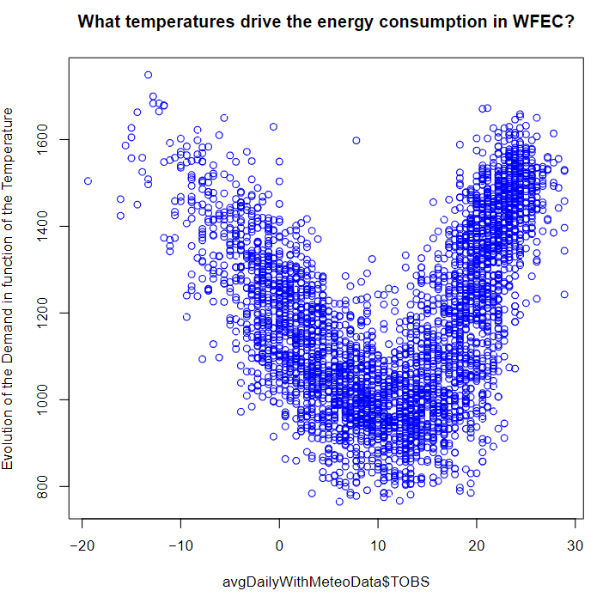
*Table 2 Performance metrics of test data set*



*Graph 2 Performance metrics of test data set*

The 3-day moving average performs the best in both data set across all metrics. Essentially, data from a tighter time frame share similar value and momentum.

1. **Data Exploration of Potential Explanatory variables with Y.WFEC**

For the data exploration, we decided to extract daily meteorological summaries (wind, temperature, snowfall, rainfall) from the NOAA website at <https://www.ncdc.noaa.gov/cdo-web/search> for the 2011 through 2021 period for specific weather stations situated in the WFEC territory, that is stations located in different parts of Oklahoma, Kansas, New Mexico, and Texas. We also extracted the daily OK WTI Cushing Oil Price per Barrel for our period on the EIA site at <https://www.eia.gov/dnav/pet/hist/RWTCD.htm>. We extracted the daily summaries because meteorological data are the first driver of energy demand, especially for WFEC where industrial and agricultural clients dominate. Daily Oil Price per Barrel (WTI Cushing) is used because it could serve as a proxy variable for the economic health of our region, which is usually correlated with the energy demand variable, Y.WFEC. Further, given the oil and gas industry in the rural belt of Oklahoma (main parts of WFEC) represents a significant portion of the economic activity and thus a significant part of the energy demand, we also used the daily oil price per barrel.

**Figure 1**

**Figure 2**

Choosing the right weather station or a combination of the weighted average of the weather stations meteorological summaries is then the next challenge, given the large size of the WFEC region and the diversity of its climate. After considering several stations that span the different locations of WFEC, we explored whether the stations have drastic differences in their temperature distribution. We noticed the distributions of observed temperature are relatively like the Blanchard’s station. The average and median are relatively similar. However, only the Blanchard station has close to complete data with only 2 missing rows. Given its central position in WFEC and a relative temperature distribution similarity to other stations, the Blanchard Weather Stations is chosen only. Once the proper manipulation is completed, we then merge the daily meteorological summaries to daily oil WTI oil prices. First, we want to see initial relationship between energy demand and observed temperature TOBS. After binning temperatures in intervals of 10 degrees from -20 degree Celsius to 30+ degree Celsius, we see, as showed in the box plot (Figure 1), that the energy demand is the lowest in the 0 to 10 and 10 to 20 degree Celsius distribution and the highest when the temperature is below 0 degree Celsius or over 20 degrees. This is potentially the sign of a heating and cooling factor respectively amongst WFEC clients. That trend is confirmed with the effective temperature. This is also confirmed by the scatter plot (Figure 2).

We also observe the following distribution for the observed temperature, rain, and oil prices over the 2011 – 2021 period.

Observed Temperature

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min | 1st Quartile | Median | Mean | 3rd Quartile | Max | NA’s |
| -19.40 | 3.30 | 12.20 | 11.34 | 20.00 | 28.90 | 17 |

Rainfall

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min | 1st Quartile | Median | Mean | 3rd Quartile | Max | NA’s |
| 0 | 0 | 0 | 2.547 | 0 | 158.2 | 15 |

Oil price (1141 missing data because no data on Saturday and Sundays)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min | 1st Quartile | Median | Mean | 3rd Quartile | Max | NA’s |
| -36.98 | 48.49 | 60.88 | 68.45 | 93.49 | 113.39 | 1141 |

We then searched for the most appropriate temperature of reference for HDD and CDD. After looking for the temperature, on the cold and warm ends of the temperature-energy demand relationship, that would make energy demand significantly increase, we concluded that the respective Tref for HDD and CDD are 5 and 20 degree Celsius. The following charts (Figure 3 and Figure 4) show us that there is a cooling and heating effect on the energy demand with a clear upward relationship for both.

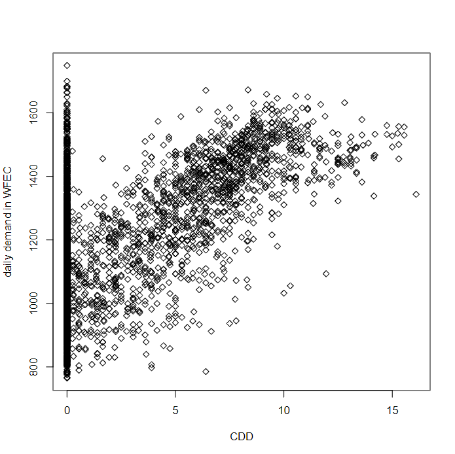
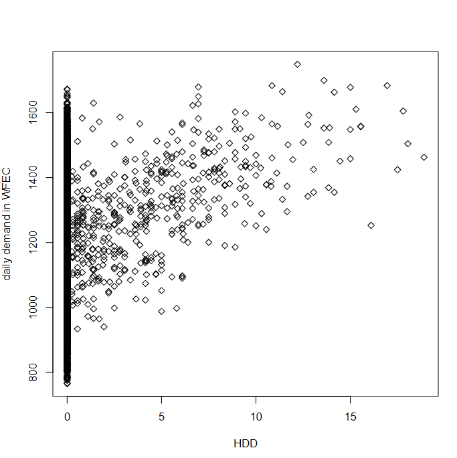


Figure 3

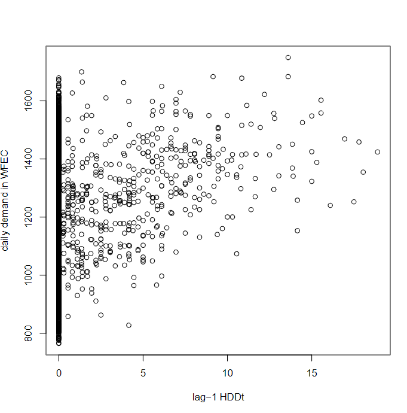
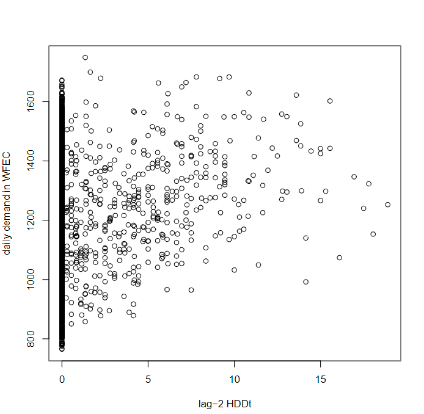
Figure 4

The lag -1 and lag-2 HDD (figure 5 and 6) and CDD charts (figure 7 and 8) show us there is also a cooling and heating effect with a lag of 1 and 2 days prior the effective date!

HDD lag effect

Figure 6

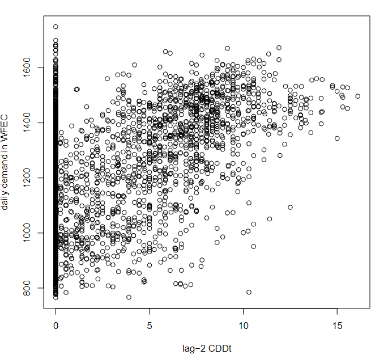
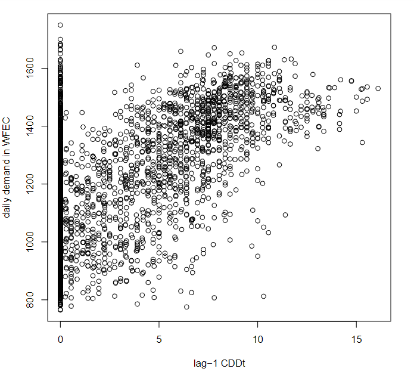
Figure 5



CDD lag Effect

Figure 7

Figure 8



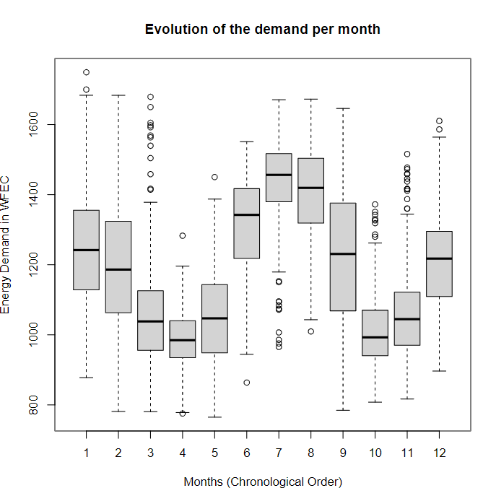
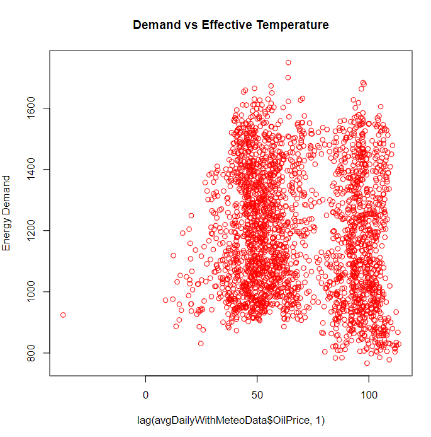
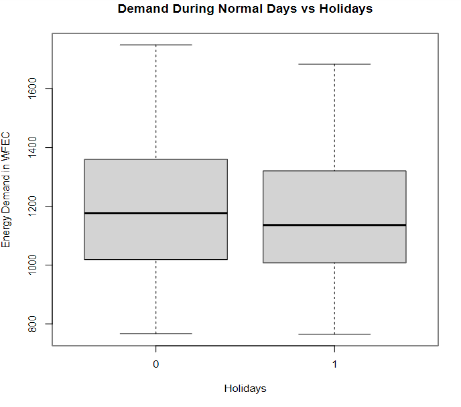
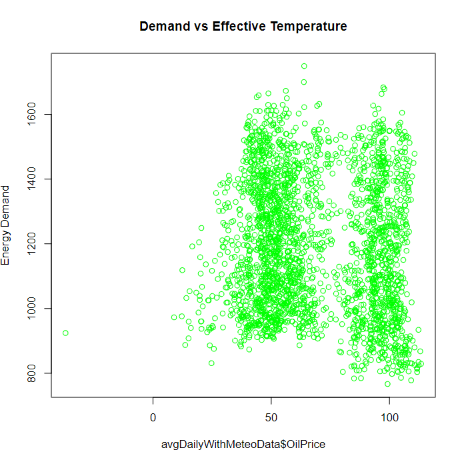
Note there is little effect of precipitation levels on the demand. The effective temperature level (0.5 \*Observed Temperature + 0.5 \* Temperature observed the day before) and the weighted average temperature ((Min +Max)/2) follow the same shape as the observed temperature. Further, we observed that the energy demand does not change significantly from one weekday to another, for instance between Monday and Saturday. Thus, there is no weekend/weekday effects in WFEC, which makes sense given the agricultural and industrial base of the clients. As showed in the box plot (Figure 9), it seems the highest demands occur in January, February, June July, and August, that is when the temperature conditions are the most extreme. Additionally, there is no significant effect of holidays on energy demand with only a slight decrease in demand as showed in the box plot (Figure 10**).** Finally, there is no apparent effect (Figure 11) or lag effect (Figure 12) of the daily WTI on energy demand. The explanation is that energy demand of oil and gas fields is not dependent on the current daily WTI prices but rather a more complex combination of supply and demand factors. An exploration of the Future contracts might be useful.

Figure 11

Figure 12

Figure 10

Figure 9