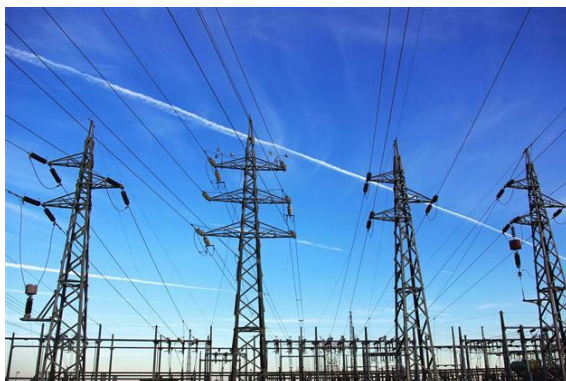


Homework 3



Electricity Usage

Name: Young, Brent

Predict 413 Section #: 55

Quarter: Winter 2018

Introduction

For homework 3, I chose “Monthly Average Residential Electricity Usage for Iowa City 1971 – 1979”. I chose this time series because I recently bought a house and I find it interesting how electricity usage rises and decreases depending on the seasons throughout the year. I am also always trying to decrease electricity usage using energy efficient bulbs or smart thermostat (Ecobee4) so understanding the ebbs/flow of energy usage can help me save money from a personal standpoint.

Exploratory Data Analysis

Structure and Size of Full data

The structure of the full dataset has 106 rows and 2 variables: The first variable, month is a factor variable and represents the date, while usage is an integer variable and represents the average residential electricity usage in Iowa City from 1971 – 1979.

Figure 1: Descriptive Statistics of Full Data

Min	1st Quartile	Median	Mean	3rd Quartile	Max
368.0	415.8	465.0	489.7	534.0	763.0

Observations: Figure 1 shows that the mean number of electricity usage is 489.7, the median electricity usage is 465.0, minimum is 368.0 and max electricity usage is 763.0.

Time Series Analysis

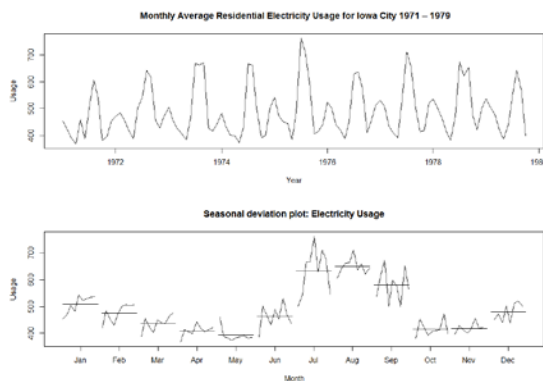


Figure 2

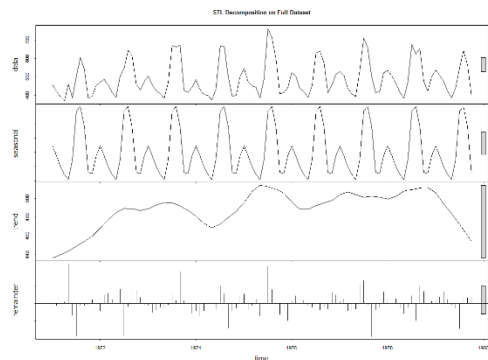


Figure 3

Observations: Prior to building my models, I created a timeplot and seasonal subseries plot (figure 2) to investigate seasonality and trend. The plots revealed no trend and seasonality (multiplicative), indicating that the variation around the trend-cycle, appears to be proportional to the level of time series. Additionally, the seasonal subseries plot shows that electricity usage

is highest in the summer months from July through September (most likely from air conditioning usage). Lastly, figure 3 shows a STL decomposition of the data. The results confirmed the seasonality and no trend that we saw from the previous plots.

Description of Training and Test Datasets (80/20 Split)

Prior to building my models, I split the full data into a training (80%) and test set (20%) using an 80/20 approach. For instance, my training dataset is from January 1971 to December 1977, while my test set is from January 1978 to October 1979.

Types of Models & Formulation of Models

Given that the EDA revealed that there is strong seasonality that corresponds to a monthly cycle, I decided to use ETS, auto.arima, neural network, and Holt-Winters method with damped trend and multiplicative seasonality to build my models. I will also use seasonal naïve as a benchmark for my models. I will then compare the performance of my models using error statistics on the train and test datasets.

ETS: MNM Model

I decided to use ETS because it has the ability to handle seasonality, estimate the necessary parameters, and then select the appropriate model objectively based on AIC. Using the ETS model selection function, ETS produced a M,N,M model, which indicates multiplicative errors, no trend, and multiplicative seasonality. The multiplicative seasonality makes sense given that the variation around the trend-cycle, appears to be proportional to the level of time series.

Auto.arima: ARIMA(1,0,0)(0,1,1)[12] with drift

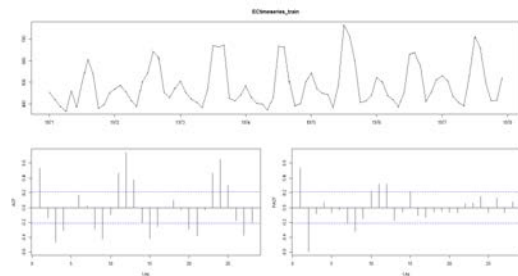


Figure 4

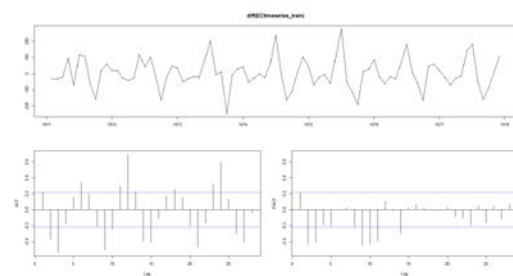


Figure 5

Observations: Prior to building the auto.arima model, I decided to use `tsdisplay` to see if the data needed differencing. Figure 4 shows the data pre-differencing and figure 5 shows the data post differencing. I also used `nsdiffs` which shows the number of differencing required for seasonal data and it showed 1.

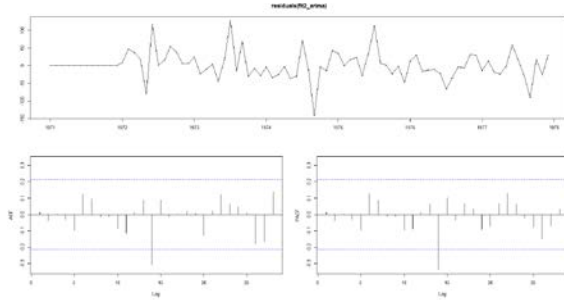


Figure 6

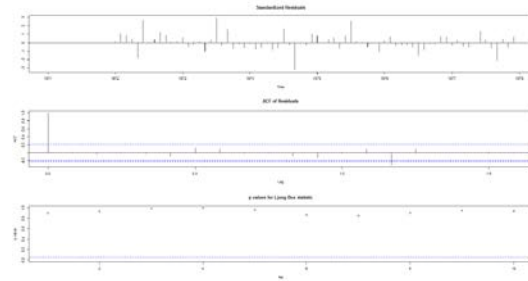


Figure 7

Observations: Similar to ETS, I decided to use `auto.arima` because it has the ability to handle seasonality, estimate the necessary parameters, and then select the appropriate model objectively based on AIC. Figures 6 and 7 shows a `tsdisplay` of the residuals, ACF of Residuals, and p-values for Ljung-Box test. The ACF of the residuals and the Ljung-Box test (p-value = 0.7566) confirmed that the model contained white noise. As a result, the $ARIMA(1,0,0)(0,1,1)[12]$ with drift, which means 1 order of autoregressive, 0 degrees of first differencing, and 0 order of the moving average part for non-seasonal, and 0 order of autoregressive, 1 degrees of first differencing, and 1 order of the moving average part for seasonal is valid.

Neural Network Model

I chose to use neural network because it has the ability to forecast time series data with seasonality. Using the `nnetar` model selection function, `nnetar` produced a $NNAR(1,1,2)$ model, which indicates 1 lagged input and 2 nodes of hidden layer.

Holt-Winters method with damped trend and multiplicative seasonality

Furthermore, I chose Holt-Winters method with damped trend and multiplicative seasonality because Hyndman mentions that it is often the single most accurate forecasting method for seasonal data.

Diagnostics

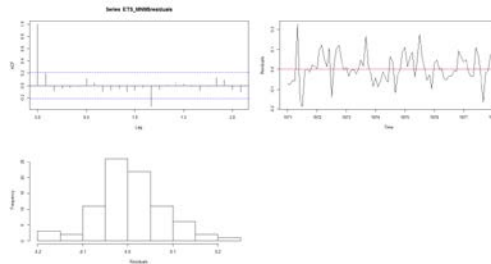


Figure 8: ETS

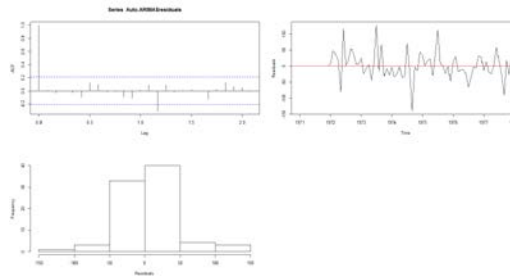


Figure 9: Auto.Arima

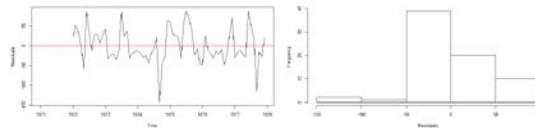


Figure 9: Neural Network

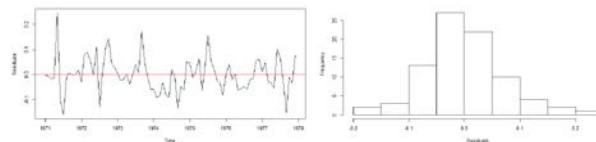


Figure 10: Holt-Winters method

Observations: In order to determine if the models above can be improved upon (e.g., no correlations between forecast errors for successive predictions), I conducted a correlogram of the in-sample forecast errors and conducted a Ljung-Box test for the ETS and auto.arima models. The Ljung-Box test and ACF plot shows that there is little evidence of non-zero autocorrelations in the in-sample forecast errors. Additionally, the time plot of forecast errors for all four models show that the forecast errors have constant variance over time. The histogram of forecast errors of all four models also shows the forecast errors are normally distributed with mean zero. Therefore, we can conclude that these models provide an adequate predictive model for the monthly average residential electricity usage for Iowa City and that the assumptions that the 80% and 95% intervals were based upon are probably valid.

Forecasts and Performance/Accuracy

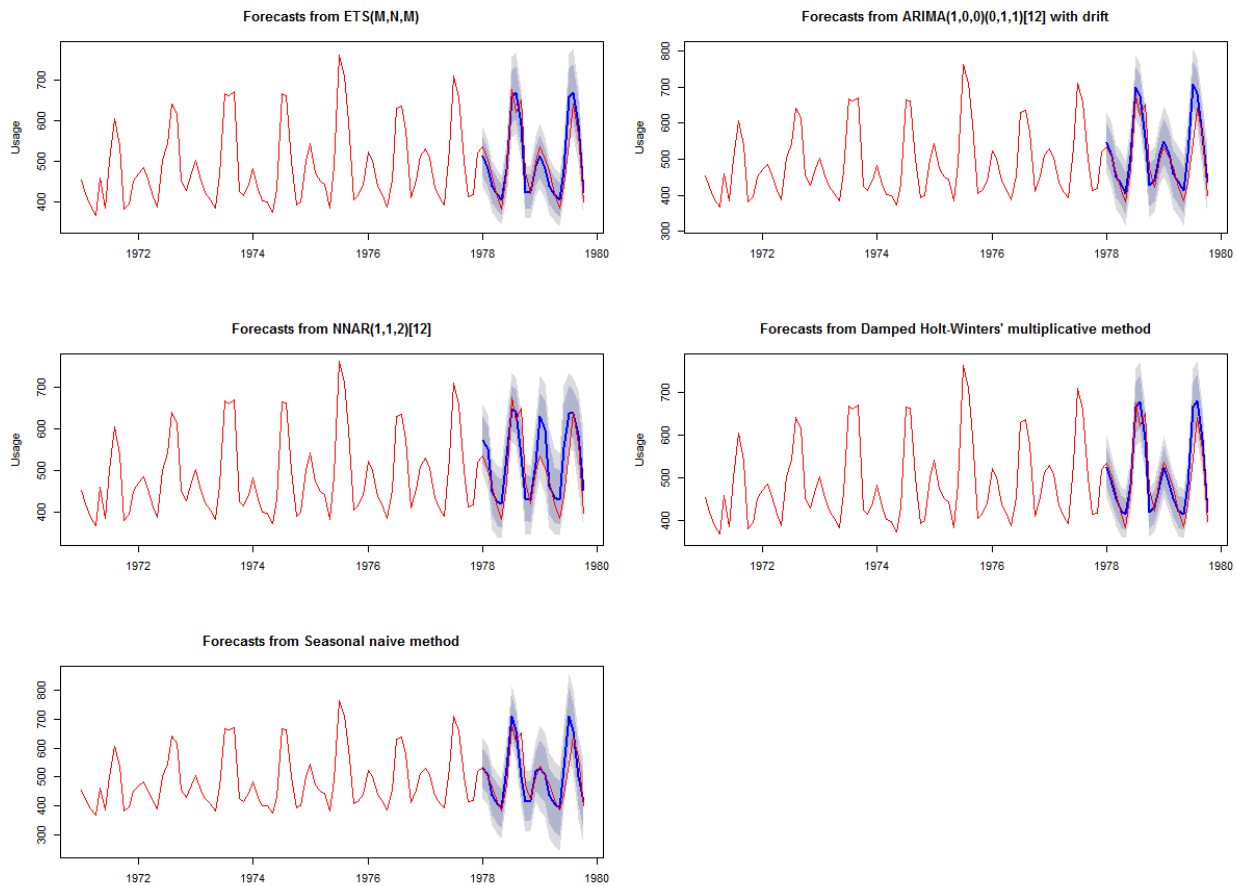


Figure 11 Forecasts on Test dataset

Model Name	Train/Test	ME	RMSE	MAE	MPE	MAPE	MASE	AIC	AICc	BIC
ETS	Training	2.303148	37.84077	27.21975	-0.0128103	5.453404	0.6790791	999.8076	1006.8664	1036.2698
	Test	-1.850765	37.24624	28.81941	-0.5066782	5.624926	0.7189873	N/A	N/A	N/A
ARIMA	Training	1.913783	39.32015	26.02641	-0.0948306	5.179075	0.6493076	756.56	757.15	765.66
	Test	-19.137473	49.49007	34.38078	-4.0404318	6.776653	0.8577326	N/A	N/A	N/A
NN	Training	-0.0107993	43.02041	34.7781	-0.7459867	7.064513	0.867645	N/A	N/A	N/A
	Test	-26.10377	57.46413	45.36837	-5.8849087	9.173367	1.131851	N/A	N/A	N/A
HW	Training	-0.5972309	35.54904	25.41077	-0.55125	5.124509	0.6339486	998.1143	1008.6374	1041.869
	Test	-7.2905775	39.05553	29.16996	-1.605136	5.734369	0.7277329	N/A	N/A	N/A
SNAIVE	Training	9.305556	52.82229	40.08333	1.3646579	7.864821	1	N/A	N/A	N/A
	Test	-3.318182	58.90786	38.22727	-0.8764943	7.286397	0.953695	N/A	N/A	N/A

Figure 12 Performance & Accuracy Metrics

Observations: Figure 11 shows the forecasts on the test dataset, while figure 12 shows the performance/accuracy metrics of the models. The results show that in regards to the error

statistics, HW produced the best error statistics (lowest RMSE, MAE, MASE) compared to the other models on the training set. ARIMA had the second best error statistics, followed by ETS, and NN. All the models performed better than our SNAIVE benchmark in the training dataset. In regards to the test dataset, ETS produced the best error statistics (lowest RMSE, MAE, MASE) compared to the other models. HW had the second best error statistics (produced error statistics that were very close to ETS), followed by ARIMA, and SNAIVE. It's also interesting to note that the prediction intervals for ETS and HW have the narrowest prediction intervals compared to all the models. Surprisingly, NN did not perform better than our SNAIVE benchmark and also performed the worse compared to our core models as well. However, this is not surprising considering that the forecasts above show that the NN model was way off in capturing the "second hump".

Best Model & Conclusion

Overall, it appears that there's some mixed accuracy results between the models, with HW, ETS, and ARIMA performing the best. However, if I had to select one model, I would choose ETS since it performed the best on the test dataset in regards to the error statistics. However, HW is a close second, which isn't surprising considering that it is often the single most accurate forecasting method for seasonal data according to Hyndman.

In regards to future work there are three areas that could help improve the models. First, it would be beneficial to explore an ensemble based approach (e.g., mixing the models). Second, it would be interesting to explore other modeling approaches such as XGBoost. Third, it would be interesting to see if there are additional external regressors that could be added that could make the auto.arima or neural network model more accurate. However, this would require additional data, in which we currently don't have. In the end, I learned a lot from this assignment because I was able to go through the whole time series forecasting process from start to finish, while also incorporating techniques that I learned throughout this course.

Here are my forecasts for the next 5 months:

