

EXECUTIVE SUMMARY

This report provides an analysis of Victorian electricity market. It aims to develop insights and forecast for electricity demand, maximum demand forecast and the first week of 2018. Various data cleaning steps are implemented to ensure the quality of modelling output. We choose ARIMA model with dummy variables and lags to explain the autocorrelation dynamics in the residual. We considered variable significance, residual performance and training error in the model selection process to ensure the accuracy. 6 candidates models are selected from the model selection process. Point forecast and maximum usage forecast performed under these 6 selected models, which are evaluated by RMSE, MAPE, MASE and ACF for point forecast, and CC-test for maximum usage forecast. The result of this model suggests that key drivers include time of the week, price, temperature and humidity. The best point forecast gives good prediction on testing data on 2017 and the interval forecast model passes the CC-test.

EXPLORATION AND PRELIMINARY ANALYSIS

PRELIMINARY ANALYSIS

The dataset shows Victorian electricity market during year 2013 to 2017. All the data provided are available on a half-hourly basis. In other words, each day consists of 48 periods.

SEASONALITY

We plotted demand, price, temperature, humidity and wind (figure 1) against time (2013-2017) to see their patterns. We can roughly see there are multiple layers of seasonality for demand, price, temperature and humidity. We then plotted them in figure 2 on a yearly basis (01-Jan-2013 to 01-Jan-2014), a quarterly basis (01-Jan-2013 to 01-April-2013) and a monthly basis level (01-Jan-2013 to 01-Feb-2013) to examine the finest level of seasonality. We can see clear yearly, quarterly, monthly, weekly and daily levels of seasonality. However, for wind series, there is no obvious seasonal pattern.

THE MAGNITUDE OF THE FLUCTUATION DOES NOT INCREASE OVERTIME

According to the plot for demand and price time series (figure 1), there is no sign of increasing magnitude of the fluctuation over time. Therefore, we do not do log transformation for these two variables.

OUTLIERS

We plotted the price time series, we can see there are around 23 abnormal spikes from 2013 to 2017. These spikes are assumed to be abnormal activities, which have huge impact on coefficients estimations. Thus, we calculate the outliers with IQR method lower boundary $=Q_1 - 1.5(Q_3 - Q_1)$

upper boundary $=Q_3 + 1.5(Q_3 - Q_1)$. Next, we use Winsorization method to replace the extreme value with boundary (upper boundary = 102.076 ; lower boundary = -9.854)

STATIONARITY

Based on the nature of the dataset and context of the case, one of the models we considered using is ARIMA. In order to produce a model that can predict demand accurately, Demand and potential variables that would be affected by time including Price, Temperature, Humidity and Wind will be processed and transformed into stationary time series first for the whole dataset.

Results for stationary tests, plots against time in different time period, ACF and PACF plots for each differenced time series are summarized I (figure 3). Demand, for example, was first test with Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test and augmented Dickey–Fuller test (ADF) tests. The tests showed contradictory results (stationary for KPSS test and non-stationary for ADF test). According to the seasonality analysis, the finest level of seasonality is daily. So, 48 periods of observation will be included in one day; so seasonal differencing with frequency of 48 are applied first; after that, we plot the ACF and PACF, it can be seen that the ACF decaying slowly, which indicates first differencing should be applied here; although the KPSS and ADF tests both shows the series is stationary already, we believed these to test are not useful when testing a large and complicated dataset like in our case, the autocorrelation is high and persist over many lags, which indicates the grey area between non-stationary and stationary. To avoid the risk of running spurious regression, we further take first

difference, and the ACF now decays exponentially and PACF also has a clear cut-off, indicating the data series is stationary.

Same approaches are applied to other 4 series. We found Price, Temperature and Humidity shows similar patterns with Demand (so same differencing methods also applied to them (apply seasonal differencing at frequency equals to 48 first, and then apply first difference on top of seasonal differencing, which will be referred as DD later).

However, Wind demonstrates a different pattern (figure 4). Both tests showed Wind series is already stationary. As mentioned before, we cannot see clear seasonal pattern on its plot. The ACF and PACF plots however, suggested that wind is not stationary. since the ACF is decaying slowly with a relatively clear PACF, we decided to apply first differencing only. The ACF showed a cut-off after first differencing indicating the series is stationary. We later test the series again with seasonal differencing applied at a frequency of 48 periods to confirm our judgement. It can be seen (figure 5) that the PACF is getting worse with clear seasonal pattern, and ACF is not getting any better compared with raw data series. Therefore, we chosen only applying first differencing to Wind series, which will be refereed as D later.

	demand	Price	temp	humid	wind
Diff 48	Y	Y	Y	Y	
Diff 1	Y	Y	Y	Y	Y

DUMMY VARIABLES

To address the spikes existed in ACF and PACF after making the demand series stationary, we decided to add more dummy variables. According to the monthly plot, demand on Monday, Thursday and Friday are considerably higher than rest of the week. Therefore, we add seven weekday dummy variables.

LAGGED TIME AND PRICE

In terms of selecting the number of lag for exogenous variable, we utilize the observation that we have on the time series plot and ACF/PACF plot. With the presence of both AR and Seasonal Autoregressive model (SAR) pattern in weather and price series, we consider the first 4 lags (2 hours) in AR term and one lag in SAR term (1 day). The price variable itself is lagged one period to avoid simultaneity issues since demand and price they can affect each other cotemporally.

TECHNICAL ANALYSIS

ASSUMPTION

First, we assume that the price data and the weather data would not change dramatically over a week. Second, we assume that the extreme value of price is abnormal activity, which does not help describe the underlying process.

MODEL CONSTRUCTION

At first, we consider different models such as halt-winter, simple OLS regression and ARIMA model for our forecasting purpose. However, halt-winter is not suitable for our task since it cannot capture multiple seasonality layers. The fact that simple OLS regression cannot capture the MA dynamic characteristics of the data make it inferior to the ARIMA specification, thus we mainly consider the ARIMA model in order to achieve more accurate prediction.

All the variables that are used in the system including demand has been differenced to ensure stationarity. The reason that we difference the variables before fitting the ARIMA model instead of using d and D parameters is because different variables have different order of integration.

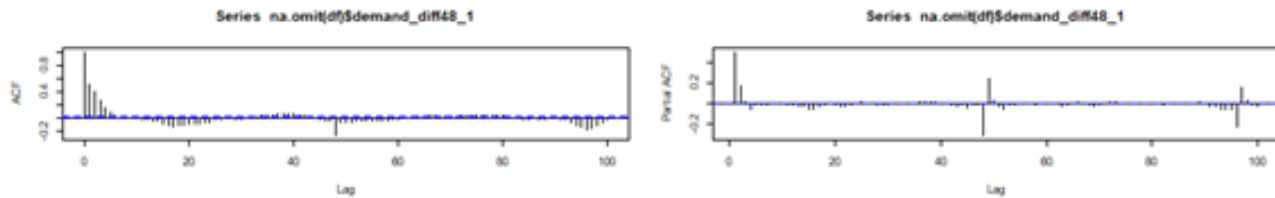
To construct our ARIMA model for forecasting electricity demand, we consider all exogenous variables such as weather and price as well as their lagged values to capture their autocorrelated nature, and ARMA term to capture the autocorrelation in residuals.

MODEL SELECTION FOR BEST FITTED MODEL

With all predictors we included for model construction, we carried out a model selection to achieve the best performing model for both point forecasting and max forecasting. We choose to build dynamic ARIMA models as it allows MA terms and captures the autocorrelation on residuals, which should have explanatory power than the linear model with lagged variables. We first get a range for likely AR and MA orders from the ACF and PACF plots of demand in figure 6. The differencing term should be 0 as all differencing has been done before we do build the model. The ACF plot of demand shows around 6 spikes that are outside the interval. Hence, we choose 4,5,6 to be possible MA orders. For AR term we

look at the PACF plot. We see there are around 4 spikes out of the interval, so we choose to include 3,4,5 for AR orders. Quickly running an auto function on ARIMA also suggests (4,0,4) is a good choice.

There are also seasonal patterns in the ACF and PACF plots. Spikes can be seen around 48 and 96 lag period, indicating that we could include 1 and 2 for the order of MA terms in the seasonal part. Similar pattern can be seen in the PACF plot on the right, hence we include 1 and 2 for the AR term's order in the dynamic ARIMA model.



To summarize in Table1, for demand lag we consider the following. In total we consider 36 combinations of these combinations. When doing model selection, we follow a backward selection process where we start with higher number of the orders and eliminate those are not significant.

Non-seasonal			Seasonal ₄₈		
P	D	Q	p	d	Q
4	0	4	1	0	1
4	0	4	1	0	2
4	0	5	1	0	1
...

We also follow a backward selection process for the consideration of exogenous variables. We start with including all the exogenous variables including price, temperature, humidity, wind, day of the week, time of the day and dummy variables such as holiday. We also consider all possible lags for price, temperature, humidity and wind. We believe the value of the previous 2 hours of these variables should give information on the next demand, as well as values of 12 hours and 24 hours before, which can be seen from the ACF and PACF plots in Figure 4. For dummies we do not include lagged variables as there is no point of doing so.

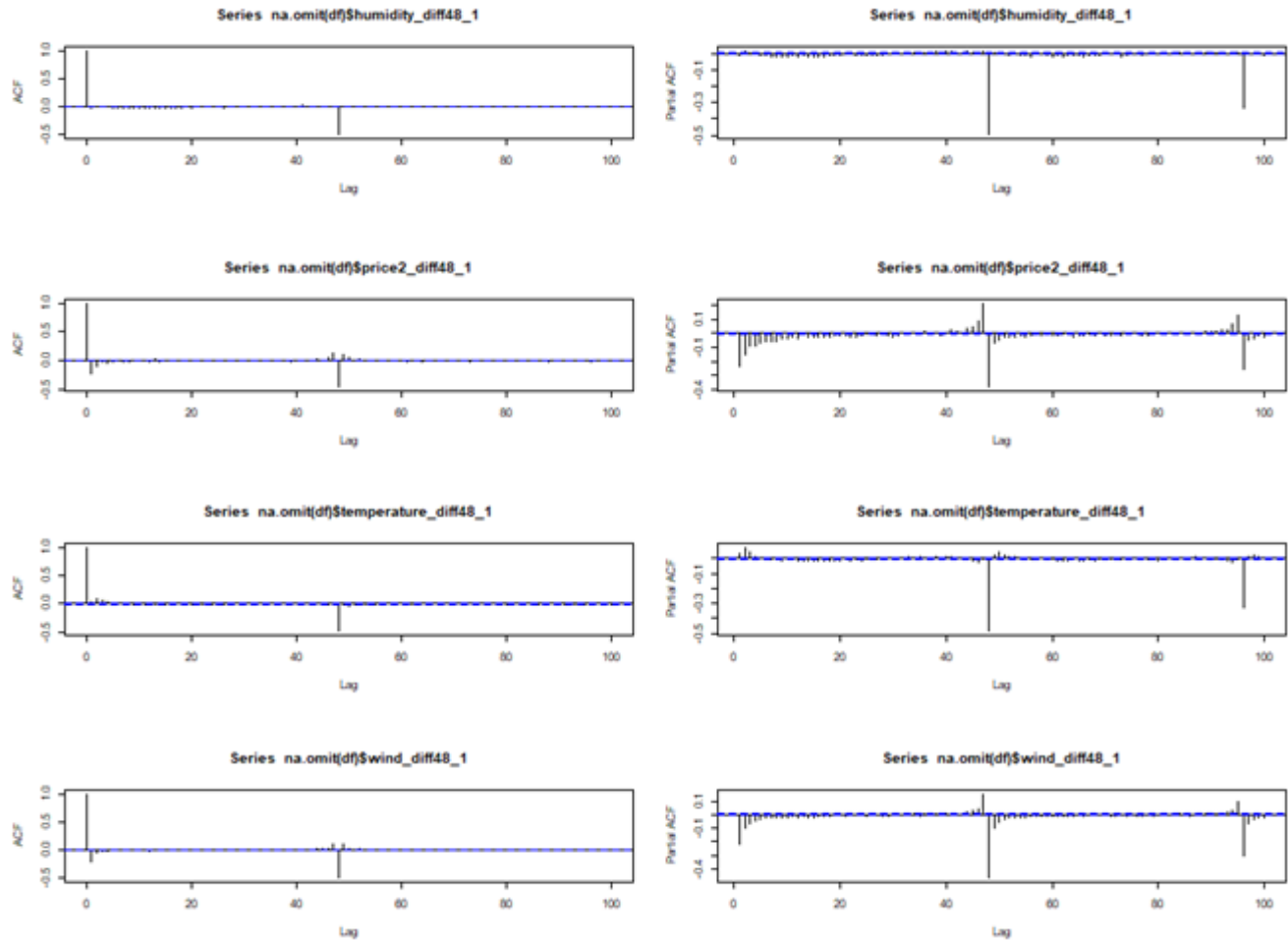


Figure4

Below we present part of our model selection journey in Table 1 in the appendix. Not all fitted models are included.

EXPANDING WINDOW FORECAST:

Since our final objective is to provide a 48×7 -steps ahead forecast, we also use the expanding with the same forecast horizon to evaluation and select our best forecast model. The mechanism is to use the expanding windows for training set so we have 48×7 new observations in each iteration. Then based on the new training set, we forecast again the next one week of demand forecast. To do the forecast in each run, we need to feed in the information about the exogenous variables. However, since we are not supposed to see the observations for the test set, we find some proxies for the exogenous variables. From *figure X* (weather plot), we can clearly observe that the weather series exhibits strong time in day fluctuation, such as the temperature is low in the morning and night. Thus, we assume that the weather

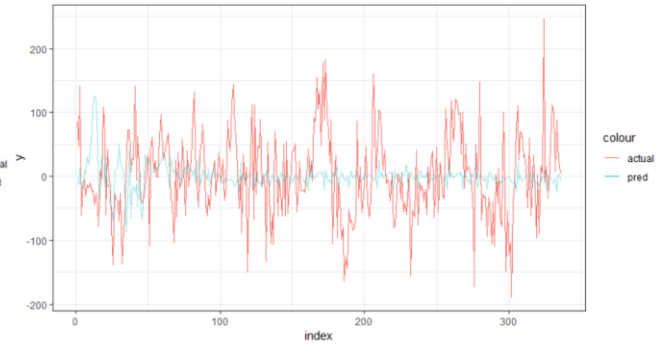
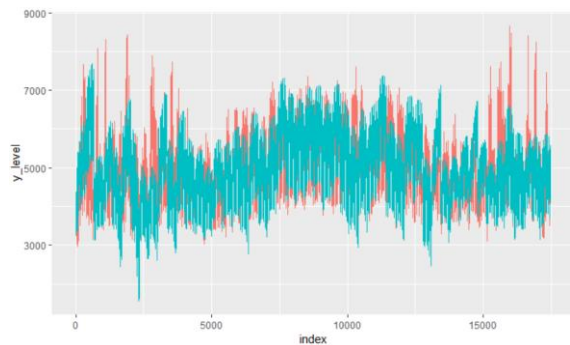
pattern tends to repeat itself a week, and two consecutive days would have similar weather values for the same time periods of day. Thus, we use the last day of our expanding training sample dynamically to proxy the next 7 days weather values. In regards to the dummy variables such as time of day and weekday, their values for the test time period are observable, so no proxies are needed.

POINT FORECAST EVALUATION

After we achieve the best fit models from the previous section, we then carry out the point forecast evaluation. The methodology has been explained in detail in the last section. Then from the 52-week expanding window testing on 2017 data, we have summarized the results in Table X. The model mapping can be found in the appendix.

	RMSE	MAPE	MASE	ACF
Model 1	84.17729	175.7477	0.9965935	0.5447099
Model 2	83.50459	173.6467	0.9911151	0.5395911
Model 3	84.18671	176.4797	0.9969679	0.5446738
Model 4	85.52303	215.7628	1.004794	0.5236853
Model 5	86.32421	223.4976	1.012874	0.5339425
Model 6	87.25232	262.1538	1.029842	0.5486842

From the table we can see the performance of all the models are very similar. Model 4 has a MASE larger than 0, indicating it is not as good as using the previous day's value. Among other models, Model 2 performs the best in terms of RMSE, MAPE and MASE. Hence, for giving point forecast we choose Model 2 to be our best forecast model. The forecast of the whole 2017 vs the actual data are presented in Figure X. The plot on the left shows the prediction and actual for the demand and the plot on the right shows the comparison after differencing.

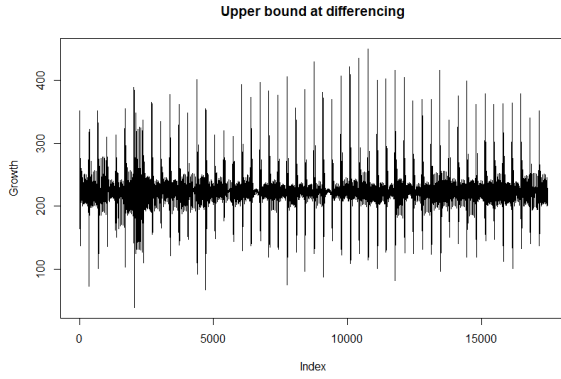


INTERVAL FORECAST EVALUATION

One main purpose of this project is to predict the expected maximum demand. To measure this, we construct an interval forecast with a 99% confidence interval and see if 99th percentile of the demand can be covered by the forecast. We then calculate the percentage and perform the CC-test to check if the model achieves the requirement of 99% percentile as a way of tail test. Results of each of the selected 5 models are listed below in Table X.

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Normal	p_hat	0.982257	0.982429	0.982143	0.983001	0.981324	0.980206
	p	1.80E-20	1.03E-19	5.51E-21	2.84E-17	1.43E-18	0.95E-15
Student T	p_hat	0.989068	0.989641	0.989183	0.988095	0.988095	0.987236
	p	0.222648	0.634996	0.283895	0.014024	0.008137	0.004679
Skewed Normal	p_hat	0.982429	0.982544	0.982315	0.983116	0.983273	0.983367
	p	1.03E-19	3.26E-19	3.23E-20	8.39E-17	9.86E-19	9.98E-20
Skewed T	p_hat	0.99027	0.990556	0.990327	0.988725	0.982634	0.981537
	p	0.718473	0.455666	0.661894	0.096877	0.123841	0.154786

Although some of our models fail the CC-test, we understand that this is a difficult task for the model. Demand by itself has quite a number of outliers, which are hard to be captured by the ARIMA model. In the future, other types of models can be leveraged to capture the outliers if the purpose of the prediction is specifically for maximum demand. In addition, as the number of observations increase, the confidence interval gets really narrow, because the standard error is $\frac{\sigma}{\sqrt{n}}$ which decreases as n get larger. This can be reflected by that \hat{p} is over 98% already but still the p-value is extremely small. Model 2 with skewed student-t error distribution has the highest percentage which means it covers the most of the actual demand data among of the selected models we have. It is the same as the best point forecast model which is Model 2. The upper bound of the 2017 prediction can be seen in Figure 2.



CONVERTING

After getting the predicted DD-demand, we need to convert it back to the level predicted demand using the formula

$$D_t = \Delta(\Delta m D_t) + D_{t-1} + D_{t-48} - D_{t-49}$$

For the fact that DD-demand is forecasted at the weekly frequency, we also update the actual demand value on a weekly basis. Using dynamic loop, we are able to use the actual data when it is available. For example, for each weekly forecasting horizon, when $t \leq 48$ we are able to observe the actual value for y_{t-48} , then we will use the predicted value $y_{t-48}(\hat{y})$ when $t > 48$.

KEY RESULT DISCUSSION

The best fitted model, the best point forecast model and the best maximum demand forecast model are the same. We choose ARIMA(4,0,6)(2,0,1)[48] errors:

$$\begin{aligned} \Delta(\Delta m D)_t = & 4.4249 + 3.6389 \text{ Mon}_t - 8.9788 \text{ Tue}_t - 5.7906 \text{ Wed}_t - 4.5095 \text{ Thurs}_t - 5.705 \text{ Fri}_t - 9.321 \text{ Sat}_t - \\ & 0.2425 \text{ Peak}_t - 0.0394 \text{ Shoulder}_t + 2.0593 \Delta(\Delta m \text{Temp})_t + 0.4265 \Delta(\Delta m \text{Humid})_t + 0.9268 \\ & \Delta(\Delta m \text{price})_t - 0.0022 \Delta(\Delta m \text{price})_{t-24} + 0.1568 \Delta(\Delta m \text{price})_{t-48} - 0.0751 \Delta(\Delta m \text{Humid})_{t-24} + 0.1404 \\ & \Delta(\Delta m \text{Humid})_{t-48} - 0.4439 \Delta(\Delta m \text{Temp})_{t-24} + 1.0673 \Delta(\Delta m \text{Temp})_{t-48} + \eta_t \end{aligned}$$

$$\begin{aligned} \eta_t = & -0.051 \eta_{t-1} + 0.076 \eta_{t-2} + 0.127 \eta_{t-3} - 0.046 \eta_{t-4} + 0.536 \varepsilon_{t-1} + 0.333 \varepsilon_{t-2} + 0.144 \varepsilon_{t-3} + \\ & 0.130 \varepsilon_{t-4} + 0.057 \varepsilon_{t-5} + 0.021 \varepsilon_{t-6} + 0.133 \eta_{t-48} - 0.072 \eta_{t-48} - 0.848 \varepsilon_{t-48} \end{aligned}$$

PRICE DEMAND SENSITIVITY

According to our fitted model, demand is sensitive to price, as the price variable is statistically significant.

KEY DRIVERS

From the fitted model, we found price, temperature, humidity and weekend (week day exclude Monday) are the key drivers of the electricity demand. We also tested the effect of holiday and working hour of the day, but these two variables are statistically insignificant. Although time of day (electricity usage among whole day (peak, shoulder, low) defined by electricity company) are not statistically significant, we believe they are important in predicting demand, so we keep them in the model. Moreover, we find that positive Thursday, Friday, peak, price, temperature and humidity are positively related with demand, while, we find that weekend, Tuesday, Wednesday, off-peak and shoulder are negatively related with demand.

RESIDUAL

Although the model failed the Box-Ljung test, according to figure x, we can see that the model is able to capture autocorrelation dynamics for the first ten lags. This might due to the nature of the complex dataset, as there are complicated autocorrelations observed.

TRAINING ERROR

Training errors are presented in figure x, MASE equals to 0.4280, indicating the model error is smaller than using sample mean. Although ACF1 test shows there are still some autocorrelation existed, this might result from the complicated nature of dataset.

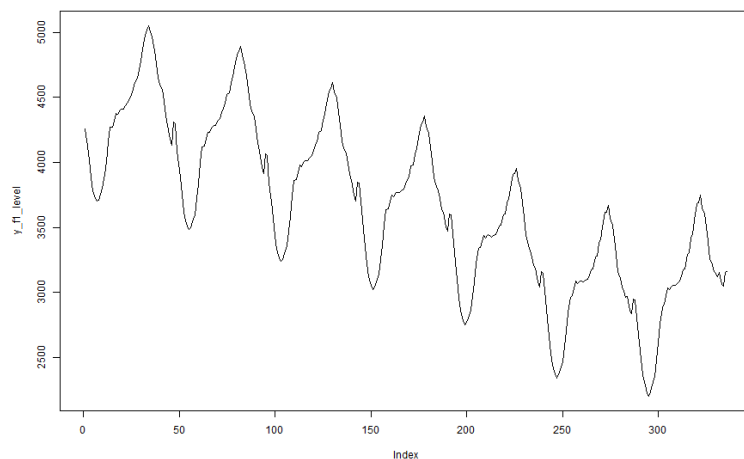
POINT FORECAST PERFORMANCE

The forecast cannot perfectly capture all the variations. Due to the large forecast horizon ($h = 336$), we can see that the fluctuation of the forecasting decays over the forecast horizon. This might result from the complicated nature of the dataset. The ACF test shows our model is able to explain some autocorrelations.

INTERVAL FORECAST PERFORMANCE

Our best forecast model, Model 2, successfully covers 99% of the demand with a confidence interval of 99%. More specifically, for 7-day expanding window testing on 2017, the best interval forecast ARIMA model captures 99.056% of the demand with a skewed student-t distribution error distribution. The forecast on test data also passes the CC-test.

PREDICTION FOR FIRST WEEK OF 2018



The above graph shows the prediction made by the best forecast model for the first week of 2018. Key insights include that it is a decreasing trend with the highest being 5100 and lowest at 2200.

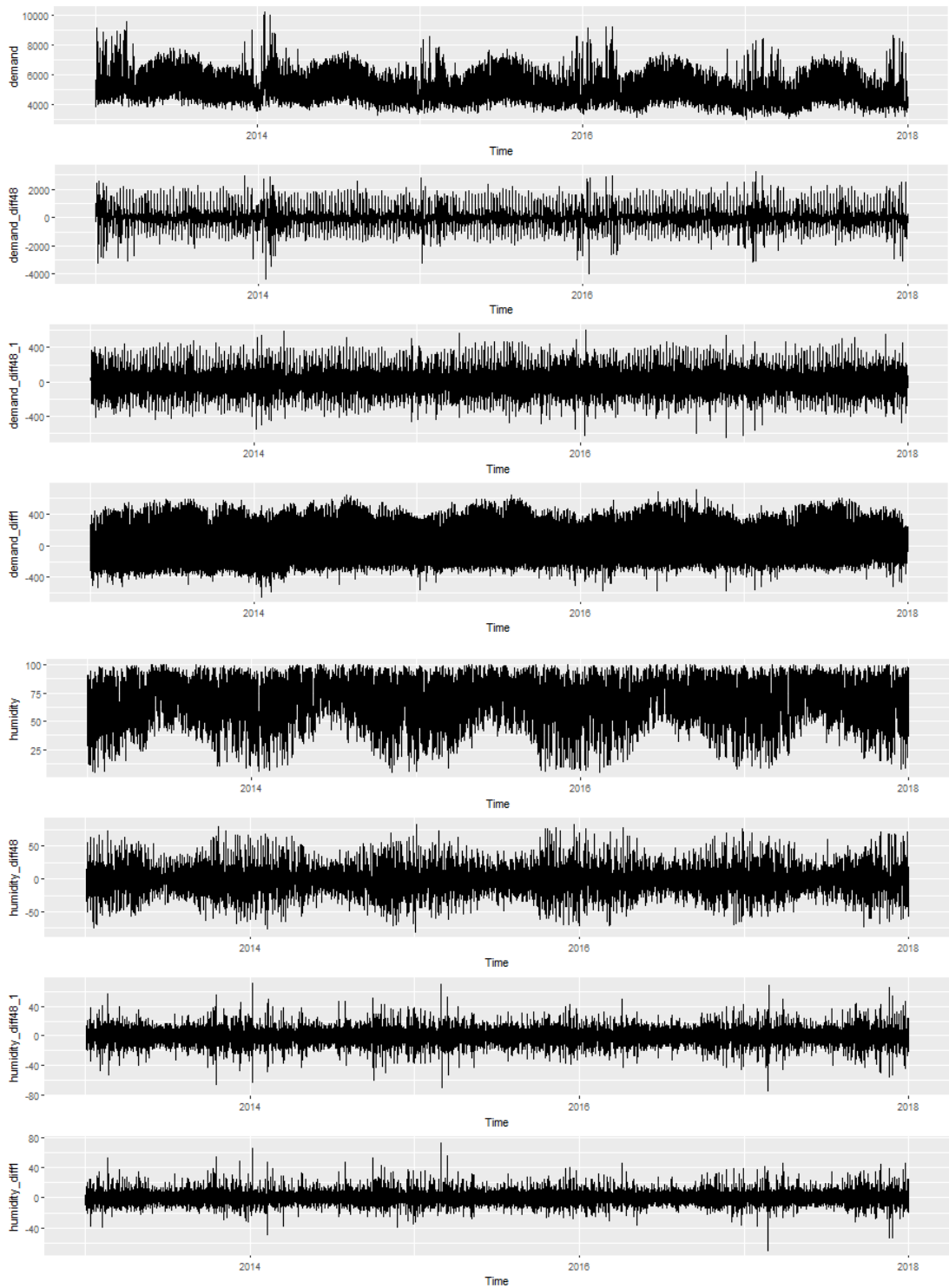
CONCLUSION & FUTURE STUDY

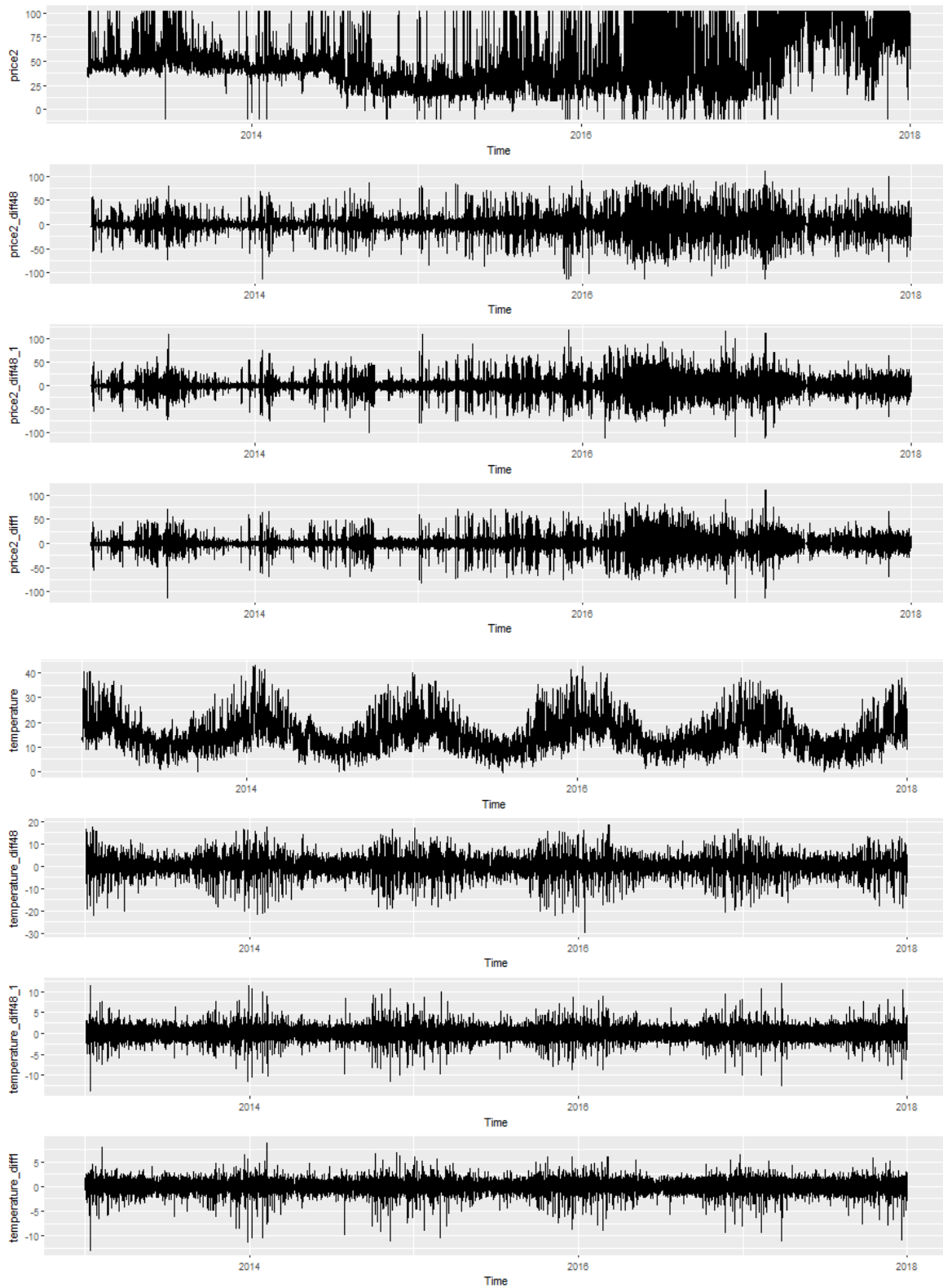
In response to the case study, the model suggests demand is sensitive to changes in price. Key drivers include weekend, Tuesday, Wednesday, price, temperature, humidity. Although time of the day (electricity usage) is not shown as statistically significant, we still include it for forecasting for better accuracy. Holiday is not significant in our model and excluded.

In the future, we could also incorporate GDP, inflation rate, industry usage and demographic features into electricity demand analysis. Monthly dummy could also be utilized to capture the fluctuation of demand over the year. Also, with the given dataset, we could consider converting temperature data into dummy variable to assess four seasons effect.

APPENDIX

Figure 1





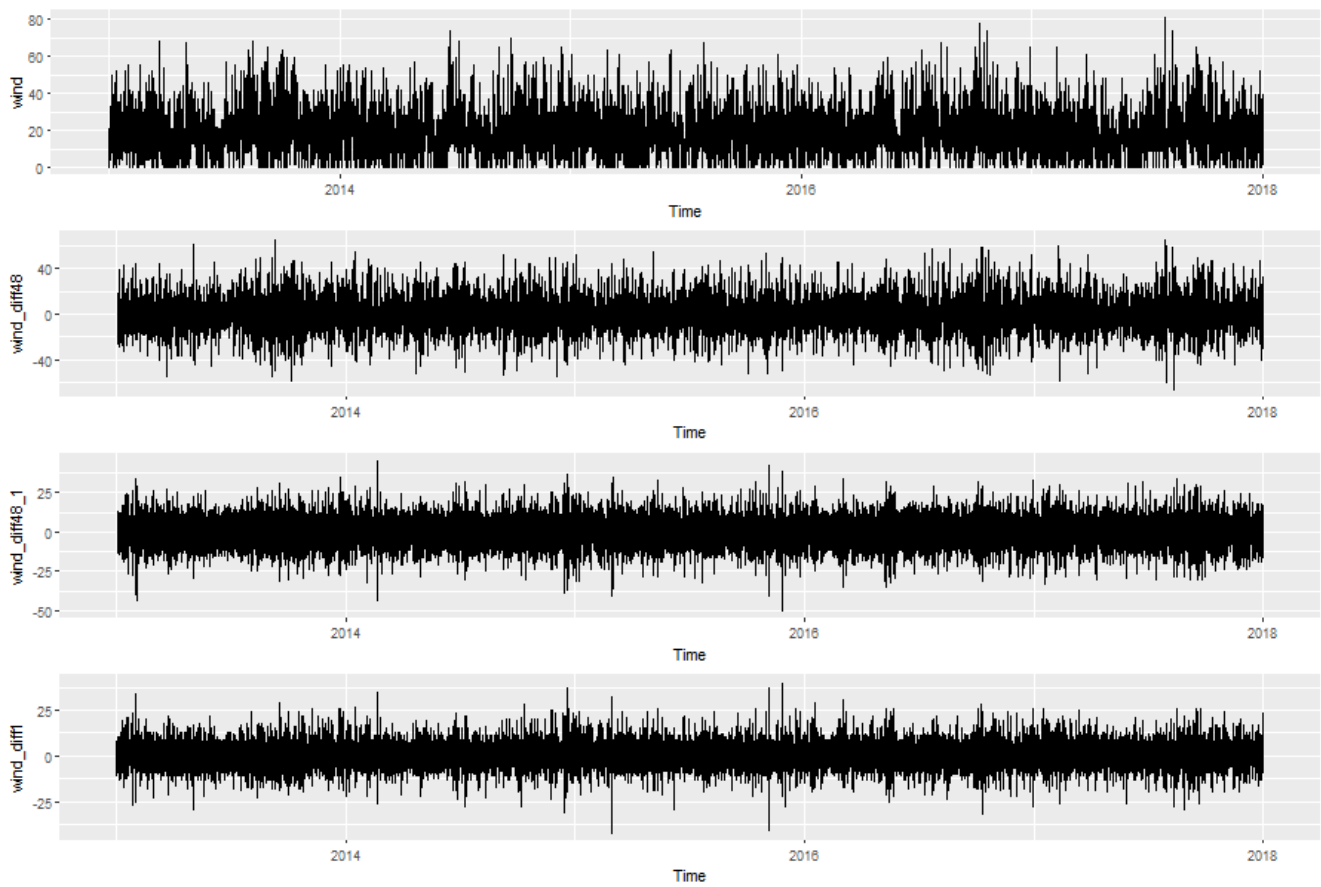
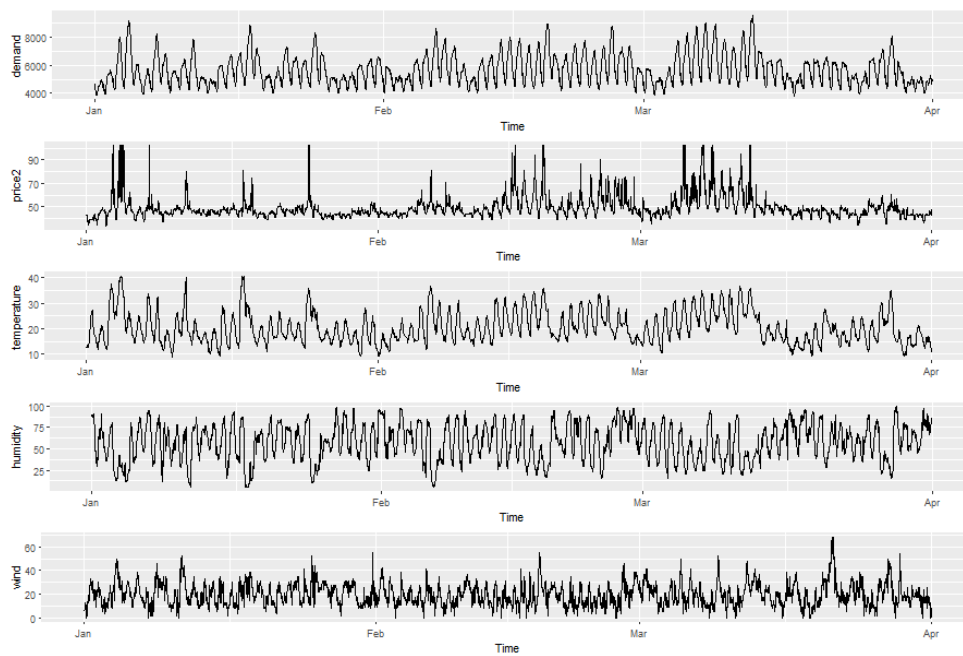


Figure2



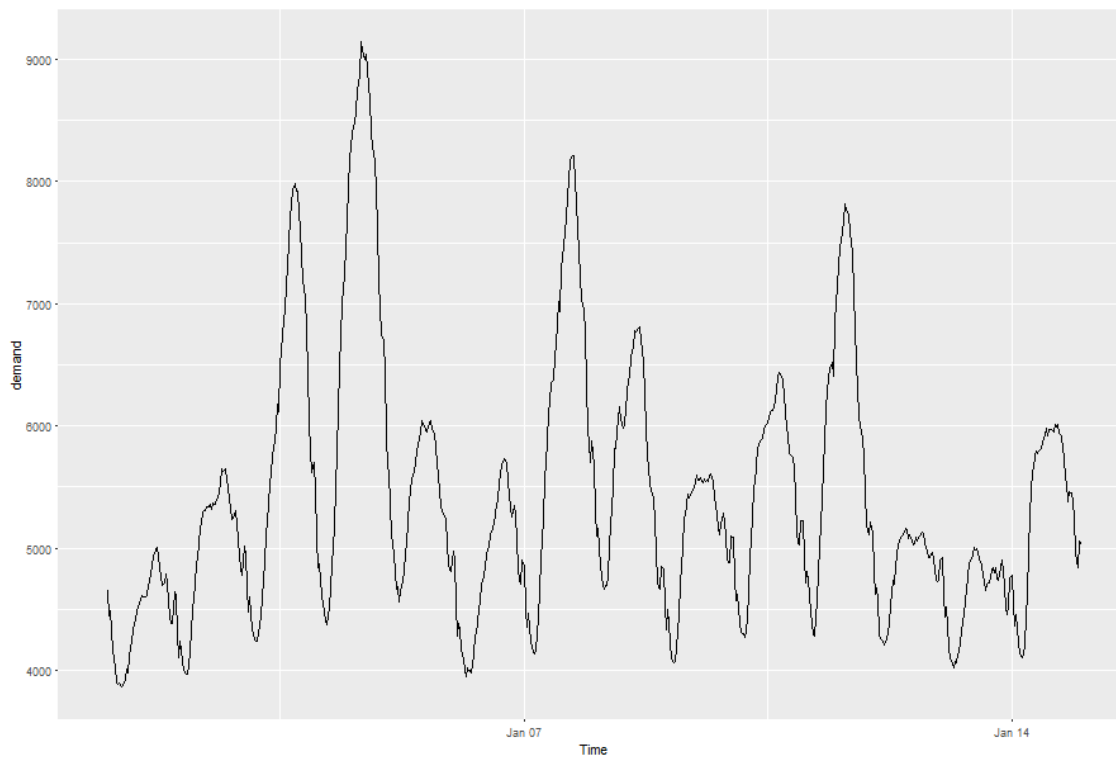
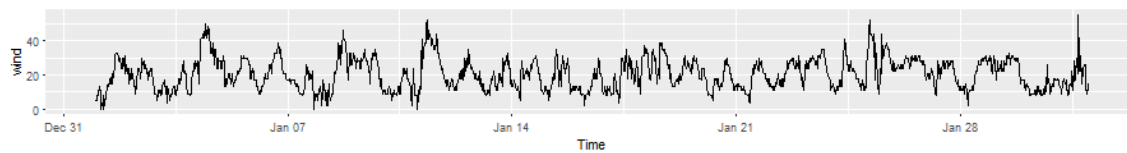
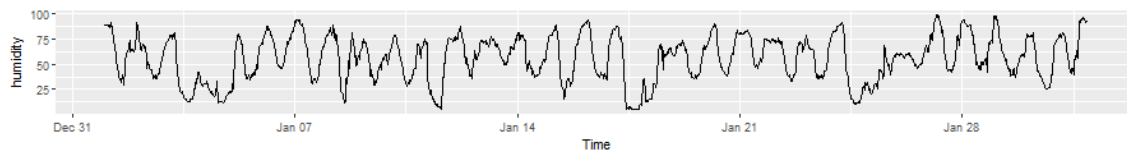
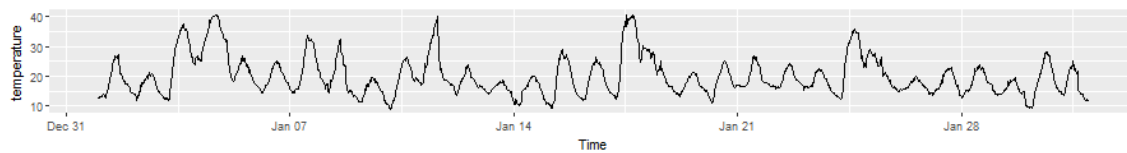
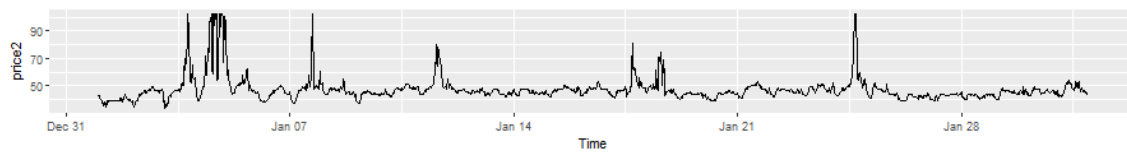
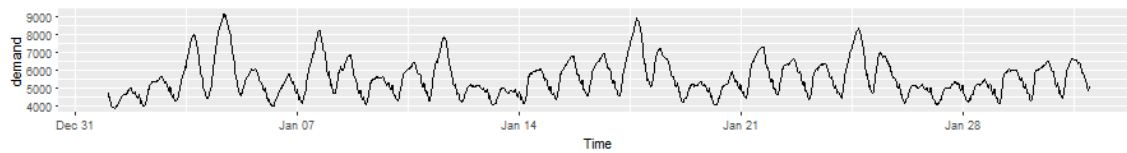
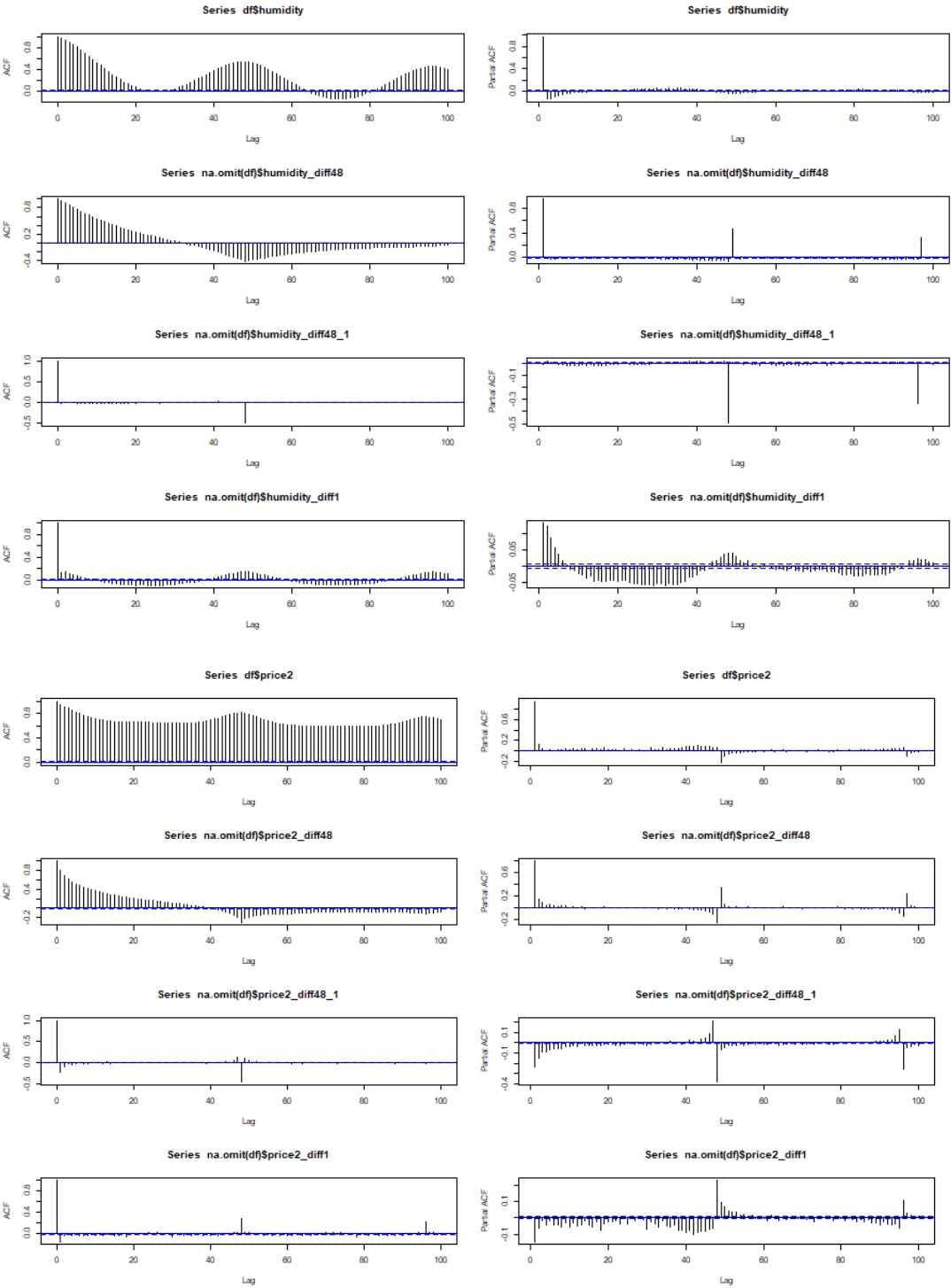
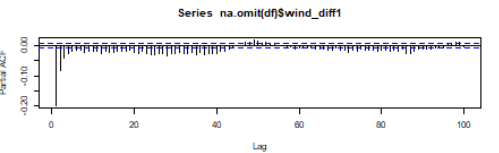
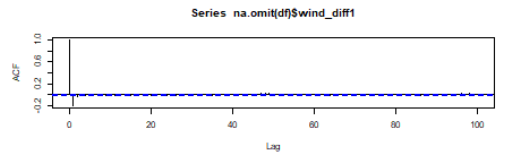
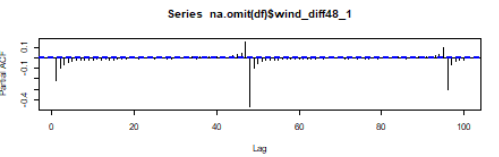
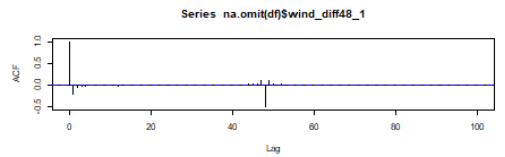
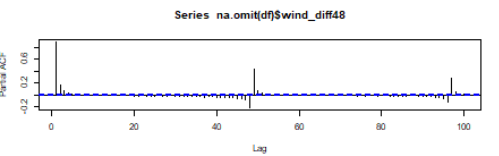
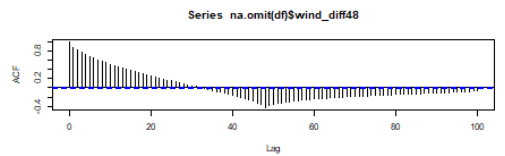
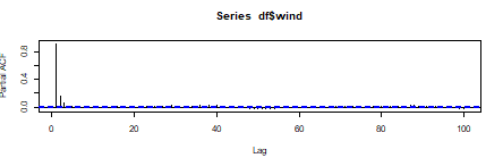
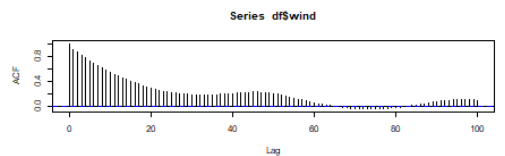
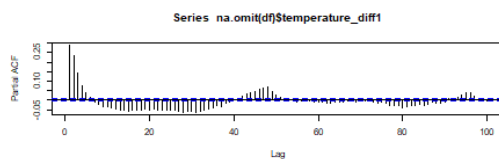
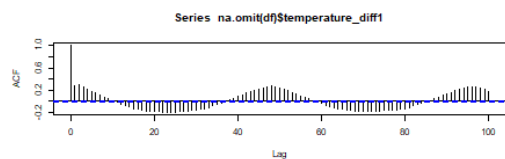
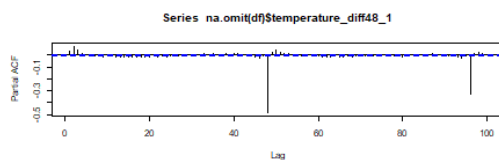
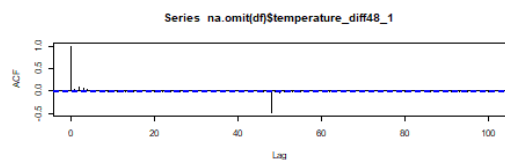
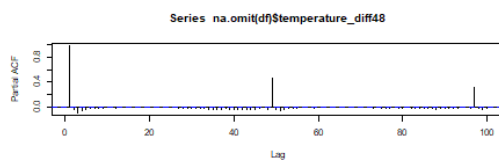
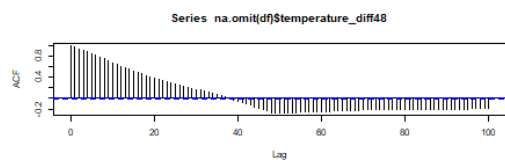
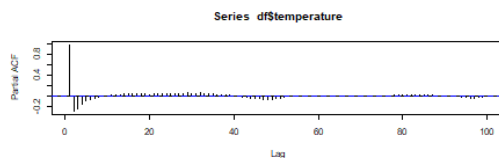
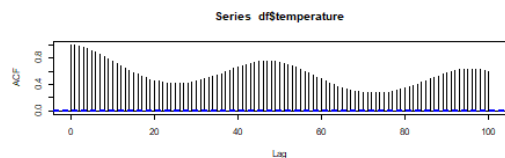
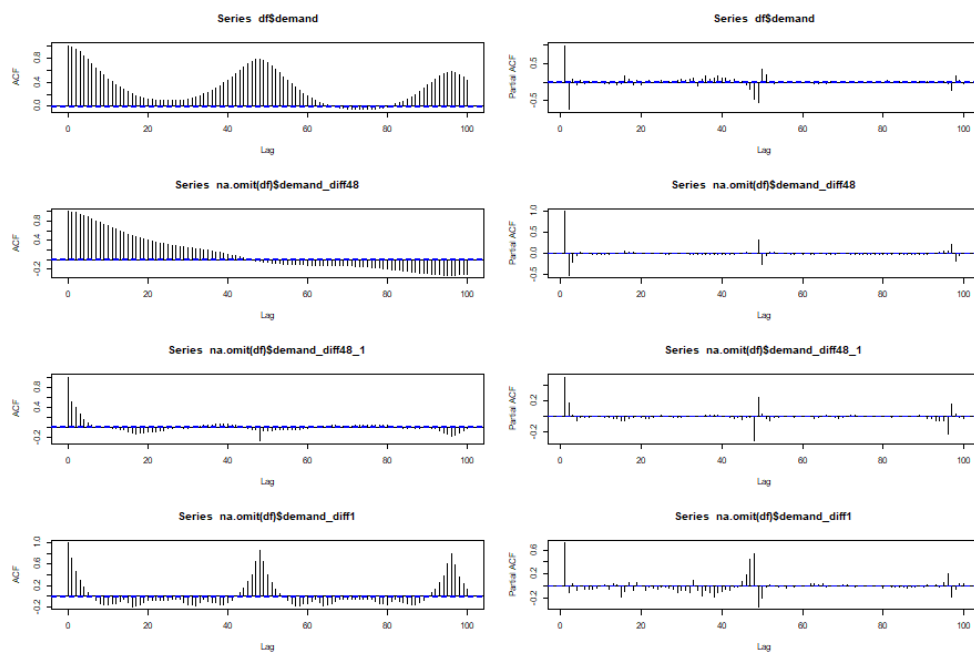


Figure 3







Model Details:

Model 1: (4,0,6), (1,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Model 2: (4,0,6), (2,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Model 3: (6,0,4), (1,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Model 4: (6,0,4), (2,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Model 5: (6,0,6), (1,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Model 6*: (6,0,6), (2,0,1) xreg="mon", "tue", "wed", "thur", "fri", "sat", "peak", "shoulder", "temperature_diff48_1", "humidity_diff48_1", "price2_diff48_1", "lagprice_24", "lagprice_48", "laghumid_24", "laghumid_48", "lagtemp_24", "lagtemp_48"

Table 1

	fit1	fit2	fit3	fit4	fit5	fit6
AR1	-0.0387 (0.830)	-0.1053 (0.525)	0.1031 (0.337)	-0.0150 (0.898)	-0.0951 (0.447)	0.0741 (0.368)
AR2	0.0698 (0.555)	0.1773 (0.119)	0.2786 (0.000)	0.1416 (0.004)	0.1300 (0.059)	0.1166 (0.022)
AR3	0.1407	0.5278 (0.000)	0.2935 (0.026)	0.5848 (0.000)	0.6202 (0.000)	0.5467 (0.000)
AR4	-0.0553	-0.295 (0.000)	-0.2327 (0.000)	-0.4474 (0.000)	-0.3787 (0.000)	-0.3939 (0.000)
AR5				0.0496 (0.269)	0.0194 (0.761)	0.0351 (0.381)
AR6				0.0072 (0.657)	-0.0061 (0.740)	0.0020 (0.885)
MA1	0.5231 (0.004)	0.5908 (0.000)	0.3810 (0.000)	0.4997 (0.000)	0.5794 (0.000)	0.4106 (0.000)
MA2	0.3328 (0.006)	0.2570 (0.184)	0.0555 (0.517)	0.2507 (0.000)	0.3001 (0.000)	0.2329 (0.001)
MA3	0.1278	-0.2855 (0.056)	-0.1806 (0.063)	-0.3657 (0.000)	-0.3626 (0.000)	-0.3458 (0.000)
MA4	0.1320 (0.000)	0.1641 (0.000)	0.1136 (0.000)	0.2830 (0.000)	0.2232 (0.034)	0.2296 (0.001)
MA5	0.0565	0.0010 (0.974)				
MA6	0.0201	-0.0090 (0.466)				
price	0.0016 (0.290)	0.0001 (0.940)		0.0013 (0.398)		
humidity	-0.0106 (0.385)			-0.0129 (0.335)		
wind	0.0195 (0.281)	0.0130 (0.401)		0.0204 (0.292)		
temperature	-0.0224 (0.427)			-0.0269 (0.259)		
Mon	3.7080 (0.105)	0.2183 (0.924)	3.6306 (0.110)	2.8484 (0.209)	2.4449 (0.281)	3.1113 (0.173)

Tue	-8.8804 (0.000)	-13.23 (0.000)	-8.9912 (0.000)	-9.9847 (0.000)	-10.449 (0.000)	-9.667 (0.000)
Wed	-5.6986 (0.000)	-8.964 (0.000)	-5.7959 (0.003)	-6.5842 (0.000)	-6.9283 (0.000)	-6.3496 (0.001)
Thu	-4.4018 (0.024)	-8.026 (0.000)	-4.5228 (0.020)	-5.4 (0.005)	-5.767 (0.003)	-5.151 (0.008)
Fri	-5.5902 (0.008)	-8.969 (0.000)	-5.7093 (0.006)	-6.5128 (0.000)	-6.9063 (0.001)	-6.2927 (0.003)
Sat	-9.1601 (0.000)	-15.148 (0.000)	-9.2746 (0.000)	-10.608 (0.000)	-11.222 (0.000)	-10.204 (0.000)
Working hour	-0.1289 (0.384)	-0.0526 (0.688)		-0.1638 (0.268)	0.0195 (0.879)	
peak	-0.3344 (0.135)	-0.2948 (0.119)	-0.2273 (0.227)	-0.3240 (0.149)	-0.2616 (0.220)	-0.2449 (0.236)
shoulder	-0.1065 (0.455)			-0.0792 (0.579)	-0.0280 (0.836)	