Homework #2

Aaron Ge, 11337837
Haotong Sun, 11337172
Danni Weng, 11304053

Problem 1

a.

Dataframe head of our data (2 lagged periods and their squared values)

	Return	Lag_1	Squared_Lag_1	Lag_2	Squared_Lag_2
2	-0.002334	0.027379	0.000750	-0.008704	0.000076
3	0.005135	-0.002334	0.000005	0.027379	0.000750
4	-0.013877	0.005135	0.000026	-0.002334	0.000005
5	0.011446	-0.013877	0.000193	0.005135	0.000026
6	-0.001554	0.011446	0.000131	-0.013877	0.000193

The Python function returns this data frame. It contains the returns of our 3 stocks (AAPL, DIS, H), along with their lagged values of 1 and 2 periods prior, and their respective squared values. We have also dropped the missing values that resulted from the lag.

b.

OLS Regression of our data(2 lagged periods and their squared values)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Return   R-squared:                       0.001
Model:                            OLS   Adj. R-squared:                  0.001
Method:                 Least Squares   F-statistic:                     1.916
Date:                Mon, 25 Mar 2024   Prob (F-statistic):              0.105
Time:                        18:58:44   Log-Likelihood:                 17323.
No. Observations:                6802   AIC:                        -3.464e+04
Df Residuals:                    6797   BIC:                        -3.460e+04
Df Model:                           4
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
const             0.0005      0.000      2.192      0.028    5.66e-05       0.001
Lag_1            -0.0261      0.012     -2.147      0.032      -0.050      -0.002
Squared_Lag_1     0.2733      0.181      1.508      0.132      -0.082       0.629
Lag_2            -0.0029      0.012     -0.238      0.812      -0.027       0.021
Squared_Lag_2     0.1401      0.181      0.774      0.439      -0.215       0.495
==============================================================================
Omnibus:                     1292.349   Durbin-Watson:                   2.000
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            36031.131
Skew:                           0.129   Prob(JB):                         0.00
Kurtosis:                      14.272   Cond. No.                         870.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

We use an OLS regression on the data in part A to find the relationship between the current return (dependent variable “Return”) and its past returns (independent variables “Lag_1, Squared_Lag_1, Lag_2, Squared_Lag_2).

The coefficient for const is 0.0005, meaning that holding all other variables constant, the return tends to be 0.0005.

For Lag_1, the coefficient is -0.0261, which suggests that returns from the previous period have a negative correlation with current period returns. The p-value is 0.028, which is statistically significant at the 5% level.

Squared_Lag_1 has a coefficient of 0.2733, and a p-value of 0.132 which is higher than 0.05, making it not statistically significant at the 5% level. There isn’t enough evidence to suggest that squared returns from the previous period correlate with current period returns.

Lag_2 and Squared_Lag_2, with coefficients -0.0029 and 0.1401 and p-values 0.812 and 0.439 respectively, are not statistically significant at the 5% level. This means that there is a lack of evidence to suggest that returns and squared returns from two previous periods are correlated with current period returns.

The Durbin-Watson test for autocorrelation gives us a value of 2, indicating no autocorrelation in the residuals of our model.

The R-squared value of our OLS is 0.001. It indicates that only 0.1% of the variance in the dependent variable (return) can be explained by the model’s independent variables (Lag_1, Lag_2, Squared_Lag_1, Squared_Lag_2). An R-squared value this low suggests that the independent variables have very little explanatory power on the dependent variable.

The adjusted R-squared value is also 0.001. It is a modified version of R-squared which is adjusted for the number of predictors in the model and offers a more accurate measure of goodness of fit. It penalizes the R-squared for each variable added to the regression. Since both adjusted R-squared and R-squared have the same value, it also suggests that the independent variables have very little explanatory power on the dependent variable.

OLS Regression of our data(50 lagged periods and their squared values)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Return   R-squared:                       0.047
Model:                            OLS   Adj. R-squared:                  0.033
Method:                 Least Squares   F-statistic:                     3.276
Date:                Wed, 27 Mar 2024   Prob (F-statistic):           3.71e-25
Time:                        05:58:38   Log-Likelihood:                 17348.
No. Observations:                6752   AIC:                        -3.449e+04
Df Residuals:                    6651   BIC:                        -3.381e+04
Df Model:                         100
Covariance Type:            nonrobust
==============================================================================
```

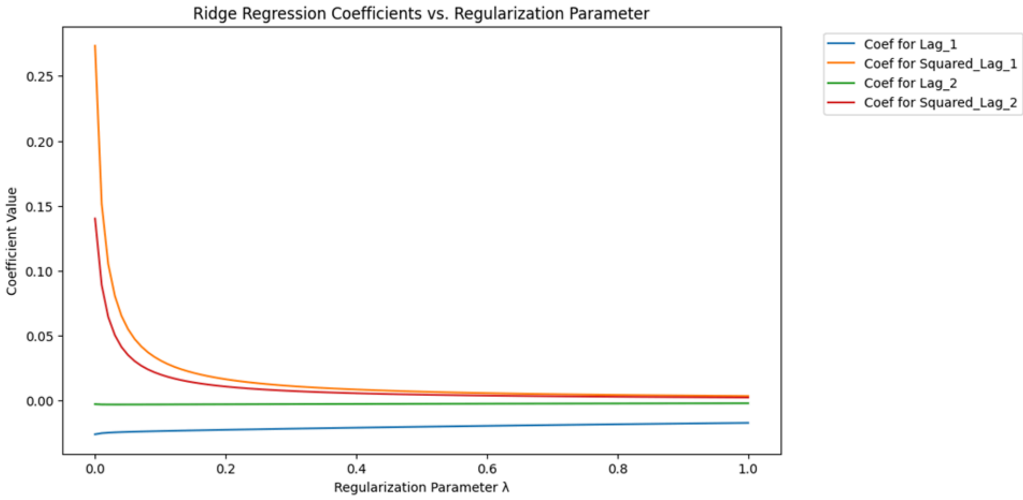
OLS Regression of our data(2000 lagged periods and their squared values)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Return   R-squared:                       0.821
Model:                            OLS   Adj. R-squared:                 -0.076
Method:                 Least Squares   F-statistic:                    0.9156
Date:                Wed, 27 Mar 2024   Prob (F-statistic):              0.949
Time:                        06:20:37   Log-Likelihood:                 16227.
No. Observations:                4802   AIC:                        -2.445e+04
Df Residuals:                     801   BIC:                             1461.
Df Model:                        4000
Covariance Type:            nonrobust
==============================================================================
```

As we add more independent variables (lagged variables and their squared values) in the regression, the R-squared will increase. Here, we have the R-squared for the regression of 50 and 2000 lagged periods with respective values of 0.047 and 0.821. Yes, we have a high R-squared but this introduces the problem of overfitting which linearly increases as we add more regressors. If we look at the adjusted R-squared, the respective values are 0.033 and -0.076. Overall, both R-squared and adjusted R-squared values are still low, and much more still cannot be explained by the model.

c.

Ridge regression plot



.

Ridge regression is a technique used to reduce overfitting through shrinkage. It imposes a penalty on the size of the coefficients; it increases as the coefficients increase. The penalty is controlled by λ which determines how much shrinkage will occur. In the plot, the X axis represents the shrinkage parameter (lambda λ) and the Y axis represents the lagged values from various periods and their squared values.

The goal is to reduce the coefficients as close to 0 as possible, but not to 0 because as they approach 0, the gain from shrinking them further also approaches 0. This technique also puts more emphasis on shrinking larger coefficients than smaller ones.

When λ = 0, the coefficients are at their highest values and close to the OLS regression. As λ increases, there's a noticeable and steep decline in the coefficients of Squared_Lag_1 (0.2734) and Squared_Lag_2 (0.1399). This shows that they are more sensitive to regularization. Their values reduce significantly, converging towards zero at a faster rate than the coefficients for Lag_1 and Lag_2. The coefficients for Lag_1(-0.0261) and Lag_2 (-0.003) are already small; they are already closer to zero and do not exhibit a steep slope upward as λ increases. This indicates that the ridge regression does not heavily penalize these terms since they are not as large or as influential.

As λ continues to increase, shrinking for all coefficients slows down and begins to flatten out. This suggests that further increases in λ will not substantially change the value of the coefficients; they are approaching their limit of shrinkage.

Towards the end of the plot, as λ approaches 1, all the coefficients are approaching closer values to each other compared to their starting points. However, they do not converge to the same value, indicating that even after the penalty, the model still recognizes the level of importance of each variable.

d.

Cross-validation is a popular method for choosing an appropriate λ. It can also be used to estimate how reliable our fitted model is. The general cross-validation can have drawbacks, such as less precise estimates since we only use T/2 instead of T observation. This can hurt the model's performance. We can avoid this problem by maximizing sample size by using other methods such as the Leave-One-Out Cross Validation (LOOCV) method and k-Fold Cross-Validation

In LOOCV, the number of folds is equal to the number of observations in the dataset; for n data points, we will have n folds. For each fold, one data point is used as the validation set, and the remaining n-1 data points are used as the training set. We repeat this procedure n times to produce n estimated test errors MSE. The LOOCV estimation is the average of the n estimated test errors.

In the 10-fold CV, we randomly split the data into 10 equal folds. For each fold, the fold is used as the validation set, and the remaining 9 folds are used as the training set. We repeat this 10 times, with each fold being used once as the validation set to produce k estimated test errors MSE. The 10-fold CV estimation is the average of the k estimated test errors.

Optimal Shrinkage Parameter and its coefficients when using LOOCV

```
Optimal shrinkage parameter (lambda): 1.0
Coefficients at optimal shrinkage parameter: [ 0.00000000e+00 -4.95295913e-04  3.57083934e-04 -5.51789443e-05
  1.83073761e-04]
```

In our model, we set the LOOCV = true to use this CV method. This gives us an optimal shrinkage parameter of 1. This means that a regularization strength of 1 provides the best balance between fitting the data well and keeping the model coefficients small to avoid overfitting.

Optimal Shrinkage Parameter and its coefficients when using 10-fold CV

```
Optimal shrinkage parameter (lambda): 1.0
Coefficients at optimal shrinkage parameter: [ 0.00000000e+00 -4.95295913e-04  3.57083934e-04 -5.51789443e-05
  1.83073761e-04]
```

Alternatively, we could set LOOCV = false to use 10-fold CV. This also gives us an optimal shrinkage parameter of 1.

Problem 2

a.

Descriptive statistics

	Zonal COMED price	System load forecast	Zonal COMED load foecast
count	52416.000000	52416.000000	52416.000000
mean	30.467624	91010.088084	11371.249561
std	19.412552	16211.499741	2200.765254
min	-6.981259	57879.000000	7249.000000
25%	22.305977	79565.750000	9792.000000
50%	27.744014	88988.000000	11125.500000
75%	34.827130	100885.250000	12470.000000
max	839.302231	157156.000000	22241.000000

Interpretation:

The descriptive statistics provide an overall insight into the electrical prices in COMED area, as well as the load forecast in the entire PJM region(system load forecast) and COMED area.

count: Our sample includes 52416 data points, as our data is on an hourly basis, we could say that our sample includes 52416 hours of data.

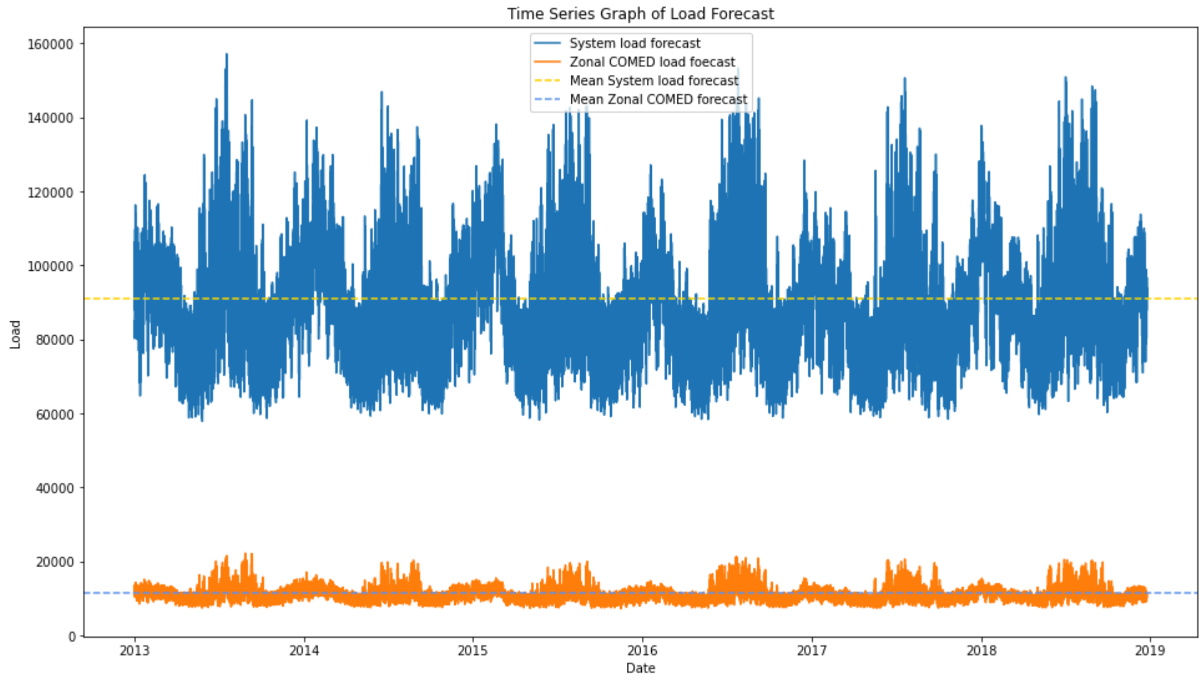
mean: Zonal COMED has a mean price of approximately 30.47 during the 6 years. The mean load of the whole region is around 91010, which is resolutely higher than the load of COMED zone solely(11371.25).

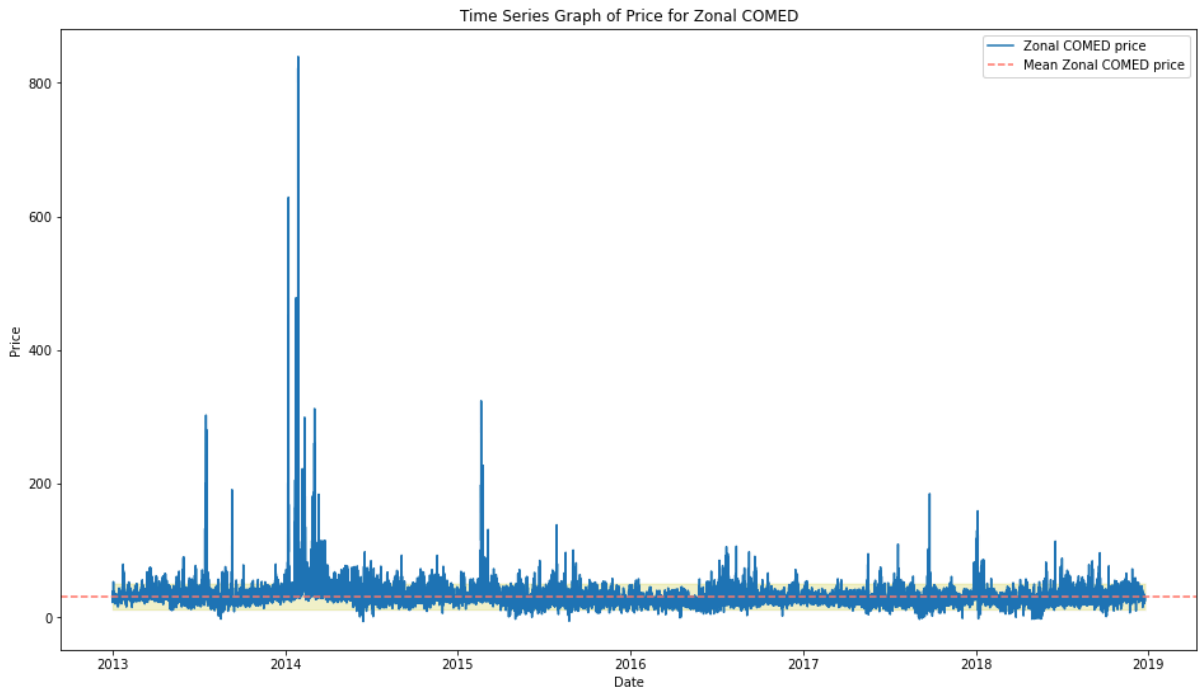
std: The standard deviation of the zonal price is smaller than the others, as the price index itself is much smaller than the load index. The zonal load is less volatile than the system load(2200.77 compared to 16211.5), same, that might be because the range of zonal load is way smaller than the whole region.

percentiles: Gives us insights into the distribution of price and loads.

max: The max value regarding zonal price is unusually large(839.30), which might be due to the data recording error.

time-series graphs





Interpretation

As our original dataset has two different types of data, we separate the graph into two, with the appropriate scale on the y-axis.

In terms of the graph of load, both system and zonal load have apparent seasonal effects. They both have regular changes according to time. For example, both of them tend to have higher loads during the mid-year.
When it comes to the price graph, we can hardly observe any seasonal effects from the graph. But we could see, that there are a few outliers, this might be because the data is not clean enough, and there might still be some data recording errors.

b.

Notation:
As our dataset is hourly, we choose 2 days periods(p=48) as our max lag.
We choose AIC as our information criterion.

Results

For Zonal COMED price:

AutoReg Model Results

Dep. Variable:	Zonal COMED price	**No. Observations:**	52416
Model:	AutoReg(47)	**Log Likelihood**	-167581.057
Method:	Conditional MLE	**S.D. of innovations**	5.936
Date:	Thu, 21 Mar 2024	**AIC**	3.564
Time:	16:09:20	**BIC**	3.572
Sample:	01-02-2013	**HQIC**	3.567
	- 12-24-2018		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	1.1602	0.068	17.102	0.000	1.027	1.293
Zonal COMED price.L1	0.9971	0.004	228.192	0.000	0.989	1.006
Zonal COMED price.L2	-0.0534	0.006	-8.660	0.000	-0.066	-0.041
Zonal COMED price.L3	-0.1115	0.006	-18.065	0.000	-0.124	-0.099
Zonal COMED price.L4	0.0446	0.006	7.193	0.000	0.032	0.057
Zonal COMED price.L5	-0.0175	0.006	-2.819	0.005	-0.030	-0.005
Zonal COMED price.L6	0.0137	0.006	2.217	0.027	0.002	0.026
Zonal COMED price.L7	0.0505	0.006	8.149	0.000	0.038	0.063
Zonal COMED price.L8	-0.0627	0.006	-10.115	0.000	-0.075	-0.051
Zonal COMED price.L9	0.0475	0.006	7.657	0.000	0.035	0.060
Zonal COMED price.L10	0.0340	0.006	5.484	0.000	0.022	0.046
Zonal COMED price.L11	0.0959	0.006	15.467	0.000	0.084	0.108
Zonal COMED price.L12	-0.0965	0.006	-15.538	0.000	-0.109	-0.084
Zonal COMED price.L13	0.1352	0.006	21.710	0.000	0.123	0.147
Zonal COMED price.L14	-0.1689	0.006	-27.026	0.000	-0.181	-0.157
Zonal COMED price.L15	-0.0383	0.006	-6.079	0.000	-0.051	-0.026
Zonal COMED price.L16	0.0436	0.006	6.926	0.000	0.031	0.056
Zonal COMED price.L17	-0.0076	0.006	-1.204	0.229	-0.020	0.005
Zonal COMED price.L18	0.0007	0.006	0.107	0.915	-0.012	0.013
Zonal COMED price.L19	-0.0005	0.006	-0.080	0.936	-0.013	0.012
Zonal COMED price.L20	-0.0125	0.006	-1.990	0.047	-0.025	-0.000
Zonal COMED price.L21	0.0524	0.006	8.343	0.000	0.040	0.065
Zonal COMED price.L22	-0.0637	0.006	-10.133	0.000	-0.076	-0.051
Zonal COMED price.L23	0.1211	0.006	19.820	0.000	0.109	0.133
Zonal COMED price.L24	0.3263	0.006	54.729	0.000	0.315	0.338
Zonal COMED price.L25	-0.3461	0.006	-56.670	0.000	-0.358	-0.334

For System load forecast:

[8]:

AutoReg Model Results

Dep. Variable:	System load forecast	**No. Observations:**	52416
Model:	AutoReg(48)	**Log Likelihood**	-421699.492
Method:	Conditional MLE	**S.D. of innovations**	760.274
Date:	Thu, 21 Mar 2024	**AIC**	13.269
Time:	16:09:26	**BIC**	13.278
Sample:	01-03-2013	**HQIC**	13.272
	- 12-24-2018		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	378.3774	27.989	13.519	0.000	323.520	433.235
System load forecast.L1	1.8526	0.004	424.237	0.000	1.844	1.861
System load forecast.L2	-1.1702	0.009	-127.230	0.000	-1.188	-1.152
System load forecast.L3	0.3463	0.011	32.906	0.000	0.326	0.367
System load forecast.L4	0.0085	0.011	0.802	0.423	-0.012	0.029
System load forecast.L5	-0.0803	0.011	-7.552	0.000	-0.101	-0.059
System load forecast.L6	0.0287	0.011	2.695	0.007	0.008	0.050
System load forecast.L7	0.0049	0.011	0.459	0.646	-0.016	0.026
System load forecast.L8	-0.0730	0.011	-6.865	0.000	-0.094	-0.052
System load forecast.L9	0.0630	0.011	5.923	0.000	0.042	0.084
System load forecast.L10	0.0731	0.011	6.876	0.000	0.052	0.094
System load forecast.L11	-0.0762	0.011	-7.162	0.000	-0.097	-0.055
System load forecast.L12	-0.0315	0.011	-2.958	0.003	-0.052	-0.011
System load forecast.L13	0.0749	0.011	7.038	0.000	0.054	0.096
System load forecast.L14	-0.0354	0.011	-3.327	0.001	-0.056	-0.015
System load forecast.L15	-0.0460	0.011	-4.321	0.000	-0.067	-0.025
System load forecast.L16	0.0117	0.011	1.103	0.270	-0.009	0.033
System load forecast.L17	0.0434	0.011	4.074	0.000	0.023	0.064
System load forecast.L18	0.0137	0.011	1.288	0.198	-0.007	0.035
System load forecast.L19	-0.0113	0.011	-1.058	0.290	-0.032	0.010
System load forecast.L20	-0.0199	0.011	-1.869	0.062	-0.041	0.001
System load forecast.L21	0.0056	0.011	0.526	0.599	-0.015	0.026
System load forecast.L22	-0.0402	0.011	-3.789	0.000	-0.061	-0.019
System load forecast.L23	0.1796	0.010	17.562	0.000	0.160	0.200
System load forecast.L24	0.5284	0.009	59.924	0.000	0.511	0.546
System load forecast.L25	-1.1986	0.009	-135.928	0.000	-1.216	-1.181
System load forecast.L26	0.6577	0.010	64.315	0.000	0.638	0.678
System load forecast.L27	-0.0719	0.011	-6.772	0.000	-0.093	-0.051
System load forecast.L28	-0.1073	0.011	-10.094	0.000	-0.128	-0.086
System load forecast.L29	0.0843	0.011	7.930	0.000	0.063	0.105
System load forecast.L30	-0.0543	0.011	-5.105	0.000	-0.075	-0.033
System load forecast.L31	0.0256	0.011	2.409	0.016	0.005	0.047
System load forecast.L32	0.0578	0.011	5.434	0.000	0.037	0.079

For Zonal COMED load forecast:

AutoReg Model Results

Dep. Variable:	Zonal COMED load foecast	**No. Observations:**	52416
Model:	AutoReg(48)	**Log Likelihood**	-329442.464
Method:	Conditional MLE	**S.D. of innovations**	130.578
Date:	Thu, 21 Mar 2024	**AIC**	9.746
Time:	16:09:32	**BIC**	9.754
Sample:	01-03-2013	**HQIC**	9.749
	- 12-24-2018		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	72.0820	4.419	16.313	0.000	63.422	80.742
Zonal COMED load foecast.L1	1.6684	0.004	382.023	0.000	1.660	1.677
Zonal COMED load foecast.L2	-0.7395	0.008	-87.063	0.000	-0.756	-0.723
Zonal COMED load foecast.L3	0.0526	0.009	5.790	0.000	0.035	0.070
Zonal COMED load foecast.L4	0.0358	0.009	3.938	0.000	0.018	0.054
Zonal COMED load foecast.L5	-0.0327	0.009	-3.602	0.000	-0.051	-0.015
Zonal COMED load foecast.L6	0.0038	0.009	0.417	0.677	-0.014	0.022
Zonal COMED load foecast.L7	-0.0114	0.009	-1.250	0.211	-0.029	0.006
Zonal COMED load foecast.L8	-0.0415	0.009	-4.563	0.000	-0.059	-0.024
Zonal COMED load foecast.L9	0.0324	0.009	3.562	0.000	0.015	0.050
Zonal COMED load foecast.L10	0.0234	0.009	2.580	0.010	0.006	0.041
Zonal COMED load foecast.L11	0.0246	0.009	2.710	0.007	0.007	0.042
Zonal COMED load foecast.L12	-0.1117	0.009	-12.304	0.000	-0.130	-0.094
Zonal COMED load foecast.L13	0.1550	0.009	17.059	0.000	0.137	0.173
Zonal COMED load foecast.L14	-0.1028	0.009	-11.279	0.000	-0.121	-0.085
Zonal COMED load foecast.L15	-0.0098	0.009	-1.069	0.285	-0.028	0.008
Zonal COMED load foecast.L16	-0.0040	0.009	-0.441	0.659	-0.022	0.014
Zonal COMED load foecast.L17	0.0629	0.009	6.892	0.000	0.045	0.081
Zonal COMED load foecast.L18	-0.0057	0.009	-0.622	0.534	-0.024	0.012
Zonal COMED load foecast.L19	0.0039	0.009	0.431	0.666	-0.014	0.022
Zonal COMED load foecast.L20	-0.0173	0.009	-1.901	0.057	-0.035	0.001
Zonal COMED load foecast.L21	-0.0257	0.009	-2.819	0.005	-0.044	-0.008
Zonal COMED load foecast.L22	-0.0148	0.009	-1.627	0.104	-0.033	0.003
Zonal COMED load foecast.L23	0.1787	0.009	19.719	0.000	0.161	0.197
Zonal COMED load foecast.L24	0.3895	0.008	46.477	0.000	0.373	0.406
Zonal COMED load foecast.L25	-0.8106	0.008	-96.721	0.000	-0.827	-0.794
Zonal COMED load foecast.L26	0.1780	0.009	19.636	0.000	0.160	0.196
Zonal COMED load foecast.L27	0.1553	0.009	17.070	0.000	0.137	0.173
Zonal COMED load foecast.L28	-0.0431	0.009	-4.723	0.000	-0.061	-0.025
Zonal COMED load foecast.L29	-0.0284	0.009	-3.107	0.002	-0.046	-0.010
Zonal COMED load foecast.L30	-0.0087	0.009	-0.958	0.338	-0.027	0.009
Zonal COMED load foecast.L31	0.0408	0.009	4.477	0.000	0.023	0.059
Zonal COMED load foecast.L32	0.0149	0.009	1.635	0.102	-0.003	0.033
Zonal COMED load foecast.L33	-0.0284	0.009	-3.117	0.002	-0.046	-0.011
Zonal COMED load foecast.L34	0.0079	0.009	0.870	0.384	-0.010	0.026
Zonal COMED load foecast.L35	0.0389	0.009	4.272	0.000	0.021	0.057
Zonal COMED load foecast.L36	-0.0684	0.009	-7.524	0.000	-0.086	-0.051
Zonal COMED load foecast.L37	0.0566	0.009	6.229	0.000	0.039	0.074

Interpretation

Our return results indicate that for Zonal COMED price, we need 47 lags, which makes our model very complex, resulting in a costly and time-consuming process. Therefore, we combined the General-to-Specific approach to examine the p-values of different lags and found that with 17 lags, the p-value is 0.229, which is greater than 0.05. So, we decided the lag of our model to be 16.

The same situation occurs with System load forecast and Zonal load forecast datasets. Both results indicate the need for 48 lags to satisfy the minimum AIC value. Combing the

General-to-Specific approach, we found that for System load forecast dataset, the p-value at lag 4 is 0.423, which is greater than 0.05, and for Zonal load forecast, the p-value at lag 6 is 0.677, also greater than 0.05. Therefore, we decided their lengths are 3 and 5 separately.

c.

We use the proper lags we conclude in part(b), as our regression model to get the forecasts. We set the hmax=24, we are forecasting the values for the coming next 24 hours from the last time point in our dataset.

Results:

For Zonal COMED price
```
array([23.67877187, 23.66970249, 24.2425586 , 25.35224288, 26.94047154,
       27.99119215, 29.18370251, 28.8301672 , 28.10722997, 27.49734325,
       27.21824324, 27.11096067, 27.19676751, 27.49579039, 27.79367389,
       28.25800606, 28.75298314, 29.23752878, 29.24476251, 29.10515699,
       28.74919161, 28.41705359, 28.34299069, 28.44804861])
```

For System load forecast
```
array([85295.82107256, 83309.20092254, 82237.14661749, 82003.65842778,
       82419.4228553 , 83263.85954633, 84333.80063122, 85467.28280166,
       86550.4723077 , 87514.4179272 , 88326.69683066, 88981.43604249,
       89489.84810196, 89872.38763961, 90152.91537556, 90354.80549581,
       90498.69197877, 90601.45797532, 90676.07577188, 90731.96043022,
       90775.57714347, 90811.12071901, 90841.15412293, 90867.14676845])
```

For Zonel load forecast
```
array([10067.35866019,  9847.11610262,  9733.32646325,  9711.18754614,
        9770.69439278,  9891.56782511, 10052.21385068, 10234.88872098,
       10425.05135458, 10611.13990523, 10784.78219989, 10940.50148489,
       11075.20621461, 11187.72159015, 11278.34803478, 11348.44436996,
       11400.06514819, 11435.66182522, 11457.84461042, 11469.2010632 ,
       11472.16613407, 11468.93640315, 11461.42074137, 11451.21991117])
```

Interpretation:

Our function returns 1-D array out-of-sample forecasts for the next 24 periods.
For Zonal COMED price: The forecasted prices for the next 24 hours range from 23.67 to 28.45.
For System load forecast: The forecasted system load for the next 24 hours ranges from 82295.82 to 90867.15.
For Zonal load forecast: The forecasted zonal load for the next 24 periods ranges from 10,067.36 to 11,451.22.
These forecasts can be used for planning and decision-making purposes related to electricity pricing and load management.

d.

Notation

To save running time, we set our rolling window as large as possible, to make sure less rollings. Same, we use the proper lags we conclude in part(b).

Results

For Zonal COMED price

```
array([[24.74944578, 24.99587822, 25.37004659, 25.54679178, 25.79685821,
        26.25907221, 26.8986477 , 27.92488398, 28.44917523, 28.4402659 ],
       [24.13165643, 24.51439446, 24.78469847, 25.1182036 , 25.7010426 ,
        26.45581243, 27.55420273, 28.18917184, 28.27565361, 27.66717885],
       [23.90329488, 24.17957633, 24.57915386, 25.22091543, 26.0609105 ,
        27.24069736, 27.92664681, 28.09135661, 27.55031253, 26.75363967],
       [23.28990273, 23.69828558, 24.43634307, 25.36221642, 26.66616556,
        27.47068867, 27.70966868, 27.28255363, 26.58407495, 26.3052116 ],
       [23.51097263, 24.25098185, 25.19723171, 26.51935595, 27.35010623,
        27.61412292, 27.20269323, 26.52823883, 26.27007114, 26.41357469],
       [26.02873291, 26.95729829, 28.08690862, 28.74596563, 28.76179675,
        28.11335905, 27.29044552, 26.8045852 , 26.7518562 , 26.91472698],
       [35.17783068, 36.2256607 , 35.99451106, 35.21642478, 33.42035281,
        31.50145672, 30.32905064, 29.22343662, 28.47895543, 28.51959941],
       [33.96897857, 33.76029848, 33.22662056, 31.64851904, 30.04468801,
        29.17316866, 28.25604412, 27.80060575, 28.09032245, 29.530064  ],
       [32.08386748, 31.56689259, 30.17036124, 28.72845654, 28.09099051,
        27.39738509, 27.08196616, 27.58639967, 29.21116657, 29.87611324],
       [31.53165482, 30.13548698, 28.69742311, 28.06338463, 27.37471315,
        27.0640022 , 27.57138718, 29.20067318, 29.86951097, 31.30744851],
       [25.79778499, 24.40293333, 24.23865701, 23.96896929, 24.26385359,
        25.3495699 , 27.34112585, 28.56551414, 30.48218118, 30.23793395],
       [23.23295337, 23.08032795, 22.9373393 , 23.34524287, 24.59430939,
        26.74185594, 28.0639551 , 30.13046547, 30.01534114, 28.40497226],
       [23.19219397, 23.04810463, 23.44392445, 24.68221715, 26.81416454,
        28.12136549, 30.17854306, 30.04909304, 28.42637736, 27.79170489],
       [20.60833919, 21.02847543, 22.53101538, 24.89869514, 26.54657069,
        28.92906262, 29.00339707, 27.69315894, 27.3277753 , 26.53909643],
       [20.98256249, 22.48555636, 24.85820056, 26.51050348, 28.89940101,
        28.97985282, 27.67344723, 27.31394357, 26.53033211, 25.89252716],
       [21.39815916, 23.78157584, 25.55151036, 28.04532469, 28.27752154,
        27.11602895, 26.84731106, 26.20296458, 25.68518066, 25.9230176 ],
       [22.3072435 , 24.09174323, 26.74495195, 27.11930039, 26.1634884 ,
        26.09119704, 25.56990895, 25.24093498, 25.64149852, 25.70783195],
       [21.89762572, 24.57267301, 25.18459494, 24.44071663, 24.67475718,
        24.44598604, 24.30026019, 24.98184575, 25.29036398, 25.58978745],
       [23.00737571, 23.63487923, 23.06050117, 23.44574662, 23.43552183,
        23.49848799, 24.31080637, 24.81981063, 25.29201535, 25.62953901],
       [22.61470975, 22.05047401, 22.54615625, 22.63444818, 22.83982473,
```

```
  23.7881291 , 24.38231786, 24.98517129, 25.43529073, 25.8267    ],
 [26.25058837, 26.70453753, 26.33810944, 26.13787482, 26.49985409,
  26.53415099, 26.78625115, 26.69846023, 26.62628387, 26.95173512],
 [29.55769689, 29.16290224, 28.65374407, 28.74017415, 28.37617194,
  28.24792834, 27.92185299, 27.48425546, 27.49477717, 28.05583562],
 [28.90785193, 28.4012119 , 28.51523727, 28.17583758, 28.08317035,
  27.79107552, 27.37476788, 27.41794909, 28.00715449, 28.66931738],
 [28.29949387, 28.41452435, 28.08613107, 28.00327715, 27.72537217,
  27.32261776, 27.37429058, 27.97652168, 28.64991042, 29.38405825],
 [26.61906466, 26.30850963, 26.42005258, 26.31552226, 26.16338764,
  26.45439825, 27.20657187, 28.10992588, 29.0422592 , 29.1016899 ],
 [24.89262806, 25.01822584, 25.06697905, 25.05154405, 25.54017092,
  26.48106784, 27.50265321, 28.61633283, 28.83205339, 28.04341937],
 [24.57326323, 24.62618494, 24.65859833, 25.18974627, 26.19234126,
  27.27295312, 28.42363877, 28.69629544, 27.95677765, 27.29277826],
 [23.97911146, 24.01791682, 24.61907726, 25.68408677, 26.85495318,
  28.09184029, 28.41850842, 27.76185925, 27.16928926, 26.65622372],
 [23.9948822 , 24.59649509, 25.66427718, 26.83775972, 28.07821556,
  28.4082067 , 27.75362229, 27.16401835, 26.65344576, 26.537638  ],
 [26.51194065, 27.56073437, 28.52676005, 29.58215931, 29.64469434,
  28.7347429 , 27.98511514, 27.22916775, 26.90194632, 27.0284549 ],
 [31.70904235, 32.63401763, 33.24032919, 32.90239469, 31.41343185,
  30.11090043, 29.00849919, 28.14989713, 27.8185121 , 28.12323086],
 [30.76889757, 31.39369948, 31.2576964 , 29.94879901, 28.90662686,
  28.05284048, 27.35002143, 27.2575491 , 27.76813356, 28.72493779],
 [30.27302537, 30.14814016, 28.96064983, 28.02676791, 27.32946109,
  26.77605266, 26.77721126, 27.43135893, 28.5118443 , 29.07349824],
 [27.50575007, 26.34447478, 25.69681091, 25.25478708, 25.07032997,
  25.42373875, 26.29860678, 27.71757353, 28.57085177, 29.56286832],
 [25.12096765, 24.48543688, 24.17590791, 24.10961501, 24.63384626,
  25.67180506, 27.19295638, 28.20295239, 29.33000215, 28.92149934],
 [23.95789922, 23.65360234, 23.64445389, 24.21965398, 25.33127789,
  26.92275523, 27.9768121 , 29.17143329, 28.82114801, 28.10103911]])
```

For System load forecast

```
array([[ 87148.51459422,  86654.18399092,  86504.6110119 ,
         86652.25353109,  87014.93490122,  87505.16242617,
         88046.61900699,  88581.40062928,  89071.21244519,
         89495.10792813],
       [ 86695.54265129,  86565.42817854,  86725.74578581,
         87093.7748494 ,  87583.04621314,  88118.93805195,
         88645.33629345,  89125.5063248 ,  89539.69020237,
         89881.25662269],
       [ 85980.35458421,  85865.57360256,  86054.57846884,
         86468.55518518,  87018.35533913,  87623.84167483,
         88222.92950343,  88773.77394448,  89252.92609074,
         89651.55643908],
       [ 85983.32759343,  86227.67425306,  86677.63894059,
         87242.53798033,  87845.16432541,  88428.27420739,
         88955.129268  ,  89406.72867952,  89777.62894177,
```

```
 90071.5865339 ],
[ 88253.62601471,  89656.10250998,  90840.78258544,
 91703.97108571,  92238.74723198,  92491.57612875,
 92531.24924035,  92428.81008074,  92246.2510352 ,
 92031.57836504],
[ 95173.35809704,  98952.11047939, 101503.28468853,
 102747.7912766 , 102869.11896648, 102162.63995148,
 100938.48487875,  99466.95015461,  97954.6432831 ,
 96540.7229425 ],
[103962.32912784, 108869.36956906, 111646.94782145,
 112412.93469617, 111587.0940473 , 109685.31534849,
 107195.03715956, 104512.06536707, 101919.61987952,
 99593.3494755 ],
[ 97392.95013771,  94774.79432828,  92029.94840452,
 89728.23805252,  88100.44572813,  87162.61773504,
 86813.17585562,  86902.01890209,  87274.89112384,
 87798.13766763],
[ 96777.0137108 ,  94973.53850688,  93284.34420672,
 91914.02866851,  90928.38244982,  90308.07591273,
 89989.79163049,  89894.87880611,  89947.24173176,
 90082.85584355],
[ 95906.29481852,  94655.60959614,  93570.60710091,
 92704.89724804,  92062.32842412,  91617.89309416,
 91333.35795845,  91167.83681646,  91084.11074808,
 91051.71785558],
[ 91104.99614432,  88350.77227009,  86398.98559836,
 85299.85115389,  84940.13704589,  85135.78672294,
 85692.06718221,  86437.73504244,  87240.29121345,
 88008.77354033],
[ 86116.10551961,  83113.73629723,  81331.03916377,
 80683.97444117,  80932.96358652,  81791.49453798,
 82991.49121732,  84316.09599953,  85610.12054583,
 86776.83325433],
[ 81840.6896845 ,  79459.48389345,  78422.98929534,
 78508.26293813,  79397.16787694,  80769.34640352,
 82352.7781233 ,  83944.21632847,  85410.33880333,
 86678.192582  ],
[ 84444.01301923,  85750.86903727,  87360.80812541,
 88890.57884982,  90143.6758411 ,  91052.80817098,
 91630.73133827,  91932.2988876 ,  92027.82390432,
 91986.28764748],
[ 79946.09691409,  78827.04810173,  78581.1340891 ,
 79087.6611863 ,  80135.13160219,  81497.93227497,
 82979.42402988,  84430.82165655,  85754.29003216,
 86897.0226671 ],
[ 80448.45344553,  80964.60388622,  81966.80503715,
 83222.44319482,  84546.24132724,  85808.18066405,
 86929.7424944 ,  87874.29306228,  88635.61260855,
 89227.02344567],
[ 80473.80219512,  81245.42894486,  82351.20692402,
```

83612.22958467, 84886.2396661 , 86074.52500771,
87119.16106551, 87995.38945607, 88702.42544937,
89254.71720617],
[81827.46529572, 83206.88886106, 84645.99327819,
85994.9213434 , 87169.41484593, 88135.44044711,
88893.44588389, 89464.60586625, 89880.08866235,
90173.49941265],
[85236.95223374, 87630.44964812, 89600.36975494,
91035.92824797, 91953.51370362, 92436.93217639,
92595.31328265, 92536.45678686, 92352.31862563,
92113.17567411],
[89196.40923507, 91902.58111849, 93817.25382023,
94936.35179592, 95382.56789183, 95329.3217814 ,
94952.24778828, 94402.40229563, 93795.06310851,
93208.63461086],
[91886.01342943, 93792.9159958 , 94906.98826947,
95351.14127681, 95298.37955923, 94923.64988332,
94377.28202535, 93773.9199028 , 93191.48479515,
92675.79995935],
[92815.36220836, 93469.86034885, 93614.98716074,
93436.51949307, 93085.1225553 , 92670.98461931,
92266.3952185 , 91912.37453681, 91626.64810766,
91411.2837286],
[94404.06057069, 94988.34848673, 95095.57716172,
94864.20112884, 94427.61565784, 93896.47999328,
93352.30275773, 92848.08792646, 92412.76794193,
92057.10406227],
[94140.40421572, 93848.96921905, 93358.20627976,
92812.61287539, 92301.79787346, 91872.44954207,
91540.8220099 , 91303.82257173, 91147.82230533,
91055.00886288],
[93206.03598869, 92412.94902846, 91670.58574502,
91076.96979681, 90662.83343067, 90418.04055436,
90311.64925608, 90305.76496732, 90364.09739977,
90456.4368207],
[92224.95937993, 91394.20080038, 90743.05371234,
90304.71008312, 90064.36798488, 89983.369079 ,
90015.67596046, 90117.90338145, 90254.43891741,
90399.16254493],
[91479.27377122, 90868.08418423, 90455.69803147,
90226.20896514, 90143.08148828, 90163.78106093,
90248.61603031, 90365.19046559, 90489.83220157,
90607.15858847],
[91279.0092748 , 91059.71658988, 90955.7668757 ,
90925.27486522, 90935.9369286 , 90964.95738643,
90997.76127758, 91026.19880579, 91046.70330133,
91058.66686569],
[91491.34942915, 91590.21019285, 91691.57311224,
91757.52350436, 91776.01567803, 91750.22107883,
91690.70496922, 91610.20011549, 91520.50512317,

```
       91430.96346902],
      [ 95264.764403  ,  97093.84662547,  98284.18549161,
        98775.68471471,  98662.7338542 ,  98106.73363518,
        97279.58015556,  96331.75050061,  95378.18931729,
        94495.85452558],
      [100355.88575571, 103080.07614584, 104569.93790797,
       104877.17506775, 104244.10322328, 102976.45258765,
       101365.94326107,  99650.60397388,  98001.13960364,
        96523.39562295],
      [ 94201.96895642,  91517.66578697,  89108.49305042,
        87332.81752746,  86275.97031083,  85865.25418403,
        85954.16491674,  86378.36012849,  86988.5426191 ,
        87666.20398093],
      [ 93166.31649167,  91532.2112643 ,  90260.8481583 ,
        89416.02518367,  88965.99696927,  88831.95374285,
        88920.97685977,  89145.99550114,  89435.88204477,
        89738.74603754],
      [ 93475.36078338,  93117.55191973,  92867.20876677,
        92667.20180888,  92486.97147221,  92313.36459624,
        92143.47351972,  91979.49960062,  91825.39118296,
        91684.88650034],
      [ 91881.98263572,  91050.74778282,  90472.73357871,
        90133.52056426,  89989.27559939,  89986.3724044 ,
        90073.49763632,  90207.97382032,  90358.03432888,
        90502.55578133],
      [ 89010.74236225,  87473.63114028,  86510.29048779,
        86103.53882931,  86149.09968601,  86511.92377028,
        87060.97940176,  87687.51774493,  88311.7930011 ,
        88882.53930802]])
```

For ZoneI load forecast

```
array([[10390.09495469, 10500.48841681, 10615.39845887, 10729.60997639,
        10841.86539377, 10948.67410444, 11046.49810571, 11133.24925678,
        11207.85090753, 11269.94890817],
       [10357.20437372, 10418.03704634, 10496.62383007, 10589.44210475,
        10692.82492189, 10800.33199122, 10906.06944919, 11005.74057626,
        11096.25708205, 11175.63151535],
       [10272.38235969, 10295.99429757, 10352.59392883, 10436.2119488 ,
        10540.22911842, 10655.8018563 , 10774.76532613, 10890.75834395,
        10999.01443429, 11096.25151923],
       [10246.05811756, 10283.80688014, 10355.00224969, 10452.23606436,
        10566.60390585, 10688.93032871, 10811.52863223, 10928.51075638,
        11035.64222497, 11130.22437089],
       [10296.50041947, 10372.49483702, 10472.89795456, 10589.00466344,
        10711.65387005, 10833.41455017, 10948.73448223, 11053.66345013,
        11145.74429466, 11223.81862373],
       [10653.31579458, 10859.69757825, 11045.6099511 , 11206.33464648,
        11334.78934383, 11431.11118867, 11498.80979732, 11541.73997207,
        11564.1024311 , 11570.28157319],
       [11973.80415135, 12580.20891795, 13017.92910932, 13297.50834695,
```

```
 13420.46010535, 13412.85658828, 13308.14327082, 13135.55990268,
 12920.75088881, 12685.76661081],
[12046.73510649, 12283.11930561, 12430.07433777, 12480.67319449,
 12460.33040132, 12391.68660705, 12289.81123447, 12168.35861488,
 12039.19872439, 11911.19831596],
[11566.1433365 , 11442.54760758, 11314.95113528, 11197.40297362,
 11111.66113382, 11058.27440636, 11031.85092951, 11028.14790981,
 11042.353243  , 11069.20287547],
[11593.86775977, 11523.36826612, 11443.42259492, 11378.19207291,
 11328.41160655, 11291.75335198, 11267.9945669 , 11255.72371086,
 11252.56323337, 11256.27860814],
[11179.82073331, 10970.22606133, 10819.59780683, 10723.23010627,
 10678.37364481, 10677.84717323, 10711.11179474, 10768.06811034,
 10839.92597425, 10919.18441912],
[10816.84746205, 10608.33494485, 10473.83644516, 10408.17675851,
 10403.98762203, 10447.62285679, 10524.90769468, 10623.60428054,
 10733.28448355, 10845.45516747],
[10104.66209564,  9780.07895007,  9589.20222899,  9516.69366418,
  9548.29299647,  9659.62423063,  9825.06857875, 10022.87744092,
 10234.94629689, 10446.756228  ],
[10145.96443774, 10093.13756483, 10111.54099823, 10192.70990688,
 10312.71977371, 10453.3677322 , 10602.6294811 , 10750.62649032,
 10889.823153  , 11015.37041868],
[ 9373.55718366,  9120.42248355,  9022.71888499,  9045.13844064,
  9168.60226249,  9366.51101096,  9609.87132195,  9874.97708702,
 10143.24934495, 10400.35343464],
[ 9392.54210457,  9397.51599285,  9487.59360845,  9647.98599054,
  9852.41995905, 10077.41494869, 10306.49061062, 10527.17982543,
 10730.3386453 , 10910.24546223],
[ 9278.90875564,  9324.22158119,  9455.11937882,  9643.45893845,
  9865.61509339, 10102.70468666, 10339.11060368, 10563.02352513,
 10766.45819455, 10944.61414136],
[ 9332.95713272,  9467.15080328,  9657.6617424 ,  9881.00328939,
 10118.30363924, 10354.12296108, 10576.88388814, 10778.79719292,
 10955.22867808, 11104.08710834],
[ 9544.11831516,  9763.67694971, 10006.15591874, 10253.8989513 ,
 10491.56122017, 10709.12475039, 10900.84431788, 11063.81503508,
 11197.4140278 , 11302.78615861],
[ 9921.44816507, 10223.47106392, 10510.442642  , 10769.51098829,
 10990.85377214, 11171.92204847, 11314.0005808 , 11420.01124695,
 11494.10869125, 11541.16132008],
[10363.45368246, 10703.25544591, 10997.12868994, 11237.46203683,
 11421.87994331, 11554.50319121, 11641.97106321, 11691.58353229,
 11710.87978169, 11707.11479369],
[10343.53090601, 10501.63975453, 10652.52914382, 10788.14343205,
 10912.15803598, 11023.92470608, 11121.19099193, 11203.41337474,
 11270.98334471, 11324.63794155],
[10575.24697732, 10753.91625492, 10907.83210234, 11041.83235794,
 11155.35964628, 11247.65253764, 11320.12211157, 11374.81372921,
 11413.86885361, 11439.62752941],
```

```
[10759.79514055, 10915.92943692, 11051.39106596, 11165.71543969,
 11258.14817628, 11330.21934441, 11384.13038386, 11422.15482768,
 11446.74490373, 11460.400163  ],
[10739.12447291, 10807.85783411, 10878.2191212 , 10946.66225761,
 11014.49950715, 11080.35113903, 11141.796031  , 11197.314083  ,
 11246.03158872, 11287.46237085],
[10694.41218415, 10721.95678397, 10762.18928591, 10814.62923395,
 10877.75709359, 10946.85478752, 11017.39070945, 11085.94300332,
 11149.86168352, 11207.2643117 ],
[10586.54230817, 10575.66366666, 10594.42352037, 10639.16257563,
 10704.99606337, 10784.65177985, 10871.11466775, 10958.69434644,
 11042.92647122, 11120.54335776],
[10574.502512  , 10592.81687966, 10637.25395401, 10702.91158416,
 10782.5176164 , 10869.03521012, 10956.74481275, 11041.15736394,
 11118.98403419, 11188.10109678],
[10534.0923812 , 10556.37673846, 10607.45038605, 10679.11212612,
 10764.25050439, 10855.95460237, 10948.17447659, 11036.30058635,
 11117.08784547, 11188.46549857],
[10771.85498602, 10904.25303737, 11029.49776033, 11143.87913996,
 11240.75246658, 11318.43138667, 11378.02595702, 11421.1308508 ,
 11449.78542973, 11466.40021348],
[11720.28760689, 12153.51042368, 12470.79650312, 12678.38434844,
 12775.6004689 , 12780.08294738, 12715.08989996, 12600.99256018,
 12455.76947407, 12294.99701403],
[11606.24031363, 11716.99162776, 11788.50504262, 11811.46739647,
 11802.84195006, 11774.80293229, 11733.1931768 , 11683.69807398,
 11631.45411698, 11580.05719557],
[10947.74375954, 10728.9731869 , 10560.70026529, 10447.72980901,
 10401.28770661, 10411.63665792, 10464.03822308, 10546.34925413,
 10647.49259189, 10757.52353067],
[10990.03937915, 10920.27619791, 10872.19797267, 10861.16320235,
 10877.75511928, 10912.52180251, 10960.25198315, 11015.73263132,
 11073.99808912, 11131.3336269 ],
[10862.53544263, 10792.66526332, 10767.27108837, 10776.02434954,
 10809.40285577, 10861.02634831, 10924.14905775, 10992.50853122,
 11061.28877084, 11126.94484112],
[10663.40049017, 10589.22974631, 10565.85052995, 10581.69770293,
 10630.23270516, 10702.09127583, 10787.57825093, 10878.97365303,
 10970.26666119, 11056.87591211]])
```

Interpretation

In our practice, to save time, we set the rolling window T0 = 52380, as our sample size is 52416, we are making 36 times of rollings here. We are estimating 10 hours ahead(h = 10). Our results are (36, 10) formed 2-D arrays.

e.

The following questions are in the context of 'price', thus we use df[' Zonal COMED price'] to finish the tests.
In terms of 300 observations in the estimation window, to save running time, we use the last 400 observations to do the rolling.

Result

```
[0.1850399603149963,
 0.09798012873135717,
 0.04873853866681943,
 0.022865670137094948,
 0.007671779418050209,
 -0.0019916038635062207,
 -0.008022380416649823,
 -0.010546892451447657,
 -0.012080274545006982,
 -0.01302684196581825,
 -0.01366756768518356,
 -0.013071622683981444,
 -0.012056079935665504,
 -0.0109540375909826,
 -0.01090717928223757,
 -0.01286363838160896,
 -0.016428053774044687,
 -0.022266592927657217,
 -0.027742375708815877,
 -0.029432093158253344,
 -0.028435794361592617,
 -0.026113653539075445,
 -0.022891700314181623,
 -0.019045976669314308]
```

Interpretation

When computing the DM statistics, I calculated d as '`d = MSFE_ar3 - MSFE_ar6`', Thus. if the DM statistic is positive, which means on average, the forecast errors of AR(3) forecasts are bigger than those of the AR(6), we could conclude that the AR(6) forecasts are better, and vice versa.
Since in the large sample, the DM statistics can be seen as a standard normal distribution, and at a 90% confidence level, the critical score for a two-tailed test is approximately 1.645. This means that if the absolute value of our DM statistic is greater than 1.645, we can conclude that the result is statistically significant at the 90% confidence level.

In terms of our results, there are no forecast horizons where the AR(6) forecasts perform significantly better than the AR(3) forecasts at the 90% confidence level. Although DM statistics of forecast horizons 1 to 5 are positive, they are smaller than 1.645. This indicates

that at the 90% confidence level, we failed to reject the null hypothesis, and we could not say AR(6) forecasts perform significantly better than the AR(3) forecasts.
Meanwhile, for those negative DM statistics (horizons from 6 to 24), they are larger than -1.645. which indicates that at the 90% confidence level, we failed to reject the null hypothesis, and we could not say AR(3) forecasts perform significantly better than AR(6) forecasts.

f.

Result

We rotated the data frame required in the description, to make it more readable.

	MDM	p-value
1	0.184103	0.853933
2	0.096479	0.923140
3	0.047482	0.962129
4	0.022032	0.982423
5	0.007308	0.994169
6	-0.001875	0.998504
7	-0.007462	0.994047
8	-0.009687	0.992271
9	-0.010952	0.991262
10	-0.011652	0.990704
11	-0.012055	0.990382
12	-0.011363	0.990934
13	-0.010324	0.991763
14	-0.009234	0.992632
15	-0.009046	0.992782
16	-0.010490	0.991631
17	-0.013162	0.989499
18	-0.017514	0.986027
19	-0.021405	0.982922
20	-0.022257	0.982243
21	-0.021056	0.983201
22	-0.018915	0.984909
23	-0.016202	0.987073
24	-0.013156	0.989503

Interpretation

The null hypothesis is that the AR(3) forecast encompasses the AR(6) forecasts. All of our modified DM statistics have large p-values. This indicates that at any confidence level, including 99%, there is insufficient evidence to reject the null hypothesis. Thus we could conclude that forecasts AR(3) encompass the AR(6) forecasts.

Problem 3

a.

Results

Lag	AIC	BIC	HQIC
1	30.757071	30.759101	30.757706
2	28.132076	28.135629	28.133187
3	27.801695	27.806770	27.803281
4	27.779157	27.785754	27.781219
5	27.744723	27.752844	27.747261
6	27.704536	27.714180	27.707550
7	27.680788	27.691954	27.684278
8	27.656220	27.668909	27.660186
9	27.607674	27.621886	27.612116
10	27.569541	27.585276	27.574459
11	27.525311	27.542569	27.530705
12	27.485807	27.504588	27.491677
13	27.462122	27.482427	27.468469
14	27.265252	27.287080	27.272075
15	27.155159	27.178509	27.162457
16	27.108660	27.133535	27.116435
17	27.072016	27.098413	27.080267
18	27.060058	27.087979	27.068785
19	27.053989	27.083434	27.063193
20	27.049269	27.080236	27.058948

Interpretation

With the function calculate_ic we calculated the OPTIMAL LENGTH FOR A VAR FOR THE VARIABLES for the given dataset. From the results we can see that the IC results are decreasing as the number of lags keeps going up, which suggests that more historical data does help the model to better capture the dynamics between the data, thus improving the model's fit to the data.

At the same time, it is noted that the BIC decreases more slowly compared to the AIC, which suggests that while a higher number of lags may reduce the value of the information criterion, it also increases the risk of model overfitting. At a lag number of 19, the HQIC stops decreasing and the IC value gradually stabilizes, suggesting that 19 lags may be a better number of lags.

b.

Results

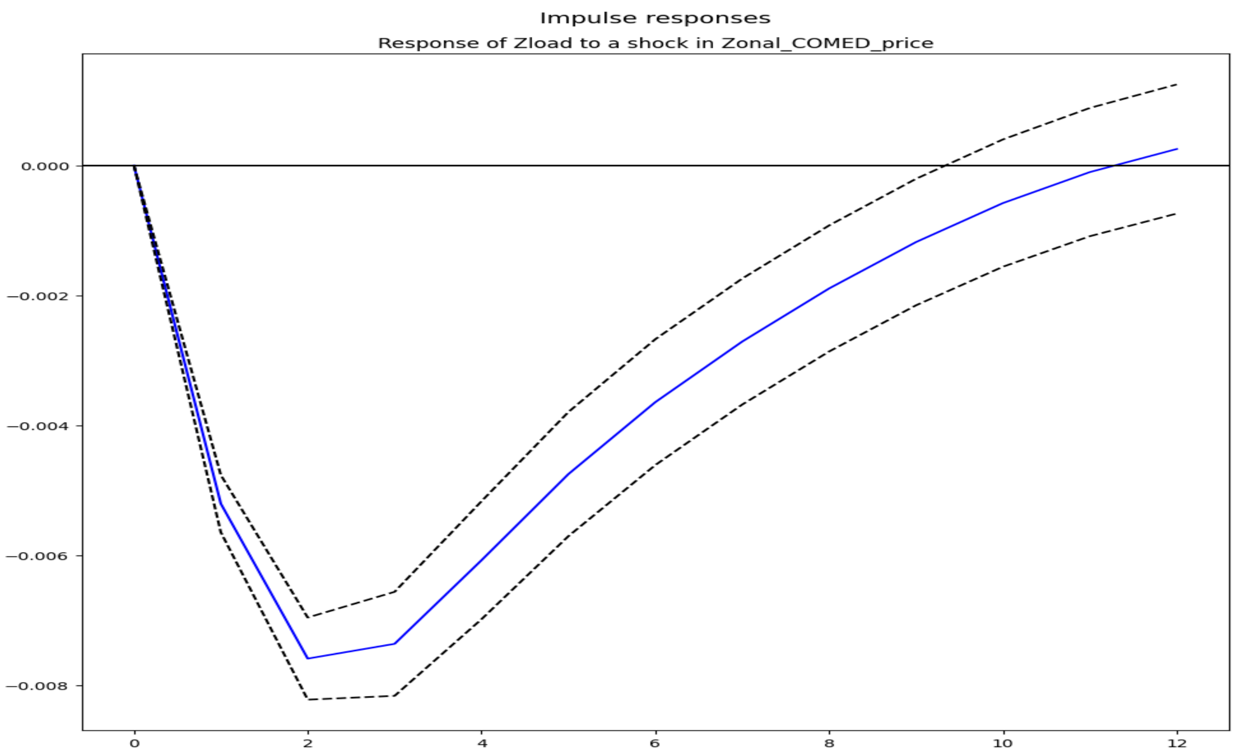
```
                          Zonal_COMED_price  System_load_forecast  \
Zonal_COMED_price              16710.417925            151.268283
System_load_forecast              19.168456         100634.209447
Zonal_COMED_load_forecast         12.695566           2555.598165

                          Zonal_COMED_load_forecast
Zonal_COMED_price                         29.201306
System_load_forecast                     324.049866
Zonal_COMED_load_forecast             148601.682234
```

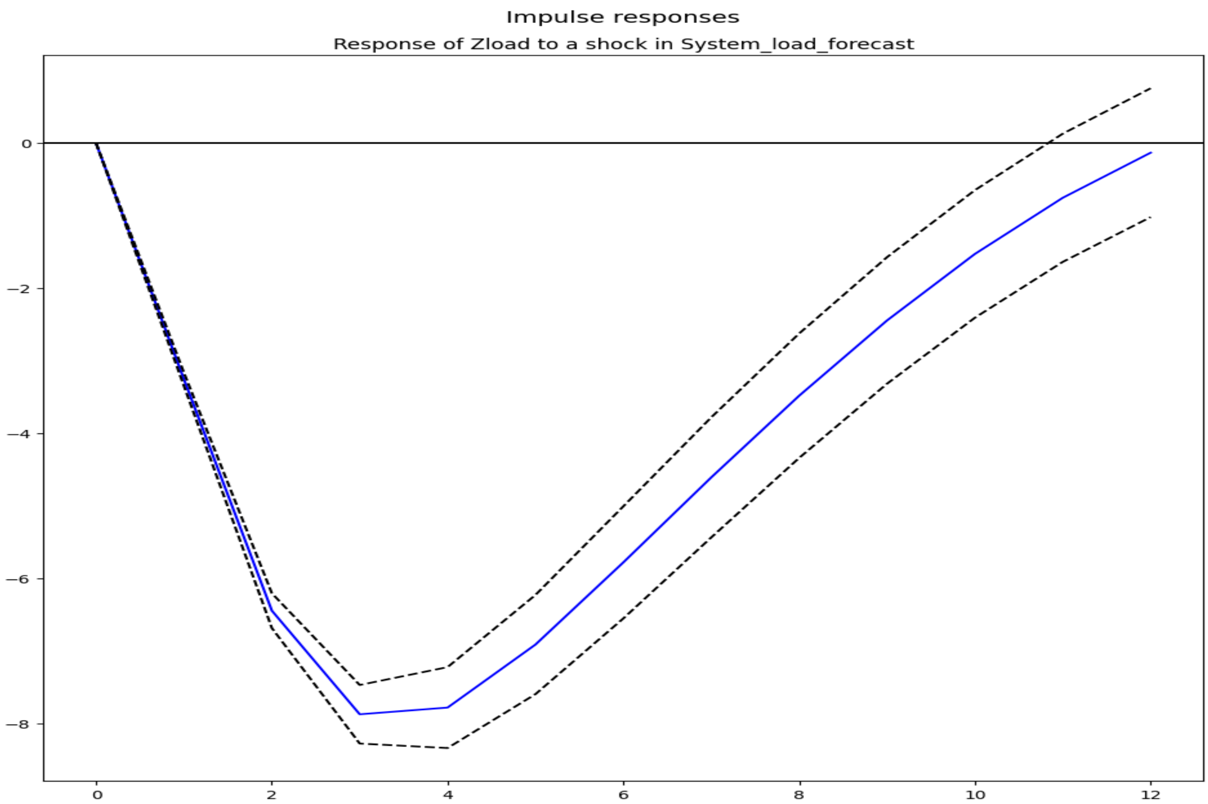
Interpretation

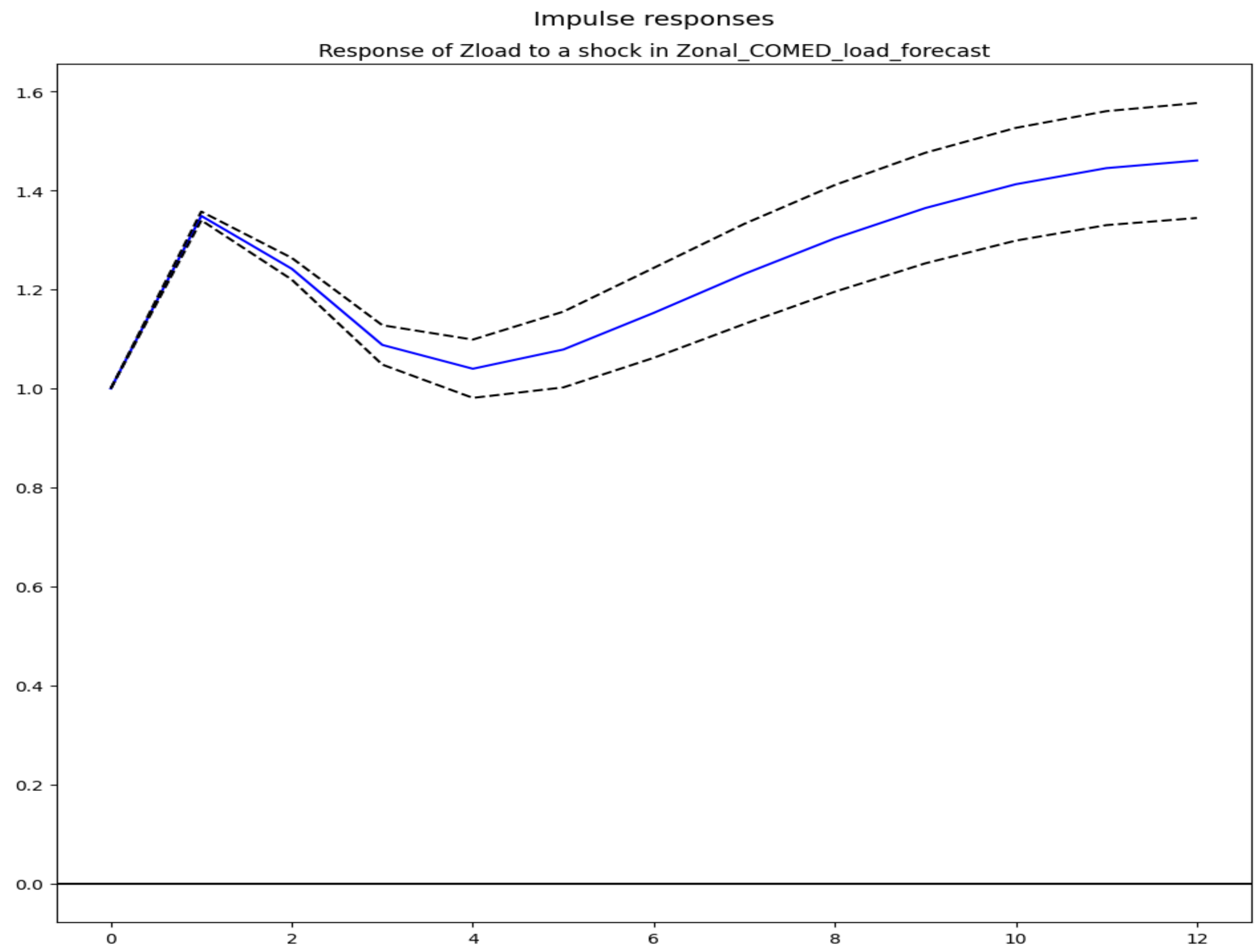
We ran the Granger Causality Test on the data using granger_causality_tests and obtained an N x N F-statistic. Where we focus on the part of each variable that interacts with the other variables, the F-statistic for Zonal_COMED_price vs. System_load_forecast is 151.268283, indicating that "System_load_forecast" significantly Granger Causality Tests "Zonal_COMED_price" to "Zonal_COMED_price". to "Zonal_COMED_price"; the F-statistic of "Zonal_COMED_load_forecast" to "System_load_forecast" is 151.268283 indicating that "System_load_forecast" is significantly Granger-induced. The F-statistic of "Zonal_COMED_load_forecast" is 324.049866, which indicates that "Zonal_COMED_load_forecast" significantly Granger causes "System_load_forecast The F-statistic of "System_load_forecast" to "Zonal_COMED_load_forecast" is 2555.598165, which indicates that "System_load_forecast" significantly Granger leads to "System_load_forecast". indicating that "System_load_forecast" significantly Granger causes "Zonal_COMED_load_forecast".
In summary, each pair of variables shows significant Granger causality. This means that, based on statistical criteria, the past value of each variable is predictive of the future value of the other variable.

c.

Results


Impulse responses
Response of Zload to a shock in Zonal_COMED_price
0.000
−0.002
−0.004
−0.006
−0.008
0
2
4
6
8
10
12


Impulse responses
Response of Zload to a shock in System_load_forecast
0
−2
−4
−6
−8
0
2
4
6
8
10
12



Interpretation

We plot the corresponding plots of each variable against Zload shocks over the next 12 cycles. We can see from the plots how each variable changes over time in response to the Zload shock.

The response to the "Zonal_COMED_load_forecast" shock starts off with a rapid decline and tends to back to zero within 12 periods. This means that the impact of the shock is not long-lasting and doesn't remain significant at the end of the 12 periods of the simulation.

The second figure shows the response of "System_load_forecast" to the "Zonal_COMED_load_forecast" shock. In this figure, a similar image can be observed that the response curve starts off with a rapid decrease, and then after about 4 periods, the decrease rate starts to decay towards 0, indicating a relatively short-lived response.

The response of Zonal_COMED_load_forecast to its own shock rises rapidly at first and then stabilizes at a high level, which indicates a longer-lived response.

So the response of "System_load_forecast" to the "Zonal_COMED_load_forecast" shock and "System_load_forecast" to the "Zonal_COMED_load_forecast" shock are relatively short-lived positive shocks.