Group Project Report (1st Draft)

Members: Minxuan Chen, Jie Min, Guangyu Xing, Xuan Yang

Tutor: Refik Soyer

Time Series Forecasting

The George Washington University

April 16th, 2018

1. Introduction

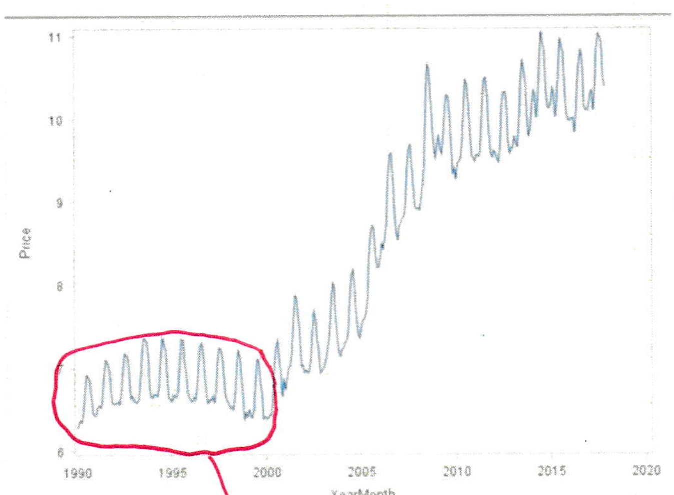
As the growing volumes of energy traded nowadays, power exchanges have taken a key role. As a result, forecasting for power exchanges has become increasingly important for market participants both for short term bidding and for long term budgeting. Electricity is one source of energy and an essential part of our daily life as well. The main goal of our analysis is to predict future values of electricity.

We have 335 observations in our dataset starts from January 1990 to November 2017. There are three variables that we used in this analysis: monthly electricity production, CPI and natural gas imports. The cost of production affects production, and then the prices of electricity. Electricity prices, as one of the fastest growing sub groups of the consumer price index (CPI) recently, is largely increased and is expected to increase continually in the next few years. Also, natural gas is one of the important part of the source of electricity. It accounted for 21% of the source of generation in 2008. We set 36 samples as our hold-out samples.

Because of deregulation reforms, electricity prices forecasting becomes very important to participants in electricity markets. On the other hand, Modeling the price of electricity is a challenging task, considering its features: electricity cannot be stored, production needs to be balanced regularly against demand, and we have to consider its seasonality, volatility and inelasticity.

2. Univariate Time-series models

Figure 2.1



O.k.

I would put a dummy to reflect changes that is, before 2000 after 2000

Pl. number pages

According to time series data view function of SAS, we graph the trend of average retail price of electricity by month. As is shown in the graph, the whole year trend of retail price at first remain stable before 2000. After 2000, the whole year trend of average retail price shoots up and keeps increasing until now.

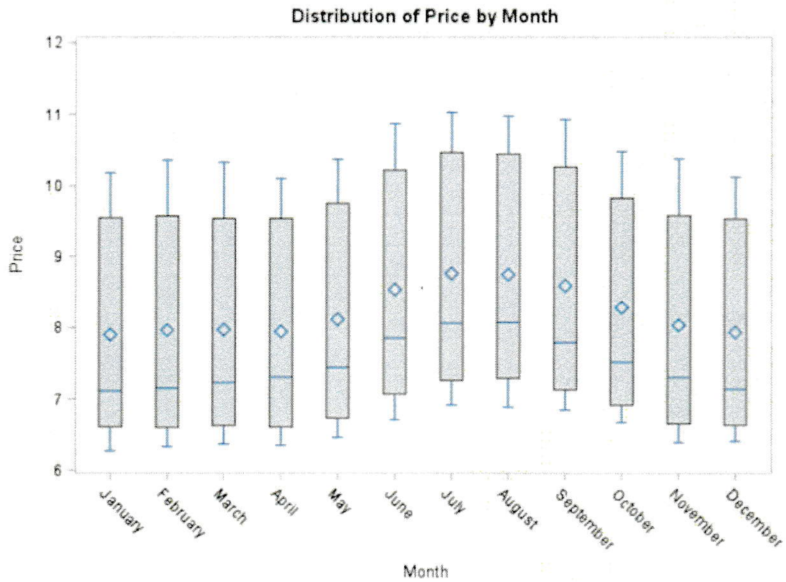


Figure 2.2

According to seasonality analysis, there is a remarkable seasonal trend run through whole year. We can see that there is a peak appearing in summer between June and September. Average price gently decreases after September till December, and increases slowly from January to May.

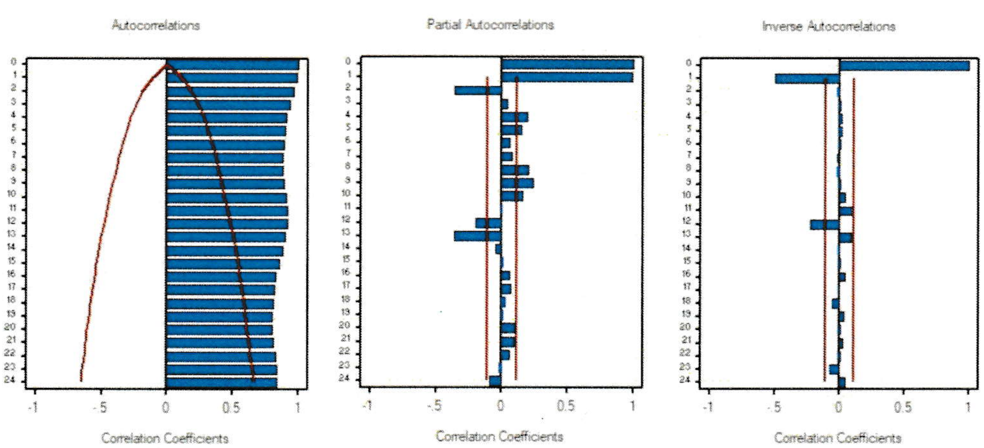


Figure 2.3

From above ACF plot we can find out that it is not stationary. (as suggested by Figures 2.1 and 2.2)
There are total of 335 observations in our data, and we chose 36 observations as our hold-out samples in all models.

2.1 Deterministic Time Series Models (Seasonal Dummies &Trend) & Error model

2.1.1 Linear Trend

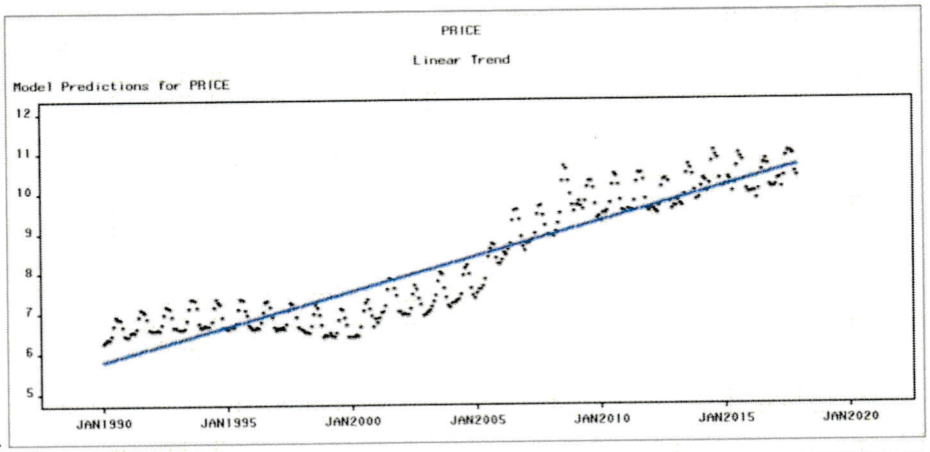


Figure 2.4

no need for this model. Pl use together with seasonal dummies

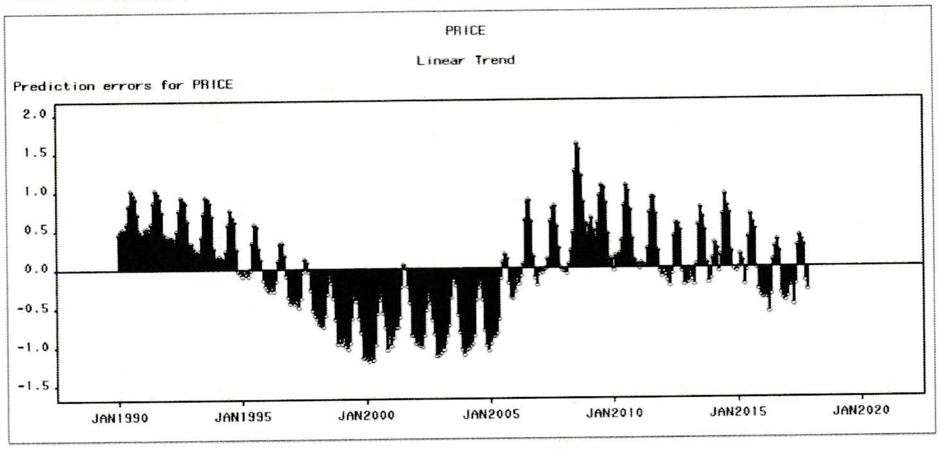


Figure 2.5

no need

After we fit a linear trend with average retail price of electricity, we find that it is not suitable due to large prediction errors between predicted values with real price data.

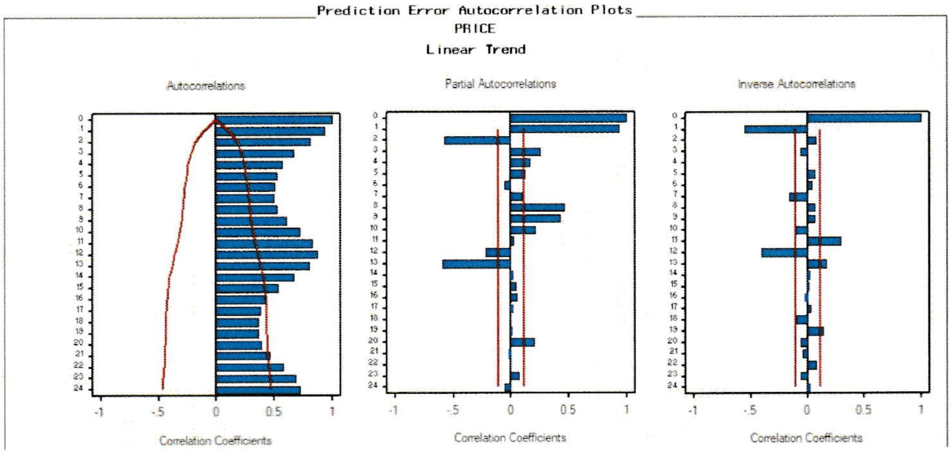


Figure 2.6

The above graph also demonstrates that linear trend model doesn't fit very well. According to sample autocorrelation function of price, we can conclude that the series is not stationary.

2. 1.2 Linear Trend + Seasonal Dummy

By observing time series data, we assume that average retail price of electricity has a prominent seasonality. Thus, we add seasonal dummy based on linear model and get outputs below.

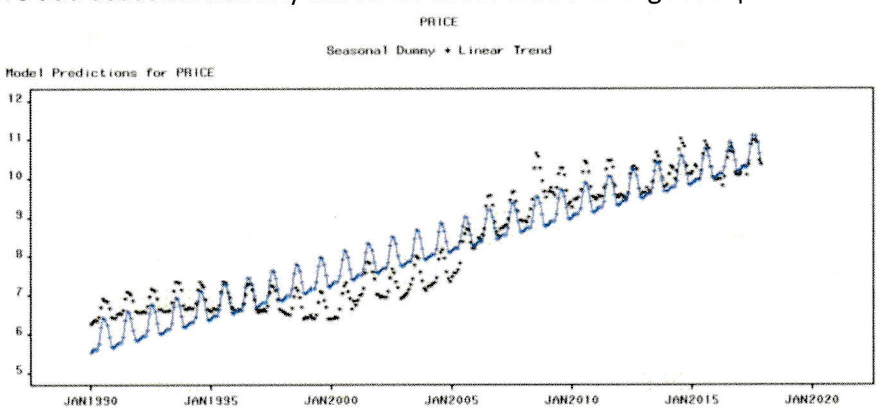


Figure 2.7

See my comment on page 1

We could find that price data got fit better after linear trend added.

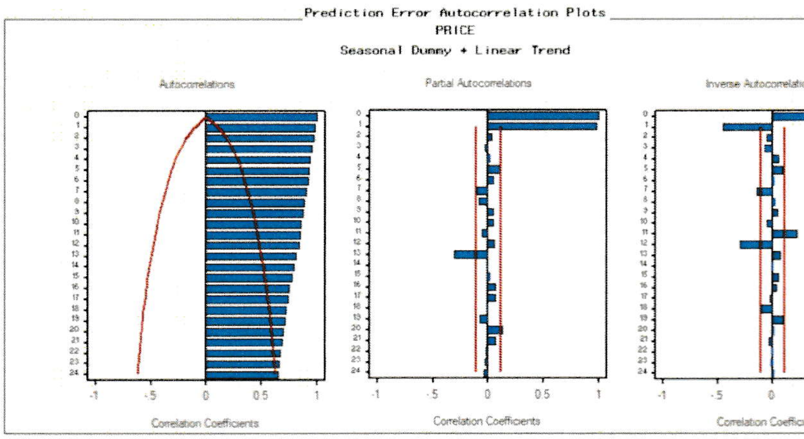


Figure 2.8

Parameter Estimates
PRICE: PRICE
Linear Trend + Seasonal Dummies

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	5.49799	0.1251	43.9395	<.0001
Linear Trend	0.01449	0.000369	39.3079	<.0001
Seasonal Dummy 1	0.03679	0.1572	0.2341	0.8170
Seasonal Dummy 2	0.08070	0.1571	0.5135	0.6125
Seasonal Dummy 3	0.07661	0.1571	0.4875	0.6305
Seasonal Dummy 4	0.05932	0.1571	0.3775	0.7093
Seasonal Dummy 5	0.20882	0.1571	1.3289	0.1969
Seasonal Dummy 6	0.60913	0.1571	3.8763	0.0008
Seasonal Dummy 7	0.82224	0.1571	5.2325	<.0001
Seasonal Dummy 8	0.79015	0.1571	5.0282	<.0001
Seasonal Dummy 9	0.61246	0.1571	3.8974	0.0007
Seasonal Dummy 10	0.31197	0.1571	1.9852	0.0592
Seasonal Dummy 11	0.03628	0.1572	0.2308	0.8195
Model Variance (sigma squared)	0.30237	.	.	.

Fit Range: JAN1990 to NOV2014

Figure 2.9

2.1.3 Linear Trend + Seasonal Dummy + Error Model

Obviously not stationary

By observing the ACF and PACF of the linear trend and seasonal dummy model, we can see that the ACF drops slowly but PACF chops off at lag 2. So, we can use an AR(1) model as the error model to further fit the series. The fitted series and residue plots are shown below.

I bet your AR component is close to 1 ⇒ like differencing

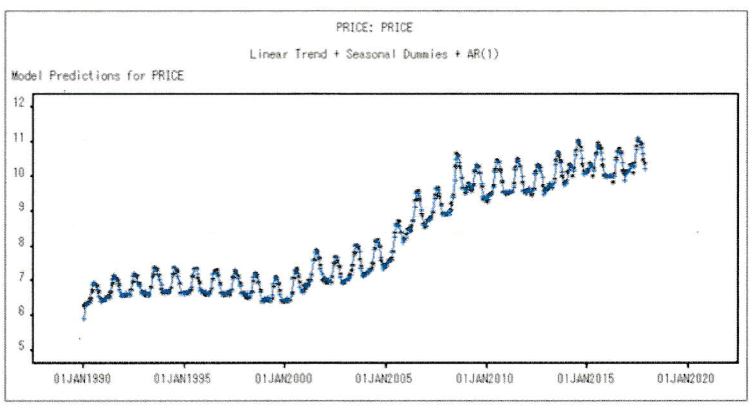


Figure 2.10

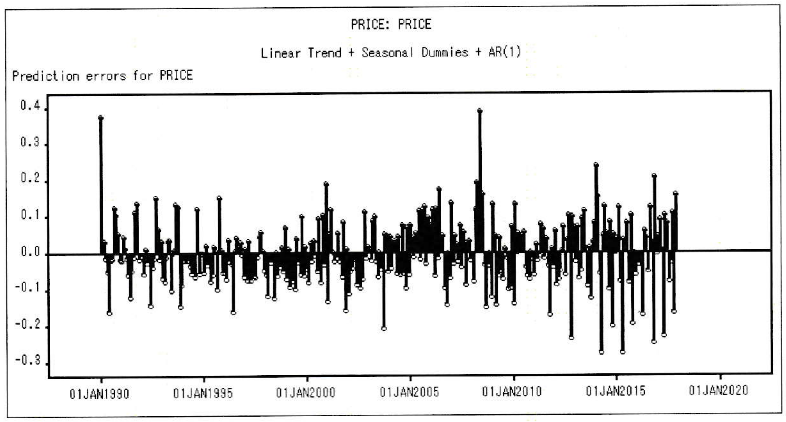


Figure 2.11

no need

The model seems to fit the series quite well. We see the ACF of this model acts almost like a white noise. But the seasonality still exists, we might need other models to better ~~fitting~~ the series.

No they're not there are seasonal AC's

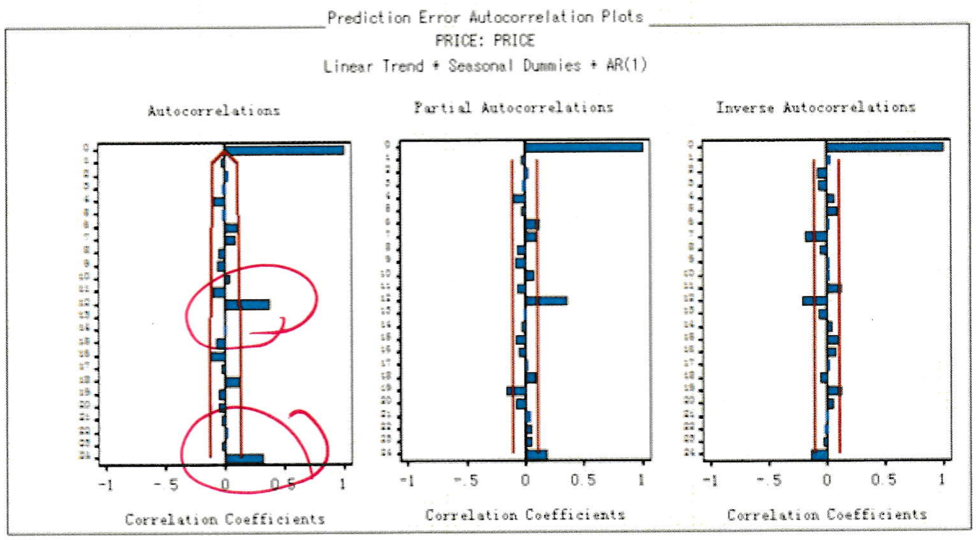


Figure 2.12

2.2 Exponential Smoothing models (only the relevant ones)

2.2.1 Simple Exponential Smoothing

GIven that the linear trend and seasonal dummy models are still not enough to fit our series, we need to estimate more models. Here is the results of a simple exponential smoothing model. From the fitted plot, we can see the series are well fitted.

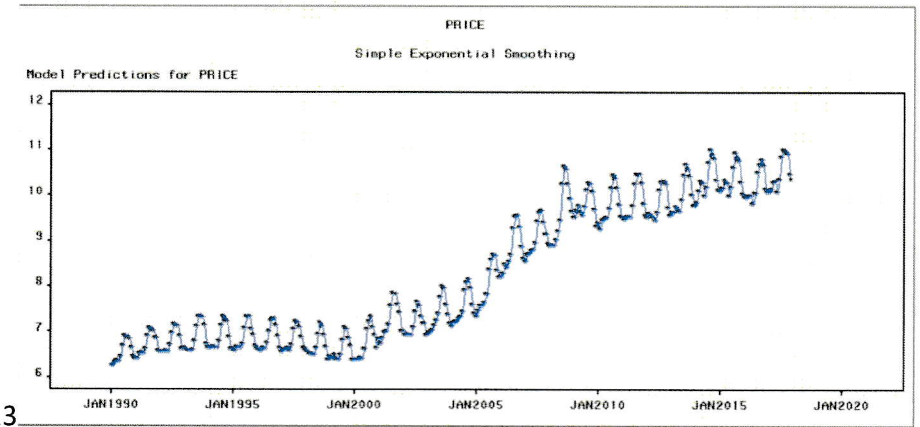


Figure 2.13

But, when we look at the residues and ACF, we can see that the seasonality exists in the residues and we need to take it into consideration.

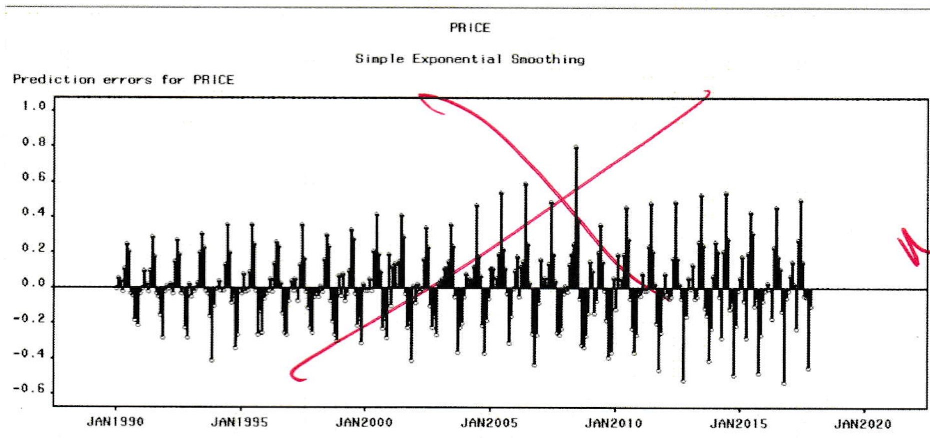


Figure 2.14

NO NEED!

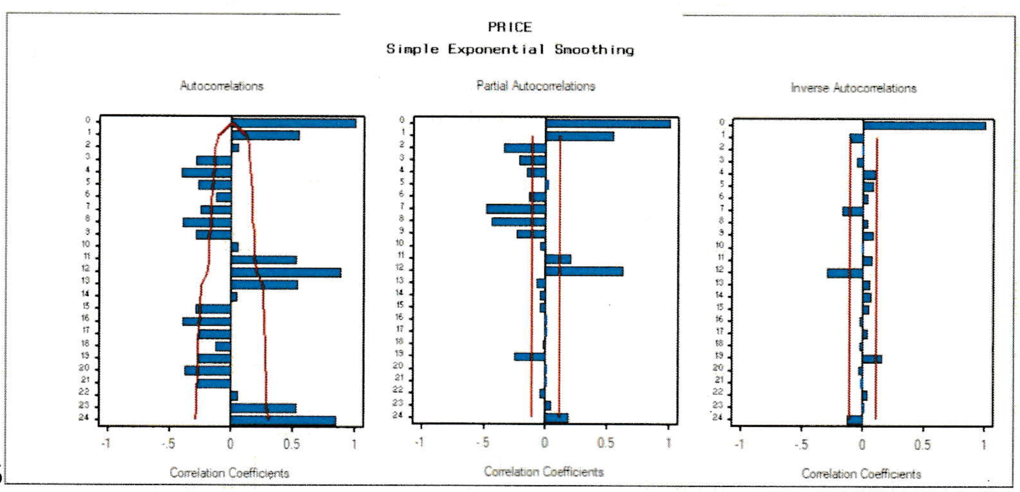


Figure 2.15

2.2.2 Seasonal Smoothing

To fit the series seasonality, we can use seasonal smoothing to model the series. The fitted series are shown below.

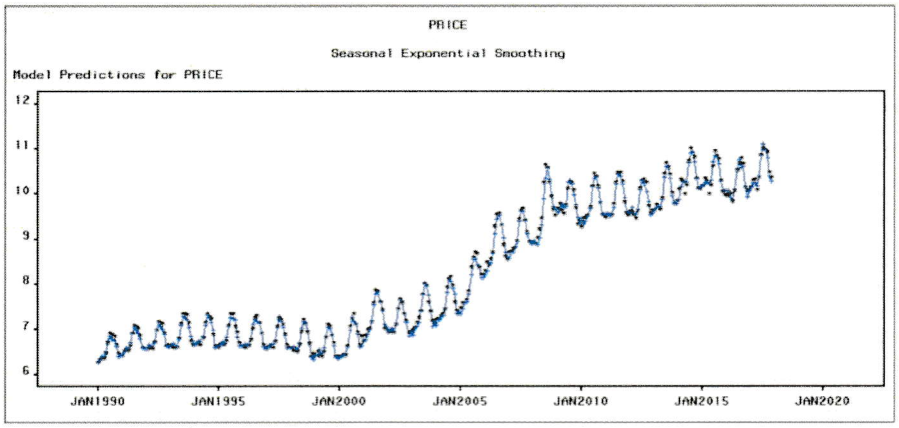


Figure 2.16

looks good

Pl. show the estimated component here (Figure 2.20 here)

We can look at the residues and ACF. It seems seasonal smoothing model performs much better than simple smoothing.

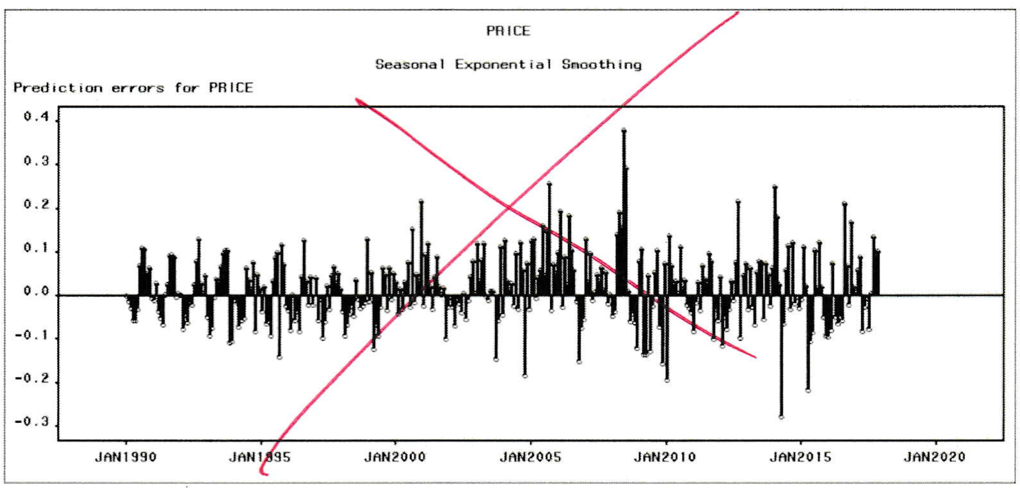


Figure 2.17

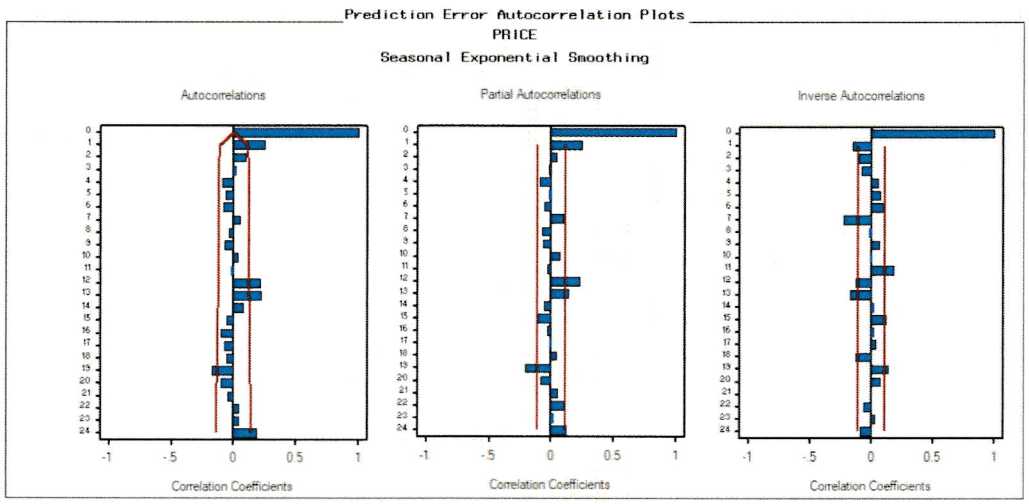


Figure 2.18

Seasonal smoothing model is a better model but it's still not enough. The model didn't capture all the seasonality behaviors in series. And, the series is not transformed to a white noise through the model. The white noise test results are shown below.

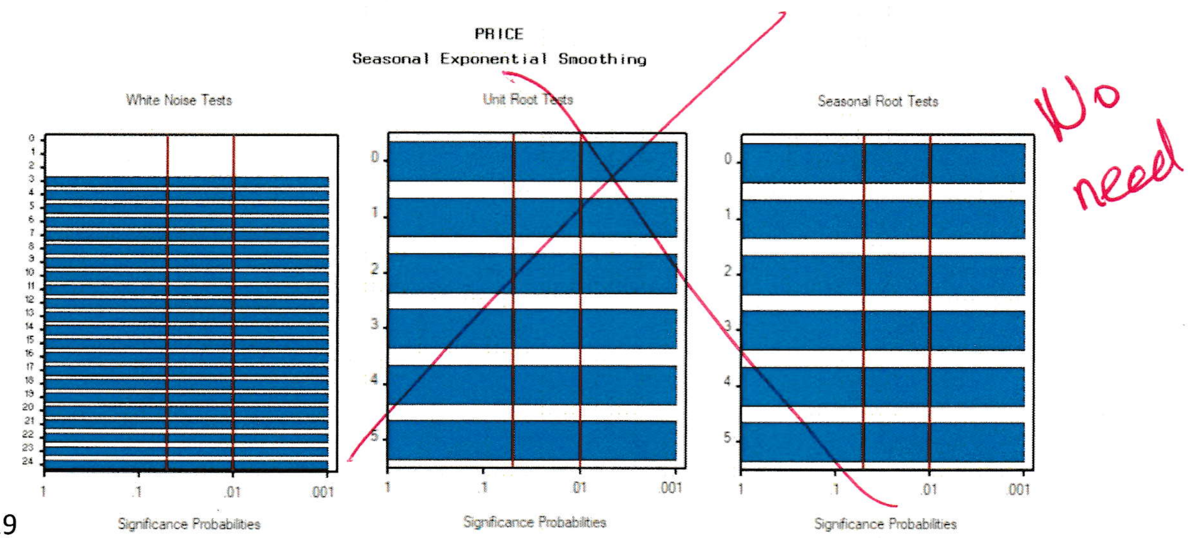


Figure 2.19

We can identify the seasonal factors in our model which is shown below. We can conclude that the series is relatively higher through May to September.

very smooth seasonal weight

PRICE

Seasonal Exponential Smoothing

Model Parameter	Estimate	Std Error	T	Prob>|T|
LEVEL Smoothing Weight	0.79880	0.0316	25.3095	<.0001
SEASONAL Smoothing Weight	0.99900	0.1902	5.2525	<.0001
Residual Variance (sigma squared)	0.00682	.	.	.
Smoothed Level	10.39063	.	.	.
Smoothed Seasonal Factor 1	-0.18088	.	.	.
Smoothed Seasonal Factor 2	-0.11495	.	.	.
Smoothed Seasonal Factor 3	-0.18582	.	.	.
Smoothed Seasonal Factor 4	-0.24443	.	.	.
Smoothed Seasonal Factor 5	0.01608	.	.	.
Smoothed Seasonal Factor 6	0.50846	.	.	.
Smoothed Seasonal Factor 7	0.69763	.	.	.
Smoothed Seasonal Factor 8	0.60169	.	.	.
Smoothed Seasonal Factor 9	0.42381	.	.	.
Smoothed Seasonal Factor 10	-0.05263	.	.	.
Smoothed Seasonal Factor 11	-0.26063	.	.	.
Smoothed Seasonal Factor 12	-0.24185	.	.	.

Figure 2.20

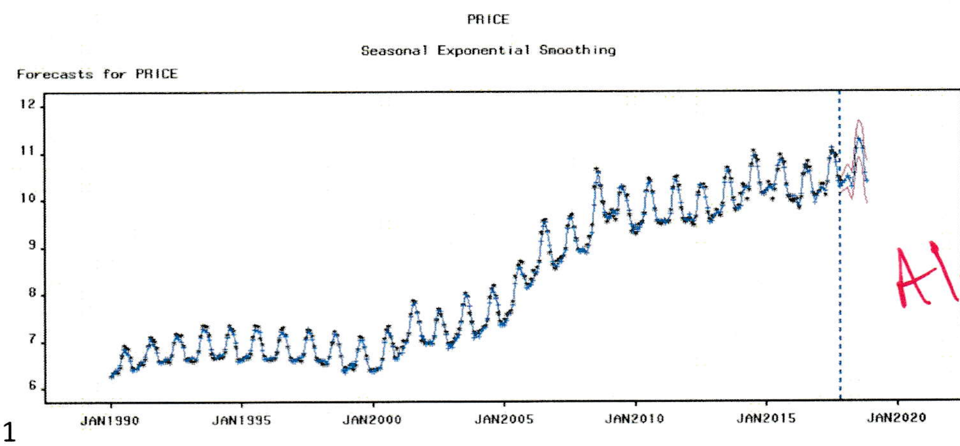


Figure 2.21

Already shown in 2.16 REDUNDANT

2.3 ARIMA models (with seasonal ARIMA components if relevant)

2.3.1 Series Differencing

2.3.1.1 Original Series

As we discussed above, the ACF and PACF of original series are not stationary and have seasonal pattern. We will try to do differencing and seasonal differencing to the series.

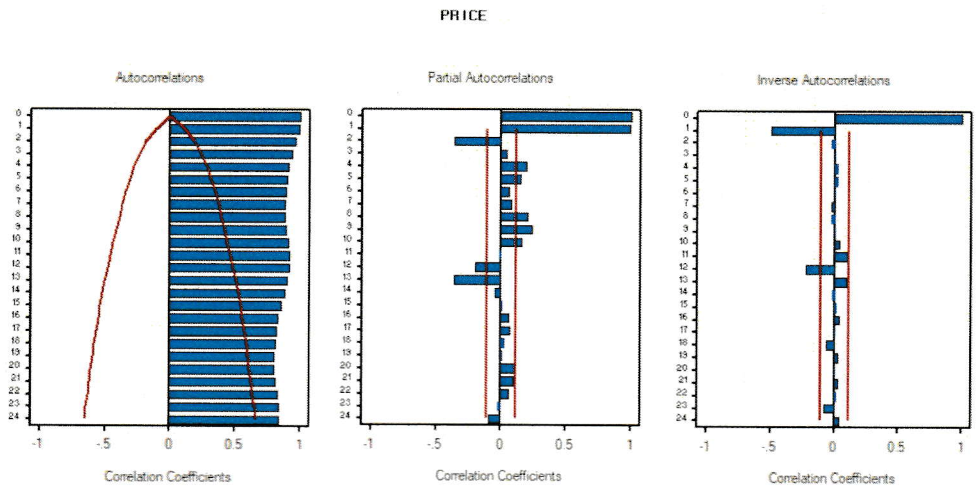


Figure 2.22

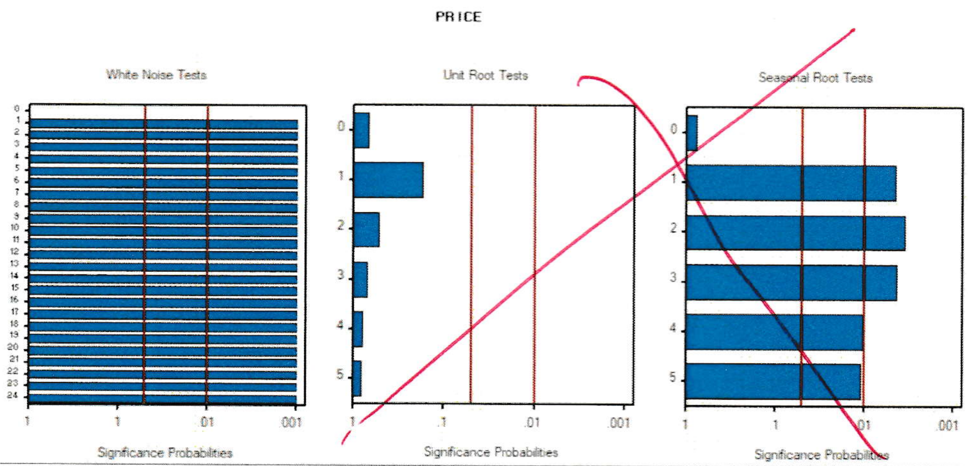


Figure 2.23

No need

2.3.1.2 First Difference

First, we do first difference to the series. We can tell an obvious seasonal trend for every 12 month from the plots below. The seasonal difference is still needed.

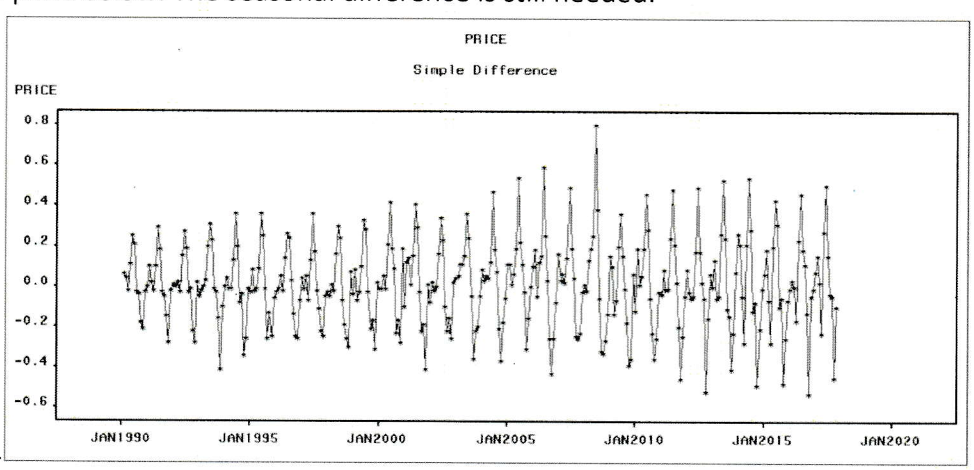


Figure 2.24

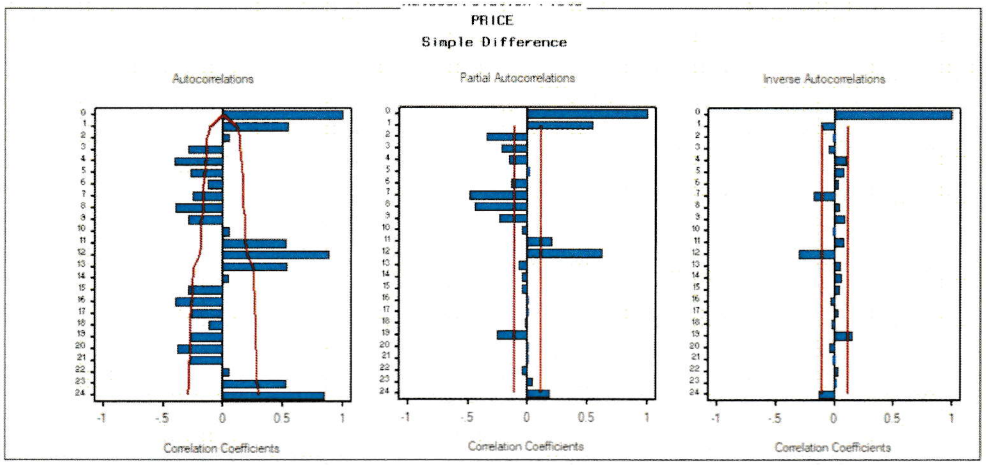


Figure 2.25

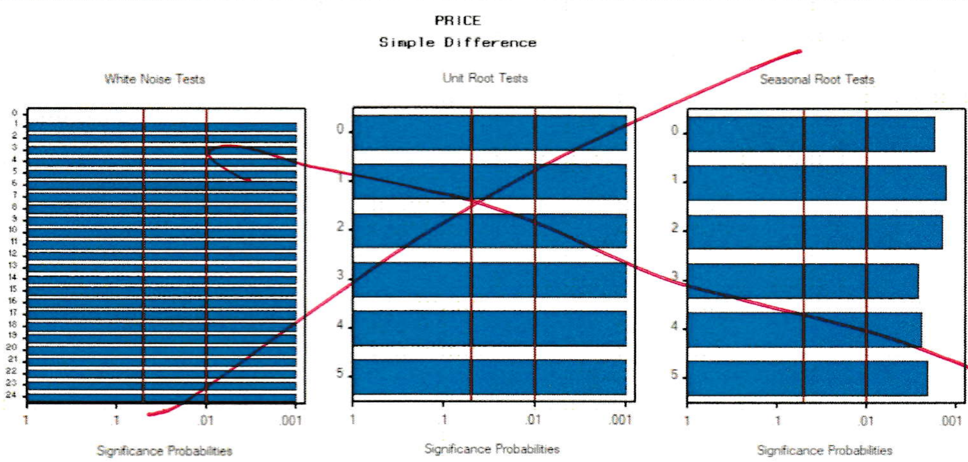


Figure 2.26

no need

2.3.1.3 First Simple Difference and Seasonal Difference

After first and seasonal difference, we can find that the series became stationary. The ACF and PACF is chopped to 0 after lag 0. For the seasonal factor, ACF cut off to 0 after first 12 lags. It suggests a pattern of ARIMA(0,1,0)(0,1,1)$_{12}$ model.

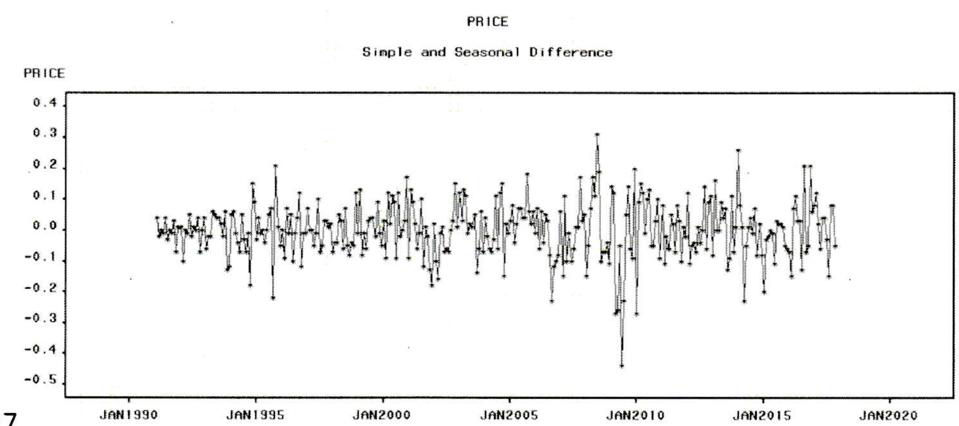


Figure 2.27

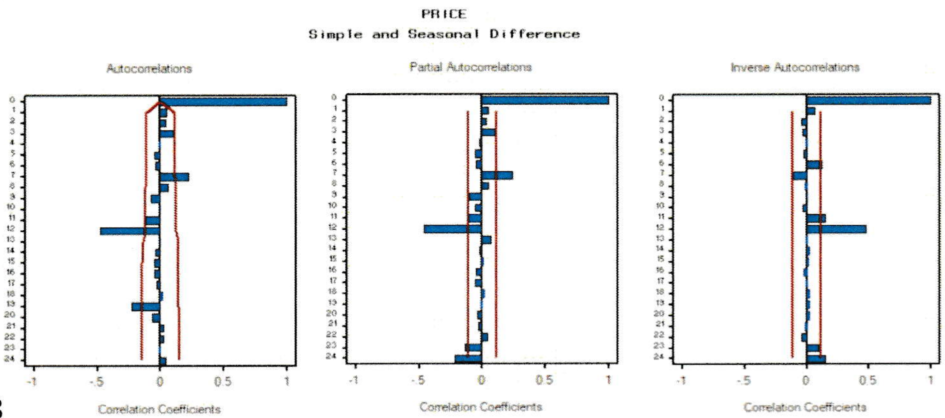


Figure 2.28

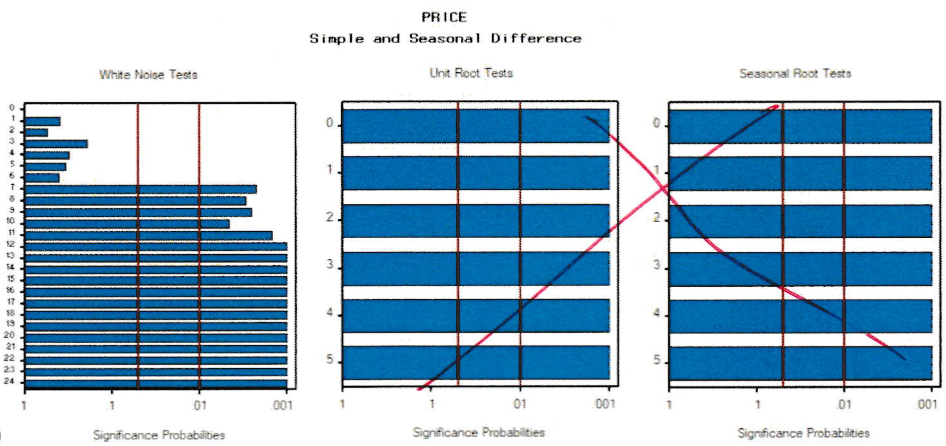


Figure 2.29

no need

2.3.1.4 ARIMA Model

nice!

We fitted the series with ARIMA(0,1,0)(0,1,1)$_{12}$ model. The model performs well in fitting the series. And the residue series is almost a white noise. The fitted series, residue and ACF plots are shown below.

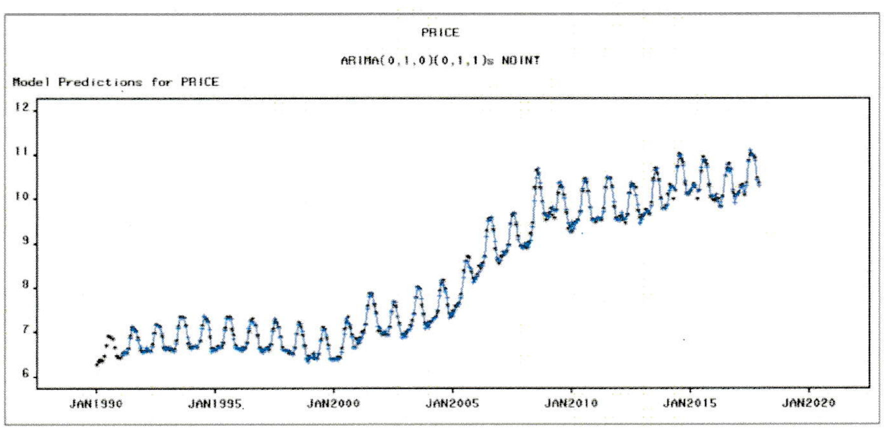


Figure 2.30

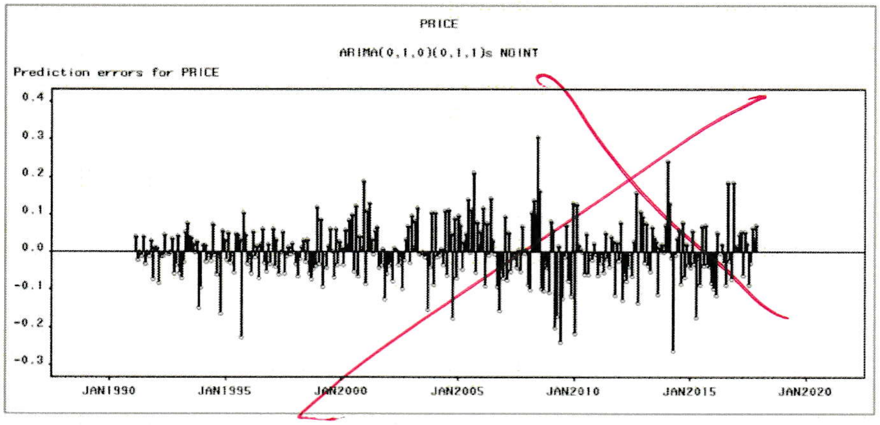


Figure 2.31

No need

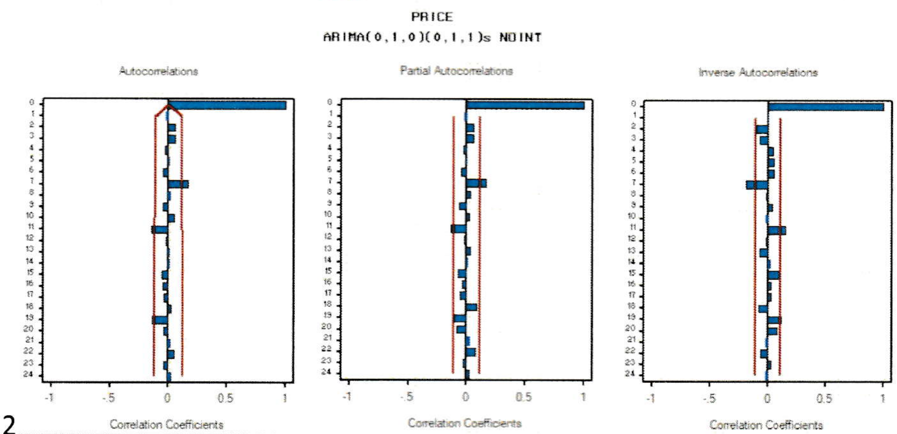


Figure 2.32

Practically it's WN, the AC's are so small

The white noise test is shown below. We can conclude that the series is not white noise.

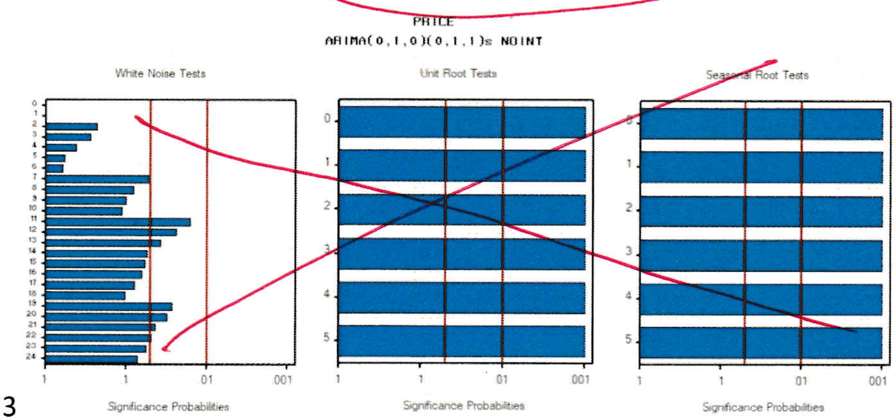


Figure 2.33

No need

The model has an MAPE of only 0.58. The model performs well and achieved low values in both MAPE and RMSE.

PRICE

ARIMA(0,1,0)(0,1,1)s NOINT

Model Parameter	Estimate	Std. Error	T	Prob>¦T¦
Seasonal Moving Average, Lag 12	0.65594	0.0513	12.7894	<.0001
Model Variance (sigma squared)	0.00583	.	.	.

Figure 2.34

Statistic of Fit	Value
Mean Square Error	0.0056614
Root Mean Square Error	0.07524
Mean Absolute Percent Error	0.58005
Mean Absolute Error	0.06015
R-Square	0.952

Figure 2.35

Here is the prediction made by ARIMA model.

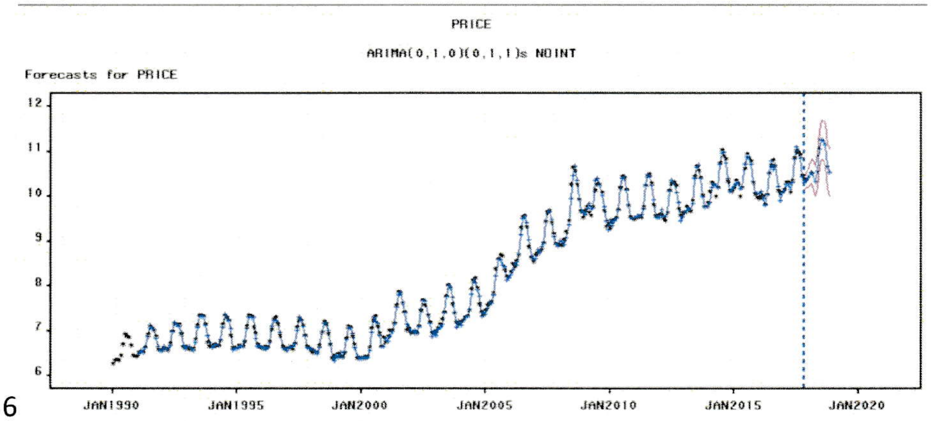


Figure 2.36

2.4 Comparison of models (in terms of fit and validation)

After analyzing the smoothing models and ARIMA models, we can find that the Seasonal Exponential Smoothing model and ARIMA(0,1,0)(0,1,1)$_{12}$ are the potential best model to fit and forecast the series.

From the forecast graph of the both models (Figure 2.21 and 2.36), we can conclude that they did a good job in forecasting for the hold-out samples.

Forecast Model	Model Title	Root Mean Square Error
	Seasonal Exponential Smoothing	0.08965
✓	ARIMA(0,1,0)(0,1,1)s NOINT	0.07524

Figure 2.37

Forecast Model	Model Title	Mean Absolute Percent Error
	Seasonal Exponential Smoothing	0.69940
✓	ARIMA(0,1,0)(0,1,1)s NOINT	0.58005

Pl: also compare model fit based model variances.

Figure 2.38

When Root Mean Square Error and Mean Absolute Percent Error are used to evaluate the two models, the ARIMA(0,1,0)(0,1,1)$_{12}$ model is the better model for it has lower errors.

As the results showed above, ARIMA(0,1,0)(0,1,1)$_{12}$ is the best model for univariate time series forecasting for US electricity price.