Electricity Transfer Function

STA6253

Lindsey Dodd
Reisa Widjaja

Introduction

This project seeks to identify a transfer function noise model for use in forecasting residential electricity sales in the US (Y) with retail price of electricity as the input (X). We first considered only the data from January 1990 to July 2009 to develop the model. We then used the model to forecast residential electricity sales in the US for the first 7 months of 2010 and compare the forecasts with the actual values for the indicated dates.

Data

"Electric_TF" data: 247 Obs, 4 Variables (Year, Month, Sales, Avg_Price)

- Define Xt = Average Retail Price Residential (c/kWh)
- Define Yt = Residential Sales (Mwh)
- Time = Monthly Data 1990JAN ~ 2009DEC (~2010JUL)
- Analyses in the next few session are using Train Data

Obs	year	month	res_sales	avg_price	Date
1	1990	1	95420231	7.17	JAN90
2	1990	2	74498370	7.48	FEB90
3	1990	3	71901767	7.57	MAR90
4	1990	4	65190618	7.69	APR90
5	1990	5	62881008	7.96	MAY90
6	1990	6	73899811	8.10	JUN90
7	1990	7	90935492	8.18	JUL90

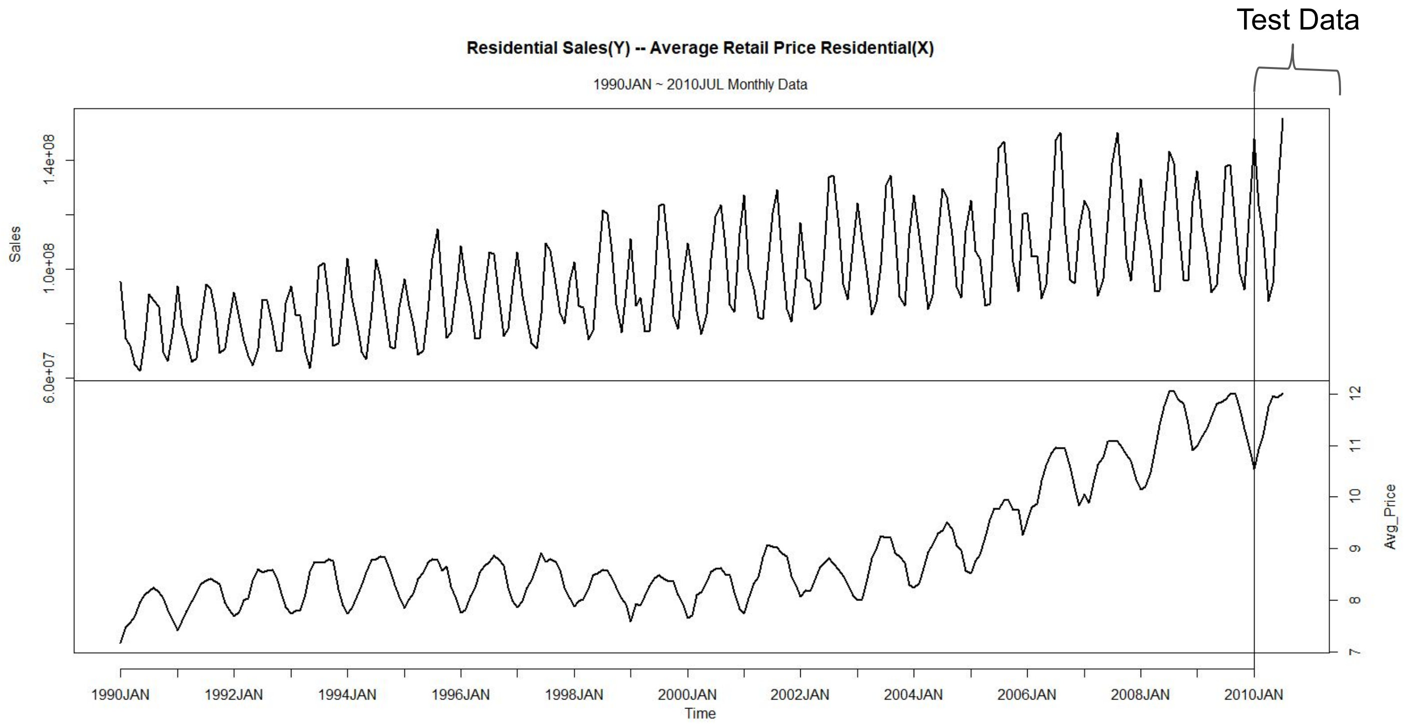
Partial view of the Train Data

Obs	year	month	res_sales	avg_price	Date
1	2010	1	147848708	10.54	JAN10
2	2010	2	123329790	10.93	FEB10
3	2010	3	112057413	11.20	MAR10
4	2010	4	88111138	11.75	APR10
5	2010	5	94776950	11.96	MAY10
6	2010	6	126974815	11.92	JUN10
7	2010	7	155325187	12.01	JUL10

Test Data

Plot of Yt & Xt

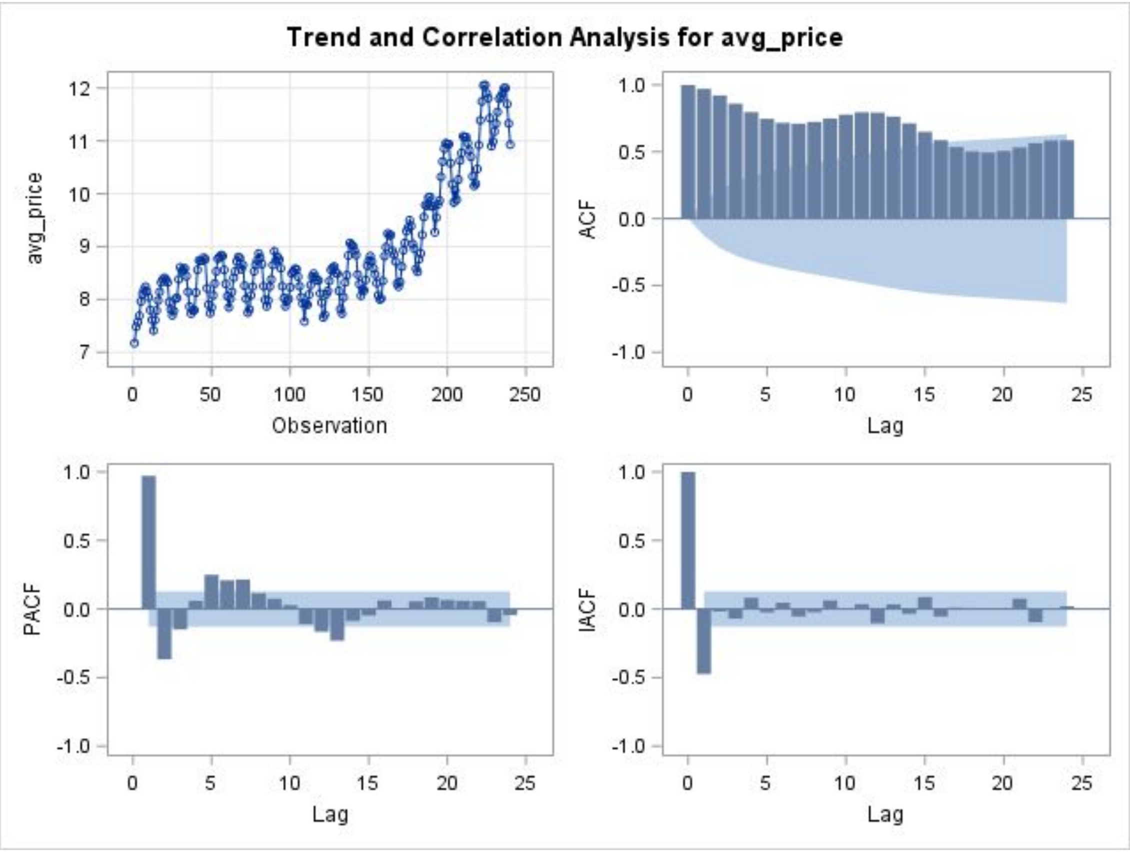
- Not Stationary
  - Increasing mean
  - Constant variance
- Seasonal Pattern



Xt series (Avg_Price)

- Not Stationary
- Seasonal Pattern
- ACF decays every 12 lags (seasonal), and decays within 12 lags (non seasonal)
- PACF dies out fast, several lags have significant values

Suggesting Taking 1st Difference & 1st Seasonal Difference for the Xt series → “Diff.diff.Xt”



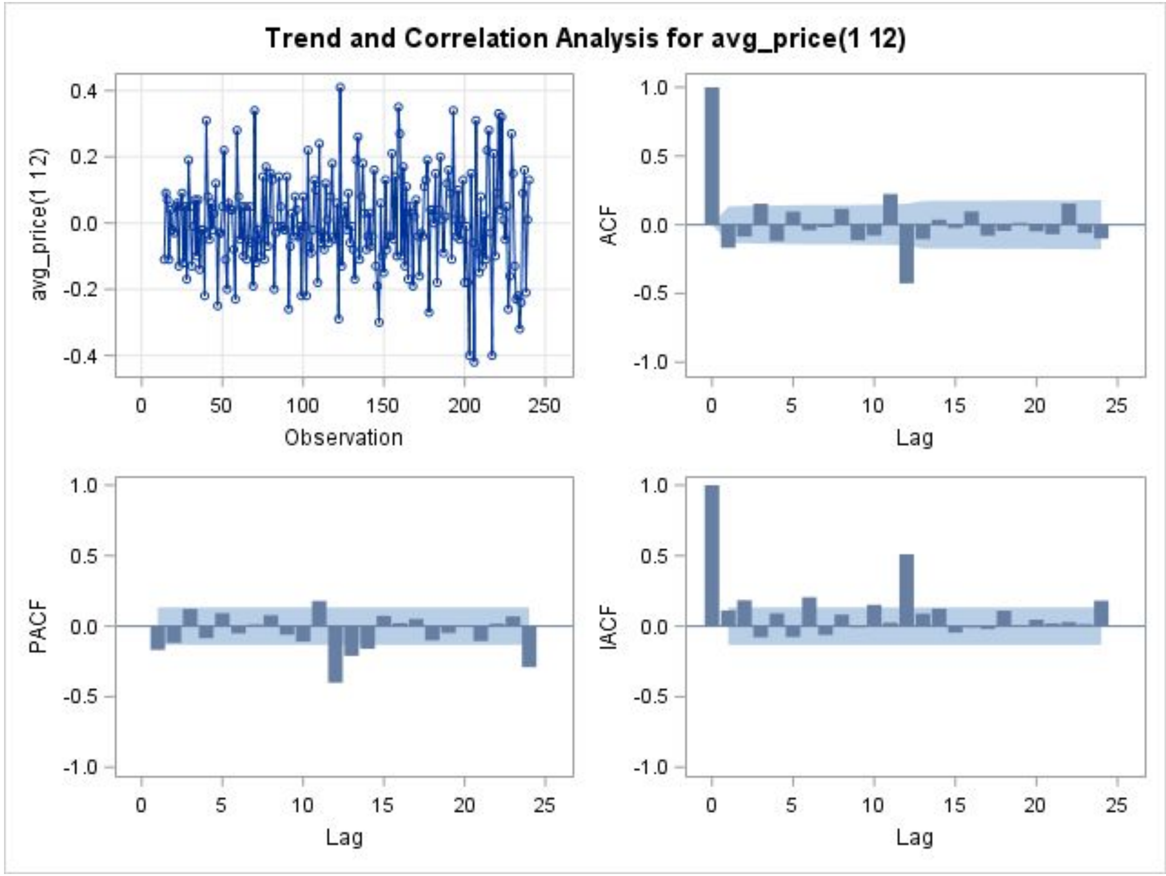
Plot of Xt Series and ACF, PACF, IACF

Diff.diff.Xt series

- Time series data looks much more stationary. (mean around zero)
- ACF has spikes at lag 11 & 12
- PACF values are generally low, with lag 1,11,12,13,14,24 are accounted as “statistically significant”
- EACF suggest ARMA(1,1) for Diff.diff.Xt series
-

```
> eacf(Diff.diff.Xtrain)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x o x o o o o o o o x  x  o  o
1 x o o o o o o o o o o  x  x  o
2 x o o o o o o o o o o  x  x  o
3 x x o o o o o o o o o  x  o  x
4 x x o o o o o o o o o  x  x  x
5 x x x o o o o o o o o  x  o  o
6 x o x x o o o o o o o  x  x  o
7 x x o o o o o o o o o  x  x  o
```

- auto.arima() in R suggested on ARIMA(2,1,1)(0,1,1)[12] for Xt data



Plot of Diff.diff.Xt Series and ACF, PACF, IACF

Identify Model for Xt (Prewhitening Process) ARIMA(2,1,1)(0,1,1)[12]

```
> auto.arima(Xtrain, trace=FALSE) Series: Xtrain
ARIMA(2,1,1)(0,1,1)[12]

Coefficients:
         ar1      ar2     ma1     sma1
     -0.9542  -0.2709  0.7621  -0.7570
s.e.  0.1164   0.0642  0.1064   0.0521

sigma^2 estimated as 0.01256:  log likelihood=171.52
AIC=-333.03   AICc=-332.76   BIC=-315.91
```

Conditional Least Squares Estimation

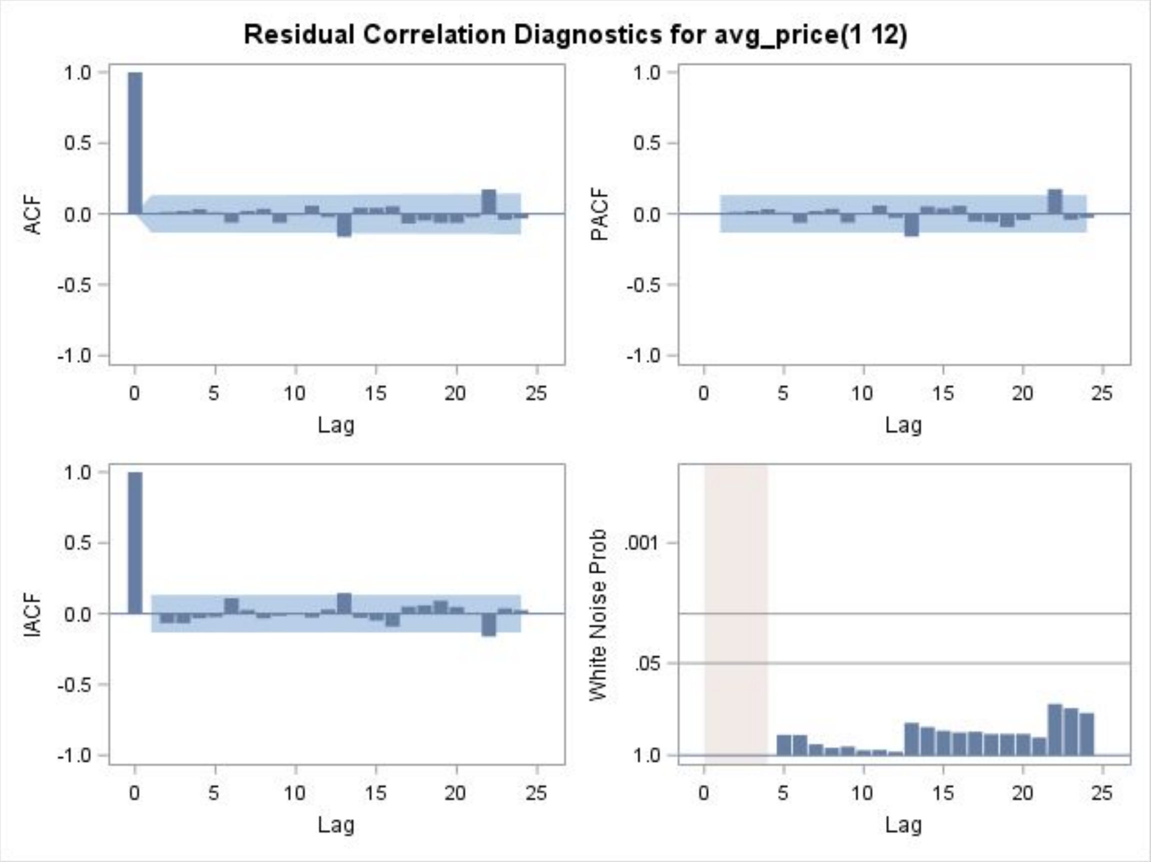
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MA1,1	-0.76565	0.14839	-5.16	<.0001	1
MA2,1	0.75750	0.04764	15.90	<.0001	12
AR1,1	-0.95934	0.15055	-6.37	<.0001	1
AR1,2	-0.27547	0.06452	-4.27	<.0001	2

$$\phi(B)W_t = \theta(B)\alpha_t \;\; or \;\; \alpha_t = \frac{\theta(B)}{\phi(B)}W_t = \frac{\theta(B)\Theta(B)}{\phi(B)}W_t = \frac{(1+0.76565B)(1-0.7575B^{12})}{(1+0.95934B+0.27547B^2)}$$

$$where \;\; W_t = \nabla\nabla_{12}X_t \rightarrow \text{Diff.diff.Xt}$$

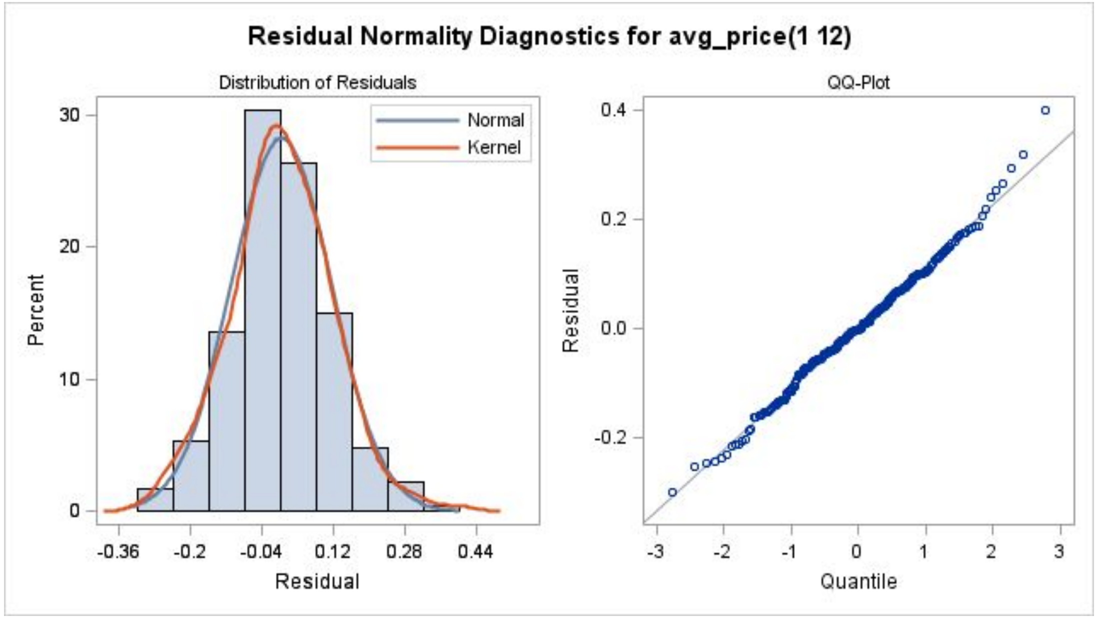
Identify Model for Xt (Prewhitening Process) Model Diagnostic

- ACF & PACF show a “statistical significant” values at lag 13 & 22 but very low. Most values are insignificant → No correlation between lags
- White Noise Probability Plot
  - p-values>0.05 → residuals is white noise



Identify Model for Xt (Prewhitening Process)
Model Diagnostic

- Check Normality
  - QQplot & Histogram → Residuals are Normal
  - Shapiro-Wilk Test: p-value=0.1152 → Residuals are Normal
- Check Independency
  - Ljung-Box Test: p-value=0.1044 → Residuals are Independent with each other.



Applying Pre-whitened Filter to Yt
Identify Impulse Response Function v(B)

Cross Correlations of Diff.diff.Yt and Diff.diff.Xt
(with ±2 standard error limits)

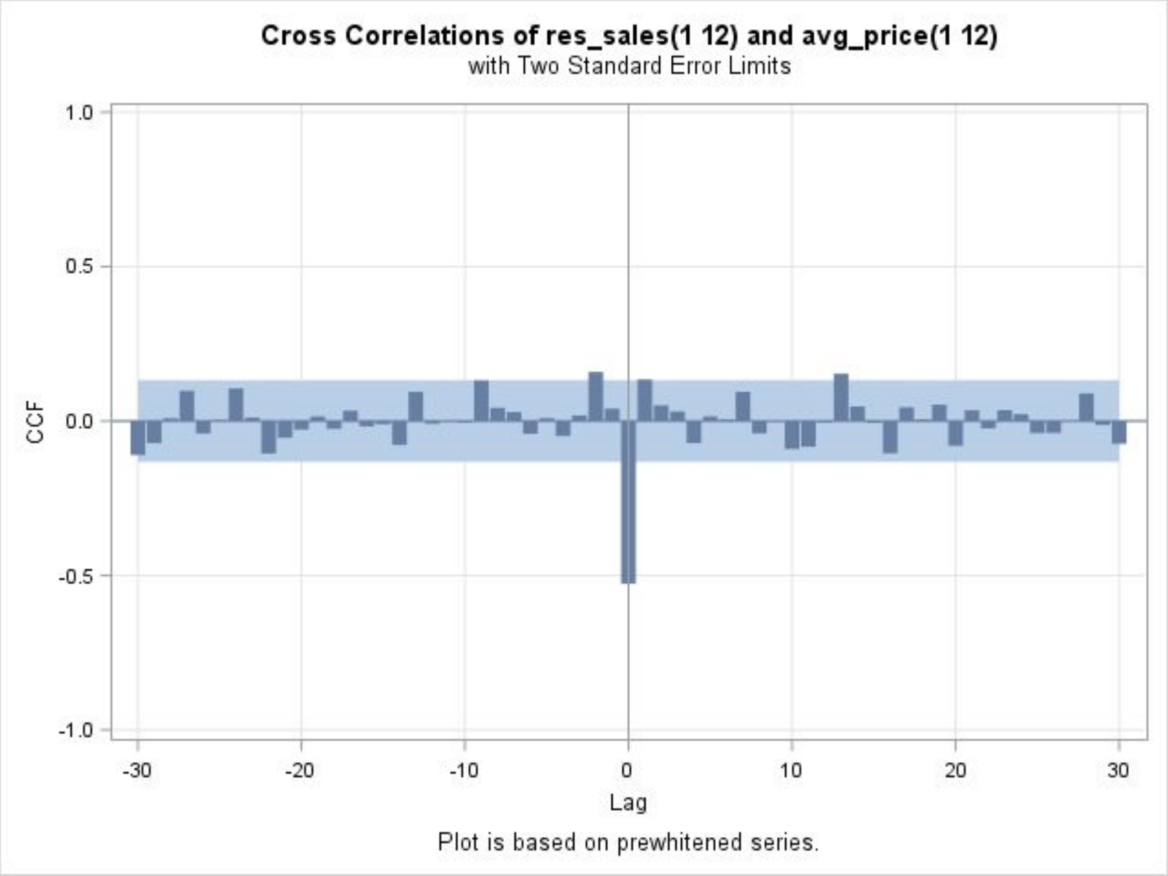
- Spikes at lag 0
- b=0, s=1, r=0

Impulse Response Function

No mean term in this model.

Input Number 1	
Input Variable	avg_price
Period(s) of Differencing	1,12

Numerator Factors	
Factor 1:	-2.32E7 + 4091158 B**(1)



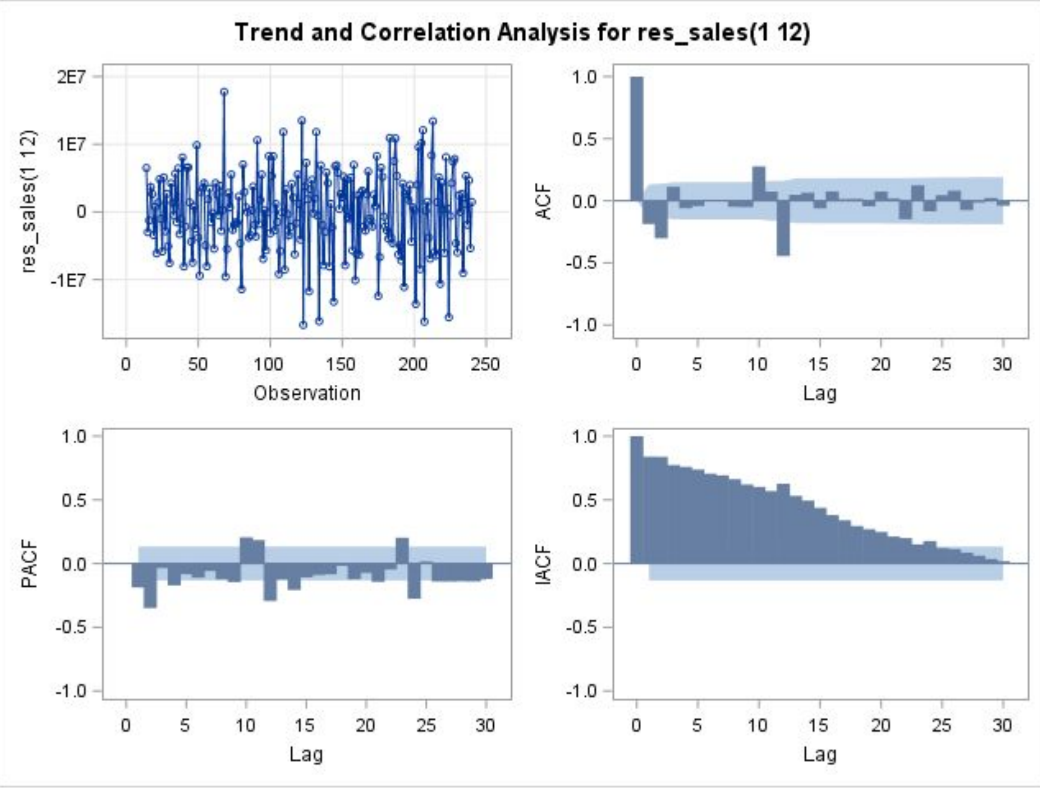
Applying Pre-whitened Filter to Yt
Identify Impulse Response Function v(B)

Cross Correlation Check of Residuals with Input Xt

- p-values > 0.05 → The parameters b, s, r for the Impulse Response Function is defined well

Crosscorrelation Check of Residuals with Input avg_price

To Lag	Chi-Square	DF	Pr > ChiSq	Crosscorrelations					
5	3.28	4	0.5118	0.007	0.031	0.047	-0.035	-0.097	0.024
11	14.06	10	0.1704	0.028	0.082	-0.002	-0.068	-0.185	-0.034
17	15.04	16	0.5216	0.017	0.018	0.045	-0.022	-0.036	0.001
23	22.15	22	0.4512	0.022	-0.057	-0.091	0.067	0.119	-0.029
29	27.26	28	0.5042	-0.004	-0.040	-0.031	-0.022	0.140	-0.000
35	39.07	34	0.2524	0.005	-0.030	0.100	0.044	-0.187	0.067
41	42.11	40	0.3798	-0.049	0.016	0.006	0.043	-0.082	-0.046



Plot of Diff.diff.Yt Series and ACF, PACF, IACF

Residuals Diagnostic for Diff.diff.Yt

Autocorrelation Check of Residuals

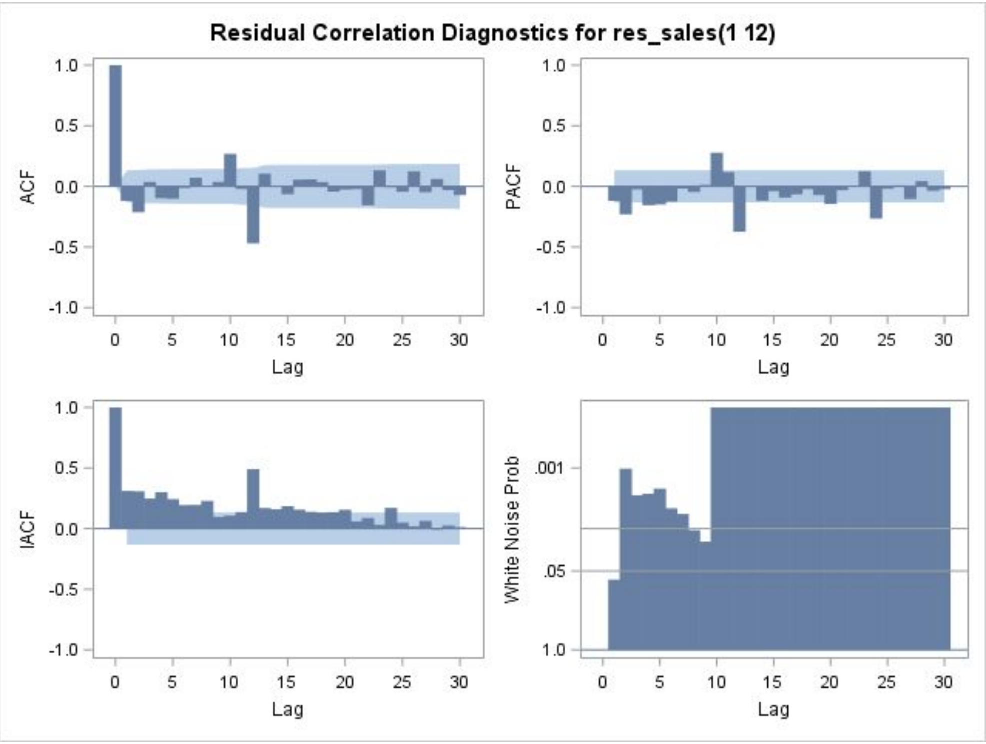
- p-values < 0.05 → there are correlation between lags

Residual Correlation Diagnostic for Diff.diff.Yt

- Several significant lags in ACF & PACF
- White Noise Probability plot
  - Most p-values < 0.05 → Residuals are not white noise

Refit & Reidentify model is necessary.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	18.72	6	0.0047	-0.120	-0.213	0.036	-0.098	-0.102	-0.013
12	90.69	12	<.0001	0.071	0.004	0.037	0.268	-0.022	-0.470
18	96.27	18	<.0001	0.105	-0.002	-0.064	0.056	0.058	0.036
24	107.65	24	<.0001	-0.044	-0.026	-0.022	-0.156	0.132	-0.008
30	115.31	30	<.0001	-0.044	0.124	-0.048	0.061	-0.031	-0.071
36	121.31	36	<.0001	0.114	0.070	-0.059	0.034	0.001	0.012
42	130.97	42	<.0001	-0.018	-0.055	0.100	-0.101	0.100	0.033



Refit & Reidentify the Model

1. Identify ARMA(p,q) model to refit the transfer function model for Diff.diff.Yt
2. Check Residuals Diagnostic
3. Repeat steps until finding the proper p & q order for the ARMA model.

- ARMA(2,1)(0,1)[12] has been chosen for Diff.diff.Yt

Residuals Diagnostic for Diff.diff.Yt

Cross Correlation Check of Residuals with Input Xt
- p-values > 0.05 → The parameters b, s, r for the Impulse Response Function is defined well

Autocorrelation Check of Residuals
- p-values > 0.05 → there are no correlation between lags

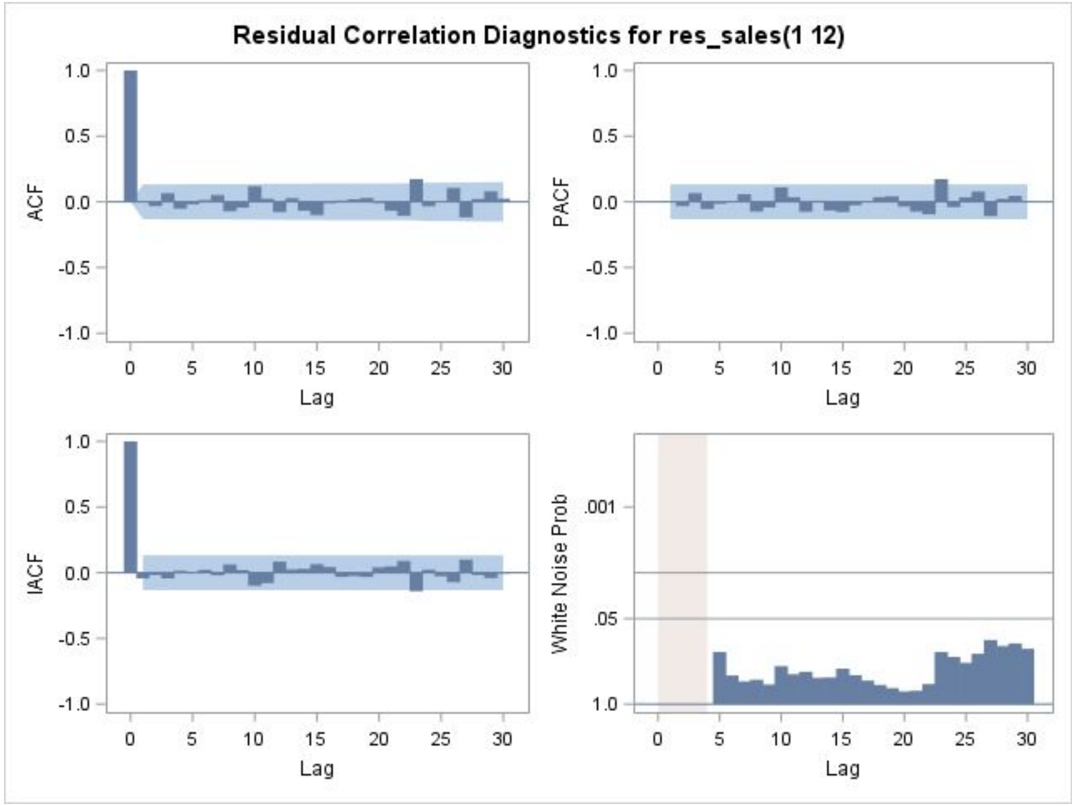
Crosscorrelation Check of Residuals with Input avg_price									
To Lag	Chi-Square	DF	Pr > ChiSq	Crosscorrelations					
5	5.73	4	0.2203	-0.083	-0.022	0.101	0.048	-0.057	0.046
11	13.56	10	0.1939	-0.006	0.146	-0.014	-0.013	-0.054	-0.101
17	17.93	16	0.3280	-0.009	0.071	0.086	0.009	-0.073	0.038
23	22.88	22	0.4085	0.027	0.030	-0.090	0.066	0.068	-0.056
29	28.34	28	0.4464	0.079	0.013	-0.009	-0.027	0.130	0.006
35	38.87	34	0.2595	0.009	-0.031	0.066	0.090	-0.165	0.077
41	41.42	40	0.4086	-0.052	0.001	-0.020	0.034	-0.051	-0.067

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	2.00	2	0.3678	0.005	-0.033	0.066	-0.051	-0.019	0.014
12	9.24	8	0.3226	0.053	-0.071	-0.044	0.119	0.025	-0.077
18	13.13	14	0.5166	0.029	-0.067	-0.100	-0.009	0.011	0.021
24	25.19	20	0.1942	0.030	-0.008	-0.066	-0.104	0.174	-0.033
30	33.63	26	0.1446	0.001	0.107	-0.116	0.023	0.081	0.025
36	39.58	32	0.1676	0.115	0.048	-0.074	0.024	-0.000	0.030
42	46.09	38	0.1723	0.021	-0.084	0.007	-0.111	0.060	-0.004

Residuals Diagnostic for Diff.diff.Yt (After refit)

Residual Correlation Diagnostic for Diff.diff.Yt

- Most ACF & PACF are not significant → no correlation between lags
- White Noise Probability plot
    - Most p-values > 0.05 → Residuals are white noise



Transfer Function Model for Diff.diff.Yt

$Z_t = \nabla\nabla_{12}Y_t$ → Diff.diff.Yt
$W_t = \nabla\nabla_{12}X_t$ → Diff.diff.Xt

Impulse Response Function

$$\nu(B) = \frac{\omega_s(B)B^b}{\delta_r(B)} = \frac{\omega_1(B)B^0}{\delta_0(B)} = \omega(B)$$

$$\nu(B) = \omega(B) = -1.62E^{-7} + 9944617B$$

Transfer Function Model

$$Z_t = \nu(B)W_t + \eta_t$$

$$Z_t = (-1.62E^{-7} + 9944617BW_t) + \eta_t$$

Where

$$\eta_t = \frac{(1-0.93983B)(1-0.68696B^{12})}{(1-0.7004B+0.16535B^2)}\varepsilon_t$$

No mean term in this model.

Autoregressive Factors	
Factor 1:	1 - 0.7004 B**(1) + 0.16535 B**(2)

Moving Average Factors	
Factor 1:	1 - 0.93983 B**(1)
Factor 2:	1 - 0.68696 B**(12)

Input Number 1	
Input Variable	avg_price
Period(s) of Differencing	1,12

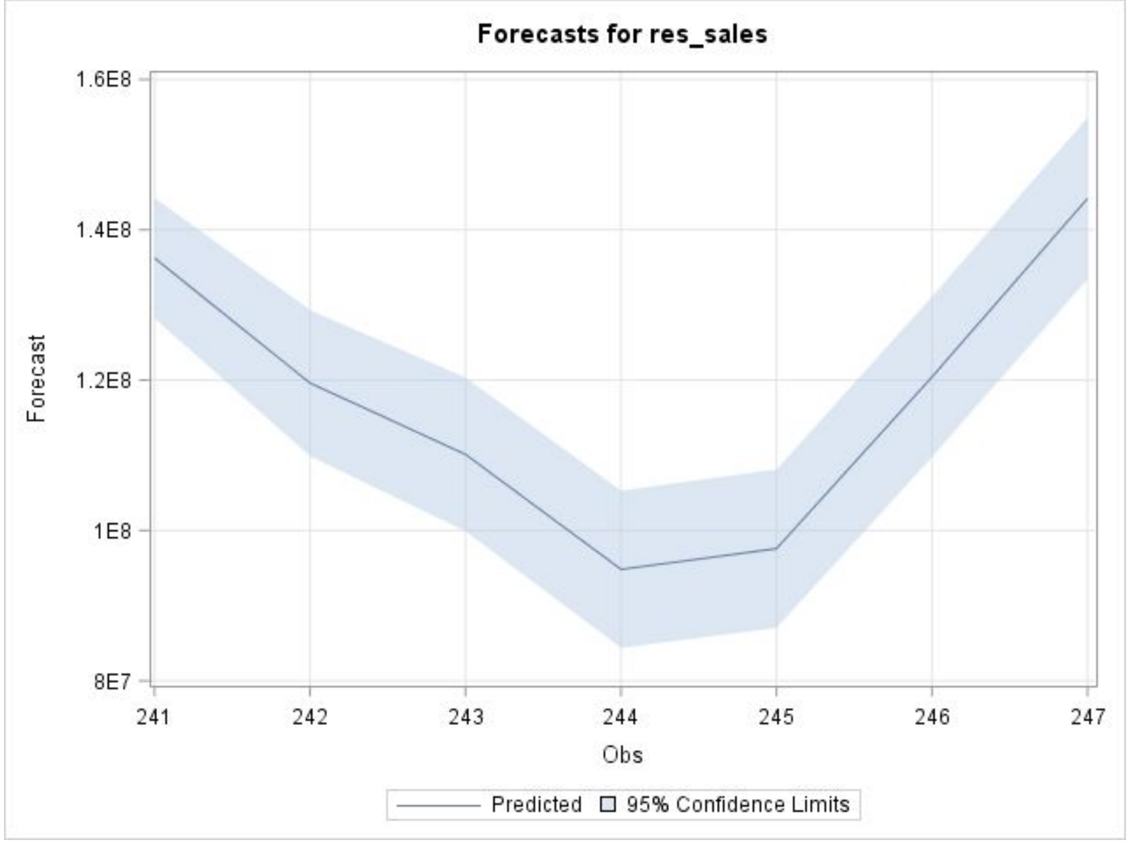
Numerator Factors	
Factor 1:	-1.62E7 + 9944617 B**(1)

Forecasting Results

Forecasts of Residential Sales for the next 7 months (JAN~JUL 2010) with 95% Confidence Limits

Forecasts for variable res_sales

Obs	Forecast	Std Error	95% Confidence Limits	
241	136232592	4088707	128218874	144246310
242	119614223	4955673	109901282	129327165
243	110127810	5211666	99913132.6	120342488
244	94838185.2	5335982	84379851.7	105296519
245	97589901.8	5380524	87044267.8	108135536
246	120437293	5440005	109775079	131099507
247	144158246	5483245	133411284	154905209



Forecasting Results

Compare with Observed Yt (Test Data)