Time Series Project

Spring 2011

Candidate: Muhammad Ashar

Ref No: 33040537

Estimation of Jet Fuel Prices

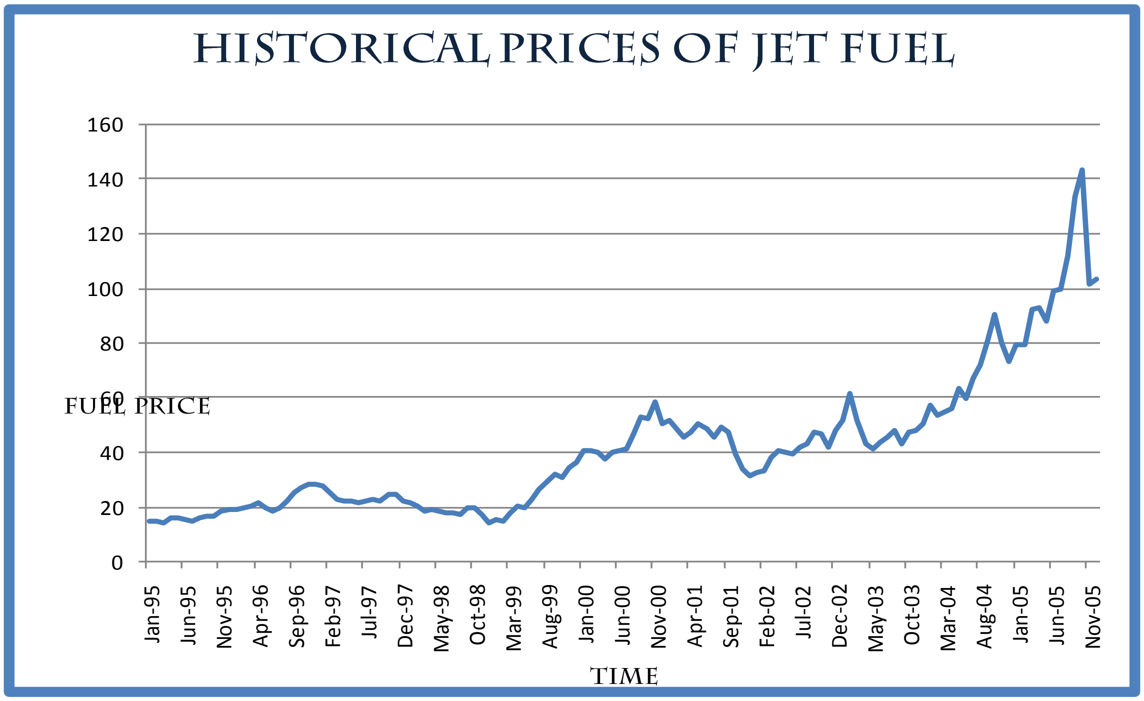
Introduction

Historically, the prices of jet fuel are showing much volatility especially over the past few years the variability of fuel prices causes serious problem for private airlines as well as military budget. If the future prices are predicted through regression analysis and time series modeling it would not only assist in lessening the stress caused due to unknown future fuel prices, but also help investors to hedge themselves against variable future prices

The project aims to model the price of jet fuel per gallon. Initially, the data is thoroughly examined and several models including AR(1), AR(2), AR(3) were analyzed to estimate the best fitted model.

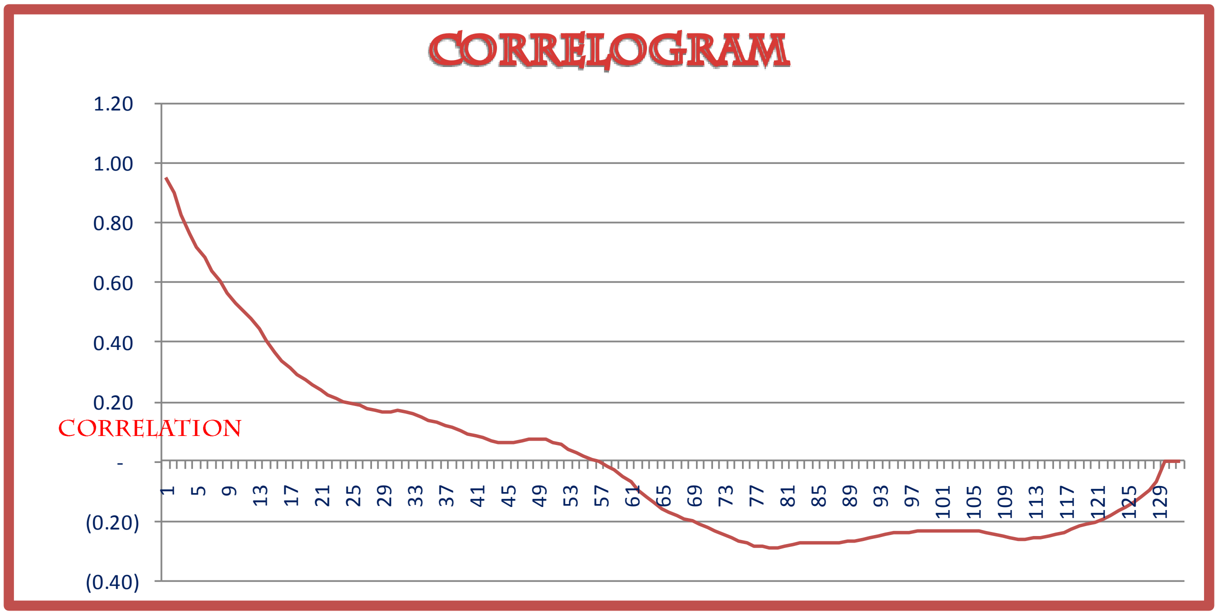
Data

The data of monthly jet fuel prices from January 1995 to December 2005 (132 data points) is used. The data is taken from http://www.indexmundi.com/commodities/?commodity=jet-fuel&months=360. The graph of the raw data is shown below:



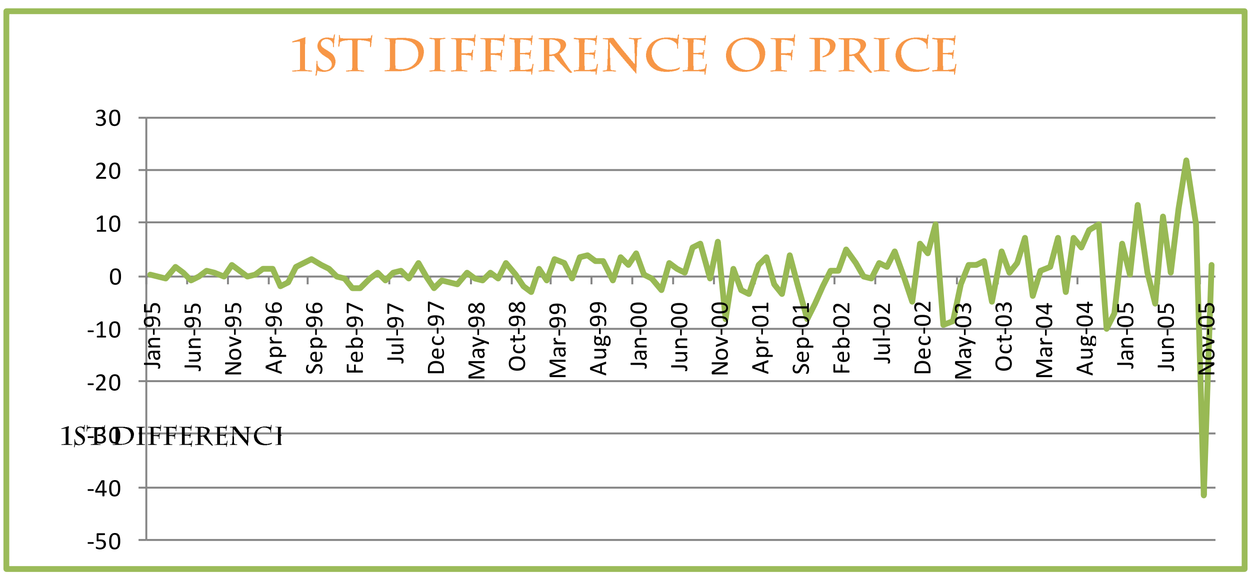
Please see the tab “data (Raw Data)” for the method to produce the graph.

The data is checked fro stationarity before attempting to fit an ARIMA model. Hence, sample autocorrelation is performed to determine the stationarity of the series.

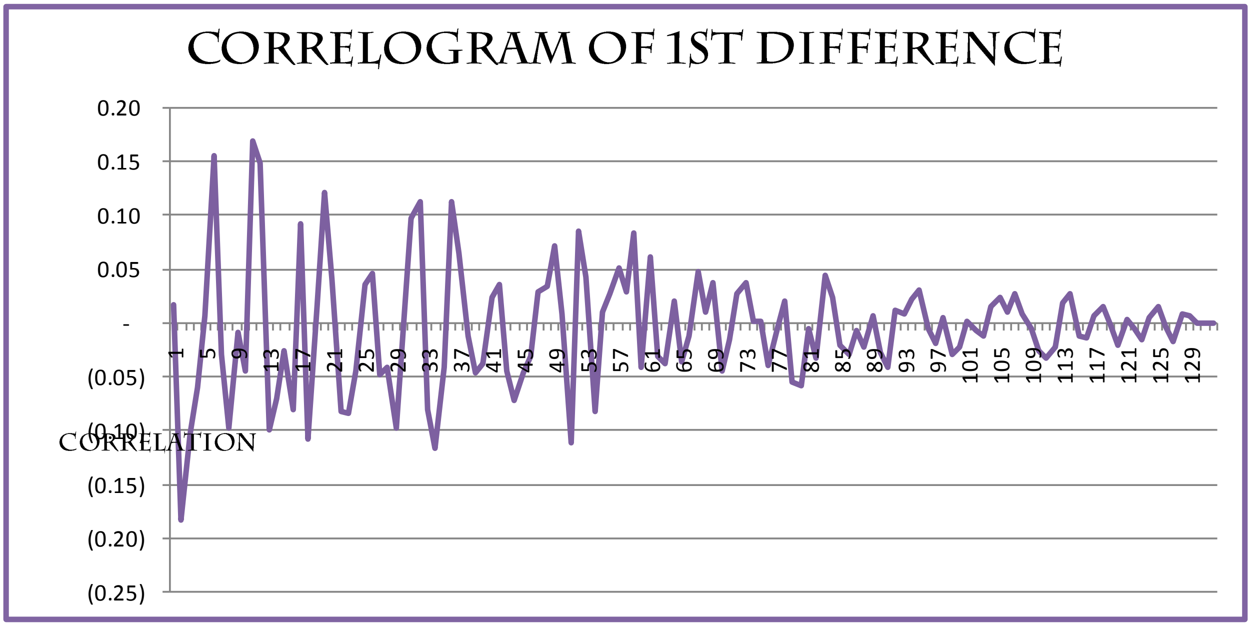


Please refer to the "data" tab in Excel. As we can see from the graph, the correlation does not reach 0 until around lag 57 (must reach 0 "quickly" which it clearly does not in this case). Then it goes below 0 and levels out to zero at lag 130. I will look at first differences to see if it is stationary.

1st Difference



And its corresponding correlograms is shown as below:



The 1st difference appears to be stationary because it does not depict any fluctuations with respect to time. The autocorrelation both go to zero quickly and oscillate around zero and decrease oscillation as lag increases.

Several models such as AR(1), AR(2) and AR(3) models will be fit to the 1st difference

Parameterization of the Model

After having stationary data, we will fit our data using an autoregressive model (AR (p)), using p = 1, 2 and 3. Since we are using the first difference, this is equivalent to ARIMA (p, 1, 0) models.
By Using Excel's regression data analysis add-in, the following is a summary of the regression results along with the resulting AR equations:

The detail results are attached in Excel sheets, (see tabs 1st Regression, 2nd Regression and 3rd Regression)

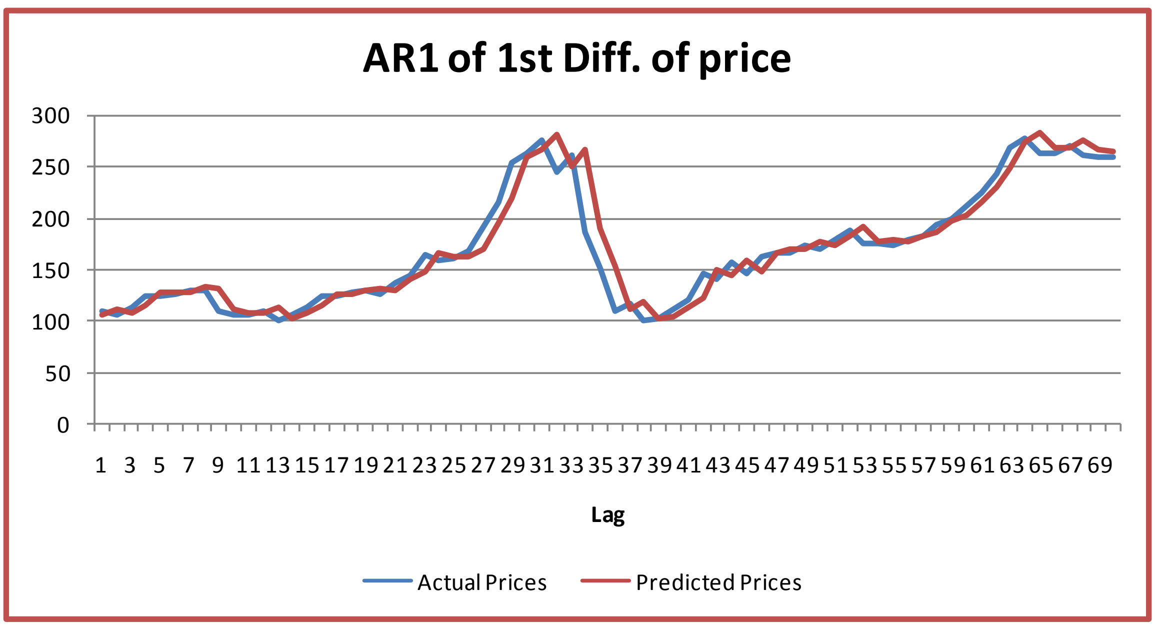
AR (1)

Regression Statistics	
Multiple R	0.017676993
R Square	0.000312476
Adjusted R Square	-0.00743704
Standard Error	5.784822444

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.6656	0.5088	1.3081	0.1932	(0.3411)	1.6722	(0.3411)	1.6722
$Y_{(T-1)}$	0.0177	0.0880	0.2008	0.8412	(0.1565)	0.1919	(0.1565)	0.1919

So the first equation is:

AR (1): $Y_t = 0.01768Y_{t-1} + 0.66558 + e_t$



The adjusted R^2 is negative showing the series shows negligible or no independence on each other.

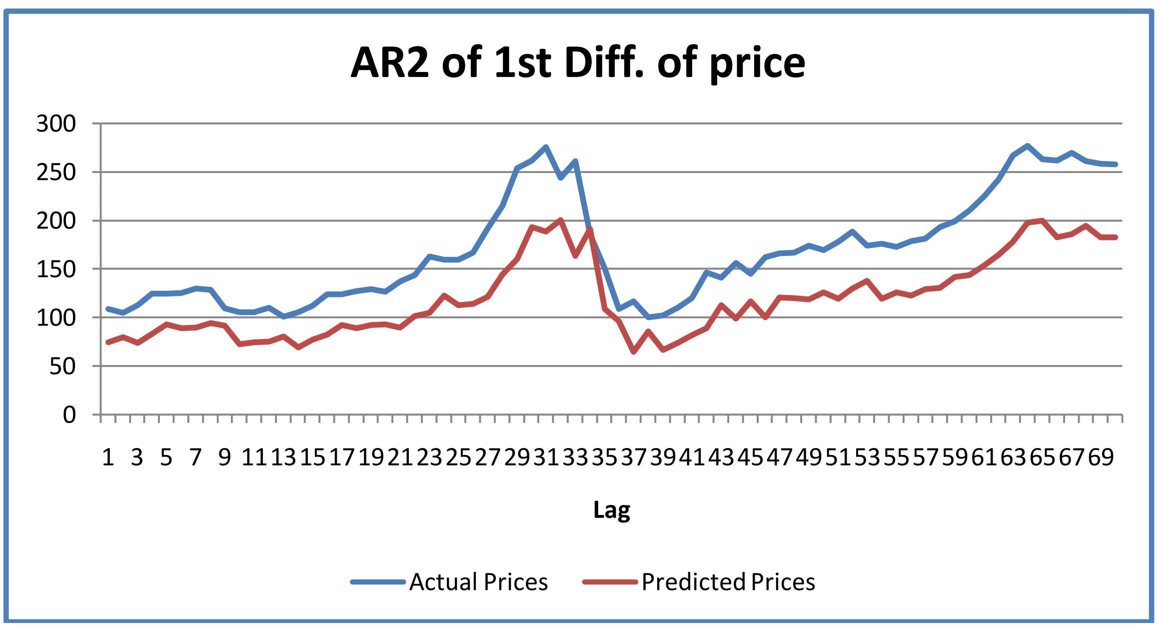
AR (2):

Regression Statistics	
Multiple R	0.243208115
R Square	0.059150187
Adjusted R Square	0.044333655
Standard Error	5.655580522
Observations	130

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.9811	0.5113	1.9190	0.0572	(0.0306)	1.9928	(0.0306)	1.9928
X Variable 1	0.0268	0.0861	0.3113	0.7561	(0.1436)	0.1973	(0.1436)	0.1973
X Variable 2	(0.3185)	0.1130	(2.8183)	0.0056	(0.5421)	(0.0949)	(0.5421)	(0.0949)

The second equation is:

AR(2): $Y_t = 0.0268Y_{t-1} - 0.3185Y_{t-2} + 0.9811 + e_t$



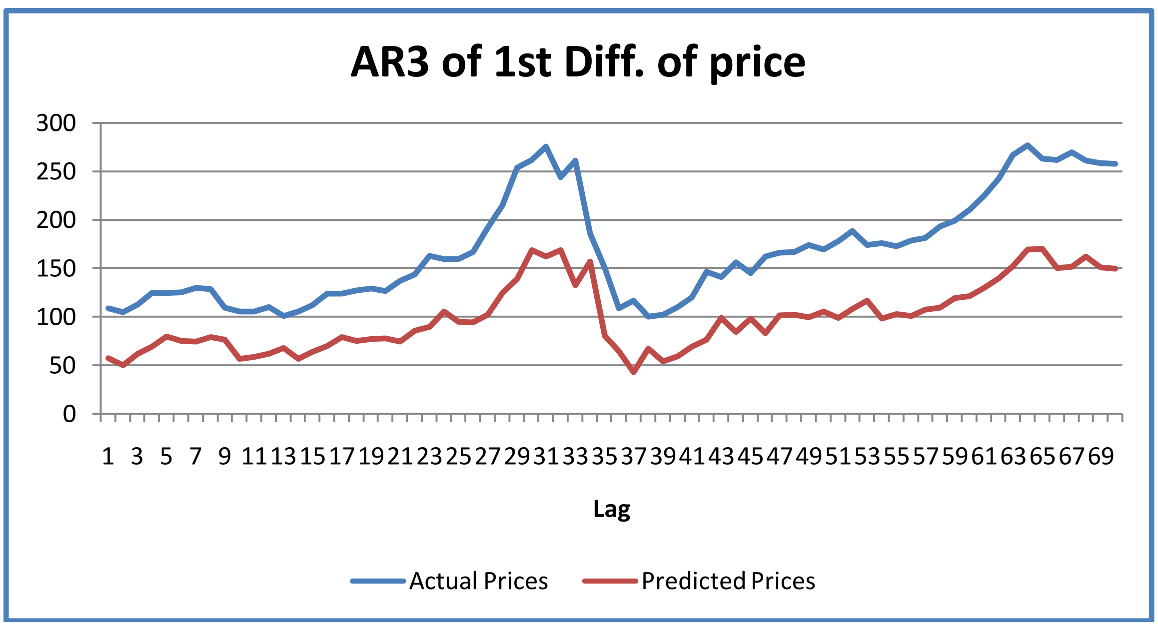
AR(3)

Regression Statistics	
Multiple R	0.259073
R Square	0.067119
Adjusted R Square	0.044729
Standard Error	5.675429
Observations	129

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1.0987	0.5254	2.0912	0.0385	0.0589	2.1384	0.0589	2.1384
X Variable 1	0.0026	0.0896	0.0291	0.9768	(0.1747)	0.1799	(0.1747)	0.1799
X Variable 2	(0.2933)	0.1162	(2.5248)	0.0128	(0.5232)	(0.0634)	(0.5232)	(0.0634)
X Variable 3	(0.1243)	0.1217	(1.0217)	0.3089	(0.3651)	0.1165	(0.3651)	0.1165

The third equation is:

AR(3): $Y_t = 0.0026Y_{t-1} - 0.2933\ Y_{t-2} - 0.1243Y_{t-3} + 1.0987 + e_t$



Results

Table	Sum of Coefficients	R-Squared	Adjusted R-Squared	Durbin-Watson Statistic	Box Pierce	Chi-Square 10%
AR(1)	0.6833	0.000313	-0.00743	1.99227	51.88	147.80
AR(2)	0.6895	0.05915	0.044333	2.04248	41.5595	147.8048
AR(3)	0.6837	0.06712	0.044729	2.01181	41.08312	147.8048

The above table shows that the sum of coefficients for each model is less than 1 which shows that the models are stationary. Moreover, the Durbin-Watson statistic is around 2 for each, suggesting no serial correlation. It is also observed thatthe Box-Pierce Q statistics are lower than Chi-Squared (10%) critical value. The null hypothesis therefore, cannot be rejected which states that the residual are formed by a white noise process.

Selection of Model:

Based on R-squared and Adjusted R-squared, AR(2) and AR(3) is much better than AR(1). Moreover, the value of box pierce test & adjusted R^2 is almost same for AR(2) and AR(3). Based on the Principle Of Parsimony which states given a choice of two almost equivalent model selection, select the simpler one, ***I will select AR(2)***

AR(2): $Y_t = 0.0268Y_{t-1} - 0.3185Y_{t-2} + 0.9811 + e_t$