Julie Rachford
Student Project
Time Series
Winter 2012

Copper Prices

Introduction

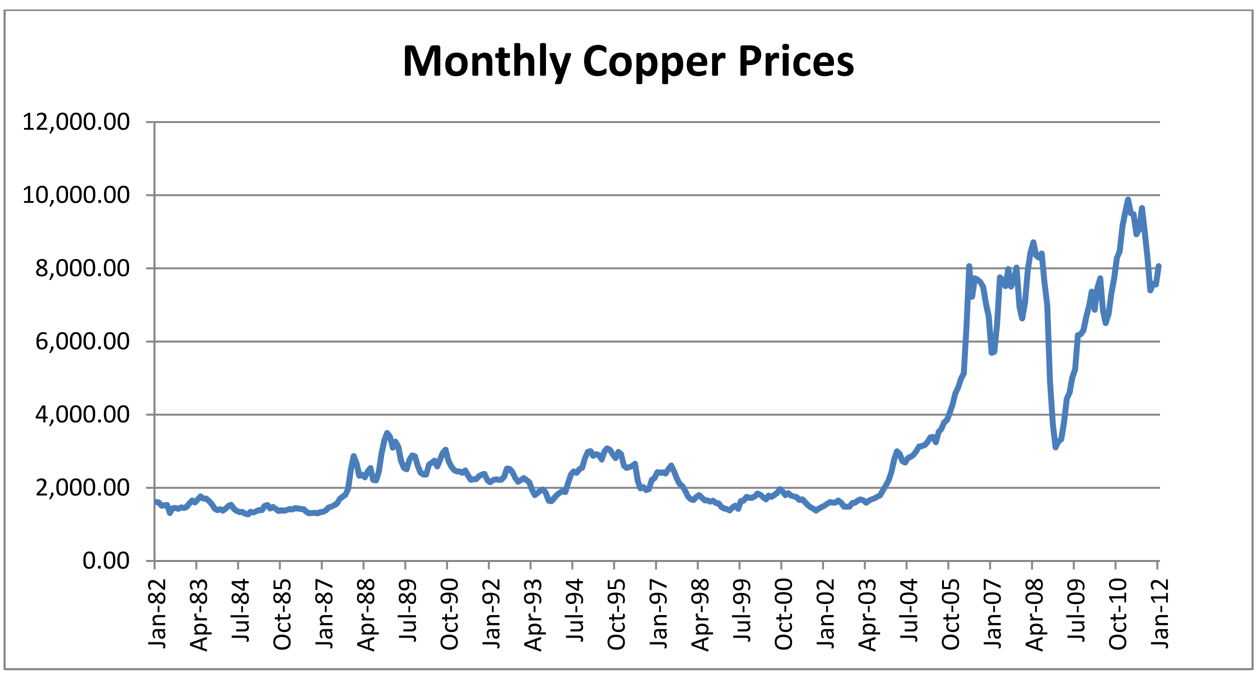
With my husband being an electrician, copper prices are something he pays close attention to. Therefore, I chose to model copper prices using International Monetary Fund prices. I selected various time series techniques to determine which model best fit copper prices. These models include AR(1), AR(2) and AR(3).

Data

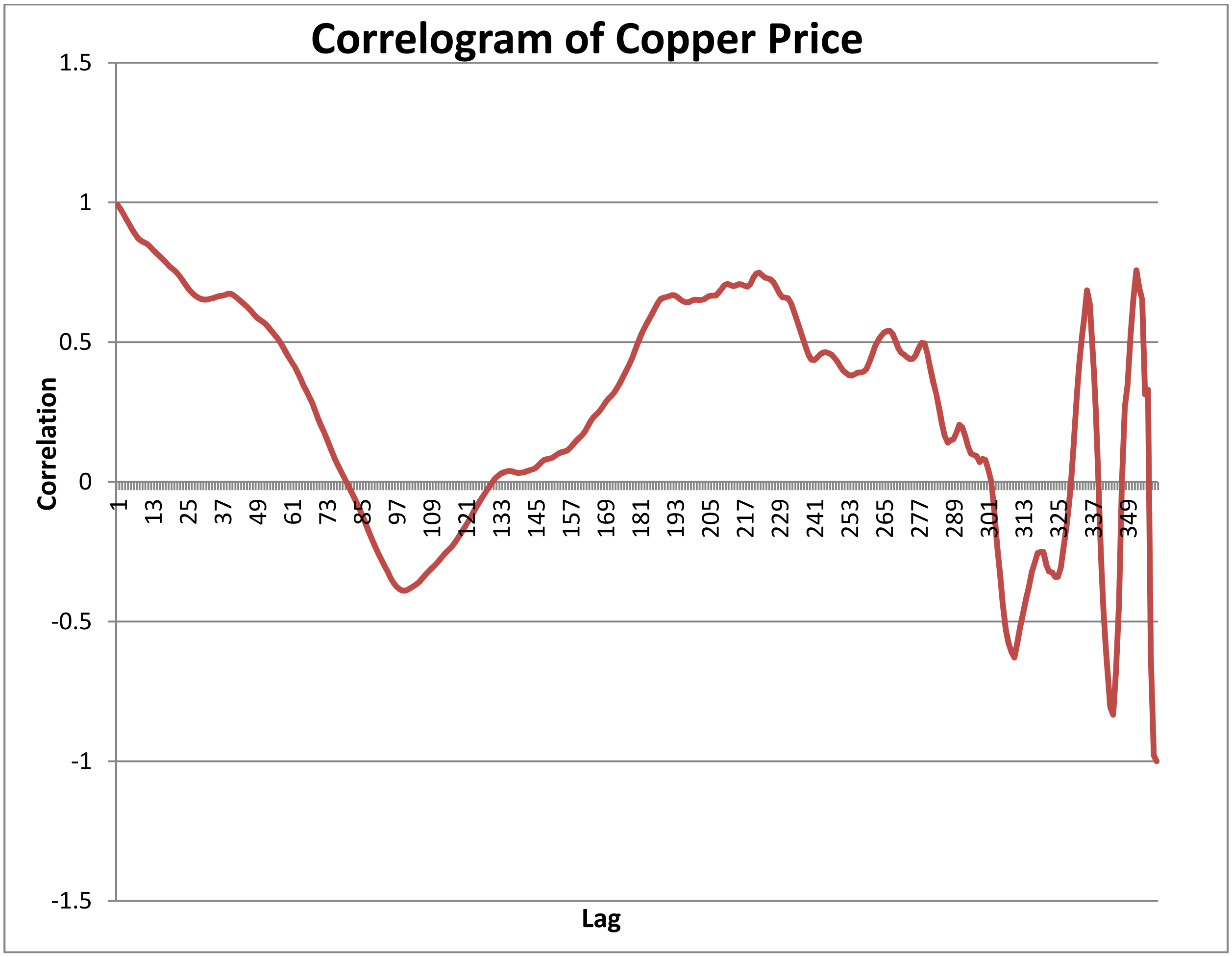
I relied upon the data from the International Monetary Fund to perform my analysis. The data can be found at the following website:

http://www.indexmundi.com/commodities/?commodity=copper&months=360

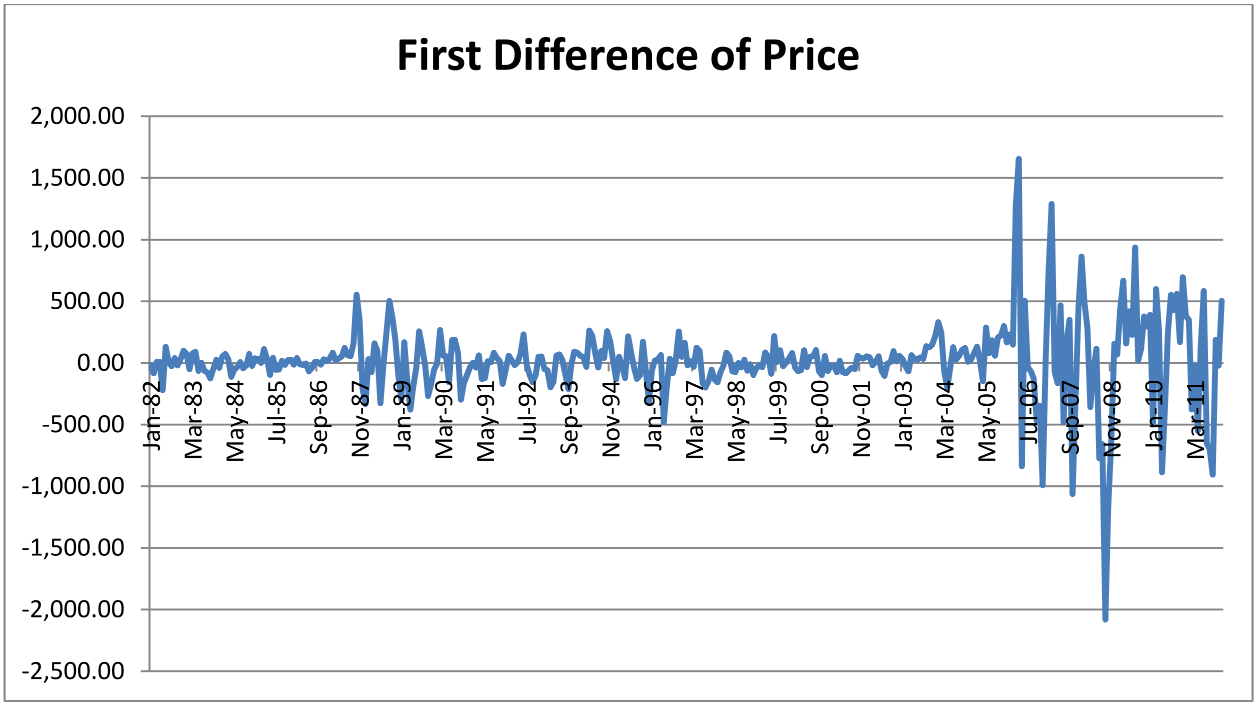
I choose to utilize data compiled on a monthly basis for the last 30 years which is measured in US dollars per metric ton. The data I compiled can be found on the attached excel spreadsheet. Below is a graph of this raw data:

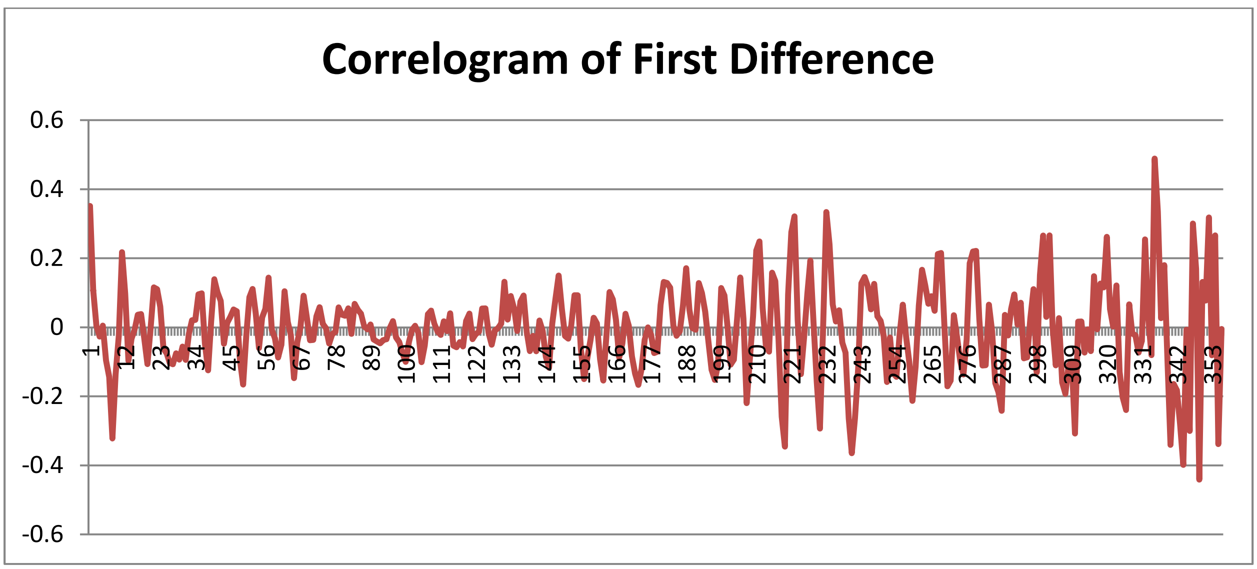


Prior to fitting the data to the ARIMA models, stationarity will need to be verified. This is done by reviewing the sample autocorrelation. Below is a graph of this sample autocorrelation:



Based on my evaluation of this graph, I do not believe that this series is stationary. The correlation does not reach zero until approximately lag 80, stays negative until approximately lag 125 when it remain positive again for quite some time. For a time series to be considered stationary, the correlation must reach zero quickly and must not fluctuate with time (mean and variance are constant). Given this conclusion, I choose to assess the graph of the first difference for stationarity. The following two graphs depict first difference of price and the resultant correlogram:





The correlogram of the first difference looks approximately stationary, as the mean and variance are more or less constant and the correlation reaches zero quickly. Based on the assumption that the first difference is a stationary process, I will now fit the data to the following autoregressive models: AR(1), AR(2) and AR(3).

Parameterization of the Model:

The first difference corresponds to the ARIMA(p,1,0) models, where utilizing Excel's regression data analysis add-in, "p" equates to 1, 2 and 3.

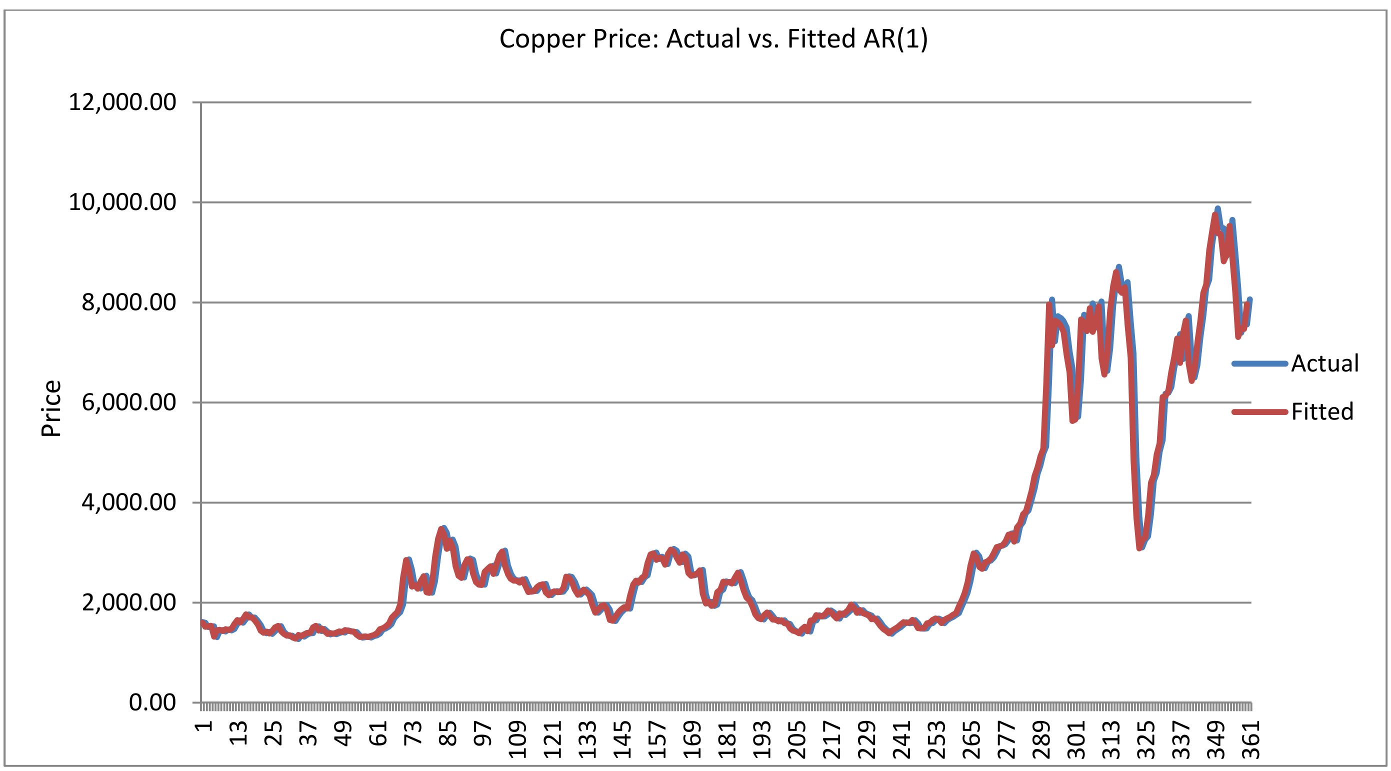
The results of the three autoregressive models are as follows:

AR(1)

Regression Statistics	
Multiple R	0.9907686
R Square	0.9816225
Adjusted R Square	0.9815712
Standard Error	297.81099
Observations	360

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	30.515229	27.0534505	1.127961	0.260092	- 22.68842	83.7188812	-22.688424	83.7188812
Y_0	0.9843555	0.007118382	138.2836	0	0.970356	0.99835462	0.97035643	0.99835462

The AR(1) model parameters: $Y_1 = 30.515229 + 0.9843555Y_{t-1} + e_t$

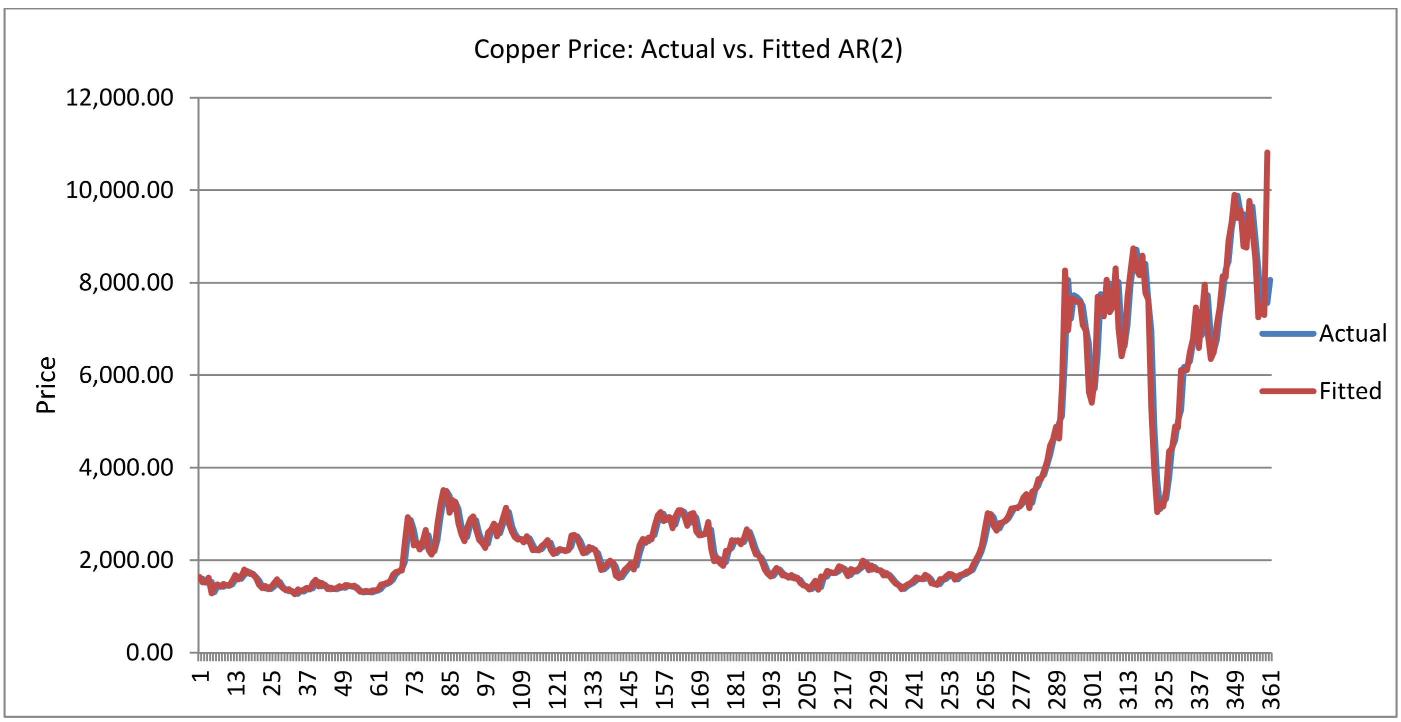


AR(2)

Regression Statistics	
Multiple R	0.991888828
R Square	0.983843447
Adjusted R Square	0.98375268
Standard Error	278.3829262
Observations	359

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	37.4242273	25.37851965	1.47464186	0.1411926	-12.48643645	87.33489106	-12.486436	87.33489106
Y_0	1.33690946	0.049403915	27.0607998	2.134E-88	1.239749252	1.434069664	1.2397493	1.434069664
Y_1	-0.35236824	0.049085219	-7.1787036	4.13E-12	-0.448901678	-0.25583479	-0.4489017	-0.25583479

The AR(2) model parameters: $Y_2 = 37.4242273 + 1.33690946Y_{t-1} + -0.35236824Y_{t-2} + e_t$

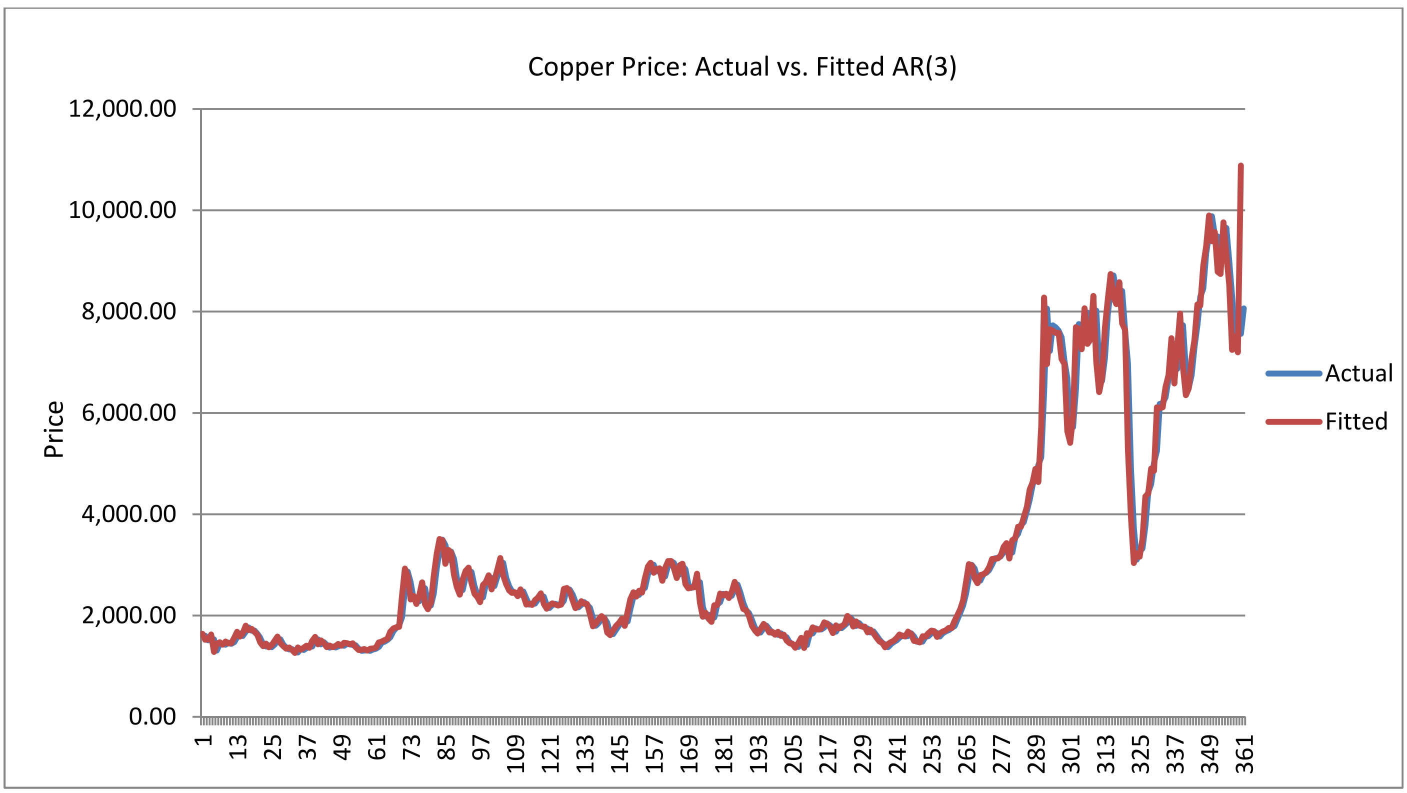


AR(3)

Regression Statistics	
Multiple R	0.991815079
R Square	0.983697151
Adjusted R Square	0.983558991
Standard Error	278.7449653
Observations	358

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	38.5773739	25.53629403	1.510688035	0.13176004	-11.6445447	88.79929245	-11.644545	88.79929245
Y_0	1.3448998	0.053069407	25.34228067	1.5333E-81	1.240528843	1.449270755	1.24052884	1.449270755
Y_1	-0.37308447	0.086497219	-4.31325393	2.0891E-05	-0.5431975	-0.20297144	-0.5431975	-0.20297144
Y_2	0.01207136	0.052595885	0.229511497	0.81860394	-0.09136833	0.115511048	-0.0913683	0.115511048

The AR(3) model parameters: $Y_3 = 38.5773739 + 1.3448998Y_{t-1} + -0.37308447Y_{t-2} + 0.01207136Y_{t-3} + e_t$



Results

Model	Sum of Coefficients	R-Squared	Adjusted R-Squared	Durbin-Watson Statistic
AR(1)	0.984355526	0.981622511	0.981571178	1.075670501
AR(2)	0.984541222	0.983843447	0.98375268	1.915154394
AR(3)	0.983886688	0.983697151	0.983558991	1.931726349

The results of the regression analysis show that the sum of the coefficients for each model is less than 1 indicating that all models are stationary. In addition, the Durbin-Watson statistic is around 2 for AR(2) and AR(3) signifying no serial correlation for these models. Based on these points, the null hypothesis, which states that the residuals are formed by a white noise process, cannot be rejected for models AR(2) and AR(3),

Model Selection

Given the fact that only models AR(2) and AR(3) indicated no serial correlation, and of these two AR(2) has a better R^2 value, I selected the AR(2) model.

$$\mathbf{Y_2 = 37.4242273 + 1.33690946Y_{t-1} + -0.35236824Y_{t-2} + e_t}$$