Introduction

As an avid milk drinker and resident of Virginia, the price of milk at my local grocery store is of interest to me. How much will I have to spend on milk gallons next month? Hot cocoa season is fast approaching; how much will this tasty treat cost to make in the winter months ahead? To help answer these questions, I will analyze historical prices of milk in Virginia. I will use techniques I learned in the Time Series to find a model that best fits this data. Perhaps then I will be able to predict the impact hot cocoa season will have on my pocketbook.

Data

I downloaded the data set of milk prices in Virginia from the USDA National Agricultural Statistics Service via the following website: *http://future.aae.wisc.edu/data/monthly_values/by_area/10?grid=true*

The monthly prices of milk in Virginia from January 1995 through September 2014 are shown below in dollars per hundredweight (cwt). This chart can also be seen in the Excel workbook on the “Milk Prices” tab.

Date	$ / cwt	Date	$ / cwt	Date	$ / cwt	Date	$ / cwt	Date	$ / cwt	Date	$ / cwt	Date	$ / cwt
Jan-95	14.3	Jan-98	15.9	Jan-01	15.9	Jan-04	15.2	Jan-07	16.2	Jan-10	19.1	Jan-13	22.9
Feb-95	13.6	Feb-98	16.6	Feb-01	15.3	Feb-04	15.3	Feb-07	16.9	Feb-10	19.0	Feb-13	22.2
Mar-95	13.5	Mar-98	15.8	Mar-01	16.0	Mar-04	16.3	Mar-07	17.7	Mar-10	18.1	Mar-13	21.6
Apr-95	13.3	Apr-98	15.9	Apr-01	16.5	Apr-04	17.3	Apr-07	17.8	Apr-10	17.0	Apr-13	21.6
May-95	13.5	May-98	15.4	May-01	17.2	May-04	20.2	May-07	19.3	May-10	17.6	May-13	21.6
Jun-95	13.0	Jun-98	15.2	Jun-01	17.8	Jun-04	19.0	Jun-07	21.6	Jun-10	18.9	Jun-13	22.2
Jul-95	13.1	Jul-98	14.3	Jul-01	17.9	Jul-04	20.7	Jul-07	24.1	Jul-10	19.6	Jul-13	22.5
Aug-95	13.6	Aug-98	16.2	Aug-01	18.6	Aug-04	16.2	Aug-07	24.9	Aug-10	19.8	Aug-13	23.0
Sep-95	13.0	Sep-98	16.9	Sep-01	18.9	Sep-04	16.9	Sep-07	24.8	Sep-10	20.5	Sep-13	23.4
Oct-95	13.5	Oct-98	18.6	Oct-01	17.9	Oct-04	17.2	Oct-07	25.3	Oct-10	21.4	Oct-13	24.0
Nov-95	14.0	Nov-98	18.7	Nov-01	17.8	Nov-04	17.2	Nov-07	25.0	Nov-10	21.6	Nov-13	24.9
Dec-95	14.2	Dec-98	15.8	Dec-01	15.0	Dec-04	17.4	Dec-07	23.7	Dec-10	20.6	Dec-13	25.2
Jan-96	14.5	Jan-99	19.2	Jan-02	15.4	Jan-05	17.9	Jan-08	23.8	Jan-11	20.1	Jan-14	25.8
Feb-96	14.6	Feb-99	18.0	Feb-02	15.0	Feb-05	16.4	Feb-08	23.1	Feb-11	21.3	Feb-14	26.9
Mar-96	14.2	Mar-99	17.4	Mar-02	14.5	Mar-05	17.4	Mar-08	20.7	Mar-11	22.6	Mar-14	27.7
Apr-96	14.0	Apr-99	12.8	Apr-02	14.2	Apr-05	16.2	Apr-08	22.1	Apr-11	22.9	Apr-14	27.4
May-96	14.4	May-99	13.7	May-02	13.9	May-05	16.0	May-08	20.8	May-11	22.8	May-14	27.3
Jun-96	15.0	Jun-99	14.2	Jun-02	13.8	Jun-05	15.1	Jun-08	21.9	Jun-11	24.0	Jun-14	26.2
Jul-96	15.7	Jul-99	14.3	Jul-02	13.6	Jul-05	16.0	Jul-08	23.9	Jul-11	24.3	Jul-14	26.6
Aug-96	15.9	Aug-99	14.4	Aug-02	13.9	Aug-05	16.6	Aug-08	22.4	Aug-11	25.2	Aug-14	27.9
Sep-96	16.2	Sep-99	16.1	Sep-02	13.8	Sep-05	16.7	Sep-08	22.2	Sep-11	25.3	Sep-14	28.3
Oct-96	16.8	Oct-99	17.1	Oct-02	13.9	Oct-05	16.9	Oct-08	20.3	Oct-11	23.4		
Nov-96	16.7	Nov-99	17.2	Nov-02	14.0	Nov-05	16.8	Nov-08	20.7	Nov-11	23.2		
Dec-96	15.1	Dec-99	13.8	Dec-02	13.8	Dec-05	16.5	Dec-08	18.0	Dec-11	22.1		
Jan-97	14.2	Jan-00	14.4	Jan-03	13.9	Jan-06	16.2	Jan-09	17.5	Jan-12	22.2		
Feb-97	14.1	Feb-00	13.8	Feb-03	13.6	Feb-06	16.2	Feb-09	14.5	Feb-12	20.1		
Mar-97	14.3	Mar-00	14.2	Mar-03	13.0	Mar-06	15.6	Mar-09	13.7	Mar-12	19.0		
Apr-97	14.2	Apr-00	14.3	Apr-03	12.7	Apr-06	14.0	Apr-09	14.0	Apr-12	18.6		
May-97	14.1	May-00	14.6	May-03	12.9	May-06	14.0	May-09	14.1	May-12	18.7		
Jun-97	13.6	Jun-00	14.8	Jun-03	12.9	Jun-06	13.9	Jun-09	13.6	Jun-12	18.7		
Jul-97	12.7	Jul-00	15.1	Jul-03	13.4	Jul-06	14.1	Jul-09	14.1	Jul-12	19.3		
Aug-97	12.8	Aug-00	15.2	Aug-03	14.4	Aug-06	14.5	Aug-09	14.2	Aug-12	20.0		
Sep-97	13.1	Sep-00	15.4	Sep-03	15.8	Sep-06	14.9	Sep-09	15.2	Sep-12	21.5		
Oct-97	14.4	Oct-00	15.2	Oct-03	16.6	Oct-06	16.0	Oct-09	16.8	Oct-12	23.2		
Nov-97	14.9	Nov-00	15.6	Nov-03	16.4	Nov-06	16.2	Nov-09	17.5	Nov-12	24.7		
Dec-97	15.0	Dec-00	15.2	Dec-03	15.8	Dec-06	16.0	Dec-09	18.6	Dec-12	24.0		

Table 1: Monthly Milk Prices from January 1995 through September 2014

Below is a line graph showing the change in price over time. This graph (and all line graphs following) can also be seen in the Excel workbook on the "Graphs" tab.

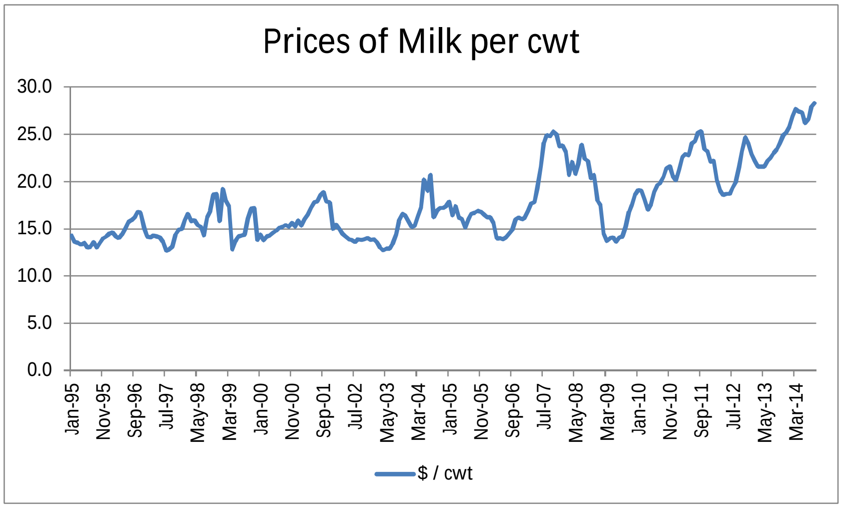

Figure 1: Monthly Milk Prices from January 1995 through September 2014

I graphed the prices over the most recent 5 years to see if there is any seasonality. From Figure 2 below, I determined that there is no seasonality in this time series since there are no apparent price patterns over the 12 month periods.

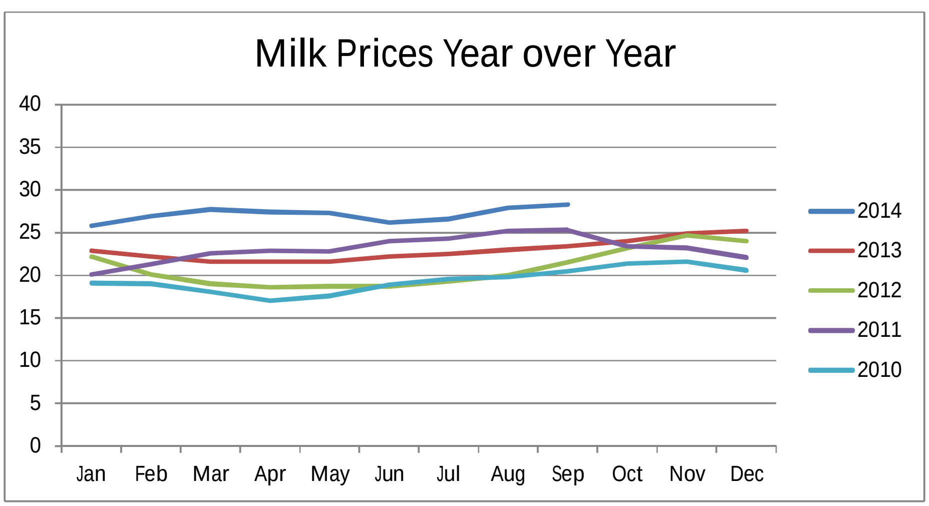

Figure 2: Monthly Milk Prices Year over Year for 2010 through YTD 2014

Analysis

From Figure 1, it looks like the price of milk has been generally increasing over the past 19 years. I hypothesized from this graph that the data is not stationary since the correlogram does not decline rapidly enough for a stationary process. I further tested this by calculating the autocorrelation values using Equation 1 below and by plotting them on the line graph seen in Figure 2.

$$r^2 = \frac{\sum_{t=k+1}^{n}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} \qquad \text{Equation 1}$$

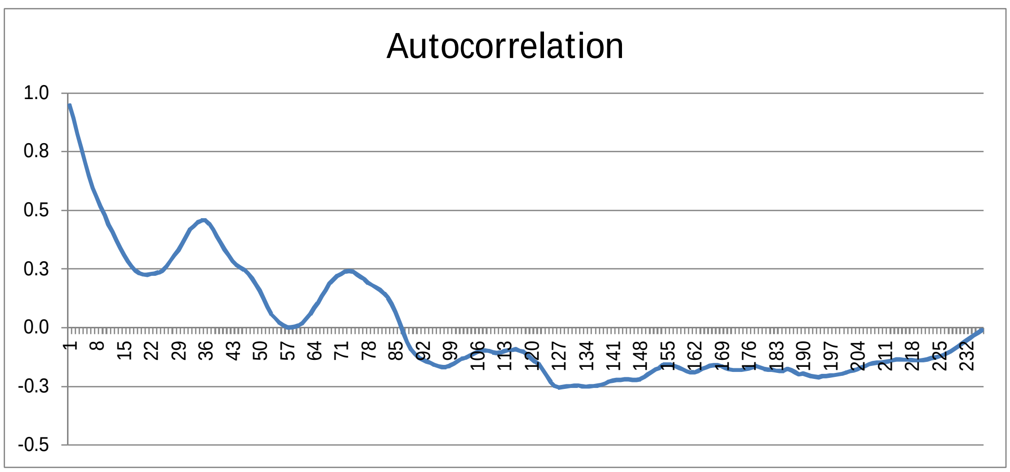

Figure 3: Autocorrelation of Milk Prices

Figure 3 shows that the autocorrelation values drop to zero at lag 87 and then fluctuate below zero before appearing to converge back up to zero at lag 237. This does not suggest stationarity since the values do not drop quickly to zero, and they do not return to zero at subsequent lags with minimum fluctuations. To further test stationarity, I transformed the data by taking first and second differences and plotted their autocorrelations as seen below.

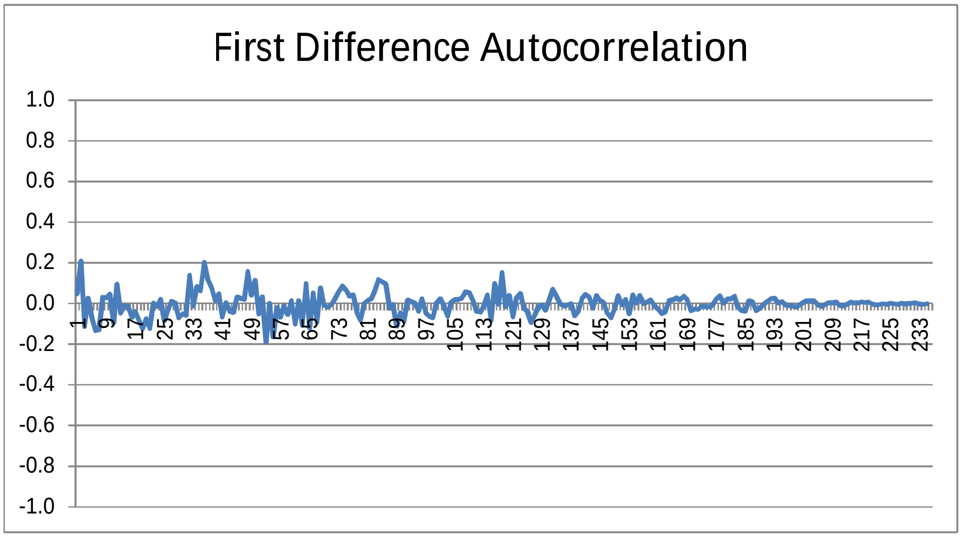

Figure 4: Autocorrelation of First Differences

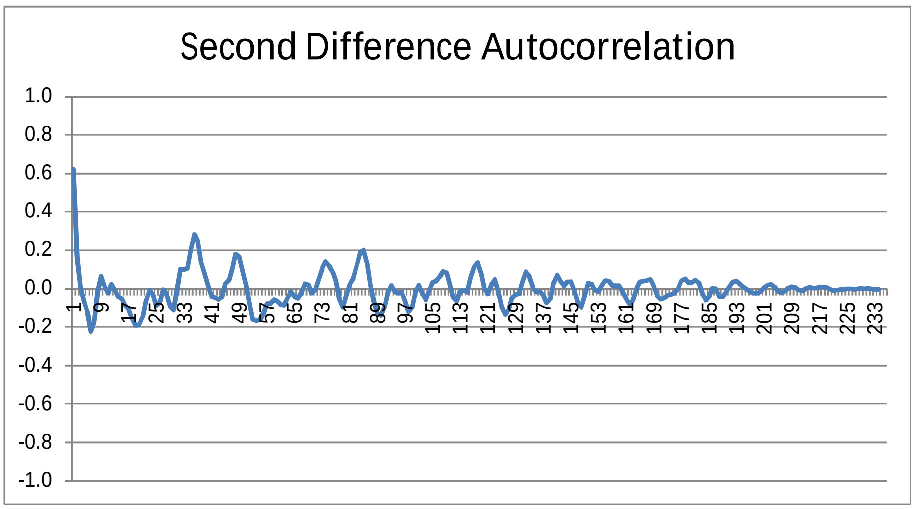

Figure 5: Autocorrelation of Second Differences

In both Figure 4 and Figure 5, the values drop quickly to zero and oscillates around zero. The oscillations in both graphs fluctuate minimally towards the end of the lags. Both of these figures suggest stationarity. I will work with the second differences since the oscillations in Figure 5 fluctuate less than the oscillations in Figure 4.

I then considered three different models and fit them the data set: AR(1), AR(2), and AR(3). I will compare the R^2 values and the standard error values to determine which model is best for this data.

For the AR(1) model, I used Equation 2 and Excel's built-in regression function. The work done for this model can be found in the Excel workbook on the "Analysis" tab and the "AR(1)" tab.

$$Y_t - Y_{t-1} = \varepsilon_t + \phi_1(Y_{t-1} - Y_{t-2}) \qquad \text{Equation 2}$$

Regression Statistics	
Multiple R	0.622854379
R Square	0.387947577
Adjusted R Square	0.38530942
Standard Error	1.17435072
Observations	234

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	202.8000436	202.8000436	147.0524984	1.54504E-26
Residual	232	319.9511102	1.379099613		
Total	233	522.7511538			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.053443529	0.076992121	0.694142835	0.488287115	-0.09824958	0.205136636	-0.09824958	0.205136636
X Variable 1	0.623842635	0.051444488	12.12652046	1.54504E-26	0.522484548	0.725200723	0.522484548	0.725200723

Figure 6: Summary Output for AR(1)

Based on the summary output in Figure 6 above, Equation 2 becomes the following:

$$Y_t - Y_{t-1} = .05344 + 0.6238(Y_{t-1} - Y_{t-2}) \qquad \text{Equation 3}$$

For the AR(2) model, I used Equation 4 and Excel's built-in regression function. The work done for this model can be found in the Excel workbook on the "Analysis" tab and the "AR(2)" tab.

$$Y_t - Y_{t-1} = \varepsilon_t + \phi_1 (Y_{t-1} - Y_{t-2}) + \phi_2 (Y_{t-2} - Y_{t-3})$$ Equation 4

Regression Statistics	
Multiple R	0.682998079
R Square	0.466486376
Adjusted R Square	0.461847127
Standard Error	1.10098449
Observations	233

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	243.7719255	121.8859627	100.5521335	4.18537E-32
Residual	230	278.798375	1.212166848		
Total	232	522.5703004			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.064988925	0.072381215	0.897870054	0.370193755	-0.077626082	0.207603932	-0.077626082	0.207603932
X Variable 1	0.846888695	0.06158497	13.75154831	8.3701E-32	0.725545871	0.968231519	0.725545871	0.968231519
X Variable 2	-0.359324088	0.061685172	-5.825129084	1.91098E-08	-0.480864342	-0.237783834	-0.480864342	-0.237783834

Figure 7: Summary Output for AR(2)

Based on the summary output in Figure 7 above, Equation 4 becomes the following:

$$Y_t - Y_{t-1} = .06499 + 0.84689(Y_{t-1} - Y_{t-2}) – 0.35932(Y_{t-2} - Y_{t-3})$$ Equation 5

For the AR(3) model, I used Equation 6 and Excel's built-in regression function. The work done for this model can be found in the Excel workbook on the "Analysis" tab and the "AR(3)" tab.

$$Y_t - Y_{t-1} = \varepsilon_t + \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \phi_3(Y_{t-3} - Y_{t-4}) \qquad \text{Equation 6}$$

Regression Statistics	
Multiple R	0.692632543
R Square	0.47973984
Adjusted R Square	0.472894311
Standard Error	1.091964649
Observations	232

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	250.690121	83.56337366	70.08076069	3.74985E-32
Residual	228	271.8641894	1.192386795		
Total	231	522.5543103			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.056075552	0.072058504	0.778194791	0.437261	-0.085910195	0.1980613	-0.085910195	0.1980613
X Variable 1	0.903235993	0.065398226	13.81132247	6.20046E-32	0.774373811	1.032098176	0.774373811	1.032098176
X Variable 2	-0.49380211	0.082781221	-5.965146497	9.24631E-09	-0.656916151	-0.330688074	-0.656916151	-0.330688074
X Variable 3	0.158654261	0.065836649	2.409816781	0.016754036	0.028928201	0.288380321	0.028928201	0.288380321

Figure 8: Summary Output for AR(3)

Based on the summary output in Figure 8 above, Equation 6 becomes the following:

$$Y_t - Y_{t-1} = .05608 + 0.90324(Y_{t-1} - Y_{t-2}) - 0.49380(Y_{t-2} - Y_{t-3}) + 0.1587(Y_{t-3} - Y_{t-4}) \qquad \text{Equation 7}$$

Let us now summarize and compare the R^2 values and the standard error values of each model.

	AR(1)	AR(2)	AR(3)
R^2	0.388	0.466	0.480
Adjusted R^2	0.385	0.462	0.473
Standard Error	1.174	1.101	1.092

Table 2: Results of the AR(p) Models

Since the R^2 value and adjusted R^2 value of the AR(3) model are closest to one, and since the standard error of the AR(3) model is the lowest, I will use AR(3) to predict milk prices.

I then fit the AR(3) model to the data and compared the predicted prices to the actual prices as shown below.

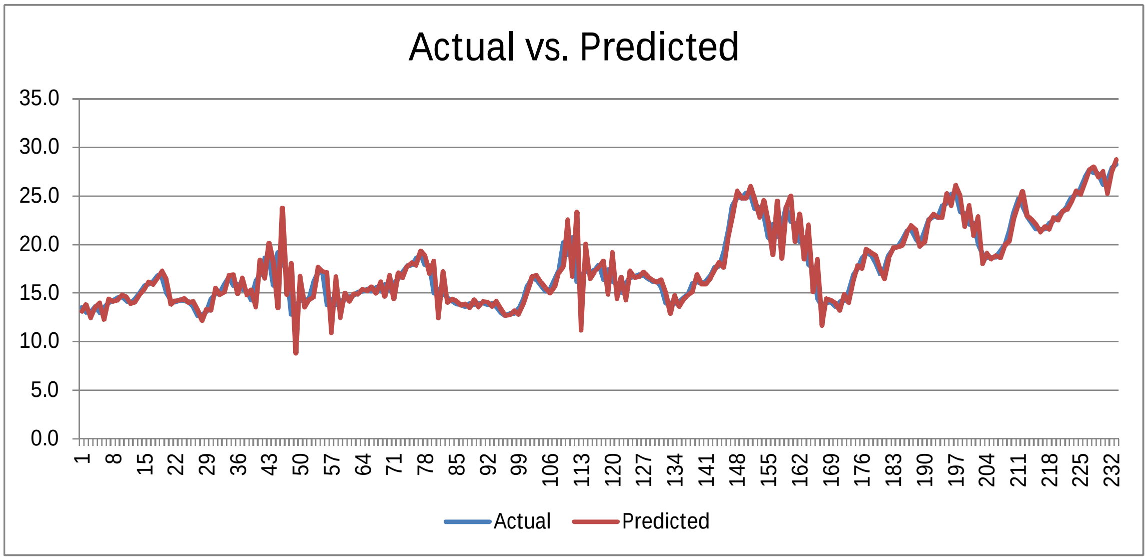

Figure 9: Results of the AR(3) Model Compared to Actual Prices

Figure 9 illustrates that the AR(3) model is a good fit for this data. The predicted monthly milk prices fall near in line to the actual monthly milk prices over the past 19 years.

Conclusion

The time series composed of historical monthly prices of milk in Virginia is not a stationary series. In order to make it stationary, I transformed it by taking first and second differences. I then considered the AR(1), AR(2), and AR(3) models to see which one would best fit the series. Using the general formulas for each model as well as the Excel regression add-in, I fit each model to the data. I then compared the R^2 values and standard error values and determined that the AR(3) model would best fit the series. I then used the AR(3) formula

$$\boldsymbol{Y_t = Y_{t-1} + .05608 + 0.90324(Y_{t-1} - Y_{t-2}) - 0.49380(Y_{t-2} - Y_{t-3}) + 0.1587(Y_{t-3} - Y_{t-4})}$$

to predict milk prices in Virginia. Figure 9 above shows that the AR(3) model is indeed a good fit.