Math 265 Final Project

Presented to

Dr. Steven M. Crunk

Department of Mathematics & Statistics

San José State University

Scott Li

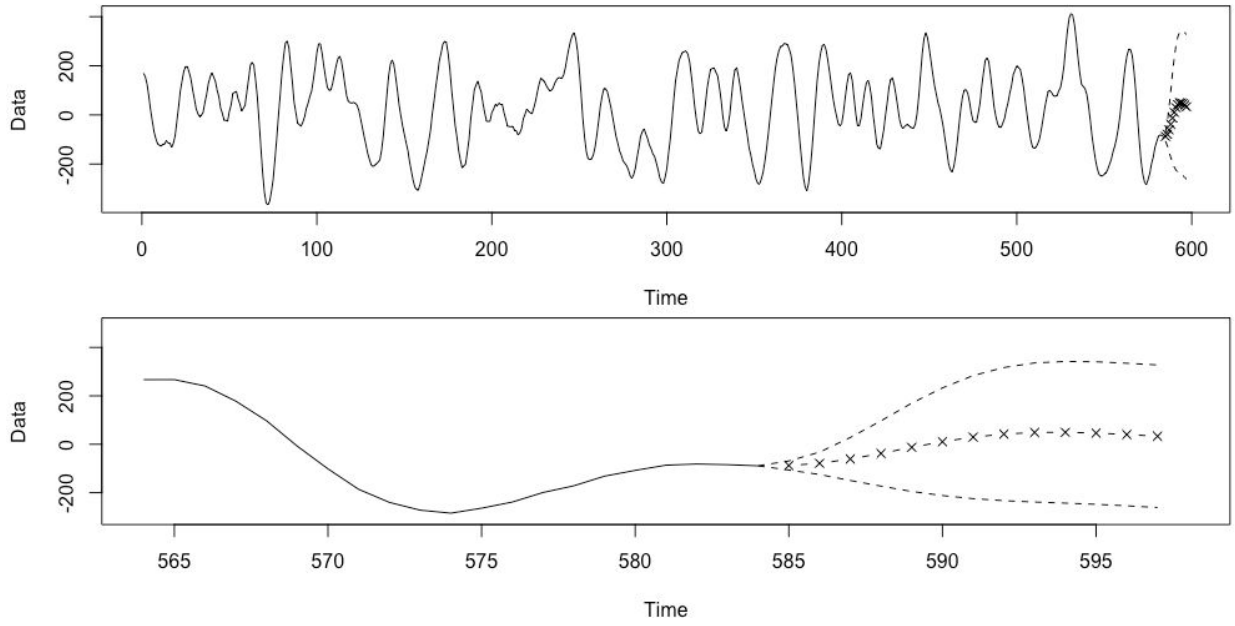
Fall 2017

Executive Summary for Dataset 1

The goal for this dataset is to fit a model to the data and predict the next thirteen values. Dataset 1 is a 584 observation time series of with unknown units for the data and time. The chosen model is a 5th order autoregressive model with a zero mean. The estimated coefficients and their standard errors are presented below.

Coefficient	α_1	α_2	α_3	α_4	α_5
Estimate	2.2609	-1.0560	-1.3543	1.6397	-0.5187
S.E.	0.0352	0.0729	0.0635	0.0730	0.0353

The entire dataset with a forecast thirteen time units ahead is plotted below, with a closer view of the forecast. The forecasted values are shown with X's and the future values are likely to stay within the dashed lines. There is no long term trend with this data. There is a moderately strong cyclical pattern about every 27 time units. There are other periodicities as well but this one is the strongest.



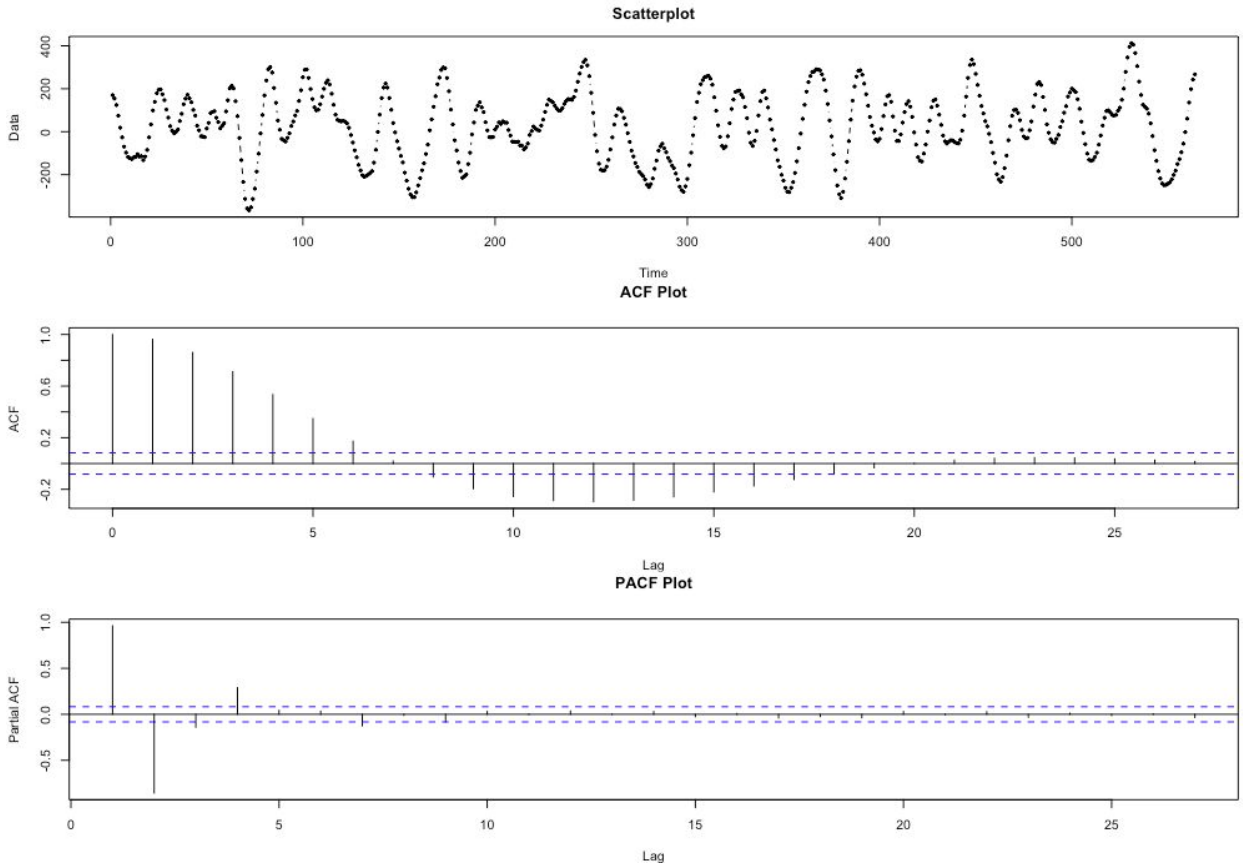
The predicted values are also tabled below with their upper and lower prediction bounds.

Time	585	586	587	588	589	590	591	592	593	594	595	596	597
Upper	-68.9	-32.5	27.9	96.8	169.8	232.8	283.8	317.0	335.9	341.9	340.6	334.6	327.7
Value	-87.6	-78.8	-60.9	-38.1	-12.6	10.5	29.6	42.0	48.7	49.3	46.3	40.2	33.2
Lower	-106.3	-125.0	-149.7	-172.9	-195.0	-211.9	-224.7	-233.0	-238.6	-243.2	-248.0	-254.2	-261.3

The cyclical nature of the data is projected to continue with projections showing a rise to about 50 units in 10 time steps, then a fall to 30 units. However, the prediction interval is quite wide. The data can end up anywhere between -260 to 330 units after 13 time steps.

Technical Appendix for Dataset 1

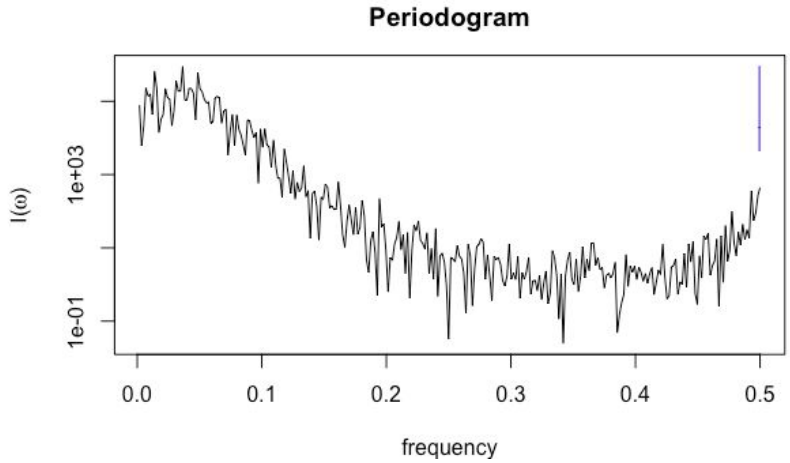
It is known this dataset, a time series of 584 observations, is generated from an unknown ARIMA(p,d,q) process. Thus, as an extra source of model validation, I will holdout the last 20 observations to compute sum of squared prediction errors. The first 564 observations will be examined first. All analysis is done in R.



The scatterplot shows no signs of a trend and at least one periodicity. The Augmented Dicky-Fuller test shows that the data is stationary, with a p-value of 0.01 (at most). The ACF plot appears to have sinusoidal decay. The PACF plot cuts off at lag 4 or lag 7. This suggests an AR(4) or AR(7) model. The raw periodogram, on the next page, shows a two peaks – one at around 0.05 and one at 0.5. This suggests an AR model with a minimum order of 3.

Automated AR model fitting procedures are also tried on the data. The Yule-Walker, Burg, and Ordinary Least Squares fitting methods suggest an AR(7) model. The Maximum Likelihood Estimation method suggests an AR(5) model.

Thus, the candidate models to test are AR models of order 4 to order 7.



An examination of the residual diagnostic plots and the significance of the coefficients show that AR(5) is the only model out of the four without major violations to the model fit. Just in case there is an MA component that is difficult to see from the plots, I fit some ARMA(4,q) and ARMA(5,q) models for orders of q from 1 to 6. From these models, ARMA(4,3) and ARMA(4,4) show no major violations to the model fit. Thus, it is time to examine some model selection metrics for these final candidates, shown in the following table.

Model	Validation SSE	σ^2	AIC	BIC
AR(5)	546518	90.7	4167.2	4197.5
ARMA(4,3)	600657	90.9	4172.6	4211.6
ARMA(4,4)	578493	90.2	4170.6	4213.9

A comparison between the model spectrums and the raw periodogram is also done. The three are very similar but the selection metrics for the AR(5) model appear to be the best. The AR(5) model has the lowest validation SSE, AIC, and BIC so it is a clear choice for the final model.

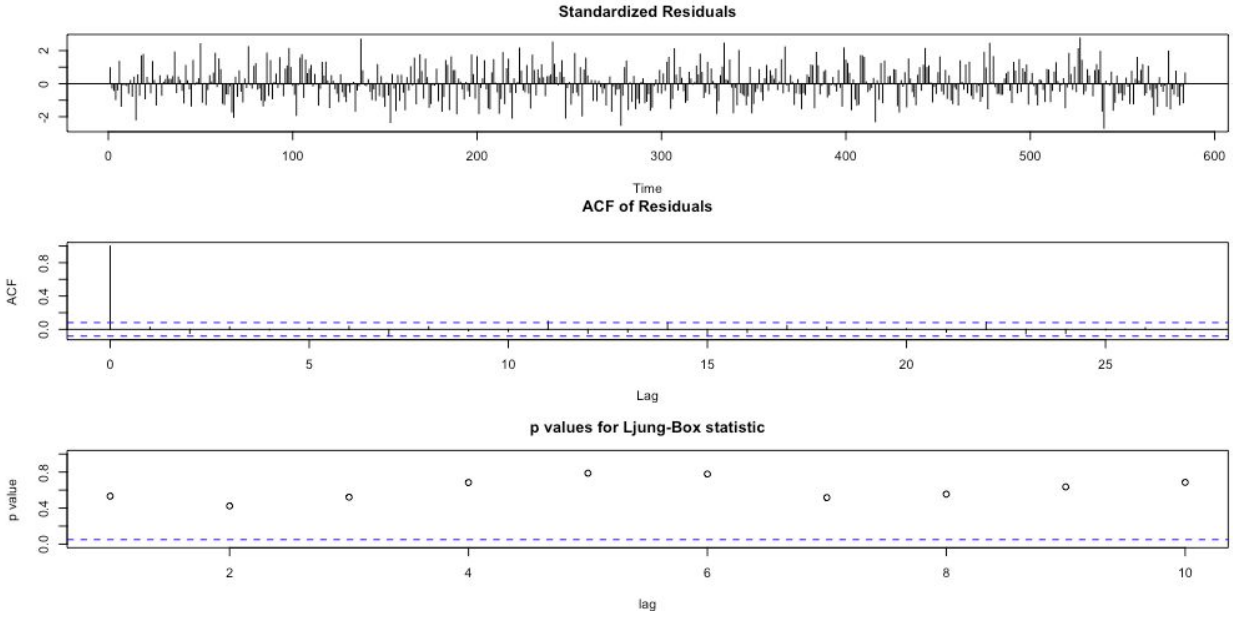
The AR(5) model is fit to the full data with the default methods of the ARIMA function. The mean is not significant so the model was fitted with zero mean.

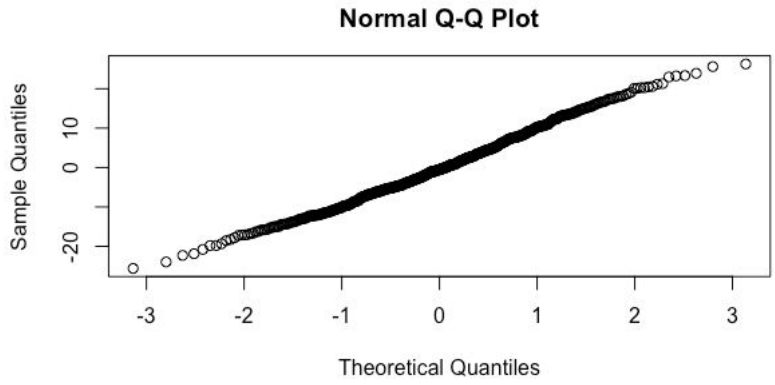
Coefficient	α_1	α_2	α_3	α_4	α_5
Estimate	2.2609	-1.0560	-1.3543	1.6397	-0.5187
S.E.	0.0352	0.0729	0.0635	0.0730	0.0353

All of the coefficients are significant. Then, residual diagnostics are checked again (on the next page). There is no indication of a lack-of-fit here since the residuals look random and

there is no evidence of autocorrelation of the residuals. The normal Q-Q plot of the residuals looks adequate as well, since it is a relatively straight line.

Thus, forecasting is done with the predict function for the next 13 observations. A rough 95% prediction interval is calculated using the prediction standard errors.



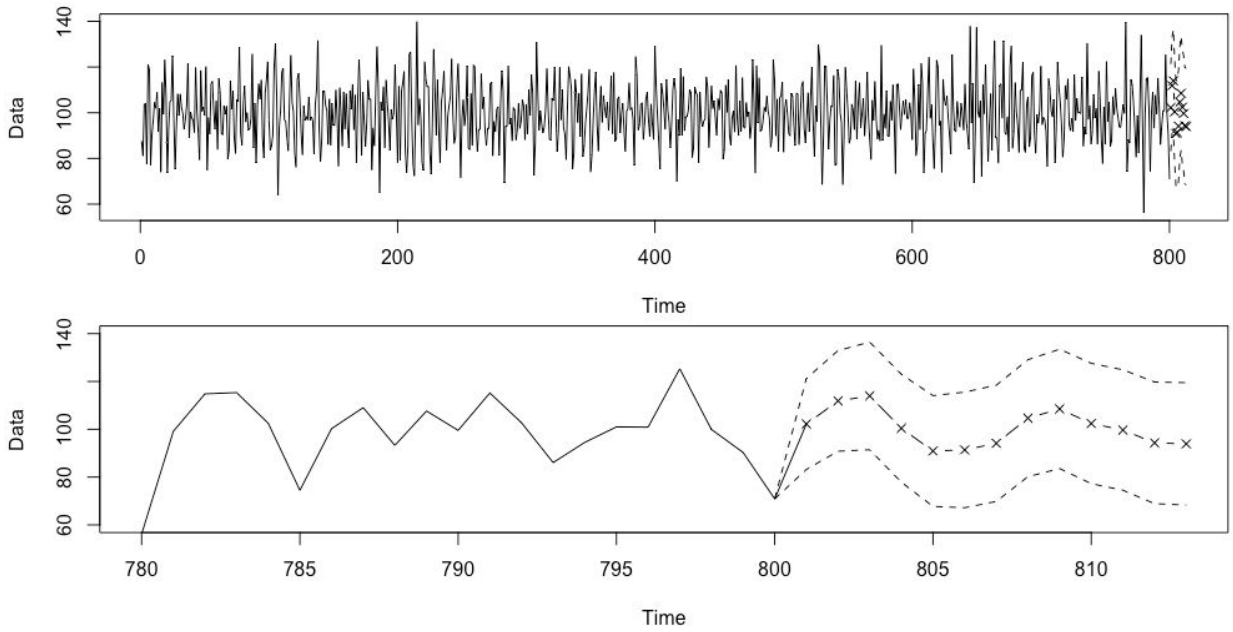


Executive Summary for Dataset 2

The goal for this dataset is to fit a model to the data and predict the next thirteen values. Dataset 2 is an 800 observation time series of with unknown units for the data and time. The chosen model is an ARIMA(4,0,4) with a non-zero mean. The estimated coefficients and their standard errors are presented below.

Coefficient	mean	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
Estimate	99.75	-0.26	-0.44	-0.03	-0.82	-0.21	-0.10	-0.36	0.88
S.E.	0.16	0.04	0.03	0.03	0.03	0.04	0.04	0.05	0.04

The entire dataset with a forecast thirteen time units ahead is plotted below, with a closer view of the forecast. The forecasted values are shown with X's and the future values are likely to stay within the dashed lines. The data has no clear trend with an average value of 99.7 and a median value of 99.4. There is a moderately strong cyclical pattern about every 6.7 time units.



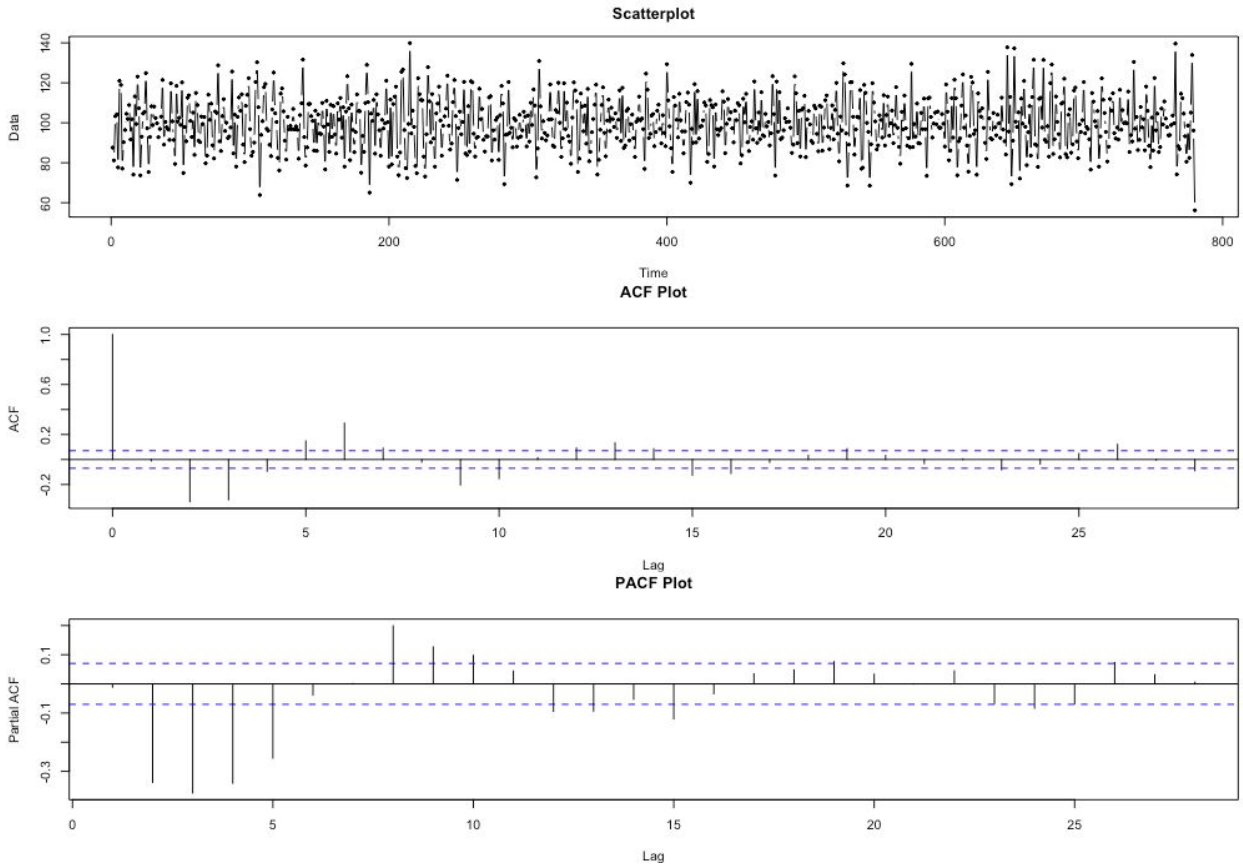
The predicted values are also tabled below with their upper and lower prediction bounds.

Time	801	802	803	804	805	806	807	808	809	810	811	812	813
Upper	121.28	132.90	136.47	122.99	114.08	115.53	118.44	129.09	133.47	127.58	124.86	119.75	119.49
Value	102.23	111.85	113.95	100.42	90.90	91.36	94.13	104.63	108.52	102.41	99.67	94.32	93.90
Lower	83.18	90.81	91.42	77.85	67.73	67.18	69.82	80.18	83.56	77.25	74.48	68.88	68.31

In summary, the data has cyclical fluctuations that are expected to continue, according to the forecast. The data is projected to rise to about 114 units in 3 time steps, fall to 91 units in 5 time steps, rise to 109 in 9 time steps, and fall again to 94 units. The prediction interval for this forecast is about 50 units wide so fluctuations within 50 units can be expected for these forecasted values.

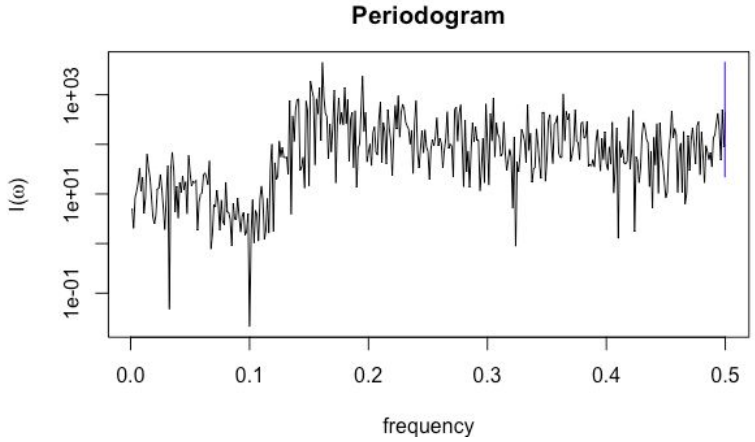
Technical Appendix for Dataset 2

It is known this dataset, a time series of 800 observations, is generated from an unknown ARIMA(p,d,q) process. Thus, as an extra source of model validation, I will holdout the last 20 observations to compute sum of squared prediction errors. The first 780 observations will be examined first. All analysis is done in R.



The scatterplot shows no trend. There may be an outlier at time 780 but in this context, I know it is due to random chance. The Augmented Dicky-Fuller test shows that the data is stationary, with a p-value of 0.01 (at most). The ACF plot looks like sinusoidal decay. The PACF plot also looks like sinusoidal decay. Where they cut off is hard to determine, but it looks like there is an AR and MA component.

The periodogram (on the next page), looks like there is a dip at around 0.1, and a peak at around 0.17. There may be other features that are covered by noise. Thus there is probably an AR component of order 2 minimum and an MA component of order 2 minimum.



The auto.arima() function from the "forecast" package is also used. This function implements the Hyndman-Khandakar algorithm for automated ARIMA model fitting. This function suggests an ARMA(2,2) model. This is a good starting point. From the ACF and PACF plots, the maximum orders I want to try are 5 for AR and 4 for MA.

After fitting all the models and examining the residual diagnostics and the significance of the coefficients, there are three models remaining: ARMA(2,2), ARMA(3,3), and ARMA(4,4). I also try out some higher order models and ARMA(5,5) and ARMA(6,6) pass these criteria as well. The some model selection metrics for these models are tabled below.

Model	Validation SSE	σ^2	AIC	BIC
ARMA(2,2)	2606.71	96.55	5793.53	5821.48
ARMA(3,3)	2622.37	96.04	5793.55	5830.83
ARMA(4,4)	2554.00	95.04	5790.37	5836.96
ARMA(5,5)	2524.41	94.47	5791.73	5847.64
ARMA(6,6)	2797.74	93.72	5791.52	5856.75

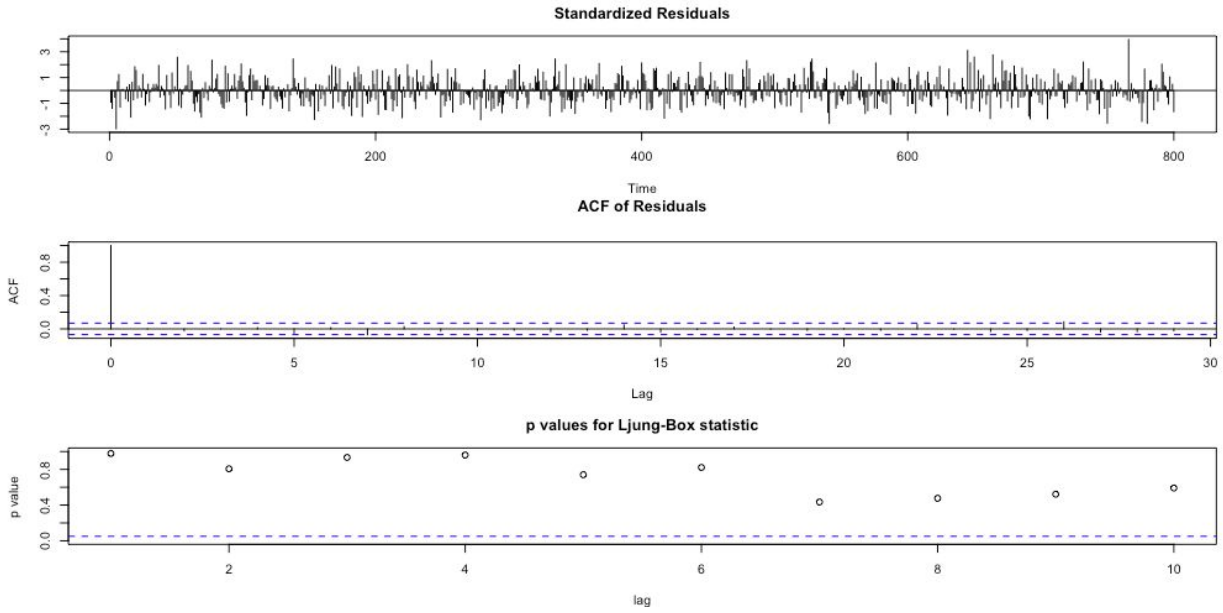
The spectrum plots are also compared to the raw periodogram. The spectrum for the ARMA(3,3) is quite off so this model may be unacceptable for the data. All other models except ARMA(2,2) may be overfitting since their spectrum plots show short spikes or dips that the raw periodogram does not have, from the 0.2 to 0.5 frequency. However, simulating random datasets with these models all produce similar periodograms since sharp peaks are easily covered by noise. The confidence interval for the periodogram makes me suspect that the activity from the 0.2 to 0.5 region is just noise. A comparison of the theoretical ACF and PACF values with the estimated ones narrows down the choices a bit but this risks fitting to noise, although the holdout method risks this as well. From this, I need to pick between ARMA(2,2) and ARMA(4,4). Based on the evaluation metrics,

ARMA(4,4) is only worse in the BIC. Thus I will choose ARMA(4,4) as the final model, with a slight risk of overfitting.

This model is fit to the full data. Some of the standard errors cannot be estimated with the CSS-ML method. Thus, the model is fit using a CSS method to calculate the standard errors, but the coefficients for the CSS-ML model are reported since the log likelihood is slightly better. They may not correspond perfectly but the variance for the CSS method is likely to be greater than the variance from the CSS-ML method, if the CSS-ML method is the default. The high order coefficients are significant.

Coefficient	mean	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
Estimate	99.7526	-0.2581	-0.4421	-0.0300	-0.8230	-0.2112	-0.1004	-0.3618	0.8804
S.E.	0.1605	0.0380	0.0340	0.0329	0.0291	0.0389	0.0426	0.0498	0.0416

The residual diagnostics are also checked. There are no major violations in these. There is no evidence of autocorrelation of the residuals. Also, the Q-Q plot looks adequate. Thus, forecasting is done with the predict function for the next 13 observations with a rough 95% prediction interval.

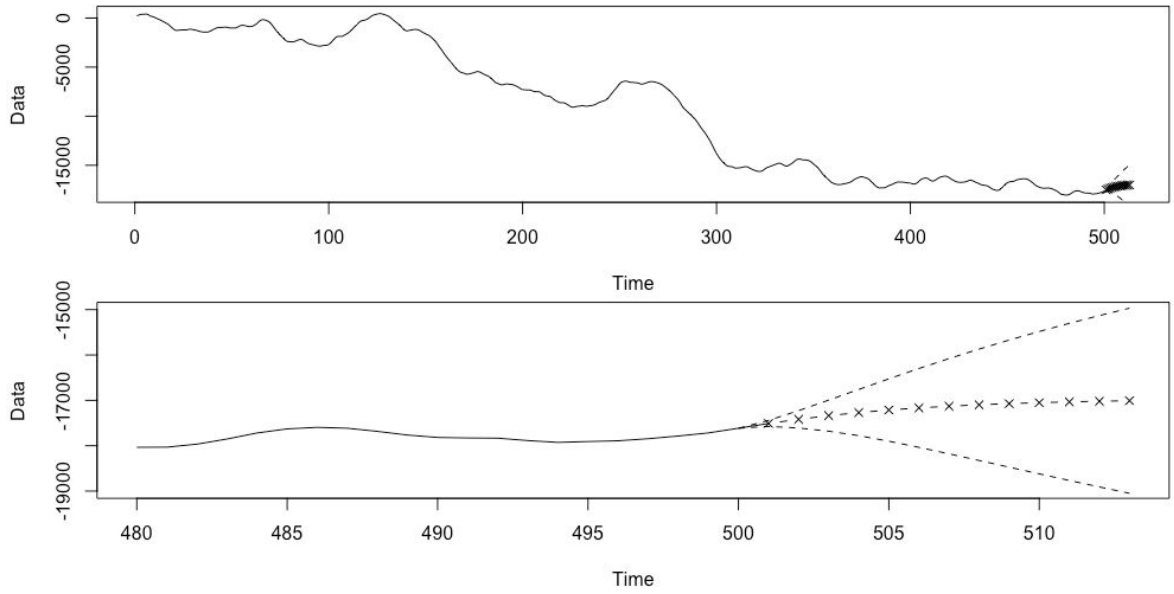


Executive Summary for Dataset 3

The goal for this dataset is to fit a model to the data and predict the next thirteen values. Dataset 3 is a 500 observation time series of with unknown units for the data and time. The chosen model is an ARIMA(1,1,4) with zero-mean. The estimated coefficients and their standard errors are presented below.

Coefficient	α_1	β_1	β_2	β_3	β_4
Estimate	0.8216	0.9185	0.1472	0.1176	0.1312
S.E.	0.0397	0.0590	0.0903	0.0809	0.0512

The entire dataset with a forecast thirteen time units ahead is plotted below, with a closer view of the forecast. The forecasted values are shown with X's and the future values are likely to stay within the dashed lines. There is an obvious long-term trend of about 35.6 units of decrease per time unit on average. There is no consistent cyclical fluctuation in the data.



The predicted values are also tabled below with their upper and lower prediction bounds. These values are in thousands.

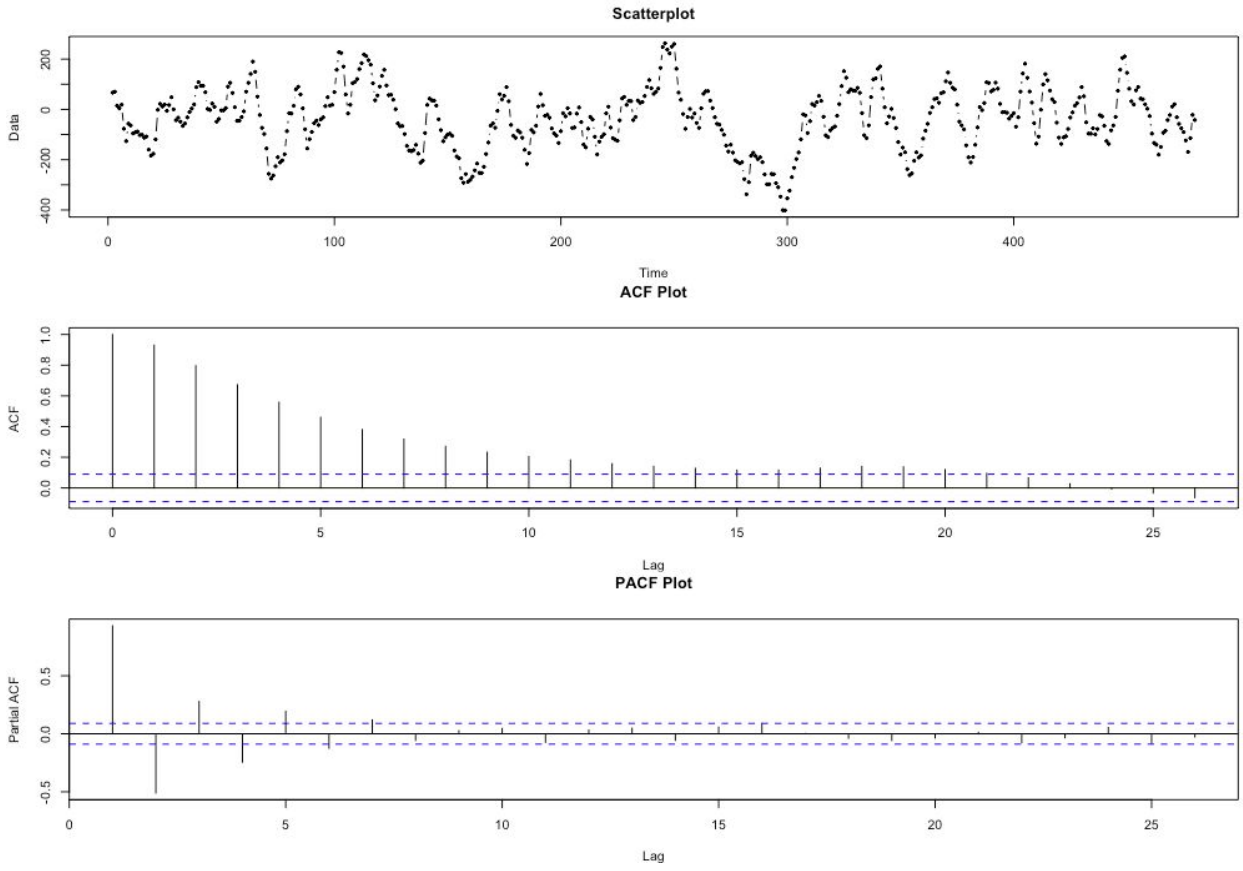
Time	501	502	503	504	505	506	507	508	509	510	511	512	513
Upper	-17.44	-17.23	-16.99	-16.76	-16.53	-16.30	-16.08	-15.87	-15.67	-15.48	-15.30	-15.13	-14.97
Value	-17.51	-17.42	-17.34	-17.27	-17.21	-17.17	-17.13	-17.10	-17.07	-17.05	-17.03	-17.02	-17.01
Lower	-17.57	-17.61	-17.68	-17.78	-17.90	-18.04	-18.18	-18.33	-18.47	-18.62	-18.77	-18.91	-19.05

In the next 13 time units, the data is projected to be relatively flat, slightly increasing from about -17500 to about -17000. However, the uncertainty of these predictions increases rapidly. At the end of 13 time steps, the data may reach an all time low of -19000 or jump about 2500 units to -15000.

Technical Appendix for Dataset 3

It is known this dataset, a time series of 500 observations, is generated from an unknown ARIMA(p,d,q) process. Thus, as an extra source of model validation, I will holdout the last 20 observations. The first 480 observations will be examined first. All analysis is done in R.

A plot of the data shows an obvious trend. The Augmented Dicky-Fuller test shows that the data is not stationary with a p-value of 0.1859. Thus the data needs to be differenced. After taking the first differences, the Dicky-Fuller test shows stationarity.



The scatterplot of the differenced data looks stationary. The ACF plot looks like it just decays. The PACF plot cuts off at lag 7 maybe. The periodogram of the differenced data shows one downward sloping trend from 0 to 0.5, which is not very informative given the ACF and PACF patterns. The Hyndman-Khandakar algorithm suggests an ARIMA(2,1,2) but I will search through more models. So first, I try some ARIMA(p, 1, 0) models with a non-zero mean until the coefficients become not significant. After checking the coefficients and residual diagnostics, the ARIMA(7,1,0) is the only model without issues. Then, I check for MA components by fitting all ARIMA(p, 1, q) models for p from 2 to 7 and q from 1 until the MA coefficients are insignificant. The diagnostic plots and the coefficients were not enough to narrow down these choices, so I compared their spectrum plots with the raw periodogram. This eliminated many models which overfit the data and showed complicated

features that the raw periodogram does not have. This narrowed the candidates to just five models. However, their spectrums are all very similar so those plots cannot be used for selection.

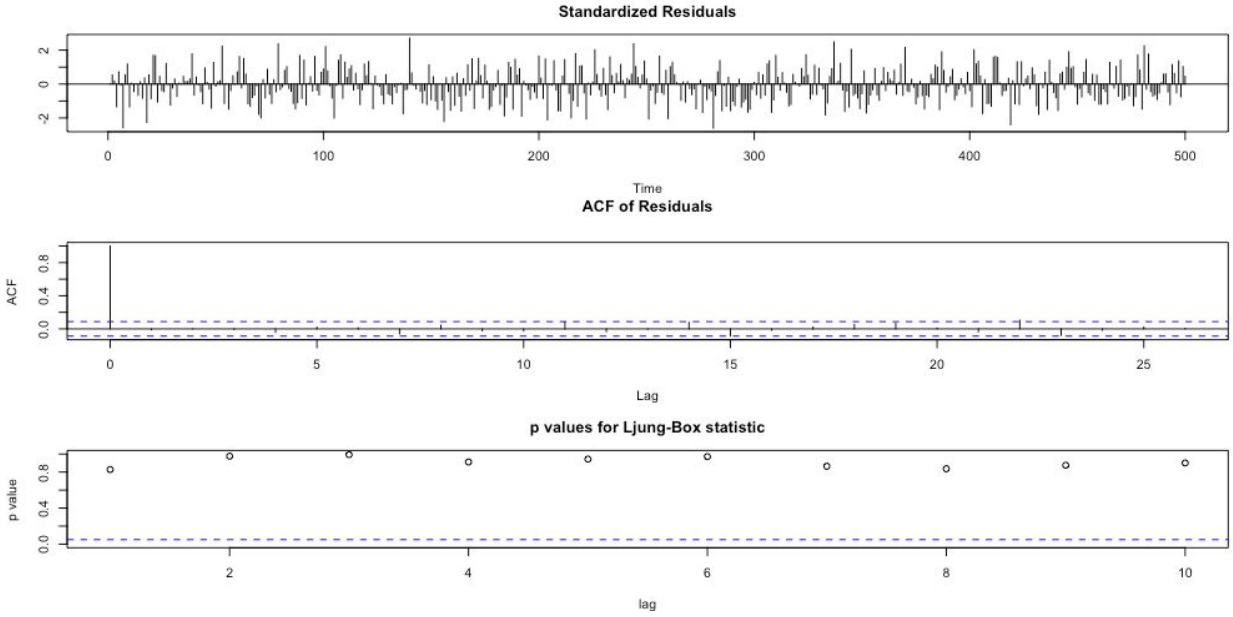
The mean squared error was also computed for the models. However, after plotting the predictions for the 20 observation holdout with the actual observations, I noticed that the predictions for these models just extends in a straight line from the previous observations. I think this makes the MSE less reliable since the actual observations may deviate from the predictions randomly, and thus there is too much dependence on where the validation set starts. This problem is true for any model but seem more apparent for this one. To try to account for this, I calculated the SSE for a 50 observation holdout and refitted the models for a 10 observation holdout. Then I fit the models for the full data to compare metrics. The results are shown below, with the best in each column highlighted.

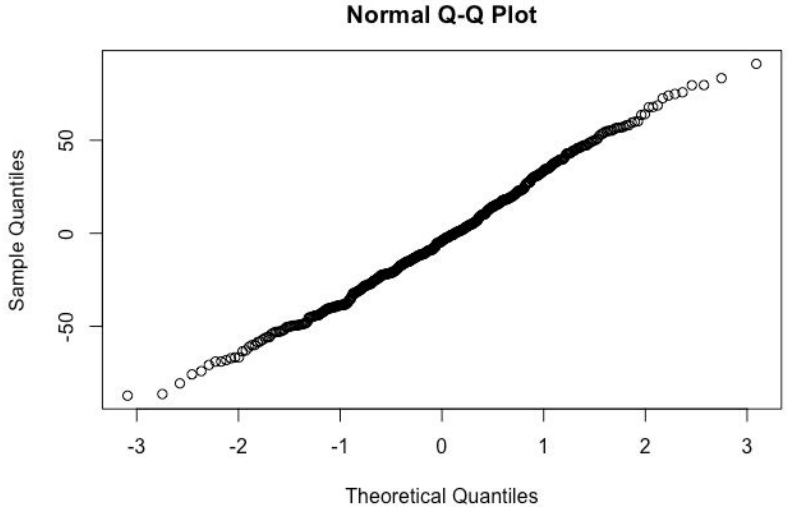
Model	SSE Last 50	SSE Last 20	SSE Last 10	σ^2	AIC	BIC
ARIMA(7,1,0)	12087424	16172564	346588	1117	4939.8	4977.7
ARIMA(2,1,2)	17444028	17153289	601288	1125	4937.2	4962.5
ARIMA(4,1,2)	8136880	14933826	521255	1114	4936.5	4970.2
ARIMA(1,1,4)	7259778	14585561	525718	1119	4936.6	4966.1

This is a tough choice. Holistically, I lean towards the ARIMA(1,1,4) model. It has relatively good prediction accuracy for these holdout sets and it has the second best AIC and BIC. One cause of concern is that the theoretical ACF of does not remain significant from lags 10 to 20 like it does with the estimated values. However, ACF alone can be misleading. Overall, this is chosen as the final model to fit on the full dataset. The coefficients and standard errors are shown below.

Coefficient	α_1	β_1	β_2	β_3	β_4
Estimate	0.8216	0.9185	0.1472	0.1176	0.1312
S.E.	0.0397	0.0590	0.0903	0.0809	0.0512

The mean is not significant so it was forced to be zero. The high order coefficients are significant. The residual diagnostic plots shown below, show no cause for concern. The Q-Q plot also looks adequate. Thus, forecasting is done with the predict function for the next 13 observations with a rough 95% prediction interval.


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag

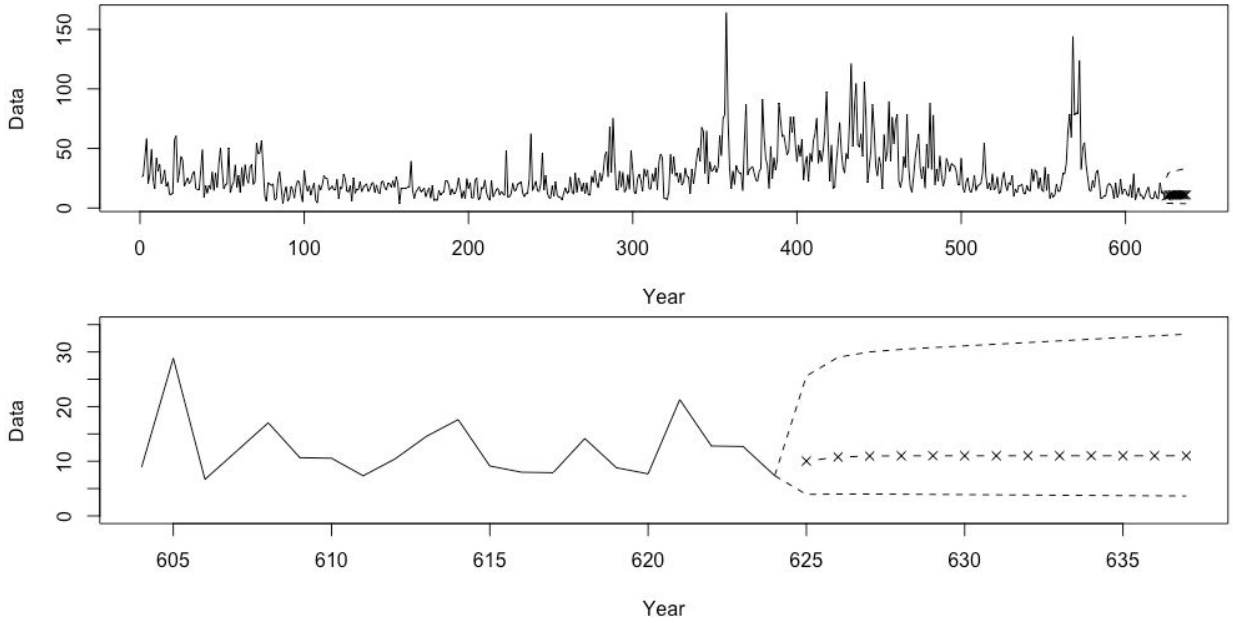

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

Executive Summary for Deposits Dataset

The goal for this dataset is to fit a time series model to the data and predict the next thirteen values. This dataset measures the thickness of sedimentary deposits for one location every year for a 624 year period. The units for the thickness are unknown. The chosen model is an ARIMA(1,1,1) model with zero-mean, on the log-transformed data. The estimated coefficients and their standard errors are presented below.

Coefficient	α_1	β_1
Estimate	0.2291	-0.8827
S.E.	0.0523	0.0295

The entire dataset with a forecast thirteen years ahead is plotted below, with a closer view of the forecast. The forecasted values are shown with X's and the future values are likely to stay within the dashed lines. The thickness of the deposits over the 624 year period have a median of 21.4 units and a mean of 28.1 units. There is no obvious trend to this data and there is no consistent cyclical fluctuation, although large spikes in the thickness is observed.



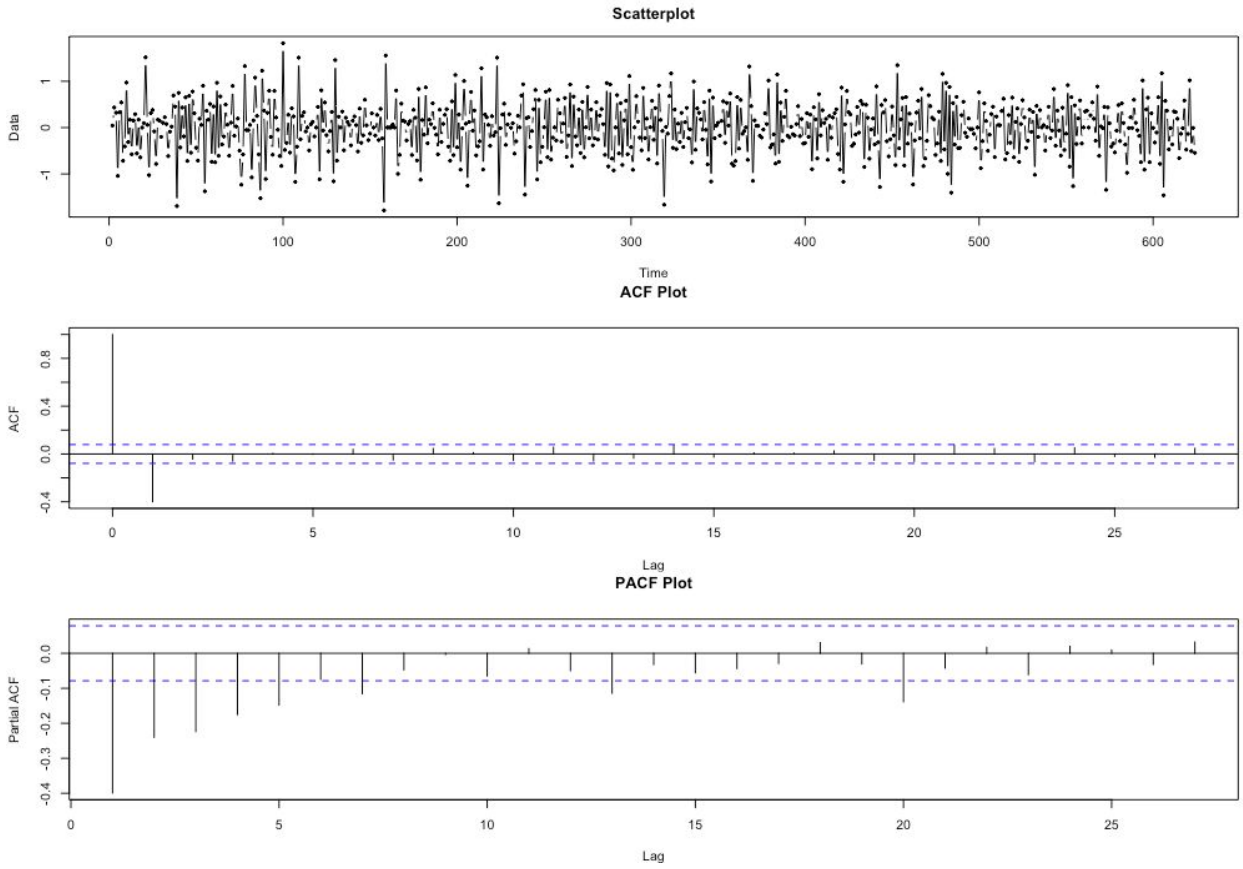
The values for the forecast and prediction interval are also tabled below for convenience.

Year	625	626	627	628	629	630	631	632	633	634	635	636	637
Upper	25.6	29.0	30.0	30.5	30.8	31.1	31.4	31.7	32.0	32.3	32.6	32.9	33.2
Value	10.0	10.8	10.9	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
Lower	3.9	4.0	4.0	4.0	3.9	3.9	3.9	3.8	3.8	3.7	3.7	3.7	3.6

In summary, although there are historical 13 year fluctuations of 70 or more units and many large spikes in the thickness, the next 13 years are projected to have have relatively stable deposit thicknesses of about 11 units. However, the thickness can vary from between about 4 and 30 units in this time span.

Technical Appendix for the Deposits Dataset

The data is the observed thickness of sediment deposits every year for 624 years. Dealing with outliers is not considered. The plotted data looks stationary in the mean but has sharp peaks which means the data may not be second-order stationary. A log transformation is the variance-stabilizing Box-Cox transformation suggested by Guerrero's (1993) method. This transformation is very sensible since it will also force the process to have positive values. However, it looks like the data still needs to be differenced after this, according to the Augmented Dicky-Fuller test and the presence of a long seasonality that an ARMA model may not account for. Thus, first differences are taken.



Now the data does not appear to have a trend or any seasonality. The Dicky-Fuller test confirms this. It looks like one model can fit this data well so I holdout the last 20 observations of the differenced data and fit models on the rest. The ACF plot appears to cut off after lag 1. The pacf may show decay or some cutoff at lag 1 or 5 or 6. The Hyndman-Khandakar algorithm suggests an ARIMA(1,1,1) to the log transformed data. Thus, some candidate models to try are ARIMA(1,1,q) models. Just in case, some MA and AR models are tried.

After checking the residual diagnostics and the significance of the coefficients, five models did not have major violations to the fit. Some evaluation metrics for the fits on the transformed data are tabled below.

Model	Validation SSE	σ^2	AIC	BIC
ARIMA(0,1,2)	3.47	0.2306	835.55	853.16
ARIMA(1,1,1)	3.75	0.2294	832.67	850.28
ARIMA(3,1,3)	3.86	0.2252	833.19	868.40
ARIMA(4,1,4)	5.27	0.2239	833.77	877.78
ARIMA(5,1,5)	5.26	0.2254	838.55	891.38

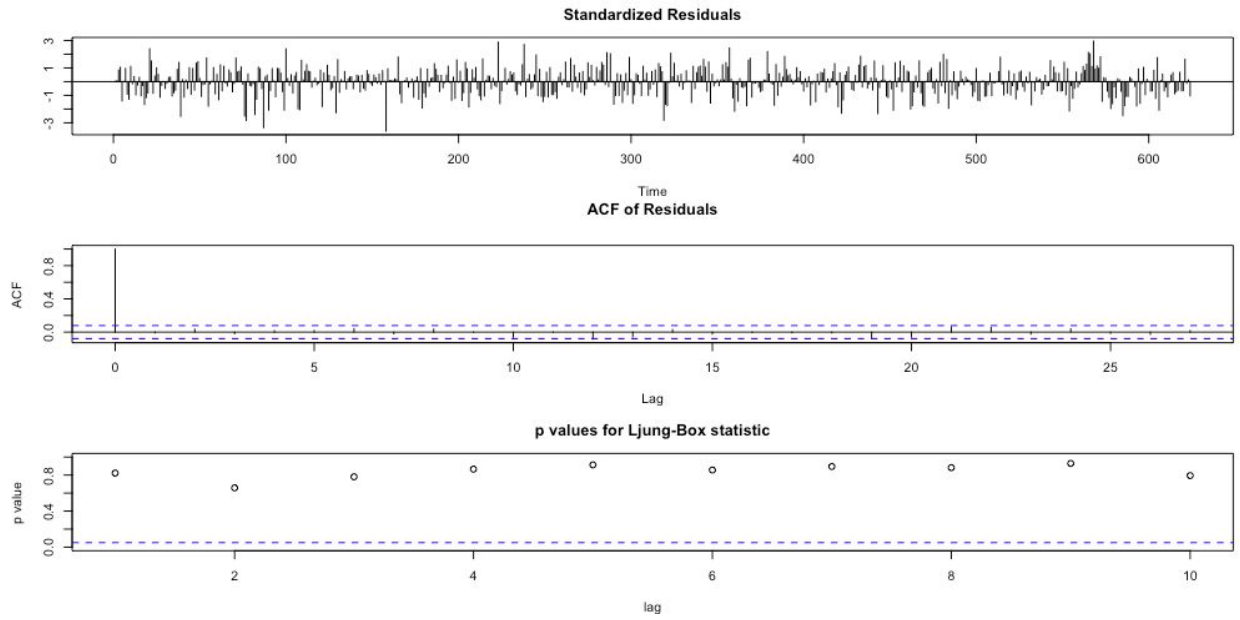
From these metrics, the ARIMA(1,1,1) is probably the best model. It has the lowest AIC, and BIC. However, it does not have the best SSE for a 20 observation holdout. Overall, there is more evidence to support the ARIMA(1,1,1) model over the ARIMA(0,1,2) model, especially because they have the same number of parameters.

Thus, the final model chosen is a zero-mean ARIMA(1,1,1), since the intercept term is not significant when fitted. This model is fit to the full log-transformed dataset.

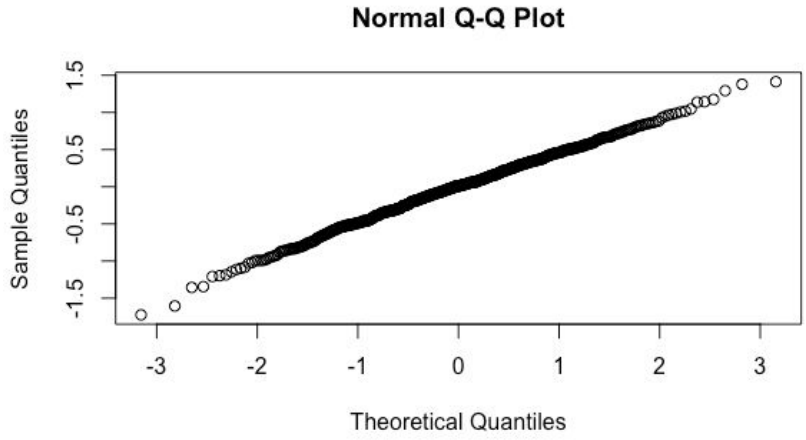
Coefficient	α_1	β_1
Estimate	0.2291	-0.8827
S.E.	0.0523	0.0295

Both coefficients are significant. The residual diagnostics are checked again (plots on the next page). The scatterplot of the standardized residuals shows little pattern but there are some areas of concern. Overall, there is no evidence of any autocorrelation of the residuals according to the ACF of the residuals and the high p-values.

Then, the log-transformed data is forecast 13 years ahead. The predicted points and approximate 95% prediction interval is exponentiated to compare with the raw deposits data.



The Q-Q plot of the residuals also shows no sign of deviation from normality as there is a relatively straight line.

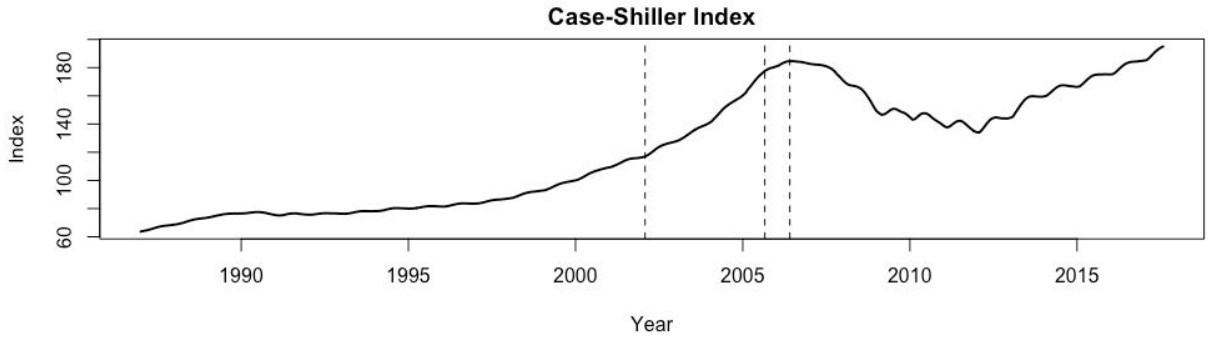


Executive Summary for the Housing Dataset

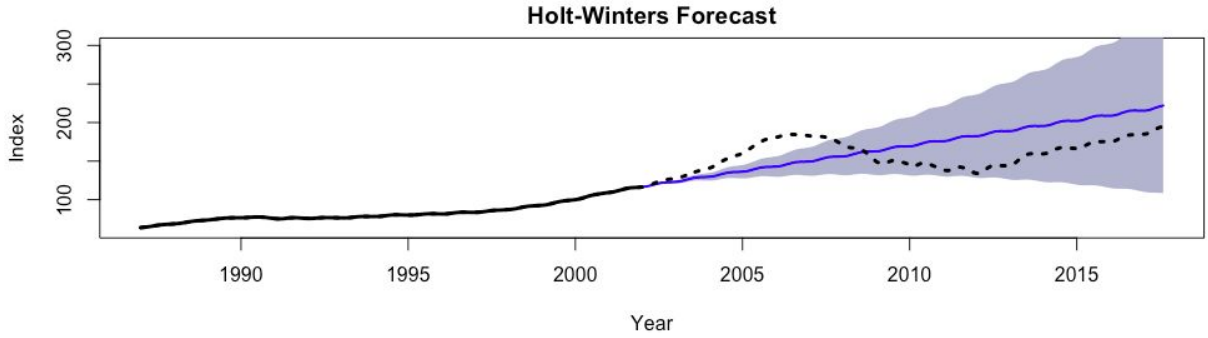
The goal of this project is to analyze the Case-Shiller Index from January 1987 to August 2017 and address the following questions:

- When did the U.S. housing bubble begin and when did it burst?
- Have we recovered from the effects of the bubble?
- What would have happened without the bubble?
- For those who are currently renting, is now a good time to buy a home?

The plot below shows the Index along with dashed lines for these dates. From a changepoint, trend, and variance analysis, the start of the housing bubble is estimated to be in February 2002. At this date, the acceleration of growth and the volatility of the Index jumped. The bubble burst during the period of September 2005 to June 2006, when growth started to slow and then decline.



We may still feel the effects of the bubble since the volatility in the index has not settled down to pre-bubble levels. However, this could be due to many factors other than the effects of the bubble. The current growth rate is similar to right before the bubble started but it is uncertain whether this is sustainable or not in the current economy. Below is a simple forecast for the "what-if" scenario without the bubble forming. The forecast is shown with a shaded prediction interval along with a dotted line for the historical data.



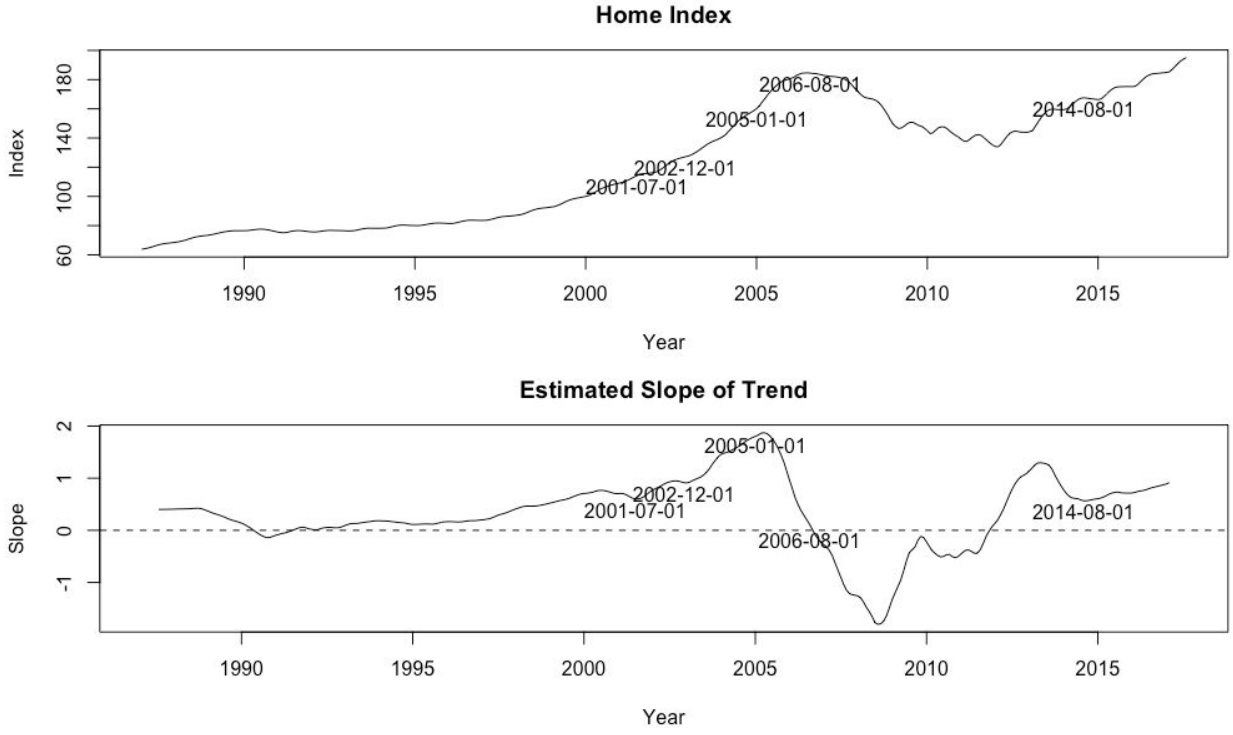
If it is assumed that the growth rate remained constant from before February 2002 onwards, the Home Index would end up higher than it is today. However, this cannot be stated with confidence as any change this trend, especially early on, shifts the forecast by a large margin. As for renters, this data cannot give them a recommendation on whether now is a good time to buy a house. There may be another bubble, or prices may continue to rise steadily for a long time. A renter needs to consider the rapidly growing rents in some locations and weigh that with the current long-term cost of a house.

Technical Appendix for the Housing Dataset

First is an exploration of the data. The data is clearly not stationary. Taking first differences of this results in an ACF plot with peaks at lag 12 and 24, indicating that there must be a yearly seasonality. In the original dataset, it is clear that housing prices rose and then fell but picking out specifics is difficult and somewhat subjective. One task is to identify key dates for the timeline of the housing bubble in the US. The housing bubble was characterized by a rapid increase in home prices and greater volatility in the prices. Thus, these dates can be determined by analyzing the trend and the variance in the data. Because the Case-Shiller index is published with a two month lag, I adjusted for this in my conclusions when possible, although I should have just shifted the dates in the dataset.

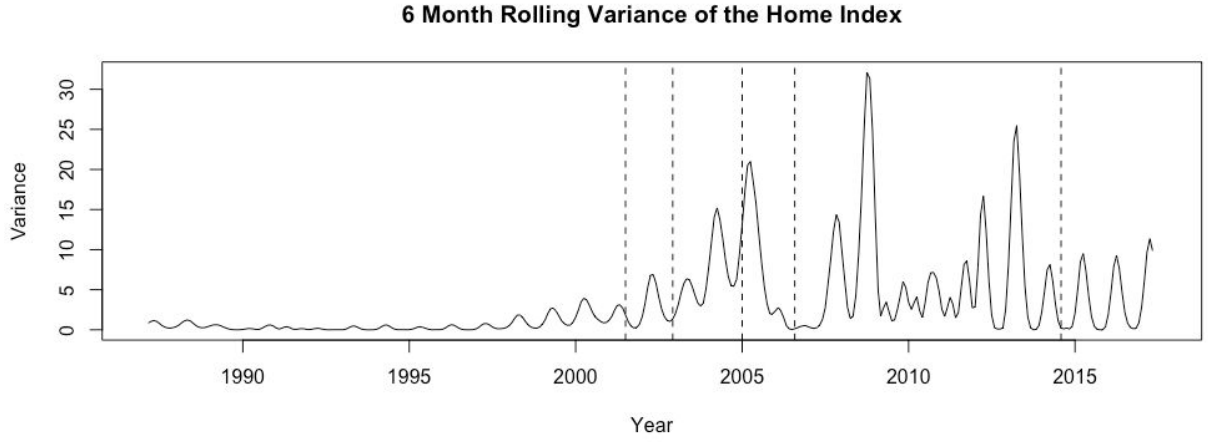
For trend analysis, I use the prophet package. This package implements a black box changepoint detection method which will give a good starting point. I used the prophet() function on the data from January 1987 to February 2006, since the housing bubble began sometime before peak of the housing index and asked for one changepoint. The model identified April 2002 as the date. This gives February 2002 as the estimated start of the bubble, accounting for the 2 month lag time for the Home Index.

For another trend analysis method, I use the stl() function with default methods, which uses Loess to decompose the time series into trend, seasonal, and residual components. I use this to isolate or smooth out the trend so it is easier to see changes in slope, and hopefully remove changes due to seasonality. I then take the first differences of this trend data to crudely estimate the slope of the trend. A plot of this along with some dates of interest are plotted below.



It is understandable why the prophet() function chose April 2002. Sometime between July 2001 and December 2002, the acceleration and rate of the housing price increased further. Thus, I will estimate that the housing bubble began in February 2002 based on this data alone. The slope of the housing index was at its peak around January 2005. Sometime between this month and August 2006, the housing bubble must have burst, since this acceleration died down immediately and then the Index peaked. I estimate that the housing bubble started to burst when the slope started to approach the slope at the start of the housing bubble. So a reasonable estimate for the period of time the bubble burst is September 2005 to June 2006. Perhaps in August 2014, the trend is back to "normal" levels, or the slope is growing to recover from the effects of the bubble. It is unclear from this plot.

Next is the analysis of variability. Originally, I fitted a GARCH model to the once differenced, seasonally differenced, differenced data to extract a measure of the conditional standard deviation, but there is a much simpler method that is suitable for the purposes of this analysis. Below is a plot of the rolling variance of the Home Index with a 6 month window. A six month window is chosen to balance the tradeoff between seeing a smoothed trend and seeing detailed features. The short oscillations are due to seasonality but their amplitudes give an measure of the variance.



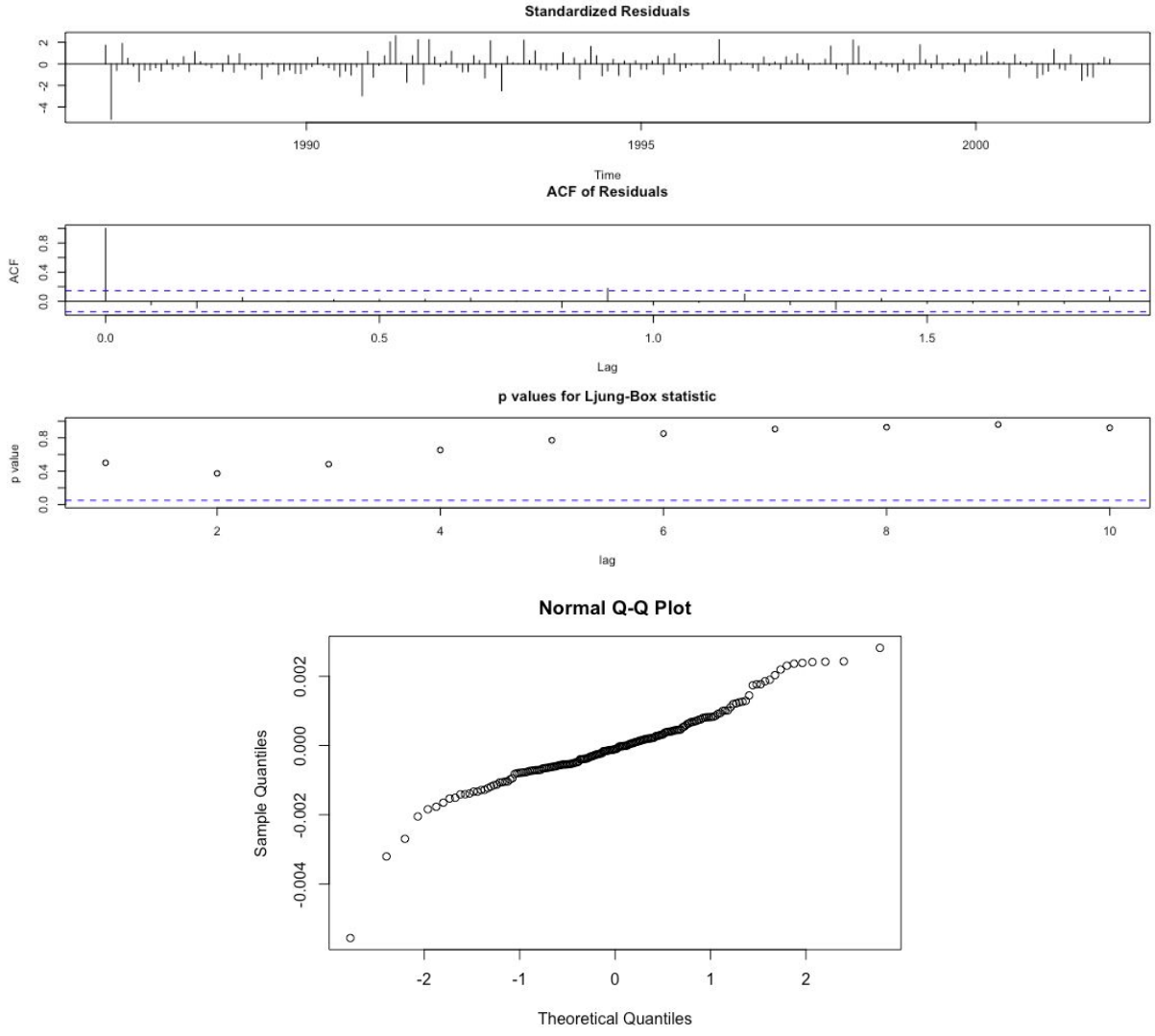
The dashed lines correspond to the dates marked on the previous plot. The dates that estimate the start of the housing bubble mark a sudden jump in the variance although the variance increases even before then. The estimates of the burst period mark a peak in the variance, followed by a reduction. The dip in the variance in 2006 may represent the proverbial "calm before the storm" as housing prices reach their critical point and inevitably fall. Without economic intuition, I am not sure how to interpret that. The effects of the burst of the housing bubble are clearly seen as the variance spikes to an all time high in around 2008.

This plot gives evidence that the US still feels the effects of the bubble in 2017. The variance has not settled down to pre-bubble levels, although it may never settle because of a complete change in people's behavior and the market. However, it is difficult to say whether another housing bubble is coming or whether the housing market has recovered. The economic environment and behavior are not the same as it was in 2002. In a sense, the

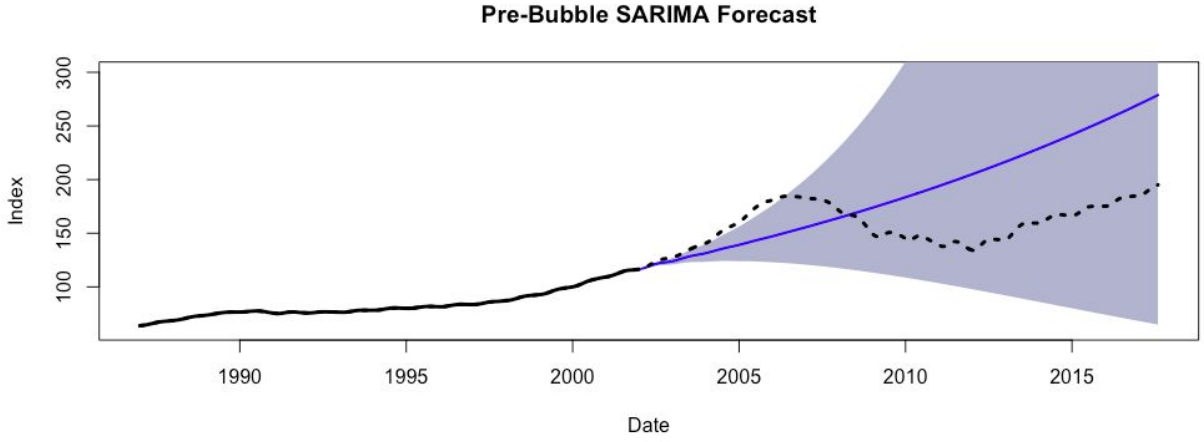
analysis of the trend and variance shows that we may still be feeling the effects of the bubble.

The next task is to forecast what would have happened to the housing index if there was no bubble. For this, I use the data prior to February 2002 for the pre-bubble model. I also use the popular auto.arima() function from the forecast package to choose the model with a variant of the Hyndman-Khandakar algorithm. I am confident with this algorithm after comparing its suggestions to my own final models for the other datasets, although the models for this dataset need to be much more complicated. I used a log transformation to force the forecasts to be positive. The suggested model is an SARIMA(3,2,1)(1,0,0)[12] for the log transformed dataset. The model coefficients and residual diagnostics are shown below.

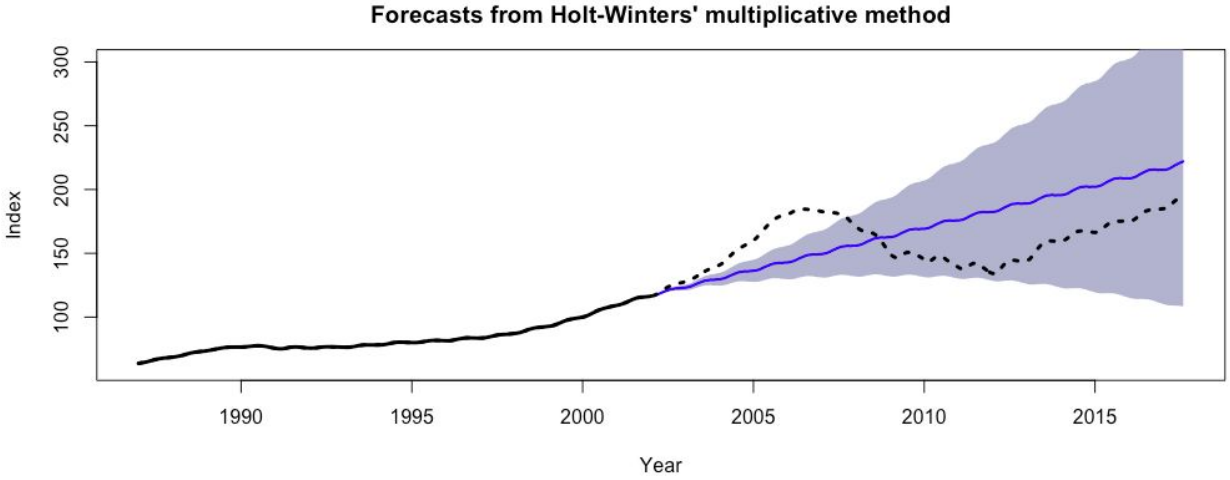
Coefficient	ϕ_1	ϕ_2	ϕ_3	θ_1	Φ_1
Estimate	0.8883	0.1623	-0.4898	-0.8144	0.4205
S.E.	0.0806	0.0964	0.0691	0.0551	0.0815



The high order coefficients are significant. There is no evidence for the autocorrelation of the residuals since the p-values for the Ljung-Box statistic are high. The Q-Q plot shows lots of deviation near the tails so this could be a cause for concern. Forecasts are made using the forecast() method to August 2017. The forecast results are plotted below with the historic data plotted in a dashed line. The forecast starts with their prediction intervals and the beginning of the dashed line.

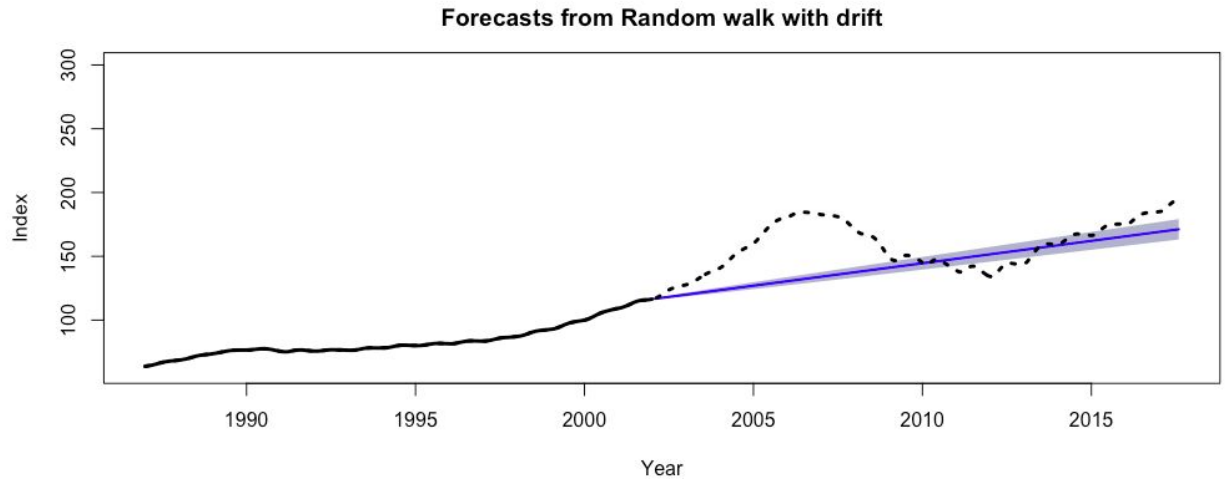


This forecast looks unacceptable because the 80% prediction interval grows without bounds into impossible Index levels. This is wrong since the Index and its growth rate must be bounded by economic principles, although forecasting 187 months into the future for economic data is may not appropriate with this method anyway. Another forecasting method to try is an automated Holt-Winters procedure. I choose the multiplicative method because a close up of the data shows that the seasonality increases in amplitude with increasing Index before the bubble.



This procedure may be too simple for this kind of data but results in a more reasonable 80% prediction interval. However, I would not trust either of these intervals given the assumptions that these forecasts require. Putting even stronger assumptions on the the

Home Index, I also plot a random walk with drift forecast on the next page, to see what the Index would look like if the overall average trend continued in a constant manner.

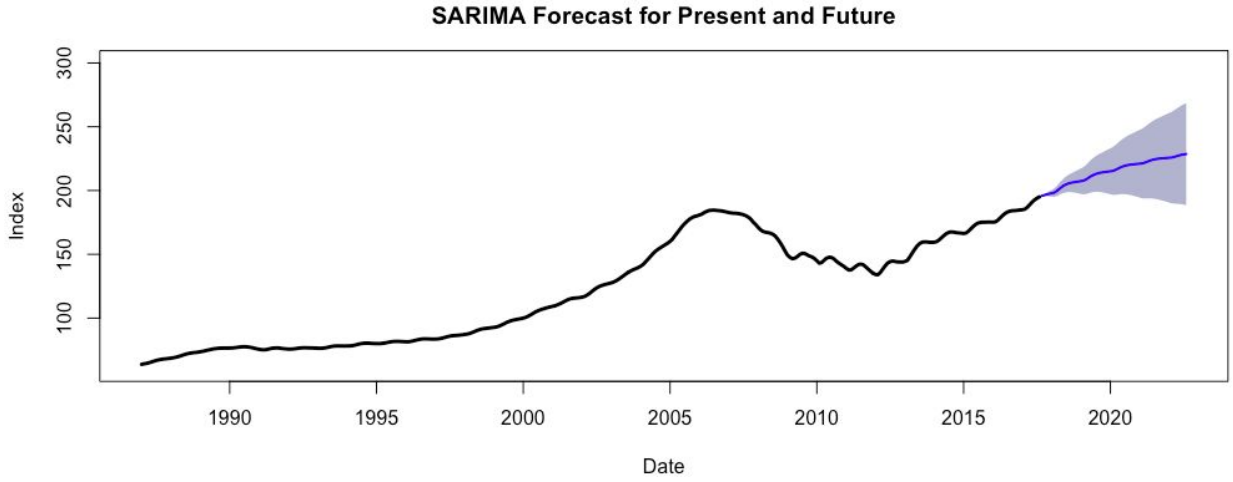


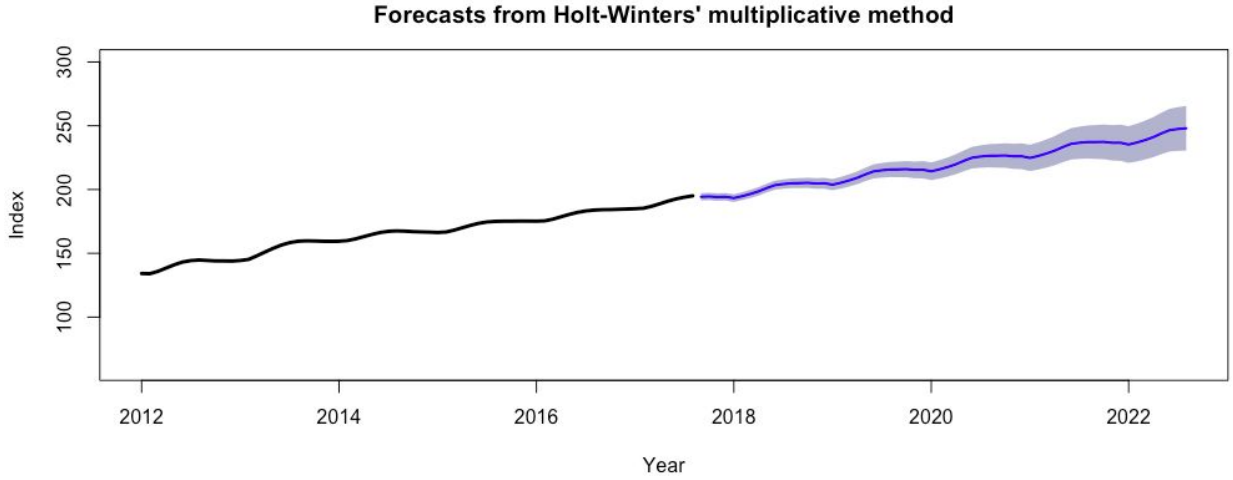
Overall, I trust the Holt-Winters forecast the most since the results are the most reasonable based on my assumptions about stable economic activity. The point forecast seems to extend the previous trend. There are many flaws to these methods overall. The forecasts are sensitive to the point at which the bubble is assumed to begin. The forecasts assume that "pre-bubble" activity is stable and predictable in that the trend just continues without interruption. Also, the forecasts for SARIMA overestimate the prediction interval without constraints put on the economic process, Holt-Winters may underestimate the prediction interval, and the last forecast definitely underestimates the prediction interval.

If the Holt-Winters forecast and related assumptions are valid, it appears that the point estimates show that housing prices would have been higher than they are now if there was no bubble. However the prediction interval is too wide to make a strong conclusion, even if the assumptions are met.

As for the question of whether now is a good time to buy a home for a renter, forecasts for the next year using any method on either the full data or post-bubble will probably show a continuation of the upward trend. There is no simple implementation of a SARIMA-GARCH model in R, so I will use the auto.arima() function on the full data and exponential smoothing on data from 2012 to 2017.

The resulting SARIMA(1,1,2)(2,0,0)[12] model has some problems with the residuals. Of course, the standardized residuals are larger during the housing burst and recession. Some p-values for the Ljung-Box statistic are around 0.3 and the Q-Q plot is rather s-shaped with one large outlier. Still, the results are plotted on the following page along with the Holt-Winters forecast for the data since 2012. Both forecasts are made to five years after August 2017.





Unfortunately, this Case-Shiller Index is not adjusted for inflation. However, it is likely that this Index will increase further beyond the highest point during the housing bubble. Again, it is difficult to tell whether another bubble is coming from this data alone since the economy has changed since 2002. Also, this is unknown territory since this Index is higher than it has ever been. It is also uncertain whether the Index will continue to rise, level off, or slowly decrease. Thus, it may not be the right time to buy a house, but rent prices are growing even faster in some areas. Thus, no clear recommendation can be made here.