Econ 144 Project 3 - Cal Parker & Jordan Youhanaie

2022-08-02

```
Loading required package: tinytex
Warning message:
In system(paste(shQuote(c(...)), collapse = " ")):
  sh: pdfcrop: command not found

MP Statistical IV mental experiment for 'quarterly'
Using seed 1234567890 ...
  ARIMA(2,1,2)(1,1,1)[4] with drift

Loading packages...
Loading tidyverse:
Loading tidyverse:
Loading lubridate:

--- IV modeling package ----------------------------------------------- Version 1.1.4

  ar1 ar2 ma1 ma2 sar1 sma1  drift
  0.7 0.1 0.3 0.4    0.9 -0.8  0.01

  sigma^2 = 1.1:  log likelihood = -1.5
  AIC=17.0  AICc=Inf  BIC=Inf

Loading packages...
```

```
>
> ccf(ts.ARIMA.res)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res"
$snames
NULL
attr(,"class")
[1] "acf"
>
> ccf(ts.ARIMA.res,ts.ARIMA.res)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res" "ts.ARIMA.res"
$snames
NULL
attr(,"class")
[1] "acf"
>
```

```
>
> ccf(ts.ARIMA.res,ts.ARIMA.res^2)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res" "ts.ARIMA.res^2"
$snames
NULL
attr(,"class")
[1] "acf"
>
```

```
>
> ccf(ts.ARIMA.res,ts.ARIMA.res^3)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res" "ts.ARIMA.res^3"
$snames
NULL
attr(,"class")
[1] "acf"
>
```

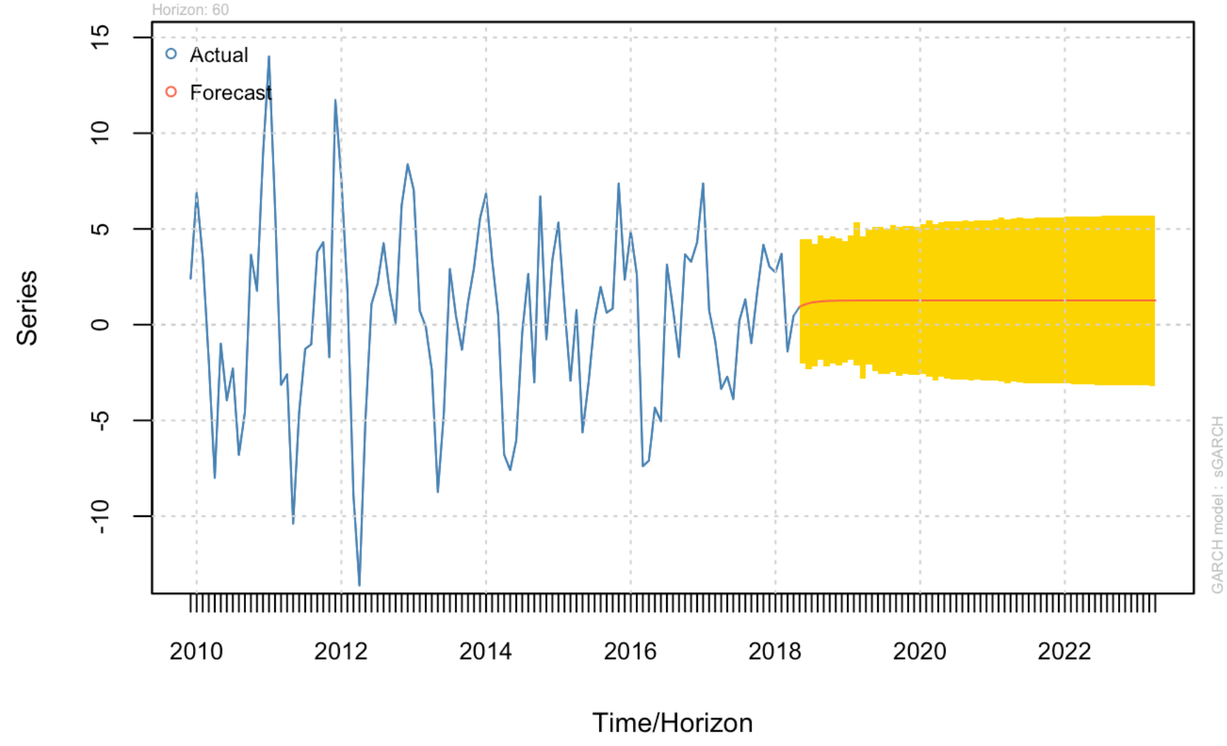
```
>
> ccf(ts.ARIMA.res,ts.ARIMA.res^4)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res" "ts.ARIMA.res^4"
$snames
NULL
attr(,"class")
[1] "acf"
>
```

```
>
> ccf(ts.ARIMA.res,ts.ARIMA.res^5)
$acf
, , 1
     [,1]
[1,]    1
attr(,"class")
[1] "matrix" "array"
$type
[1] "correlation"
$n.used
[1] 1000
$lag
, , 1
     [,1]
[1,]    0
attr(,"class")
[1] "matrix" "array"
$series
[1] "ts.ARIMA.res" "ts.ARIMA.res^5"
$snames
NULL
attr(,"class")
[1] "acf"
>
```

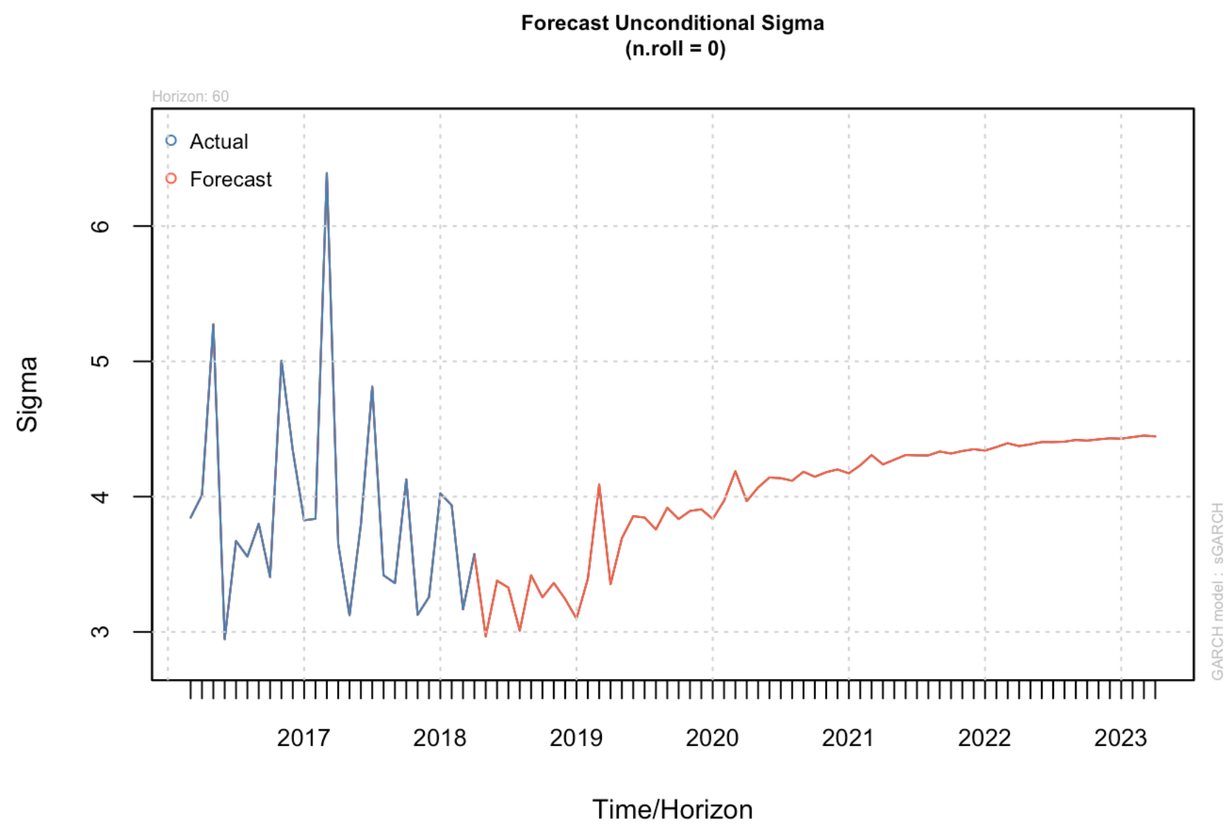
As implied by the residual plots for our ARIMA model and our cross-correlation models, we do not have residuals that are iid as we saw/know is the model, but we can see some of the residuals for all our other models. With the exception of our ARIMA model, all these models residuals seem like white noise. Although, the residuals for our ARIMA model are not iid, we can still use our model to predict future values because we have accounted for the correlation in our data. However, even conducting a prediction test for these models, we see that they are all correlated. This is unfortunate, as it means our models are not independent of each other. Although, we can still use our models to predict future values, we cannot use the residuals of the equal data, as these are not iid and correlated. Hence, we will be transforming the data to see if this fixes the problem.

```
Warning: Coefficients are not stationary.
Warning: Coefficients are not stationary.
>
> ts.ARIMA.res.sq=ts.ARIMA.res^2
> ts.ARIMA.res.sq[1:5]
[1] 0.000000 0.000000 0.000000 0.000000 0.000000
```

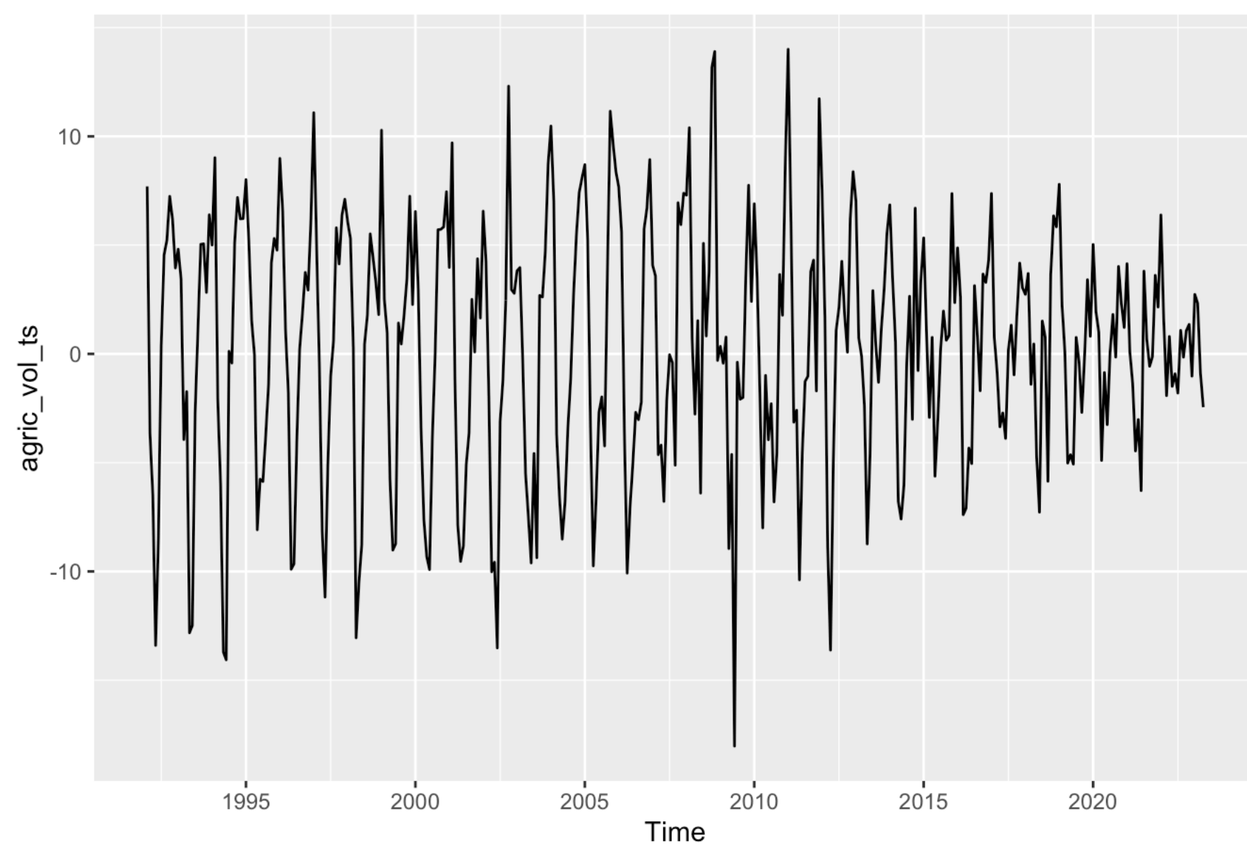
```
mv: cannot rename /home/mw/.local/share/jupyter/kernelspec/ir: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js.map: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js.LICENSE.txt: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js.gz: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js.gz.LICENSE.txt: Permission denied
mv: cannot rename /home/mw/.local/share/jupyter/nbextensions/nbextensions_configurator/tree_tab/main.js.gz.map
```




```
plot(modelforecastArch12, which = 3)
```




```
#Recreate our volatility data for the whole time series so we can compare our forecast to the acutal data
ln_agric_ts <- log(agric_ts)
ln_agric_ts_Lag <- stats::lag(ln_agric_ts,-1)
agric_vol_ts <- 100*(ln_agric_ts-ln_agric_ts_Lag)
autoplot(agric_vol_ts)
```



Comparing our forecast series with the actual data, the point forecast obviously does not match the variance of the actual data, but the range of the confidence band appears to be relatively close to the volatility of the actual data.

Conclusions and Future Work

To wrap up our project, we conclude by reviewing the most important findings of our project. From the data, we found that MAPA and prophet provided substantially successful model forecasts compared to the other forecasting methods, including two forecasts that combined individual forecasting methods. The RMSE for our prophet (5th order) forecast was 189 and our MAPA forecast RMSE was 171.These two values are much lower than the other model fitting methods, including our combined forecast methods which were 216 for combining all model forecasts, and 228 for combining the top 4 forecast methods, excluding prophet. We used as many forecasting methods as we could in the beginning and then tried to test the combination by including only the best forecasting methods for creating an improved combination, or so we thought that it would be better. We believe that the combination may have been better with more forecasting methods than with selected methods because the extreme forecast values of the fully combined forecasting method had balanced out the expected value and been placed closer to the realized values whereas the selected forecasts were slightly biased and did not contain the extreme weights to balance the expected value. However, this accuracy is likely to come at a cost since the variance for the combined forecast may be greater than the selected forecast combination because there were more extreme predictions. While we may have had many successful mean-forecasting models, all of our models indicated non-covariance stationary residuals after fitting the forecasting model to the testing data shown by our checkresiduals functions and Ljung-Box tests. After creating models to use for forecasting the mean of our data, we moved on to forecasting volatility and transformed our data using a first order logarithmic difference of our data set. To forecast our volatility, we used a GARCH and ARCH process with an automatically specified ARIMA function to create the volatility forecasts. Next we tested a first order ARCH model and a (1,1) GARCH model to gauge the our difficulty for fitting an accurate volatility forecast. After, we viewed the ACF SSR plots of our data and our model fits. We had made decent progress at first. We tested different orders of our models and found that increasing the GARCH model orders didn't substantially affect the ACF SSR plot of our fitted values, but increasing the order of our ARCH model order seemed to be effective as shown by our ACF SSR plot of our model fit. Ultimately, we decided to select an ARCH(12) model to forecast volatility because it had the best ACF SSR plot. For future work, we could improve our volatility forecast by finding a GARCH or ARCH order that better explains the seasonally ordered volatility shown by the ACF plot of our Squared Standardized Residuals (SSR). We had thought the SSR would be explained by the mean forecasting method, but that did not resolve the issue after many attempts. Also, we had trouble finding a mean forecasting model that provided covariance stationary residuals after fitting the forecasting model to the testing data. From these results, there seems to be a way to more accurately forecast our models that we did not included in the selected forecasting models. Considering that our combined forecast model did not present the best results out of all the mean-forecasting methods that we used, there may have been an additional model fitting method that we are not exposed to yet and explains the data better for forecasting data similar to ours.

References

FRED Data for 'Manufacturers' Total Inventories: Pesticide, Fertilizer, and Other Agricultural Chemical Manufacturing': https://fred.stlouisfed.org/series/U25ATI US Census Bureau for 'Pesticide, Fertilizer, and Other Agricultural Chemical Manufacturing: U.S. Total': https://www.census.gov/econ/currentdata/dbsearch?programCode=M3&startYear=1992&endYear=2023&categories[]=25A&dataType=VS&geoLevel=US&adjusted=1&notAdjusted=1&errorData=0