

AAEC 4484/AAEC(STAT) 5484: Applied Economic Forecasting

Your Name Here

Homework #6 - Spring 2024

Instructions: In all cases, please ensure, where necessary, that your graphs and visuals have proper titles and axis labels. Refer to the output, whenever appropriate, when discussing the results. **Lastly, remember that creativity (coupled with relevance) will be rewarded.** **Instructions:** Please ensure that your graphs and visuals have proper titles and axis labels, where necessary. Refer to the output, whenever appropriate, when discussing the results. **Lastly, remember that creativity (coupled with relevance) will be rewarded.**

Question 1: Forecasting Delinquency Rate on Loans to Finance Agricultural Production

A local bank named N-U-I-TRUST has requested that you to predict the delinquency rate of future loans. However, they have not yet provided any additional information beyond this request. You remembered that the FRED database contains some useful macro variables and decided to explore further. While looking into the database, you came across the **Delinquency Rate on Loans to Finance Agricultural Production, All Commercial Banks (DRFAPGACBN)** series. You think that this series would be an appropriate starting point for your analysis. You can find more details about this series at the following URL: <https://fred.stlouisfed.org/series/DRFAPGACBN> You created an outline and decided to proceed with forecasting as follows.

1. Using the `quantmod` package, pull the `DRFAPGACBN` series from the FRED website. Be sure to declare this series as a `tsibble()` object. Save the final variable as `delrate`. ****Ensure that `delrate` has the appropriate index and frequency.**
2. **Use appropriate graphs to visualize the data and briefly comment** on any relevant properties that you deem pertinent.
 - In particular, does there appear to be seasonality and/or a trend?
 - If there is seasonality, does it appear to be multiplicative or additive?

Note:

(i) Be sure to explain how you used your graphs to arrive at your conclusion. (ii) I expect you to use/present at least three (3) graphs to supplement your answers. Be sure to grid.arrange your figures.

Now, let us back-test a few strategies to see which one would have performed best on past data under our assumption that the patterns observed will continue into the future.

3. Subset the data into a training set, `delrate.train`, and a test set, `delrate.test`. Assign the last five (5) years of data (`n = 20`) to the test set.

Plot the training, test, and full data to ensure that the split was done correctly. Add the original data as **points** and the test and training data as **lines**. To ensure you can see the data (points) clearly, set the `alpha`(transparency) parameter to 0.5 and the `size` to 0.5.

Note: Instead of using mental math or using the `filter_index` command window command, I would like you to use the `slice_head()` and `slice_tail` functions, as appropriate, to split the data.

A quick Google search should help you with the syntax. My aim is for you to be comfortable exploring unfamiliar functions or those not explained in our class notes.

4. You thought about it for a while and decided that you will pursue seven (7) distinct forecasting models to begin with:

- Seasonal Naïve
- Simple Exponential Smoothing (SES)
- Holt (Additive Trend)
- Holt (Additive Damped-Trend)
- Holt-Winters (Additive: Error, Trend, Seasonality)
- Holt-Winters (Additive: Error, Seasonality; Additive: Damped-Trend)
- Holt-Winters (Multiplicative: Error, Seasonality; Additive: Trend)

In a single step, produce the model fits for each of the models above and store the results as `model.fit`. Next, produce the forecasts over the test period and store them in a variable called `full.fore`. Ensure your models are given appropriate names to help you identify them later. **Do not produce any plots at this stage.**

5. Produce a plot that includes the *test* data and *point forecasts* of your model results. *Ensure your series are properly labeled.*
6. Briefly comment on how well your forecast appears to stack up against the true data. Do you believe the point at which you subset the data actually mattered for your model predictions above? Explain. **You do not need to conduct formal testing, and four (4) sentences will suffice.**
7. Which of the forecasting models stored above provides the better forecast?

Use the RMSE, MAPE, and MSE to make your **justify** your final determination. Note that you might have to manually create the MSE since it is not a default output of the `accuracy` function.

You are required to create a table using the `kable` command here. Please ensure that your table is correctly formatted to 2 decimal places and the models are appropriately labeled.

8. Using the `gg_tsresiduals()` command of your preferred model, comment on whether the residuals appear to be white noise. **For your answers, be sure to conduct a Ljung-Box test (with 8 lags), discuss its null and your conclusion at the 1% and 5% levels of significance.**
9. Write a brief **nontechnical** report to the Manager of N-U-I-TRUST reporting your findings.

Your report should include an explanation of how you arrived at the preferred model, but try to avoid boring them with all the technical details involved.

- Since you have a preferred model, use it on the **full data** to provide forecasts for the next three (3) years ($h = 12$) and potentially frame your discussion around that. What do the 90% Prediction intervals around your estimate look like?
- What does your forecast imply about the delinquency rate on loans to the Ag sector in the next three (3) years? Do you expect a(n) decline/increase?

In the main, this question is free-form, and my attempt to have you explain your model results to a non-scientific audience.