STA 137 Time Series Analysis: Central African Republic's Exports (1960-2017)

Nisha Thiagaraj, Adithi Sumitran, Bradley Parmer-Lohan

Introduction

The Central African Republic (CAR) has gone through a lot of changes in the last 60 years. The CAR gained independence from France in 1960, but still received support from the foreign nation. French support persisted over the next 40 years and 4 dictators, until the Bangui Accords were signed in 1997. These accords led to France withdrawing from the country, the United Nations starting a peacekeeping mission, and ultimately, another coup in 2003. A new constitution was drafted in 2015, and a new president was elected in 2016. These political events ultimately tie in to the value of exports that we see in our time series analysis.

The CAR's exported materials also went through major changes in this 57 year period. While they have primarily exported raw materials, their automotive production capabilities have recently expanded to complement their economy. This constantly-shifting export demographic in both material and destination is important to consider in our data analysis. Different areas of the economy will have different short and long term impacts on the overall wellbeing of the nation, and different nations may be more or less consistent purchasers of materials, leading to good fortunes one year, and poor investment the next.

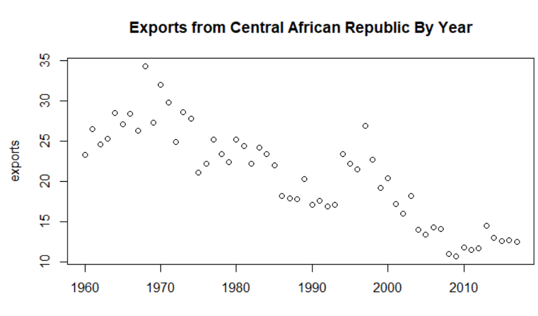
Data Analysis

Step One: Plotting the Data

As mentioned previously, looking at the time series graph allows us to see interesting spikes in the data. These line up with the shifting political landscape of the CAR we highlighted in the introduction. For example, the peak in exports is in the mid 1960s, which coincides with CAR receiving independence from France. Another peak is in the late 1990s, around the time of the Bangui Accords, which ended fighting between the government and rebel forces and promoted economic stability. The lowest dip is in the early 2000s, which fits with the political uncertainty at the time, given the 2003 coup.

So, we must keep this in mind as an important context to our data analysis. The exports of any nation are highly dependent not just on other countries' desire to invest in the CAR, but also the government's ability to properly stabilize the means of production and provide confidence to investors that orders will be fulfilled.

Figure 1: Time Series Plot of Central African Republic Exports



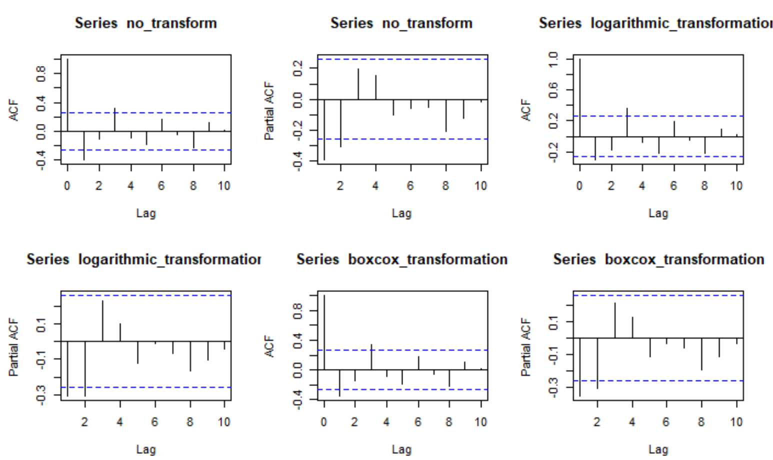
Step Two: Transforming the Data

Before applying any models or estimations to our data, we needed to apply a transformation to it to ensure we were getting the clearest picture possible. We applied two different transformations: logarithm and box-cox. We will compare these three transformation methods based on how significant the residuals are, how close they are to normality, and how well they fit our applied models. The logarithmic transformation is useful in this context because logarithmic transformations effectively compress data, keeping the relative distances from the mean while eliminating the impacts of extreme values. If our data has extreme values or unequal variance, a log transformation may allow us to make a better model. We also used the box-cox transformation, which is useful because it fits data to a normal distribution. If the data does not follow this typical distribution, then a box-cox transformation will likely help mitigate this issue. For the box-cox transformation, we calculated the lambda value which fit the data best to be 0.5051.

Step Three: Plotting the ACF and PACF

The ACF and PACF are crucial to compare when analyzing time series data. When comparing the ACFs and PACFs of different transformations, we are primarily looking for an indication as to what order of AR and MA modeling to apply. The ACF will point us in the right direction for the MA application, while the PACF can give us an indication as to what AR model should apply. This is because in ACF, AR(p) decays exponentially, while MA(q) cuts off at lag q, so the MA model can be determined from the ACF plot. Similarly, in PACF, MA(q) decays exponentially, and AR(p) cuts off at lag p, so the AR model can be found from PACF. Looking at the plots of ACF and PACF for the three transformations, we can see that the cutoff on the ACF plots is at lag 3, indicating that we should investigate up till MA(3). Furthermore, the PACF plots have significant cutoffs till lag 2, indicating we should investigate up till AR(2). We will therefore look at models such as MA(3) and AR(2) before choosing one outright.

Figure 2: ACF and PACF Plots of The Various Transformation



Step Four: Comparing AIC and BIC

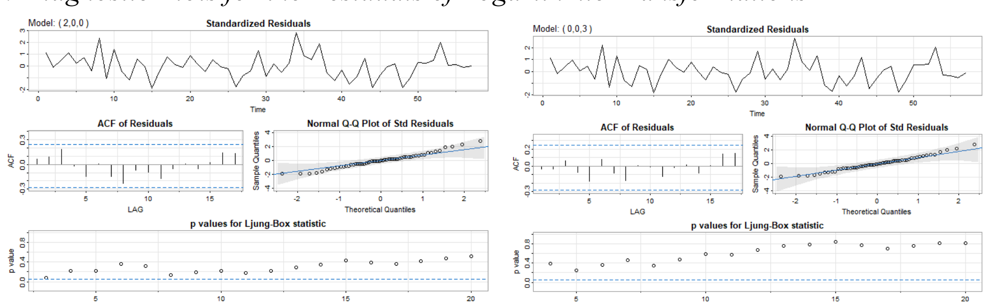
Our next step in our data analysis process is to compare the AIC and BIC of our different models. Generally, the lower the AIC, the better the model. Additionally, we prefer models with lower complexities, as this can lower computation costs with no major impact on our forecasting capabilities. So, we applied AR 1 and 2, MA 1-3, and ARMA(1,1) to all three of our data transformations. Within each

data transformation, AR(2) and MA(3) consistently came out on top with the lowest AIC values. MA(3) had a slightly higher AIC, but never by more than about 0.01, which we do not consider a significant difference. The other models had differences of around .03 or higher, so these two were the clear winners by the metric of AIC. For example, in the logarithmic transformation, the AIC of AR(2) was -1.32 and MA(3) was -1.31. For AR(1), MA(1), MA(2), and ARMA(1,1), none of the AICs were below -1.3.
A similar trend was observed for BIC values. BIC is used for model selection, like AIC, but it factors in model complexity, which means having a higher lag order will hurt the BIC score. In all the transformations, AR(2) also always had the lowest BIC score. As such, based on the purely AIC and BIC values, AR(2) was consistently selected as the best fitting model.

Step Five: Comparing Residuals

The next metric we need to check for both of our models is how well the residuals fit, while looking at other assumptions. To look at residuals, we can start with the ACF plots of residuals. Of all the model types, only AR(2) and MA(3) had residuals which did not go past the cutoffs, which signifies homogeneity of residuals. Next, the QQ-plots show normality of residuals, with plots that most closely follow the line being very normally distributed. Of all the plots, AR(2) and MA(3) plots had the least outliers, indicating normality of residuals. We also looked at the p-values of the Ljung-Box statistic, which is a test for independence. Only in the AR(2) models and MA(3) models were the p-values insignificant, meaning that the null hypothesis of independence of residuals could not be rejected. As such, the AR(2) and MA(3) models best fit with the necessary assumptions for residuals.

Figure 3: Diagnostic Plots for the Residuals of Logarithmic Transformations



Step 6: Selecting a Model

Finally, before we can make a confident forecast on our time series, we have to pick between AR(2) and MA(3). We ultimately selected AR(2) with a logarithmic transformation. We believe that this model represents the best of both worlds when forecasting our data. Our PACF plots for one, resembled AR(2). Additionally, the AIC and BIC values for the AR(2) model were the lowest in each transformation group. While comparing residuals, we did see that the p-values for MA(3) were higher than those of AR(2), but felt that because the p-values were insignificant regardless, we needed to pick the less complex model. We also settled on the logarithmic model because of the residual plots. Looking at the QQ plots, we saw that the normality condition for AR(2) was best met in the logarithmic model, making this the ideal choice.

For selecting the model, we set $y_t = \log x_t - \log x_{t-1}$. Then, we look at the coefficients from our code output. At the $\alpha = 0.05$ significance level, the xmean output is not significant, so it is dropped from the model, but the AR(1) and AR(2) coefficients are significant. Thus, we fit the model $y_t = -0.4035y_{t-1}$ - $-0.3037y_{t-2} + w_{t.}$

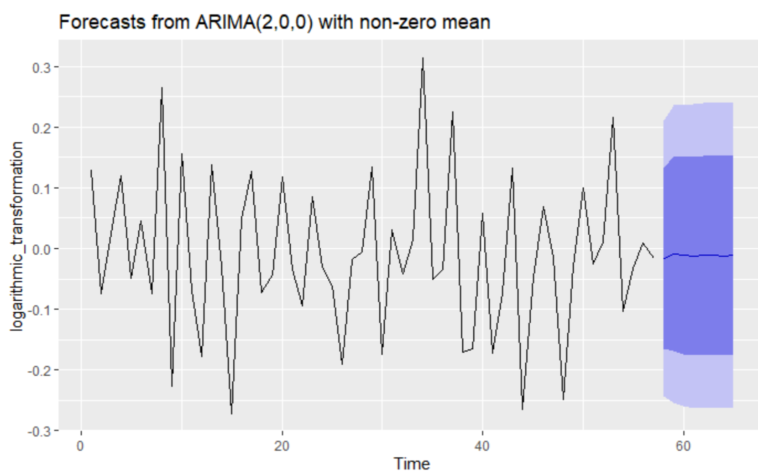
Figure 4: Coefficients of AR(2) Model

```
Coefficients:
      Estimate     SE t.value p.value
ar1    -0.4035 0.1263 -3.1957  0.0023
ar2    -0.3037 0.1247 -2.4351  0.0182
xmean  -0.0117 0.0091 -1.2895  0.2027
```

Step 7: Forecasting

We are now able to create a forecast for our data, and have displayed it below. We use forecasting in order to predict how future cycles of the data will look. In this case, we would like to see the future of CAR exports, which could be useful for those who are hoping to invest in the Central African Republic's economy. This forecast is based on an AR(2) model following a logarithmic transformation of the data.

Figure 5: Forecast of ARIMA (2,0,0)



Conclusions

As we can see from our forecast, the Central African Republic's exports are equally likely to grow or shrink by a factor of about 0.15. The time series plot of the CAF does remain very cyclical after transformation, and as such, we are unable to confidently say that the exports will continue to move in one direction or another, based on the data that we have. The CAR has a significant amount of natural resources which make it a popular country to trade with, but unfortunately, its relatively weak and unpredictable government may not be able to keep that opportunity stable enough for consistent growth.

Code:

```
library(astsa) ## textbook package
library(forecast) ## forecast package
library(MASS)

df=finalPro_data
exports=finalPro_data$Exports
year=finalPro_data$Year

ts.plot(exports)
plot(y=exports, x=year)

choice_one=diff(exports)

choice_two=diff(log(exports))

bc=boxcox(exports~year)
lambda=(bc$x)[which.max(bc$y)]
choice_three <- diff((exports^lambda - 1) / lambda)

par(mfrow=c(2,3))

acf(choice_one,lag.max = 10)
pacf(choice_one,lag.max = 10)

acf(choice_two,lag.max = 10)
pacf(choice_two,lag.max = 10)

acf(choice_three, lag.max = 10)
pacf(choice_three, lag.max = 10)


sarima(choice_one,1,0,0) # AIC = 4.893757  AICc = 4.897655  BIC = 5.001286
sarima(choice_one,2,0,0) # AIC = 4.829009  AICc = 4.836953  BIC = 4.972381
sarima(choice_one,0,0,1) # AIC = 4.856507  AICc = 4.860406  BIC = 4.964036
sarima(choice_one,0,0,2) # AIC = 4.858377  AICc = 4.866322  BIC = 5.001749
sarima(choice_one,0,0,3) # AIC = 4.838539  AICc = 4.852034  BIC = 5.017754
sarima(choice_one,1,0,1) # AIC = 4.883527  AICc = 4.891471  BIC = 5.026899


sarima(choice_two,1,0,0) # AIC = -1.260289  AICc = -1.25639  BIC = -1.15276
sarima(choice_two,2,0,0) # AIC = -1.323425  AICc = -1.315481  BIC = -1.180053
```

```
sarima(choice_two,0,0,1) # AIC = -1.297257  AICc = -1.293358  BIC = -1.189728
sarima(choice_two,0,0,2) # AIC = -1.264288  AICc = -1.256344  BIC = -1.120916
sarima(choice_two,0,0,3) # AIC = -1.312767  AICc = -1.299271  BIC = -1.133552
sarima(choice_two,1,0,1) # AIC = -1.262494  AICc = -1.254549  BIC = -1.119122

sarima(choice_three,1,0,0) # AIC = 1.816965  AICc = 1.820864  BIC = 1.924494
sarima(choice_three,2,0,0) # AIC = 1.753953  AICc = 1.761897  BIC = 1.897325
sarima(choice_three,0,0,1) # AIC = 1.778947  AICc = 1.782845  BIC = 1.886476
sarima(choice_three,0,0,2) # AIC = 1.79935  AICc = 1.807294  BIC = 1.942722
sarima(choice_three,0,0,3) # AIC = 1.763922  AICc = 1.777417  BIC = 1.943137
sarima(choice_three,1,0,1) # AIC = 1.763922  AICc = 1.777417  BIC = 1.943137


ARMAtoMA(ar=-0.3065,ma=0,10)

l=forecast(arima(choice_two,order=c(2,0,0)),h=8)
autoplot(l)

l=forecast(arima(choice_two,order=c(0,0,3)),h=8)
autoplot(l)
```

References:

https://www.britannica.com/topic/history-of-Central-African-Republic
https://www.cia.gov/the-world-factbook/countries/central-african-republic/#economy