Forecasting International Price of Coffee Beans

ECON 4395 Time Series Analysis and Forecasting

TABLE OF CONTENTS

Figures

Tables

1 Introduction

Coffee beans have been increasing their presence in many international markets. Traditionally, people in the middle east and Europe appreciated the beverage. Today, the United States and European nations are the primary consumers; yet, with economic developments, some markets in Asia, such as Japan, South Korea, China, and India, which traditionally consumed tea, have been nurturing their coffee culture. The market may observe a similar pattern in other developing countries as local people become more financially independent. The price of coffee beans affects the global economy mainly for three reasons: consumer behavior, economic impact, and small farmers' standard of living; forecasting enables businesses to analyze and comprehend the market in depth.

While coffee beans tend to stay inelastic, changes in the price can alter buyers' behaviors to purchase substitution goods such as energy drinks or tea in the long run. (ICO, 2019). Many people appreciate the beverage as an essential part of their lives mainly due to the caffeine. The caffeine market includes many substitution goods, challenging the multi-billion industry. The coffee market generates so much money, and some nations heavily depend on it.

About half of the leading coffee exporters fail under the least developed countries category (ICO, 2019), and a change in coffee price can impact local people. While economic development requires industrialization, an increase in agricultural productivity contributes to reducing poverty. (World Bank 2008). Many nations that produce coffee beans lack other non-agricultural industries they can depend on. The demand for cheap coffee matters far more for producers than consumers.

Not diversifying income sources and depending on coffee alone expose small farmers to the international price. According to Coffee Development Report 2019, producer price and international price are co-integrated, meaning the two prices share similar price movements in the long run. (ICO, 2019). The same report also says,

"Formal risk management tools, such as hedging in futures markets, are often too complex and too costly for small farmers and remain a viable option only for larger or aggregated producers." (ICO, 2019).

Consequently, forecasting possesses the potential to protect small farmers from international prices and stabilize their revenue. Essential forecasting has been providing some advantages for larger farmers.

Researchers can utilize ICO price indicators to comprehend the market at different levels: consumers in developed nations, the economy in developing countries, and producers in developing nations. Forecasting is an advantageous technique to avoid risk but far too complicated for small farmers. This paper attempts to discover the best forecasting methods for ICO indicator prices.

2 Literature Review

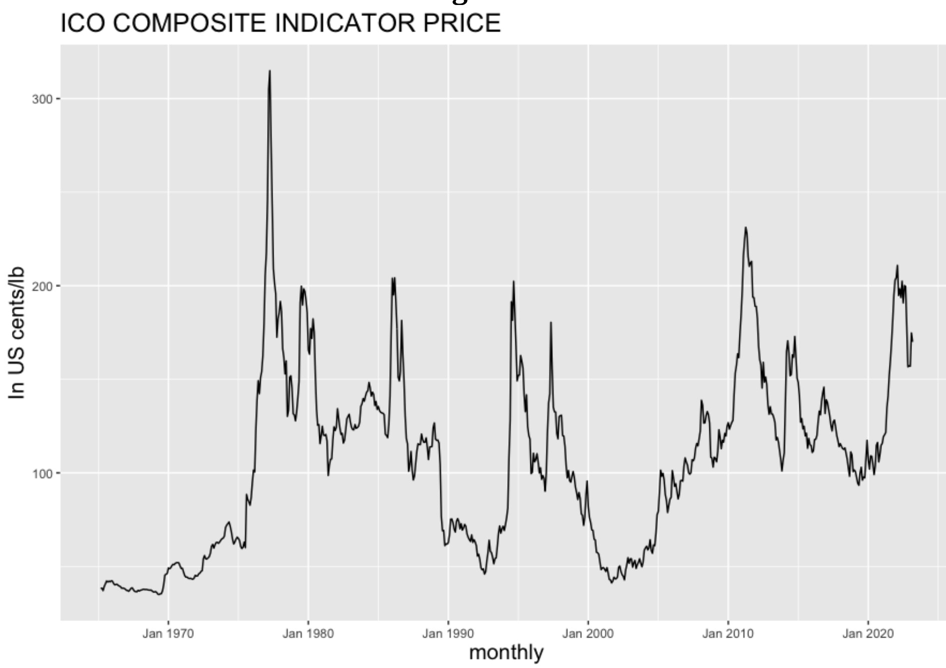
Dynamic Modeling and Forecasting of Data Export of Agricultural Commodity by Vector Autoregressive Model, a journal of Southwest Jiaotong University, studied three commodities in Indonesia, coffee beans, cacao beans, and tobacco, to determine the best agricultural crops to sell between 2007 and 2018. The study concluded that VAR (2) model most accurately predicted the price by using other similar commodities to forecast each other for the VAR.

Using Artificial Neural Networks and Optimal Scaling Model to Forecast Agriculture Commodity Price: An Ecological-economic Approach by Roberto Louis Forestal and Shih-Ming Pi employs an even more advanced method by considering the social aspects when forecasting. The research used three data types (cocoa, coffee, and crude oil) and divided the process into three steps (input, process, and output). The input part includes the commodity supply, price of complementary/substitute, and ecological indicators. The process part adjusts for the social influence to sophisticate the forecasting. The output is the commodity price. The research adopted optimal scaling regression with nonlinear transformations and artificial neural networks and reported the advantage of both methods. Forestal and Pi wrote that the optimal scaling regression method could explain the variability better than the neural networks, but the neural networks can predict prices more accurately.

3 Forecasting International Coffee Beans Prices: Empirical Analysis

3.1 Data Analysis

Figure 1



The data above describes the international coffee bean price change from March 1965 to March 2023 using International Coffee Organization (ICO) data. The ICO composite indicator is one of the most trusted and oldest data sets for the price of international coffee beans. ICO employs an original system to determine costs by weighing nations' coffee consumption. Countries that consume a lot of coffee can influence the ICO indicator more than nations that destroy only a little bit. The United States and European nations contribute the most to the indicator. While the data is historical, it is not adjusted for inflation.

Figure 2

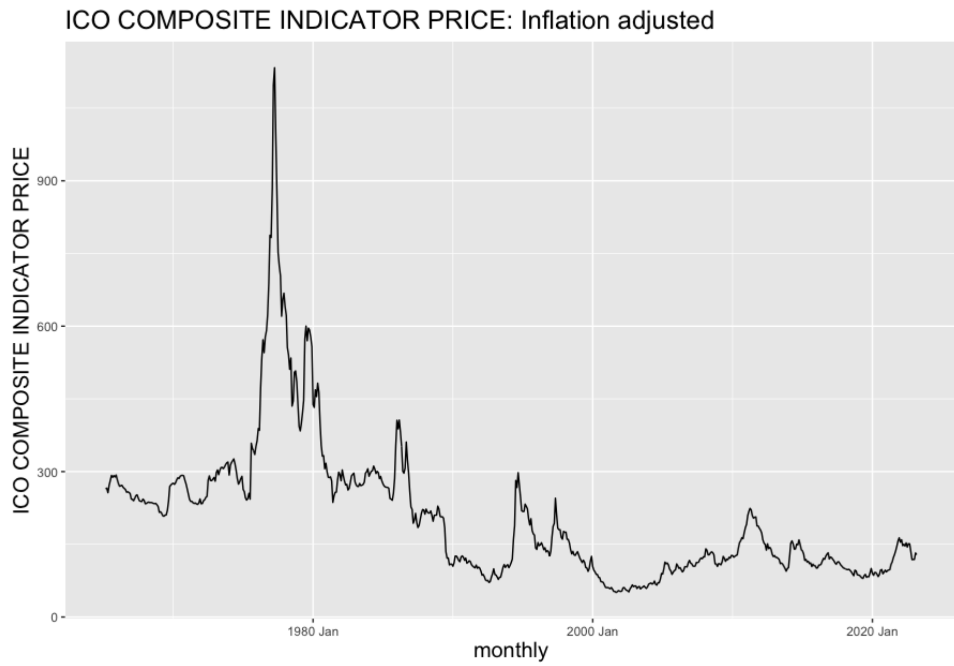


Figure 2 incorporates the Consumer Price Index (CPI) with CPI 100 in 2010. The transformation highlights the significant spike around 1975. The New York Times article lists international events, such as Angola's civil war, worldwide inflation, energy cost, and frost in Brazil. ICO mentions only the frost in Brazil, which destroyed two-thirds of Brazil's coffee bushes. After its peak in 1975, the price gradually decreased and stabilized. It hit its lowest price in 2002.

The paper focuses on the Figure 1 basic data but mentions some other results with adjustments, inflation and ignoring the 1975 frost in Brazil, in section 4 Adjustments. Originally, I used the same data set from Figure 2 for the forecasting, yet none of the models' residuals showed white noise characteristics. I conducted two adjustments to keep the residual white noise and construct an accurate model. First, I eliminated the old data to filter the outlier around 1975. (Figure 3). The overly special event might impair the overall quality. Second, I returned to the original data without the CPI adjustments: from Figure 2 to Figure 1. CPI adjustment clarifies the real price change and is helpful for understanding the basics, yet the adjustment seemed to overcomplicate the forecasting process and return inaccurate results with high autocorrelation in the residuals. Figure 4 data sets combine the two methods, inflation and time trimming.

Figure 3

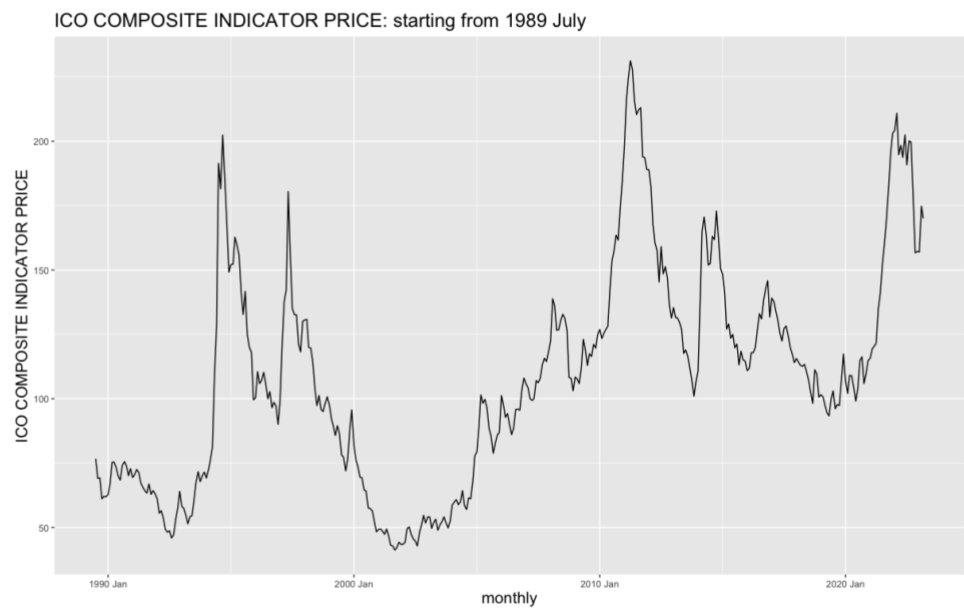
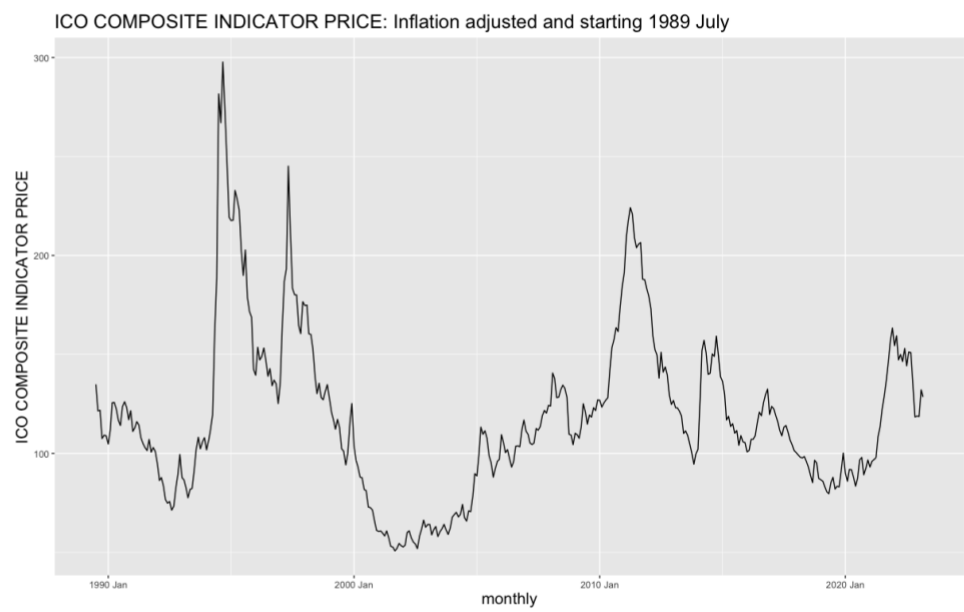
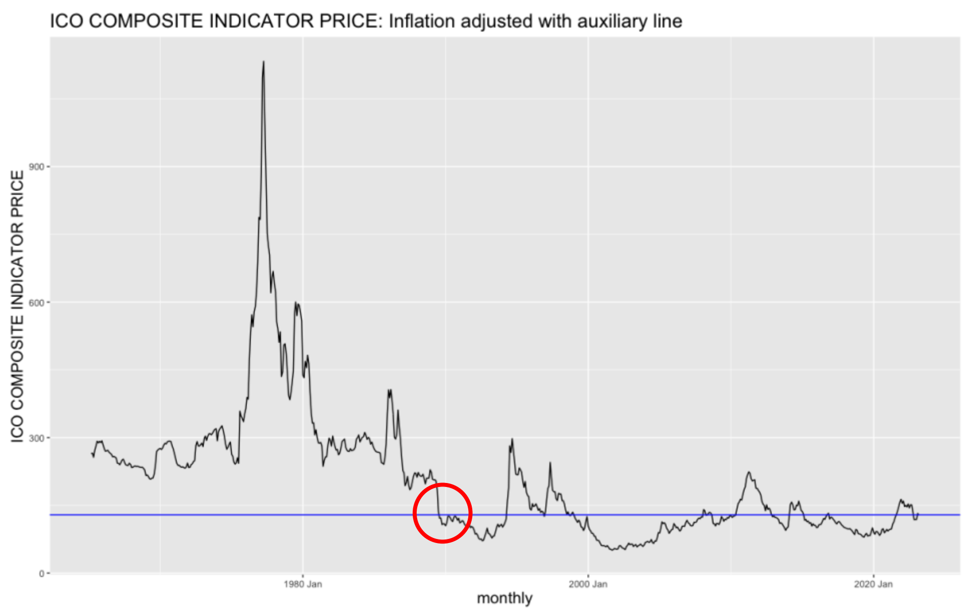


Figure 4



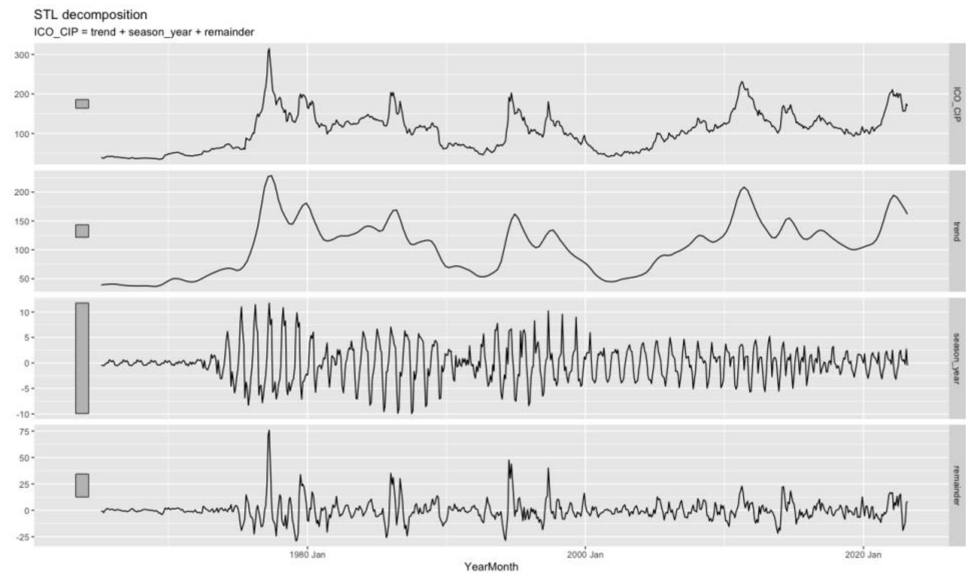
I added a straight auxiliary line to Figure 2 to find the starting point where the real price becomes stable (Figure 5). The blue line indicates the March 2023 price of 128.5, and the blue line does not overlap the data before July 1989: the red circle, and data do not possess drastic outliers after July 1989. By choosing the March 2023 price, I intended to minimize the trend effect and keep the new data stationary. The paper analyzes the adjustments in section 4 Adjustments. The rest of this section utilizes the untouched data sets from Figure 1.

Figure 5



3.2 Data Description

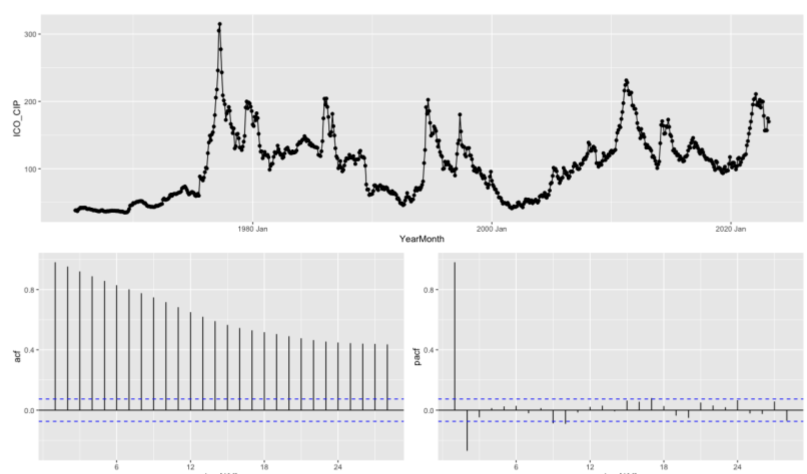
Figure 6



The STL decomposition shows a weak but long trend and small seasonality. (Figure 6). To further understand the data set, I reviewed the ACF and PACF. The ACF shows a steady trend without obvious seasonality, a unit root and non-white-noise features. I must difference the data for some models. (Figure 7).

Figure 7

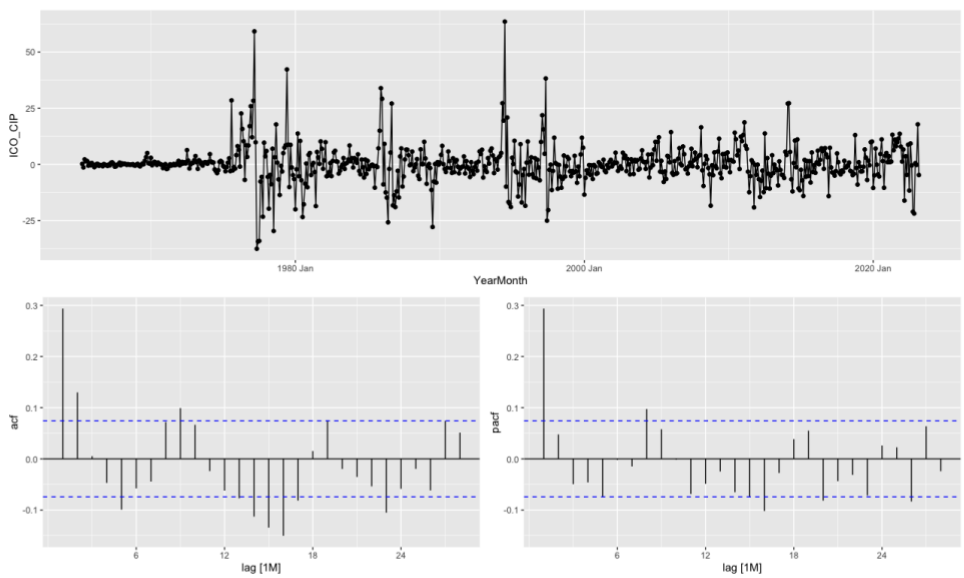
ACF and PACF test



For some models, non-stationary data requires differencing. Figure 8 is the first differenced version. Some spikes in ACF exceed the confidence interval, non a white noise data set, but the kpss test did not require a second difference. Among the five methods, Exponential Smoothing, Time Series Decomposition, ARIMA, VAR, and Neural Network, only three passed the Ljung-Box test-- ARIMA, VAR, and Neural Network.

Figure 8

ACF and PACF tests after differenced



ARIMA

Figure 4 indicates a differencing, so the estimated ARIMA models derived from Figure 8 are I(1) and AR(1) or MA(2)—ARIMA(1,1,0) and ARIMA(0,1,2). The step wise method suggested ARIMA(0,1,2), and the search method suggested ARIMA(2,1,3)(0,0,1). Table 1 describes the ARIMA models. ARIMA(2,1,3)(0,0,1) lowered the RMSE the most. As for the information criterion, for AICc, ARIMA(0,1,2) and, for BIC, ARIMA(0,1,1) performed the best, respectively. I have chosen ARIMA(2,1,3)(0,0,1) for my model's overall good performance. Figure 9 describes the residuals. While some spikes surpass the confidence intervals, the Ljung-Box test's p-value is 0.623, which implies no autocorrelation.

Table 1

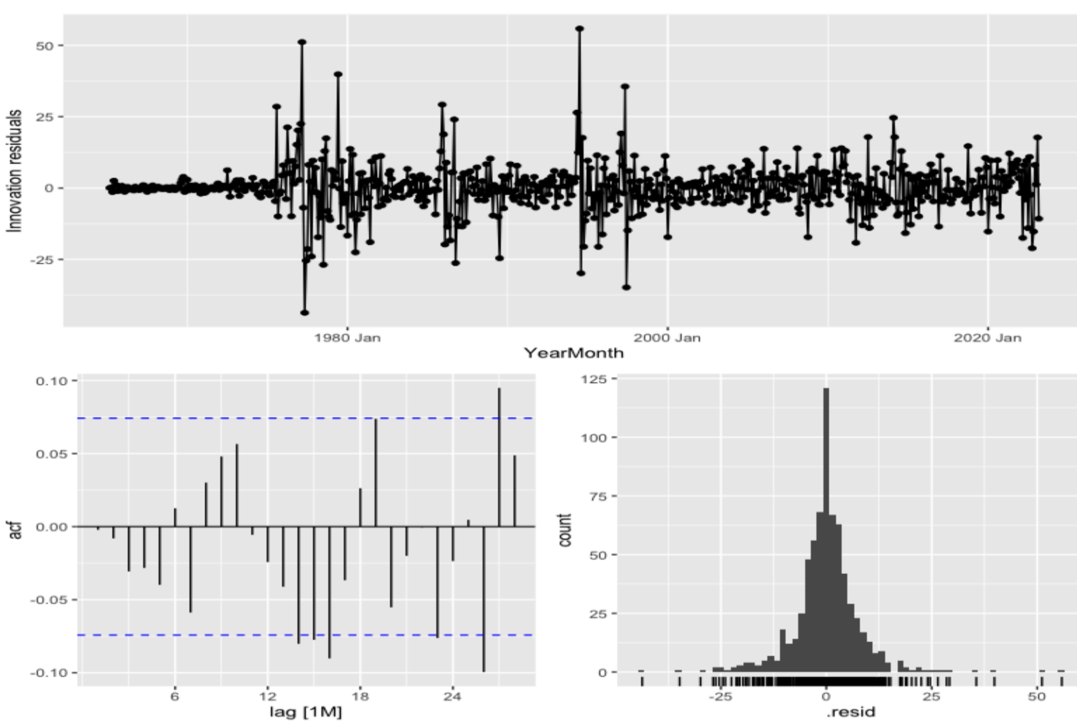
ARIMA RMSE and IC

	.type	RMSE		.model	.type	RMSE
arima011	Training	8.27		arima011	Test	18.3
arima012	Training	8.26		arima012	Test	18.5
stepwise012	Training	8.26		search213001	Test	17.8
search213001	Training	8.15		stepwise012	Test	18.5

.model	AICc	BIC
arima011	4836	4845
arima012	4837	4850
search213001	4837	4850
stepwise012	4827	4859

Figure 9

Innovation residuals for ARIMA

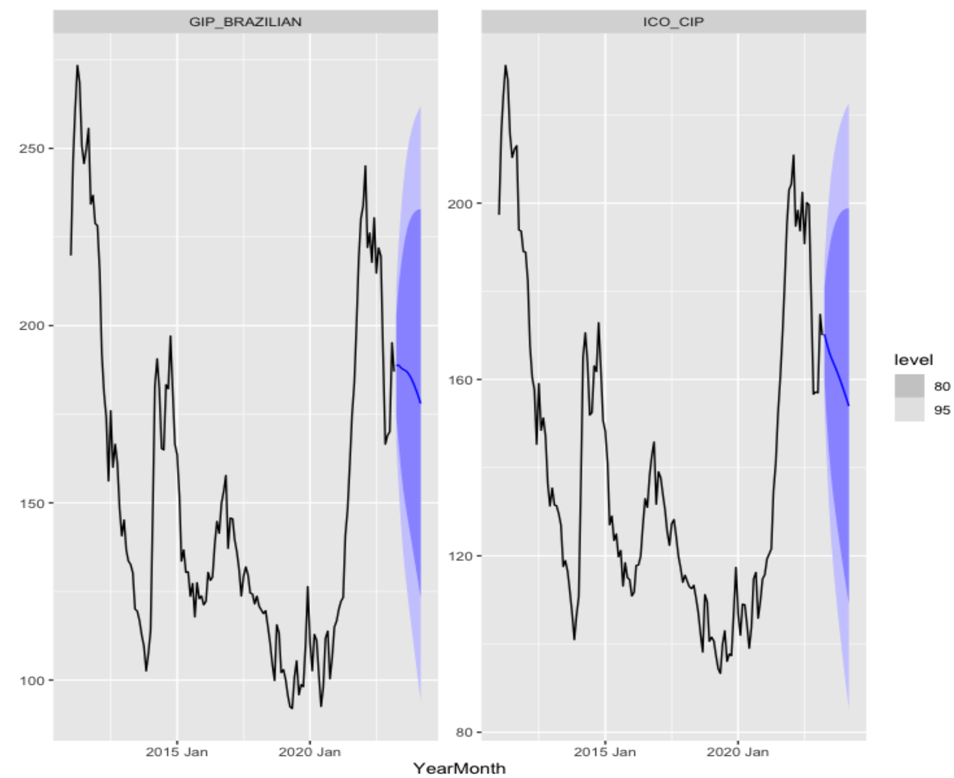


VAR

I selected the Brazilian beans' price for two reasons. First, the journal of Southwest Jiaotong University mentioned that the international price fluctuates with Brazilian coffee beans. Second, the two move together – co-movement. Figure 10 demonstrates the similar movement between the two.

Figure 10

VAR Forecast



The R programming calculated VAR(5) for aicc and VAR(3) for bic. Table 2 reveals that aicc-- VAR(5)-- outperformed in all testing in-sample RMSE, out-of-sample RMSE, and Information criteria (AICc and BIC). The p-value for the Ljung-Box test was 0.301 (ICO) and)00962(Bra). I have presented the model regardless of the autocorrelation in the Brazil data set since our primary focus is the ICO price (Figure 10). Figure 11 suggests no autocorrelation in the residuals. AICC is my choice for VAR.

Table 2

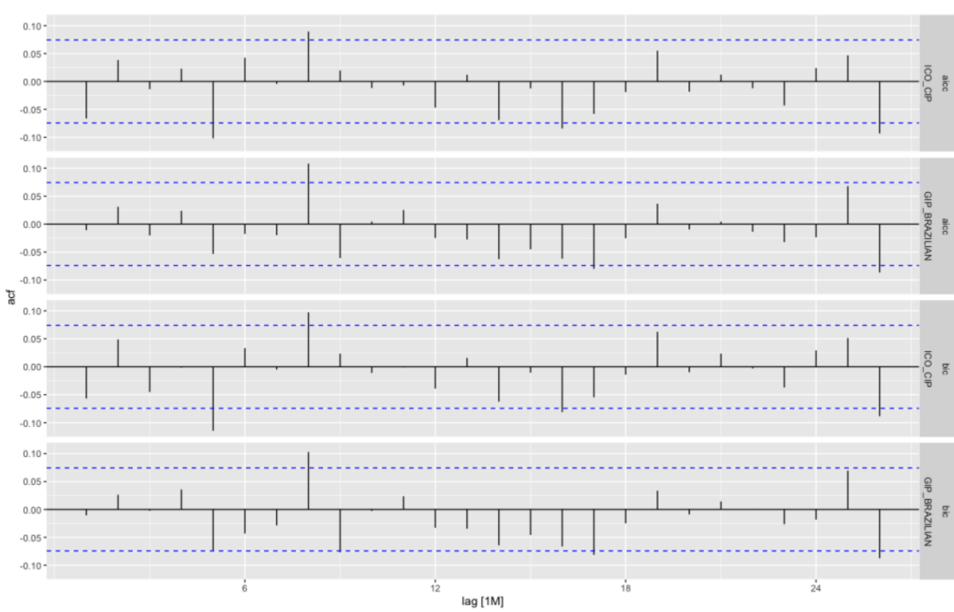
VAR RMSE and IC

	.type	RMSE		.model	.type	RMSE
aicc	Training	8.1		aicc	Test	17.5
aicc	Training	11.0		aicc	Test	14.3
bic	Training	8.13		bic	Test	17.0
bic	Training	11.2		bic	Test	14.9

.model	AICc	BIC
aicc	9429	9544
bic	9461	9542

Figure 11

VAR residual ACF



Neural Network Model

As Roberto Louis Forestal and Shih-Ming Pi concluded Neural Network Model scored the low RMSE values, yet the model does not excel at explaining the data set. NNAR (13,1,7)

[12] received 5.88 in-sample RMSE and 14.3 for out-of-sample RMSE. The model's residuals do not correlate to each other-- p-value 0.955 using the Ljung-Box test (Figure 12).

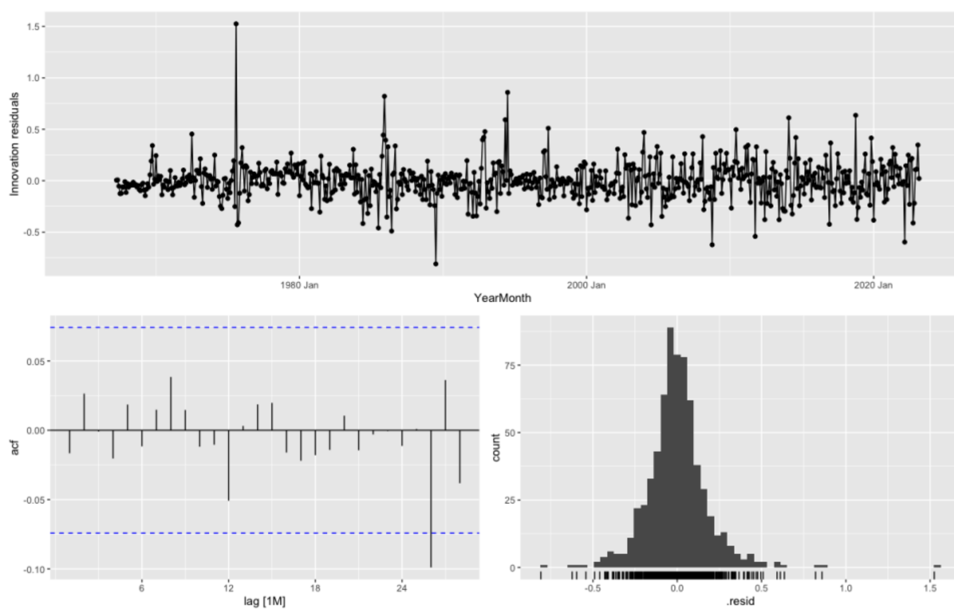
Table 3

NNETAR RMSE and IC

	.type	RMSE		.model	.type	RMSE
NNETAR(sqrt(ICO_CIP))	Training	8.1		NNETAR(sqrt(ICO_CIP))	Test	14.4

Figure 12

NNETAR residual ACF



3.3 Model Comparison

Table 4

Model Comparison

No CPI from 1965 March	Figure 1		
	ARIMA	VAR	Neural Network
Estimated Model	ARIMA(213)(001)	aicc (VAR(5))	NNETAR(sqrt(ICO_CIP))
BIC	4859	9544	na
AICc	4827	9429	na
RMSE (In-sample)	8.15	(ICO)8.1(Bra)11.0	5.88
RMSE (out-of-sample)	17.8	(ICO)17.5(Bra)14.3	14.3
ljung_box	0.623	(ICO)0.301(Bra)0.00962	0.955

Table 4 summarizes all the RMSE numbers, and the Neural Network model resulted in the best model among the three with the lowest RMSE for the in-sample and the out-of-sample and low correlation among the residuals. Both ARIMA and VAR can still forecast the data well, but the high correlation in residuals in VAR Brazil (0.00962) is concerning. For the final result comparison with realized data, I used the Neural Network.

4 Adjustments

Figure 13

Four data sets

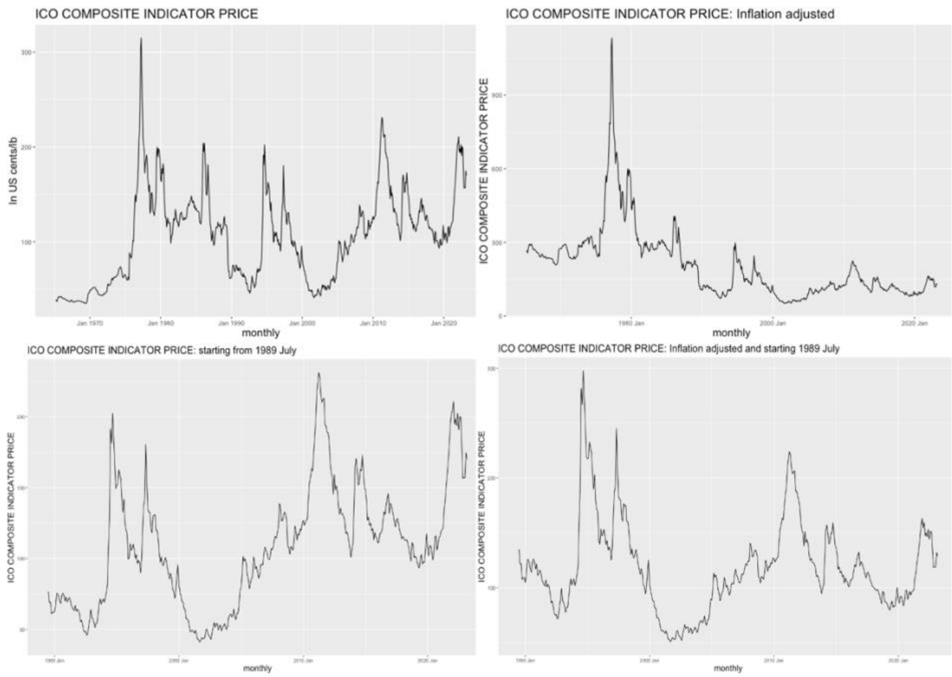


Figure 13 includes Figures 1, 2, 3, and 4 from the top left corner. The RMSE for Figures 1 and 4 are in Table 5. RMSE is scale dependent, but all four data sets use cents as their unit, which keeps the residuals comparable. Models generated by data sets from Figures 2 and 3 did not perform better than the untouched Figure 1 data. For the ARIMA and VAR models, Figure 4, which combines two adjustments—inflation and omitting outliers—generated smaller RMSE for the out-of-sample data compared to unadjusted Figure 1 data. The smaller RMSE for the test indicates that Figure 4 avoided overfitted issues. The difference between in and out-of-sample RMSE shrunk. In addition to the out-of-sample RMSE, Ljung-Box test results imply less autocorrelation among the residuals. In contrast, the transformation disvalued Neural Network's RMSE for the out-of-sample.

Table 5

No CPI from 1965 March	Figure 1		
	ARIMA	VAR	Neural Network
Estimated Model	ARIMA(213)(001)	aicc (VAR(5))	NNETAR(sqrt(ICO_CIP))
BIC	4859	9544	na
AICc	4827	9429	na
RMSE (In-sample)	8.15	(ICO)8.1(Bra)11.0	5.88
RMSE (out-of-sample)	17.8	(ICO)17.5(Bra)14.3	14.3
ljung_box	0.623	(ICO)0.301(Bra)0.00962	0.955

CPI from 1989 July	Figure 4		
	ARIMA	VAR	Neural Network
Estimated Model	ARIMA(1,0,5)	aicc (VAR(5))	NNETAR(sqrt(ICO_CIP))
BIC	2973	5193	na
AICc	3004	5093	na
RMSE (In-sample)	9.28	(ICO)9.30(Bra)11.6	3.16
RMSE (out-of-sample)	10.9	(ICO)10.4(Bra)12.9	23.6
ljung_box	0.997	(ICO)0.641(Bra)0.952	0.999

Despite the positive Figure 4 data's ARIMA and VAR performances, people should use the models created from the untouched Figure 1 data for three reasons. First, I could not find the reasons behind the sudden price drop around 1989. Second, the current CPI data I am using is not perfect and has room for improvement. Choosing the appropriate CPI increases reliability. Third, the models from unadjusted Figure 1 data do not possess residual autocorrelation or other issues. However, once we overcome the first and second reasons, we might be more confident to use the Figure 4 data than the Figure 1 data.

5 Conclusion

Forecasting the international price of coffee beans allow people to construct an estimation for local markets since many local coffee beans' prices for producer and international price are co-integrated. In many cases, the ability to forecast divides successful and unsuccessful farmers due to the difference in risk management performances, such as hedging future markets. Data provided by the International Coffee Organization (ICO) enabled us to peek the future prices.

For a simple prediction, I concluded that Neural Network performs the best, yet the model does not provide insights to comprehend the data sets. On the other hand, ARIMA(2,1,3)(0,0,1) can explain the characteristics such as unit root, trend, and seasonality. The VAR model is acceptable, but a better result from the Ljung-Box test for Brazil can increase confidence in the model.

I attempted three transformations to the data to further improve the data to obtain lower out-of-sample RMSE. Simply applying inflation or removing the outliers did not improve the models, yet simultaneous adjustments for inflation and outliers dwarfed the out-of-sample RMSE for ARIMA and VAR. While the results imply that appropriate transformation can aid the forecasting, the transformation requires further investigation of the international coffee price history and an improved CPI selection method.

6 References

International Coffee Organization [ICO]. (2019). *Coffee Development Report 2019*. International Coffee Organization. http://www.ico.org/documents/cy2021-22/coffee-development-report-2019.pdf

World development report, 2008: agriculture for development. (2008). *Choice Reviews Online*, *45*(09), 45–4765. https://doi.org/10.5860/choice.45-4765

SOUTHWEST JIAOTONG UNIVERSITY [西南交通大学学报]. (2020). DYNAMIC MODELLING AND FORECASTING OF DATA EXPORT OF AGRICULTURAL COMMODITY BY VECTOR AUTOREGRESSIVE MODEL. *JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY*, *Vol. 55*(No. 3). https://doi.org/10.35741/issn.0258-2724.55.3.41

Forestal, R. L., & Pi, S. (2021). Using Artificial Neural networks and Optimal Scaling Model to Forecast Agriculture Commodity Price: An Ecological-economic Approach. *Advances in Management and Applied Economics*, 29–55. https://doi.org/10.47260/amae/1133

Times, N. Y. (1975, August 4). Frost in Brazil Sending Coffee Prices Up. *The New York Times*. https://www.nytimes.com/1975/08/04/archives/frost-in-brazil-sending-coffee-prices-up-frost-in-brazil-raises.html

7 Appendix

Table 6

NNETAR(sqrt(ICO_CIP)) forecast			
.model	YearMonth	ICO_CIP	.mean
NNETAR(sqrt(ICO_CIP))	2023 Apr	sample[5000]	164.801875
NNETAR(sqrt(ICO_CIP))	2023 May	sample[5000]	155.723088
NNETAR(sqrt(ICO_CIP))	2023 Jun	sample[5000]	151.342628
NNETAR(sqrt(ICO_CIP))	2023 Jul	sample[5000]	147.40429
NNETAR(sqrt(ICO_CIP))	2023 Aug	sample[5000]	141.564875
NNETAR(sqrt(ICO_CIP))	2023 Sep	sample[5000]	140.878677
NNETAR(sqrt(ICO_CIP))	2023 Oct	sample[5000]	144.862851
NNETAR(sqrt(ICO_CIP))	2023 Nov	sample[5000]	147.094116
NNETAR(sqrt(ICO_CIP))	2023 Dec	sample[5000]	147.735596
NNETAR(sqrt(ICO_CIP))	2024 Jan	sample[5000]	150.780964
NNETAR(sqrt(ICO_CIP))	2024 Feb	sample[5000]	159.643531
NNETAR(sqrt(ICO_CIP))	2024 Mar	sample[5000]	163.112506

Table 6 is the Neural Network, yet it does not show an interval forecast. Thus, I added Table 7 ARIMA(2,1,3)(0,0,1).

Table 7

ARIMA213001 forecast				
.model	YearMonth	ICO_CIP		.mean
search213001	2023 Apr	N(172	68)	171.957241
search213001	2023 May	N(172	179)	172.478032
search213001	2023 Jun	N(171	319)	171.094073
search213001	2023 Jul	N(168	476)	168.339181
search213001	2023 Aug	N(167	640)	167.246392
search213001	2023 Sep	N(169	796)	168.749509
search213001	2023 Oct	N(171	937)	170.968045
search213001	2023 Nov	N(172	1075)	171.506917
search213001	2023 Dec	N(170	1222)	170.010655
search213001	2024 Jan	N(168	1379)	168.284151
search213001	2024 Feb	N(168	1538)	168.151437
search213001	2024 Mar	N(170	1688)	169.555979

Table 8

Realized data in April and predictions	
Realized data	178.71
ARIMA213001	171.96
VAR(5)	170

For the inflation-adjusted models, I need to convert the realized data as well. 178.71 cents with CPI = 100 at the April 2010 level is 135.08. I used CPI for 2022 (132.3) for 2023 as well. 178.71/132.3*100=135.08.

Table 9

Adjusted realized data in April and models	
Realized data	135.08
ARIMA 105	132
VAR (5)	131