Crop Price Prediction: A Comparison of the Recursive and Direct Forecasting Strategies on SARIMAX Models

Karishma Harrykissoon and Patrick Hosein
Department of Computing and Information Technology
The University of the West Indies
St. Augustine, Trinidad and Tobago
karishma.harrykissoon@my.uwi.edu, patrick.hosein@sta.uwi.edu

***Abstract*—External factors, like weather conditions, cause crop price fluctuations, which negatively affect farmers both financially and emotionally. It is therefore important to forecast crop prices so farmers can monitor changes in prices and determine better planting, harvest and selling dates, to maximise profits. The goal of this study is to develop statistical time series forecasting models to predict the future prices of tomato, banana, orange, ginger, hot pepper, lettuce, pumpkin, cabbage cucumber, and cassava. Variations of the Seasonal AutoRegressive Integrated Moving Average with exogenous factors (SARIMAX) were used, where the exogenous variables were volume traded, precipitation, maximum temperature and minimum temperature. To improve forecasting accuracy, the recursive and direct forecasting strategies were compared using the Mean Absolute Percentage Error (MAPE) performance metric. One-step forecasting was also performed and compared with the Naïve Seasonal Mean Model. For seven crops, the direct strategy outperformed the recursive strategy. Also, the recursive one-step forecasts outperformed the Naïve Seasonal forecasts for seven of the crops.**

***Index Terms*—ARIMA, SARIMAX, direct forecasting, recursive forecasting, MAPE**

I. Introduction

In the 1970s, Trinidad and Tobago's agriculture sector contributed to approximately 7% of the nation's Gross Domestic Product (GDP), however, it reached to a minimum of 0.37% in 2007. There is a slow increase in this value since, with the GDP now being about 1%, as at the second quarter of 2020 [1]. This study focuses on eight locally produced crops, tomato, orange, ginger, hot pepper, lettuce, pumpkin, cucumber, and cassava, and two imported crops, banana, and cabbage.

One of the major factors affecting crop prices is the volume produced, as this is linked to demand and supply. As at the first quarter of 2021, there has been a decrease in total vegetable production by 29.5%, with significant reductions in cucumber, tomatoes, and pumpkin. Some of the contributing factors were excessive rainfall and the COVID-19 pandemic. The latter had a greater impact on the demand for crops due to decreased operations of retail and wholesale markets and temporary closure of the food industry [2].

Arising out of the pandemic is global inflation. As at January 2022, the food inflation value was 6.6%. Also, a decrease in foreign exchange and increases in shipping and freight costs, impacted the prices of imported crops and fertilizers. This caused interruptions in the value chain analysis for crop production, resulting in consumers paying more for fresh produce. Furthermore, majority of the food supply is imported with the annual food import bill being around TT$5 billion in 2022. Hence, the government has been encouraging consumers to purchase locally in order to boost crop production [3].

Trinidad and Tobago has a dry season from January to May, and a wet season from June to December. In recent years, there has been changes in these seasons owing to climate change. The islands experienced a drought between 2009 and 2010, which caused water shortages and agricultural losses. This caused a 6.9% increase in food prices in March, 2010 compared to 2.7% in January, 2010 [4]. Recently, there has been an increase in annual rainfall causing more frequent flash flooding across the country. This also negatively impacts the farmers and crop prices on a yearly basis [5].

By developing forecasting models to predict crop prices, government can monitor such and implement necessary measures to better assist the farmers. These include adjusting the market prices for crops so farmers can obtain fair prices for their harvests, and regulating the import and export industry so more crops are sold locally and tariffs are removed on some imported crops. Farmers will also be able to monitor the prices so they can make more informed decisions on planting, harvesting and selling dates to obtain maximum profits. Given that external factors, like weather, play a significant role in crop cultivation, farmers will also be able to monitor weather conditions for crops that are affected by weather fluctuations.

In this study, statistical time series forecasting models are used to forecast the prices of the ten crops, over a horizon of 12 months. These models include the AutoRegressive Integrated Moving Average (ARIMA) and Seasonal AutoRegressive Integrated Moving Average with exogenous factors (SARIMAX), with the latter taking into account external factors such as rainfall, temperature and volume traded. With the aim of improving forecasting accuracy, the recursive and direct forecasting strategies are compared for multi-step forecasting, and one-step forecasting is also done, using the Mean Absolute Percentage Error (MAPE) performance metric.

II. Literature Review

In [6], univariate ARIMA models were developed to forecast cotton prices in six states in India. Model parameters were determined via AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF) plots, and model performances were assessed using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and MAPE. Although all of the MAPE scores were in the range of 4.88% to 7.40%, the models failed to incorporate external factors, like the climatic conditions of the Indian monsoon.

Tomato prices were forecasted for five states in India, and the entire country, using Seasonal AutoRegressive Integrated Moving Average (SARIMA) models. Model parameters were determined by ACF and PACF plots and models were selected based on BIC, Coefficient of Determination (R^2), Root Mean Square Error (RMSE), MAPE and Mean Absolute Error (MAE). The MAPE ranged from 18.1% to 47.7%. However, no external factors were considered for these models [7].

The prices of areca nuts in India were forecasted using a SARIMA model, Holt-Winter's seasonal method, and Long Short-Term Memory (LSTM). The SARIMA(1,0,0)(0,1,2,12) was determined by AIC, BIC, and Akaike's Information Corrected Criterion (AICc), with RMSE as the performance evaluation metric. Although, statistical methods are commonly used to predict crop prices, LSTM performed the best. This study did not consider any exogenous variables [8].

The daily sales of banana, in a retail store in Germany was forecasted using SARIMA and SARIMAX models. The parameters for the SARIMA(0,0,3)(1,0,1,6) was determined from ACF and PACF plots. The exogenous variables included promotional price reductions and holiday effects. Using MAPE, RMSE and adjusted R^2 as performance evaluation metrics, the SARIMAX model outperformed the SARIMA. SARIMAX with all of the exogenous variables gave a reduced MAPE score of 5.87%, suggesting that external factors do have an influence on forecasting accuracy [9].

The daily hotel demand for a hotel in the United States was forecasted using SARIMAX, seasonal naïve, Holt-Winters' Triple Exponential Smoothing, ARIMA, Artificial Neural Network- Multi-Layer Perceptron (ANN-MLP), Generalized AutoRegressive Conditional Heteroscedasticity (sGARCH), and Glosten–Jagannathan–Runkle GARCH (GJR-GARCH) models. The external regressors included temperature and holidays. Multi-step forecasting was done for seven horizons using direct, or rolling, forecasting. Pairwise comparisons of the six statistical models with ANN-MLP, using relative MAPE scores, showed that the SARIMAX model outperformed the others for six of the seven horizons [10].

There are currently three strategies for multi-step forecasting, namely the recursive, direct, and DirRec, or joint, strategies, with the latter combining the advantages of the first two. Four of the aforementioned papers used the recursive strategy, with [10] implementing the direct approach. Although there is insufficient literature on comparing these strategies to traditional statistical models, there are instances where they were applied to machine learning models.

Reference [11] implemented these three strategies on data from the M3 and NN5 forecasting competitions, using the k Nearest Neighbours (kNN) algorithm. For M3, the recursive strategy was better than the direct. The DirRec gave slightly better Symmetric Mean Absolute Percentage Error (SMAPE) and MAPE scores than the recursive method. For NN5, the direct strategy was better than the recursive, possibly due to this dataset having more non-linearity than the M3. The DirRec gave marginally better metric scores to the direct method. Although both the direct and recursive methods can be used, it was suggested for the DirRec to be used in real world applications to avoid having to make a choice between the other two strategies.

The performances of these methods were compared using the kNN algorithm and Forward-Backward input selection for a horizon of 100 time steps, on the Santa Fe and Poland Electricity Load time series. For both datasets, the Mean Squared Error (MSE) reduced by approximately two-thirds using the direct strategy. The DirRec method was the best performing, giving an approximate 20% reduction in the MSE compared to the direct strategy. Although the DirRec strategy gave a particularly good accuracy, and the direct method had a greater computational time than the recursive, it was suggested that the direct strategy be utilized if computational time is of concern [12].

The major disadvantage of using the SARIMAX model for multi-step forecasting is that the exogenous variables need to be forecasted first, before the target variable. Since more predictions are required, this leads to accumulation of prediction errors. Hence, the predictions quickly degrade and the model loses forecasting accuracy as the horizon increases. In order to avoid this, it is recommended to forecast a single timestep, wait for the values of the exogenous variables to be observed, and then forecast the other timestep [13].

In this study, time series forecasting of ten crop prices are done using the ARIMA, SARIMA and SARIMAX models, taking into account the recursive and direct forecasting strategies for multi-step forecasting. The exogenous variables considered are volume traded, rainfall, and minimum and maximum temperatures. ACF and PACF plots are not ideal for determining the order of the parameters q and p respectively, since this study considers models that are not purely moving average (MA) or autoregressive (AR). These plots will either display a decaying or sinusoidal pattern, which is indicative of a combination of AR and MA processes. As such, the parameters will be determined by fitting several models with various combinations of the parameter values and then selecting the best one based on AIC values and residual analyses [13].

Considering the disadvantage of the SARIMAX model, one-step forecasting is also applied to each of these models, using the recursive strategy. The MAPE performance evaluation metric is used in both instances. Finally, a decision is made on whether to forecast a single price for the month of January (one-step forecasting), or to forecast the prices over a horizon of 12 months, from January to December (multi-step forecasting).

III. Relevant Algorithms

This section describes the different techniques which are relevant to the research.

A. ARIMA and SARIMA

The ARIMA is a statistical time series forecasting model and a type of Box-Jenkins model that is used to predict values for non-stationary time series. The current value of a differenced time series depends on its past values, which emanate from the AR(p) portion, and its past errors, which stem from the MA(q) portion. The model is denoted as ARIMA(p,d,q), where p is the number of lagged values, d is the order of integration, and q is the number of lagged errors. It is mathematically represented as:

$$y'_t = C + \Phi y'_{t-1} + \ldots + \Phi_p y'_{t-p} + \mu + \theta_1 \epsilon'_{t-1} + \ldots + \theta_q \epsilon'_{t-q} + \epsilon_t \quad (1)$$

where y'_t is the current value of the differenced time series, C is a constant, $\Phi_p y'_{t-p}$ are the past values of the differenced time series, μ is the mean of the difference time series, $\theta_q \epsilon'_{t-p}$ are the past error terms, and ϵ_t is the current error term.

The SARIMA model takes into consideration the seasonal component of the time series and is denoted as SARIMA(p,d,q)(P,D,Q,m), where the first three parameters are the same as the ARIMA(p,d,q), P is the order of the seasonal AR(P) process, D is the seasonal order of integration, Q is the order of the seasonal MA(Q) process, and m is the frequency.

B. SARIMAX

The SARIMAX model is an extension of the SARIMA model that caters for external variables. It has the same parameters as the SARIMA model and adds a linear combination of the external factors to the SARIMA model. It is represented by:

$$y_t = SARIMA(p,d,q)(P,D,Q,m) + \sum_{i=1}^{n} \beta_i X_t^i \quad (2)$$

where y_t is the current value, and X_t is an exogenous variable.

C. Recursive Forecasting Strategy

In this strategy, the same time series model is used repeatedly to make predictions. The prediction of the next time step is calculated by using the predictions of previous time steps, along with original observations as inputs into the model. For multi-step forecasting, one-step ahead predictions are applied recursively until forecasts are obtained for the entire horizon. However, since more predictions are used as inputs, as the horizon increases, then the cumulative prediction error increases and so the performance of the model quickly degrades. The *forecast()* and *predict()* methods of the SARIMAX model imported from *statsmodels* use this type of forecasting.

D. Direct Forecasting Strategy

Also known as the rolling forecast strategy, a separate model is created for each prediction as a new observation is received from the training set. The prediction of the next time step is calculated by using original observations as inputs only, and so only the normal prediction error is present in the model. Since no prior predictions are used then there is no accumulation of prediction errors and better results are obtained for multi-step forecasting, in some scenarios. However, since the model is re-created for each time step, then this causes an increase in the computational time.

IV. Methodology

A. Description of Dataset

The dataset comprised of two files from the National Agricultural Market Information System (NAMIS) website and one from the Trinidad and Tobago Meteorological Service (TTMS) website. The first contained the average wholesale prices of commodities and the second entailed the total volume of commodities traded monthly at the Norris Deonarine Wholesale Market, Macoya, Trinidad and Tobago, from January 2006 to December 2021. A total of ten of the most purchased crops were selected. The third file consisted of daily precipitation, minimum temperature, and maximum temperature from January 1st, 1981 to December 31st, 2021. Only data from January 1st, 2006 to December 31st, 2021 were selected and the monthly averages were computed. After cleaning, the three files were combined to produce the final dataset which has 192 rows and 26 columns.

B. Experiments

For each of the ten crops, the following were done:

1) Naïve Seasonal Mean forecasts: this was done using the training dataset containing the monthly average prices from January 2006 to December 2020. The corresponding MAPE value was calculated.
2) Time series decomposition.
3) Select exogenous variables: these were determined by testing combinations of the variables and choosing the combination that gave the lowest MAPE score.
4) Stationarity tests: the Augmented Dickey-Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests were used to determine the parameters for non-seasonal (d) and seasonal (D) differencing.
5) Determine model parameters: combinations of the ARIMA(p,d,q), SARIMA(p,d,q)(P,D,Q,m) or SARIMAX(p,d,q)(P,D,Q,m) parameters were tested. The model with the lowest AIC, or one with similar AIC but less parameters to be estimated, was selected.
6) Model identification: for qualitative residual analysis the residual plot, a histogram of the residuals, the Q-Q plot, and the correlogram were examined, while the Ljung-Box test was used for quantitative residual analysis.
7) Recursive forecasts: this was done by using the *get_prediction()* method from the SARIMAX function.

8) Direct forecasts: this was executed by using a user-defined function.
9) Determine the best model: the model with the lowest MAPE score was selected.
10) One-step forecasts: the price for January, 2021 was forecasted, using the Naïve Seasonal Mean Model and the recursive strategy, and the MAPE was calculated.

V. Results

Table 1 summarises the statistical forecasting models for each crop and the MAPE scores. The exogenous variables for tomato were volume traded, lettuce were volume traded, precipitation, maximum and minimum temperatures, and cabbage were precipitation, maximum and minimum temperatures.

The recursive strategy gave better predictions for tomato, ginger and cabbage, which indicates that adding new observations to the models had little or no effect on the forecasted prices. Hence, the prices can be considered as somewhat stationary and therefore follow the same pattern every year. This was evident in the type of differencing done. For these models no seasonal differencing was done despite the time series decompositions indicating seasonality. However, in predicting the prices of cabbage, the MAPE score for the recursive model was only 0.28% better than the direct. For banana, orange, hot pepper, lettuce, pumpkin, cucumber and cassava, the direct method outperformed the recursive method.

Table 2 provides a summary of the MAPE scores calculated for the one-step forecasts using the Naïve Seasonal Mean model and recursive model. For tomato, banana, ginger, lettuce, pumpkin, cabbage, and cassava, the MAPE values for the recursive one-step forecasts were lower than the Naïve Seasonal Mean forecasts. Also, for these same seven crops, the MAPE scores were lower than 10%, suggesting that the models using the recursive method gave highly accurate forecasts for one timestep ahead. For orange, hot pepper and cucumber the Naïve Seasonal Mean model gave lower MAPE scores than the recursive forecasts. Overall, the recursive models performed better.

VI. Discussion

The recursive strategy is biased for non-linear models and its performance tends to degrade for long forecasting horizons.

TABLE I
MAPE Scores for Multi-step Forecasts

Crop	Model	Naïve	Recursive	Direct
Tomato	SARIMAX(0,1,3)(2,0,0,12)	30.78	17.82	19.20
Banana	SARIMA(2,0,1)(0,1,1,12)	12.53	5.72	3.72
Orange	SARIMA(1,0,0)(0,1,1,12)	18.50	23.51	13.97
Ginger	SARIMA(2,1,1)(3,0,3,12)	25.15	14.46	22.47
Hot pepper	ARIMA(1,1,2)	32.28	26.50	16.91
Lettuce	SARIMAX(1,0,3)(0,1,1,12)	17.89	13.11	11.95
Pumpkin	ARIMA(3,1,1)	41.13	39.37	22.50
Cabbage	SARIMAX(0,1,2)(1,0,0,12)	20.44	10.32	10.60
Cucumber	SARIMA(3,1,1)(1,0,1,12)	30.58	30.61	26.36
Cassava	SARIMA(3,0,1)(2,0,3,12)	17.61	9.60	4.66

TABLE II
MAPE Scores for One-step Forecasts

Crop	Naïve	Recursive
Tomato	38.63	0.06
Banana	8.47	1.76
Orange	16.79	23.36
Ginger	36.12	3.93
Hot pepper	11.82	15.78
Lettuce	27.93	9.76
Pumpkin	7.57	0.07
Cabbage	22.59	3.99
Cassava	11.90	2.83

This is due to the error made at one timestep being transferred to consequent timesteps causing an increase in the cumulative errors and bias of the model. However, the direct method performs better for data with non-linearity [11]. Since majority of the crop prices were non-stationary and exhibited strong seasonality and trend components, then this can be indicative of non-linear data. As a result, the corresponding MAPE values were lower for the direct method. Hence, it is better to use this approach and retrain the model with each additional observation since real world data are not stationary.

The MAPE performance metric does not always meet the validity criterion because the percentage error distribution is positively skewed for data that contain outliers in this side of the distribution. The presence of outliers occurs due to external factors and can bias the forecasting model [14]. From exploratory data analysis, nine of the crops had a positively skewed distribution. Therefore, this could have contributed to the MAPE values being greater than 10%.

On comparing the performance of the Naïve Seasonal Mean model and the recursive model for one-step forecasting, the recursive models gave overall better performances for seven crops. The MAPE scores were within the range of 0.06% to 9.76%, indicating highly accurate forecasts, compared to the MAPE values for the multi-steps forecasts which were significantly higher. Specifically, the one-step forecasts were better for the SARIMAX models. This is as a result of reduced prediction error since only one value needed to be forecasted for each of the target and exogenous variables. There was only one prediction error, as opposed to multiple errors being accumulated for a 12 month horizon.

The statistical forecasting models developed will be beneficial, especially to the farmers, as monitoring of the future crop prices can be done. The farmers can determine when will be the optimal time to plant a particular crop so that it can be harvested in a period where maximum profits can be realised, therefore minimising, or avoiding, the negative financial and emotional impacts. In order to give the farmers full access to the benefits, a user-friendly application can be developed to make forecasts based on real-time data.

The prediction of crop prices will also be beneficial to the government. If crop prices are expected to soar locally for a particular period in the future, then the government will be able to implement measures to prevent this, as too high prices can lead to wastage due to lower demand. Some of these include suitable storage of crops that were already harvested in order to stabilise price fluctuations during the year, and regulating the import and export industry. In both these instances, the farmers will be able to obtain fair prices for the harvests.

There can also be an expansion of the various types of local crops planted, as seen in the analysis of cabbage. Since cabbage is an imported crop but the climatic conditions of Trinidad and Tobago had an influence on the price forecasting, then this suggested that green cabbages were imported from a region with similar weather conditions. As a result, this variety can be cultivated locally which will allow farmers to earn additional incomes and will also aid in reducing the food import bill.

VII. Recommendations

There were some limitations of the statistical models developed, as such, recommendations for future work are given to improve forecasting accuracy. The presence of outliers can bias the models. Since majority of the crop price distributions were right skewed, then removal of outliers would lead to more accurate forecasts. Reference [14] acknowledged the suggestion made by Makridakis (1993) to eliminate outliers when computing the MAPE score. A new metric was proposed called the R-MAPE, which includes the outliers but prevents them from dominating the measure of error.

Given that the direct method tends to perform better for non-linear data, and it outperformed the recursive method, then non-linear models should be considered. A neural network approach can be applied to the time series, such as by using LSTM, or the N-BEATS model.

In addition, the DirRec strategy can be used along with an appropriate input selection method since this method gave improved results in its previous applications. It also works well for non-linear data as it reduces the bias and variance of the forecasting model [11], [12]. Although it was only tested on machine learning models, it can be applied to statistical models and compared with the recursive and direct methods.

Structural breaks in a time series can cause large forecasting errors leading to unreliable models. If the date of the break is known, then the Chow test is used to test for the presence of the structural break. If the date is unknown, then unit roots test are used, such as the Zivot-Andrew's and Perron-Vogelsang tests for the presence of one break, and the Clemente-Montanes-Reyes test for two breaks [15], [16].

VIII. Conclusion

Based on this research, for banana, orange, hot pepper, lettuce, pumpkin, cucumber, and cassava, the direct strategy outperformed the recursive strategy. The models developed for banana and cassava were the most accurate, with MAPE scores of 3.72% and 4.66%, respectively. The only advantage of the recursive strategy is that it requires less computational time.

In addition, the recursive one-step forecasts outperformed the Naïve Seasonal forecasts for tomato, banana, ginger, lettuce, pumpkin, cabbage, and cassava. The SARIMAX models for tomato, lettuce and cabbage gave lower MAPE scores for the one-step forecasts than the multi-step forecasts over a 12 month horizon, indicating that it is better to perform one-step forecasting when there are external regressors.

The decision on whether to use one-step or multi-step forecasting depends on its application. One-step forecasting is useful for short term crops that are affected by external factors, like lettuce, which has a harvest time of four to six weeks. Multi-step forecasting is useful for the long term crops, like tomato, which has a harvest time of two to three months.

References

[1] The World Bank. "Agriculture, forestry, and fishing, value added (% of GDP) - Trinidad and Tobago." https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?end=2020&locations=TT&most_recent_year_desc=false&start=1966&view=chart (accessed Jul. 20, 2022).

[2] Ministry of Finance, "Review of the Economy 2021," 2021. [Online]. Available: https://www.finance.gov.tt/wp-content/uploads/2021/10/Review-of-the-Economy-2021.pdf

[3] Ministry of Trade and Industry, "Government Proactive in Mitigating Impact of Rising Global Food Prices," Mar. 30, 2022. [Online]. Available: http://www.news.gov.tt/content/government-proactive-mitigating-impact-rising-global-food-prices#.YugKo3bMJPa (accessed Aug. 05, 2022).

[4] S. L. Beharry, D. Gabriels, D. Lobo, and R. M. Clarke, "A 35-year meteorological drought analysis in the Caribbean Region: case study of the small island state of Trinidad and Tobago," SN Applied Sciences, vol. 1, no. 10, Sep. 2019, doi: 10.1007/s42452-019-1296-4.

[5] L. Fontes de Meira and W. Phillips, "An economic analysis of flooding in the Caribbean The case of Jamaica and Trinidad and Tobago," Studies and Perspectives series - ECLAC subregional headquarters for the Caribbean, vol. 78, 2019.

[6] A. Darekar and A. A. Reddy, "Cotton Price Forecasting in Major Producing States," Economic Affairs, vol. 62, no. 3, pp. 373–378, 2017, doi: 10.5958/0976-4666.2017.00047.x.

[7] A. A. Reddy, "Price Forecasting of Tomatoes," International Journal of Vegetable Science, vol. 25, no. 2, pp. 176–184, Jul. 2018, doi: 10.1080/19315260.2018.1495674.

[8] K. M. Sabu and T. K. M. Kumar, "Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala," Procedia Computer Science, vol. 171, pp. 699–708, 2020, doi: 10.1016/j.procs.2020.04.076.

[9] N. S. Arunraj, D. Ahrens, and M. Fernandes, "Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry," International Journal of Operations Research and Information Systems, vol. 7, no. 2, pp. 1–21, Apr. 2016, doi: 10.4018/ijoris.2016040101.

[10] A. Ampountolas, "Modeling and Forecasting Daily Hotel Demand: A Comparison Based on SARIMAX, Neural Networks, and GARCH Models," Forecasting, vol. 3, no. 3, pp. 580–595, Aug. 2021, doi: 10.3390/forecast3030037.

[11] S. B. Taieb and R. J. Hyndman, "Recursive and direct multi-step forecasting: the best of both worlds," 2012, [Online]. Available: https://www.monash.edu/business/econometrics-and-business-statistics/research/publications/ebs/wp19-12.pdf

[12] A. Sorjamaa and A. Lendasse, "Time Series Prediction using DirRec Strategy," presented at the European Symposium on Artificial Neural Networks, Belgium, 2006.

[13] M. Peixeiro, Time Series Forecasting in Python. New York: Manning Publications Co., 2022. [Online]. Available: https://www.manning.com/books/time-series-forecasting-in-python-book

[14] J. J. M. Moreno, A. P. Pol, A. S. Abad, and B. C. Blasco, "Using the R-MAPE index as a resistant measure of forecast accuracy," Psicothema, vol. 25, no. 4, pp. 500–506, 2013, doi: doi: 10.7334/psicothema2013.23.

[15] Z. Parpiev. Advanced Applied Econometrics. Class Lecture, Topic: "Lecture 9 Testing for unit roots and structural breaks." Westminster International University in Tashkent, Uzbekistan, 2016.

[16] R. Susmel. FINA 8397: Econometrics II: Quantitative Methods in Finance II. Class Lecture, Topic: "Lecture 16 Unit Root Tests." C.T. Bauer College of Business, Houston, TX, 2019.