International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(5s) 2026 pp. 483-495
https://doi.org/10.70917/ijcisim-2026-2747
Article

Predicting 20 years of Agricultural Stability and Security in the Philippines: A Comparative Machine Learning and Time Series Forecasting Approach for Food Security Planning

Rowell Marquez Hernandez

College of Informatics and Computing Sciences
Batangas State University – The National Engineering University
E-mail: rowell.hernandez@g.batstate-u.edu.ph
ORCID ID: https://orcid.org/0000-0002-8748-6271

Abstract: Food security remains a persistent national concern in the Philippines, driven by rapid population growth, climate variability, geopolitical disruptions, and fluctuating agricultural productivity. To support early interventions and strategic planning, it is essential to develop reliable forecasting tools capable of predicting food shortages and surpluses. This study presents a comprehensive time series forecasting model employing multi-decadal historical data (1961–2024) sourced from government agencies and international bodies, covering key indicators related to food production, supply, and consumption. The research focuses on the top 15 major agricultural crops in the Philippines, which collectively represent a significant portion of the national food system and economy. Multiple forecasting algorithms were implemented to identify optimal predictive models for each crop. These include Gradient Boosting Regressor, Decision Tree Regressor, K-Nearest Neighbors, Exponential Smoothing, Grand Means Forecaster, Naive Forecaster, Theta Forecaster, Auto ARIMA, and Prophet. Model performance was rigorously evaluated using statistical metrics such as RMSE, MAE, MAPE, and R^2 to determine accuracy, consistency, and analytical robustness. Results show that several models particularly Gradient Boosting, Prophet, Theta, and Auto ARIMA, successfully captured long-term trends and production shocks, long-term trends, and production shocks. To further determine and investigate the best fitting model in the dataset, PyCaret is also utilize in this study. The forecasting outputs demonstrate strong predictive capability for future crop production and availability. The findings reinforce the importance of data-driven approaches in agricultural planning, especially for countries highly vulnerable to climate and market instability. The proposed forecasting framework serves as a decision-support tool for policymakers, local government units, agricultural agencies, and food industry stakeholders. By accurately predicting potential shortages and surpluses, the model can assist in resource allocation, importation strategies, crop diversification, farm support programs, and long-term food security planning. Ultimately, this research contributes to sustainable agricultural development, improved food resilience, and the broader national goal of ensuring stable and affordable food for all Filipinos.
Keywords: . Crop yield prediction, Auto ARIMA, Prophet forecasting, Predictive analytics, Sustainable agriculture

1. Introduction

The study explores the crucial concept of food security, emphasizing its significance and the challenges it presents, particularly focusing on the situation in the Philippines. With a population exceeding 110 million [1] and a high poverty rate, the country faces difficulties in providing access to sufficient quantities of affordable, nutritious food [2]. This predicament is exacerbated by the Philippines' dependence on imported food, leaving it vulnerable to price fluctuations and external factors that affect food availability and affordability [3][4].

This research provides an overview of time series forecasting methods that can be used for food security applications. We first review the different types of time series data that are relevant for food security forecasting. We then describe various time series forecasting methods, including traditional statistical methods [5] and more recent machine learning approaches [6]. Finally, we discuss some challenges and future directions in time series forecasting for food security. Time series forecasting is a powerful tool that can be used to predict future trends and patterns in a wide range of applications, including finance, economics, marketing, and healthcare [7]. In recent years, machine learning algorithms such as XGBoost have gained significant attention for time series forecasting due to their ability to handle large and complex datasets smoothing, and Moving Average are examples of traditional methods [8][9], however models as diverse as XGBoost [7], RNN [10], Transformer [11] can also be used.

This research applies time series forecasting methods to the Philippines to predict future trends in food production and consumption. We find that the Philippines is facing a number of challenges to food security, including a growing population, a high level of poverty, dependence on imported food, and vulnerability to natural disasters [12][13][14]. However, we also find that there are a number of opportunities to improve food security in the Philippines, including investing in agricultural infrastructure and technology, promoting the development of local food systems, and implementing programs to support small-scale farmers and rural communities [15][16][17][18].

2. Data Preprocessing

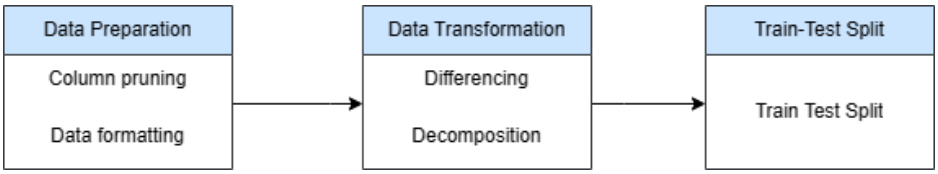

Figure 1. Data Processing.

Data preprocessing plays a crucial role in the success of forecasting models by ensuring the quality, relevance, and compatibility of the data. In this section, we present our data preprocessing approach using Google Colab and the Pycaret library for creating and comparing various forecasting models [20]. The data preprocessing pipeline [20][21] is a series of steps that are used to prepare data for analysis. The steps in the pipeline include:

Data preparation:

- Column pruning: This step involves removing columns from the data that are not relevant to the analysis.
- Data formatting: This step involves ensuring that the data is in a format that is suitable for analysis.
- Data transformation:
- Differencing: This step involves taking the difference between consecutive values in the data. This can be used to remove the trend from the data [22].
- Decomposition: This step involves breaking the data down into its trend, seasonality, and noise components. This can be useful for understanding the different factors that are affecting the data [23][7].
- Train-Test split:
- Train-test split: This step involves splitting the data into a training set and a test set. The training set is used to train the model, and the test set is used to evaluate the model's performance [24]. The training set should be at least 80% of the total data, and the test set should be at least 20% of the total data.

Data

Table 1. Sample Rice Production Dataset

	Domain	Area	Element	Item	Year	Unit	Value
0	Crops and livestock products	Philippines	Production	Rice	1961	tonnes	3910100.0
1	Crops and livestock products	Philippines	Production	Rice	1962	tonnes	3966980.0
2	Crops and livestock products	Philippines	Production	Rice	1963	tonnes	3842860.0
3	Crops and livestock products	Philippines	Production	Rice	1964	tonnes	3992400.0
4	Crops and livestock products	Philippines	Production	Rice	1965	tonnes	4072636.0

The dataset used in this paper is sourced from the Food and Agriculture Organization of the United Nations (FAO) [25] and the Philippine Statistics Authority (PSA), covering the period from 1961 to 2024. It primarily focuses on the top 15 agricultural crops in the Philippines, as identified by the Philippine Statistics Authority in 1996 [25]. These crops include rice, corn, coconut, sugarcane, banana, pineapple, coffee, cacao, cassava, sweet potato, peanut, onion, garlic, eggplant, and cabbage. The dataset comprises 7 features and 60 instances, all of which are univariate annual crop production datasets.

Data Preparation

The In the data preparation phase, column pruning was performed to remove unnecessary or unwanted columns. We utilized the pandas library to drop the columns "Domain," "Area," "Element," "Unit," and "Item," retaining only the "Year" and "Value" columns. This step was executed to streamline the dataset and focus on the relevant variables.

Additionally, data formatting was applied to convert the "Year" column to the datetime datatype. This conversion enables better handling and manipulation of datetime variables. The pandas function pd.to_datetime() was used to transform the "Year" column to the desired datetime format, specified as '%Y'.

Data Transformation

In the data transformation phase, we utilized the Pycaret library to automate the process of differencing and decomposition, as well as to generate visualizations.

- Differencing: This technique involves taking differences between consecutive observations to achieve stationarity [26].
- Decomposition: This process separates the time series into its constituent components, including trend, seasonality, and residual [7].

Train-Test Split

For the train-test split, we employed a commonly used split ratio of 80% for the training set and 20% for the test set [27][28][29][30]. By applying an 80-20 train-test split ratio [31], we divided our dataset into an 80% training set and a 20% test set, allowing us to train our model on a sizable amount of the data and assess its performance on an additional, unknown dataset.

3. Methodology

Data Gathering Procedure

The dataset is provided by two different organizations, the United Nations Food and Agriculture Organization Corporate Statistical Database (UN FAOSTAT). The UN FAOSTAT database provides free access to food and

agriculture data for over 245 countries and territories and covers all FAO regional groupings from 1961 to the most recent year available 2024 [32]. The Philippine Statistics Authority OPENSTAT (PSA OPENSTAT). The PSA OPENSTAT database is presented in the statistics generated and compiled by the PSA at the national and sub-national levels. These are limited to summarized and/or aggregated data which are organized into three (3) major domains, namely, (1) Demographic and Social Statistics, (2) Economic Statistics, and (3) Environment and Multi-domain Statistics.

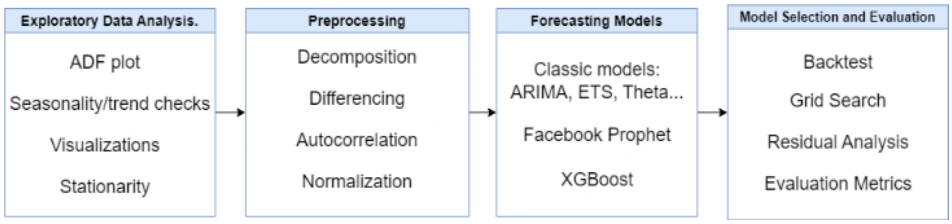

Figure 2. Time Series Forecasting process flow

Time Series Forecasting process flow

Exploratory Data Analysis

Exploratory Data Analysis (EDA), which entails a number of crucial processes, is the first step in the time series forecasting process [33][34][35][36]. First, stationarity is evaluated using an ADF plot [37]. To test whether a time series is steady, ADF charts are utilized. EDA also involves tests for trends and seasonality in the data. This stage is crucial for carrying out stationarity tests and comprehending the underlying patterns and properties of the time series. The remaining steps of the forecasting process are guided by these EDA-derived insights.

I. Preprocessing

Next in the forecasting process is Data Preprocessing, which encompasses several important techniques such as decomposition, differencing, and normalization [38]. Time series decomposition has been employed to separate a time series into its constituent components: trend, seasonality, and residual [24][27].

II. Forecasting models

Moving forward in the time series forecasting process, PyCaret was employed to train and compare the most suitable forecasting models for the given available crops. While PyCaret is predominantly intended for general machine learning, it can be adapted to handle time series forecasting by incorporating relevant methodologies Huber, AdaBoost, and Extra Trees to name a few [39][40].

III. Model Selection and Evaluation

In the model selection and evaluation process, PyCaret utilizes various techniques [20] to enhance the effectiveness of this stage. It employs backtesting for the train and validation split, grid search for hyperparameter tuning, and leverages available evaluation metrics to improve the performance of the models. During the train and validation split, PyCaret applies backtesting, which involves dividing the available data into training and validation sets [41]. This approach allows for the evaluation of the models on unseen data, providing insights into their generalization capabilities and performance on new observations. For hyperparameter tuning, PyCaret employs grid search. Grid search systematically explores a predefined grid of hyperparameter values to find the combination that yields the best performance. By optimizing the model's hyperparameters, PyCaret fine-tunes the models, enhancing their ability to accurately forecast future values.

In the final stage of the process, PyCaret leverages the available evaluation metrics to assess the performance of the forecasting models. It offers a variety of metrics, such as mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE)[41][42], and others. By carefully selecting the appropriate evaluation metric

based on the specific requirements of the forecasting task, PyCaret enables effective comparison and evaluation of the models [43].

From the datasets obtained from the FAO Statistics and PSA, this study employed a data gathering procedure to collect crop production values from 1961 to 2024, crop utilization per capita from 1990 to 2024, and the population dataset of the Philippines. These datasets are essential for estimating crop consumption.

The researchers utilized two forecasting models, namely ARIMA and Prophet, which can handle both seasonal and non-seasonal data. Auto ARIMA was employed to determine the optimal values for the parameters ARIMA The mathematical expressions for ARIMA are as follows:

Autoregression (AR): This model shows a changing variable that regresses on its own lagged, or prior, values [44].

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \quad (1)$$

Integrated (I): This represents the differencing of raw observations to make the time series stationary [44].

$$y_t = \Delta^d x_t = (1 - L)^d x_t \quad (2)$$

Moving average (MA): This model incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations [44].

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (3)$$

To select the best model between ARIMA and Prophet, the researchers evaluated the Mean Absolute Percentage Error (MAPE), a commonly used metric for forecast accuracy. If the MAPE exceeds 10% (0.1) [45], the researchers fine-tuned the auto ARIMA parameters such as max_p, max_q, and max_order values. The use of MAPE as an evaluation metric is advantageous when comparing the performance of a model on train and test datasets since it provides a relative measure of the error in percentage terms, making it suitable for comparing different datasets with varying scales and magnitudes [46].

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \quad (4)$$

The forecasted values were saved in a CSV format and transferred to a spreadsheet for further analysis of the expected deficit or surplus production of crops.

For determining the forecasted values of production, per capita consumption, and population using ARIMA, the following steps were followed:

1) Load the dataset.
2) Preprocess the data.
3) Split the data into 80% for training and 20% for testing.
4) Plot the data to identify any seasonality.
5) Check if the data is stationary by visualization or using statistical tests like the Augmented Dickey Fuller Test (ADF). If the data is non-stationary, perform differencing to make it stationary and determine the order of differencing.
6) Use auto ARIMA to find the best-fit model, adjusting the default parameters if necessary.
7) Create and fit the auto ARIMA model.

8) Evaluate the forecast accuracy using MAPE. If MAPE exceeds 10%, adjust the auto ARIMA parameters for a more optimized search.

9) Save the forecasted values in a CSV format.

10) Similarly, for determining the forecasted values using Prophet, the following steps were taken:

11) Load the dataset.

12) Preprocess the data.

13) Create and fit the Prophet model.

14) Save the forecasted values in a CSV format.

It is worth noting that Prophet requires less processing as it is designed to be a robust and easy-to-use forecasting tool that automatically detects and adjusts for various types of seasonality.

After obtaining the predicted values of production, per capita consumption, and population, the estimated consumption was calculated using the formula:

Estimated consumption = Crops per Capita * Number of Population

The estimated consumption values were tabulated in another CSV file, and the estimated deficit/surplus of the crops was calculated using the formula:

Estimated Crops = Forecasted Production - Estimated Consumption

In addition to ARIMA and Prophet, this study also utilized Pycaret [43], a Python-based low-code machine learning library, to streamline the machine learning workflow. Pycaret facilitated the application of various time series forecasting models to the data, allowing for efficient model comparison and evaluation. The models considered in this study included ARIMA, Prophet, XGBoost, Exponential Smoothing State Space Model (ETS), SARIMA, Gradient Boosting Regressor, Exponential Smoothing, Decision Tree Regressor, K Nearest Neighbors, Grand Means Forecaster, Theta Forecaster.

4. Results and Discussion

All of the top 15 agricultural crops show no observable seasonality based on their decomposition plots. This indicates that production levels do not follow a consistent recurring annual pattern. However, abrupt increases or decreases in production values occurring at specific time periods can still influence the forecasting models, since irregular fluctuations become part of the residual component of the time series. These sudden shifts may be attributed to climate disturbances, pest infestation, supply chain disruptions, or policy interventions, and they can significantly affect short-term predictive accuracy if not properly captured by the model.

Out of the 15 crops analyzed, only four; garlic, groundnuts or peanuts, onions, and sweet potatoes, showed a deficit in projected production when compared with estimated consumption levels. This implies that domestic supply may not be sufficient to meet national demand for these commodities in the forecasted period. It must be noted that Garlic, Eggplant, Cabbage have production forecasts but no FAO Food Balance Sheet per-capita supply data, so surplus/deficit cannot be computed. They appear in all forecast files but are excluded from the food security analysis.

Based on the predicted results of production versus demand, the model provides a data-driven basis for determining whether the country should engage in importation or exportation. For example, the forecast shows that onions may experience a production deficit of approximately 150,000 tonnes in 2024. This insight allows the government to prepare importation strategies ahead of time, mitigating potential price surges or supply shortages. Conversely, crops predicted to produce a surplus can be allocated for export, value-adding processes, or market stabilization activities. In this way, the forecasting model becomes a valuable tool for agricultural planning, economic decision-making, and food security management.

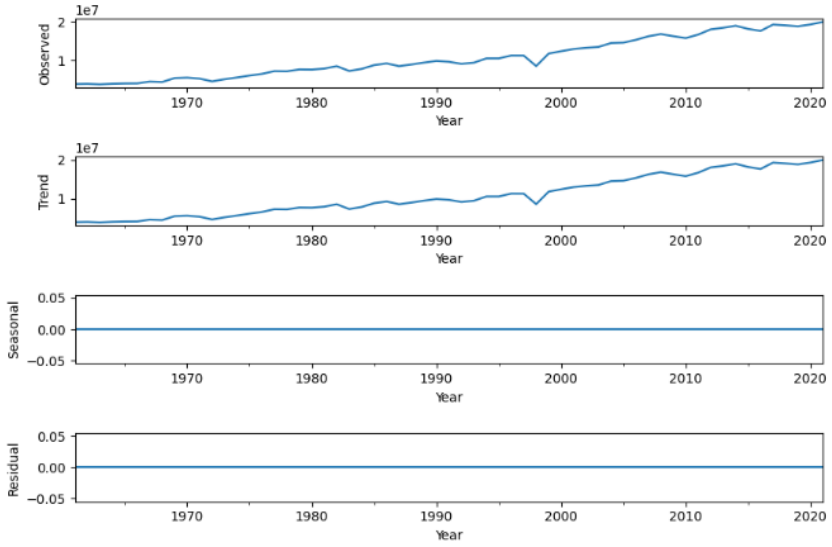

Figure 3. Sample Decomposition Plot of Rice Dataset

The decomposition plot helps in understanding the underlying patterns in the data by separating it into its individual components and providing insights into trends, seasonality, and residuals. From the specific sample in figure 3, the decomposition plot provides insights into the observed plot from the year 1961 to 2024. The plot shows an uptrend in the data, indicating an increasing pattern.

Differencing is a common technique used to remove trends and seasonality from a time series, making it more stationary and suitable for modeling. The difference plot in PyCaret helps visualize the transformed time series data after differencing and provides insights into the trends, autocorrelation, and partial autocorrelation patterns within the series. The difference plot displays the transformed actual values of the time series in an upward or downward trend.

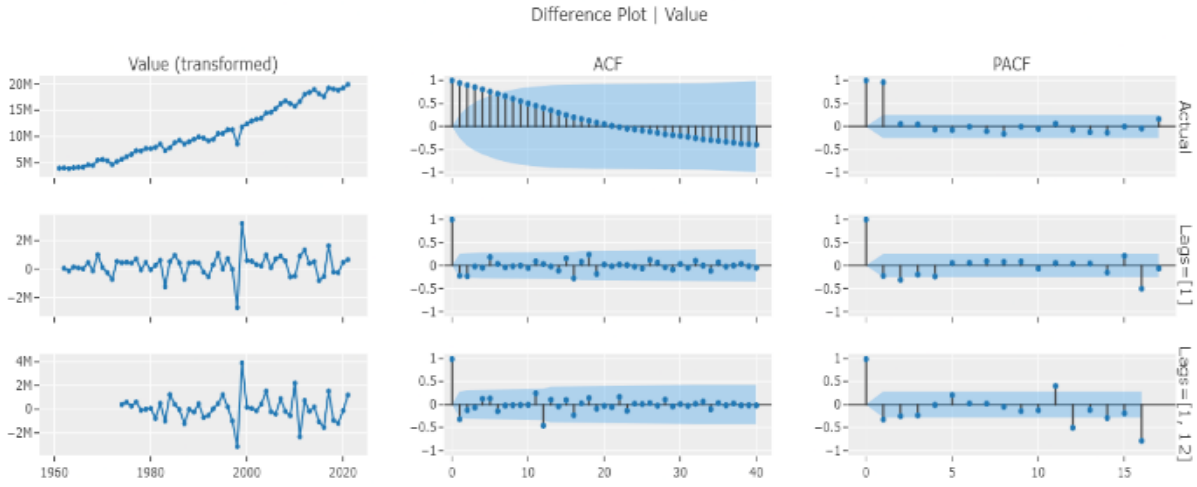

Figure 4. Sample Differencing Plot of Rice Dataset

As shown in Figure 5, the actual value exhibits both downward and upward trends. Among the forecasting models used, the Auto ARIMA model shows the closest resemblance to the actual value. It intersects with the actual value at three points, indicating a good fit and capturing the general trend pattern.

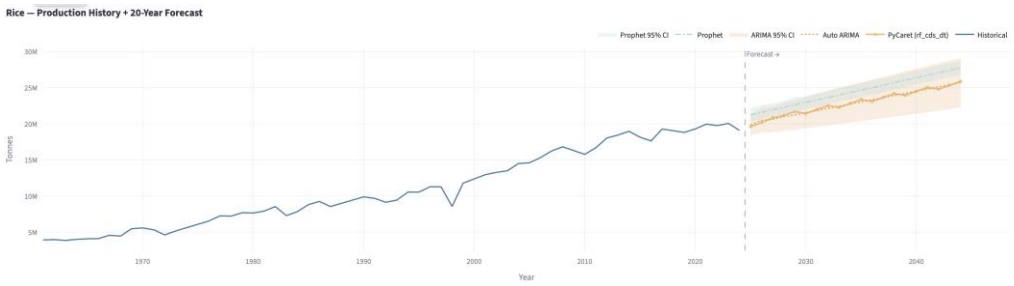

Figure 5. Sample Train Validation Test Split of Rice Production Dataset

As shown in Figure 6, the graphical representation of the actual value showcases a dynamic pattern characterized by a sequence of downtrend, uptrend, and subsequent downtrend. This variability suggests the influence of multiple underlying factors contributing to the observed fluctuations. Among the forecasting models evaluated, the Naive forecaster demonstrates the closest alignment with the actual value. This model captures the general trend and exhibits a flat or horizontal trend that closely follows the direction of the actual value. Its ability to maintain proximity to the actual value highlights its effectiveness in capturing the underlying patterns and providing reliable forecasts.

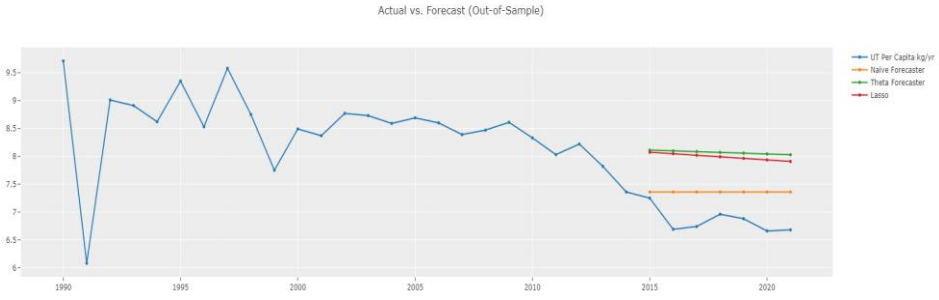

Figure 6. Sample Train Validation Test Split of Coconut Consumption Dataset

Figure 7. Actual vs 20 Years Forecast of Onion Production Dataset

In Figure 7, the plotted actual value of the onion production dataset against the 20-year forecasted horizon. The plot shows that the actual values exhibit a pattern of both sudden uptrend and downtrend, followed by a continuation of an uptrend. The Huber model's plot is near the actual value and shows a slow uptrend direction that is still within the range of the actual values. This indicates that the Huber model provides an acceptable forecast for the

onion production dataset. onion is in deficit in the near-term (2024) but is forecast to shift to surplus by a specific later year

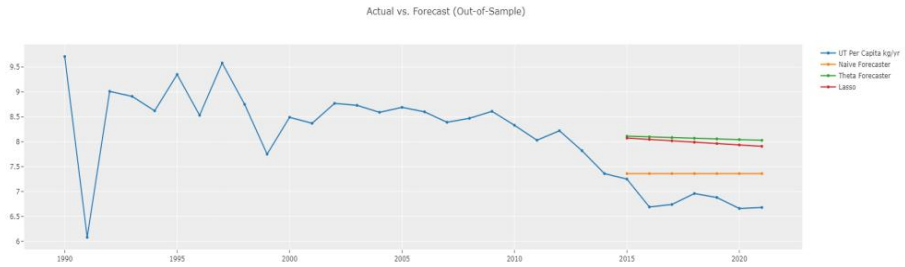

Figure 8. Actual vs 20 Years Forecast of Rice Production Dataset

In Figure 8, the plotted actual value of the rice production dataset against the 20-year forecasted horizon. The plot shows that the actual values exhibit a pattern of slow uptrend with some small downtrend, followed by a continuation of an uptrend. The Auto ARIMA model's plot is nearest to the actual value and merges with them, indicating that it provides a close and accurate forecast. The Auto ARIMA model also has the lowest and acceptable values for various evaluation metrics, including MASE (Mean Absolute Scaled Error) of 1.3772, RMSSE (Root Mean Squared Scaled Error) of 1.3223, MAE (Mean Absolute Error) of 590365.1130, RMSE (Root Mean Squared Error) of 745660.3059, MAPE (Mean Absolute Percentage Error) of 0.0543, and SMAPE (Symmetric Mean Absolute Percentage Error) of 0.0532.

	Model	MASE	RMSSE	MAE	RMSE	MAPE	SMAPE	R2	TT (Sec)
auto_arima	Auto ARIMA	1.3772	1.3223	590365.1130	745660.3059	0.0543	0.0532	0.5583	1.4933
ada_cds_dt	AdaBoost w/ Cond. Deseasonalize & Detrending	1.5294	1.5188	714757.3040	941275.3355	0.0597	0.0590	0.2855	0.3000
et_cds_dt	Extra Trees w/ Cond. Deseasonalize & Detrending	1.5400	1.5852	711904.2759	971297.1959	0.0597	0.0591	0.2462	0.6500

Figure 9. Actual vs 20 Years Forecast of Rice Production Dataset

As shown in Figure 9, The metrics demonstrate that the Auto ARIMA model performs well and provides accurate forecasts compared to the other models considered. The second-best suitable or fit model for the rice production dataset is AdaBoost. The third-best forecasting model for the rice production dataset is Extra Trees.

Figure 10. Top 10 models by median MAPE across all crops Darker = lower MAPE = better

In the Production Dataset, the dataset that achieved the lowest MAPE (highest accuracy) is Rice using the Auto ARIMA model, Additionally, the most frequently used and best model among the 15 datasets is the Auto ARIMA model.

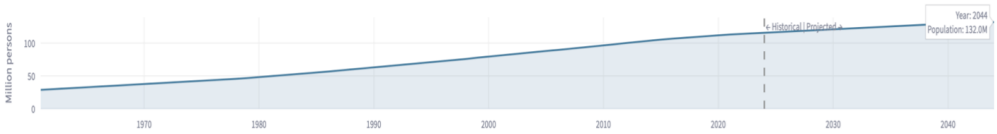

Figure 11. Philippine Population Projection for 2044

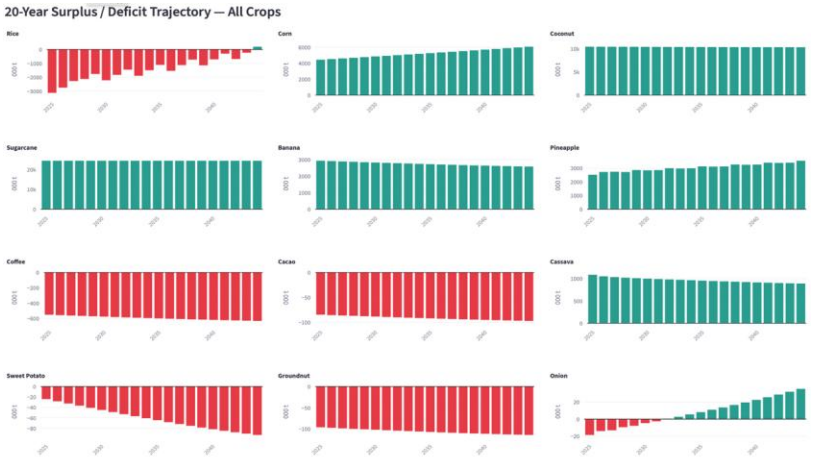

Figure 12. 20 Years Forecast of all crops for 2025-2044

The projected trajectories of crop surplus and deficit over the 2025–2044 period must be interpreted against the backdrop of sustained demographic expansion. According to the population projection shown in figure 11, the national population is expected to rise from approximately 114 million in 2025 to 132.0 million by 2044, representing a net addition of nearly 18 million persons, or an increase of roughly 16 percent, over the two-decade horizon. Although the annual population growth rate is projected to moderate over time, absolute population continues to climb steadily through the terminal year of the projection, indicating that aggregate food demand will continue to expand even as growth decelerates. This sustained demographic pressure provides critical context for interpreting the divergent commodity-level trends observed in Figure 12.

Against this backdrop of continuous population growth, the widening deficits observed for coffee, cacao, groundnut, and particularly sweet potato take on added significance, as these trajectories imply that production growth for these commodities is not only failing to close existing supply gaps but is doing so against a denominator of demand that is itself expanding by nearly a sixth over the same period. The deficit in sweet potato, which nearly quintupled in magnitude across the projection window, is therefore likely to be compounded rather than offset by population-driven demand growth, reinforcing the urgency of productivity interventions for this crop. Conversely, the narrowing deficits observed for rice and the transition of onion from deficit to surplus are particularly notable achievements when viewed alongside rising population, since these improvements in self-sufficiency are occurring despite, rather than in the absence of, an expanding consumer base. For commodities such as banana and cassava, which remain in surplus but show a gradually eroding margin, the concurrent rise in population suggests that per capita surplus is contracting even more sharply than the aggregate figures indicate, implying that these buffer stocks may be depleted sooner than the raw trajectories alone would suggest once population-adjusted, or per capita, demand is taken into account. Overall, the juxtaposition of crop-level production trends with population projections underscores that food security outcomes for the Philippines over the next two decades will depend not merely on whether production trajectories are improving

in absolute terms, but on whether they can outpace a population base projected to grow by approximately 18 million additional persons by 2044.

5. Conclusion

This initial study demonstrated the potential of machine learning-enhanced time series forecasting as a decision-support tool for agricultural planning and food security management in the Philippines. By utilizing over six decades of historical production, consumption, and demographic data (1961–2024) sourced from FAOSTAT and the Philippine Statistics Authority, and by benchmarking a diverse suite of forecasting algorithms, including Auto ARIMA, Prophet, Gradient Boosting Regressor, Theta Forecaster, and several tree-based and distance-based learners. The research identified Auto ARIMA as the most consistently accurate and frequently selected model across the fifteen major crop datasets evaluated, achieving the lowest error metrics among all candidates tested. This finding reinforces the continued relevance of classical statistical forecasting approaches, even in an era increasingly dominated by machine learning, particularly for univariate agricultural time series characterized by long-term trend behavior rather than strong seasonality.

The 20-year forecasts generated through this framework, when reconciled against projected per capita consumption and population growth, revealed that the majority of the fifteen crops examined are expected to maintain production levels sufficient to meet domestic demand through 2044. However, garlic, groundnut, onion, and sweet potato were identified as commodities at risk of sustained production deficits, a concern that is further compounded when viewed against the Philippines' projected population growth from approximately 114 million in 2025 to 132 million by 2044. This near 16-percent expansion in the national population underscores that self-sufficiency for these vulnerable crops cannot be assumed to improve passively; without targeted productivity interventions, rising demand alone is likely to widen these existing gaps rather than close them. Conversely, crops projected to sustain healthy surpluses provide clear opportunities for export expansion, value-adding industries, and price stabilization efforts, illustrating the dual utility of the proposed forecasting framework in flagging both areas of vulnerability and areas of opportunity within the national food system. Taken together, these findings affirm that data-driven, forecasting-based approaches offer meaningful value to policymakers, local government units, and agricultural stakeholders tasked with anticipating shortages, planning importation and exportation strategies, and allocating resources for farm support programs. Future research should consider expanding the modeling framework to incorporate exogenous variables such as climate indices, farm input costs, and international commodity prices, which may further improve forecast granularity and responsiveness to short-term shocks. Moreover, extending the analysis to sub-national or regional levels would allow for more localized food security planning, given the significant variation in agricultural productivity and population growth rates across Philippine regions. Lastly, this research contributes a replicable, scalable forecasting methodology that supports the broader national goal of achieving stable, resilient, and affordable food security for all Filipinos in the decades ahead.

6. Acknowledgement

The author would like to acknowledge Project ISTACK, where he served as project leader, as the inspiration in drafting the content of this study, particularly in underscoring the importance of identifying and measuring the impact of predicting the most in-demand crops in the Philippines over the next two decades. The author also acknowledges the support of his university in facilitating the conduct of the experiments within the Digital Transformation Laboratory at STEERHUB.

Authors' Profiles

Rowell M. Hernandez is a researcher and academic leader from Batangas State University – The National Engineering University, specializing in data-driven solutions, emerging technologies, and applied research for national development. His work spans artificial intelligence, digital twins, food security forecasting, and technology-enabled policy systems. He is notably involved in the development of the Bioethanol Industry Online Monitoring System (BIOMS), a national digital platform that supports the implementation of the Philippine Biofuels Act by enabling real-time data reporting, compliance tracking, and analytics for government agencies and industry partners. ISTACK and BARACO. Dr. Hernandez has produced research on forecasting frameworks, intelligent monitoring systems, applying statistical and machine-learning methodologies to social, agricultural, and industrial problems. He also contributes to academic publications, curriculum development in AI and data science, and the mentoring of student research.

REFERENCES

1. Philippines: Population Expected to Reach 100 Million Filipinos in 14 Years | Philippine Statistics Authority. (n.d.). https://psa.gov.ph/content/philippines-population-expected-reach-100-million-filipinos-14-years
2. Hunger and food insecurity. Hunger | FAO |Food and Agriculture Organization of the United Nations. https://www.fao.org/hunger/en/
3. Nguyen, T. (2022). Why the Philippines Is So Vulnerable to Food Inflation. Carnegie Endowment for International Peace. https://carnegieendowment.org/2022/07/13/why-philippines-is-so-vulnerable-to-food-inflation-pub-87467
4. Briones, R., Antonio, E., Habito, C., Porio, E., & Songco, D. (2017). Food Security and Nutrition in the Philippines. Food Security and Nutrition Strategic Review in the Philippines. https://docs.wfp.org/api/documents/WFP-0000015508/download/
5. Dancker, J. (2023, February 6). A Brief Introduction to Time Series Forecasting Using Statistical Methods. Medium. https://towardsdatascience.com/a-brief-introduction-to-time-series-forecasting-using-statistical-methods-d4ec849658c3
6. Ibrahim, M. (2022). How to Use XGBoost and LGBM for Time Series Forecasting? 365 Data Science. https://365datascience.com/tutorials/python-tutorials/xgboost-lgbm/
7. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts. https://otexts.com/fpp2/
8. Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 785–794. https://doi.org/10.1145/2939672.2939785
9. Munarsih, E., & Saluza, I. (2020). Comparison of exponential smoothing method and autoregressive integrated moving average (ARIMA) method in predicting dengue fever cases in the city of Palembang. Journal of Physics, 1521(3), 032100. https://doi.org/10.1088/1742-6596/1521/3/032100
10. Time Series Analysis - XGBoost for Univariate Time Series - Michael Fuchs Python. (2020, November 10). https://michael-fuchs-python.netlify.app/2020/11/10/time-series-analysis-xgboost-for-univariate-time-series/
11. Khankary, S. (2022, January 29). Univariate Time Series Forecasting using RNN(LSTM) - MLearning.ai - Medium. Medium. https://medium.com/mlearning-ai/univariate-time-series-forecasting-using-rnn-lstm-32702bd5cf4
12. Klingenbrunn, N. (2022, April 10). Transformer Implementation for TimeSeries Forecasting | by Natasha Klingenbrunn | MLearning.ai | MLearning.ai. Medium. https://medium.com/mlearning-ai/transformer-implementation-for-time-series-forecasting-a9db2db5c820
13. Project to address food security issues in PH via improvement of vegetable value chain begins | Philippines | Countries & Regions | JICA. (n.d.). https://www.jica.go.jp/philippine/english/office/topics/news/220121_02.html
14. The FAO Legislative Advisory Group-Philippines (FLAG-PH) initiative | Department of Economic and Social Affairs. (n.d.). https://sdgs.un.org/partnerships/fao-legislative-advisory-group-philippines-flag-ph-initiative
15. Challenging the Change: The Growing Impact of Climate Change on PH Food Security and Livelihoods. (n.d.). Philippines. https://philippines.un.org/en/158099-challenging-change-growing-impact-climate-change-ph-food-security-and-livelihoods
16. World Bank Group. (2020, September 8). PHILIPPINES: Vibrant Agriculture is Key to Faster Recovery and Poverty Reduction. World Bank. https://www.worldbank.org/en/news/press-release/2020/09/09/philippines-vibrant-agriculture-is-key-to-faster-recovery-and-poverty-reduction
17. Sanchez, F. (2015). Challenges faced by Philippine agriculture and UPLB's [University of the Philippines Los Baños] strategic response towards sustainable development and internalization. AGRIS: International Information System for the Agricultural Science and Technology. https://agris.fao.org/agris-search/search.do?recordID=PH2016000388
18. SDGs - Philippines. (2022, September 8). Goal 2 – Zero Hunger - End hunger, achieve food security and improved nutrition and promote sustainable agriculture SDGs - Philippines. SDGs - Philippines -. https://sdg.neda.gov.ph/goal-2/

19. Pawlak, K., & Kołodziejczak, M. (2020). The Role of Agriculture in Ensuring Food Security in Developing Countries: Considerations in the Context of the Problem of Sustainable Food Production. Sustainability, 12(13), 5488. https://doi.org/10.3390/su12135488
20. Yıldırım, S. (2021, December 15). Machine Learning Made Easy by PyCaret - Towards Data Science. Medium. https://towardsdatascience.com/machine-learning-made-easy-by-pycaret-5be22394b1ac
21. Data Preprocessing - Docs. (n.d.). https://pycaret.gitbook.io/docs/get-started/preprocessing
22. Sundararaman, B. (2021, December 15). PyCaret Machine Learning Data Preprocessing ML | Towards Data Science. Medium. https://towardsdatascience.com/pycaret-the-machine-learning-omnibus-part-2-87c7d0756f2b
23. Cerqueira, V. (2023, March 15). Understanding Time Series Trend - Towards Data Science. Medium. https://towardsdatascience.com/understanding-time-series-trend-addfd9d7764e
24. Brownlee, J. (2020). How to Decompose Time Series Data into Trend and Seasonality. MachineLearningMastery.com. https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/
25. FAOSTAT. (n.d.). https://www.fao.org/faostat/en/#home
26. Philippine Statistics Authority. (1996). Table 18. Philippine agricultural crops by area, production and value in 1994. Fao. https://www.fao.org/3/W6928E/w6928e0t.htm
27. Howell, E. (2022, December 28). Time Series Decomposition - Towards Data Science. Medium. https://towardsdatascience.com/time-series-decomposition-8f39432f78f9
28. Gholamy, A. (n.d.). Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. ScholarWorks@UTEP. https://scholarworks.utep.edu/cs_techrep/1209/
29. Detective, D. (2021, December 13). The 80/20 Split Intuition and an Alternative Split Method. Medium. https://towardsdatascience.com/finally-why-we-use-an-80-20-split-for-training-and-test-data-plus-an-alternative-method-oh-yes-edc77e96295d
30. Sachin. (2023, January 24). What Is The 80/20 Rule In Machine Learning? TechMediaToday. https://www.techmediatoday.com/what-is-the-80-20-rule-in-machine-learning/
31. Verma, J. (2020, October 13). Split Training and Testing Data Sets in Python - AskPython. AskPython. https://www.askpython.com/python/examples/split-data-training-and-testing-set
32. Statistics. (n.d.). Food and Agriculture Organization of the United Nations. https://www.fao.org/statistics/en/#:~:text=chevron_right-,FAOSTAT,the%20most%20recent%20year%20available.
33. Biswal, A. (2023). What is Exploratory Data Analysis? Steps and Market Analysis. Simplilearn.com. https://www.simplilearn.com/tutorials/data-analytics-tutorial/exploratory-data-analysis
34. What is Exploratory Data Analysis? | IBM. (n.d.). https://www.ibm.com/topics/exploratory-data-analysis
35. Exploratory Data Analysis (EDA): Types, Tools, Process. (n.d.). https://www.knowledgehut.com/blog/data-science/eda-data-science
36. Mahadevan, M. (2022). Step-by-Step Exploratory Data Analysis (EDA) using Python. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2022/07/step-by-step-exploratory-data-analysis-eda-using-python/
37. G, V. K. (2023). Statistical Tests to Check Stationarity in Time Series. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/06/statistical-tests-to-check-stationarity-in-time-series-part-1/
38. Subasi, A. (2020). Data preprocessing. In Practical Machine Learning for Data Analysis Using Python (pp. 27–89). Academic Press. https://doi.org/10.1016/b978-0-12-821379-7.00002-3
39. Ali, M. (2022, January 7). Time Series Forecasting with PyCaret Regression Module. Medium. https://towardsdatascience.com/time-series-forecasting-with-pycaret-regression-module-237b703a0c63
40. Zvorničanin, E. (2023). Time Series Projects: Tools, Packages, and Libraries That Can Help. neptune.ai. https://neptune.ai/blog/time-series-tools-packages-libraries
41. Malik, N., & Agarwal, B. (2022). Time Series Nowcasting of India's GDP with Machine Learning. In 2022 International Conference on Artificial Intelligence of Things (ICAIoT), Istanbul, Turkey, pp. 1–6. https://doi.org/10.1109/icaiot57170.2022.10121883
42. Mathangpeddi. (2021, December 15). Leverage the power of PyCaret - Towards Data Science. Medium. https://towardsdatascience.com/leverage-the-power-of-pycaret-d5c3da3adb9b
43. PyCaret. (n.d.). PyCaret. Retrieved from https://pycaret.org/
44. Sun, M., Kim, G., Lei, K., & Wang, Y. (2022). Evaluation of Technology for the Analysis and Forecasting of Precipitation Using Cyclostationary EOF and Regression Method. Atmosphere, 13(3), 500. https://doi.org/10.3390/atmos13030500
45. Learn Mean Absolute Percentage Error | Vexpower. (n.d.). https://www.vexpower.com/brief/mean-absolute-percentage-error
46. Trueblood, R. P., & Lovett, J. N. (2001). *Data mining and statistical analysis using SQL* (Vol. 1). Berkeley, CA: Apress.