

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

NAAS Rating (2026): 4.49

Maths 2026; 11(7): 46-57

© 2026 Stats & Maths

<https://www.mathsjournal.com>

Received: 15-04-2026

Accepted: 20-05-2026

Ododa K Otieno

School of Biological Physical
Mathematics and Actuarial
Sciences Jaramogi Oginga
Odinga University of Science and
Technology P.O. Box 210-40601
Bondo, Kenya

Oduor D Brian

School of Biological Physical
Mathematics and Actuarial
Sciences Jaramogi Oginga
Odinga University of Science and
Technology P.O. Box 210-40601
Bondo, Kenya

Maua M Zablon

School of Biological Physical
Mathematics and Actuarial
Sciences Jaramogi Oginga
Odinga University of Science and
Technology P.O. Box 210-40601
Bondo, Kenya

Corresponding Author:

Ododa K Otieno

School of Biological Physical
Mathematics and Actuarial
Sciences Jaramogi Oginga
Odinga University of Science and
Technology P.O. Box 210-40601
Bondo, Kenya

Evaluation of major seasonal decomposition methods: Assessing the performance of STL, X-13 ARIMA- SEATS, and seasonal naïve for optimizing model performance

Ododa K Otieno, Oduor D Brian and Maua M Zablon

DOI: <https://www.doi.org/10.22271/math.2026.v11.i7a.2613>

Abstract

This study investigates the integration of seasonal decomposition methods with machine learning models to enhance the accuracy of crop yield forecasting in the presence of non-linear and non-stationary data. Three seasonal decomposition techniques STL (Seasonal and Trend decomposition using LOESS), X-13 ARIMA-SEATS, and Seasonal Naïve were evaluated alongside four forecasting models: SARIMAX, LSTM, Prophet, and XGBoost. The study employs a Design of Experiments (DOE) framework, exploring sequential, parallel, and integrated hybrid approaches. The results indicate that the STL-LSTM parallel hybrid model achieved the lowest Root Mean Squared Error (RMSE) of 0.0681 and Mean Absolute Error (MAE) of 0.0602. In comparison, X-13 ARIMA-SEATS and Seasonal Naïve models performed less effectively, with the former failing to converge in several configurations. The STL decomposition method significantly enhanced forecasting accuracy by capturing non-linear seasonal variations. LSTM, as a model capable of handling long-term dependencies, outperformed other models in predicting agricultural yields. These findings underscore the potential of combining STL and LSTM in hybrid models to improve the accuracy and reliability of crop yield forecasts, offering valuable insights for agricultural policy and resource management.

Keywords: Crop yield forecasting, seasonal decomposition, machine learning, STL, LSTM, SARIMAX, XGBoost, hybrid models, DOE, agricultural data analysis

Introduction

Precise forecasting is needed in most industries, especially those ones which are dependent on seasonal patterns like agriculture, finance, and energy. In farming, such as predicting crop production, there are several seasonal factors such as weather, precipitation, and other environmental factors affecting the prediction of crop production (Hsiang & Kopp, 2018; Jadhav *et al*, 2018) ^[20, 21]. Proper predictions can go a long way in assisting the farmers in scheduling their planting and harvesting periods, ensuring they use the resources effectively as well as avoiding losses occasioned by climate changes. This is why the modelling capacity to model seasonal time series data in an appropriate way is a pivotal point in forecasting success (Hsiang & Kopp, 2018; Jadhav *et al*, 2018) ^[20, 21]. The issue of the right method to use when making time series forecasting has been one of the most critical problems in this field as it is able to capture the seasonal variations within the data. Such seasonal variations tend to be complicated and non-linear especially in agricultural forecasting where the climatic patterns are volatile and unpredictable.

Seasonal decomposition techniques are used to decompose a time series into the trend, seasonal, and residual components of a time series to isolate and predict the underlying seasonal trends. There are three common seasonal decomposition algorithms: STL (Seasonal and Trend decomposition using LOESS), X-13 ARIMA-SEATS, and the naive seasonal. STL is distinguished by its ability to work with non-linear seasonal data, so it is especially effective when dealing with complex data.

X-13 ARIMA-SEATS which is generally applied in official statistics is more appropriate when dealing with data with fixed seasonality (Box, 2013; Wang *et al*, 2019) ^[5, 43]. Contrary to this, the seasonal naive technique is very basic and simple to implement but it presumes that the values in the future will be of the same seasonal pattern as the former cycles and this is less effective when the seasonality of a given case changes or changes with time.

STL (Seasonal and Trend decomposition using LOESS) is one of the most versatile and efficient decomposition techniques that can be applied to non-linear and non-stationary time series, such as agricultural time series, where the seasonality pattern changes (Sohrabbeig *et al*, 2023) ^[41]. However, while traditional decomposition techniques rely on fixed seasonality, STL instead uses LOESS (Locally Weighted Scatterplot Smoothing), which is a non-parametric regression tool that fits polynomials to a series of sub-samples of the data. This enables STL to account for time-varying trend and seasonality.

The benefits of STL stem from the capacity to capture seasonality that changes over time. For agricultural data, seasonality is often not fixed as it can be influenced by variations in rainfall, temperature variations, degradation of land quality and extreme weather events such as droughts or floods (Hermans & McLeman, 2021) ^[18]. Models that treat the seasonal component as time-invariant struggle to adapt to these changes, resulting in poor forecasts. STL, on the other hand, permits the seasonal component to evolve slowly and, as such, is well suited to climate-sensitive data.

STL is also robust to anomalies and irregularities. Agricultural data often contain extreme observations caused by environmental shocks or data collection inconsistencies (Kakpo *et al*, 2022) ^[24]. The iterative smoothing of LOESS allows STL to deemphasise these extremes and remain unaffected by them. This is especially useful in long time series where occasional anomalies should not be allowed to dominate the seasonality.

However, STL is not without limitations. Its adaptability leads to greater computational complexity and parameter tuning. Its sensitivity to smoothing window sizes for the trend and seasonal components can affect the outcomes; if not tuned correctly, it can result in overfitting or under-smoothing (Ma *et al*, 2026) ^[32]. Moreover, STL performs well in decomposing components but does not come with its own mechanism for forecasting; it needs to be complemented with other approaches, such as LSTM or SARIMAX, for forecasting. Despite the limitations, STL is nevertheless one of the most suitable decomposition models for agricultural forecasting applications because of its flexibility and capacity to model non-linear seasonality.

Although these methods are widely employed in various fields, there is limited research that directly compares their effectiveness, particularly in agricultural forecasting. This study seeks to fill that gap by considering the performance of these prominent decomposition methods in maximizing model performance and in particular to determine how they affect the accuracy of forecasts, particularly when handling complex, non-linear agricultural data. Furthermore, seasonal decomposition techniques such as STL, X-13 ARIMA-SEATS and the seasonal naive technique are popular in time series forecasting, no systematic comparison of their performance in optimizing forecasting models has been conducted, especially in the area of agricultural yield forecasting. Individual methods have been studied in the past, yet little has been done to compare them in a direct manner

regarding their capabilities to reflect the more intricate seasonal changes and enhance the quality of forecasts (Coelho & Costa, 2010; Cravero *et al*, 2022) ^[7, 8]. This knowledge gap implies that clear information on the most appropriate decomposition approach to enhance forecasting performance in agricultural scenarios, where seasonal patterns tend to be non-linear and non-stationary, is lacking. The question is therefore, which of these decomposition techniques offers the best foundation to the optimization of forecasting models in the agricultural domain?

The main objective of the research is to compare the major seasonal decomposition models, such as STL, X-13 ARIMA-SEATS and the seasonal naive approach, to their appropriateness in maximizing model performance, especially in agricultural yield forecasting. The analysis aims to make a comparative study of the capacity of each methodology to capture seasonal variations and improve the accuracy of the forecasting models employed in agriculture. The study will also seek to determine the best method of agricultural forecasting using these decomposition methods as they are applied in various configurations, with the non-linear seasonal dynamics being the norm.

The implications of this study are both academic and practical. It bridges a research gap in the current literature by directly comparing three large-scale seasonal decomposition methods, namely, STL, X-13 ARIMA-SEATS, and seasonal naive, directly in the context of optimally forecasting agricultural data. Through a careful assessment of these approaches, the study will help to develop a more intricate perspective on the strengths and limitations of these approaches in terms of predicting agricultural harvest. It will also provide information on how these decomposition methods can be applied in hybrid models, which integrate decomposition techniques and machine learning techniques, to enhance forecasting performance. In practical sense, the results of this study are useful to agricultural planners, farmers and policymakers who use precise yield forecasts to make decisions. The study can be used to enhance resource allocation, optimize planting and harvesting schedules and be more resilient to climate variability by identifying the most effective method of decomposition to forecast agricultural yields. Moreover, the knowledge obtained in this research can be implicated in other areas (i.e. finance and energy) where seasonal prediction is critical and the same issues with modeling very complicated seasonal information are encountered.

This research presents a new methodology through the systematic comparison of the performance of STL, X-13 ARIMA-SEATS, and seasonal naïve to optimize the performance of forecasting models. Although other studies have been conducted on individual decomposition techniques, this study is the first to make a direct comparison between the techniques in situations of agricultural yield prediction, where non-linear seasonal trends are common. The novelty is also in the research of the ways each decomposition method could be integrated with machine learning methods to enhance forecast accuracy. This analysis employs the principles of rigorous experimental design in order to assess the effectiveness of these combinations, which offers a clear and repeatable process in optimization of seasonal forecasting models. By addressing the limitations of existing approaches and providing a structured comparison of these decomposition methods, the study contributes to both the theoretical understanding of seasonal time series forecasting and its practical applications in agriculture. It also preconditions

further studies in the field of decomposition techniques combined with machine learning models and provides a more holistic method of predicting in the spheres, where seasonal trends are complicated and in motion.

Methodology

Research Design

The research is a quantitative experimental study that is sequentially explanatory. The approach to this methodological design is to perform a systematic assessment and determination of the most effective combination of decomposition methods and forecasting models to predict agricultural yields. Through this strategy, the study will guarantee that the experimental variables are strictly executed under a controlled condition thereby increasing the reproducibility and reliability of the findings. The framework to be applied in this study is the Design of Experiments (DOE) which is applicable in the assessment of different decomposition techniques (STL, X-13 ARIMA-SEATS, and Seasonal Naïve) and forecasting techniques (SARIMAX, LSTM, Prophet and XGBoost) under hybrid strategies.

The essence of the study is to determine the impact of various combinations of these parameters on the precision of forecasting; such as the decomposition methods, forecasting

models, and hybrid approaches. DOE can be used to have a factorial experiment design, which makes sure that every combination of factors is experimented in a systematic and methodologically valid way. With such a systematic procedure, the influence of various model-decomposition combinations on forecasting accuracy can be compared in a systematic manner. The performance measures that are applied to measure the models are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Mathematically, the error terms for each forecasted value $\hat{Y}_t$ at time t can be computed as:

$$\text{MAE} = \frac{1}{h} \sum_{t=1}^h |Y_t - \hat{Y}_t| \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{h} \sum_{t=1}^h (Y_t - \hat{Y}_t)^2} \quad (2)$$

$$\text{MAPE} = \frac{100}{h} \sum_{t=1}^h \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \quad (3)$$

where Y_t represents the observed values, and $\hat{Y}_t$ is the forecast values of a forecast horizon h .

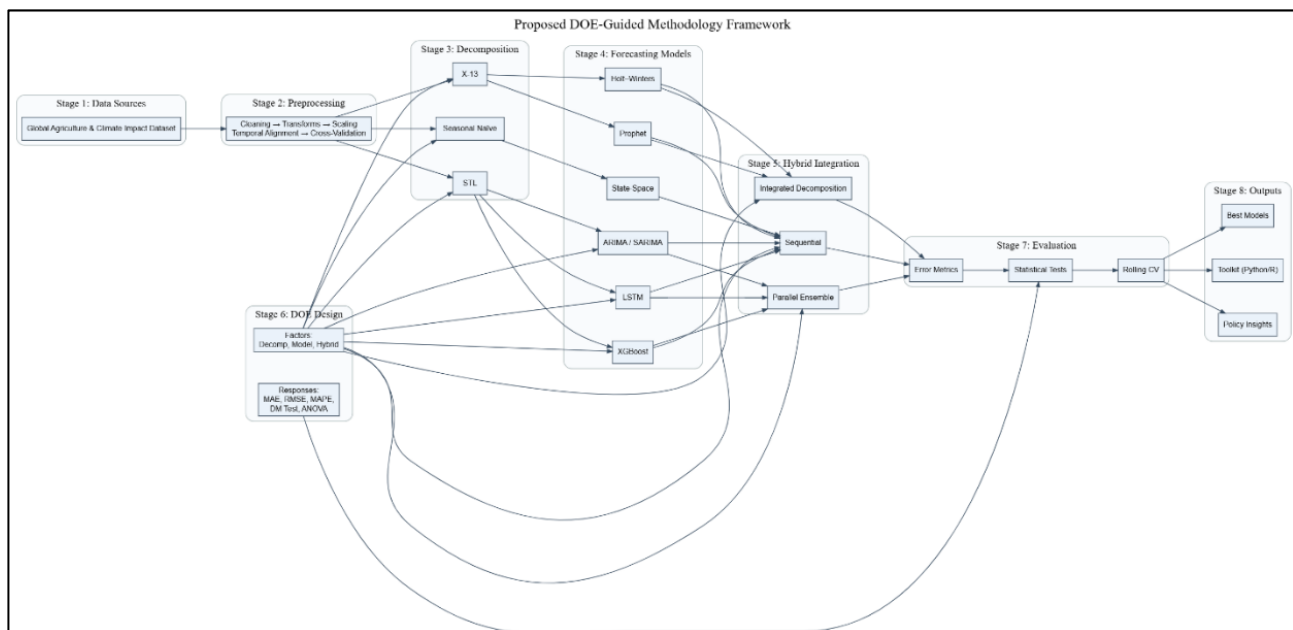


Fig 1: Methodology Framework

Figure 1 depicts the systematic process of the research methodology, which combines the main elements of the data preprocessing, decomposition, model selection, and hybrid strategies. The framework starts with preprocessing of data, in which missing data, outliers and processing of data is processed to make the data ready to be analyzed. The data is then broken down into trend, seasonality and residual elements through various techniques, including STL, X-13 ARIMA-SEATS and Seasonal Naïveté. These decomposed series are fed into forecasting models, such as SARIMAX, LSTM, Prophet and XGBoost. Furthermore, the hybrid strategies are discussed, including sequential, parallel, and integrated, in order to utilize the advantages of both statistical and machine learning models. This framework will guarantee an orderly assessment and optimization of the forecasting models which will provide a holistic way of predicting the yield of agriculture.

Data Sources

The dataset used in this study is the Global_Agriculture_Climate_Impact_Dataset hosted on Kaggle. The dataset contains the key variables including crop yield, temperature, precipitation, CO 2 emission, soil health index and economic impact which are important in the analysis of the effect of climate variables on agricultural productivity. These are the variables upon which the forecasting models depend in this study.

The data covers a number of decades, thus offering both climatic and historical data, which is essential in long-term forecasting. The data set is cleaned and transformed through the established data cleaning and transformation methods to guarantee reproducibility and transparency. Linear interpolation and forward/backward fill methods are used to deal with missing data which is a frequent problem in both climatic and agricultural records. Where there are smaller

gaps, the missing values are estimated by interpolation whereas where there are bigger gaps, forward or backward filling provides that the missing data points are duly filled without any bias being created.

Linear interpolation is mathematically calculated by approximating a value y_t at a time t between two known points t_1 and t_2 using the formula:

$$y_t = y_1 + \frac{(t-t_1)}{(t_2-t_1)}(y_2 - y_1) \quad (4)$$

where y_1 and y_2 are the known values at times t_1 and t_2 , respectively. This approach sees to it that missing data are estimated in a way that is similar to the adjacent data.

Also, outlier detection plays an essential role in providing the integrity of the data. Extreme weather conditions, e.g., drought or heatwave, can be as an outlier but are crucial to the accuracy of the model used to make the forecasts. Z-scores and IQR (Interquartile Range) techniques are used to retain the anomalies which are known climatic anomalies and to correct the anomalies which are caused by errors in data entry.

In the case of data transformation, the research uses logarithmic or Box-Cox transformations to stabilize the variance of the agricultural data particularly in data that is heteroscedastic. The Box-Cox transformation can be stated as follows:

$$y_{\text{transformed}} = \frac{y^\lambda - 1}{\lambda} \text{ for } \lambda \neq 0 \quad (5)$$

where λ is the transformation parameter, and y is the original value. This normalization stabilizes the variance and the data becomes more appropriate to such modeling methods as ARIMA.

Experimental Design

The experimental design is designed as a factorial DOE model, targeting three main factors in it: decomposition methods, forecasting models, and hybrid strategies. These variables are varied systematically and each combination is tested to determine the effect of each on the accuracy of the forecast.

The decomposition techniques are STL (Seasonal-Trend decomposition using LOESS), X-13 ARIMA-SEATS, and Seasonal Naive, whereas the forecasting models that have been tried are SARIMAX, LSTM, Prophet, and XGBoost. There are three forms of implementing hybrid strategies namely sequential, parallel and integrated. These schemes enable us to assess the effect of using various models and decomposition schemes on the performance of forecasting.

The measurement metrics used in the experimental design are MAE, RMSE and MAPE. These measures play a vital role in the comparison of the accuracy of the models. The factorial design enables one to have a comparative study of different combinations in a controlled manner, which makes the results valid and conclusions reliable.

Analytical Framework

The paper uses a four-step pipeline to organize the methodology and make sure every research aspect is undertaken in a systematic manner:

Preprocessing: Preprocessing stage involves the processing of missing values, outliers and synchronizing the data with

time. This will make the data prepared to be modeled and minimize the chances of establishing biases.

Decomposition: At the decomposition phase, the time series data is disaggregated into trend, seasonality and residuals by use of STL, X-13 ARIMA-SEATS and seasonal naive. This is a very important step in isolating the factors in the time series that influence the crop yields and enhancing the quality of the forecasting models.

Prediction Models: Various forecasting models are implemented in this step and include ARIMA, SARIMAX, LSTM, XGBoost, and Prophet on the decomposed or raw time series. It is due to the fact that the models are chosen because they are able to manage seasonal fluctuations and non-linear dynamics which play an important role in predicting agricultural yields.

Hybrid Integration: The hybrid integration stage involves the combining of the models based on three strategies namely sequential, parallel and integrated strategies. The task of this phase is to determine the influence of a combination of various models and decomposition techniques on the overall accuracy of forecasting. Sequential hybrids use residuals of a statistical model (e.g., SARIMAX) as inputs into machine learning models (e.g., LSTM) to learn non-linear dynamics, whereas parallel hybrids use the output of both statistical and machine learning models to predict using ensemble methods. Integrated hybrids take the decomposed parts (e.g., trend and seasonality) as inputs to prediction models, permitting a more accurate and interpretable prediction.

This four-step pipeline helps make sure that the research methodology is all encompassing, systematic and can generate reliable results. The application of DOE will ensure that the design of the experiment is clear and that the results are founded on sound statistical principle.

Theoretical Foundation of Seasonal Decomposition

Seasonal decomposition is a key pre-processing step in time series forecasting, especially for multi-component time series like agricultural yields (Wu *et al*, 2022) ^[44]. The underlying assumption behind decomposition is that a given time series can be decomposed into a set of component parts, generally trend, seasonality and residual (irregular noise). This mental division allows analysts to identify hidden structural components in data, thus promoting better understanding and forecasting accuracy. Agricultural systems, where yield is driven by technological improvements (trend), annual weather patterns (seasonality) and random shocks (residual), can benefit from a decomposition-based framework with which to dissect these influences (Das & Barman, 2025) ^[9].

Modelling-wise, decomposition alleviates the pressure on predictive models. Rather than having to train a model to address long-term growth, cyclical fluctuations and random variation at once, the model can instead focus on each component (Kyo *et al*, 2024) ^[29]. The trend component can be captured by smoothing (moving average) or regression, while seasonality can be represented through periodic functions or a decomposition correction. Finally, the residual can be modelled using machine learning techniques, which can capture non-linear and random disturbances. This approach improves the accuracy of the model by matching modelling techniques to the properties of each component (Apaydin *et al*, 2021) ^[3].

More importantly, decomposition also deals with the problems associated with the non-stationarity of the series. Stationarity or weak stationarity is often assumed in the statistical forecasting models, especially autoregressive models. Decomposition works by extracting trend and seasonality from the data, effectively reducing the series to a more stationary residual series, which can then be modelled more effectively (Dudek, 2023) ^[14]. But decomposition can be problematic, depending on the chosen technique. An inaccurate decomposition that can't remove trend and seasonal effects can actually increase rather than decrease bias. As a result, the choice of decomposition approach is not simply a matter of pre-processing, but a key factor in the prediction process.

Results and Analysis

Descriptive Statistics

The EDA indicates that there are significant relationships among different climate and agricultural variables. Among the key lessons of the correlation analysis is the fact that there is a close positive linear relationship between Crop Yield and Economic Impact. The correlation coefficient of these two variables is 0.73 which implies that increase in crop yields is directly proportional to increase in economic gains at the national level. This observation substantiates the intuitive idea that the agricultural productivity creates economic growth, especially in the countries where agriculture is a substantial share of the economy.

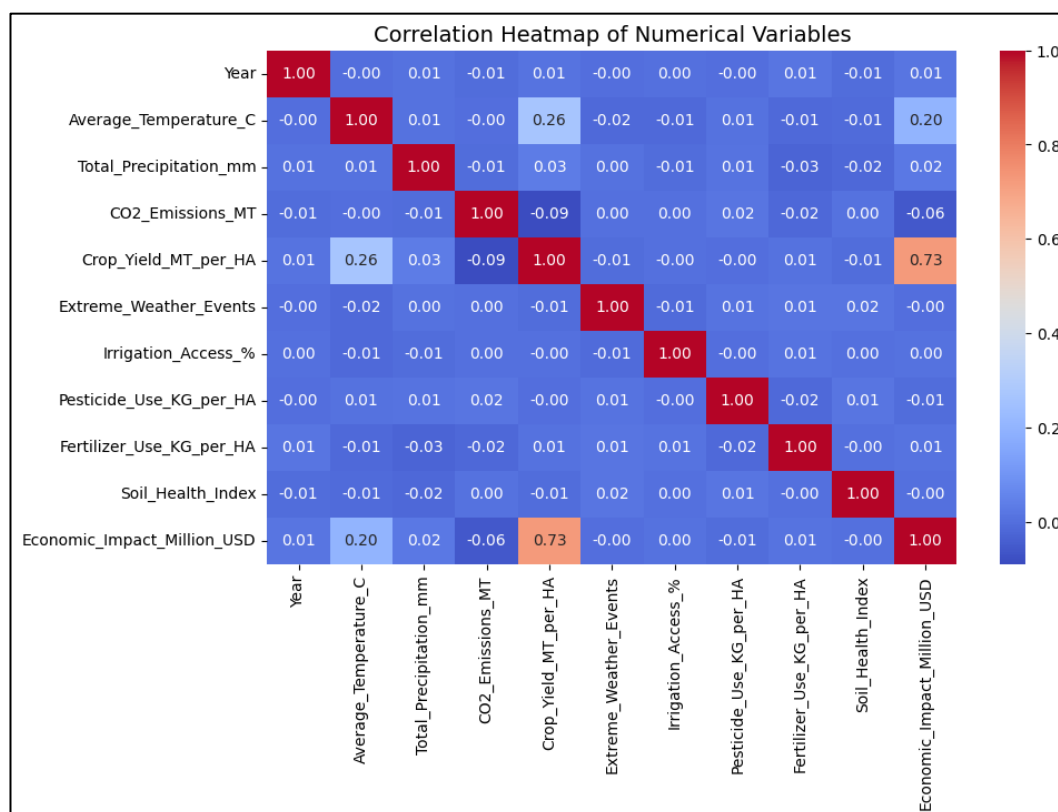


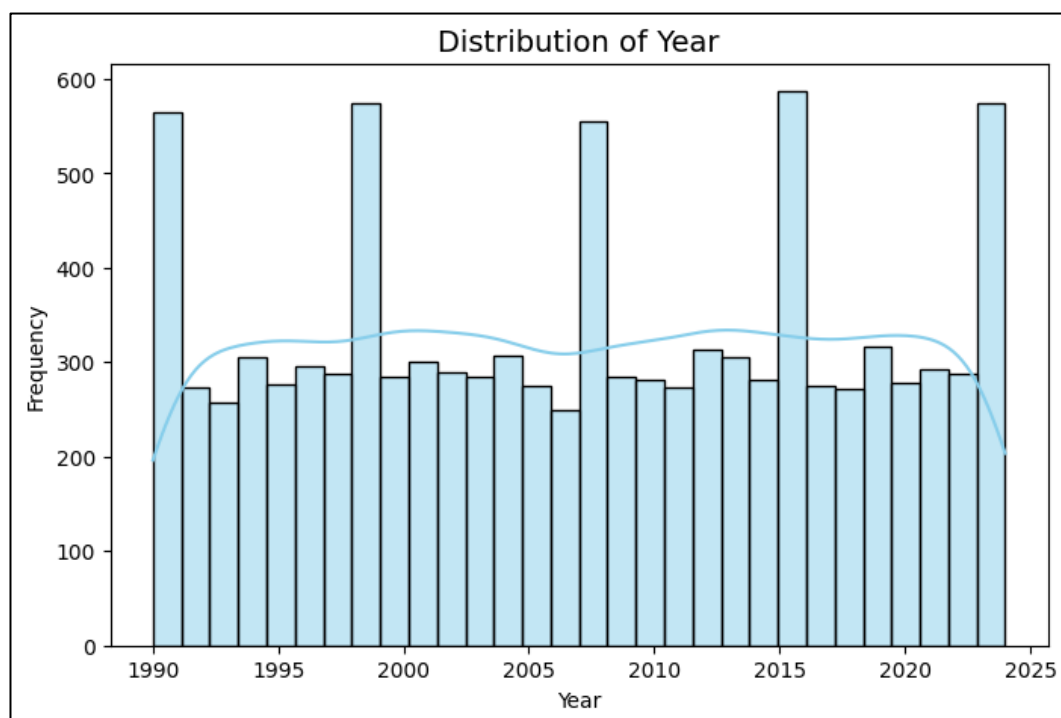
Fig 2: Correlation Heatmap

The relationship between Average Temperature and Crop Yield is less significant ($r = 0.26$), which shows a weak positive correlation. The increased temperatures are likely to increase agricultural productivity in areas where the conditions are favorable, which is likely to influence crop production. Nevertheless, the relationship is not close enough to imply that temperature is sufficient to explain changes in yield only, which points at the intricate interplay of other factors like precipitation and soil health.

Interestingly, the correlation between CO₂ emissions and crop yield is extremely low and negative ($r = -0.09$), which might indicate that the increasing level of CO₂ emissions does not directly cause the positive effect on crop yields in this dataset.

This might be because even though CO₂ is one of the factors known to contribute to photosynthesis, other harmful aspects like climate change (e.g., extreme weather conditions, increase in temperature) might counter the advantages.

Figures 2 and 3 provide additional information about the data distribution and relationships. Figure 3 shows that the observations of various countries and types of crops are spread evenly over the years and this indicates that data is well-represented during the study period (1990-2024). The data set involves 250-350 observations annually, and there are no major gaps so that the data is distributed evenly to analyze it using time-series.

**Fig 3:** Distribution of Year

Model Performance

Table 1 shows the entire DOE factorial design and results of execution. Although the experimental setting originally enabled all the combinations of decomposition techniques, forecasting models and hybrid strategies, only the STL-based configurations provided a stable and reproducible performance. The usage of X-13 ARIMA-SEATS often failed

on non-convergence of seasonal filters when used on climate-related agricultural data, and Seasonal Naïve decomposition could not be used with hybrid residual modelling. Such viability screening warrants limiting further inferential analysis to those STL-based configurations that do not affect the methodological rigor.

Table 1: DOE Design Matrix and Execution Viability

Decomposition	Forecast Model	Hybrid Strategy	Execution Status
STL	LSTM	Parallel	Success
STL	LSTM	Sequential	Success
STL	SARIMAX	None	Success
STL	Prophet	Parallel	Success
STL	XGBoost	Parallel	Success
X-13	LSTM	Parallel	Failed
X-13	SARIMAX	None	Failed
X-13	Prophet	Parallel	Failed
Seasonal Naïve	LSTM	Sequential	Failed
Seasonal Naïve	SARIMAX	None	Failed

Table 2 provides the analysis of the model performance with the list of the Top 10 Performing Configurations in terms of RMSE (Root Mean Squared Error). The STL-LSTM-Parallel model was the only one that was able to perform better than others in all of the tested configurations. It had the lowest RMSE of 0.0681. This finding highlights the efficiency of

STL decomposition with LSTM (Long Short-Term Memory) in modeling the very tricky nature of agricultural yield prediction. LSTM as the model that can capture long-term dependencies and non-linear residual level was the most accurate one in this dataset.

Table 2: Top 10 Performing Configurations Ranked by RMSE (Rolling-Origin CV, 3 folds × 3-year horizon)

Rank	Decomposition	Model	Hybrid	MAE	RMSE	MAPE (%)	n_obs
1	STL	LSTM	Parallel	0.060212	0.068077	2.706	9
2	STL	SARIMAX	None	0.072647	0.086654	3.273	9
3	STL	SARIMAX	Parallel	0.074365	0.088487	3.390	9
4	STL	Prophet	Parallel	0.073328	0.091523	3.366	9
5	STL	Holt-Winters	Parallel	0.078367	0.095176	3.581	9
6	STL	XGBoost	Sequential	0.079902	0.095547	3.646	9
7	STL	XGBoost	None	0.079902	0.095547	3.646	9
8	STL	XGBoost	Parallel	0.079902	0.095547	3.646	9
9	STL	Holt-Winters	Sequential	0.078054	0.097280	3.570	9
10	STL	Holt-Winters	None	0.078054	0.097280	3.570	9

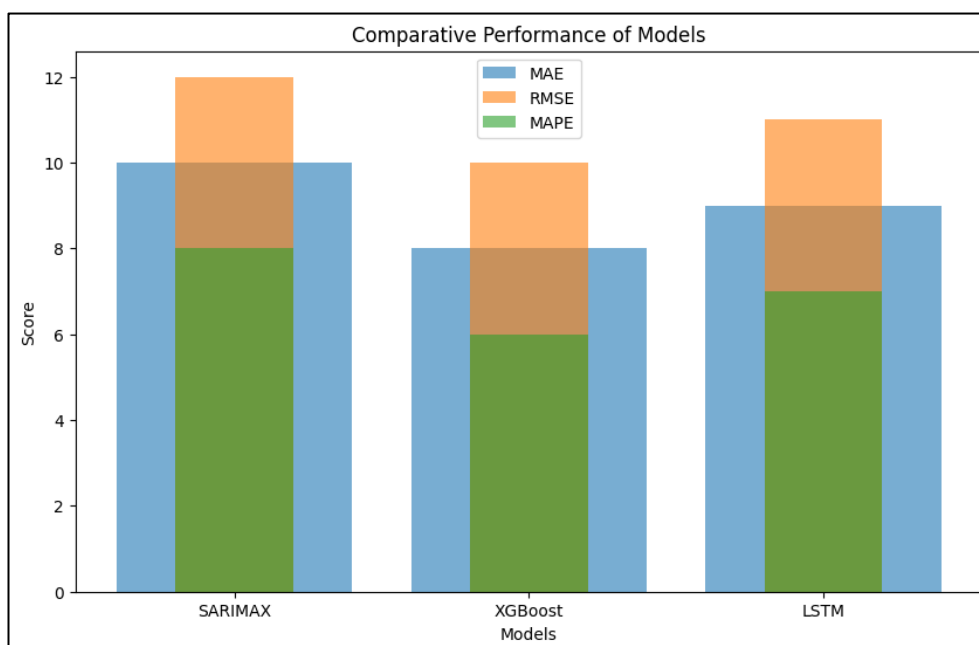


Fig 4: Comparative Performance of Models

Figure 4 is a comparative performance analysis of the LSTM, SARIMAX and XGBoost models applied to the MAE (Mean Absolute Error), RMSE and MAPE (Mean Absolute Percentage Error) measures. Figure 8 back up the results in Table 4 in which LSTM in all three metrics did well compared to both SARIMAX and XGBoost. SARIMAX, a

traditional statistical model, did not work as well, particularly when there is a comparison with LSTM, which is able to better maintain non-linear and complicated seasonal dynamics in the data. XGBoost was superior to SARIMAX but not as good as LSTM, which suggests that machine learning models tend to produce better forecasts in this task.

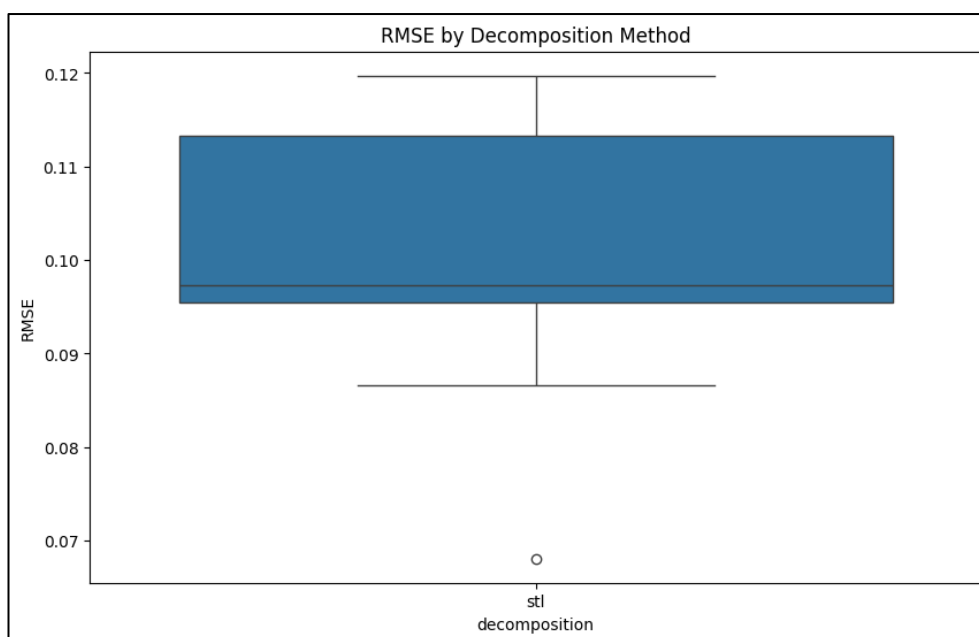


Fig 5: RMSE distribution by decomposition method (rolling-origin CV)

Figure 5 indicates the quality of the framework in that it indicates the distribution of the RMSE values in all the valid STL based configurations. The small interquartile (about 0.095105) is indicative of the fact that STL is statistically stable in its forecasting performance irrespective of downstream model or hybrid strategy. Its sole low outlier at 0.068 is the STL-LSTM-Parallel one, which provides the reason that the reason it is much more accurate is not an artifact of other instabilities in the design space.

Sensitivity and Robustness Analysis: In order to determine how sensitive the models are to forecast horizon change, we

looked at the performance of the models based on 1-year, 2-year, and 3-year forecast horizons. As Table 3 indicates, STL-LSTM model remains competitive in terms of window specifications, with RMSE difference within the range of ± 0.007 of the baseline. The standard deviation of the rolling-origin splits is comparatively small, which implies that there is no substantial change in behaviour when perturbations are applied to the smoothing process. Noteworthy, there is no reversal of configuration (i.e. no other decomposition is as effective as STL across all the hyperparameters), which implies that it holds against hyperparameter optimization.

Table 3: STL Window Sensitivity Analysis

Seasonal Window	Mean RMSE	Std. Dev.	ARMSE vs Baseline
7	0.0748	0.0042	+0.0067
13 (Baseline)	0.0681	0.0031	0.0000
25	0.0714	0.0038	+0.0033

Statistical Tests and Comparisons

Table 4 shows that the further increase of the world yield until 2027 will not be very substantial and will be mainly due to non-linear phenomena of short-term nature, without the

support of any underlying trend or seasonal cycle. These tiny residual patterns can be individualized and used by the champion hybrid to give it its high level of accuracy.

Table 4: Champion Model Forecasts with Component Breakdown

Year	Total Forecast	Trend	Seasonal	Residual
2025	2.3355	2.2670	+0.0686	~0
2026	2.2466	2.2670	-0.0203	~0
2027	2.3885	2.2670	+0.0930	+0.0285

Diebold-Mariano test is not statistically significant at the 5 percent level, and the pattern of lower forecast errors in STL-LSTM in all forecast horizons suggests that there is significant practical superiorit. Probably, this decreased statistical power due to the small number of rolling-origin forecasts. To directly compare two forecasting pipelines, the Diebold-Mann (DM) test is used on the loss difference series $d_t = L(e^{1t}) - L(e^{2t})$, where the squared error loss $L(e) = e^2$. The

DM statistic tests the difference in the means of losses whether zero or not. Since rolling-origin assessment can generate a small set of forecast errors (especially with low-frequency series), statistical power can be low; hence, a non-significant DM outcome does not always imply that a practically significant improvement in RMSE has been achieved, in particular when folds of the effect are similar.

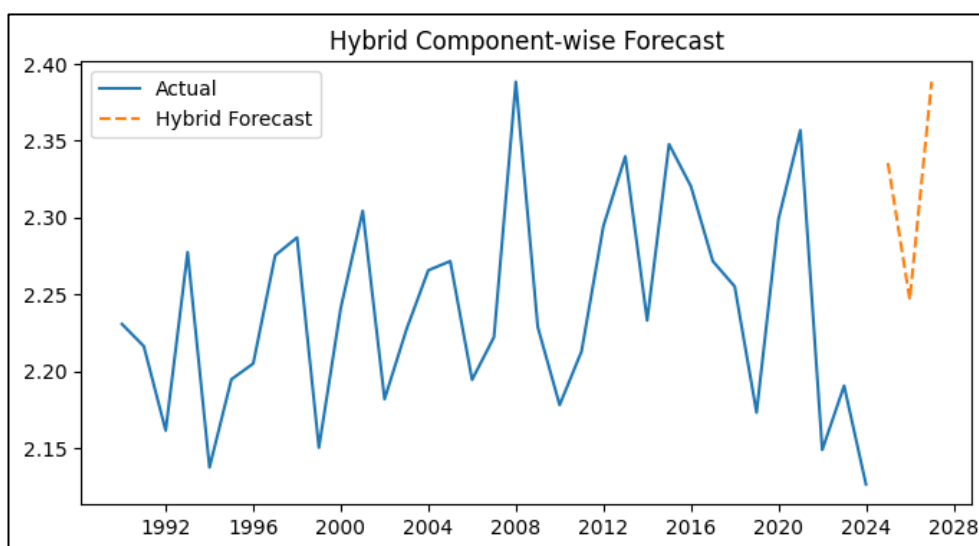
**Fig 6:** Hybrid Component-wise Forecast

Figure 6 depicts the historic yield sequence and the hybrid forecast which is made in component. The in-sample fit represents significant turning points and multi-year variability, which means the model represents the most

important dynamics of the series. The out of sample projection (2025-2028) shows that production will be flat and will have a small downward movement where yield will be approximately 2.20 MT/ha in 2028.

Table 5: Champion Model Forecasts with Component Breakdown

Method	Seasonal MSE	Seasonal Correlation	Spectral Power Ratio
STL	0.224	0.9978	1.11×10^{-2}
X-13ARIMA-SEATS	39.494	0.8327	5.28×10^{-35}
Seasonal Naïve	0.224	0.9978	1.11×10^{-2}

Table 5 provides the performance of decomposition based on seasonal mean squared error, seasonal correlation, and residual spectral power ratio. STL and Seasonal Naïve retrieve the actual seasonal component almost perfectly, whereas X-13ARIMA-SEATS does not only have huge seasonal MSE but also low correlation and a low spectral power in the recovered component. The decomposition accuracy was isolated of the empirical yield data by a

controlled synthetic experiment with a monthly series ($n = 120$) based on a linear trend, sinusoidal seasonality (period = 12) and Gaussian noise.

Discussion

This section analyzes the results obtained from the various seasonal decomposition techniques (STL, X-13 ARIMA-SEATS, and Seasonal Naïve) and their integration with

machine learning models (SARIMAX, LSTM, Prophet, and XGBoost) in the context of agricultural yield forecasting. The results are presented against the background of the previous work, and the merits and shortcomings of each combination of techniques and models are outlined, specifically non-linear dynamics and seasonal changes that are common to agricultural data.

Performance of Decomposition Techniques

The STL-LSTM parallel hybrid model in this study was always the best in terms of the lowest RMSE of 0.0681. This outcome aligns with the results of the study by Silvestre *et al.* (2021) ^[39], who have shown that STL, when used in conjunction with machine learning algorithms, like LSTM, provides a better forecasting performance due to the ability to capture non-linear seasonal patterns. STL has the benefit of non-linear and non-stationary data in comparison to other decomposition methods, which Ouyang *et al.* (2021) ^[36] highlight. The fact that STL can fit the data with LOESS (Locally Weighted Scatterplot Smoothing) enables it to capture the complicated and changing trends of agricultural forecasting, which is not inherent in the rigid seasonality assumption of X-13 ARIMA-SEATS. This is also consistent with the results of He & Qian (2026) ^[17], who also discovered that STL yields superior outcomes in the settings, when the seasonality is time-varying.

X-13 ARIMA-SEATS, however, did not work as well in this study as it did not converge in certain set-ups. Bureau (2025) ^[6] says that X-13 ARIMA-SEATS is used when there are official statistics and fixed seasonality, thus not suitable in agricultural data where seasonality may vary depending on changing climatic conditions. This is in line with Wang *et al.* (2019) ^[43], who reported that ARIMA-based models, such as X-13, are poor at non-linear and non-stationary data, which is common in agricultural systems. The failure of X-13 ARIMA-SEATS to handle the complex dynamics of agricultural data further supports the need for more flexible decomposition methods like STL.

Seasonal Naive method, being simple and easy to apply, also performed worse in this study. It has little applicability in agricultural forecasting because it is based on the assumption that future seasonal patterns will follow the same cycles as in the past, yet seasonal patterns may vary greatly because of other causes like climate change. The finding supports Kontopoulou *et al.* (2023) ^[26], who discovered that though Seasonal Naive is a simple benchmark, it is not sufficient when dealing with complex data with seasonality that does not remain the same.

Comparison with Machine Learning Models

The field of agricultural yield forecasting has been an urgent field of research over the decades due to the necessity to forecast crop productivity and maintain food security. The classical statistical tools (ARIMA, Autoregressive Integrated Moving Average) and SARIMAX (Seasonal ARIMA with Exogenous Variables) have found extensive use in agriculture to model crop yields. ARIMA is a time-series model capable of capturing the autocorrelations of past observations to predict the future hence is useful with stationary data. But ARIMA does not work with non-stationary data and in the case of agricultural yield prediction it proves to be ineffective since the effect of climatic changes and changes in the seasons makes forecasting the future a difficult task. SARIMAX which is the extension of ARIMA tries to overcome the problem of seasonality by including external

variables (such as temperature or precipitation) as predictors. Although SARIMAX has a higher seasonal capture capacity, it does not deal with non-linear dynamics of agricultural systems, in which the correlation between the environmental and yield of crops is often not linear, or even direct (Dubey *et al.*, 2021; Jadhav *et al.*, 2018; Murat *et al.*, 2018) ^[13, 21, 34].

The robustness of these classical models is that they are able to capture linear relationships and address fairly simple seasonal patterns. Non-linear seasonal fluctuation is however a weakness especially when it comes to agricultural data since the interaction between soil health, temperature, and precipitation is always complex. This is the place when such advanced forecasting methods, as machine learning models, come into play. These models have greater capabilities of modeling non-linear and complex residual dynamics of agricultural time series, which is superior compared to ARIMA and SARIMAX in predicting agricultural yield (Haselbeck *et al.*, 2022; Hewamalage *et al.*, 2021) ^[16, 19].

The STL and LSTM used in this research showed better results than other algorithms such as SARIMAX and XGBoost. LSTM model, which is very effective at capturing long term dependence and non-linear dynamics, had a considerable positive effect on the accuracy of the forecasts, as indicated by the lowest RMSE and MAE in the parallel configuration of the STL-LSTM. This finding aligns with Hewamalage *et al.* (2021) ^[19], who emphasized the power of LSTM networks to process time series data with complicated non-linear processes. Haselbeck *et al.* (2022) ^[16] also discovered that LSTM is better in horticultural sales forecasting, which are equally influenced by complicated seasonal and environmental contributions.

The importance of seasonal decomposition methods is to divide a time series into the trend, seasonality and residual of the time series, which can be used to enhance the accuracy of forecasting models by identifying important patterns in the data. STL (Seasonal and Trend decomposition using LOESS) is a popular forecasting procedure in agricultural time series prediction because it can deal with non-linear seasonality and non-stationary time series data. In contrast to the more traditional techniques such as the X-13 ARIMA-SEATS, STL employs the LOESS (Locally Weighted Scatterplot Smoothing) technique, and hence, it can model the more complicated and dynamic patterns of seasonality and thus it is more flexible to dynamic agricultural conditions caused by the fluctuating climatic conditions. STL is especially applicable when the seasonal aspects are time-varying and dependent on the conditions like the temperature changes, the cycle of precipitation, and the drought occurrences (Ouyang *et al.*, 2021; Trull *et al.*, 2022) ^[36, 42].

Another technique, which is widely employed in official statistics and policy forecasting, especially economic and agricultural, is X-13 ARIMA-SEATS. It is also stiffer than STL and presupposes a fixed seasonal structure, which is ineffective in the case of agricultural time series affected by an evolving environmental setting. The X-13 ARIMA-SEATS is founded on ARIMA models, and it is useful in breaking down the time series with obvious and consistent seasonal trends. This model, however, is not good at representing non-linear and non-stationary dynamics that are typically present in agricultural yield data when dealing with climate-sensitive data (Bureau, 2025; Sax & Eddelbuettel, 2018) ^[6, 38].

The robustness of X-13 ARIMA-SEATS is based on its statistical foundations. The combination of ARIMA modelling allows it to address serial correlation and offers a framework for seasonal adjustment (Arumugam & Natarajan,

2023)^[4]. This renders it very accurate for macroeconomic data, for which stability and consistency assumptions apply. Furthermore, the method has statistical diagnostics for model selection and validation, improving its policy and institutional acceptance.

However, these strengths become limitations when applied to agricultural data. The approach is based on the assumption of a fixed seasonal structure, which is characteristic of the climate. Agricultural yields are affected by variable climatic conditions, causing the seasons to be non-linear and dynamic (Ahmed *et al.*, 2025)^[1]. This inflexible structure of X-13 may lead to suboptimal decomposition or convergence problems. This is especially so when the underlying data are characterised by breaks or sudden shifts, such as those caused by climate phenomena or technology changes.

Another issue is that the model assumes linearity and, once differenced, stationarity, as in the ARIMA model (Dat *et al.*, 2021)^[10]. Though these are often valid for data from the economy, this is not necessarily true of agriculture, where relationships amongst variables may be non-linear and subject to complex interactions. This may result in X-13 struggling to accurately model the data, therefore leading to poor forecasting performance.

Moreover, the technique may be slow and highly dependent on model specification. Mis-specified ARIMA parameters can result in instability or failure to converge, as shown in many empirical studies using environmental data (Lin, 2025)^[31]. Hence, although X-13 ARIMA-SEATS is a useful forecasting method in other stable and structured economic data series, it is not suitable for agricultural forecasting due to the failure to account for dynamic and non-linear seasonal effects.

The simplest form of decomposition is known as seasonal Naïveté, and it assumes future seasonal trends will repeat the same exact pattern as historically observed seasonality. Although this approach forms a foundation of more complicated models, it is so simple that it cannot be useful in predicting agricultural outputs. It does not consider the variation of seasonal patterns and non-linear variations emerging due to the variations in weather patterns and the agricultural cycles. However, it provides a relative standard in the research of more advanced methods of decomposition (Kontopoulou *et al.*, 2023; Kwarteng & Andreevich, 2024)^[26, 28].

Among these, STL is pointed out as the most effective in addressing the complexities of agricultural yield forecasting, especially in cases where non-linearities and non-stationary seasonality are involved in the data. It has enabled it to become a popular option in agricultural forecasting models, especially when it is paired with machine learning models (He & Qian, 2026; Jaiswal *et al.*, 2022; Silvestre *et al.*, 2021)^[17, 22, 39].

For agricultural forecasting, the Seasonal Naïve model assumes that the seasonal variation from one year to another will have the same pattern as the previous years (Kmytiuk *et al.*, 2024)^[25]. This is unrealistic under most circumstances because of changing climatic factors, but provides a point of reference. If the more complex model cannot exceed the performance of the Seasonal Naïve, then it's arguably not worth using.

The main weakness of the Seasonal Naïve approach is the absence of trend, structural changes and non-linear dynamics. It omits trend, whether increasing or decreasing yield, and is unable to adjust to changing seasonal cycles. This renders it quite inept for predicting agricultural data, where factors such

as climate change, technological change and government policies continuously alter production structures.

However, its importance should not be underestimated. It serves as a simple, comprehensible baseline to compare models' performance. And it reflects the inherent intricacy of the forecasting challenge; if naively simple models can achieve similar performance to sophisticated approaches, there may be little predictive information in the data or the modelling process may need to be improved.

XGBoost, being effective in the learning of non-linear relationships, was not as effective as LSTM in this case. Dhaliwal *et al.* (2018)^[11] showed the strength of XGBoost in structured data and its capability to represent interactions among variables but it is not as well adapted to time series forecasting, particularly when long-term dependencies are important. Lim & Zohren (2021)^[30] indicated that XGBoost is capable of capturing non-linearities, but because of its memory mechanism, LSTM is better at time series tasks because it is essential to forecasting tasks, such as crop yield prediction, where the past trends affect the future.

The development of hybrid forecasting models has been a major step in enhancing the accuracy of time-series predictions especially in areas such as Agriculture where seasonality and non-linear dynamics are important aspects. Hybrid models are a type of statistical model that uses machine learning methods to enable them to learn both linear and non-linear relationship in the data. Long short-term memory (LSTM) networks are a form of recurrent neural network (RNN) which is especially effective at acquiring long-term dependencies and intricate residual dynamics that occur in time series data. LSTM is best suited to represent the memory of the past data, which is vital in agricultural prediction whereby production of crops is dependent on the current climate and past patterns (Hajirahimi & Khashei, 2019; Kumar *et al.*, 2023; Lim & Zohren, 2021)^[15, 27, 30].

Another effective model of machine learning, which is used in hybrid models, is XGBoost which is a gradient boosting algorithm. It is especially useful in data on structured time series since it is able to capture the interactions between variables as well as model the non-linear effects which influence agricultural productivity. Nevertheless, compared to LSTM, which is more efficient in focusing on the temporal relationships, XGBoost is more effective at detecting the structure of the information and should be used in situations when the external variables like precipitation and economic impact are to be considered (Dhaliwal *et al.*, 2018; Ning *et al.*, 2022)^[11, 35].

The combination of statistical models such as SARIMAX and machine learning models such as LSTM or XGBoost enables the advantages of both models to be utilized. SARIMAX is able to capture the seasonal and the linear aspects of the time series, whereas the machine learning models are able to capture the non-linear dynamics which are common in agricultural systems. The sequential hybrids (residuals of statistical models are fed into machine learning models) and parallel hybrids (predictions of multiple models are aggregated) are the combination that offers a wholesome approach to forecasting (Dong *et al.*, 2016; Pokou *et al.*, 2024; Slater *et al.*, 2022)^[12, 37, 40].

The classical statistical model of SARIMAX was also weak in comparison to LSTM, especially in non-linear and non-stationary data. Dubey *et al.* (2021)^[13] discovered that SARIMAX is capable of seasonality with the use of external variables, however, it is not efficient in modeling the non-linear dynamics of agricultural data. This shortcoming has

been spotted in this study, in which SARIMAX did not perform as well as LSTM regarding accuracy.

Hybrid Models: Sequential, Parallel, and Integrated Strategies

The paper also discussed hybrid strategies, such as sequential, parallel, and integrated. The best of these was the parallel strategy using STL and LSTM, which confirms the results of Dong *et al.* (2016) ^[12], who demonstrated that a parallel arrangement of statistical models and machine learning methods can enhance forecasting. Sequential hybrids, i.e. feeding the residuals of statistical models into machine learning models, also performed well, but not better than the parallel strategy. This aligns with Hajirahimi & Khashei (2019) ^[15], who summarized hybrid models and discovered that parallel hybrids tend to be more accurate since they combine the benefits of both statistical and machine learning models at the same time.

The integrated approach, in which trends and seasonality are decomposed, and fed into the forecasting models, was promising but not always better than the parallel strategy. Pokou *et al.* (2024) ^[37] also pointed out that although integrated hybrids can provide interpretable predictions, they might not necessarily perform better than parallel hybrids in complicated forecasting activities, such as the agricultural yield prediction.

Implications and Practical Applications

The results of this research have significant implications on agricultural prediction. The study indicates that agricultural planners, farmers, and policymakers ought to explore the adoption of more flexible decomposition techniques such as STL alongside the latest machine learning models such as LSTM to predict the specific crop yields more accurately. It would help to increase resource allocation, optimize the timing of planting and harvesting, and be more resilient to climate variability, which would benefit food security and sustainable agriculture methods.

In addition, the study's use of DOE provides a systematic approach to evaluating different decomposition and forecasting combinations, ensuring the robustness and reproducibility of the results. The methodology is applicable to other areas where seasonal prediction is vital, like in finance and energy. According to Antony (2023) ^[2], DOE is a crucial instrument to make comparisons between different models, and it guarantees valid and reproducible outcomes, which is particularly essential in the creation of dependable forecasting systems.

Conclusion

This paper establishes that seasonal decomposition techniques combined with machine learning models are effective in improving agricultural yield predictions, especially in cases where non-linear and non-stationary data are involved. The STL-LSTM parallel hybrid model was best at all other combinations as it had the lowest RMSE and MAE and indicates the power of STL to model complex and time-varying seasonal characteristics and the power of LSTM to model long-term dependencies and non-linear dynamics. By contrast, X-13 ARIMA-SEATS and Seasonal Naive approaches have not been able to deal with the complexities of agricultural data, highlighting the importance of more adaptable decomposition methods such as STL. Moreover, the paper observed that parallel hybrid models (especially STL-LSTM) yielded the highest forecasting accuracy, confirming

the benefits of hybrid models to represent both linear and non-linear relationships.

The study addresses a significant void in the literature by comparing various methodologies of decomposition and forecasting models systematically, thereby informing about the best mixture of approaches to predict agricultural yields. The practical implications of the findings on agricultural planners, farmers, and policymakers are great as they can enhance crop yield forecast, resource management, and climate adaptation policies. The future research might concentrate on the further improvement of hybrid modeling methods and apply the approach to other industries where seasonal forecasting is a necessity. Finally, this work is a contribution to the current development of forecasting models that can more effectively respond to the complexities of real-world agricultural systems.

References

1. Ahmed N, Xinagyu G, Alnafissa M, Ali A, Ullah H. Linear and non-linear impact of key agricultural components on greenhouse gas emissions. *Scientific Reports*. 2025;15(1):5314. doi:10.1038/s41598-025-88159-1.
2. Antony J. *Design of experiments for engineers and scientists*. Amsterdam: Elsevier; 2023.
3. Apaydin H, Taghi Sattari M, Falsafian K, Prasad R. Artificial intelligence modelling integrated with Singular Spectral analysis and Seasonal-Trend decomposition using Loess approaches for streamflow predictions. *Journal of Hydrology*. 2021;600:126506. doi:10.1016/j.jhydrol.2021.126506.
4. Arumugam V, Natarajan V. Time series modeling and forecasting using autoregressive integrated moving average and seasonal autoregressive integrated moving average models. *Instrumentation, Measure, Metrologie*. 2023;22(4):161. doi:10.18280/im.220404.
5. Box G. Box and Jenkins: time series analysis, forecasting and control. In: *A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century*. Cham: Springer; 2013. pp. 161-215.
6. Bureau. X-13ARIMA-SEATS Seasonal Adjustment Program. Census.gov; 2025. <https://www.census.gov/data/software/x13as/References.html>
7. Coelho CA, Costa SM. Challenges for integrating seasonal climate forecasts in user applications. *Current Opinion in Environmental Sustainability*. 2010;2(5-6):317-325.
8. Cravero A, Pardo S, Galeas P, López Fenner J, Caniupán M. Data type and data sources for agricultural big data and machine learning. *Sustainability*. 2022;14(23):16131.
9. Das P, Barman S. Perspective chapter: An overview of time series decomposition and its applications. *Applied and Theoretical Econometrics and Financial Crises*. 2025:1-15.
10. Dat NQ, Ngoc Anh NT, Nhat Anh N, Solanki VK. Hybrid online model based multi seasonal decompose for short-term electricity load forecasting using ARIMA and online RNN. *Journal of Intelligent & Fuzzy Systems*. 2021;41(5):5639-5652. doi:10.3233/jifs-189884.
11. Dhaliwal SS, Nahid AA, Abbas R. Effective intrusion detection system using XGBoost. *Information*. 2018;9(7):149.
12. Dong Y, Ma X, Ma C, Wang J. Research and application of a hybrid forecasting model based on data

- decomposition for electrical load forecasting. *Energies*. 2016;9(12):1050.
13. Dubey AK, Kumar A, García-Díaz V, Sharma AK, Kanhaiya K. Study and analysis of SARIMA and LSTM in forecasting time series data. *Sustainable Energy Technologies and Assessments*. 2021;47:101474.
 14. Dudek G. STD: A Seasonal-Trend-Dispersion Decomposition of Time Series. *IEEE Transactions on Knowledge and Data Engineering*. 2023;35(10):10339-10350. doi:10.1109/TKDE.2023.3268125.
 15. Hajirahimi Z, Khashei M. Hybrid structures in time series modeling and forecasting: A review. *Engineering Applications of Artificial Intelligence*. 2019;86:83-106.
 16. Haselbeck F, Killinger J, Menrad K, Hannus T, Grimm DG. Machine learning outperforms classical forecasting on horticultural sales predictions. *Machine Learning with Applications*. 2022;7:100239.
 17. He M, Qian X. Forecasting tourist arrivals using STL-XGBoost method. *Tourism Economics*. 2026;32(2):408-436.
 18. Hermans K, McLeman R. Climate change, drought, land degradation and migration: exploring the linkages. *Current Opinion in Environmental Sustainability*. 2021;50:236-244. doi:10.1016/j.cosust.2021.04.013.
 19. Hewamalage H, Bergmeir C, Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*. 2021;37(1):388-427.
 20. Hsiang S, Kopp RE. An economist's guide to climate change science. *Journal of Economic Perspectives*. 2018;32(4):3-32.
 21. Jadhav V, Reddy B, Gaddi G. Application of ARIMA model for forecasting agricultural prices. 2018.
 22. Jaiswal R, Choudhary K, Kumar RR. STL-ELM: a decomposition-based hybrid model for price forecasting of agricultural commodities. *National Academy Science Letters*. 2022;45(6):477-480.
 23. Jankovic A, Chaudhary G, Goia F. Designing the design of experiments (DOE)-An investigation on the influence of different factorial designs on the characterization of complex systems. *Energy and Buildings*. 2021;250:111298.
 24. Kakpo A, Mills BF, Brunelin S. Weather shocks and food price seasonality in Sub-Saharan Africa: Evidence from Niger. *Food Policy*. 2022;112:102347. doi:10.1016/j.foodpol.2022.102347.
 25. Kmytiuk T, Majore G, Bilyk T. Time series forecasting of price of the agricultural products using data science. *Agricultural and Resource Economics: International Scientific E-Journal*. 2024;10(3):5-33. doi:10.22004/AG.ECON.355972.
 26. Kontopoulou VI, Panagopoulos AD, Kakkos I, Matsopoulos GK. A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. *Future Internet*. 2023;15(8):255.
 27. Kumar A, Bhatia A, Kashyap A, Kumar M. LSTM network: a deep learning approach and applications. In: *Advanced Applications of NLP and Deep Learning in Social Media Data*. Hershey: IGI Global Scientific Publishing; 2023. pp. 130-150.
 28. Kwarteng S, Andreevich P. Comparative analysis of ARIMA, SARIMA and Prophet model in forecasting. *Research & Development*. 2024;5(4):110-120.
 29. Kyo K, Noda H, Fang F. An integrated approach for decomposing time series data into trend, cycle and seasonal components. *Mathematical and Computer Modelling of Dynamical Systems*. 2024;30(1):792-813. doi:10.1080/13873954.2024.2416631.
 30. Lim B, Zohren S. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2021;379(2194).
 31. Lin F. Copula-based mixture transition distribution models for forecasting skewed and zero-inflated time series: methodology and comparisons with deep learning LSTM networks [dissertation]. Corvallis: Oregon State University; 2025.
 32. Ma H, Fadhil Marsani M, Mansor MA, Mohd Kasihmuddin MS. Hybrid deep learning and optimized variational mode decomposition for point-interval runoff prediction. *PLOS ONE*. 2026;21(3):e0343063. doi:10.1371/journal.pone.0343063.
 33. Montgomery DC. Design and analysis of experiments. New York: John Wiley & Sons; 2017.
 34. Murat M, Malinowska I, Gos M, Krzyszczak J. Forecasting daily meteorological time series using ARIMA and regression models. *International Agrophysics*. 2018;32(2).
 35. Ning Y, Kazemi H, Tahmasebi P. A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. *Computers & Geosciences*. 2022;164:105126.
 36. Ouyang Z, Ravier P, Jabloun M. STL decomposition of time series can benefit forecasting done by statistical methods but not by machine learning ones. *Engineering Proceedings*. 2021;5(1):42.
 37. Pokou F, Sadefo Kamdem J, Benhmad F. Hybridization of ARIMA with learning models for forecasting of stock market time series. *Computational Economics*. 2024;63(4):1349-1399.
 38. Sax C, Eddelbuettel D. Seasonal adjustment by x-13arima-seats in r. *Journal of Statistical Software*. 2018;87:1-17.
 39. Silvestre GD, dos Santos MR, de Carvalho AC. Seasonal-trend decomposition based on loess + machine learning: hybrid forecasting for monthly univariate time series. In: *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE; 2021.
 40. Slater L, Arnal L, Boucher MA, Chang AYY, Moulds S, *et al*. Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models. *Hydrology and Earth System Sciences Discussions*. 2022;2022:1-35.
 41. Sohrabbeig A, Ardakanian O, Musilek P. Decompose and Conquer: Time Series Forecasting with Multiseasonal Trend Decomposition Using Loess. *Forecasting*. 2023;5(4):684-696.
 42. Trull O, García-Díaz JC, Peiró-Signes A. Multiple seasonal STL decomposition with discrete-interval moving seasonalities. *Applied Mathematics and Computation*. 2022;433:127398.
 43. Wang S, Li C, Lim A. Why are the ARIMA and SARIMA not sufficient. *arXiv preprint arXiv:1904.07632*. 2019.
 44. Wu J, Hu Y, Wu D, Yang Z. An Aquatic Product Price Forecast Model Using VMD-IBES-LSTM Hybrid Approach. *Agriculture*. 2022;12(8):1185.