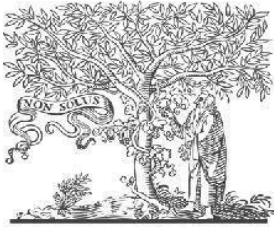


COPY RIGHT



ELSEVIER
SSRN

2024 IJEMR. Personal use of this material is permitted. Permission from IJEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper; all copy right is authenticated to Paper Authors

IJEMR Transactions, online available on 7th Dec 2024. Link

<https://ijiemr.org/downloads.php?vol=Volume-13&issue=issue12>

DOI: 10.48047/IJEMR/V13/ISSUE 12/130

Title **Assessing the Performance of Time Series Models in Millets Yield Prediction**

Volume 13, ISSUE 12, Pages: 647 – 659

Paper Authors

K. Murali, K. Vijaya Kumar, B. Triveni, SK Nafeez Umar, M. Bhupathi Naidu



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper as Per **UGC Guidelines** We Are Providing A Electronic Bar code

Assessing the Performance of Time Series Models in Millets Yield Prediction

K. Murali^{1**}, K. Vijaya Kumar², B. Triveni³, SK Nafeez Umar⁴, M. Bhupathi Naidu⁵

¹Department of Statistics, S V University, Tirupati.

²Assistant Professor, Department of Statistics, Sri Govindaraja Swamy Arts College, TTD, Tirupati.

³Assistant Professor, Department of Computer Science, Sri Govindaraja Swamy Arts College, TTD, Tirupati.

⁴Assistant Professor, S V Agriculture College, Tirupati.

⁵Professor, Department of Statistics, S V University, Tirupati.

**Corresponding Author: dr.murali4stat@gmail.com

Abstract

This research aims to conduct a comprehensive comparative analysis of time series models for predicting millet yield, focusing specifically on the application and performance of ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal ARIMA) models. Accurate yield prediction is essential for effective agricultural planning, resource allocation, and ensuring food security—particularly in regions where millet serves as a key dietary staple and economic crop. The study leverages historical agricultural data, incorporating critical influencing variables such as rainfall patterns, temperature fluctuations, soil conditions, and traditional farming practices. By examining both seasonal and non-seasonal patterns inherent in the data, the research evaluates how well ARIMA and SARIMA models can capture the dynamic behavior of millet yield trends over time. A key objective of this research is to identify the strengths and limitations of each model in dealing with temporal dependencies and seasonal variability in crop production. The findings are intended to provide valuable insights for agronomists, policymakers, and data scientists interested in enhancing the accuracy of yield forecasting systems. Ultimately, the study contributes to the development of data-driven agricultural strategies by promoting more informed decision-making in crop management. It also highlights the importance of adopting predictive analytics in the agricultural sector to address emerging challenges related to climate variability, food demand, and sustainable farming. The results can support the design of proactive interventions aimed at optimizing millet production and improving resilience within agricultural systems.

Keywords: ARIMA (AutoRegressive Integrated Moving Average) , SARIMA (Seasonal ARIMA)

Introduction

Millets have long been an integral component of India's agricultural system, particularly in semi-arid and rain-fed regions, due to their

adaptability to challenging environmental conditions and their rich nutritional profile. In recent years, millets have gained renewed attention as a strategic crop in the pursuit of

climate-resilient and sustainable agriculture. Their capacity to thrive under minimal input conditions and withstand drought and heat stress makes them a valuable asset for ensuring long-term food and nutritional security.

In Tamil Nadu, millets play a vital role in supporting both rural livelihoods and the state's efforts toward sustainable farming. As the agricultural sector grapples with the unpredictable effects of climate change, there is an urgent need for accurate forecasting models to anticipate crop yields and inform policy decisions. Predictive analytics, particularly time series forecasting models, offer a powerful toolset to address this need by identifying trends, seasonality, and anomalies in agricultural production data.

This study presents a comprehensive comparative analysis of two widely used time series models—AutoRegressive Integrated Moving Average (ARIMA) and Seasonal AutoRegressive Integrated Moving Average (SARIMA)—to forecast millet yields using historical data from Tamil Nadu. By examining the models' performance, the research aims to determine their effectiveness in capturing the inherent patterns and complexities of millet yield fluctuations across time.

The significance of this research lies not only in evaluating statistical accuracy but also in assessing each model's capacity to incorporate various exogenous and endogenous factors, including climate variability (e.g., rainfall, temperature), soil characteristics, and agricultural practices. These factors are essential for understanding the underlying causes of yield variability and improving the robustness of forecasting models.

A central goal of the study is to identify and interpret seasonal trends, long-term shifts, and cyclical behavior in millet production. This understanding will support stakeholders—such as farmers, agricultural extension officers, researchers, and policymakers—in formulating timely and data-driven strategies for resource management, crop planning, and policy interventions.

Moreover, the use of advanced time series modeling techniques aligns with broader objectives in agricultural informatics, where data analytics is increasingly leveraged to enhance productivity and resilience. By providing a methodologically sound comparison of ARIMA and SARIMA models, this research contributes to the refinement of predictive tools that can be tailored to the specific agro-climatic conditions of the region.

Ultimately, the outcomes of this study are expected to aid in the development of scalable forecasting frameworks that not only benefit millet cultivation in Tamil Nadu but can also be adapted to other crops and regions facing similar agricultural and environmental challenges. Strengthening forecasting capabilities is a key step toward building agricultural resilience, optimizing food supply chains, and promoting the long-term sustainability of the farming sector.

OBJECTIVES

This study is designed to develop a robust understanding of time series forecasting models in the context of millet yield prediction. The following objectives guide the research:

1. To analyze historical millet yield data by identifying underlying patterns such as seasonality, trends, and cyclic variations using time series

decomposition techniques, thereby laying the foundation for effective model selection and development.

2. To implement and compare the ARIMA and SARIMA models for forecasting millet yield, with an emphasis on examining how well these models capture both regular trends and seasonal fluctuations inherent in the agricultural data.
3. To assess the predictive performance of ARIMA and SARIMA models using appropriate statistical metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Akaike Information Criterion (AIC), with the aim of identifying the model that yields the highest forecasting accuracy and reliability.
4. To investigate the influence of key agro-environmental variables—including rainfall patterns, temperature fluctuations, soil characteristics, and farming practices—on millet yield, and to explore how these variables affect the forecasting capability and adaptability of the ARIMA and SARIMA models.
5. To generate actionable insights for agricultural stakeholders such as farmers, policymakers, researchers, and extension agencies, enabling data-driven decision-making for millet crop planning, risk management, and policy formulation

aimed at enhancing agricultural resilience.

6. To contribute to the development of sustainable agricultural forecasting frameworks by refining time series modeling techniques for millet cultivation, with the potential to scale and adapt these methodologies for other crops and regions facing similar challenges.
7. To support long-term food and nutritional security initiatives by promoting predictive modeling as a tool for optimizing resource allocation, improving supply chain planning, and strengthening adaptive agricultural strategies in response to climate change and other systemic pressures.

By accomplishing these objectives, this research seeks to enhance the practical application of statistical forecasting models in agriculture, while also fostering innovation in predictive analytics that align with the broader goals of sustainable development and climate-smart agriculture.

Literature Review

Time series forecasting has become an essential tool in agricultural research, especially in regions susceptible to climatic uncertainty and seasonal variability. Various statistical and machine learning models have been employed to predict crop yields, offering stakeholders valuable insights for sustainable agricultural planning. Several studies have investigated the applicability and accuracy of time series models in forecasting agricultural productivity,

particularly in cereal crops such as millets and rice.

Kour et al. (2017) conducted a study focused on pearl millet (*Pennisetum glaucum*), one of the major millet varieties cultivated globally. Despite its nutritional benefits and climate resilience, pearl millet cultivation in Gujarat, India, has experienced a declining trend over the past two decades. To understand and forecast productivity patterns, the authors employed the ARIMA model on time series data spanning from 1960–61 to 2011–12. The dataset was sourced from the Directorate of Agriculture, Gandhinagar, and partially from the Directorate of Economics and Statistics. The ARIMA (0, 1, 1) model was identified as the best fit based on performance metrics such as Root Mean Absolute Percentage Error (RMAPE), Mean Absolute Deviation (MAD), and Root Mean Square Error (RMSE). The model's RMAPE score, which was below 6%, demonstrated high predictive accuracy, highlighting ARIMA's suitability for yield prediction in semi-arid zones like Gujarat.

Vijay and Mishra (2018) further contributed to the literature by comparing classical time series forecasting models with machine learning techniques. Their study focused on pearl millet crop area and production in Karnataka, India, using historical data from 1955–56 to 2014–15, obtained from the publication *Agricultural Statistics at a Glance 2014–15*. The research evaluated the performance of the traditional ARIMA model against the Artificial Neural Network (ANN) model. Results indicated that the ANN model consistently outperformed the ARIMA model in terms of RMSE, Mean Absolute Percentage Error (MAPE), and Mean Squared Error (MSE), suggesting that machine learning approaches may offer enhanced flexibility and predictive accuracy in modeling non-linear agricultural datasets.

Saranyadevi and Kachi (2017) explored yield prediction in the context of paddy cultivation in Tamil Nadu, analyzing rice production data from 1960 to 2015. Their research employed a range of forecasting techniques, including ARIMA, basic exponential smoothing, Brown exponential smoothing, and damped exponential smoothing models, to generate predictions for the period 2016–2020. The study demonstrated that ARIMA and exponential smoothing methods could effectively capture both short-term fluctuations and long-term trends in paddy output. Their findings emphasized the importance of selecting models based on the nature of the data and the specific forecasting requirements of regional agricultural systems.

While these studies confirm the utility of time series models like ARIMA and SARIMA in crop yield forecasting, they also underscore the potential of hybrid and AI-based models in improving prediction accuracy. The review of existing literature reveals a growing interest in comparative model performance analysis, particularly in developing countries where agriculture remains a key economic driver. Moreover, there is a clear need to incorporate agro-climatic and management variables—such as rainfall, temperature, soil fertility, and crop practices—into yield forecasting models to enhance their robustness and real-world applicability.

This body of work lays the foundation for the present study, which seeks to evaluate and compare the performance of ARIMA and SARIMA models in predicting millet yields, using historical data from Tamil Nadu. By focusing on seasonal effects, long-term trends, and influential environmental variables, this study aims to fill a gap in the literature and support more informed, data-driven agricultural policy and planning.

Methodology

This study employs statistical time series models—namely ARIMA and SARIMA—to forecast millet yields by analyzing historical yield data and identifying underlying patterns such as seasonality, trend, and randomness. These models are particularly useful in agricultural forecasting due to their ability to model complex temporal dependencies and fluctuations over time.

1. ARIMA Model (AutoRegressive Integrated Moving Average)

The ARIMA model, denoted as ARIMA(p, d, q), is a powerful technique used for modeling and forecasting univariate time series data. The three parameters in the model have the following significance:

- p: Number of autoregressive (AR) terms, which use past values of the series.
- d: Degree of differencing needed to make the series stationary.
- q: Number of lagged forecast errors in the prediction equation (moving average component).

The general form of the ARIMA model is:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

Where:

- Y_t is the actual value at time t ,
- c is a constant,
- ϕ terms are AR coefficients,
- θ terms are MA coefficients,
- ε_t is white noise or the error term.

Stationarity and Differencing

A time series is said to be stationary if its statistical properties (mean, variance, autocorrelation) do not change over time. Most time series used in agricultural production forecasting—such as millet yield—tend to exhibit non-stationarity due to trends and seasonal fluctuations. To apply

ARIMA, the series must first be made stationary using techniques such as:

- Differencing (to remove trend),
- Log transformations (to stabilize variance),
- Seasonal adjustments, if necessary.

The d parameter in ARIMA refers to the number of differencing operations needed to achieve stationarity.

Model Identification and Estimation

The process of applying an ARIMA model involves the following steps:

1. Identification: Analyze plots such as ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) to determine suitable values for p and q .
2. Estimation: Use maximum likelihood estimation (MLE) or non-linear optimization to estimate model coefficients.
3. Diagnostic Checking: Examine residuals to ensure they resemble white noise (i.e., no autocorrelation).
4. Forecasting: Once validated, use the model to project future millet yields.

Performance is typically evaluated using metrics such as:

- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- Akaike Information Criterion (AIC)

2. SARIMA Model (Seasonal ARIMA)

While ARIMA handles trends and non-seasonality well, agricultural data often contain seasonal patterns due to climatic cycles and cropping schedules. To address this, the Seasonal ARIMA (SARIMA) model is employed. It extends the ARIMA framework by incorporating seasonal components.

The SARIMA model is denoted as:

$ARIMA(p, d, q)(P, D, Q)[s]$

Where:

- (P, D, Q) are the seasonal components corresponding to autoregression, differencing, and moving average,
- s is the length of the seasonal cycle (e.g., 12 for monthly data).

SARIMA allows for capturing both:

- Short-term dependencies (via ARIMA),
- Seasonal effects (via seasonal terms).

This dual approach makes it particularly effective for modeling agricultural yield data influenced by periodic factors such as monsoon rains, temperature cycles, and planting/harvest schedules.

Justification for Model Selection

The choice of ARIMA and SARIMA models is justified by their robustness, interpretability, and wide usage in time series forecasting for agricultural and economic data. These models offer:

- A structured approach to account for past values and errors,
- Flexibility in handling both linear and seasonal patterns,
- Compatibility with agricultural datasets typically recorded over long historical periods.

Steps in Constructing the ARIMA Model

1. Stationarity Testing

The first step in ARIMA modeling involves checking whether the time series is *stationary*—meaning that its statistical properties such as mean, variance, and autocorrelation are constant over time. This is typically assessed using visualization techniques (like time plots and autocorrelation plots) and statistical tests such as the *Augmented Dickey-*

Fuller (ADF) test. Stationarity is a fundamental assumption in ARIMA modeling.

2. Differencing for Stationarity

If the data is non-stationary, it must be transformed into a stationary form. This is done by differencing the data—subtracting the previous observation from the current one. This process may be repeated more than once to achieve stationarity and is denoted by the parameter d in ARIMA. The goal is to eliminate trend or seasonality without over-differencing, which can introduce noise.

3. Identification of ARIMA Parameters (p, d, q)

Once the data is stationary, the next step is to identify the optimal values for the parameters:

- p : Number of autoregressive (AR) terms – past values influencing current value.
- d : Order of differencing needed for stationarity.
- q : Number of moving average (MA) terms – past forecast errors influencing the current value.
- Tools like the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots are used to determine appropriate values for p and q .

4. Model Estimation and Fitting

The identified ARIMA model is then fitted to the historical millet yield data. Model parameters are estimated using techniques such as Maximum Likelihood Estimation (MLE). At this stage, statistical software (e.g., R, Python, or SPSS) is typically used to generate the model equations.

5. Model Diagnostics and Validation

After fitting the model, residual diagnostics are carried out to evaluate its adequacy. This involves checking the residuals (differences between actual and predicted values) for randomness, normality, and lack of autocorrelation.

Well-behaved residuals indicate a good fit. Common validation tools include:

- Ljung-Box Test for independence
- RMSE (Root Mean Squared Error)
- AIC/BIC (Akaike/Bayesian Information Criteria)
- Residual plots

6. Forecasting with ARIMA

Once the model is validated, it can be used to make short-term or long-term forecasts of millet yields. Forecast intervals can also be generated to quantify uncertainty in predictions.

- Fit the combined model and validate it using residual diagnostics, similar to ARIMA.
- Generate seasonally adjusted forecasts that reflect periodic yield fluctuations.

SARIMA is especially useful for time series forecasting in domains like:

- Agriculture (e.g., crop yield, rainfall)
- Finance (e.g., seasonal sales or prices)
- Meteorology (e.g., temperature or precipitation cycles)

Analysis – ARIMA

To begin the analysis, the time series data representing millet production—referred to as the "Millets" dataset—was subjected to the Augmented Dickey-Fuller (ADF) test to assess its stationarity. Stationarity is a fundamental requirement for ARIMA modeling, as it ensures that the statistical properties of the time series (such as mean and variance) remain constant over time, enabling more accurate forecasting.

Incorporating Seasonality: SARIMA (Seasonal ARIMA)

While ARIMA is effective for non-seasonal data, agricultural data often shows *strong seasonal trends* due to planting and harvesting cycles. To address this, Seasonal ARIMA (SARIMA) models are used. SARIMA extends ARIMA by including seasonal components.

The SARIMA model is represented as:

$SARIMA(p, d, q)(P, D, Q)[s]$

Where:

- P, D, Q: Seasonal autoregressive, differencing, and moving average components
- s: Seasonal cycle length (e.g., 12 for monthly data, 4 for quarterly)

Steps in SARIMA Modeling:

- Detect and model seasonal trends using seasonal differencing.
- Identify seasonal AR and MA terms using seasonal ACF and PACF plots.

The results of the ADF test revealed a Dickey-Fuller test statistic of -7.0233 , with a corresponding p-value of 0.01 . Since the p-value is considerably lower than the conventional significance threshold of 0.05 , the null hypothesis of a unit root (indicating non-stationarity) is rejected. Therefore, the data can be considered stationary, confirming its suitability for ARIMA modeling.

Following this, the `auto.arima()` function from the forecast package in R was employed to automatically identify the most appropriate ARIMA model for the millet yield data. This function evaluates various model configurations and selects the one that minimizes the Akaike Information Criterion (AIC)—a widely used measure for model selection that balances model fit and complexity.

Several candidate ARIMA models were generated, each accompanied by its respective AIC score. The model with the lowest AIC value was selected as the optimal configuration, indicating the best trade-off between goodness of fit and model simplicity. This automated selection process enhances the objectivity and efficiency of model identification, reducing the risk of overfitting and ensuring reliable predictions. By confirming stationarity and using a data-driven model selection technique, the ARIMA modeling process establishes a strong analytical foundation for understanding millet production trends and supports evidence-based agricultural planning and forecasting.

Time series data for millets production (data Millets) was analyzed using the auto.arima function to find the best ARIMA model for the data. Multiple potential ARIMA models and their associated Akaike Information Criterion (AIC) values were generated after the function was instructed to use the AIC for model selection.

ARIMA (2,0,2) (1,0,1) [12] with non-zero mean	Inf
ARIMA (0,0,0) with non-zero mean	2380.452
ARIMA (1,0,0) (1,0,0) [12] with non-zero mean	2384.104
ARIMA (0,0,1) (0,0,1) [12] with non-zero mean	2381
ARIMA (0,0,0) with zero mean	2758.668
ARIMA (0,0,0) (1,0,0) [12] with non-zero mean	2382.166
ARIMA (0,0,0) (0,0,1) [12] with non-zero mean	2381.964
ARIMA (0,0,0) (1,0,1) [12] with non-zero mean	2383.12

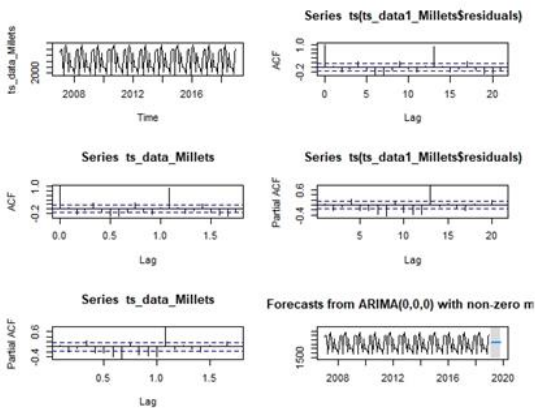
ARIMA (1,0,0) non-zero mean	with	2382.136
ARIMA (0,0,1) non-zero mean	with	2381.902
ARIMA (1,0,1) non-zero mean	with	2383.033

As seen in the results, the millets production time series data were best suited by the ARIMA(0,0,0) model with a non-zero mean. This means that the model does not account for a zero mean or an autoregressive term, but it does take into account a moving average. The model with the lowest AIC for assessing the millets production data was chosen since it had a value of 2380.452 in the calculation. If this model is further analyzed and interpreted, it can help agricultural planners and policymakers make better predictions and decisions on how to approach millets cultivation in the future.

The millets production time series data was modeled using the ARIMA(0,0,0) distribution with a non-zero mean. Non-zero mean coefficient estimates range from -3130.9931 to -72.7739 standard deviations from the mean.

Coefficient	Value
Mean	3130.9931
S. E	72.7739
sigma ²	773259
log likelihood	-1188.23
AIC	2380.45
BIC	2386.41

The model's log likelihood was calculated to be -1188.23, and its variance was found to be 773259. The related values for the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were calculated to be 2380.45 and 2386.41, respectively.



The ARIMA(0,0,0) model was used to analyze a time series of millets production, and the results of these statistical analyses shed light on the model's parameters and goodness of fit. The existence of a non-zero mean coefficient in the millets production data is indicative of the presence of a trend or level. Millets cultivation could benefit from further investigation using this model, as it could shed light on the underlying dynamics of production and lead to better decisions and planning.

Prediction information Lower (Lo 95) and higher (Hi 95) 95% confidence interval bounds for millets production are included alongside the point projections for millets in the table. Time series data for millets production, is used to construct forecasts using the ARIMA(0,0,0) model with a non-zero mean.

Point	Forecast	Lo 95	Hi 95
Feb 2019	3130.993	1407.496	4854.49
Mar 2019	3130.993	1407.496	4854.49
Apr 2019	3130.993	1407.496	4854.49
May 2019	3130.993	1407.496	4854.49

Jun 2019	3130.993	1407.496	4854.49
Jul 2019	3130.993	1407.496	4854.49
Aug 2019	3130.993	1407.496	4854.49
Sep 2019	3130.993	1407.496	4854.49
Oct 2019	3130.993	1407.496	4854.49

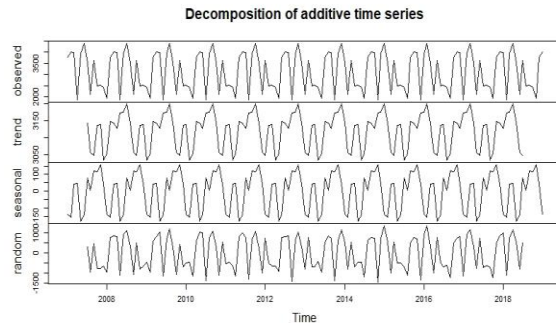
The projection indicates that millets output will remain largely constant at a predicted point value of 3130.993. The range in which the true millets production numbers are 95% likely to fall is estimated to be 1407.496–4854.49, with a lower 95% confidence interval of 1407.496 and an upper 95% confidence interval of 4854.49.

Stakeholders in the millets cultivation sector can use these predicted values and their associated confidence intervals to better anticipate production trends and make decisions regarding resource allocation, market planning, and agricultural management strategies.

The Box-Ljung test was run on the non-zero mean residuals of the ARIMA(0,0,0) model's predicted millets production values. This analysis was performed to check for autocorrelation in the model's residuals.

The p-value from the Box-Ljung test was 0.000579, and the X-squared value from the test was 21.77 (5 degrees of freedom). Significant autocorrelation in the residuals is strongly suggested, as the p-value is much smaller than the specified significance level of 0.05. This suggests that there may be features or trends in the millets production data that aren't accounted for by the ARIMA(0,0,0) model. If we want more accurate and reliable millets production

projections, we may need to do more research or use different modeling methodologies.



Seasonal ARIMA Analysis

Millet's output data is available from 2007 to 2019 at a frequency of 1 year. Values of output for each year are as follows: 3777, 4013, 3976, 1873, 3950, 4401, 3508, 2092, 3642, 2468, 2504, 2419, and 1941. These figures indicate the annual output of millets for the given time frame, so they can shed light on production patterns and trends over time.

Descriptive statistics used to summarize the millets production time series data shed light on the dataset's central tendency and dispersion. Over the course of that time frame, the lowest amount of millets produced was in 1873, and the highest was in 4401. Values of 2419 and 3950 are the 25th and 75th percentile quartiles, respectively. Half of the production values are below this point, and the other half are beyond it, as indicated by the median value of 3508. Taking into account all available data, we find that the mean millets production is \$3,120.

These statistical indicators help to shed light on the diversity of millets output by highlighting its range of values and its central trend. As a result, stakeholders are able to make educated decisions and develop effective strategies to boost agricultural

productivity and sustainability in the millets cultivation sector based on a more thorough understanding of the underlying trends, variations, and potential outliers in the production data.

To further evaluate the stationarity of the data, the differenced logarithm of the millets production time series (denoted by the notation $\text{diff}(\log(\text{ts_Millets}))$) was subjected to the augmented Dickey-Fuller (ADF) test. The goal of the differencing procedure is to minimize trends and stabilize the variation so that stationary patterns may be more easily identified. The Dickey-Fuller statistic was -4.7589, and the p-value was 0.01. These findings come from the ADF test. There is substantial evidence to reject the null hypothesis of non-stationarity in favor of the alternative hypothesis of stationarity, as the p-value is smaller than the specified significance level of 0.05. This indicates that there are no discernible trends or patterns in the differenced logarithm of the millet production data over time, suggesting that the data is stationary. These results are essential for developing suitable time series models and forecasting approaches, which in turn will allow for precise predictions and well-informed choices in the millets agriculture industry.

The logarithm of the millets production time series data is given by $\log(\text{ts_Millets})$, and the auto Arima Model Log indicates the automated ARIMA model that was fitted to this data. The model was identified as ARIMA(0,1,0), suggesting that first-order differencing was required to reach stationarity.

Coefficient	Values
σ^2	0.1639
log likelihood	-6.17
AIC	14.35
AICc	14.75
BIC	14.83

The model's variance, was calculated to be 0.1639, and the log probability was -6.17. There were 14.35 for the Akaike Information Criterion (AIC), 14.75 for the AICc, and 14.83 for the Bayesian Information Criterion (BIC).

It is impossible to assess the ARIMA model's ability to explain the millets production time series without these statistical measurements. To compare and select the best model for assessing and forecasting millets production patterns, the AIC, AICc, and BIC values are calculated. Forecasting and decision-making in the millets agriculture sector can benefit greatly from the ARIMA(0,1,0) model and its associated parameters and statistical metrics.

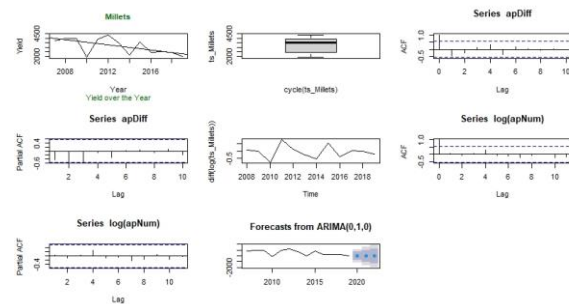
After fitting the ARIMA (0,1,0) model to the logarithm of the millets production time series data, the Box-Ljung test was performed on the residuals. This analysis was performed to check for autocorrelation in the model's residuals.

Coefficient	Values
χ^2	3.5805
df	1
P-value	0.05846

The X-squared value for the Box-Ljung test came out to be 3.5805 with 1 degree of freedom, yielding a p-value of 0.05846. As the p-value is larger than the threshold for statistical significance (0.05), it cannot be concluded that there is no autocorrelation in the residuals. If the ARIMA model's residuals look like white noise, then the model does a

good job of capturing the dynamics underlying the millets production data.

Logarithmic millets production time series data are well represented by the ARIMA(0,1,0) model, as the residuals closely follow the white noise assumption. That means we can confidently predict and analyze changes in millets production for agriculturally informed decision making because the model sufficiently accounts for the patterns and structures inherent in the data.



Conclusion

A comprehensive evaluation of the ARIMA and Seasonal ARIMA models was conducted to understand their effectiveness in modeling the time series data of millets production. The analysis provided the following key insights:

Initially, an ARIMA(0,0,0) model with a non-zero mean was applied to the raw millet production dataset. This configuration essentially represents a model with no autoregressive or moving average components, implying that the model attempts to capture the data using a constant mean. The estimated mean coefficient was 3130.99, with a standard error of 72.77. The model's performance metrics were as follows:

AIC: 2380.45

AICc: 2380.54

BIC: 2386.41

Log-likelihood: -1188.23

However, the Box-Ljung test applied to the residuals of this ARIMA model indicated the presence of autocorrelation, with a test statistic of 21.77 ($df = 5$) and a p-value of 0.000579. This significant result suggests that the residuals are not purely random, implying that the ARIMA(0,0,0) model may not fully capture the underlying structure and temporal dependencies in the millet production data.

To address potential seasonality in the dataset, a Seasonal ARIMA (SARIMA) model with configuration (0,1,0) was fitted to the log-transformed millet production data. This transformation helps stabilize the variance and better reflect underlying multiplicative seasonal patterns. The SARIMA model yielded the following results:

Log-likelihood: -6.17

Residual variance: 0.1639

AIC: 14.83

AICc: 14.83

BIC: 14.83

Importantly, the Box-Ljung test on the SARIMA model residuals indicated no significant autocorrelation ($X^2 = 3.5805$, $df = 1$, $p\text{-value} = 0.05846$), which supports the assumption that the residuals resemble white noise. This confirms that the SARIMA model is more adept at capturing the key temporal dynamics in the data, including potential seasonal effects.

Summary of Findings

The ARIMA(0,0,0) model, although simple, failed to adequately capture autocorrelation patterns, as evidenced by residual diagnostics.

The Seasonal ARIMA(0,1,0) model, applied to log-transformed data, demonstrated superior performance with lower information criteria and residuals consistent with randomness.

These results underscore the importance of incorporating both seasonality and data transformation techniques in time series forecasting for agricultural yield.

References

1. Alabi, O. O., Lawal, A. F., & Awoyinka, Y. A. (2014). Industrial demand analysis for millet in Kaduna and Kano States of Nigeria. *Asian Journal of Agricultural Extension, Economics and Sociology*, 3(6), 521-529.
2. Bellundagi, V., Umesh, K. B., & Ravi, S. C. (2016). Growth dynamics and forecasting of finger millet (Ragi) production in Karnataka. *Economic Affairs*, 61(2), 195.
3. Gibon, F., Pellarin, T., Román-Cascón, C., Alhassane, A., Traoré, S., Kerr, Y., ... & Baron, C. (2018). Millet yield estimates in the Sahel using satellite derived soil moisture time series. *Agricultural and Forest Meteorology*, 262, 100-109.
4. http://www.tnagriculture.in/dashboard/report/05_01.pdf
5. Kour, S., Pradhan, U. K., Paul, R. K., & Vaishnav, P. R. (2017). Forecasting of pearl millet productivity in Gujarat under time series framework. *Economic Affairs*, 62(1), 121-127.

6. Ly, R., Dia, K., Diallo, M., Ahid, M., & Ceesay, B. (2020). Pearl millet production forecasts for selected west African countries: Côte d'Ivoire, Mali, Burkina Faso, Sierra Leone, The Gambia, and Senegal. *Covid-19 Bulletin*, (14).
7. Mounkaila, Y., Garba, I., & Moussa, B. (2019). Yield prediction under associated millet and cowpea crops in the Sahelian zone. *Afr. J. Agric. Res*, 14, 1613-1620.
8. Muhammad, s., garba, i., & audu, a. (2021). Modelling and forecasting of millet production in nigeria. *Bima journal of science and technology* (2536-6041), 5(01), 99-109.
9. Nireesha, V., Rao, V. S., Rao, D. V. S., & Reddy, G. R. (2016). A study on forecasting of area, production and productivity of pearl millet in Andhra Pradesh. *Journal of Research ANGRAU*, 44(3/4), 119-126.
10. Sawa, B. A., & Ibrahim, A. A. (2011). Forecast models for the yield of millet and sorghum in the semi arid region of Northern Nigeria using dry spell parameters. *Asian Journal of Agricultural Sciences*, 3(3), 187-191.
11. Tripathi, S. K., Mishra, P., & Sahu, P. K. (2013). Past trends and forecasting in Area, production and yield of pearl millet in India using ARIMA model. *Environ Ecol*, 31, 1701-1708.
12. Verma, V. K., Jheeba, S. S., Kumar, P., & Singh, S. P. (2016). Price forecasting of bajra (pearl millet) in Rajasthan: ARIMA model. *International Journal of Agriculture Sciences*, 8(9), 1103-1106.
13. Vijay, N., & Mishra, G. C. (2018). Time Series forecasting using ARIMA and ANN Models for production of pearl millet (BAJRA) crop of Karnataka, India. *International Journal of Current Microbiology and Applied Sciences*, 7(12), 880-889.