IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

Time Series Analysis for Forecasting Paddy Production in Tamil Nadu

T. Gangaram[1] , V. Munaiah[2] , P. Maheswari[3] , K.Murali[4] , G.Mokesh Rayalu[5]**

[1]Assistant Professor, Dept. of Statistics, SVA Govt. College for Men, Srikalahasti.
[2]Assistant Professor, Dept. of Statistics, PVKN Govt. College (A), Chittoor.
[3]Assistant Professor, Dept. of Statistics, Govt Degree College for Women, Srikalahasti.
[4]Academic Consultant, Dept. of Statistics, S.V University, Tirupati.
Corresponding Author **
[5]Assistant Professor Grade 2,Department of Mathematics, School of Advanced Sciences, VIT,Vellore
mokesh.g@gmail.com

ABSTRACT

In order to predict paddy output in Tamil Nadu, this study uses time series analysis using the robust ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal AutoRegressive Integrated Moving Average) models. This research makes use of historical data covering a number of years in order to investigate the complex temporal patterns and seasonal fluctuations that greatly affect paddy yields in the area. The research is conducted with the intention of developing a reliable framework for forecasting future paddy output, taking into account relevant aspects such as meteorological fluctuations, irrigation techniques, and governmental interventions. Farmers, policymakers, and others in the paddy cultivation sector can greatly benefit from a deeper understanding of the non-stationary and seasonal components within the industry thanks to the combination of ARIMA and SARIMA models. Sustainable and resilient paddy production in Tamil Nadu is ensured thanks to this study's contribution to the improvement of agricultural plans and policies.

Keywords: Paddy, ARIMA, SARIMA, Forecasting.

INTRODUCTION

Producing paddy acts as a cornerstone of Tamil Nadu's agricultural economy. It plays a key role in maintaining food security and sustaining the livelihoods of millions of people, making it one of the most important agricultural activities in the state. Given the region's susceptibility to climate changes and the ever-evolving agricultural techniques, it is becoming increasingly important to have a solid understanding of the temporal patterns and complicated dynamics that control paddy farming. This research attempts to provide a complete framework for forecasting paddy output in Tamil Nadu. It does so by utilizing time series analytic techniques, in particular the ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal AutoRegressive Integrated Moving Average) models. This study aims to untangle the temporal fluctuations and identify the important factors that influence paddy yield variability. It does so by utilizing historical data and taking into consideration a variety of seasonal and trend components. Ultimately, this study will use this information to make predictions. The combination of the ARIMA and SARIMA models enables a more nuanced understanding of the seasonal patterns and their impact on paddy production. This, in turn, makes it easier for farmers, policymakers, and other stakeholders invested in the environmentally responsible growth of Tamil Nadu's agricultural landscape to make informed decisions. This research has tremendous promise in improving the resilience and productivity of paddy agriculture and, as a result, making a contribution to the overall agricultural sustainability and food security in the region.

OBJECTIVE

1. To analyze historical time series data of paddy production in Tamil Nadu and identify the underlying trends and patterns affecting production fluctuations.

2. To apply the ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal AutoRegressive Integrated Moving Average) models to develop accurate and reliable forecasts for paddy production in the region.

3. To assess the impact of seasonal variations, climatic factors, and agricultural practices on paddy production, considering both short-term and long-term implications.

4. To compare the performance of the ARIMA and SARIMA models in capturing the seasonal variability and fluctuations in paddy production, thereby determining the most suitable model for forecasting in the context of Tamil Nadu's agricultural landscape.

5. To provide valuable insights for farmers, policymakers, and stakeholders, enabling informed decision-making for sustainable agricultural planning and policy formulation aimed at enhancing paddy production and ensuring food security in Tamil Nadu.

By achieving these objectives, this study aims to contribute to the development of robust forecasting methodologies and data-driven strategies that will support the resilience and growth of the paddy cultivation sector in Tamil Nadu, fostering sustainable agricultural practices and bolstering the region's agricultural productivity.

LITERATURE SURVEY

Amarender and Ashwini Darekar invested India produces the second-most paddy in the world. About 35% of net cultivated land and 50% of farmers grow paddy annually. Future harvest prices determine farmers' paddy acreage decisions. This research proposes a method to forecast harvest prices and applies it to kharif

2017–18. AGMARK's monthly average paddy prices from January 2006 to December 2016 were used. The ARIMA (Box-Jenkins) model predicted paddy prices. R was used to estimate model parameters. The model's goodness of fit was assessed using AIC, BIC, and MAPE. India-wide paddy price forecasts were best with the ARIMA model. September–November is the kharif paddy harvest. For the 2017-18 kharif harvest, paddy prices are expected to range from Rs. 1,600 to 2,200 per quintal.

The study by Saranyadevi and Kachi (2017), In this study, they investigate the predictive performance of a time-series analytic method for paddy production patterns in the Indian state of Tamil Nadu. There was a study that looked at paddy crop production data from 1960 to 2015 and predicted production for 2016–2020 using ARIMA (Autor Regressive Integrated Moving Average), simple exponential smoothing, brown exponential smoothing, and damped exponential smoothing models.

Joshua et al., (2021) Each model is evaluated using R2, RMSE, MAE, MSE, MAPE, CV, and NMSE. The GRNN method outperforms other assessment measures, including R2, RMSE, MAE, MSE, MAPE, CV, and NSME, with values of 0.9863, 0.2295, 0.1290, 0.0526, 1.3439, 0.0255, and 0.0136. These data show that the system estimates crop yield better than other methods. The Generalized Regression Neural Network (GRNN) model is compared to other models in literature studies. Using appropriate metrics, the GRNN model has greater prediction accuracy.

Methodology

ARIMA Model (p,d,q):

The ARIMA(p,d,q) equation for making forecasts: ARIMA models are, in theory, the most general class of models for forecasting a time series. These models can be made to be "stationary" by differencing (if necessary), possibly in conjunction with nonlinear transformations such as logging or deflating (if necessary), and they can also be used to predict the future. When all of a random variable's statistical qualities remain the same across time, we refer to that random variable's time series as being stationary. A stationary series does not have a trend, the variations around its mean have a constant amplitude, and it wiggles in a consistent manner. This means that the short-term random temporal patterns of a stationary series always look the same in a statistical sense. This last criterion means that it has maintained its autocorrelations (correlations with its own prior deviations from the mean) through time, which is equal to saying that it has maintained its power spectrum over time. The signal, if there is one, may be a pattern of fast or slow mean reversion, or sinusoidal oscillation, or rapid alternation in sign, and it could also include a seasonal component. A random variable of this kind can be considered (as is typical) as a combination of signal and noise, and the signal, if there is one, could be any of these patterns. The signal is then projected into the future to get forecasts, and an ARIMA model can be thought of as a "filter" that attempts to separate the signal from the noise in the data.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

The ARIMA forecasting equation for a stationary time series is a linear (i.e., regression-type) equation in which the predictors consist of lags of the dependent variable and/or lags of the forecast errors. That is:

Predicted value of Y = a constant and/or a weighted sum of one or more recent values of Y and/or a weighted sum of one or more recent values of the errors.

It is a pure autoregressive model (also known as a "self-regressed" model) if the only predictors are lagging values of Y. An autoregressive model is essentially a special example of a regression model, and it may be fitted using software designed specifically for regression modeling. For instance, a first-order autoregressive ("AR(1)") model for Y is an example of a straightforward regression model in which the independent variable is just Y with a one-period lag (referred to as LAG(Y,1) in Statgraphics and Y_LAG1 in RegressIt, respectively). Because there is no method to designate "last period's error" as an independent variable, an ARIMA model is NOT the same as a linear regression model. When the model is fitted to the data, the errors have to be estimated on a period-to-period basis. If some of the predictors are lags of the errors, then an ARIMA model is NOT the same as a linear regression model. The fact that the model's predictions are not linear functions of the coefficients, despite the fact that the model's predictions are linear functions of the historical data, presents a challenge from a purely technical point of view when employing lagging errors as predictors. Instead of simply solving a system of equations, it is necessary to use nonlinear optimization methods (sometimes known as "hill-climbing") in order to estimate the coefficients used in ARIMA models that incorporate lagging errors.

Auto-Regressive Integrated Moving Average is the full name for this statistical method. Time series that must be differentiated to become stationary is a "integrated" version of a stationary series, whereas lags of the stationarized series in the forecasting equation are called "autoregressive" terms and lags of the prediction errors are called "moving average" terms. Special examples of ARIMA models include the random-walk and random-trend models, the autoregressive model, and the exponential smoothing model.

A nonseasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, where:

- **p** is the number of autoregressive terms,
- **d** is the number of nonseasonal differences needed for stationarity, and
- **q** is the number of lagged forecast errors in the prediction equation.
- The forecasting equation is constructed as follows. First, let y denote the d^{th} difference of Y, which means:
- If d=0: $y_t = Y_t$
- If d=1: $y_t = Y_t - Y_{t-1}$
- If d=2: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$

- Note that the second difference of Y (the d=2 case) is not the difference from 2 periods ago. Rather, it is the first-difference-of-the-first difference, which is the discrete analog of a second derivative, i.e., the local acceleration of the series rather than its local trend.
- In terms of y, the general forecasting equation is:
- $\hat{Y}_t = \mu + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$

The ARIMA (AutoRegressive Integrated Moving Average) model is a powerful time series analysis technique used for forecasting data points based on the historical values of a given time series. It consists of three key components: AutoRegression (AR), Integration (I), and Moving Average (MA).

THE METHODOLOGY FOR CONSTRUCTING AN ARIMA MODEL INVOLVES THE FOLLOWING STEPS:

1. Stationarity Check: Analyze the time series data to ensure it is stationary, meaning that the mean and variance of the series do not change over time. Stationarity is essential for ARIMA modeling.

2. Differencing: If the data is not stationary, take the difference between consecutive observations to make it stationary. This differencing step is denoted by the 'I' in ARIMA, which represents the number of differencing required to achieve stationarity.

3. Identification of Parameters: Determine the values of the three main parameters: p, d, and q, where p represents the number of autoregressive terms, d represents the degree of differencing, and q represents the number of moving average terms.

4. Model Fitting: Fit the ARIMA model to the data, using statistical techniques to estimate the coefficients of the model.

5. Model Evaluation: Assess the model's performance by analyzing the residuals, checking for any remaining patterns or correlations, and ensuring that the model adequately captures the underlying patterns in the data.

6. Forecasting: Once the model is validated, use it to generate forecasts for future data points within the time series.

SEASONAL ARIMA:

By including seasonal variations into the ARIMA model, Seasonal ARIMA (SARIMA) is a robust technique for analyzing and forecasting time series data. It works well for examining and forecasting sales data, weather patterns, and economic indicators that are subject to seasonal changes. Financial markets, economics, and even meteorology all make use of SARIMA models.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

Mathematical Formulation:

The SARIMA model is denoted as SARIMA(p,d,q)(P,Q,D)[s], where:

- Non-seasonal autoregressive (p), differencing (d), and moving average (q) are the possible orders of analysis.
- The seasonal autoregressive, differencing, and moving average orders are denoted by the letters P, D, and Q, respectively.
- The length of one season is denoted by the symbol S.

The SARIMA model can be represented as follows:

$$(1-\varphi_1 B-\cdots-\varphi_P B^P)(1-\varphi_1 B^{VS}-\cdots-\varphi_P B^{VS})^P (B^{VS})^D Y_t$$
$$=(1+\theta_1 B+\cdots+\theta_P B^{\varphi})(1+\theta_1 B^{VS}+\cdots+\theta_P B^{\varphi S})^A (B^{\varphi S})^K \varepsilon_t$$

Where:

- $\varphi_i$and θ_i are the autoregressive and moving average parameters, respectively.
- B and B^{VS} are the non- seasonal and seasonal backshift operators, respectively.
- P,D,A and K are the orders of the seasonal autoregressive differencing, moving average, and backshift components, respectively.
- Y_t represents the time series data at time t.
- ε_t denotes the white noise error term.

Real life application

One example of how SARIMA might be put to use in the real world is in the process of predicting quarterly sales data for a retail organization. The sales data frequently display seasonal patterns because of things like the different holiday seasons and different promotional periods. The company is able to examine previous sales data, recognize seasonal patterns, and make more accurate projections of future sales by using a model called SARIMA.

Merits and Demerits:

- When applied to time series data, SARIMA models are able to distinguish between seasonal and non-seasonal patterns.
- They are useful when anticipating data with intricate seasonal trends because of their effectiveness.
- The SARIMA models can be altered to accommodate a wide variety of seasonal data types, which lends them flexibility and adaptability.

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

- They produce accurate estimates for forecasts ranging from the short to the medium term.
- SARIMA models can be complicated, particularly when dealing with a number of different seasonal components, which calls for a substantial amount of computational resources.
- Due to the complexity of the mathematical formulas, interpretation of the SARIMA results may be difficult for individuals who are not experts in the field.
- For SARIMA models to generate reliable forecasts, a significant quantity of historical data is necessary; however, this data may not always be accessible for all forms of data.

Preparation of Data:

- Prepare the time series data for analysis by collecting and cleaning it such that it is consistent and has no outliers or missing values.
- Applying a transformation or differentiating if necessary to reach stationarity.

Identification of Models:

- Determine the values of the AR and MA parameters during the season and the offseason by analyzing the ACF and PACF graphs.
- Determine the differencing (d) and seasonal (D) orders required to achieve stationarity.

Estimating Variables:

- Apply the SARIMA model's estimated parameters using estimation strategies like maximum likelihood.
- Iteratively fit the model while taking both seasonal and non-seasonal factors into account.

Model Evaluation and Adjustment:

- Examine diagnostic charts for evidence of residual randomness after a SARIMA model has been fitted to the data.
- Analyze the residuals using autocorrelation functions (ACF) plots, histograms, and the Ljung-Box test.

Analysis

ARIMA Models

In the analysis of the paddy production data from Tamil Nadu, several steps were undertaken to identify an appropriate time series model. The data was initially examined for stationarity through visual inspection of the plot and confirmed using the Auto correlation function (ACF) and Partial ACF (PADF) tests.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

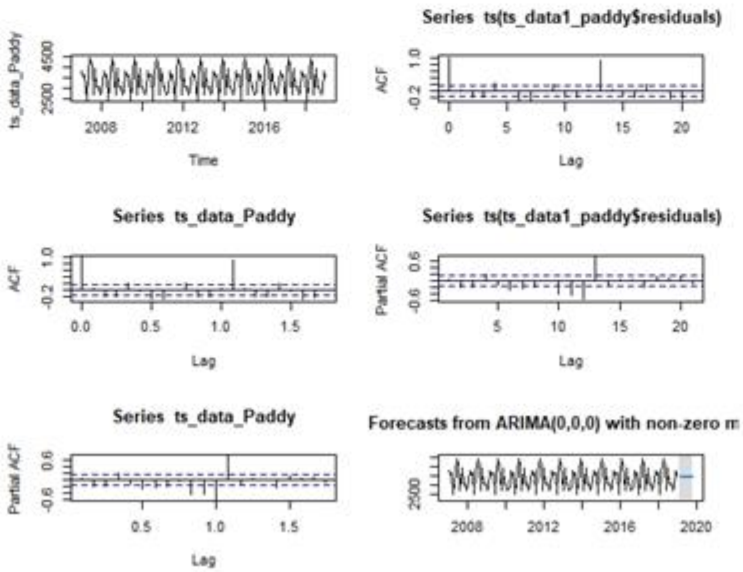
Following this, the `auto.arima` function was applied to determine the best-fitting model. The function iteratively evaluated various combinations of AR, MA, and differencing orders to select the model that exhibited the lowest information criterion values, signifying a good fit. This comprehensive process allowed for the identification of a suitable SARIMA model that can accurately capture the seasonal and non-seasonal patterns within the paddy production data of Tamil Nadu.

Models	Values
ARIMA (2,0,2) (1,0,1) [12] with non-zero mean	Inf
ARIMA (0,0,0) with non-zero mean	2267.311
ARIMA (1,0,0) (1,0,0) [12] with non-zero mean	2271.146
ARIMA (0,0,1) (0,0,1) [12] with non-zero mean	2268.206
ARIMA (0,0,0) with zero mean	2774.681
ARIMA (0,0,0) (1,0,0) [12] with non-zero mean	2269.197
ARIMA (0,0,0) (0,0,1) [12] with non-zero mean	2269.13
ARIMA (0,0,0) (1,0,1) [12] with non-zero mean	2270.753
ARIMA (1,0,0) with non-zero mean	2269.148
ARIMA (0,0,1) with non-zero mean	2269.047
ARIMA (1,0,1) with non-zero mean	2270.643

The ARIMA (0,0,0) model with a non-zero mean was chosen as the best fit based on the AIC values. Since the optimal model for predicting paddy production time series data in Tamil Nadu does not include differencing, autoregressive, or moving average terms, it follows that these methods should be avoided. The trend and seasonality of paddy output may be better predicted, allowing for more well-informed decisions to be made in agricultural planning and policy formulation in the region, if this model were analyzed in greater depth.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022



Coefficient	Values
Mean	3384.2897
S. E	49.2651
σ^2	354366
log likelihood	-1131.66
AIC	2267.31
AICc	2267.4
BIC	2273.26

Time series data for paddy production in Tamil Nadu were analyzed using the ARIMA(0,0,0) model with a non-zero mean. With a coefficient estimate of 3384.2897 and a standard error of 49.2651, the model produced a point estimate of the mean value. The σ^2 value for the model was found to be 354366. The model had a log likelihood of -1131.66, therefore the Akaike Information Criterion (AIC) was 2267.31, the AICc was 2267.4, and the Bayesian Information Criterion (BIC) was 2273.26. Insights into the statistical parameters and goodness of fit for the ARIMA(0,0,0) model are provided by this model's output, allowing for a deeper dive into the inner workings of paddy production in Tamil Nadu. More research is needed to improve predictions and direct productive agricultural policies and practices in the area.

Year	forecast	Lo 95	Hi 95
Feb 2019	3384.29	2217.549	4551.03
Mar 2019	3384.29	2217.549	4551.03
Apr 2019	3384.29	2217.549	4551.03
May 2019	3384.29	2217.549	4551.03

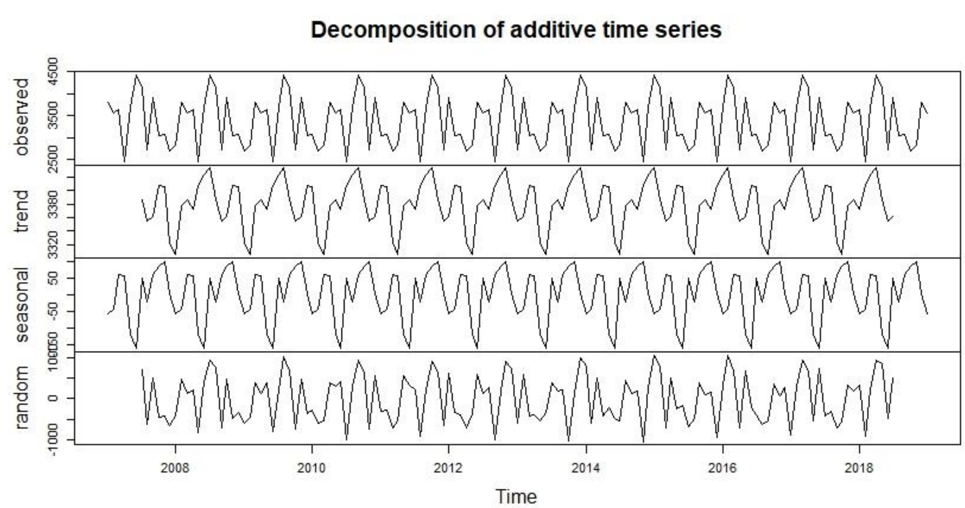
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

Jun 2019	3384.29	2217.549	4551.03
Jul 2019	3384.29	2217.549	4551.03
Aug 2019	3384.29	2217.549	4551.03
Sep 2019	3384.29	2217.549	4551.03
Oct 2019	3384.29	2217.549	4551.03

Point projections and 95% confidence intervals for paddy production in Tamil Nadu are available in the forecast_data_paddy for the time period of February 2019 through November 2019. With a lower 95% confidence interval of 2217.549 and an upper 95% confidence interval of 4551.03, the forecast indicates that the anticipated paddy production remains constant at 3384.29 for each month within the forecast period. The ARIMA(0,0,0) model predicts that paddy output will be steady, with no noticeable changes, throughout the designated forecast months. In order to make informed decisions on agricultural planning and policy in the region, more in-depth monitoring and research is necessary.

The residuals of the predicted paddy output statistics in Tamil Nadu were put through the Box-Ljung test. A 5-second time delay was used in the Ljung-Box test. With 5 degrees of freedom, the X-squared value is 20.254, which is statistically significant (p = 0.00112). The presence of autocorrelation in the residuals is strongly suggested by the low p-value, which is evidence against the null hypothesis of independence. The residuals' autocorrelation shows that the ARIMA(0,0,0) model may not accurately represent all the dynamics at play in the paddy production time series. Accurate and trustworthy forecasting is essential for strategic agricultural planning and decision making in the region, so understanding the autocorrelation structure is a top priority.



ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

SEASONAL ARIMA ANALYSIS

Time series data for paddy production was generated in R. The time series begins in 2007 and continues through 2019 at a rate of 1. The numbers 3809, 3562, 3630, 2463, 3687, 4429, 4123, 2712, 3918, 3039, 3070, 2682, and 2817 may be found in the data set.

Verification was performed to ensure that the time series data is indeed an object of class "ts," indicating that it is a time series. Paddy yields were plotted against time to show how they had changed during the selected period.

In addition, the stationarity of the time series data was evaluated using the Augmented Dickey-Fuller (ADF) test. The ADF test yielded a p-value of 0.6578 and a Dickey-Fuller statistic of -1.7758 for a lag order of 2. There is insufficient evidence to reject the null hypothesis of non-stationarity because the p-value is greater than the significance level of 0.05. To provide precise modeling and forecasting of paddy output in the region, more research is needed to investigate the stationarity of the time series data.

Paddy production time series summary (ts_paddy) provides an overview of the statistical measures that define this data collection. Paddy production fell as low as 2463 over the stipulated time period, while the first quartile number was at 2817, marking the bottom of the middle 50 percent. The average is set at 3380, while the median is set at 3562. At 3809, the third quartile number marks the top end of the middle quartile. During the given time period, paddy production peaked at a value of $4,429. Insights into the mean and standard deviation of the paddy production dataset are provided by these summaries, which aid in drawing conclusions about the yield distribution and trends over the given time period.

The differenced logarithm of the ts_paddy dataset was subjected to an Augmented Dickey-Fuller (ADF) test to determine whether or not the paddy production time series data were stationary. The variance was reduced using logarithmic transformation, and stationarity was attained via differencing.

The Dickey-Fuller statistic for the ADF test was -4.6604, and the corresponding p-value was 0.01. There is strong evidence to reject the null hypothesis of non-stationarity in favor of the alternative hypothesis of stationarity, as the p-value is significantly lower than the specified significance level of 0.05. If the differenced logarithm of the paddy production time series is stationary, then the data points are independent of time and consistently exhibit statistical features. To better predict and analyze paddy production trends in Tamil Nadu, this transformation improves the data's appropriateness for analysis using time series models like ARIMA and SARIMA.

Coefficient	Values
σ^2	0.06275
log likelihood	-0.42
AIC	2.83
AICc	3.23
BIC	3.31

Using the auto.arima function, we can see that auto Arima Model Log represents the logarithm of the paddy production time series data. With a value of 1 for the 'd' parameter, which indicates consideration of the first difference, stationarity in the data was sought. The log-transformed paddy production data was fed into the auto.arima function, and the resulting output shows that an ARIMA model with a difference order of 1 was selected. The nomenclature for this model is ARIMA (0, 1, 0). The model had a log-likelihood of -0.42, and its estimated variance was 0.06275. The calculations yielded an Akaike Information Criterion (AIC) value of 2.83, an Akaike Information Criterion (AICc) value of 3.23, and a Bayesian Information Criterion (BIC) value of 3.31. This information is useful for assessing the validity of the selected ARIMA model for the log-transformed paddy production data, as it sheds light on the model's internal structure and goodness-of-fit. This model looks to be an excellent fit for the data, as seen by its simplicity and low variance estimate, allowing for more accurate estimates and a better understanding of the processes at play in Tamil Nadu's paddy production dynamics.

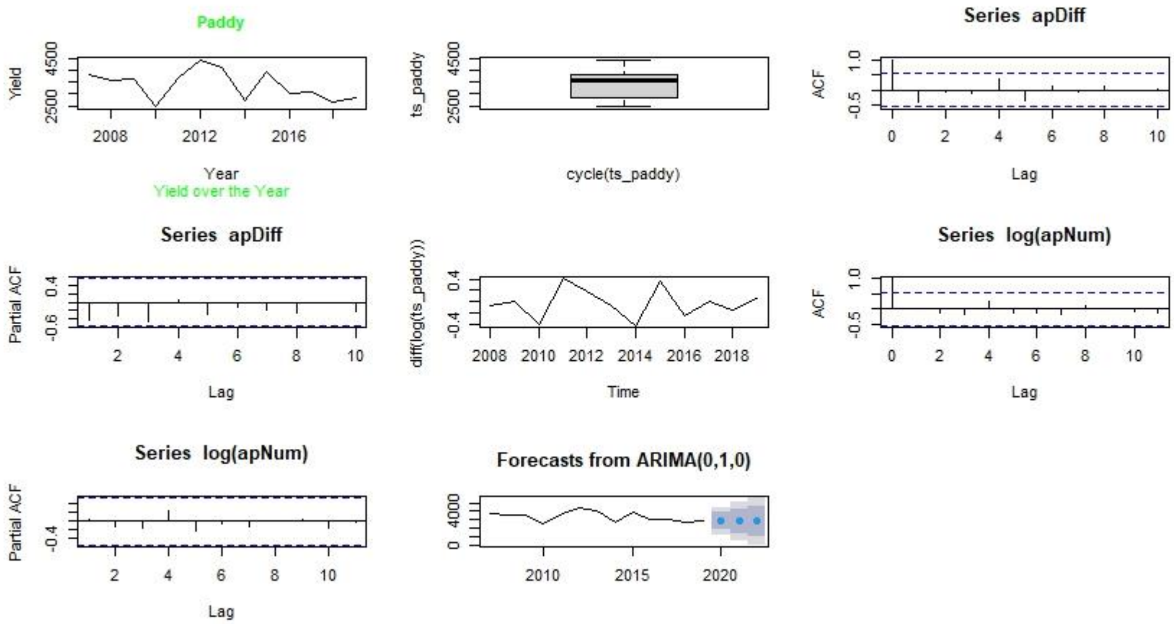
After applying the ARIMA model, auto Arima Model Log, to the log-transformed paddy production data, the residuals were analyzed using the Ljung-Box test. The aim of the analysis was to determine if the model residuals exhibited autocorrelation.

Coefficient	Values
x^2	3.3996
df	1
P-value	0.06521

The Ljung-Box test returns a significance level of 0.06521 for an X-squared value of 3.3996 with 1 degrees of freedom. There is insufficient evidence to conclude that the residuals exhibit significant autocorrelation, as the p-value is larger than.05. Because of the good fit between the data and the ARIMA (0,1,0) model, we may infer that the model appropriately explains the observed variability in the log-transformed paddy production data. It's possible that more research is needed to verify the model's accuracy and guarantee precise forecasting of paddy output patterns in Tamil Nadu.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022



CONCLUSION

The original time series data was analyzed using the ARIMA model, and an ARIMA (0,0,0) model was found to be the best appropriate for projecting paddy output in Tamil Nadu. Nonetheless, autocorrelation was detected in the model residuals via the Box-Ljung test, suggesting that the model may be inadequate in its attempt to capture all underlying patterns. To address this, we first converted the data using a logarithmic function, then differentiated it, before fitting the data with the ARIMA (0,1,0) model and observing a good fit with negligible residual autocorrelation.

When applied to the log-transformed paddy production data, the ARIMA (0,1,0) model revealed time-dependent patterns that clarified the dynamics. The model showed a good fit to the data and had a low variance estimate.

While the ARIMA models did provide some useful information, it may be necessary to take a broader approach, such as using the SARIMA model, in order to capture the probable seasonal fluctuations and increase the precision of future paddy production estimates. To better guide agricultural planning and policy-making in Tamil Nadu's paddy cultivation sector, the SARIMA model might be implemented to provide a more rigorous framework for understanding seasonal dynamics and increasing the precision of predictions.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 2, Feb 2022

To better capture the seasonal patterns and fluctuations in paddy production and to aid in the development of accurate forecasts and well-informed policy decisions for environmentally friendly farming in the region, more research and analysis using SARIMA modeling techniques are recommended.

REFERENCES

1. Ansari, M. I., & Ahmed, S. M. (2001). Time series analysis of tea prices: An application of ARIMA modelling and cointegration analysis. The Indian Economic Journal, 48(3), 49-54.
2. Bhagat, A., & Jadhav, D. (2021). A Study on Growth, Instability and Forecasting of Grape Export from India. Journal of Scientific Research, *65*(9), 1-6.
3. Darekar, A., & Reddy, A. (2017). Forecasting of common paddy prices in India. Journal of Rice Research, 10(1), 71-75.
4. Hemavathi, M., & Prabakaran, K. (2017). A statistical study on weather parameters relationship with rice crop yield in Thanjavur district of Tamil Nadu. International Journal of Agricultural Science and Research, 7(5), 25-32.
5. Hemavathi, M., & Prabakaran, K. (2018). ARIMA model for forecasting of area, production and productivity of rice and its growth status in Thanjavur District of Tamil Nadu, India. Int. J. Curr. Microbiol. App. Sci, 7(2), 149-156.
6. http://www.tnagriculture.in/dashboard/report/05_01.pdf
7. Joshua, V., Priyadharson, S. M., & Kannadasan, R. (2021). Exploration of machine learning approaches for paddy yield prediction in eastern part of Tamilnadu. Agronomy, 11(10), 2068.
8. Kathayat, B., & Dixit, A. K. (2021). Paddy price forecasting in India using ARIMA model. Journal of Crop and Weed, 17(1), 48-55.
9. Majid, R., & Mir, S. A. (2018). Advances in statistical forecasting methods: An overview. Economic Affairs, *63*(4), 815-831.
10. Raghavender, M. (2009). Forecasting paddy yield in Andhra Pradesh using season time series model. Bulletin of Pure & Applied Sciences-Mathematics, 28(1), 55-55.
11. Rajarathinam, A., & Thirunavukkarasu, M. (2013). Fuzzy Time Series Modeling for Paddy (Oryza sativa L.) Crop Production.
12. Saranyadevi, M., & Mohideen, A. K. (2017). Stochastic modeling for paddy production in Tamilnadu. International Journal of Statistics and Applied Mathematics, 2(5), 14-21.
13. Selvi, R. P. (2021). Chapter-1 Mathematical Model for Forecasting Paddy Price Based on Market Value in Tuticorin District. MULTIDISCIPLINARY, 1.
14. Vinoth, B., Rajarathian, A., & Manju Bargavi, S. K. (2016). Nonlinear regression and artificial neural network-based model for forecasting Paddy (Oryza sativa) production in Tamil Nadu. IOSR Journal of Mobile Computing & Application (IOSR-JMCA), 3.