

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

NAAS Rating (2026): 4.49

Maths 2026; 1(03): 136-143

© 2026 Stats & Maths

<https://www.mathsjournal.com>

Received: 02-01-2026

Accepted: 25-02-2026

Vankudoth Kumar

Ph.D. Scholar, Agricultural
Statistics, Division of
Agricultural Statistics, ICAR-
Indian Agricultural Statistical
Research Institute, New Delhi,
Delhi, India

A S Kamble

Assistant Professor, Department
of Agricultural Statistics, College
of Agriculture (COS),
Chintamani, University of
Agricultural Sciences (UAS),
Bengaluru, Karnataka, India

S N Manjunatha

Ph.D. Scholar, Department of
Horticulture, UAS, GKVK,
Bengaluru, Karnataka, India

Veena B

Assistant Professor, Department
of Food Technologist, ICAR-
AICRP on PHET, UAS, GKVK,
Bengaluru, Karnataka, India

Ashok Dodamani

Assistant Professor, Department
of Agricultural Extension, UAS,
GKVK, Bengaluru, Karnataka,
India

Corresponding Author:

Vankudoth Kumar

Ph.D. Scholar, Agricultural
Statistics, Division of
Agricultural Statistics, ICAR-
Indian Agricultural Statistical
Research Institute, New Delhi,
Delhi, India

Modelling and forecasting of pulse crop area and production in selected districts of Telangana using generalized additive models

**Vankudoth Kumar, AS Kamble, SN Manjunatha, Veena B and Ashok
Dodamani**

DOI: <https://www.doi.org/10.22271/math.2026.v11.i3b.2292>

Abstract

A statistical modelling framework was developed using a 30-year time-series dataset (1993-2022) to analyse and forecast the area, production, and productivity of pulse crops in three districts of Telangana state, namely Adilabad, Mahbubnagar, and Nizamabad. Several competing models, including linear, quadratic, exponential, monomolecular, and Generalized Additive Models (GAM), were evaluated for forecasting performance. Model selection was based on diagnostic criteria such as the coefficient of determination (R^2), Root Mean Squared Error (RMSE), and Mean Absolute percentage Error (MAPE). The results consistently indicated that the Generalized Additive Model outperformed the other models across all districts and variables. The GAM effectively captured the non-linear trends in pulse crop cultivation, providing reliable forecasts of area, production, and productivity. The findings highlight the suitability of GAM for agricultural time-series forecasting and offer valuable insights into the dynamics of pulse crop cultivation in Telangana, emphasizing the importance of robust statistical models for informed planning and policy formulation.

Keywords: Generalized Additive Model (GAM), forecasting, pulse crops, R^2 , RMSE, MAPE, Telangana

Introduction

Rapid climatic changes, increasing population pressure, and shrinking arable land pose significant challenges to global food and nutritional security. In this context, food legumes play a crucial role due to their high protein content and ability to supplement cereal-based diets, particularly in developing countries such as India (Savadatti, 2017) ^[9]. Pulses not only contribute to nutritional security but also enhance soil fertility through biological nitrogen fixation, thereby supporting sustainable agricultural systems (Darekar and Reddy, 2017) ^[4]. India occupies a dominant position in global pulse production, accounting for nearly 90% of the world's red gram (pigeon pea) cultivation and approximately 75% of Bengal gram (chickpea) production. Major pulse-producing states in the country include Maharashtra, Uttar Pradesh, Rajasthan, Madhya Pradesh, Andhra Pradesh, and Telangana (Vishwajith *et al*, 2014) ^[12]. In Telangana, pulse crops occupy a substantial share of the cultivated area. During 2021-22, pulses accounted for about 16.1% of the total cropped area and nearly 31.6% of the area under food grains. The total area under pulse cultivation in the state was approximately 5.17 lakh hectares, with a production of 5.49 lakh tonnes. Among the pulses grown, Bengal gram, red gram, black gram, and green gram together contributed nearly 94.33% of the state's total pulse production.

Bengal gram is one of the most important pulse crops in Telangana, widely used in various forms such as whole grain and flour for the preparation of diverse food products. During 2021-22, Bengal gram was cultivated over 1.28 lakh hectares in the state, producing about 2.38 lakh tonnes with an average productivity of 1,668 kg ha⁻¹. Red gram also occupies a prominent position in pulse cultivation, covering an area of about 2.95 lakh hectares with a production of 1.75 lakh tonnes and an average productivity of 596 kg ha⁻¹ during the same period.

Accurate forecasting of agricultural variables such as area, production, and productivity is essential for effective planning and policy formulation (Rahman *et al*, 2013) [8]. Forecasting enables stakeholders to anticipate future trends based on historical data, thereby assisting farmers, researchers, and policymakers in making informed decisions related to resource allocation, crop planning, and market interventions (Mishra *et al*, 2021) [6]. Given the inherent variability in agricultural systems, both linear and non-linear statistical models are increasingly employed to capture complex trends in crop performance.

In this backdrop, the present study aims to analyse historical trends and forecast future patterns in the area, production, and productivity of Bengal gram and red gram in three major pulse-growing districts of Telangana, namely Adilabad, Mahbubnagar, and Nizamabad. The study employs a range of statistical models, including linear, quadratic, exponential, monomolecular, and Generalized Additive Models (GAM), to identify the most suitable approach for pulse crop forecasting. The findings of this study are expected to provide valuable insights for agricultural planning, policy formulation, and future research aimed at enhancing pulse production and ensuring sustainable agricultural growth in the region.

Materials and Methods

A. Data Collection

The present study is based on secondary time-series data on area, production, and productivity of Bengal gram and red gram for a period of 30 years (1993-2022). District-wise data for Adilabad, Mahbubnagar, and Nizamabad of Telangana state were collected from the Statistical Abstracts of Telangana, the Directorate of Economics and Statistics (DES), and the Indiatat database (www.indiatat.com).

B. Statistical Models

To analyse growth patterns and generate forecasts, both linear and non-linear statistical models were employed. The models considered in the study include linear, quadratic, exponential, monomolecular, and Generalized Additive Models (GAM). These models were fitted separately for area, production, and productivity of each crop and district.

C. Generalized Additive Model (GAM)

The Generalized Additive Model was fitted using the local scoring algorithm, which iteratively estimates weighted additive models through the backfitting procedure. The backfitting algorithm is a Gauss-Seidel method that fits additive components by repeatedly smoothing partial residuals. This approach separates the parametric and non-parametric components of the model and estimates the parametric part using weighted least squares within the iterative framework (Ahmed *et al*, 2020) [1].

A generalized additive model can be expressed as a sum of smooth functions of predictor variables (Hastie and Tibshirani, 1987) and is represented as:

$$Y_t = f(t) + \varepsilon$$

Where,

Y_t = represents area or production in time period t , $f(t) = \alpha_0 + b_t(t)\alpha$, represents the smooth function of the time t , where α is some parameter and $b_t(t)$ are the basis function. The following are the three nonlinear growth curves, which were used to describe the growth of present time-series.

D. Model Adequacy and Diagnostic Criteria

Model performance was evaluated using multiple diagnostic statistics to ensure robustness and reliability of forecasts.

E. Coefficient of Determination (R^2)

The coefficient of determination (R^2) measures the proportion of variability in the dependent variable explained by the fitted model and serves as an indicator of goodness of fit (Marcillo *et al*, 2021) [5]. It is defined as:

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$$

Where,

ESS = Error sum of squares, RSS = Regression sum of squares and TSS = Total sum of squares

The value of R^2 ranges between 0 and 1, with values closer to 1 indicating a better model fit. Adjusted R^2 was also considered to account for model complexity.

F. Mean Absolute % age Error (MAPE)

Mean Absolute %age Error (MAPE) is a widely used, scale-independent measure for evaluating forecast accuracy. It expresses prediction errors as %ages, allowing comparison across different series. MAPE is given by:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100$$

Where, Y_i = Actual values, $\hat{Y}_i$ = Predicted values and n = number of observations

The model with the lowest MAPE value was considered superior (Chattopadhyay and Mitra, 2018) [2].

G. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) measures the average magnitude of prediction errors and reflects the model's ability to capture variability in the data (Sullivan and Hileman, 2020). It is computed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_t - \hat{Y}_t)^2}{n}}$$

Where,

Y_t = Observed value,

$\hat{Y}_t$ = Predicted value and n = Number of observations

Lower RMSE values indicate better predictive performance.

H. Model Selection and Forecasting

The best-fitting model for each crop, district, and variable was selected based on the combined performance of R^2 , RMSE, and MAPE values. The selected models were then used to generate forecasts for area, production, and productivity up to the year 2027.

Results and Discussion

Bengal Gram Area and Production

Adilabad District

Among the competing models, the Generalized Additive Model (GAM) emerged as the best-performing model for forecasting both area and production of Bengal gram in Adilabad district, as evidenced by the highest R^2 values and the lowest RMSE, MSE, and MAPE values. The superior

performance of GAM indicates its ability to capture non-linear temporal trends present in the historical data (Solanki *et al*, 2017) ^[10]. Based on GAM projections, the area under Bengal gram is expected to increase steadily from 32.11 thousand hectares in 2023 to 36.09 thousand hectares by 2027. Similarly, production is projected to rise from 21.90 thousand tonnes to 25.79 thousand tonnes during the same period. The upward trend suggests favourable conditions and increasing adoption of Bengal gram cultivation in the district. The observed and forecasted trends further demonstrate the smooth growth pattern captured by the GAM, reinforcing its

suitability for time-series forecasting of pulse crops.

Table 1: Forecasted values of Bengal gram area and production of Adilabad district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	32.109	21.902
2024	33.105	22.874
2025	34.101	23.846
2026	35.097	24.818
2027	36.093	25.791

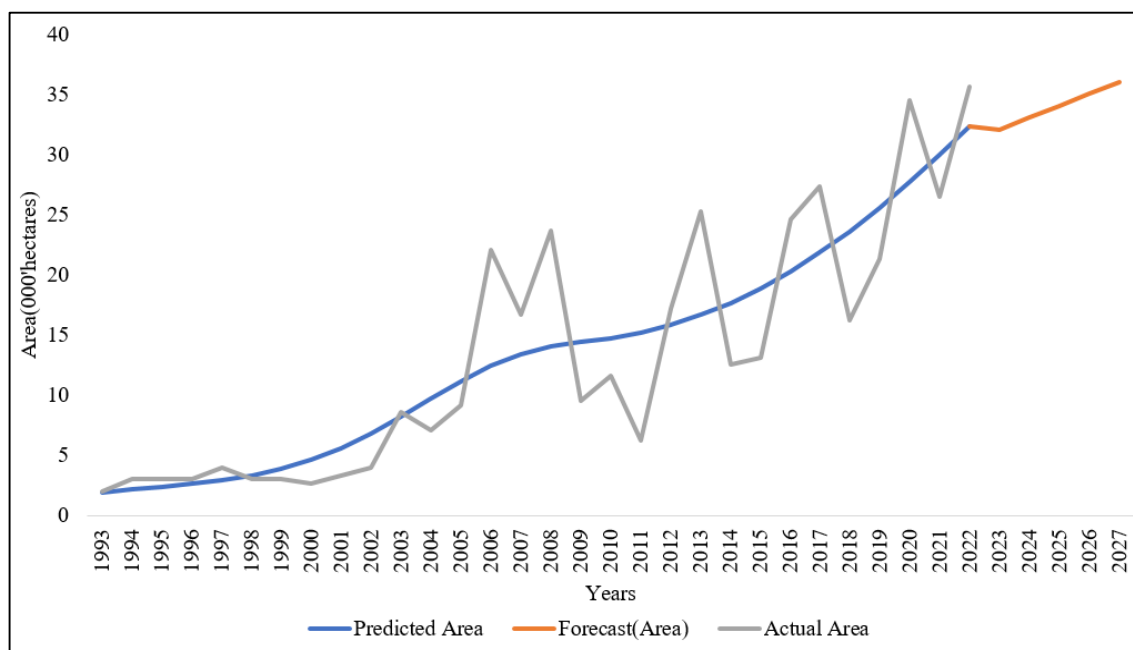


Fig 1: Observed and forecasted trends of Bengal gram area in Adilabad district

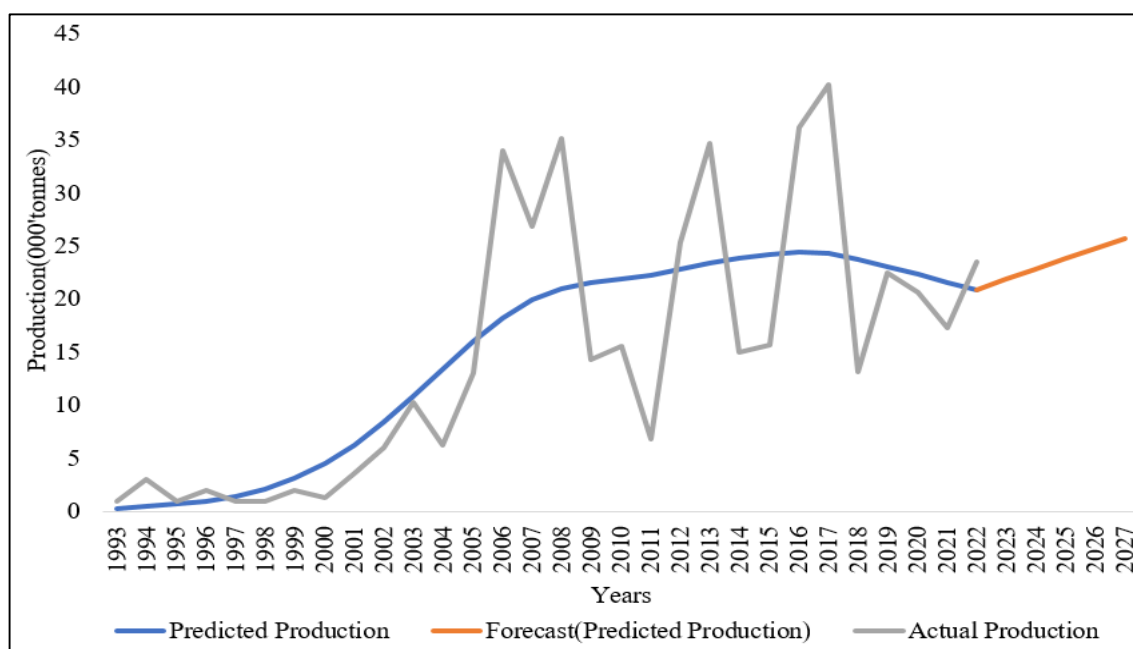


Fig 2: Observed and forecasted trends of Bengal gram production in Adilabad district

Mahbubnagar District

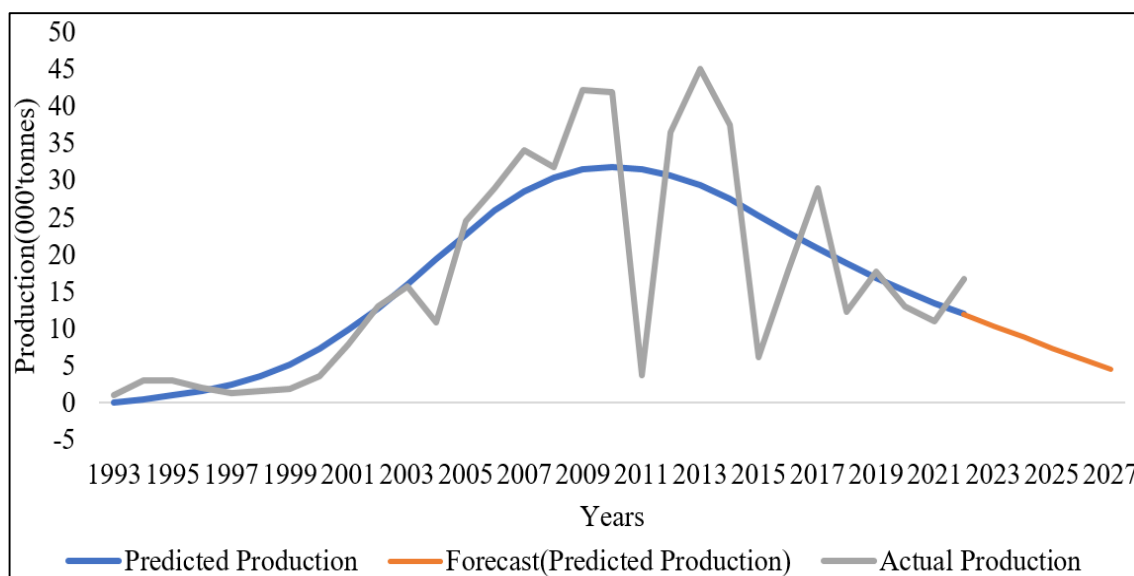
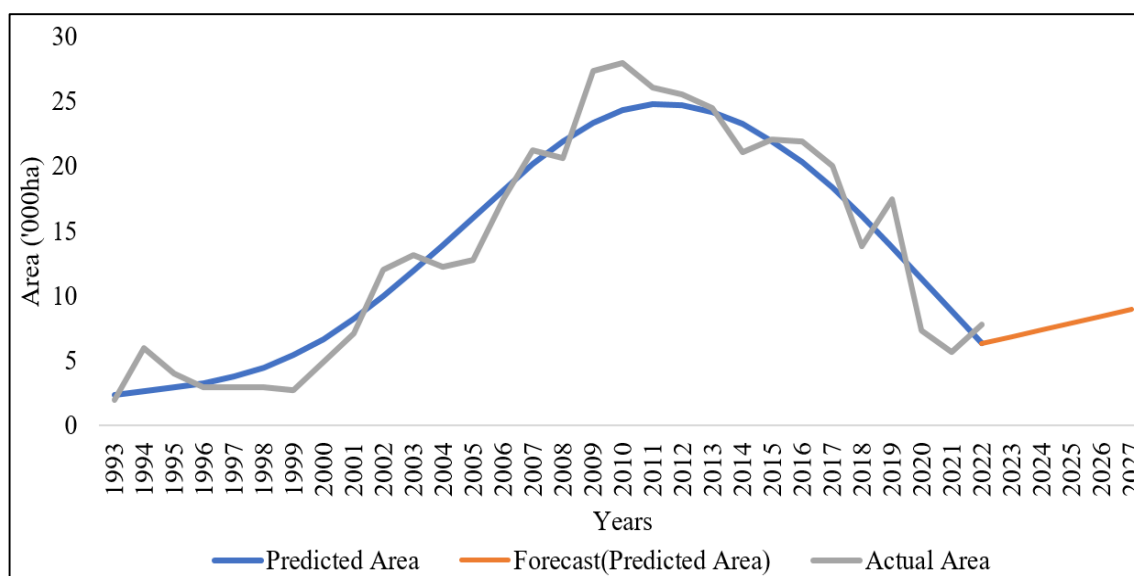
In Mahbubnagar district, GAM was identified as the most appropriate model for forecasting Bengal gram area and production due to its superior diagnostic performance, similar to the results of Chen *et al*, 2019 ^[3]. The projected area under Bengal gram is expected to increase from 6.88 thousand hectares in 2023 to 8.97 thousand hectares by 2027.

Correspondingly, production is forecasted to rise from 131.09 thousand tonnes to 146.23 thousand tonnes.

The increasing trend reflects gradual expansion of Bengal gram cultivation in the district, possibly driven by improved agronomic practices and market demand. The GAM's flexibility allowed it to effectively model the underlying non-linear growth pattern observed in the historical data.

Table 2: Forecasted values of Bengal gram area and production of Mahbubnagar district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	6.882	131.09
2024	7.403	134.26
2025	7.925	142.34
2026	8.446	138.92
2027	8.967	146.23

**Fig 3:** Observed and forecasted trends of Bengal gram area in Mahbubnagar district**Fig 4:** Observed and forecasted trends of Bengal gram production in Mahbubnagar district

Nizamabad District

For Nizamabad district, GAM again outperformed other models in forecasting Bengal gram area and production, similar to the results of Rahman and Imam, 2008. The area under Bengal gram is projected to increase from 8.80 thousand hectares in 2023 to 11.28 thousand hectares by

2027. Production is expected to rise correspondingly from 16.79 thousand tonnes to 21.65 thousand tonnes. The results indicate a consistent upward trend, suggesting increasing importance of Bengal gram cultivation in the district. The GAM-based forecasts highlight gradual but sustained growth, aligning with observed trends during the study period.

Table 3: Forecasted values of Bengal gram area and production of Nizamabad district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	8.795	16.789
2024	9.417	18.005
2025	10.039	19.221
2026	10.661	20.437
2027	11.283	21.653

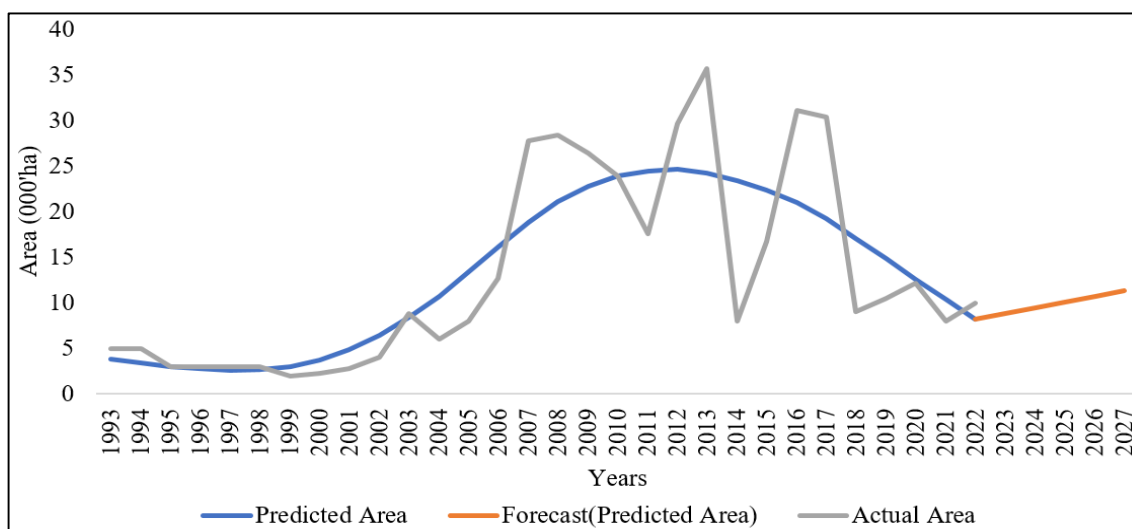


Fig 5: Observed and forecasted trends of Bengal gram area in Nizamabad district

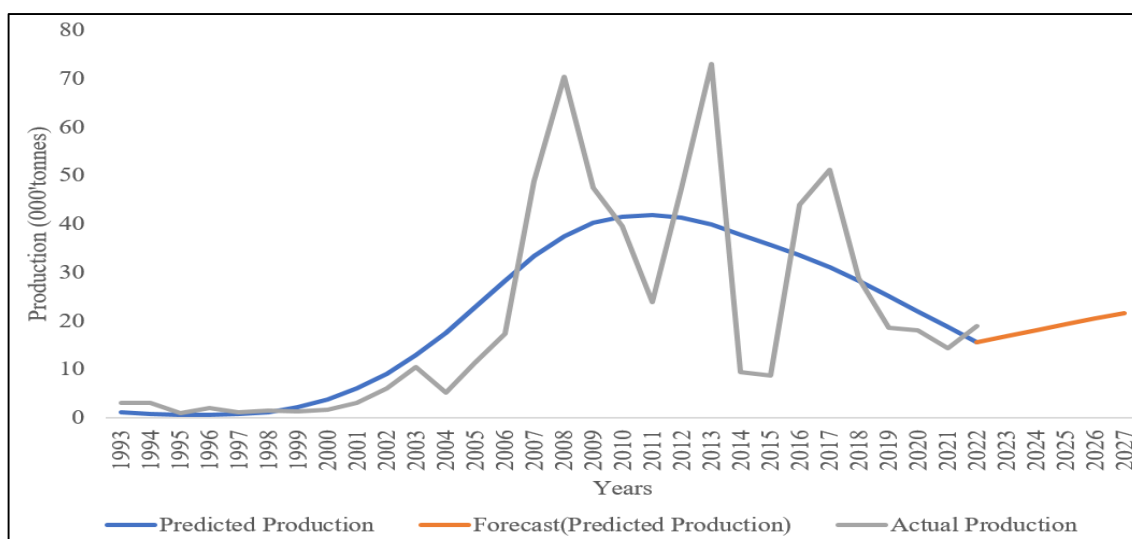


Fig 6: Observed and forecasted trends of Bengal gram production in Nizamabad district

Red Gram Adilabad District

The GAM model was found to be the best-fitting model for forecasting red gram area and production in Adilabad district. Forecasts indicate a marginal but steady increase in area from 44.38 thousand hectares in 2023 to 44.49 thousand hectares by 2027. Production is projected to rise from 42.21 thousand tonnes to 47.69 thousand tonnes over the same period. Although the expansion in cultivated area is relatively modest, the increase in production suggests improvements in productivity or crop management practices. The GAM effectively captured these subtle trends, which were not adequately represented by linear or other non-linear models.

Mahbubnagar District

In Mahbubnagar district, GAM demonstrated the highest forecasting accuracy for red gram. The area under red gram is projected to increase substantially from 146.34 thousand hectares in 2023 to 162.18 thousand hectares by 2027.

Production is expected to rise from 69.26 thousand tonnes to 72.75 thousand tonnes during the forecast period.

The pronounced upward trend reflects the district's growing reliance on red gram cultivation. The flexibility of GAM allowed it to model the accelerating growth pattern effectively, making it a robust tool for medium-term agricultural forecasting.

Nizamabad District

For Nizamabad district, the GAM model again outperformed other models for red gram forecasting. The projected area under red gram is expected to increase from 4.42 thousand hectares in 2023 to 4.81 thousand hectares by 2027. Production is forecasted to rise from 3.89 thousand tonnes to 4.45 thousand tonnes. Although the scale of cultivation is comparatively smaller, the consistent upward trend highlights the gradual expansion of red gram in the district. The GAM successfully captured these incremental changes, demonstrating its suitability even for districts with relatively smaller cultivated areas.

Table 4: Forecasted values of Red gram area and production of Adilabad district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	44.381	42.21
2024	44.411	43.58
2025	44.438	44.95
2026	44.466	46.23
2027	44.494	47.69

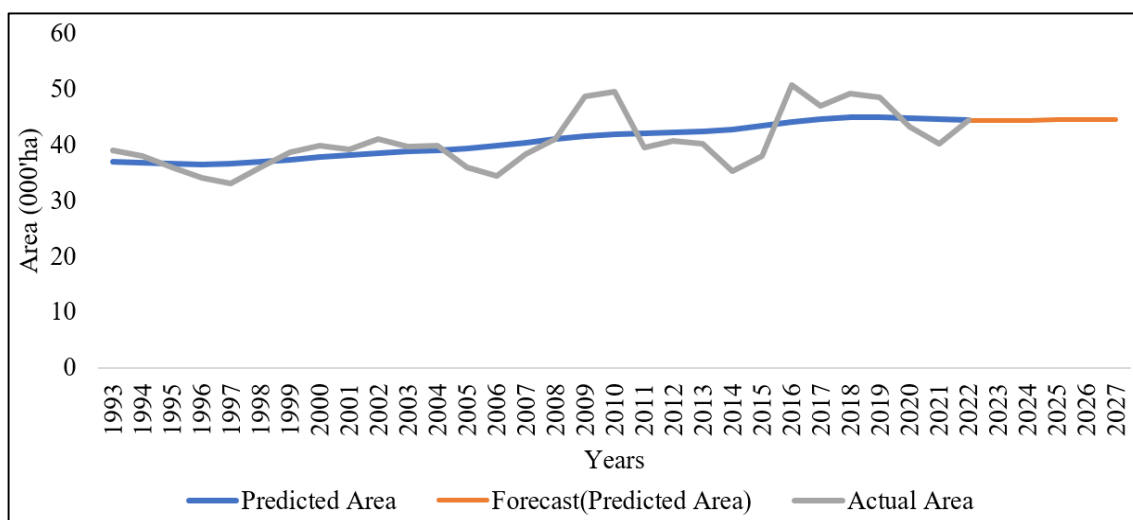


Fig 7: Observed and forecasted trends of Red gram area in Adilabad district

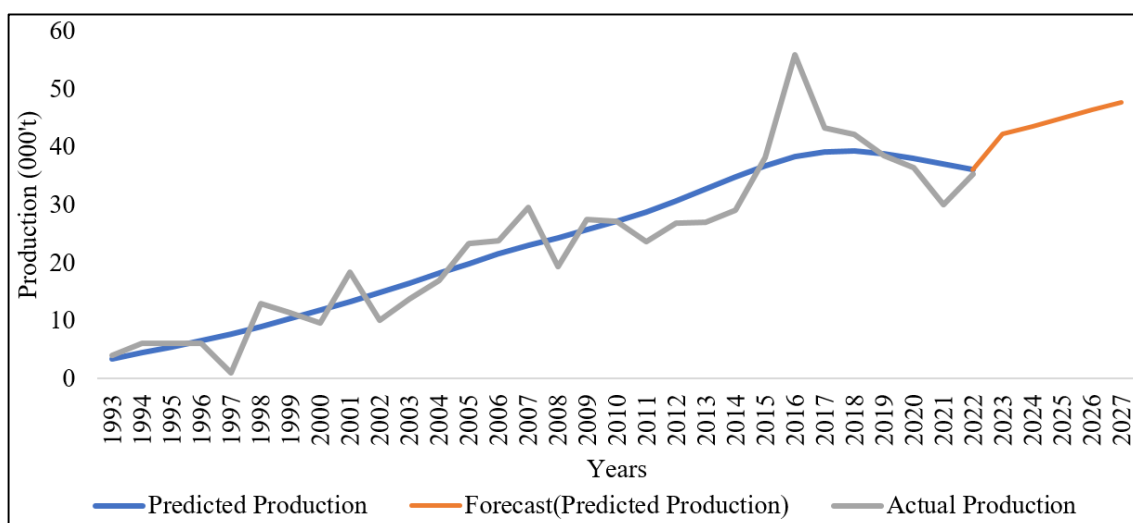


Fig 8: Observed and forecasted trends of Red gram production in Adilabad district

Table 5: Forecasted values of red gram area and production of Mahbubnagar district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	8.795	16.789
2024	9.417	18.005
2025	10.039	19.221
2026	10.661	20.437
2027	11.283	21.653

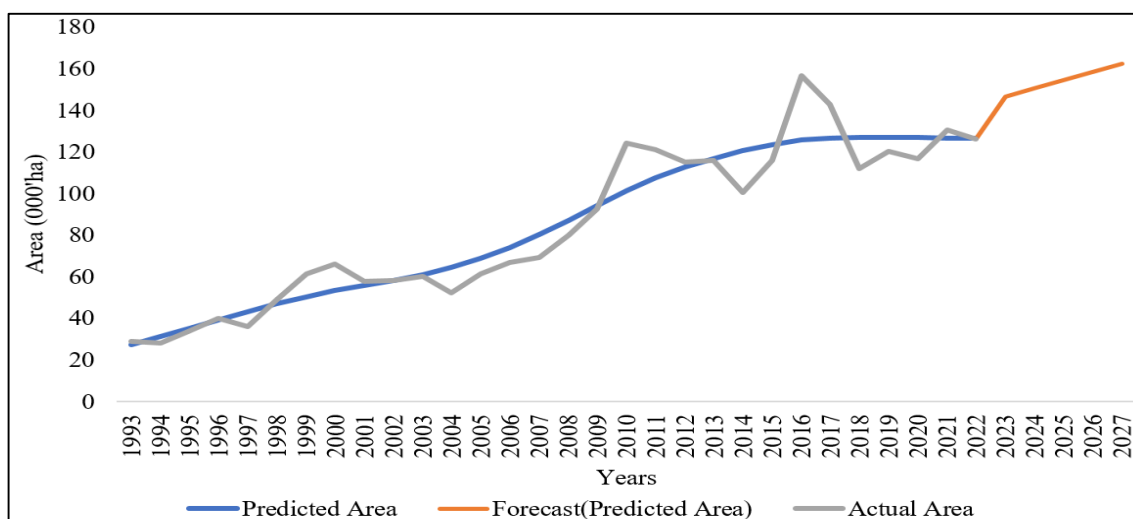


Fig 9: Observed and forecasted trends of Red gram area in Mahbubnagar district

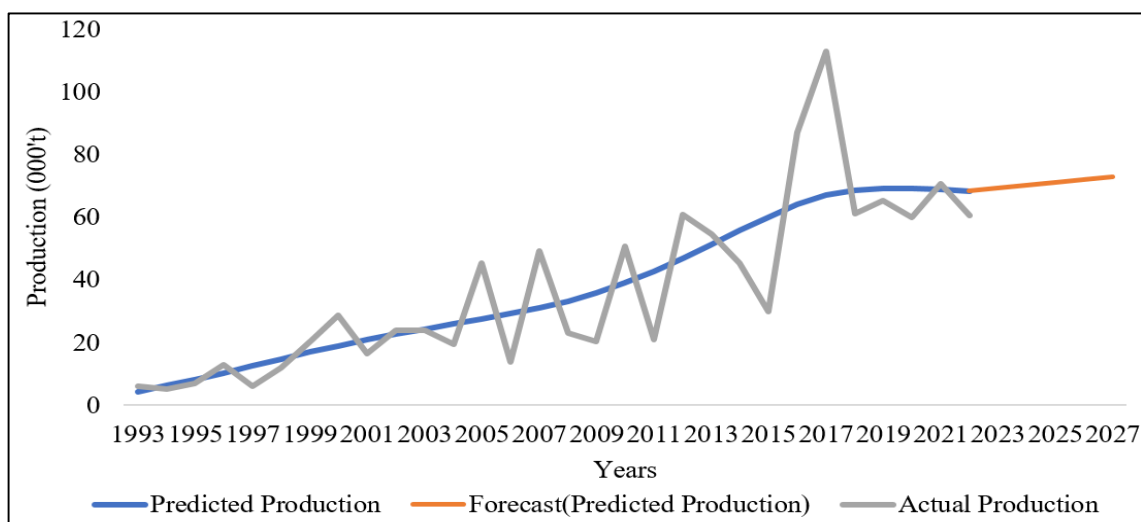


Fig 10: Observed and forecasted trends of red gram production in Mahbubnagar district.

Table 6: Forecasted values of red gram area and production of Nizamabad district by GAM model

Year	Forecasted Area ('000 hectares)	Forecasted Production ('000 tonnes)
2023	4.421	3.892
2024	4.519	4.031
2025	4.616	4.171
2026	4.714	4.308
2027	4.811	4.447

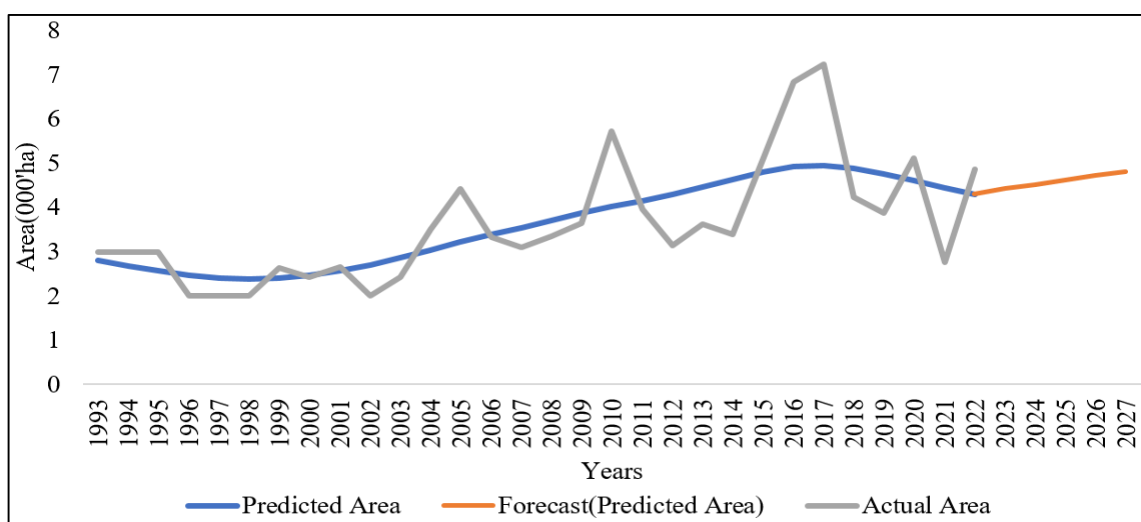


Fig 11: Observed and forecasted trends of Red gram area in Nizamabad district

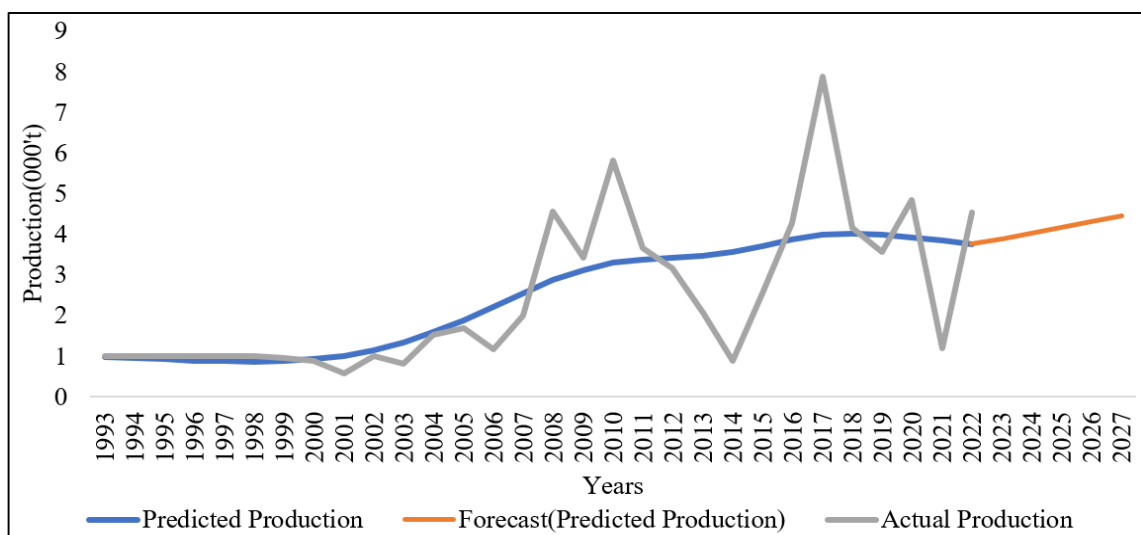


Fig 12: Observed and forecasted trends of Red gram production in Nizamabad district.

Overall Discussion

Across all districts and crops, the Generalized Additive Model consistently outperformed linear, quadratic, exponential, and monomolecular models. Its superior performance can be attributed to its ability to model complex, non-linear relationships without imposing rigid functional forms. The findings align with earlier studies that reported the effectiveness of GAM in agricultural and environmental forecasting.

The increasing trends observed for both Bengal gram and red gram across the three districts underscore the growing importance of pulse crops in Telangana's agricultural landscape. Reliable forecasts generated through GAM can support informed decision-making related to crop planning, resource allocation, and policy formulation.

Summary

The present study examined the temporal trends and forecasted the area and production of two major pulse crops, Bengal gram and red gram, across three important pulse-growing districts of Telangana state—Adilabad, Mahbubnagar, and Nizamabad. A 30-year district-level time-series dataset (1993-2022), obtained from official sources including the Directorate of Economics and Statistics, Telangana, and Indiatat, was analysed using various statistical growth models.

Both linear and non-linear models, namely linear, quadratic, exponential, monomolecular, and Generalized Additive Models (GAM), were fitted separately for area and production of each crop and district. Model performance was evaluated using diagnostic criteria such as the coefficient of determination (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Based on these criteria, the GAM consistently emerged as the most appropriate model across crops and districts.

Using the best-fitting GAM, forecasts were generated up to the year 2027. In Adilabad district, Bengal gram area and production are projected to reach 36.09 thousand hectares and 25.79 thousand tonnes, respectively. In Mahbubnagar district, the corresponding forecasts indicate an area of 8.96 thousand hectares and a production of 4.41 thousand tonnes by 2027. In Nizamabad district, Bengal gram area and production are expected to reach 11.28 thousand hectares and 21.65 thousand tonnes, respectively.

For red gram, GAM-based projections indicate that in Adilabad district, the area and production are expected to reach 44.49 thousand hectares and 47.69 thousand tonnes by 2027. In Mahbubnagar district, red gram is projected to cover 162.18 thousand hectares with a production of 72.75 thousand tonnes. In Nizamabad district, the projected area and production of red gram are 4.81 thousand hectares and 4.45 thousand tonnes, respectively.

Conclusion

The study demonstrates that the Generalized Additive Model is a robust and reliable approach for modelling and forecasting pulse crop area and production under non-linear and dynamic agricultural conditions. Compared to traditional linear and parametric non-linear models, GAM exhibited superior performance in capturing temporal variability, as reflected by higher R^2 values and lower RMSE, MSE, and MAPE values.

The forecast results indicate a consistent upward trend in both Bengal gram and red gram cultivation across all three districts

of Telangana, highlighting the increasing importance of pulse crops in the state's agricultural system. These projections provide valuable insights for policymakers, extension agencies, and researchers in planning future crop strategies, resource allocation, and market interventions.

Overall, the study underscores the importance of adopting flexible and data-driven modelling techniques such as GAM for agricultural forecasting. The findings can serve as a useful reference for future research aimed at enhancing pulse productivity and ensuring sustainable agricultural development in Telangana and similar agro-climatic regions.

References

1. Ahmed M, Ahmad S, Hayat R, Raza MA. Application of generalized additive model for rainfall forecasting in rainfed Pothwar, Pakistan. *Systems Modeling*. Singapore. 2020;1-10.
2. Chattopadhyay M, Mitra SK. Assessing the predictability of different kinds of models in estimating impacts of climatic factors on food grain availability in India. *Opsearch*. 2018;55(1):50-64.
3. Chen K, O'Leary RA, Evans FH. A simple and parsimonious generalized additive model for predicting wheat yield in a decision support tool. *Agric Syst*. 2019;173:140-150.
4. Darekar A, Reddy AA. Price forecasting of pulses: case of pigeonpea. *J Food Legumes*. 2017;30(3):212-216.
5. Marcillo GS, Martin NF, Diers BW, Da Fonseca Santos M, Leles EP, Chigeza G, Francischini JH. Implementation of a generalized additive model (GAM) for soybean maturity prediction in African environments. *Agronomy*. 2021;11(6):1043.
6. Mishra P, Yonar A, Yonar H, Kumari B, Abotaleb M, Das SS, Patil SG. State of the art in total pulse production in major states of India using ARIMA techniques. *Curr Res Food Sci*. 2021;4:800-806.
7. Rahman NMF, Imam MF. Growth, instability and forecasting of pigeon pea, chickpea and field pea pulse production in Bangladesh. *Bangladesh J Agric Econ*. 2008;31:81-95.
8. Rahman NMF, Aziz MA, Rahman MM, Mohammad N. Modeling on grass pea and mung bean pulse production in Bangladesh using ARIMA model. *IOSR J Agric Vet Sci*. 2013;6(1):20-31.
9. Savadatti PM. Trend and forecasting analysis of area, production and productivity of total pulses in India. *Indian J Econ Dev*. 2017;5(12):1-10.
10. Solanki HU, Bhatpuria D, Chauhan P. Applications of generalized additive model (GAM) to satellite-derived variables and fishery data for prediction of fishery resources distributions in the Arabian Sea. *Geocarto Int*. 2017;32(1):30-43.
11. Sullivan RM, Hileman JP. Time series modeling and forecasting of a highly regulated riverine system: implications for fisheries management. *Calif Fish Wildl*. 2020;106:221-259.
12. Vishwajith KP, Dhekale BS, Sahu PK, Mishra P, Noman MD. Time series modeling and forecasting of pulses production in India. *J Crop Weed*. 2014;10(2):147-154.