P-ISSN: 2618-0723
E-ISSN: 2618-0731

NAAS Rating (2025): 5.04
www.extensionjournal.com

International Journal of Agriculture Extension and Social Development

Volume 8; Issue 9; September 2025; Page No. 553-556

Received: 07-07-2025
Accepted: 09-08-2025

Indexed Journal
Peer Reviewed Journal

Determinants of agricultural development in Maharashtra

[1]Hake AD, [2]Thorat VA and [3]Borate HV

[1]Junior Research Assistant, Department of Agriculture Economics, Dr. Balasaheb Sawant Konkan Krishi Vidyapeeth, Dapoli, Maharashtra, India

[2]Ex. Professor (CAS), Department of Agriculture Economics, Dr. Balasaheb Sawant Konkan Krishi Vidyapeeth, Dapoli, Maharashtra, India

[3]Professor (CAS), Department of Extension Education, Dr. Balasaheb Sawant Konkan Krishi Vidyapeeth, Dapoli, Maharashtra, India

DOI: https://www.doi.org/10.33545/26180723.2025.v8.i9h.2460

Corresponding Author: Hake AD

Abstract

Agriculture has historically played a pivotal role in the economy of Maharashtra; however, its relative contribution has gradually declined over time. This decline is indicative of major structural changes within the state's economy, as reflected in the decreasing share of agriculture in the Gross State Domestic Product (GSDP), signaling a gradual transition away from an agrarian base. The regional dimensions of this transformation, as well as its long-term sustainability, necessitate a systematic exploration of sectoral performance and corresponding public policy interventions. Given the growing consensus in India on the importance of micro-level planning, it becomes both relevant and insightful to assess development at the divisional level. Such an assessment enables the identification of inter-divisional disparities and provides a comparative perspective on the relative standing of each division.

Consequently, research in this area is essential for quantifying the status of agricultural development across divisions, considering multiple sectors as well as aggregate development. Empirical analysis of the determinants of development indicates that four key sectors animal husbandry, transportation, area under crops, and input supply have made a positive contribution to agricultural growth, whereas the remaining sectors exhibited either negligible or negative contributions. Since agricultural development is inherently a multidimensional process, coordinated and efficient management of all relevant sectors is crucial to minimizing inter-divisional disparities and fostering balanced agricultural and overall economic development in the state.

Keywords: Determinants, lasso regression, spearman correlation, shrinking coefficient

Introduction

Agricultural development constitutes an integral component of overall economic development in any country. It not only ensures the provision of food for human consumption and raw materials for industries but also generates employment for a substantial proportion of the population. Economic development, in turn, is a process through which a nation enhances the efficiency of its resources to achieve a sustained increase in total income and per capita output. This process enables citizens to access desired goods and services, thereby improving their standard of living and overall well-being. Development is inherently dynamic, involving continuous changes in strategies, structures, and procedures to achieve targeted outcomes.

The Indian economy comprises several critical sectors that collectively contribute to national income, among which agriculture has historically been the most significant. The prosperity of agriculture, therefore, plays a pivotal role in determining the general prosperity of the nation. Agricultural development in India is shaped by various interrelated factors such as population growth, land utilization patterns, irrigation infrastructure, animal husbandry, transportation, cropping patterns, input supply systems, and institutional support mechanisms.

Pattern of Agricultural Development

Agricultural development in India has been driven by the implementation of diverse projects and programs aimed at achieving food self-sufficiency and increasing production. The development strategy has emphasized the production of food grains to meet domestic demand and the cultivation of commercial crops such as cotton, sugarcane, oilseeds, tea, coffee, and rubber to supply raw materials for industries and to generate export earnings.

Between the 1950s and 1980s, agricultural growth was primarily fueled by an expansion in the gross cropped area. However, since the mid-1980s, a structural transformation in agricultural production has been observed. While the total cropped area registered a moderate decline, productivity per hectare increased significantly due to technological advancements. These advancements included the introduction of high-yielding and hybrid seed varieties, increased use of chemical fertilizers, adoption of modern cultivation and harvesting techniques, expansion of irrigation facilities, and the development of institutional mechanisms such as the Agricultural Prices Commission to ensure remunerative prices for farmers. Complementary reforms such as land redistribution, enhanced agricultural credit and insurance systems, increased investments in agricultural research and extension services, and improvements in rural infrastructure also contributed to this transformation.

Moreover, the production of livestock products including milk, meat, eggs, and related value-added products has expanded to meet domestic consumption needs and support export growth. The creation of irrigation potential for rainfed areas and water-intensive crops such as sugarcane, fruits, and vegetables has been a major focus. Additionally, electricity generation from hydro, thermal, and nuclear sources has been harnessed to support agricultural operations and agro-based industries such as textiles, sugar, and food processing, further strengthening the sector's role in economic development.

Methodology

Sources of data

The time series data on public expenditure, population, education, health, agriculture, transport, banking, animal husbandry industries input supply etc. were collected from the annual publication, department of Economics and Statistics, Government of Maharashtra as well as Maharashtra census for the period from 2000-01 to 2019-20.

Determinants of development

Initially it was proposed to employ structural equation model to identify the determinants of agricultural development. However, it was found that there was high multicollinearity among the variables included in the model as a result Structural Equation Model failed to estimate the coefficients. Hence lasso regression was tried.

1. Lasso Regression

Lasso (Least Absolute Shrinkage and Selection Operator) regression is a type of regularization technique, commonly used to reduce the risk of overfitting. It builds upon the concept of linear regression by introducing a regularization term into the standard regression equation. While linear regression aims to minimize the sum of squared differences between actual and predicted values by fitting a line (or, in higher dimensions, a plane or hyperplane) to the data, real-world datasets often face the problem of multicollinearity—where features are highly correlated with one another. In such cases, Lasso regression proves useful. By adding a penalty term to the model, regularization helps control complexity and prevents overfitting.

Bias-variance tradeoff in Lasso regression

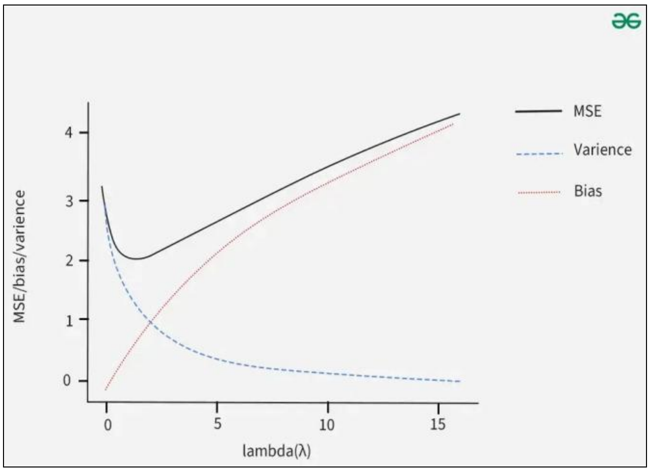
The balance between bias (errors caused by overly simplistic model assumptions) and variance (errors caused by excessive sensitivity to small fluctuations in training data) is referred to as the bias–variance tradeoff.

- In Lasso regression, the penalty term (L1 regularization) plays a key role in reducing model variance by shrinking the coefficients of less important features toward zero. This process helps limit overfitting, ensuring the model focuses less on noise in the training data and more on meaningful patterns.
- However, if the regularization strength is too high, the model can become overly simplified, leading to increased bias and a weaker ability to capture the true underlying relationships in the data.

Thus, like other regularization methods, Lasso regression involves a tradeoff between bias and variance. Achieving the optimal balance typically requires minimizing the overall prediction error (MSE), often through techniques such as cross-validation to tune the regularization parameter.

How does lasso regression works

Lasso regression can be seen as an extension of traditional linear regression. In linear regression, the objective is to find the best-fitting line by minimizing the sum of squared differences between observed and predicted values. However, this approach does not fully account for the complexity of real-world data, especially when dealing with a large number of features.



Ordinary list square (OLS)

Regression: Lasso regression is particularly useful in such cases because it introduces a penalty term while minimizing the sum of squared differences. This penalty is based on the absolute values of the predictors' coefficients.

The formula for OLS is:

Min RSS=$\sum(yi-yi)^2$

Where,

$Y_i$is the observed value.

Y_i is the predicted value for each data point i.

Penalty term for lasso regression

The Ordinary Least Squares (OLS) equation is modified by adding a penalty term. This term, known as L1 regularization, is calculated as the sum of the absolute values of the coefficients. The objective then becomes minimizing the combined value of this penalty term and the sum of squared differences.

RSS+$\lambda\times\sum|\beta i|$

B_i represents the coefficients of the predictors.

λ (lambda) acts as the tuning parameter that determines the strength of the penalty. A larger value of λ increases the regularization effect, causing more coefficients to shrink toward zero.

Shrinking coefficients
A unique feature of the penalty term in Lasso regression is its ability to shrink the coefficients of less important variables all the way to zero. When a coefficient becomes zero, the corresponding feature is effectively removed from the model, allowing Lasso to perform automatic variable selection. This property is especially valuable when working with high-dimensional datasets, where the number of predictors is large relative to the number of observations.
By reducing or eliminating the influence of unimportant predictors, Lasso regression creates simpler models that are less prone to overfitting. This not only enhances the model's interpretability but also improves its ability to generalize to new data.

Selecting the optimal λ
In Lasso regression, choosing the appropriate tuning parameter λ (lambda) is crucial. Cross-validation is often used to identify the optimal value of λ that balances predictive accuracy with model simplicity. The main goal of Lasso regression is to minimize the Residual Sum of Squares (RSS) while adding a penalty term proportional to the sum of the absolute values of the coefficients.

$$\hat{y}_i = W_0 + \sum_{j=1}^{m} X_{ij} W_j$$

$$j\,(W) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + a \sum_{j=1}^{m} |W_j|$$

$$||W||^2 = \sum_{j=1}^{m} |W_j|^2$$

In the plot, the Lasso regression equation combines the Residual Sum of Squares (RSS) with the L1 penalty applied to the coefficients βi\beta_iβi. The RSS measures the squared differences between the predicted values and the actual observed values.

L1 penalty
Lasso regression penalizes the absolute values of the coefficients, shrinking some of them to zero and thereby simplifying the model. The strength of this L1 penalty is controlled by the parameter λ (lambda). Larger values of λ impose stronger penalties, which can increase the RSS but also force more coefficients to zero, resulting in a sparser model.

When to use lasso regression
When dealing with high-dimensional datasets that include a large number of features—some of which may be redundant or irrelevant—Lasso regression proves especially useful. In addition, regression techniques can be applied in the following situations:

Features of selection
Lasso regression automatically performs feature selection by shrinking the coefficients of less important variables to zero. This is particularly useful when working with a large number of features and aiming to identify the most significant ones.

Collinearity: Lasso regression can also be useful in situations involving multicollinearity, where predictor variables are highly correlated with each other. By shrinking the coefficients of correlated variables and selecting only one of them, Lasso helps reduce redundancy and simplifies the model.

Regularization: By penalizing large coefficients, Lasso regression helps prevent overfitting. This becomes especially important in cases where the number of predictors is close to or even exceeds the number of observations.

Interpretability: Compared to traditional linear regression models that include all features, Lasso regression often produces sparse models with fewer non-zero coefficients. This results in a simpler and more interpretable final model.

Results and Discussion
1. Lasso Regression Analysis
Lasso (Least Absolute Shrinkage and Selection Operator) regression was employed to estimate the regression model due to the presence of multicollinearity among the explanatory variables. Lasso regression is particularly suitable for such cases as it performs both variable selection and regularization, thereby improving the prediction accuracy and interpretability of the model.
Lambda (λ) is a key tuning parameter in lasso regression, controlling the degree of shrinkage applied to the regression coefficients. A higher value of λ results in greater shrinkage, potentially reducing some coefficients to zero, thus excluding the corresponding predictors from the model. The optimal value of λ was determined using k-fold cross-validation, which minimizes the mean squared error (MSE) on the validation dataset.
Table 1 presents the results of the cross-validation procedure, where the minimum MSE was observed at λ = 0.01681. This value was therefore selected as the best-fit parameter for the model, as it minimizes prediction error and balances model complexity with predictive performance.

Table 1: Cross-Validation Results for Lasso Regression

Lambda	Index	Measure (MSE)	SE	Non zero Coefficients
min	0.01681	0.007050	0.001933	4
1se	0.01844	0.007902	0.001608	4

The estimated regression coefficients for the selected λ are presented in Table 2. Among the explanatory variables, only four animal husbandry, transport, area under crops, and input supply were retained in the final model. All other variables were eliminated as their coefficients were shrunk to zero, indicating their negligible influence on agricultural development in Maharashtra.

Table 2: Results of Lasso Regression

Particulars	Coefficient	Intercept	Lambda
Animal Husbandry	2.683×10^{-4}	-2.779×10^{-2}	0.0168
Transport	2.057×10^{-6}		
Area under Crops	1.574×10^{-6}		
Input Supply	5.987×10^{-8}		

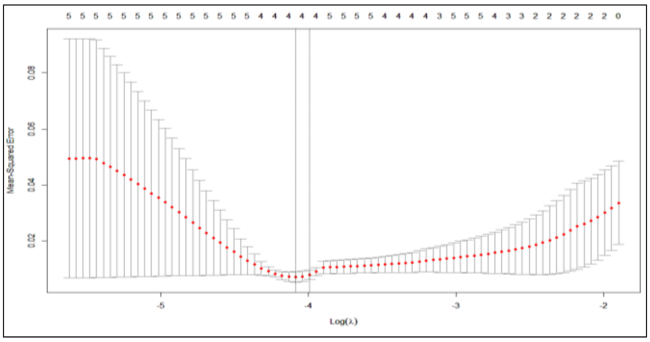


Fig 1: Lasso regression plot

The lasso regression plot (Fig.1) further confirmed the shrinkage path, demonstrating that the retained variables had the greatest predictive importance for the composite index of agricultural development.

The curve displays a typical U-shape, where MSE first decreases with increasing λ due to reduced over fitting, reaches a minimum at λ = 0.0168 (λ min), and then rises again as excessive shrinkage introduces bias.

Two vertical dashed lines are highlighted:

- **Left dashed line (λ_min):** Represents the value of λ that minimizes cross-validated MSE and was selected as the optimal parameter.
- **Right dashed line (λ_1se):** Represents the largest λ value within one standard error of the minimum MSE, which would yield a slightly simpler model with comparable predictive accuracy.

The numbers shown at the top of the plot indicate the number of non-zero coefficients for each λ. At λ_min, four variables retained nonzero coefficients, confirming their importance in explaining the composite index.

2. Spearman Correlation Analysis

To validate the results obtained from the lasso regression, a Spearman rank correlation analysis was conducted between the four selected variables and the composite index of agricultural development. The results are presented in Table 3.

Table 3: Spearman Correlation of Selected Variables with Composite Index

Variable	Correlation Coefficient
Animal Husbandry	0.94**
Transport	0.94**
Area under Crops	0.94**
Input Supply	0.71**

(**Significant at 5% level of significance)

The analysis revealed a very high and statistically significant positive correlation between animal husbandry, transport, area under crops, and the composite index of agricultural development (ρ = 0.94). Input supply also exhibited a strong positive correlation (ρ = 0.71). These results corroborate the findings of the lasso regression, confirming that these variables play a critical role in determining the level of agricultural development in Maharashtra.

The combined use of lasso regression and Spearman correlation provides robust evidence that animal husbandry, transport infrastructure, area under crops, and input supply are the most influential factors driving agricultural development in the state. The lasso model not only improved the predictive accuracy by handling multicoilinearity but also simplified the model by eliminating less relevant variables. The high correlation coefficients further substantiate the importance of these variables, reinforcing their inclusion in development planning and policy formulation.

Summary

Lasso regression was applied to address multicollinearity among explanatory variables and to identify the most significant determinants of agricultural development in Maharashtra. The optimal penalty parameter (λ = 0.01681) was selected using k-fold cross-validation, minimizing mean squared error (MSE) and ensuring a balance between model complexity and predictive performance. The final model retained four key predictors—animal husbandry, transport, area under crops, and input supply—while coefficients of all other variables were shrunk to zero, indicating negligible influence.

The lasso regression plot confirmed the selection of λ, illustrating that these four variables consistently retained non-zero coefficients at the optimal point. To validate the results, Spearman rank correlation analysis was conducted, revealing strong and statistically significant positive correlations between the selected variables and the composite index of agricultural development (ρ = 0.94 for three variables and ρ = 0.71 for input supply).

Overall, the combined evidence from lasso regression and correlation analysis highlights animal husbandry, transport, area under crops, and input supply as the most influential drivers of agricultural development in Maharashtra, providing a robust empirical basis for policy prioritization and resource allocation.

References

1. Mishra B. Inter-State disparities in the reduction of multidimensional deprivations in India. Indian J Hum Dev. 2019;13(3):239-253.
2. Mor KD, Devi S. Regional disparities in socio-economic development in post reform era: A study of Indian economy. Int J Eng Technol Sci Res. 2017;4(9).
3. Narain P, Sharma SD, Rai SC, Bhatia VK. Dimensions of socio-economic development in Jammu & Kashmir. J Ind Soc Agric Stat. 2005;59(3):243-250.
4. Preethi VP. Disparities in socio-economic development in Kerala: A disaggregated analysis. Econ Aff. 2022;67(3):153-159.