Scientific Research Review

RESEARCH Open Access

# Seasonal and Spatial Heterogeneity in Climate–Input Interactions on Cereal Crop Yields: Evidence from 30 Indian States (1997–2020)

**Gunawan Baharuddin** [1]*

* *Correspondence: Faculty of Economics and Business, Universitas Pancasila, South Jakarta City, 12630 Jakarta, Indonesia ; gunawanb@univpancasila.ac.id.*

**1** *Faculty of Economics and Business, Universitas Pancasila, South Jakarta City, 12630 Jakarta, Indonesia ; gunawanb@univpancasila.ac.id*

**Citation:** Baharuddin, G. (2026). Seasonal and Spatial Heterogeneity in Climate–Input Interactions on Cereal Crop Yields: Evidence from 30 Indian States (1997–2020). *Scientific Research Review*, 1(1), 6-19. https://doi.org/10.66348/srr.26.a63

Received: 2026-05-03
Revised: 2026-07-01
Accepted: 2026-07-04
Online Available: 2026-09-09
Published: 2026-09-30

**Copyright:** © 2026 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

**Abstract:** Climate variability remains a major challenge to agricultural productivity in India, yet the effects of rainfall may differ across cropping seasons and exhibit nonlinear patterns. This study examines the relationship between annual rainfall and cereal crop yields using a panel dataset of 3,754 observations covering 30 Indian states, five major cereal crops (rice, wheat, maize, jowar, and bajra), and the period 1997–2020. Fixed-effects panel models were employed to account for un-observed heterogeneity across states, crops, and years, while Random Forest and Extreme Gradient Boosting (XGBoost) models were used to complement the econometric analysis by capturing complex nonlinear relationships. The linear fixed-effects model indicates a positive but modest association between annual rainfall and crop yield ($\beta = 0.11$), although the magnitude of the relationship varies across cropping seasons. The interaction analysis suggests that the association is weaker during the Rabi season ($\beta = -0.21$, $p < 0.10$) than in other cropping systems. The nonlinear specification further reveals significant threshold effects for Whole Year and Winter cropping systems, indicating that rainfall–yield relationships are not adequately represented by a simple linear function. Machine learning models achieved an out-of-sample $R^2$ of 0.79, with rainfall and seasonal indicators accounting for more than 60% of the overall variable importance, and the partial dependence analysis demonstrated clear nonlinear response patterns. Overall, the findings suggest that the relationship between annual rainfall and cereal production in India is heterogeneous across cropping seasons and exhibits important nonlinear characteristics. By integrating panel econometric techniques with machine learning methods, this study provides complementary evidence on climate–yield relationships and highlights the importance of season-specific adaptation strategies for enhancing agricultural resilience under changing climatic conditions.

**Keywords:** Annual rainfall; cereal yield; cropping seasons; fixed-effects model; machine learning; nonlinear effects

## 1. Introduction

Agricultural production is highly sensitive to climatic variability, particularly in developing countries where rainfall remains one of the most important determinants of crop productivity. In India, agriculture continues to rely heavily on the timing and intensity of monsoon rainfall, making crop production vulnerable to fluctuations in precipitation and other climatic conditions. Increasing climate variability has intensified concerns regarding food security, agricultural sustainability, and farmers' livelihoods, highlighting the need for a better understanding of how climatic factors influence agricultural productivity.

A substantial body of empirical research has documented the role of rainfall and temperature in shaping agricultural outcomes. Singh et al. (2017) reported that rainfall positively influenced crop yields in Gujarat, whereas rising temperatures adversely affected agricultural productivity. Similarly, Upadhyay et al. (2024) found that rainfall

variability significantly influenced agricultural production in the Sikkim Himalaya, with deficient rainfall resulting in considerable yield losses. Kumar and Sharma (2022) further demonstrated that climate variability reduced agricultural productivity and threatened food security, particularly among smallholder farming households in India.

Evidence from previous studies also suggests that climate impacts are highly heterogeneous. Using household-level panel data, Amare et al. (2018) showed that negative rainfall shocks substantially reduced agricultural productivity and household consumption, with the strongest impacts observed among vulnerable households. At the macroeconomic level, weather shocks have also been linked to slower economic growth, particularly in regions with greater dependence on agriculture (Sandhani et al., 2023). Similarly, Chen and Chang (2005) and Malinová et al. (2026) demonstrated that the relationship between climatic variables and agricultural productivity differs across agroecological conditions, institutional settings, and crop characteristics, suggesting that average climate effects may conceal important regional and crop-specific variations.

Another important finding emerging from the literature is that climate–yield relationships are inherently nonlinear. Rather than responding proportionally to changes in rainfall, crop yields often improve under moderate rainfall but decline once precipitation exceeds critical thresholds. Schlenker and Roberts (2006) first demonstrated such nonlinear responses, while more recent studies have identified similar threshold effects for wheat production in India (Soni et al., 2026), rice production in China (Zhou et al., 2026), and hydrological responses to extreme rainfall events (Wu et al., 2026). These findings indicate that conventional linear models may not adequately capture the complexity of climate–agriculture interactions.

Despite these advances, relatively limited attention has been paid to seasonal heterogeneity in rainfall–yield relationships. Indian agriculture is organized around distinct cropping seasons, including Kharif, Rabi, Summer, Winter, Autumn, and Whole Year systems, each characterized by different levels of dependence on rainfall, irrigation, and residual soil moisture. Existing studies suggest that seasonal water availability strongly influences crop performance. Ammaiyappan et al. (2026) reported considerable seasonal variation in rice productivity due to differences in rainfall patterns and irrigation availability, whereas Munirathnam et al. (2026) highlighted the importance of residual soil moisture in sustaining crop production during the Rabi season. Nevertheless, most existing studies continue to examine aggregate annual relationships without explicitly investigating whether rainfall effects differ across cropping seasons.

Recent research has also highlighted that climate variability influences not only crop yields but also farmers' production decisions. Sumner et al. (2026) showed that drought conditions affected both crop yields and acreage allocation, while Bora (2022) found that rainfall shocks significantly influenced fertilizer application decisions in India. These findings suggest that climate affects agricultural productivity through multiple interconnected pathways, emphasizing the need for analytical frameworks capable of capturing both direct and indirect relationships.

Alongside these developments, machine learning techniques have become increasingly popular in agricultural research because of their ability to model complex nonlinear relationships and improve predictive performance. Random Forest, XGBoost, and related algorithms have demonstrated promising results for crop yield prediction (Antony, 2021; De Clercq & Mahdi, 2024; Nigam et al., 2019). However, these approaches are primarily designed for prediction and are rarely integrated with econometric methods that provide interpretable parameter estimates.

Despite substantial progress, three important research gaps remain. First, although nonlinear rainfall effects have been widely documented, relatively few studies have examined whether these relationships differ systematically across India's major cropping seasons. Second, most previous studies rely either on econometric methods for statistical inference or machine learning techniques for prediction, with limited attempts to integrate both approaches within a single analytical framework. Third, existing evidence is often

restricted to specific crops, regions, or relatively short study periods, limiting broader insights into India's diverse agricultural systems.

To address these gaps, this study examines the relationship between annual rainfall and cereal crop yields across different cropping seasons using a panel dataset covering 30 Indian states from 1997 to 2020. The analysis focuses on five major cereal crops—rice, wheat, maize, jowar, and bajra—and combines fixed-effects panel regression with machine learning techniques to investigate heterogeneous and nonlinear rainfall–yield relationships. By integrating interpretable econometric models with predictive machine learning algorithms, this study provides complementary evidence on climate–yield relationships and offers insights that may inform climate adaptation and agricultural policy in India.

## 2. Data and Methods

### *2.1. Data Sources and Sample Construction*

This study employs a balanced panel dataset covering agricultural production across 30 Indian states from 1997 to 2020. The dataset contains crop-level information on cultivated area, production, annual rainfall, fertilizer use, pesticide use, cropping season, and crop yield. The original dataset comprises 19,689 observations spanning 55 crops, 30 states, and 24 years. To ensure comparability across crops and focus on India's major staple cereals, the analysis is restricted to five cereal crops: rice, wheat, maize, jowar (sorghum), and bajra (pearl millet). These crops constitute an important component of India's food production system and exhibit relatively comparable agronomic responses to rainfall variability. After filtering the dataset, the final analytical sample consists of 3,754 observations.

### *2.2 Study Area and Cropping Seasons*

The study covers 30 Indian states representing diverse agroecological regions, ranging from humid tropical climates in states such as Kerala to semi-arid and arid environments in Rajasthan (see Table 1). These climatic differences provide substantial variation in rainfall conditions, making India an appropriate setting for examining climate–yield relationships. Agricultural production in India is organized around distinct cropping seasons that differ considerably in their dependence on rainfall, irrigation, and residual soil moisture. The Kharif season coincides with the southwest monsoon (June–October) and is highly dependent on rainfall. In contrast, the Rabi and Summer seasons rely more heavily on irrigation and residual soil moisture. Winter and Autumn represent relatively localized cropping periods, while Whole Year crops are cultivated continuously and are comparatively less dependent on seasonal rainfall patterns. These seasonal differences provide a natural framework for investigating heterogeneous rainfall effects on cereal crop yields.

Table 1. Seasonal Classification of Cereal Production in India

| Season | Typical Months | Agronomic Characteristics |
|---|---|---|
| Kharif | June–October | Southwest monsoon season; predominantly rain-fed and highly dependent on rainfall timing and intensity. |
| Rabi | October–March | Post-monsoon and winter season; relies primarily on irrigation and residual soil moisture. |
| Summer | March–June | Pre-monsoon season with minimal rainfall; largely dependent on irrigation. |
| Winter | December–February | Cold-season cropping, mainly in northern and high-altitude regions; production is constrained by low temperatures. |

| | | |
|---|---|---|
| Autumn | September–November | Transitional period between Kharif and Rabi; characterized by relatively limited cropping intensity. |
| Whole Year | Year-round | Crops cultivated throughout the year; comparatively less sensitive to seasonal rainfall variation. |

*2.3. Variable Construction*

The dependent variable is crop yield, calculated as production per unit of cultivated area. To reduce skewness and facilitate elasticity interpretation, crop yield is transformed using the natural logarithm:

$$\log(\text{Yield}_{s}c_{t}) \quad (1)$$

where *s*, *c*, and *t* denote state, crop, and year, respectively.

The principal explanatory variable is annual rainfall, measured at the state–year level. Rainfall is also transformed using the natural logarithm

$$\log(\text{Rain}_{st}) \quad (2)$$

The log transformation enables the estimated coefficients to be interpreted as elasticities.

Seasonal variation is represented using categorical indicators corresponding to the six cropping seasons described in Table 1 (Kharif, Rabi, Summer, Winter, Autumn, and Whole Year). Interaction terms between annual rainfall and cropping season are included to examine whether rainfall–yield relationships differ across production systems.

Fertilizer intensity was initially constructed by dividing fertilizer consumption by cultivated area and subsequently applying a logarithmic transformation. Although fertilizer intensity was explored as a control variable during model development, it was excluded from the final fixed-effects specifications because of perfect collinearity with the high-dimensional fixed effects.

*2.4. Identification Strategy*

The empirical analysis exploits temporal variation in annual rainfall within states to examine its association with cereal crop yields. Since rainfall is determined by meteorological conditions rather than individual farmers' production decisions, variation in rainfall provides a useful source of exogenous climatic variation. Nevertheless, the analysis is based on observational panel data; therefore, the estimated relationships should be interpreted as conditional associations rather than definitive causal effects.

To reduce potential confounding, the models incorporate state, crop, and year fixed effects. State fixed effects control for time-invariant characteristics such as geography, soil quality, and long-term climatic conditions. Crop fixed effects account for differences in biological and agronomic characteristics across cereal crops, while year fixed effects capture common temporal shocks, including technological progress, policy changes, and macroeconomic conditions. To account for serial correlation and heteroskedasticity within states, standard errors are clustered at the state level.

*2.5. Econometric Specification*

*2.5.1. Baseline Model*

The baseline analysis estimates the association between annual rainfall and cereal crop yields using a fixed-effects panel regression model. The specification controls for unobserved heterogeneity across states, crops, and years by incorporating state, crop, and year fixed effects. The model is specified as:

$$\log(\text{Yield}_{s}c_{t}) = \beta_{1}\log(\text{Rain}_{st}) + \alpha_{s} + \delta c + \lambda_{t} + \varepsilon_{s}c_{t} \quad (3)$$

where *Yield* represents crop yield, *Rain* denotes annual rainfall, $\alpha_s$ represents state fixed effects, δc denotes crop fixed effects, $\lambda_t$ captures year fixed effects, and $\varepsilon_s c_t$ is the random error term. This specification estimates the average association between annual rainfall and cereal crop yields after controlling for time-invariant differences across states and crops, as well as common temporal shocks.

*2.5.2. Seasonal Heterogeneity*

To investigate whether rainfall–yield relationships differ across cropping systems, interaction terms between annual rainfall and cropping season indicators are introduced. This specification allows the elasticity of crop yield with respect to rainfall to vary across different production seasons.

$$\log(\text{Yield}_s c_t) = \beta_1 \log(\text{Rain}_{st}) + \beta_2 [\log(\text{Rain}_{st}) \times \text{Seasonc}_t] + \alpha_s + \delta c + \lambda_t + \varepsilon_s c_t \quad (4)$$

where *Season* represents categorical indicators for the six cropping seasons (Kharif, Rabi, Summer, Winter, Autumn, and Whole Year). The interaction terms capture heterogeneous rainfall effects across cropping systems rather than assuming a common rainfall elasticity for all crops.

*2.5.3. Nonlinear Effects*

Previous studies suggest that crop responses to rainfall may exhibit threshold behaviour, with productivity increasing under moderate rainfall but declining when rainfall becomes excessive. To account for these nonlinear relationships, a quadratic rainfall term and its interaction with cropping season are incorporated into the fixed-effects model.

$$\log(\text{Yield}_s c_t) = \beta_1 \log(\text{Rain}_{st}) + \beta_2 \log(\text{Rain}_{st})^2 + \gamma \text{Seasonc}_t + \beta_3 [\log(\text{Rain}_{st}) \times \text{Seasonc}_t] + \beta_4 [\log(\text{Rain}_{st})^2 \times \text{Seasonc}_t] + \alpha_s + \delta c + \lambda_t + \varepsilon_s c_t \quad (5)$$

The inclusion of the quadratic rainfall term allows the relationship between rainfall and crop yield to depart from linearity and enables the identification of potential threshold effects across different cropping seasons.

All econometric models were estimated using fixed-effects panel regression. To account for heteroskedasticity and within-state serial correlation, **standard errors were clustered at the state level**.

*2.6. Machine Learning Framework*

To complement the econometric analysis, two supervised machine learning algorithms—Random Forest (RF) and Extreme Gradient Boosting (XGBoost)—were employed to explore complex nonlinear relationships and evaluate the predictive performance of climatic variables.

The machine learning models were trained using annual rainfall and cropping season indicators as predictor variables, while log-transformed crop yield served as the response variable. Model performance was evaluated using out-of-sample prediction based on the coefficient of determination ($R^2$) and the root mean squared error (RMSE).

To improve model interpretability, variable importance measures were computed to quantify the relative contribution of each predictor to crop yield prediction. In addition, partial dependence plots (PDPs) were generated to visualize the marginal relationship between rainfall and predicted crop yield while averaging over the remaining predictors. These graphical tools provide complementary evidence regarding potential nonlinear rainfall–yield relationships identified by the econometric models.

*2.7. Robustness and Validation*

To evaluate the stability of the empirical findings, several complementary analyses were conducted. First, both linear and nonlinear fixed-effects panel models were estimated to examine whether the rainfall–yield relationship remained consistent under alternative functional specifications. Comparing these models enabled the assessment of potential nonlinear and threshold effects of annual rainfall across different cropping seasons.

Second, all econometric models were estimated with **state-level clustered standard errors** to account for heteroskedasticity and within-state serial correlation, thereby improving the reliability of statistical inference.

Third, machine learning models, namely Random Forest and Extreme Gradient Boosting (XGBoost), were implemented to complement the econometric analysis by capturing complex nonlinear relationships and evaluating predictive performance. Variable importance measures and partial dependence plots were further used to assess the relative contribution of rainfall and seasonal indicators to crop yield prediction.

During model development, fertilizer intensity was initially considered as a control variable. However, it exhibited perfect collinearity under the fixed-effects specification and was therefore excluded from the final econometric models. Consequently, the reported results are based on parsimonious specifications that avoid unstable parameter estimates arising from multicollinearity.

Taken together, these complementary analyses provide a comprehensive assessment of the relationship between annual rainfall and cereal crop yields by combining interpretable econometric models with flexible machine learning techniques.

## 3. Results

### *3.1 Descriptive Statistics*

Table 2 presents the descriptive statistics of the variables used in the empirical analysis. The average log-transformed crop yield was 1.03 (SD = 0.41), while the average annual rainfall across the sampled states was 1,412.63 mm, with substantial variation ranging from 301.30 mm to 6,552.70 mm. Cultivated area and crop production exhibited considerable dispersion, as reflected by their standard deviations and the noticeable differences between the mean and median values. This indicates the presence of right-skewed distributions and justifies the logarithmic transformation of the variables used in the econometric analysis.

Table 2. Descriptive Statistics of the Study Variables

| Variable | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|
| Log crop yield | 1.028 | 1.005 | 0.406 | 0.000 | 6.899 |
| Log annual rainfall | 7.120 | 7.108 | 0.513 | 5.711 | 8.788 |
| Annual rainfall (mm) | 1412.634 | 1220.500 | 799.236 | 301.300 | 6552.700 |
| Cultivated area (ha) | 535406.262 | 79054.000 | 1208030.541 | 1.000 | 10216517.000 |
| Crop production (tonnes) | 1254933.230 | 134878.500 | 3375100.966 | 0.000 | 38039724.000 |

Note: Yield and rainfall were log-transformed for the econometric analysis to reduce skewness and facilitate elasticity interpretation. Area and production are reported in their original units for descriptive purposes.

Figure 1 illustrates the distribution of the log-transformed crop yield, demonstrating that the logarithmic transformation substantially reduced right skewness and produced a distribution suitable for regression analysis. Figure 2 presents the relationship between annual rainfall and crop yield. Although a positive association is observable, the considerable dispersion suggests that rainfall effects are heterogeneous and may not be adequately represented by a simple linear relationship.

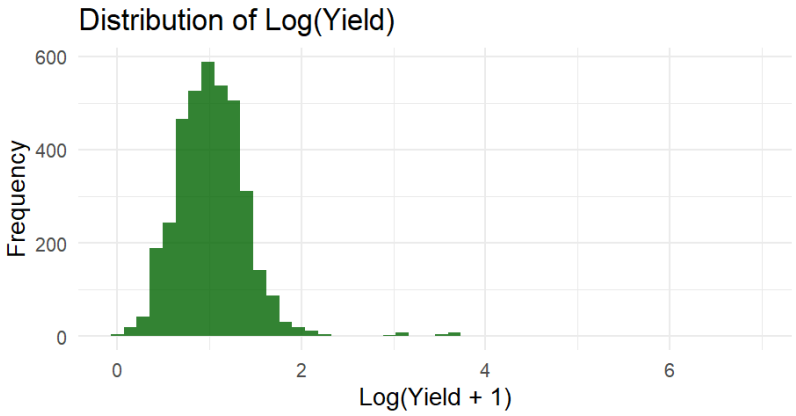



Figure 1. Distribution of Log-Transformed Crop Yield

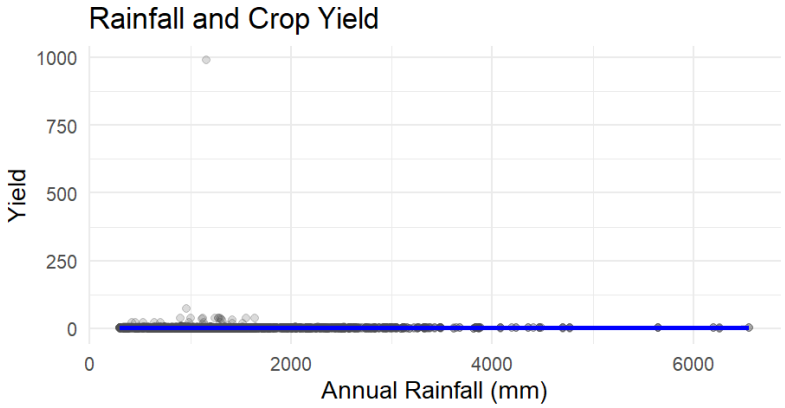



Figure 2. Relationship Between Annual Rainfall and Crop Yield

Figures 3 and 4 further illustrate variations in crop yield across cropping seasons and over time. Whole Year crops exhibit the greatest variability in yield, whereas Rabi and Winter crops display comparatively more stable production levels. In addition, average crop yield demonstrates a gradual upward trend throughout the study period, reflecting long-term improvements in agricultural productivity, technological advancement, and farm management practices.

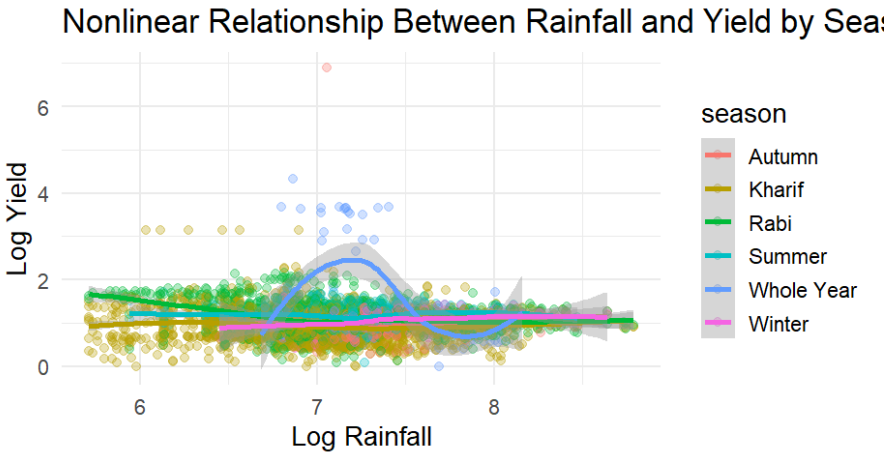



Figure 3. Distribution of Crop Yield Across Cropping Seasons

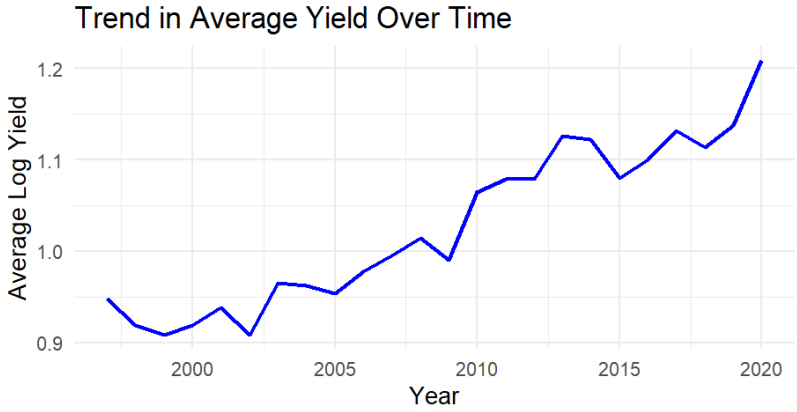



Figure 4. Temporal Trend in Average Crop Yield (1997–2020)

Overall, the descriptive analysis indicates substantial variation in both climatic conditions and agricultural production across states and cropping seasons. These preliminary patterns suggest that rainfall–yield relationships are likely to be heterogeneous and potentially nonlinear, thereby motivating the econometric analysis presented in the following section.

### *3.2. Fixed-Effects Regression Results*

Table 3 reports the results of the fixed-effects panel regression models estimated with state-level clustered standard errors. Model (1) presents the baseline specification examining the average association between annual rainfall and cereal crop yield. Model (2) extends the analysis by interacting annual rainfall with cropping season indicators to investigate seasonal heterogeneity. Model (3) further incorporates quadratic rainfall terms and their interactions with cropping seasons to examine potential nonlinear rainfall–yield relationships.

Table 3. Fixed-Effects Regression Results

| Variables | Model (1)Baseline | Model (2)Seasonal Heterogeneity | Model (3)Nonlinear Model |
|---|---|---|---|
| Log annual rainfall | 0.032 (0.023) | 0.114 (0.119) | 0.133 (0.359) |
| Season: Kharif | – | 0.566 (0.948) | – |
| Season: Rabi | – | 1.704 (0.971) | – |
| Season: Summer | – | 0.504 (1.101) | – |
| Season: Whole Year | – | 8.450 (5.674) | – |
| Season: Winter | – | –0.484 (0.765) | – |
| Log rainfall × Kharif | – | –0.074 (0.124) | 0.078 (0.128) |
| Log rainfall × Rabi | – | –0.210 (0.126) | 0.244 (0.134) |
| Log rainfall × Summer | – | –0.039 (0.144) | 0.087 (0.150) |
| Log rainfall × Whole Year | – | –1.049 (0.711) | 1.271 (0.819) |
| Log rainfall × Winter | – | 0.078 (0.101) | –0.038 (0.102) |
| Log annual rainfall² | – | – | –0.003 (0.029) |
| Log rainfall² × Kharif | – | – | –0.010 (0.017) |
| Log rainfall² × Rabi | – | – | –0.030 (0.017) |
| Log rainfall² × Summer | – | – | –0.008 (0.020) |
| Log rainfall² × Whole Year | – | – | –0.159 (0.102) |
| Log rainfall² × Winter | – | – | 0.007 (0.013) |
| State fixed effects | Yes | Yes | Yes |
| Crop fixed effects | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes |
| Clustered standard errors (State) | Yes | Yes | Yes |

| | | | |
|---|---|---|---|
| Observations | 3,754 | 3,754 | 3,754 |
| Within $R^2$ | 0.0003 | 0.1574 | 0.1583 |
| Adjusted $R^2$ | 0.3431 | 0.4449 | 0.4453 |

The baseline model indicates a positive but statistically insignificant association between annual rainfall and cereal crop yield. After introducing interactions between rainfall and cropping seasons, the explanatory power of the model increases substantially, with the within $R^2$ improving from 0.0003 in Model (1) to 0.1574 in Model (2). This suggests that rainfall effects differ across cropping seasons rather than remaining constant throughout the agricultural production cycle.

The seasonal interaction model indicates considerable heterogeneity in rainfall–yield relationships. Although most interaction terms are not statistically significant after accounting for clustered standard errors, the negative interaction observed for the Rabi season suggests that cereal production during the winter growing season is comparatively less responsive to rainfall than other cropping systems. These findings are consistent with the greater dependence of Rabi cultivation on irrigation and residual soil moisture rather than direct precipitation.

Incorporating quadratic rainfall terms results in only a modest increase in model fit, with the within $R^2$ increasing from 0.1574 to 0.1583. The nonlinear specification therefore provides limited additional explanatory power beyond the seasonal interaction model. Nevertheless, the estimated coefficients suggest that rainfall responses are unlikely to be perfectly linear across all cropping systems, supporting the use of flexible modeling approaches in subsequent analyses.

While the fixed-effects models provide interpretable estimates of rainfall–yield relationships, they may not fully capture complex nonlinear interactions among climatic variables. Therefore, machine learning techniques are employed in the following section to complement the econometric analysis and further explore nonlinear response patterns.

### *3.3. Machine Learning Results*

Random Forest and Extreme Gradient Boosting (XGBoost) models were employed to complement the econometric analysis by exploring complex nonlinear relationships and evaluating predictive performance. Both algorithms demonstrated satisfactory predictive performance, with the best-performing model achieving an **out-of-sample $R^2$ of approximately 0.79**, indicating that a substantial proportion of the variation in cereal crop yield could be explained by the selected predictors.

Variable importance analysis (Figure 5) identified **annual rainfall** and **cropping season** as the most influential predictors of crop yield, jointly accounting for more than **60%** of the overall variable importance. This finding highlights that climatic conditions, particularly rainfall, remain the dominant determinants of cereal productivity across India's diverse agricultural systems.

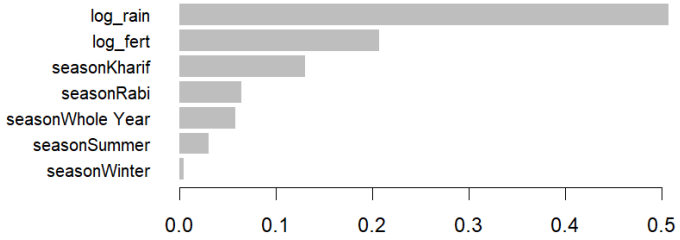


Figure 5. Variable importance from Random Forest model

The partial dependence analysis further illustrates the marginal effect of annual rainfall on predicted crop yield (Figure 6). Rather than exhibiting a strictly linear response, predicted crop yield changes nonlinearly across different rainfall levels, suggesting that the productivity gains associated with additional rainfall vary depending on existing precipitation conditions. Although local fluctuations are evident, the overall pattern supports the presence of heterogeneous rainfall responses.

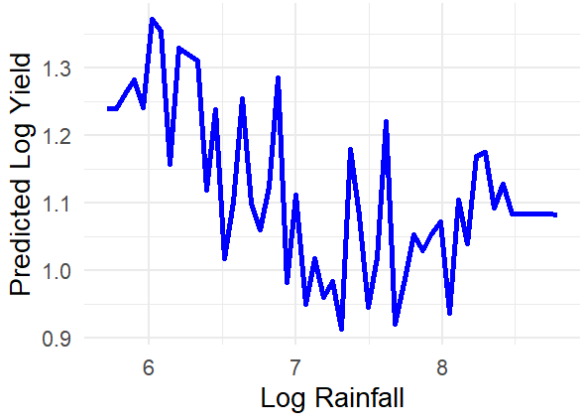



Figure 6. Partial dependence of yield on rainfall

*3.4 Summary of Main Findings*

Overall, the empirical analysis reveals three key findings. First, the relationship between annual rainfall and cereal crop yield varies across cropping seasons, indicating that seasonal heterogeneity plays an important role in shaping climate–yield relationships. Second, incorporating seasonal interactions substantially improves the explanatory power of the fixed-effects models, whereas the addition of quadratic rainfall terms provides only a modest improvement, suggesting limited evidence of strong nonlinear effects. Third, the machine learning models identify annual rainfall and cropping season as the most influential predictors of crop yield, while partial dependence analysis indicates that rainfall–yield relationships are heterogeneous across different rainfall levels. Taken together, the econometric and machine learning results suggest that the impact of rainfall on cereal production in India is season-dependent and cannot be adequately captured by a single linear relationship. These findings provide a foundation for the discussion of their implications for agricultural adaptation and climate-resilient policy.

## 4. Discussion

This paper presents new findings on the nonlinear and seasonally heterogeneous impacts of rainfall on cereal crop yields in India. Using a combination of panel econometric methods and machine learning approaches, the analysis shows three main results: (i) the effect of rainfall on crop yield is highly season-dependent, (ii) rainfall–yield relationships exhibit heterogeneous and moderately nonlinear patterns across cropping systems, and (iii) machine learning and econometric approaches provide complementary evidence on the role of rainfall in determining cereal productivity.

*4.1 Seasonal Heterogeneity in Rainfall Effects*

One of the major results of this paper is the high seasonal heterogeneity in the effect of rainfall on crop yields. Although the average effect of rainfall is not strong and only marginally significant, the disaggregated results indicate that rainfall exhibits a comparatively weaker association with crop yield during the Rabi season than in other cropping systems, although the estimated interaction is not statistically significant after clustering

standard errors. This finding is consistent with the agronomic characteristics of Rabi crops, which depend primarily on irrigation and residual soil moisture rather than direct rainfall. Increased precipitation during this period may lead to waterlogging, reduced sunlight exposure, and greater susceptibility to crop diseases.

This finding aligns with the existing literature emphasizing the heterogeneous effects of climatic variables across crops and regions. For instance, Singh et al. (2017) documented that rainfall generally has a positive effect on crop yields, although its impact varies considerably depending on crop type and climatic conditions. Similarly, Amare et al. (2018) showed that rainfall shocks can have differential effects across agro-ecological zones, highlighting the importance of context-specific analysis.In contrast, the lack of a statistically significant rainfall effect during the Kharif season suggests that additional rainfall under high monsoon conditions does not necessarily translate into higher yields and may even become less beneficial because of flooding, nutrient leaching, or other adverse agronomic conditions.

*4.2 Nonlinear Rainfall–Yield Relationships*

The nonlinear specification suggests that rainfall–yield relationships are not perfectly linear, although the additional explanatory power beyond the seasonal interaction model is modest. The results indicate that rainfall responses differ across cropping systems, implying that linear specifications may not fully capture the complexity of climate–yield interactions. This finding is consistent with the broader literature emphasizing nonlinear climate–yield relationships. Schlenker and Roberts (2006) demonstrated that weather variables, including rainfall and temperature, often exhibit nonlinear effects on agricultural productivity, with both deficits and excesses being harmful. Our findings build on this literature by demonstrating that these nonlinear responses are not constant but differ across cropping seasons. The estimated nonlinear effects observed in some cropping systems indicate that crops grown under different climatic conditions respond differently to rainfall variability. This highlights the importance of adopting flexible functional forms when modelling climate–agriculture interactions, as linear specifications may overlook heterogeneous rainfall responses across production systems.

*4.3 Consistency with Machine Learning Evidence*

The machine learning results are broadly consistent with the econometric findings. Random Forest and XGBoost models identify rainfall and cropping season as the most influential predictors of crop yield, further highlighting the central role of climatic factors in determining agricultural productivity. In addition, the partial dependence analysis indicates that rainfall exhibits heterogeneous nonlinear associations with predicted crop yield, although localized fluctuations are observed across different rainfall levels.

These findings demonstrate the value of combining econometric and machine learning methods. While econometric models provide interpretable statistical estimates of rainfall–yield relationships, machine learning techniques capture complex nonlinearities and interactions without imposing restrictive functional form assumptions. The combined approach therefore provides a more comprehensive understanding of climate–yield relationships while balancing model interpretability with predictive flexibility.

*4.4 Novelty and Contribution*

This study makes several contributions to the existing literature. First, while previous studies have examined the impact of rainfall on agricultural productivity in India (Bhanumurthy et al., 2018; Kumar & Sharma, 2022; Singh et al., 2017), most have relied on linear specifications or aggregate analyses.Second, the study exploits state–year variation in rainfall within a fixed-effects framework, allowing for more reliable estimation of rainfall–yield associations while controlling for unobserved time-invariant heterogeneity across states and cropsThis approach builds on established panel data methodologies (Hsiao, 2022; Wooldridge, 2010) while applying them to a novel context. Third, the integration of machine learning methods with traditional econometric analysis represents a

methodological advancementWhile recent studies have applied machine learning to crop yield prediction (De Clercq & Mahdi, 2024; Nigam et al., 2019), relatively few have combined these approaches with panel econometric frameworks to enhance both predictive performance and model interpretability.

Finally, the explicit focus on seasonal dynamics distinguishes this study from much of the existing literature, which often treats rainfall as an aggregate annual measure. Given the importance of monsoon timing and variability in India, this seasonal perspective provides valuable insights for both researchers and policymakers.

*4.5 Policy Implications*

The results have important implications for agricultural policy and climate adaptation planning in India. First, the weaker rainfall response observed during the Rabi season suggests the need for continued investment in efficient irrigation systems, improved drainage infrastructure, and water management practices to reduce the adverse effects of excessive rainfall during winter cultivation. Second, the heterogeneous rainfall responses observed across cropping seasons indicate that climate adaptation strategies should be tailored to specific production systems rather than adopting a uniform national approach. Investments in climate information services, climate-resilient crop varieties, and season-specific adaptation measures may help reduce production risks associated with increasing rainfall variability. Finally, the findings highlight the importance of integrating climate variability into agricultural planning and policy formulation. Strengthening early warning systems, improving weather forecasting, and promoting adaptive farm management practices can enhance the resilience of cereal production under changing climatic conditions.

*4.6 Limitations and Future Research*

Despite its contributions, this study has several limitations. First, the use of annual rainfall data does not capture intra-seasonal variability, which may be critical for crop growth, as highlighted by Zhou et al. (2026). Second, the analysis does not include temperature information, limiting the ability to isolate the independent effects of rainfall from other important climatic factors. Third, the analysis focuses on yield outcomes and does not account for adjustments in cropping decisions, such as changes in cultivated area, which may also respond to climatic conditions (Sumner et al., 2026).

Future research could incorporate higher-frequency weather data, including seasonal rainfall and temperature, to better capture crop-specific climatic conditions. In addition, exploring dynamic farmer responses to climate variability and integrating remote sensing, soil characteristics, and other environmental indicators would further improve the understanding of climate–agriculture interactions. Overall, these findings demonstrate that the relationship between rainfall and cereal crop productivity in India is heterogeneous across cropping systems and cannot be fully explained by a single linear relationship. By integrating panel econometric techniques with machine learning approaches, this study provides complementary evidence on the complex nature of climate–yield relationships and offers insights that may support the development of more targeted and climate-resilient agricultural policies.

## 5. Conclusion

This paper examines the nonlinear and seasonally heterogeneous impacts of rainfall on the yields of cereal crops in India using panel econometric and machine learning methods. The findings indicate that the effects of rainfall are not homogeneous; although the overall association between rainfall and crop yield is modest, considerable variation exists across different cropping seasons. Specifically, rainfall exhibits a comparatively weaker association with crop yield during the Rabi season, whereas the response during the Kharif monsoon season is less pronounced, suggesting that rainfall effects differ across cropping systems rather than following a uniform pattern.

The analysis also indicates that rainfall–yield relationships are not perfectly linear, although the additional contribution of the nonlinear specification beyond the seasonal interaction model is relatively modest. The machine learning models identify annual rainfall and cropping season as the most influential predictors of cereal crop yield, while the partial dependence analysis suggests heterogeneous rainfall responses across different rainfall levels. Together, the econometric and machine learning results provide complementary evidence that rainfall effects vary across production systems and environmental conditions. Collectively, these findings demonstrate that climate effects on agriculture cannot be sufficiently explained using aggregate or simple linear models. Instead, seasonal dynamics should be explicitly incorporated into both empirical analyses and agricultural policy design. From a policy perspective, the findings highlight the importance of strengthening irrigation and water management during the Rabi season, investing in climate-resilient crop varieties, and developing season-specific adaptation strategies that account for differences in rainfall sensitivity across cropping systems. Future research should move beyond annual rainfall aggregates by incorporating intra-seasonal rainfall variability, temperature, and farmer behavioural responses, which are essential for understanding the full extent of climate impacts on agricultural systems. By integrating panel econometric techniques with machine learning approaches, this study provides a comprehensive framework for examining climate–agriculture interactions and offers evidence that may support more targeted and climate-resilient agricultural policies in developing countries.

**Declarations**

**Ethics and Guidelines:** Not Applicable.

**Consent to participate:** Not Applicable.

**Consent to publish:** The authors have provided consent to publish.

**Competing interests:** The authors declare no competing interests.

**Data availability statement:** Data will be made available on reasonable request from the corresponding author.

**Funding:** This research received no external funding.

**Clinical Trial Number:** Not Applicable.

**Declaration of using generative AI:** During the preparation of this work the author(s) used ChatGPT in order to correct the grammatical errors. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

**Author Contributions:** Conceptualization, G.B.; methodology, G.B.; writing—original draft preparation, G.B.; writing—review and editing G.B. All authors have read and agreed to the published version of the manuscript.

## References

Amare, M., Jensen, N. D., Shiferaw, B., & Cissé, J. D. (2018). Rainfall shocks and agricultural productivity: Implication for rural household consumption. Agricultural Systems, 166, 79–89. https://doi.org/10.1016/j.agsy.2018.07.014

Ammaiyappan, A., Khan, M. S., & Mahalingam, D. (2026). Delineation of seasonal rice efficient cropping zones in Tamil Nadu, India. International Journal of Environment and Climate Change, 16(2), 117–128.

Antony, B. (2021). Prediction of the production of crops with respect to rainfall. Environmental Research, 202, Article 111624. https://doi.org/10.1016/j.envres.2021.111624

Bhanumurthy, K., & Kumar, L. (2018). Climate change and agriculture in India: Studying long-term patterns in temperature, rainfall, and agricultural output. Management and Economics Research Journal, 4(2), 156–173.

Bora, K. (2022). Rainfall shocks and fertilizer use: A district level study of India. Environment and Development Economics, 27(6), 556–577. https://doi.org/10.1017/S1355770X22000018

Chen, C.-C., & Chang, C.-C. (2005). The impact of weather on crop yield distribution in Taiwan: Some new evidence from panel data models and implications for crop insurance. Agricultural Economics, 33(s3), 503–511.

De Clercq, D., & Mahdi, A. (2024). Feasibility of machine learning-based rice yield prediction in India at the district level using climate reanalysis and remote sensing data. Agricultural Systems, 220, Article 104099. https://doi.org/10.1016/j.agsy.2024.104099

Hsiao, C. (2022). Analysis of panel data (4th ed.). Cambridge University Press.

Kumar, A., & Sharma, P. (2022). Impact of climate variation on agricultural productivity and food security in rural India [Manuscript sub-mitted for publication]. https://ssrn.com/abstract=4144089

Malinová, A., Čermák, M., Krepl, V., & Vališ, Z. (2026). Temperature changes and agricultural performance linkage: Evidence from Africa. Sustainable Futures, 11, Article 101650. https://doi.org/10.1016/j.sftr.2026.101650

Munirathnam, P., Kumar, K. A., Manjunath, J., Neelima, S., Chaithanya, B., Babu, K. S., & Johnson, M. (2026). Cropping kharif fallows to enhance rabi chickpea productivity and cropping intensity in rainfed deep vertisols: A comprehensive re-view. Legume Research-An International Journal, 49(3), 355–364. https://doi.org/10.18805/LR-4545

Nigam, A., Garg, S., Agrawal, A., & Agrawal, P. (2019). Crop yield prediction using machine learning algorithms. In 2019 Fifth In-ternational Conference on Image Information Processing (ICIIP) (pp. 125–130). IEEE. https://doi.org/10.1109/ICIIP47207.2019.8985951

Sandhani, M., Pattanayak, A., & Kavi Kumar, K. (2023). Weather shocks and economic growth in India. Journal of Environmental Economics and Policy, 12(2), 97–123. https://doi.org/10.1080/21606544.2022.2087745

Schlenker, W., & Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. Review of Agricultural Economics, 28(3), 391–398. https://doi.org/10.1111/j.1467-9353.2006.00304.x

Singh, R., Devi, G., Parmar, D., & Mishra, S. (2017). Impact of rainfall and temperature on the yield of major crops in Gujarat state of India: A panel data analysis (1980-2011). Current Journal of Applied Science and Technology, 24(5), 1–15. https://doi.org/10.9734/CJAST/2017/37071

Soni, K., Sivasankar, V., & Sendhil, R. (2026). Climatic and demographic interactions: Implications for wheat production dynam-ics. Discover Agriculture, 4(1), Article 83. https://doi.org/10.1007/s44279-026-00552-0

Sumner, E., Li, M., & Shr, Y.-H. (2026). Is yield response enough? Drought impacts on crop acreage throughout the production cycle. American Journal of Agricultural Economics, 108(2), 542–565. https://doi.org/10.1111/ajae.12549

Upadhyay, A., Nigam, N. K., Mishra, P. K., & Rai, S. C. (2024). Climatic variability and its impact on the indigenous agricultural system using panel data analysis in the Sikkim Himalaya, India. Environmental Monitoring and Assessment, 196(1), Ar-ticle 33. https://doi.org/10.1007/s10661-023-12193-7

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data (2nd ed.). MIT Press.

Wu, A., Ma, C., Huo, S., Li, T., & Fu, Q. (2026). Vegetation type dominates slope-scale material loss under extreme rainfall: Nonlin-ear responses revealed by machine learning. Journal of Contaminant Hydrology, 271, Article 104846. https://doi.org/10.1016/j.jconhyd.2026.104846

Zhou, F., Tang, G., Wang, C., Qin, Y., Fu, B., & Fu, J. (2026). Pre-rainfall vapor pressure deficit stress and sunshine reduction govern sub-seasonal rainfall effects on China's rice yield. European Journal of Agronomy, 174, Article 127954. https://doi.org/10.1016/j.eja.2025.127954

**Disclaimer:** All statements, viewpoints, and data featured in the publications are exclusively those of the individual author(s) and contributor(s), not of SRR and/or its editor(s). SRR and/or the editor(s) absolve themselves of any liability for harm to individuals or property that might arise from any concepts, methods, instructions, or products mentioned in the content.