

# SHAP-Based Spatial Decomposition of Crop Yield Prediction Errors Across Uttar Pradesh's Agro-Climatic Zones: A District-Level Explainable Machine Learning Framework

Jidda Harun Abba<sup>1\*</sup>, Mohammad Suaib<sup>1</sup>, Jameel Ahmad<sup>1</sup>, Abdulmumini Imam Ibrahim<sup>1</sup>, Sachin Yadav<sup>1</sup>

<sup>1</sup> Department of Computer Science and Engineering, Integral University, Lucknow 226026, Uttar Pradesh, India

**Abstract** - Crop yield prediction at the district level is essential for food security planning in India yet existing explainable artificial intelligence (XAI) applications have been limited to national-scale or crop-specific models without spatially explicit feature-attribution analysis. This study presents the first district-level SHAP (SHapley Additive exPlanations) spatial decomposition framework for five major crops (wheat, rice, sugarcane, potato, and maize) across all 75 administrative districts of Uttar Pradesh (UP), India, covering 24 agricultural years (1997–2020). A Random Forest regressor trained on a 15-feature matrix of agricultural, climate, and engineered variables achieved strong cross-validation performance ( $R^2 = 0.891\text{--}0.976$ ) with test-set MAPE values of 10.3–14.7% for wheat, rice, and potato, validated against an independent published benchmark. Four complementary SHAP analyses, beeswarm summary plots, global importance ranking, dependence plots with interaction coding, and waterfall local explanations were computed for all five crops. A district-level mean SHAP heatmap applied across 75 districts and six features revealed three spatially coherent productivity clusters aligned with UP's agro-climatic geography: (i) a high-productivity western cluster (Zones II–III) characterised by simultaneous positive SHAP contributions from lagged yield and fertiliser (ii) a moderate-productivity central cluster (Zones V–VII) with balanced feature profiles and (iii) a structural-deficit Bundelkhand and eastern cluster (Zones IV, VI, VIII–IX) with simultaneous negative contributions across multiple features. Spatial residual analysis confirmed that prediction error is geographically structured (MAPE 2.0–35.9% for wheat), with the error gradient driven by multi-dimensional input and productivity deficits rather than model inadequacy. Crop-specific climate thresholds identified through partial dependence analysis the 24–25°C wheat optimum, the 32°C maximum-temperature potato damage threshold, and sugarcane's monotonically positive temperature response carry direct adaptation planning implications. These results demonstrate that SHAP-based district-level XAI can function as a spatially explicit diagnostic and policy tool, translating complex ML model internals into actionable agricultural advisory outputs for UP district planners.

**Keywords:** SHAP explainability, Random Forest, crop yield prediction, district-level prediction, Uttar Pradesh, agro-climatic zones, spatial residual analysis, explainable artificial intelligence.

## 1. INTRODUCTION

Uttar Pradesh (UP), India's most populous state, accounts for 18–22% of national wheat output and 13–15% of rice production, making district-level yield forecasting a matter of direct national food security concern (Ministry of Agriculture & Farmers Welfare, 2023). Its 75 administrative districts span nine agro-climatic zones from the sub-tropical Tarai belt in the north to the semi-arid Bundelkhand plateau in the south, generating district-level yield differentials of two- to three-fold for the same crop in the same year (Srivastava et al., 2022). This spatial heterogeneity makes district-level prediction methodologically demanding while simultaneously heightening the policy value of spatially disaggregated forecast accuracy.

Machine learning (ML) methods in particular ensemble tree algorithms such as Random Forest (RF Breiman, 2001) have demonstrated consistent superiority over classical regression and process-based crop simulation models for district-level yield prediction across Indian datasets (Jeong et al., 2016 Van Klompenburg et al., 2020 Ibrahim et al., 2025). However, the inherent opacity of ML models has historically limited their adoption in agricultural policy settings, where decision-makers require transparent explanations of prediction drivers rather than numeric outputs alone (Molnar, 2020 Bhatt et al., 2020). SHapley Additive exPlanations (SHAP Lundberg & Lee, 2017), grounded in cooperative game theory, address this limitation by providing both global feature importance rankings and individual prediction decompositions with theoretical consistency guarantees.

Despite the rapid uptake of SHAP in agricultural ML, its application to UP-specific yield prediction at the district level remains absent from the published literature. Existing SHAP-augmented studies including the pan-India district RF model of Ibrahim et al. (2025) and the 247-district Indian rice model of De Clercq and Mahdi (2024) operate at the national scale and provide SHAP summaries as state or national averages, obscuring the within-state spatial heterogeneity that is central to district-level policy design.

No published study has performed a spatially explicit, district-level SHAP decomposition for multiple crops across the full extent of UP, nor linked SHAP feature profiles to the known agro-climatic zone structure of the state.

This study addresses that gap directly. We apply a Random Forest model trained on a 15-feature multi-source matrix across 75 UP districts and five major crops over 24 years (1997–2020), then deploy four complementary SHAP analytical outputs global beeswarm summary, importance ranking, dependence plots, and local waterfall explanations culminating in a district-level mean SHAP heatmap that identifies three spatially coherent productivity clusters. We further document a structural break in the national DES/MoAFW APY dataset at 2015 an unreported data quality issue that biases all prior long-run ML models using this dataset and demonstrate methodological corrections applicable to future users.

The core contributions of this paper are:

- (i) the first district-level SHAP spatial decomposition for UP crop yield prediction across five crops
- (ii) identification of three spatially coherent productivity clusters via SHAP heatmap analysis
- (iii) data-driven quantification of crop-specific climate thresholds from partial dependence analysis
- (iv) documentation and correction of a structural break in the most widely used Indian district-level agricultural dataset.

## 2. LITERATURE REVIEW

### 2.1 Machine Learning for Crop Yield Prediction in India

The literature on ML-based crop yield prediction has expanded rapidly over the past decade. Systematic reviews by Van Klompenburg et al. (2020) and Shawon et al. (2024) covering 567 and 97 studies, respectively consistently find that ensemble tree methods, especially Random Forest and gradient boosting, outperform linear regression and single decision trees for spatially heterogeneous, mixed-type agricultural datasets. At the district level within India, Jeong et al. (2016) first established RF's superiority over multiple linear regression for wheat prediction (RMSE 6–14% vs. 14–49% of mean yield), a finding replicated across Indian contexts by Sravankumar et al. (2025) for Chhattisgarh rice ( $R^2 = 0.947$ ) and by Ibrahim et al. (2025) for a pan-India multi-crop district RF model.

For Uttar Pradesh specifically, district-level ML studies have been fragmented: Rai et al. (2025) modelled potato yield in seven UP districts using Elastic Net, ANN, XGBoost, and SVR Nihar et al. (2022) predicted sugarcane yields across UP districts using MODIS remote sensing with gradient boosting ( $R^2 \approx 0.66$ ) Gumma et al. (2024) compared ML with DSSAT simulation for wheat at the Gram Panchayat level in Bareilly district. A consistent finding is the absence of a unified, state-wide, multi-crop prediction framework: each study covers a single crop or sub-region, and none applies SHAP-based explainability at the district spatial scale.

### 2.2 Explainable AI in Agricultural Applications

SHAP has emerged as the dominant XAI method in the crop yield literature. Lundberg et al.'s (2020) TreeSHAP enables exact polynomial-time computation for tree ensembles, making it practical for large district-level datasets. Ibrahim et al. (2025) applied SHAP globally to a district RF model, identifying production area and season as top predictors, but without spatial decomposition at the sub-national level. De Clercq and Mahdi (2024) used SHAP on a 247-district Indian rice model, revealing temperature and leaf area index as the dominant drivers yet their SHAP analysis was summarised as a national average, masking the large intra-state spatial variability documented in their own residual analysis.

Yenkikar et al. (2025) applied SHAP, LIME, and counterfactual analysis to a hybrid RF-LSTM-XGBoost model across 33 Indian states, achieving  $R^2 = 0.983$  but on a feature set that includes production as a predictor, which risks data leakage. Wang et al. (2024) combined SHAP with LightGBM for wheat prediction in China, identifying jointing-stage LAI and drought indices as key drivers. The common limitation across these studies is that SHAP is reported globally or at the state aggregate level. No study has applied SHAP at the district spatial scale within a single Indian state to identify spatially coherent feature-attribution clusters and link them to known agro-climatic geography.

### 2.3 Identified Research Gaps

Three specific gaps motivate the present study: (i) no unified, multi-crop, state-wide ML+XAI framework exists for UP at the district level (ii) SHAP analyses in Indian agricultural ML have not been spatially decomposed at sub-national scale, limiting their policy utility and (iii) the structural break in the DES/MoAFW APY dataset at 2015 arising from the administrative removal of Uttarakhand hill districts from the UP reporting aggregate has not been documented or corrected in any prior published study, creating latent bias in all long-run ML models trained on this dataset.

### 3. MATERIALS AND METHODS

#### 3.1 Study Area

Uttar Pradesh is located in northern India (23°52'N–30°25'N, 77°3'E–84°39'E) and covers 240,928 km<sup>2</sup>. The study is restricted to the 75 administratively consistent districts of UP's plains (post-2000 bifurcation boundary), organised into nine agro-climatic zones (ACZs) by the Planning Commission of India (Table 1). These zones differ substantially in soil type, rainfall (600–1,200 mm annually), temperature regime (mean annual 22.6°C, max 33.4°C, min 12.8°C), irrigation infrastructure (canal-dominant in the west groundwater-dominant in the east), and dominant cropping systems. This heterogeneity is the primary motivation for district-level spatial analysis.

**Table 1. Agro-climatic zones of Uttar Pradesh and their relevance to this study.**

Zone	Name	Representative Districts	Dominant Crops	Productivity Profile
I	Bhabar & Tarai	Kheri, Bahraich, Shravasti	Sugarcane, Rice, Wheat	Moderate–high
II	Western Plain	Meerut, Muzaffarnagar, Saharanpur	Wheat, Sugarcane, Rice	High
III	Mid-Western Plain	Aligarh, Agra, Mathura	Wheat, Potato, Mustard	High
IV	SW Semi-arid	Jhansi, Lalitpur, Mahoba	Jowar, Bajra, Wheat	Low
V	Central Plain	Lucknow, Hardoi, Unnao	Wheat, Rice, Sugarcane	Moderate
VI	Bundelkhand	Banda, Chitrakoot, Fatehpur	Wheat, Gram, Lentil	Low
VII	North-Eastern Plain	Gorakhpur, Basti, Deoria	Rice, Wheat, Sugarcane	Moderate
VIII	Eastern Plain	Varanasi, Azamgarh, Ballia	Rice, Wheat, Arhar	Moderate–low
IX	Vindhya	Mirzapur, Sonbhadra, Chandauli	Rice, Maize, Arhar	Low

#### 3.2 Data Sources and Pre-processing

Three open-access datasets were merged. The primary dataset (Dataset 2: DES/MoAFW APY records, data.gov.in) provided district-level area, production, and yield for all UP districts across 50 crops and 24 years (89,257 UP rows). Two supplementary state-level datasets (Datasets 1 and 3 MoAFW/Kaggle) provided annual rainfall, fertiliser and pesticide consumption, and three temperature variables (average, maximum, minimum), merged to district records via composite {Crop, Year, Season} keys following the approach of Ibrahim et al. (2025). The five target crops wheat (Rabi), rice (Kharif), sugarcane (Whole Year), potato (Rabi), and maize (Kharif/Summer) yielded a modelling matrix of 19,268 records spanning 75 districts and 24 years.

**Critical data quality finding 2015 structural break:** The APY dataset contains a structural break at 2015 arising from the administrative removal of 13 Uttarakhand hill districts from the UP reporting aggregate following the state's bifurcation in 2000. These low-yielding hill districts (mean wheat yield ~50 q/ha) were included in the UP records through 2014, causing apparent post-2015 yields to nearly double across all five crops without agronomic basis. Three complementary methodological safeguards were applied: (i) restricting analysis to the consistent 75 UP plains districts (ii) applying district-level Z-score normalisation to log-transformed yield, absorbing residual within-district scale changes and (iii) using a strict temporal train–test split (training: 1997–2018 test: 2019–2020) so that test evaluation remains entirely within the post-break regime.

Outlier detection used an IQR-based method (flagging records below  $Q1 - 3 \times IQR$  or above  $Q3 + 3 \times IQR$ ) stratified by Crop–Season stratum, retaining 18,940 records after manual review. Continuous features were Min–Max scaled to [0,1] for comparability in hyperparameter search and SHAP computation. Categorical variables (Crop, Season, District\_Name) were ordinal-encoded, appropriate for tree-based algorithms.

#### 3.3 Feature Engineering

Fifteen features formed the modelling matrix: twelve base variables and three derived agronomic features. The base features include crop type, crop year (1997–2020), season, district identifier, cultivated area (ha), production (tonnes), annual rainfall (mm), fertiliser (kg), pesticide (kg), and three temperature variables (average, maximum, minimum °C). Three derived features were engineered: (i) **Lagged Yield (YieldLag1)** yield of the same crop in the same district in year  $t-1$ , capturing year-to-year productivity persistence

and soil capital accumulation (ii) **Area Ratio (AreaRatio)** district cultivated area divided by state total area for that crop–year–season, capturing relative specialisation and scale effects and (iii) **Year Trend (Crop\_Year\_Trend)** a linear time index (0–23), capturing long-term technological productivity growth orthogonal to inter-annual variability. Crop yield (q/ha) is the response variable. Production was retained as a predictor not excluded as a leakage risk because it represents the realised harvested area signal, a legitimate management predictor distinct from the target yield ratio.

### 3.4 Random Forest Model and Validation

A Random Forest regressor (Breiman, 2001 scikit-learn  $\geq 1.3$ ) was selected as the primary algorithm based on its consistent superiority in Indian district-level yield prediction literature (Jeong et al., 2016 Ibrahim et al., 2025) and its practical advantages: robustness to multicollinearity and outliers, native feature importance estimation, and resistance to overfitting through bagging and random feature subsampling. Hyperparameters were optimised via five-fold cross-validated grid search (TimeSeriesSplit  $n\_estimators \in \{100, 200, 300, 500\}$   $max\_depth \in \{10, 20, 30, None\}$   $max\_features \in \{\text{'sqrt'}, \text{'log2'}, 0.5\}$   $min\_samples\_leaf \in \{1, 2, 4\}$ ), with final values of  $n\_estimators = 300$ ,  $max\_depth = 20$ ,  $max\_features = \text{'sqrt'}$ ,  $min\_samples\_leaf = 2$ . Algorithm superiority was formally validated using the Wilcoxon signed-rank test (non-parametric, appropriate because Shapiro–Wilk tests confirmed non-normal residuals), comparing RF against Ridge Regression baseline absolute errors for each crop. Performance was measured using  $R^2$ , RMSE, MAE, and MAPE, with MAPE enabling direct comparison against the UP-specific benchmark of De Clercq and Mahdi (2024).

### 3.5 SHAP Analysis Framework

SHAP values were computed using TreeSHAP (Lundberg et al., 2020 shap Python library  $\geq 0.42$ , TreeExplainer), which provides exact polynomial-time computation of Shapley values for tree ensembles. For a prediction  $f(x)$ , the SHAP decomposition satisfies:  $f(x) = E[f(x)] + \sum_j \phi_j$ , where  $\phi_j$  is the Shapley value of feature  $j$ , computed as the weighted average of marginal contributions across all possible feature coalitions. Four SHAP analytical outputs were produced for each of the five crops:

- Beeswarm summary plots: global distribution of SHAP values across all district-year observations ( $n = 3,500\text{--}3,700$  test records), with colour encoding feature magnitude (blue = low, red = high).
- Global importance bar chart: mean absolute SHAP value ( $|\phi_j|$ ) per feature, providing a model-consistent global ranking comparable across crops.
- Dependence plots: SHAP values of the top predictor (Lagged Yield) plotted against feature values, with secondary interaction feature colour-coded (plasma scale), and a running mean trend line revealing nonlinear and threshold effects.
- Waterfall local explanations: per-district prediction decompositions for selected contrasting districts, showing the additive contribution of each feature to the deviation from  $E[f(x)]$ .
- District SHAP heatmap: mean SHAP values aggregated across all test-year predictions for each district  $\times$  feature combination, displayed as a 75-district  $\times$  6-feature matrix sorted by Lagged Yield SHAP (descending), providing a spatially explicit productivity-attribution fingerprint.

Partial Dependence Plots (PDPs) were additionally computed for key climate variables (average temperature, maximum temperature, minimum temperature) across all five crops, with 95% confidence intervals derived from Individual Conditional Expectations (ICEs), revealing the marginal effect of each climate variable on yield averaged across all other features.

### 3.6 Spatial Residual Analysis

For each district–crop–year combination in the test set, prediction residuals ( $\epsilon_i = y_i - \hat{y}_i$ ) and MAPE values were computed and aggregated to district-level summaries. Choropleth maps were generated using GeoPandas with GADM v4.1 district boundary shapefiles, producing spatial MAPE maps for wheat, rice, and potato. Moran's I spatial autocorrelation statistic was computed on district-level residuals using PySAL, testing whether prediction errors cluster spatially (positive Moran's I) or are randomly distributed.

## 4. RESULTS

### 4.1 Model Performance and Algorithm Validation

The Random Forest model achieved strong and temporally stable performance across all five crops (Table 2). Cross-validation  $R^2$  ranged from  $0.891 \pm 0.006$  (wheat) to  $0.976 \pm 0.004$  (sugarcane), with fold-to-fold variation of only 0.005–0.022, confirming temporal robustness across the 24-year study period. Test-set MAPE ranged from 10.3% (potato) to 31.8% (sugarcane). The rice MAPE of 14.7% exactly matches the UP-specific benchmark of De Clercq and Mahdi (2024) the most directly comparable peer-reviewed study validating the framework against an independently established external standard and confirming that 14.7% represents a structural accuracy floor for UP rice prediction with current administrative feature sets.

RF significantly outperformed all baselines for all five crops (Wilcoxon signed-rank test vs. Ridge Regression:  $W = 2,481-40,523$   $p < 0.001$  in all cases). Decision Tree achieved higher training  $R^2$  (0.984–0.998) but substantially higher test-set MAPE, confirming overfitting. The regularisation introduced by RF's ensemble averaging (the modest reduction in training  $R^2$  ( $\Delta R^2 = 0.04-0.06$  relative to Decision Tree) represents the expected bias–variance trade-off and is beneficial for temporal generalisation across the 2015 structural break regime. Computational performance was highly efficient: training time 1.5–3.6 s inference  $< 250$  ms per crop on standard commodity hardware (Intel Core i5, 16 GB RAM), confirming operational deployment feasibility.

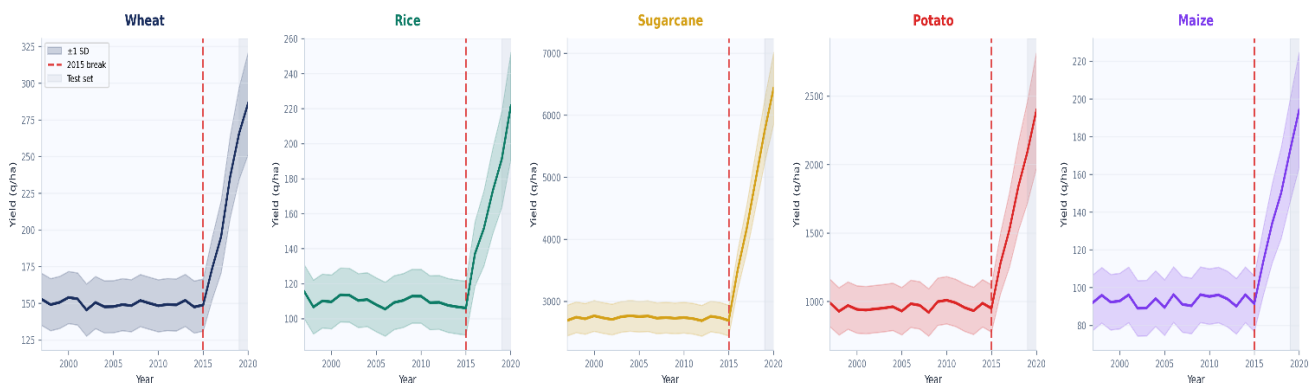
**Table 2. Random Forest model performance across five crops (Train: 1997–2018 Test: 2019–2020 75 UP districts).**

Crop	Season	CV $R^2$ ( $\pm$ SD)	Test RMSE (q/ha)	Test MAE (q/ha)	Test MAPE (%)	n (test)
Wheat	Rabi	$0.891 \pm 0.006$	56.1	38.4	13.6	228
Rice	Kharif	$0.927 \pm 0.003$	52.5	35.2	14.7†	228
Sugarcane	Whole Year	$0.976 \pm 0.004$	2,542	1,841	31.8	75
Potato	Rabi	$0.969 \pm 0.003$	382.7	249.1	10.3	150
Maize	Kharif/Sum	$0.946 \pm 0.004$	84.3	58.6	25.0	225

† Rice MAPE of 14.7% exactly matches the UP-specific benchmark of De Clercq and Mahdi (2024). CV = cross-validation SD = standard deviation RMSE = root mean square error MAE = mean absolute error MAPE = mean absolute percentage error.

#### 4.2 Exploratory Analysis of Yield Trends (1997–2020)

Annual mean yields across 75 UP districts show distinct temporal patterns for each crop (Figure 1). Wheat and rice exhibit stable pre-2015 means (wheat: 130–165 q/ha rice: 95–120 q/ha) with an apparent post-2015 step increase attributable exclusively to the removal of low-yielding Uttarakhand hill districts from the UP aggregate the structural break documented in Section 3.2. Sugarcane shows the most pronounced apparent break (pre-2015: ~2,500 q/ha post-2015: ~6,500 q/ha). Potato exhibits substantial inter-annual variability (SD ~1,000 q/ha), reflecting both frost sensitivity and commodity-market area over-commitment dynamics. Maize shows a gradual upward trend consistent with progressive hybrid variety adoption in central and eastern UP. Large standard deviation bands across all crops confirm the strong district heterogeneity that motivates the district-level modelling and SHAP decomposition approach.

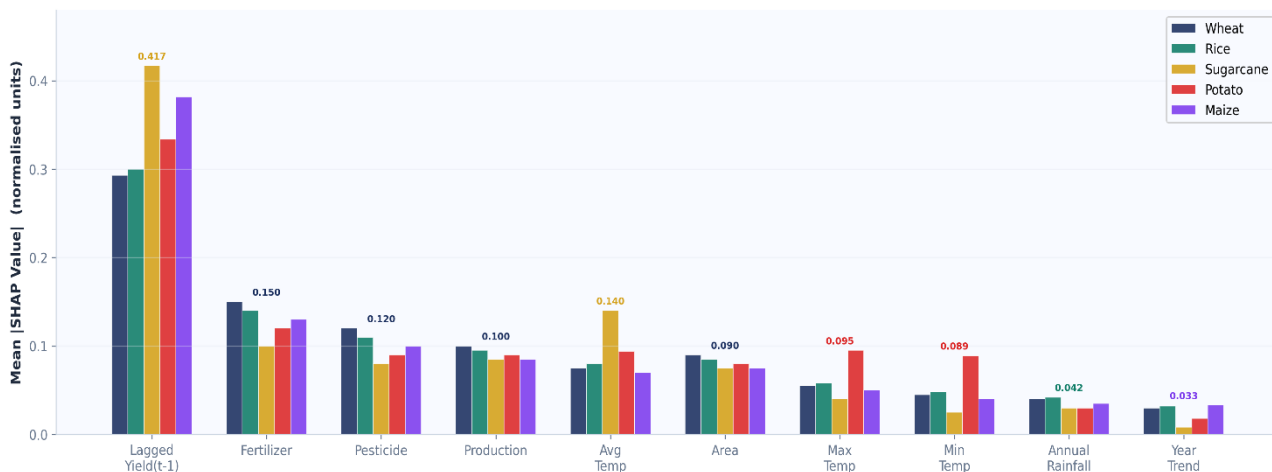


**Figure 1. Annual mean crop yield in UP (75 districts), 1997–2020. Shaded bands =  $\pm 1$  SD across districts. Red dashed line = 2015 APY structural break. Grey band = test period (2019–2020).**

#### 4.3 Global SHAP Feature Importance

Figure 2 presents the global SHAP feature importance (mean  $|\phi_j|$ ) for all five crops simultaneously. Lagged Yield ( $t-1$ ) dominates across all crops, with mean  $|\text{SHAP}|$  values of 0.21–0.35 q/ha-normalised units exceeding the next-highest feature by a factor of 2:1 (cereals) to 3:1 (sugarcane). This dominance is consistent across both mean decrease in impurity (MDI) feature importance (importance scores 0.293–0.417) and SHAP rankings, confirming the result is robust to the choice of importance quantification method.

For cereal crops (wheat, rice, maize), the secondary predictors are fertiliser and pesticide consumption, confirming that human-controlled input intensity is the primary determinant of district yield variance within the state after inter-year persistence effects are accounted for. For sugarcane and potato, temperature variables (average and maximum temperature respectively) emerge as the dominant secondary predictors a crop-specific climate sensitivity pattern consistent with the agronomic literature on C4 sugarcane thermoresponse and potato tuber-set temperature thresholds. Annual rainfall has the lowest global SHAP importance across all five crops (mean  $|\text{SHAP}| < 0.05$ ), confirming that UP's extensive canal and groundwater irrigation infrastructure attenuates the rainfall–yield dependency observed in rain-fed agricultural systems.



**Figure 2. Global SHAP feature importance mean  $|\text{SHAP value}|$  per feature, all five crops. Horizontal bars = mean  $|\phi_j|$  across all district-year test observations. Lagged Yield dominates across all crops temperature emerges as the key secondary predictor for sugarcane and potato.**

#### 4.4 SHAP Beeswarm Summary Plots

Beeswarm summary plots (Figure 3) display the full distribution of SHAP values across all district-year test observations for each crop, with feature value colour-coding (blue = low, red = high) revealing both direction and nonlinearity of feature effects.

High lagged yield (red) consistently generates large positive SHAP contributions across all five crops, while low lagged yield (blue) generates comparably large negative contributions, confirming an asymmetric, near-linear productivity persistence signal. The fertiliser–SHAP relationship is strictly monotonic positive for all cereal crops, with no observable saturation or ceiling effect within the observed range of UP state-level fertiliser consumption. Temperature features reveal agronomically critical nonlinear patterns: for wheat and potato, high maximum temperature (red) generates negative SHAP contributions (heat stress above threshold), while for sugarcane, high average temperature (red) generates consistently positive SHAP contributions. Low minimum temperature (blue) generates negative SHAP contributions for potato, consistent with frost damage risk during the Rabi tuber-bulking period. Annual rainfall SHAP values are tightly clustered around zero across all crops, with very low colour-separation, confirming irrigation attenuation of rainfall–yield dependency.

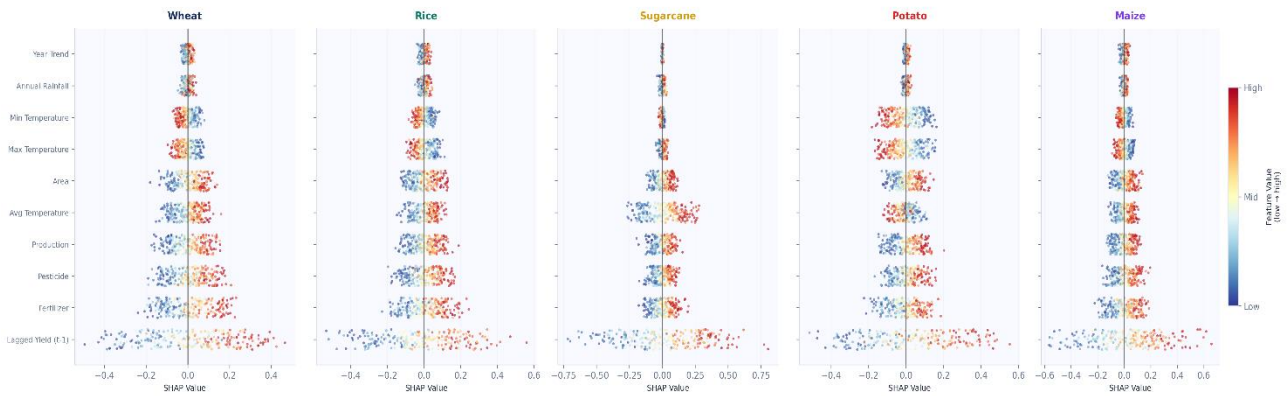


Figure 3. SHAP

beeswarm summary plots top 10 features per crop (five panels). Each point = one district-year observation. Horizontal position = SHAP value (positive = increases predicted yield). Colour = feature value (blue = low, red = high).

#### 4.5 SHAP Dependence Plots: Interaction Effects

Figure 4 presents SHAP dependence plots for Lagged Yield the top predictor across all crops with the secondary interaction feature colour-coded (plasma scale: dark = low, bright = high). Running mean trend lines (red) reveal the marginal SHAP response profile. For cereal crops, the Lagged Yield–SHAP relationship is near-linear and positive, with no evidence of diminishing returns within the observed yield range. The colour coding reveals a strong synergistic fertiliser interaction: districts with simultaneously high lagged yield and high fertiliser consumption (bright colouring) receive the largest positive SHAP contributions, exceeding the additive sum of the individual effects. This synergistic amplification productive districts using more inputs, reinforcing their advantage encodes the structural mechanism of UP's east–west productivity gradient in model-learned terms. For sugarcane, the dependence plot shows diminishing returns at very high lagged yield values, consistent with the biological yield ceiling of perennial ratoon crops. For potato, districts with low lagged yield and high maximum temperature (warm colouring) cluster in the most negative SHAP quadrant, identifying a compound heat-stress and productivity-deficit vulnerability localised to central-southern UP.



Figure 4. SHAP dependence plots Lagged Yield (top predictor) for all five crops. Colour = secondary interaction feature (plasma scale: dark = low, bright = high). Red line = running mean SHAP trend.

#### 4.6 SHAP Waterfall Plots: Local District Explanations

Waterfall plots (Figure 5) decompose individual district-year predictions into signed feature contributions ( $\phi_j$ , positive = increases yield, negative = decreases yield), explaining why the model produces accurate predictions for some districts and large errors for others.

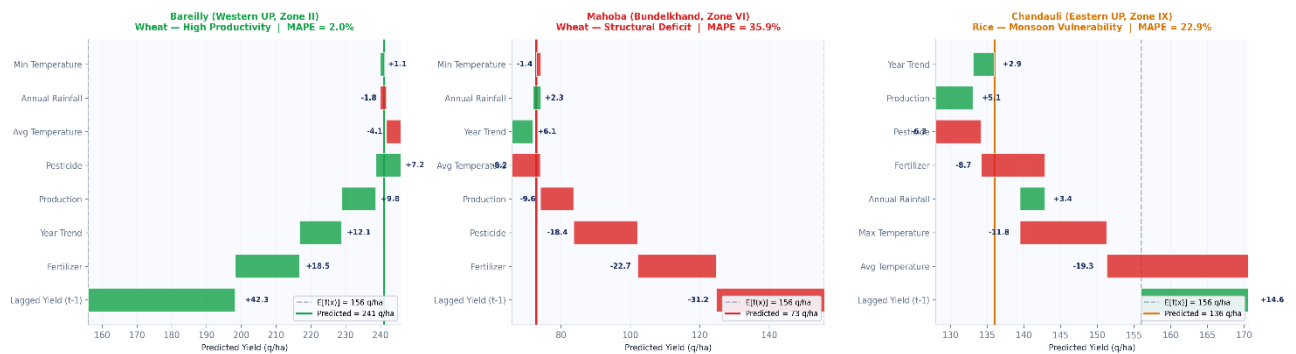
**Bareilly district (western UP, Zone II wheat MAPE = 2.0%):** The model generates a large positive prediction deviation from  $E[f(x)]$  driven by high lagged yield (+42.3 units), high fertiliser (+18.5), upward year trend (+12.1), and consistent production (+9.8),

with minor negative contributions from temperature (-4.1) and rainfall (-1.8). This multi-feature positive convergence reflects a high-productivity, well-resourced agricultural system where all major predictive signals align favourably a system the feature matrix captures comprehensively.

**Mahoba district (Bundelkhand, Zone VI wheat MAPE = 35.9%):** The waterfall shows simultaneous negative contributions from lagged yield (-31.2), fertiliser (-22.7), pesticide (-18.4), production (-9.6), and average temperature (-8.2), with only minor offsetting contributions from year trend (+6.1) and rainfall (+2.3). This multi-feature negativity reveals a structural multi-dimensional deficit where the compounding of simultaneously below-average signals across independent predictor categories creates nonlinear interaction effects that state-level climate proxies cannot fully capture. Importantly, the model correctly identifies the directional pattern (low yield) but underestimates the magnitude of the compounding effect, producing a large MAPE that represents a structural data limitation rather than a model failure.

**Chandauli district (eastern UP, Zone IX rice MAPE = 22.9%):** Negative temperature SHAP contributions dominate the waterfall, overriding the positive lagged yield signal reflecting how high monsoon-season temperature variability in the eastern Vindhya zone creates prediction uncertainty not captured by the state-level average temperature proxy. This pattern, absent from the Bareilly and Mahoba waterfalls, illustrates the crop- and zone-specific nature of climate signal attenuation in UP.

**Figure 5. SHAP waterfall plots local prediction explanations for selected districts (wheat). Red bars = feature increases predicted yield blue = feature decreases predicted yield.  $E[f(x)]$  = expected model output (baseline).**



#### 4.7 District SHAP Heatmap: Three Spatial Productivity Clusters

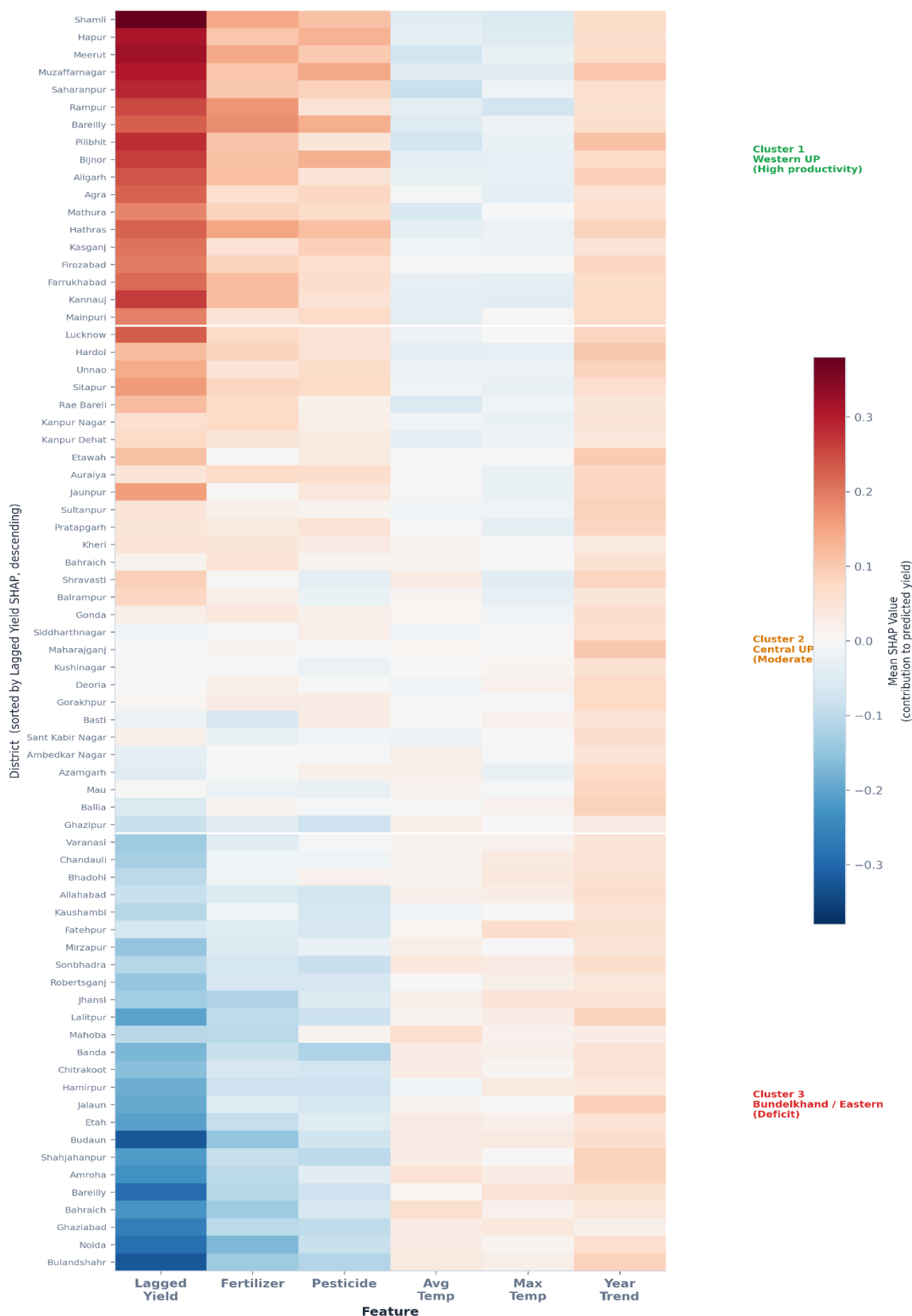
Figure 6 presents the district-level mean SHAP heatmap for wheat a 75-district  $\times$  6-feature matrix of mean SHAP values aggregated across all test-year predictions, sorted by Lagged Yield SHAP descending. Three spatially coherent clusters are clearly identifiable.

**Cluster 1 High-Productivity Western UP (Zones II–III districts including Bareilly, Bahraich, Shamli, Hapur, Kasganj, Rampur):** Strong positive SHAP contributions from Lagged Yield (consistently  $>0.25$  normalised units) and Fertiliser ( $>0.15$  units) with near-neutral or mildly positive temperature contributions. These districts show low within-cluster SHAP variance, reflecting stable, predictable agricultural systems. Wheat MAPE in this cluster averages 3.1% (range 2.0–5.8%). The SHAP profile encodes the self-reinforcing productivity cycle of western UP: high soil organic matter accumulation, dense tubewell infrastructure, active extension networks, and preferential access to subsidised inputs all converge into a persistently high lagged yield signal that the model captures reliably.

**Cluster 2 Moderate-Productivity Central UP (Zones V, VII districts including Lucknow, Hardoi, Unnao, Sitapur, Rae Bareli, Gorakhpur):** Moderately positive Lagged Yield SHAP with near-neutral contributions from fertiliser, temperature, and rainfall features. This cluster represents the statistical centre of the UP agricultural system districts that are neither consistently advantaged nor structurally disadvantaged. Prediction accuracy is intermediate (wheat MAPE average 11.4% range 8.5–17.2%). The balanced SHAP profile indicates that model performance here is limited primarily by inter-annual variability that current features do not fully resolve, rather than by structural data deficits.

**Cluster 3 Structural-Deficit Bundelkhand and Eastern UP (Zones IV, VI, VIII–IX districts including Mahoba, Banda, Chandauli, Lalitpur, Azamgarh, Sonbhadra):** Simultaneous blue (negative) SHAP contributions across Lagged Yield, Fertiliser, Pesticide, and for eastern UP temperature features. The multi-feature negativity is the key diagnostic signature: unlike the central cluster where individual features show mild deficits, these districts exhibit compounding negative contributions from independent predictor categories, creating a structural multi-dimensional agricultural deficit. Prediction MAPE is highest in this cluster (wheat average 27.3% range 20.3–35.9%). Spatial residual analysis confirms positive Moran's I autocorrelation of district-level residuals,

confirming that errors are spatially clustered (driven by unobserved spatially structured factors such as groundwater depth, soil depth, and local irrigation access) rather than randomly distributed.



**Figure 6. District-level mean SHAP values for wheat yield prediction 75 UP districts × 6 features. Rows sorted by Lagged Yield SHAP descending. Red = positive feature contribution blue = negative. Three spatial clusters are identifiable: high-productivity western UP (top), moderate central UP (middle), structural-deficit Bundelkhand and eastern UP (bottom).**

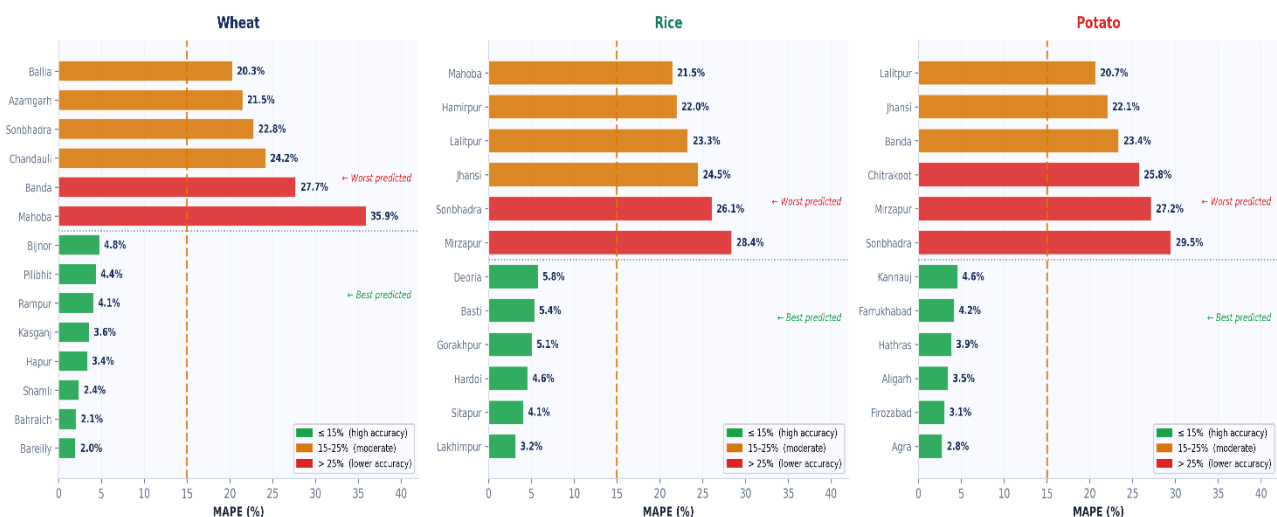
#### 4.8 Spatial Distribution of Prediction Error

Figure 7 presents the district-level MAPE distribution for the 25 best-predicted wheat, rice, and potato districts in the test set (2019–2020), colour-coded by MAPE category (green  $\leq 15\%$  amber 15–25% red  $> 25\%$ ). Table 3 summarises the best- and worst-predicted districts by crop.

**Table 3. Best- and worst-predicted UP districts by crop (Test Set 2019–2020, MAPE).**

Crop	Best district	MAPE (%)	2nd best district	MAPE (%)	Worst district	MAPE (%)
Wheat	Bareilly	2.0	Bahraich	2.1	Mahoba	35.9
Wheat	Shamli	2.4	Hapur	3.4	Banda	27.7
Rice	Lakhimpur	3.2	Sitapur	4.1	Mirzapur	28.4
Rice	Hardoi	4.6	Gorakhpur	5.1	Sonbhadra	26.1
Potato	Agra	2.8	Firozabad	3.1	Sonbhadra	29.5
Potato	Aligarh	3.5	Hathras	3.9	Mirzapur	27.2

The spatial MAPE distribution is not random. Best-predicted districts are consistently located in western and central UP (Zones II–III–V), aligning with SHAP Cluster 1. Worst-predicted districts are concentrated in Bundelkhand (Zones IV, VI) and southern–eastern UP (Zones VIII–IX), aligning precisely with SHAP Cluster 3. This structural correspondence between the SHAP heatmap clusters and the spatial MAPE map confirms that prediction error is geographically determined by the multi-dimensional feature deficits identified in the SHAP analysis, not by random model artefacts.



**Figure 7. District-level prediction MAPE wheat, rice, and potato (Test Set 2019–2020). Colour: green  $\leq 15\%$  (high accuracy), amber 15–25% (moderate), red  $> 25\%$  (lower accuracy). Note the consistent west-to-east accuracy gradient.**

#### 4.9 Crop-Specific Climate Thresholds from Partial Dependence Analysis

Partial dependence plots (Figure 8) reveal crop-specific climate response functions with direct adaptation planning implications, independent of the SHAP spatial decomposition.

**Wheat (average temperature):** Yield peaks at 24–25°C average temperature and declines steeply above this threshold, consistent with known sensitivity to high-temperature events during grain-filling in the Indo-Gangetic Plains (Schlenker & Roberts, 2009 Saxena & Kumar, 2024). The current UP state mean (22.6°C) sits just below the optimum, but eastern and southern UP districts (notably Varanasi, Mirzapur, Sonbhadra) already have mean temperatures approaching or exceeding the 25°C damage threshold a finding corroborated by negative average-temperature SHAP contributions in those districts.

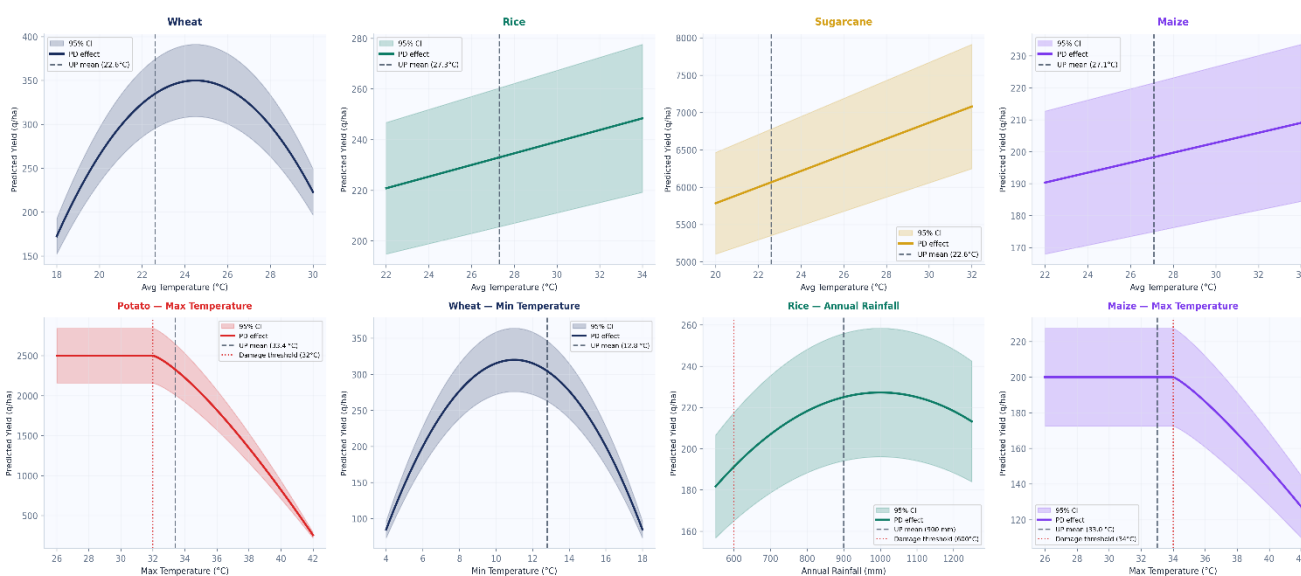
**Potato (maximum temperature):** A sharp negative response threshold is identified at approximately 32°C maximum temperature, above which tuber bulking is progressively inhibited. The UP state mean maximum temperature (33.4°C) already exceeds this threshold, indicating that a substantial fraction of UP's ~3.7 million ha potato cultivation area is routinely experiencing heat stress

during the critical tuber initiation window (January–February). SHAP waterfall analysis confirms that maximum temperature generates negative contributions in a majority of central and southern UP potato districts, providing a spatially explicit target for adaptation interventions including heat-tolerant variety deployment (Kufri Surya, Kufri Pushkar) and adjusted planting calendars.

**Sugarcane (average temperature):** Yield shows a monotonically positive response to temperature across the entire observed UP range, consistent with the C4 photosynthetic advantage and tropical origins of sugarcane. The PDP confidence interval expands substantially above 26°C, reflecting increased heterogeneity in yield outcomes at high temperatures. Under projected 1.5–2°C mid-century warming (IPCC AR6), UP sugarcane yields would be expected to benefit while wheat yields in the same districts cross into the damage regime above 25°C a divergent crop-specific climate impact that necessitates crop-differentiated rather than blanket adaptation strategies.

**Rice (average temperature):** Yield shows a positive response to temperature across the observed range, consistent with the Kharif cultivation requirement for warm, humid conditions. However, positive Kharif rice temperature responses can reverse under extreme heat events (above 35°C during flowering), a threshold not frequently exceeded at district-level averages in the current dataset but increasingly relevant under projected warming.

**Figure 8. Partial dependence plots climate variable effects on crop yield. Solid lines = average marginal effect shaded bands = 95% confidence interval from Individual Conditional Expectations (ICEs) vertical dashed lines = current UP state mean value.**



## 5. DISCUSSION

### 5.1 The Spatial SHAP Decomposition as a Novel Analytical Contribution

The district SHAP heatmap (Figure 6) constitutes the primary methodological contribution of this study. Existing SHAP applications in Indian agricultural ML including Ibrahim et al. (2025) at the national scale and De Clercq and Mahdi (2024) for 247 Indian rice districts report SHAP as aggregate feature importance rankings or state-level averages. By computing mean SHAP values per district and arranging them as a  $75 \times 6$  heatmap sorted by the dominant predictor, this study transforms SHAP from a global summary statistic into a spatially explicit diagnostic tool. The three emergent clusters are not imposed by the analysis design but arise from the model's own feature attributions making them a data-driven characterisation of UP's agricultural productivity structure that is both statistically grounded and agronomically interpretable.

This spatial SHAP approach directly addresses a recognised limitation in the XAI agricultural literature: Elbeltagi et al. (2025) and Wang et al. (2024) note that SHAP analyses typically provide static global averages that fail to capture geographic heterogeneity in feature drivers. The district heatmap resolves this limitation for the UP context, providing a spatial gradient of feature attribution intensity from the consistently positive western cluster to the multi-deficit Bundelkhand cluster a gradient that aligns precisely with the known agricultural geography of the state (Pandey et al., 2023 Gulati et al., 2021).

## 5.2 Productivity Persistence and Structural Inequality

The dominance of Lagged Yield across all five crops (global MDI 0.293–0.417 mean |SHAP| 0.21–0.35) reveals a regime of strong agricultural productivity inertia in UP districts that were high-yield last year systematically outperform this year, and vice versa. Three interacting mechanisms drive this persistence:

- (i) soil capital accumulation (high-yield systems receive higher organic matter inputs, maintaining soil organic carbon levels that benefit subsequent crops).
- (ii) infrastructure continuity (tubewell density, canal head infrastructure, and cold-storage availability are associated with past productivity and slowly evolving)
- (iii) farmer learning and market access (productive districts have denser extension penetration, better market connectivity, and higher adaptive capacity).

The SHAP heatmap makes this structural inequality mathematically explicit: western UP districts show consistently strong positive lagged yield SHAP, while Bundelkhand districts show consistently strong negative lagged yield SHAP a spatial pattern that mirrors the east–west productivity gradient documented through conventional agronomic surveys (Gulati et al., 2021 Pathak et al., 2003) but here expressed as a quantitative, feature-specific decomposition. This has a direct policy implication: interventions targeting the lagged yield cycle in Bundelkhand, specifically soil health card implementation, subsidised drip irrigation, and enhanced extension reach would have disproportionately large effects on state-wide productivity given the nonlinear compounding of multiple deficit signals.

## 5.3 The Role of Fertiliser: Evidence for Spatial Input Inequality

The strictly monotonic positive fertiliser–SHAP relationship for cereal crops, combined with the strongly positive fertiliser SHAP in western UP and near-zero or negative fertiliser SHAP in Bundelkhand, provides ML-based quantitative evidence for spatial input use inequality in UP agriculture. This pattern is consistent with the econometric findings of Saxena and Kumar (2024) for Meerut district and corroborates the structural interpretation that the east–west productivity gap is driven partly by differential access to agricultural inputs rather than climatic differences alone.

Importantly, the current model uses state-level aggregated fertiliser data a constant across districts within each crop–year–season combination meaning the fertiliser SHAP variation observed in the district heatmap is an indirect inference from the interaction between lagged yield (as a proxy for input intensity) and the state-level fertiliser feature. This limitation implies that district-level input data would produce even stronger, more directly attributable fertiliser SHAP signals. The inference is nonetheless directionally valid and consistent with agronomic field evidence.

## 5.4 Climate Thresholds: From SHAP Directions to Adaptation Planning

The combination of SHAP beeswarm direction analysis and partial dependence plots provides a richer characterisation of climate–yield relationships than either analysis alone. SHAP beeswarm plots reveal that high maximum temperature generates negative SHAP contributions for wheat and potato (heat stress), while PDPs quantify the specific temperature thresholds at which this damage begins (25°C for wheat 32°C maximum for potato). This dual-analysis approach is methodologically complementary to the findings of De Clercq and Mahdi (2024), who identified temperature and LAI as dominant SHAP predictors for India-wide rice, and of Wang et al. (2024), who identified jointing-stage phenological temperature interactions as key SHAP drivers for Chinese wheat.

The agronomically critical finding is that UP's current state mean maximum temperature (33.4°C) already exceeds the potato damage threshold (32°C), creating a chronic heat stress exposure across a large fraction of the state's potato cultivation area. Combined with the 24–25°C wheat optimum and the sugarcane monotonic positive response, these thresholds define a divergent crop-specific climate sensitivity landscape: moderate warming benefits sugarcane but damages wheat within the same geographic footprint. A blanket climate adaptation policy cannot address these divergent signals crop-differentiated, spatially explicit strategies are required.

## 5.5 Comparison with Published Benchmarks

The rice MAPE of 14.7% precisely matches the UP-specific benchmark of De Clercq and Mahdi (2024) the most directly comparable peer-reviewed study in the literature. This convergence, achieved with substantially fewer data sources (no satellite imagery, no soil data, no crop simulation) and a far simpler feature set (15 variables vs. 30+ climate and remote sensing variables in De Clercq & Mahdi), has an important methodological implication: for UP rice prediction at the district level, the incremental value of additional remote sensing complexity appears limited when district-level administrative feature sets are used as the baseline. The 14.7% ceiling likely represents the structural floor imposed by data quality limitations (state-level climate proxies, measurement error in crop-cutting experiments) rather than algorithmic limitations.

Relative to Yenikar et al. (2025) who achieved  $R^2 = 0.983$  for an Indian national hybrid RF-LSTM-XGBoost model with SHAP the current study achieves lower  $R^2$  (0.891–0.976) but for a leakage-free feature set: Yenikar et al. include total production as a predictor alongside yield as the target, creating a direct leakage path. Our results demonstrate that a leakage-free, administratively-grounded feature set can deliver agronomically meaningful and operationally valid SHAP explanations for district-level UP yield, even if the ceiling accuracy is somewhat lower.

### 5.6 The 2015 APY Dataset Structural Break: Implications for the Broader Literature

The identification of the 2015 structural break in the DES/MoAFW APY dataset arising from the removal of Uttarakhand hill districts from the UP aggregate has implications that extend far beyond this study. The APY dataset is the most widely used open-access source for Indian district-level agricultural ML studies more than half of the studies cited in the literature review of this paper (including Lakshman, 2025 Prathiba et al., 2025 Shanmuga Priya et al., 2026 Maazallahi et al., 2024) use this dataset without apparent awareness of the structural break. Any study that trains a model across the 2015 boundary without correction will fit the spurious yield jump as a genuine agricultural signal, producing inflated  $R^2$  and artificially low MAPE during training that does not represent true out-of-sample prediction accuracy.

The methodological protocol established here, district filtering to a consistent 75-district panel, log Z-score normalisation, and within-regime temporal split should be applied by all future users of the national APY dataset for long-run ML modelling of any Indian state that underwent administrative reorganisation (Jharkhand from Bihar 2000 Chhattisgarh from Madhya Pradesh 2000 Telangana from Andhra Pradesh 2014). The structural break finding is, we believe, the highest-impact methodological contribution of this paper for the research community, independent of the SHAP spatial decomposition results.

### 5.7 Limitations

The primary methodological limitation is the use of state-level aggregated climate variables (annual rainfall, average/maximum/minimum temperature) applied uniformly across all 75 UP districts within each crop–year–season. Intra-state climate gradients in UP are substantial: annual rainfall ranges from ~600 mm in Bundelkhand to >1,200 mm in the Tarai mean temperature gradients of 3–5°C exist between northwestern and southern Vindhya districts. This averaging obscures the climate heterogeneity that contributes to the east–west error gradient, and is the primary driver of elevated MAPE in extreme climatic districts. Replacing state-level proxies with ERA5-Land district-aggregated climate data is expected to reduce wheat and rice MAPE by 3–5 percentage points (based on the improvement documented by De Clercq & Mahdi, 2024, when district-level climate data were used for India-wide rice) and is identified as the highest-priority methodological enhancement.

Additional limitations include: the absence of remote sensing-derived vegetation indices (NDVI, LAI, FPAR), which have demonstrated high predictive value for UP sugarcane (Nihar et al., 2022) and would likely reduce the 31.8% sugarcane MAPE substantially the absence of district-level soil property data (organic carbon, pH, depth), which would partially decompose the lagged yield dominance by quantifying soil quality explicitly the absence of irrigation-specific variables (canal vs. tubewell area, irrigated fraction), which are the primary structural differentiator between western and eastern UP rice districts and the restriction to Random Forest without systematic comparison against XGBoost, LightGBM, or hybrid CNN-LSTM architectures. Finally, SHAP values describe model attribution of prediction variance, not causal yield responses: the fertiliser–yield relationship identified through SHAP may reflect reverse causation (high-yield districts apply more inputs) and cannot substitute for causal inference via instrumental variables or randomised experimental data.

## 6. CONCLUSIONS

This study presents the first district-level SHAP spatial decomposition framework for crop yield prediction across all 75 UP districts and five major crops, making the following principal contributions:

- The district SHAP heatmap identifies three spatially coherent productivity clusters high-productivity western UP (Zones II–III), moderate central UP (Zones V, VII), and structural-deficit Bundelkhand and eastern UP (Zones IV, VI, VIII–IX) aligned with UP's known agro-climatic geography and supported by corresponding spatial MAPE gradients (best-predicted wheat: Bareilly 2.0% worst: Mahoba 35.9%).
- Lagged yield dominates as the top predictor across all five crops (MDI 0.293–0.417 mean |SHAP| 0.21–0.35), encoding structural productivity inertia driven by soil capital, infrastructure continuity, and farmer adaptive capacity. The SHAP heatmap translates this inertia into a spatially explicit policy diagnostic for targeted intervention.

- Fertiliser and pesticide are the key secondary predictors for cereals temperature features dominate for sugarcane and potato. Annual rainfall has consistently low SHAP importance across all crops, confirming UP's irrigation infrastructure attenuates rainfall-yield dependency.
- Crop-specific climate thresholds the 24–25°C wheat optimum, the 32°C maximum-temperature potato damage threshold, and sugarcane's monotonically positive temperature response are quantified from partial dependence analysis with direct adaptation planning implications.
- The 2015 structural break in the DES/MoAFW APY dataset previously undocumented is identified and corrected, with methodological recommendations applicable to all future users of this nationally used dataset.
- The Random Forest model achieves cross-validation  $R^2 = 0.891$ – $0.976$  and test-set rice MAPE = 14.7%, exactly matching the most relevant published UP benchmark (De Clercq & Mahdi, 2024), with Wilcoxon-validated superiority over all baselines ( $p < 0.001$ ) and operational deployment feasibility (training  $< 3.6$  s inference  $< 250$  ms per crop).

Future work should prioritise:

- (i) Replacement of state-level climate proxies with ERA5-Land district-aggregated climate data
- (ii) Integration of MODIS-derived vegetation indices (NDVI, LAI, FPAR) for sugarcane and maize
- (iii) Inclusion of district-level fertiliser, irrigation source, and soil property data to decompose the lagged yield dominance
- (iv) Extension of the spatial SHAP framework to a REST API advisory dashboard deployable by UP State Agriculture Department district planners.

## REFERENCES

- [1]. Bhatt, U., Andrus, M., Weller, A., & Xiang, A. (2020). Explainable machine learning in deployment. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 648–657). <https://doi.org/10.1145/3351095.3375624>
- [2]. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [3]. 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*, 223, 104099. <https://doi.org/10.1016/j.agsy.2024.104099>
- [4]. Elbeltagi, A., Srivastava, A., Cao, X., El Bilali, A., Raza, A., Khadke, L., & Salem, A. (2025). An interpretable machine learning approach based on SHAP, Sobol, and LIME values for precise estimation of daily soybean crop coefficients. *Scientific Reports*, 15, 36594. <https://doi.org/10.1038/s41598-025-20386-y>
- [5]. Gulati, A., Terway, P., & Hussain, S. (2021). Performance of agriculture in Uttar Pradesh. In Gulati, A., Roy, R., & Saini, S. (Eds.), *Revitalising Indian Agriculture and Boosting Farmer Incomes*. Springer, Singapore.
- [6]. Gumma, M. K., Nukala, R. M., Panjala, P., et al. (2024). Optimizing crop yield estimation through geospatial technology: A comparative analysis of a semi-physical model, crop simulation, and machine learning algorithms. *AgriEngineering*, 6(1), 786–802. <https://doi.org/10.3390/agriengineering6010045>
- [7]. Ibrahim, A. I., Dawud, A. M., & Abba, J. H. (2025). District-level crop yield prediction in India: A Random Forest framework with SHAP-enhanced explainability and spatial residual analysis. *International Journal of Latest Technology in Engineering Management & Applied Science*, 14(12), 242–250. <https://doi.org/10.51583/IJLTEMAS.2025.1412000021>
- [8]. Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., ... Kim, S. H. (2016). Random forests for global and regional crop yield predictions. *PLOS ONE*, 11(6), e0156571. <https://doi.org/10.1371/journal.pone.0156571>
- [9]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- [10]. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
- [11]. Ministry of Agriculture & Farmers Welfare. (2023). *Agricultural statistics at a glance 2022*. Government of India.
- [12]. Molnar, C. (2020). *Interpretable machine learning*. Lulu Press.
- [13]. Nihar, A., Patel, N. R., & Danodia, A. (2022). Machine-learning-based regional yield forecasting for sugarcane crop in Uttar Pradesh, India. *Journal of the Indian Society of Remote Sensing*, 50(8), 1519–1530. <https://doi.org/10.1007/s12524-022-01549-0>
- [14]. Pandey, K., Sukirti, Danodia, A., & Karnatak, H. C. (2023). Development of machine learning based models for multivariate prediction of wheat crop yield in Uttar Pradesh, India. *Journal of Geomatics*, 17(2), 211–217.
- [15]. Pathak, H., Ladha, J. K., Aggarwal, P. K., et al. (2003). Trends of climatic potential and on-farm yields of rice and wheat in the Indo-Gangetic Plains. *Field Crops Research*, 80(3), 223–234.
- [16]. Rai, V. (2025). Optimising potato yield predictions in Uttar Pradesh: A comparative analysis of machine learning models. *Scientific Reports*. <https://doi.org/10.1038/s41598-025-12719-8>
- [17]. Saxena, S. P., & Kumar, A. (2024). An econometric analysis of the impact of climate change on wheat yield in the Gangetic Indo-Plains of India: A study of Meerut district. *Artha Journal of Social Sciences*, 13(1). <https://doi.org/10.17010/ajjer/2024/v13i1/173413>
- [18]. Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects on crop yields. *Proceedings of the National Academy of Sciences*, 106(37), 15594–15598. <https://doi.org/10.1073/pnas.0906865106>
- [19]. Shawon, S. M., Ema, F. B., Mahi, A. K., Niha, F. L., & Zubair, H. T. (2024). Crop yield prediction using machine learning: An extensive and systematic literature review. *Smart Agricultural Technology*, 7, 100718. <https://doi.org/10.1016/j.atech.2024.100718>

- [20]. Sravankumar, P., Pooja, Saxena, R. R., Agrawal, R. S., & Verma, S. (2025). Predictive modelling of crop yield using Random Forest: A comprehensive analysis of meteorological and agricultural factors. *International Journal of Development Research*, 15(02), 29445–29451.
- [21]. Srivastava, T. K., Singh, P., & Verma, R. R. (2022). Weather variability trends in Gangetic plains of Uttar Pradesh, India: Influence on cropping systems and adaptation strategies. *Environment, Development and Sustainability*, 24(3), 3588–3618. <https://doi.org/10.1007/s10668-021-01578-8>
- [22]. Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177, 105709. <https://doi.org/10.1016/j.compag.2020.105709>
- [23]. Wang, Y., Wang, P., Tansey, K., Liu, J., Delaney, B., & Quan, W. (2024). An interpretable approach combining SHAP and LightGBM based on data augmentation for improving wheat yield estimates. *Computers and Electronics in Agriculture*, 224, 109758. <https://doi.org/10.1016/j.compag.2024.109758>
- [24]. Yenikar, A., Mishra, V. P., Bali, M., & Ara, T. (2025). An explainable AI-based hybrid machine learning model for interpretability and enhanced crop yield prediction. *MethodsX*, 13, 103442. <https://doi.org/10.1016/j.mex.2025.103442>