engineering proceedings

MDPI

Proceeding Paper

Interpretable Machine Learning Framework for Early Potato Yield Prediction Using Agro-Meteorological Data †

MD Jiabul Hoque *, Tanjila Abedin Tajin and Tashfia Sultana

Department of Computer & Communication Engineering, International Islamic University, Chittagong 4318, Bangladesh; e233438@ugrad.iiuc.ac.bd (T.A.T.); e233442@ugrad.iiuc.ac.bd (T.S.)

* Correspondence: jiabul.hoque@iiuc.ac.bd

† Presented at the 13th International Electronic Conference on Sensors and Applications, 18–20 November 2026; Available online: https://sciforum.net/event/ECSA-13.

Abstract

Forecasting crop yield at an early stage of the growing season is crucial to reinforcing climate-smart agriculture and enhancing food security in nations with high vulnerability, including Bangladesh. This study creates an interpretable, ML-based framework for the early forecasting of potato yield and climate-driven recommendations using historical crop modelling, field production, climatic factors, and agro-meteorological data. Monthly data for rainfall and temperature were combined with yearly potato yield values covering the 1970/71 to 2024/25 agricultural years. The dataset was augmented through agronomically relevant approaches, such as noise injection, masking, mixup, pattern mixing, and shuffling, to create 21,440 records from the original 2144 observations. The proposed framework consists of two complementary components. Prediction using KNN, SVR, and RF; and recommendation through RF feature importance, ANOVA F regression analysis to identify influential weather variables, and Pearson and Spearman correlation analysis with KDE to identify optimal climatic conditions. The analysis confirms that yield is highly sensitive to stage-specific climatic conditions in terms of tuber initiation and tuber bulking. Decadal climate indices indicated that rainfall in December, October, and January made large contributions (34%, 32%, and 24%, respectively) to yield variability, while excess rainfall had generally negative associations with yield. The ideal climatic window for high yield was low rainfall and cool temperatures, especially in November and December. From the three aforementioned models, RF provided the best predictive performance with a mean MAE of 0.0803 ton_ha^{-1} and RMSE of 0.1014 ton_ha^{-1} as compared to KNN and SVR. This provides a tangible framework for decision support that is interpretable and can be used to facilitate irrigation planning, climate risk mitigation, and precision agriculture driven by IoT networks in Bangladesh and comparable agro-climatic regions.

Keywords: early potato yield prediction; machine learning; random forest; agro-meteorological data; climate-smart agriculture; precision agriculture

Check for updates

Academic Editor: Francisco Falcone

Published: 29 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Agricultural production is increasingly influenced by climate variability and changing environmental conditions [1]. Variations in rainfall and temperature can affect crop establishment, growth, productivity, and final yield, making reliable early-season forecasting important for agricultural planning, food-security preparedness, resource allocation, and climate-risk management [2,3].

Potato is an important food and commercial crop whose productivity is strongly influenced by climatic conditions throughout its growth cycle [4,5]. Rainfall and temperature affect crop establishment, vegetative development, tuber initiation, tuber bulking, and maturation [6,7]. However, these effects are stage-dependent; a climatic condition that is favourable during one stage may be unfavourable during another. Consequently, temporally resolved agro-meteorological variables may provide more meaningful predictive information than aggregated seasonal indicators [8,9].

Machine learning offers considerable potential for crop-yield prediction because climate–yield relationships often involve nonlinear responses, interactions, and threshold effects. KNN, SVR, and RF represent complementary ML approaches for modelling such relationships [10–14]. However, predictive accuracy alone is insufficient for practical agricultural decision-making [15,16]. Agricultural stakeholders require information not only about the amount of yield expected but also about which environmental variables influence yield, whether their effects are positive or negative, and which climatic conditions are associated with favourable productivity [17,18].

To address this need, the present study develops an integrated framework containing two complementary components. The first component predicts potato yield using agro-meteorological information. The second extracts interpretable climate-based knowledge through feature importance, statistical correlation, and density-based analysis.

The main contributions of this study are therefore:

1. A structured agro-meteorological ML pipeline integrating long-term climate and potato-yield information through extraction, transformation, preprocessing, augmentation, and feature engineering.
2. A dual analytical framework containing parallel crop-yield prediction and crop-yield recommendation systems.
3. An interpretable feature-identification approach combining RF feature importance with ANOVA F-regression to identify influential climatic drivers.
4. A climate-impact analysis using Pearson and Spearman correlations to examine both linear and monotonic relationships between monthly weather variables and potato yield.
5. A high-yield climatic recommendation mechanism using KDE to identify monthly rainfall and temperature conditions associated with favourable yield outcomes.
6. A comparative ML evaluation of KNN, SVR, and RF using optimized hyperparameters, cross-validation, and independent performance evaluation.

By combining prediction with climate-based interpretation, the proposed framework extends beyond conventional yield forecasting and provides a foundation for future climate-smart and IoT-enabled agricultural decision-support systems.

The remainder of the paper is organized as follows. Section 2 describes the data and proposed methodology. Section 3 presents and discusses the analytical and predictive results. Finally, Section 4 summarizes the principal findings, limitations, and future research directions.

2. Methodology

2.1. Overall Framework

Figure 1 presents the overall workflow of the proposed framework. The system follows a sequential pipeline consisting of data collection and ETL, preprocessing, data augmentation and feature engineering, parallel recommendation and prediction analysis, hyperparameter optimization, model evaluation, and final model selection.

https://doi.org/10.3390/engproc2026163053

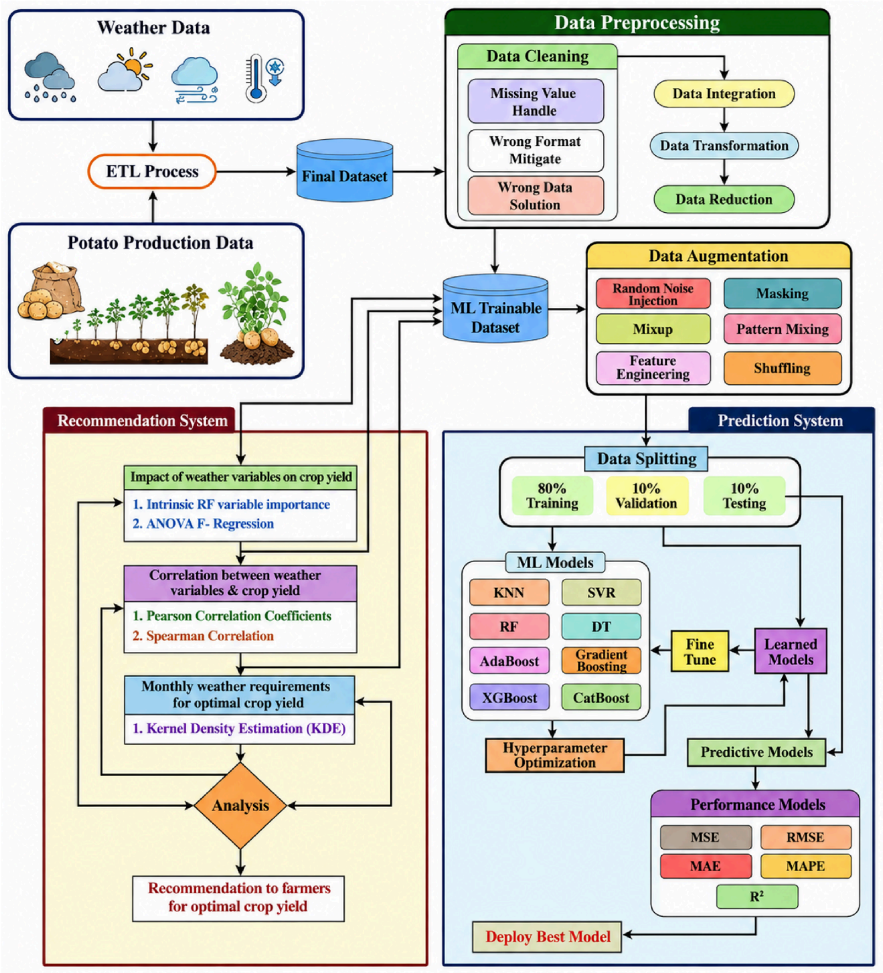


Figure 1. Workflow of the proposed interpretable potato-yield prediction and recommendation framework.

The framework is designed to preserve the temporal and agronomic relevance of agro-meteorological information while improving the robustness of ML-based yield prediction. Its final output contains both a predicted potato yield and interpretable climate-based recommendations.

2.2. Data Collection and ETL Process

The study integrates two temporally aligned data streams covering the agricultural years 1970/71–2024/25. The first consists of agro-meteorological observations obtained from the Bangladesh Meteorological Department (BMD) and the Bangladesh Climate Information Management System (CIMS). The selected variables are monthly total rainfall R_m (mm) and monthly mean air temperature, T_m (°C) for the four potato-growing months: $m \in \{1,2,3,4\} \equiv \{Oct, Nov, Dec, Jan\}$. The second data stream consists of potato-production information obtained from the Bangladesh Bureau of Statistics (BBS) and the Bangladesh Agricultural Research Institute (BARI). Potato yield was calculated as:

$$Y = \frac{P}{A} \tag{1}$$

where P denotes total production (tons) and A denotes harvested area (ha).

The integrated dataset is represented as:

$$D_{final} = \left\{ \left(X^{(y)}, Y^{(y)} \right) \right\}_{y=1}^{N} \tag{2}$$

The study produced a final dataset of 2144 observations shown in Table 1. The reported potato yield ranged from 7.83 to 19.89 t ha^{-1}, with an overall mean of 12.04 t ha^{-1}. The dataset also exhibited substantial monthly and interannual climatic variability, providing the basis for learning climate–yield relationships.

Table 1. Statistical summary of the dataset.

	rain_1m	rain_2m	rain_3m	rain_4m	temp_1m	temp_2m	temp_3m	temp_4m	Yield
count	2144	2144	2144	2144	2144	2144	2144	2144	2144
mean	34.62	9.01	7.37	19.87	23.58	19.63	18.36	20.81	12.04
Sd.	37.38	12.65	6.90	13.98	0.68	0.73	0.73	1.01	3.20
min	0.00	0.00	0.00	0.00	22.02	17.95	16.55	18.50	7.83
25%	6.46	0.51	1.10	7.13	23.12	19.07	17.84	20.04	9.65
50%	21.44	2.95	4.84	18.72	23.60	19.66	18.33	20.73	10.98
75%	49.711	15.73	11.91	29.67	24.09	20.23	18.91	21.47	13.70
max	141.19	44.58	23.68	51.10	25.39	21.31	20.61	23.89	19.89

2.3. Data Preprocessing

The integrated dataset was processed through a structured preprocessing pipeline, such as Missing value handling (Equation (3)), Outlier mitigation (Equation (4)), Temporal and spatial alignment (Equation (5)), Normalization (Equation (6)), and Variance stabilization and feature screening (Equation (7)) to improve data quality, numerical stability, and learning efficiency.

$$x_j^{(y)} \leftarrow \begin{cases} \mu_j, \ if\ x_j^{(y)} is\ missing \\ x_j^{(y)}, Otherwise \end{cases} \tag{3}$$

$$x_j^{(y)} \leftarrow min\left(max\left(x_j^{(y)}, Q_{1,j} - 1.5QR_j\right), Q_{3,j} + 1.5IQR_j\right) \tag{4}$$

$$\tilde{x}_{j,m}^{(y)} = \frac{1}{S}\sum\nolimits_{s=1}^{S} x_{j,m,s}^{(y)} \tag{5}$$

$$x_{j,norm}^{(y)} = \frac{x_j^{(y)} - min\left(x_j\right)}{max\left(x_j\right) - min\left(x_j\right)} \tag{6}$$

$$x_j^{(y)} \leftarrow \log\left(1 + x_j^{(y)}\right) \tag{7}$$

2.4. Data Augmentation and Feature Engineering

To enhance robustness, mitigate overfitting, and improve learning under climatic variability, the preprocessed dataset D_{Prep} was expanded using six complementary operations such as Random Noise Injection (Equation (8)), Masking (Equation (9)), Mixup (Equations (10) and (11)), Pattern Mixing (Equation (12)), Shuffling (Equation (13)), and Feature Engineering (Equations (14)–(17)).

$$\tilde{X}^{(i)} = X^{(i)} + \epsilon, \epsilon \sim \aleph\left(O, \sigma^2 I\right) \tag{8}$$

$$\tilde{x}_j^{(i)} = \begin{cases} 0, \ with\ probability\ \pi \\ x_j^{(i)}, \ otherwise \end{cases} \tag{9}$$

$$\tilde{X} = \lambda X^{(i)} + (1 - \lambda) X^{(k)} \tag{10}$$

$$\tilde{y} = \lambda y^{(i)} + (1 - \lambda) y^{(k)} \tag{11}$$

$$\tilde{X} = \left[X_{M_1}^{(i)}, X_{M_2}^{(k)}\right] \tag{12}$$

$$D \leftarrow Shuffle(D) \tag{13}$$

$$R_{cum} = \sum_{m=1}^{4} R_m \tag{14}$$

$$\overline{T} = \frac{1}{4}\sum_{m=1}^{4} T_m \tag{15}$$

$$\sigma_T^2 = \frac{1}{4}\sum_{m=1}^{4}\left(T_m - \overline{T}\right)^2 \tag{16}$$

$$\emptyset_m = R_m \cdot T_m \tag{17}$$

These derived variables provide additional representations of nonlinear climate–yield relationships without altering the underlying agricultural information. After the sequential augmentation and feature-engineering procedures, the dataset was expanded to 21,440 ML-ready records.

$$D_{ml} = \left\{\left(X_{aug}^{(i)}, y_{aug}^{(i)}\right)\right\}_{i=1}^{21,440} \tag{18}$$

2.5. Crop-Yield Recommendation System

The recommendation system operates as an interpretable knowledge-extraction layer complementary to the predictive model. It transforms model-ready climate data into information regarding influential weather variables, their relationships with yield, and climatic conditions associated with favourable productivity. The system consists of three connected analytical components.

2.5.1. Identification of Influential Weather Variables

RF feature importance (Equation (19)) was used to capture nonlinear variable relevance and interactions. The importance of feature x_j was derived from its contribution to variance reduction across the RF ensemble. To complement the nonlinear RF analysis, ANOVA F-regression (Equation (20)) was applied to evaluate the statistical explanatory relevance of individual features. Variables ranked highly by both approaches were treated as key climate–yield drivers and passed to the subsequent interpretive analyses. This combined approach reduces dependence on a single feature-selection mechanism and captures complementary nonlinear and statistical perspectives.

$$Imp_{RF}\left(x_j\right) = \frac{1}{T}\sum_{t=1}^{T}\sum_{n \in N_{j,t}} \Delta I(n,t) \tag{19}$$

$$F_j = \frac{SSR_j/1}{SSE_j/(N-2)} \tag{20}$$

2.5.2. Analysis of Weather-Variable Impacts

Pearson (Equation (21)) and Spearman (Equation (22)) correlation coefficients were calculated for each selected climatic variable. The combined analysis enables identification of positive or negative relationships and comparison between linear and monotonic climate–yield patterns. This is particularly relevant to agro-meteorological data, where relationships may be nonlinear, affected by thresholds, or influenced by non-normal distributions.

$$r_j = \frac{\sum_{i=1}^{N}\left(x_j^{(i)} - \overline{x}_j\right)\left(y^{(i)} - \overline{y}\right)}{\sqrt{\sum_i\left(x_j^{(i)} - \overline{x}_j\right)^2}\sqrt{\sum_i\left(y^{(i)} - \overline{y}\right)^2}} \tag{21}$$

$$\rho_j = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2-1)} \tag{22}$$

2.5.3. Monthly Climatic Conditions Associated with High Yield

For each climatic variable, KDE (Equation (24)) was applied after selecting High-yield observations (Equation (23)). The resulting density peaks identify climatic ranges most frequently associated with high-yield observations. These ranges are subsequently translated into interpretable recommendations, including potential irrigation targets and climate-risk alerts.

$$\mathcal{H} = \left\{ x_j^{(i)} \middle| y^{(i)} \geq \tau \right\} \tag{23}$$

$$\hat{f}_j(x) = \frac{1}{|\mathcal{H}|h} \sum_{x^{(i)} \in \mathcal{H}} K\left(\frac{x - x^{(i)}}{h} \right) \tag{24}$$

2.6. Crop-Yield Prediction System

The predictive component evaluates three regression models: $M = \{KNN, SVR, RF\}$. KNN (Equation (25)) predicts yield from observations with similar agro-meteorological characteristics. SVR (Equations (26) and (27)) provides a kernel-based nonlinear regression approach, while RF (Equation (28)) uses an ensemble of decision trees to capture nonlinear relationships, interactions, and threshold effects.

$$\hat{y}(x) = \frac{1}{k} \sum_{i \in N_k(x)} y^{(x)} \tag{25}$$

$$f(x) = w^T \emptyset(x) + b \tag{26}$$

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \tag{27}$$

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f_t(x) \tag{28}$$

The ML-ready dataset was divided according to the experimental design into: 70% training data; 15% validation data; and 15% testing data. For each model, hyperparameter configurations (Equation (29)) were evaluated using the validation procedure. The configuration with the lowest validation error was retained.

$$\Theta^* = arg\min_{\Theta} \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}\left(y^{(k)}, \hat{y}_{\Theta}^{(k)} \right) \tag{29}$$

The optimized models were subsequently retrained using the combined training and validation data and evaluated on the test data using:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{30}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{31}$$

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{32}$$

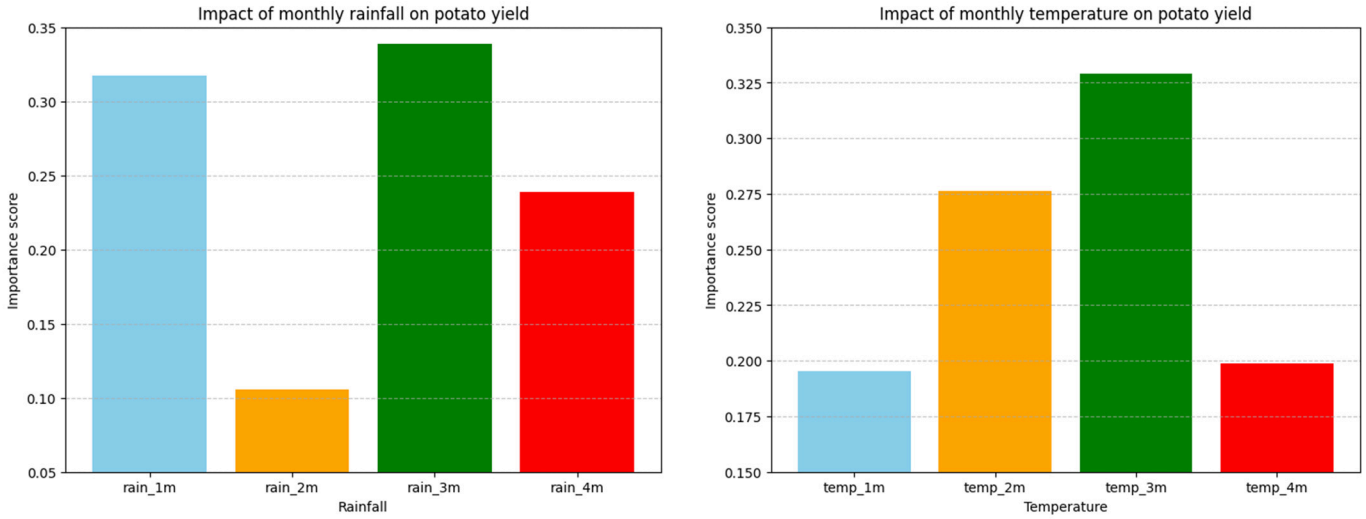
$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \tag{33}$$

The final predictor was selected by minimizing MAE, RMSE, and MAPE while maximizing R^2.

3. Results and Discussion

3.1. Important Climatic Drivers of Potato Yield

The combined feature-importance analyses (shown in Figure 2) indicate that the influence of climate variables varies substantially across the potato-growing period. RF analysis (Figure 2a) identified December rainfall as the most influential rainfall variable, followed by October rainfall and January rainfall, while November rainfall showed comparatively lower importance. The reported contribution pattern highlights the importance of stage-specific precipitation conditions. Temperature variables (Figure 2b) generally demonstrated more moderate contributions. However, November and December temperatures showed greater relevance than the temperature variables corresponding to the earlier and later stages.



(**a**) (**b**)

Figure 2. Feature importance of monthly rainfall (**a**) and temperature (**b**) variables based on RF and ANOVA F-regression. Different colours are used to distinguish the individual monthly climatic variables; each colour corresponds to the respective month shown on the horizontal axis.

The agreement and differences between RF feature importance and ANOVA F-regression provide useful complementary information. RF can identify nonlinear relationships and interactions, whereas ANOVA evaluates individual statistical explanatory relevance. Their combined use therefore strengthens the identification of influential climatic drivers.

3.2. Rainfall–Yield Relationships

The correlation analysis (illustrated in Table 2) indicates that rainfall was generally negatively associated with potato yield. October rainfall exhibited the strongest negative Pearson relationship among the monthly rainfall variables ($r = -0.1790$). November and January rainfall also showed negative relationships, while December rainfall demonstrated a weaker but still negative association. The overall pattern suggests that excessive precipitation during the potato-growing season may be associated with unfavourable yield conditions. The effect is likely stage-dependent, as rainfall can influence crop establishment, root-zone conditions, nutrient availability, and disease risk differently throughout crop development.

Table 2. Correlation between monthly rainfall and potato yield.

Months	Rainfall (mm)	Pearson Correlation	*p* Value	Spearman Correlation
rain_1m	34.62	−0.1790	6.86×10^{-17}	−0.1715
rain_2m	9.01	−0.1597	1.02×10^{-13}	−0.0954
rain_3m	7.37	−0.0675	1.78×10^{-03}	−0.0047
rain_4m	19.87	−0.1634	2.67×10^{-14}	−0.1559

The correspondence between Pearson and Spearman results provides additional evidence that the observed negative relationships are not exclusively dependent on strictly linear patterns.

3.3. Temperature–Yield Relationships

Temperature effects were more strongly dependent on the timing of the growing stage (illustrated in Table 3). November temperature showed the strongest positive association with potato yield, with Pearson and Spearman correlations of 0.2391 and 0.2032, respectively. October and January temperatures also demonstrated weak positive relationships. In contrast, December temperature exhibited a negative association with yield. This pattern suggests that relatively cooler conditions during this period may be associated with more favourable potato productivity.

Table 3. Correlations between monthly temperature and potato yield.

Months	Temperature (°C)	Pearson Correlation	*p* Value	Spearman Correlation
temp_1m	23.58	0.0594	5.95×10^{-03}	0.0705
temp_2m	19.63	0.2391	3.02×10^{-29}	0.2032
temp_3m	18.36	−0.0781	2.92×10^{-04}	−0.0286
temp_4m	20.81	0.1135	1.38×10^{-07}	0.1080

These results demonstrate that temperature cannot be treated as a single uniformly acting predictor. Its influence differs across stages of crop development, supporting the use of monthly variables and nonlinear ML models.

3.4. Climatic Conditions Associated with High Potato Yield

KDE analysis (illustrated in Figures 3 and 4) provides an additional interpretive layer by identifying monthly climatic ranges associated with high-yield observations. During October, corresponding to planting and crop establishment, high-yield observations were concentrated around rainfall of approximately 1–21 mm and temperatures of 22.8–23.7 °C. These conditions may provide sufficient moisture for crop establishment without excessive rainfall. During November, associated with vegetative development and tuber initiation, high-yield observations were concentrated under relatively dry conditions of approximately 0–5 mm rainfall and temperatures of 18.8–19.5 °C. During December, corresponding to an important tuber-development period, the high-yield climatic window was comparatively narrow: approximately 1–5 mm rainfall and 17.6–18.5 °C. During January, associated with later crop development and maturation, high-yield observations were concentrated around rainfall of 5–15 mm and temperatures of 20.2–21.1 °C. Collectively, the results indicate that favourable potato-yield outcomes in the analyzed dataset are generally associated with

relatively limited rainfall and seasonally appropriate temperatures. These KDE-derived ranges provide an evidence-based basis for the recommendation component of the proposed precision-agriculture framework.

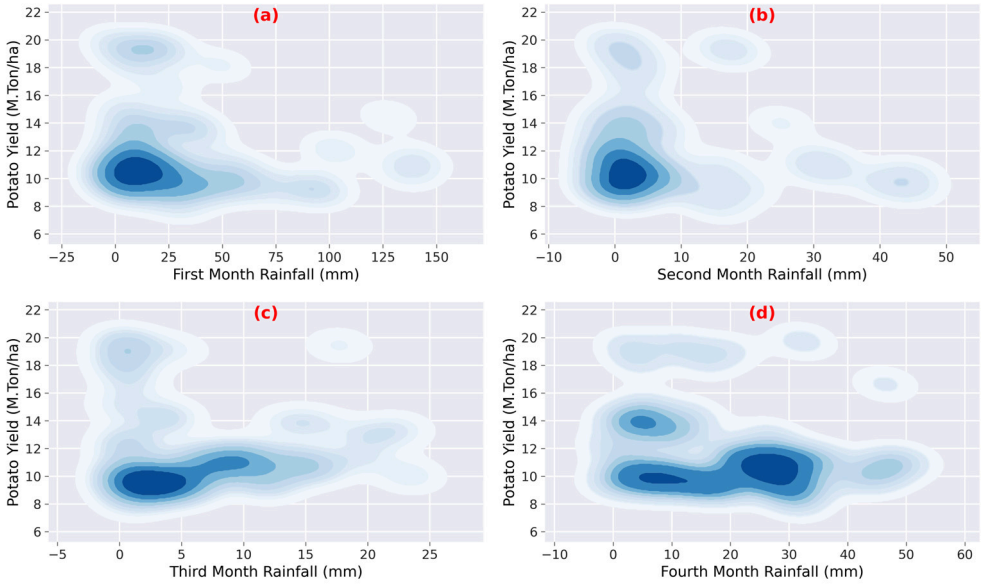


Figure 3. Monthly rainfall conditions associated with high-yield potato observations, (**a**) First month rainfall vs potato yield, (**b**) Second month rainfall vs potato yield, (**c**) Third month rainfall vs potato yield, and (**d**) Fourth month rainfall vs potato yield.

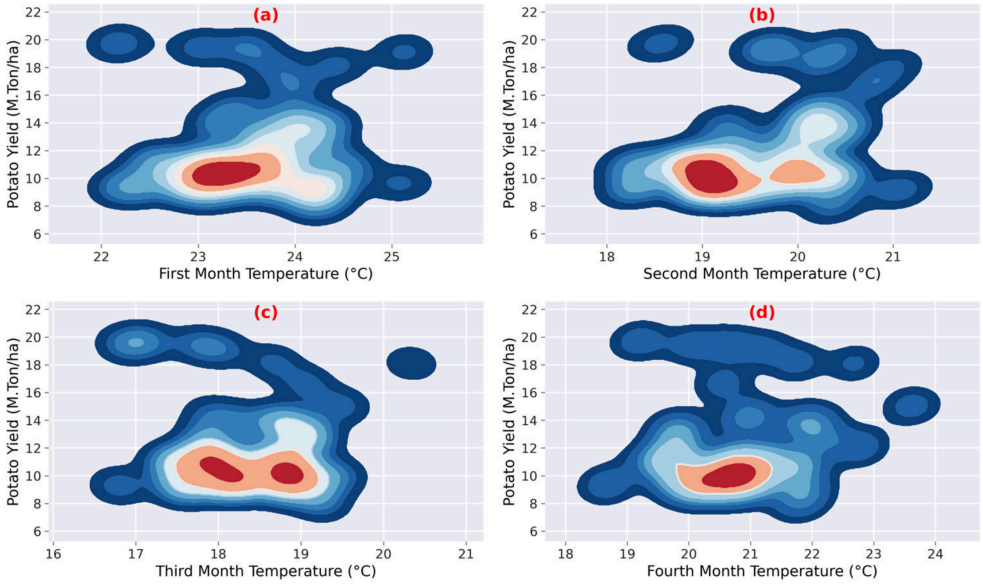


Figure 4. Monthly temperature conditions associated with high-yield potato observations, (**a**) First month temperature vs potato yield, (**b**) Second month temperature vs potato yield, (**c**) Third month temperature vs potato yield, and (**d**) Fourth month temperature vs potato yield.

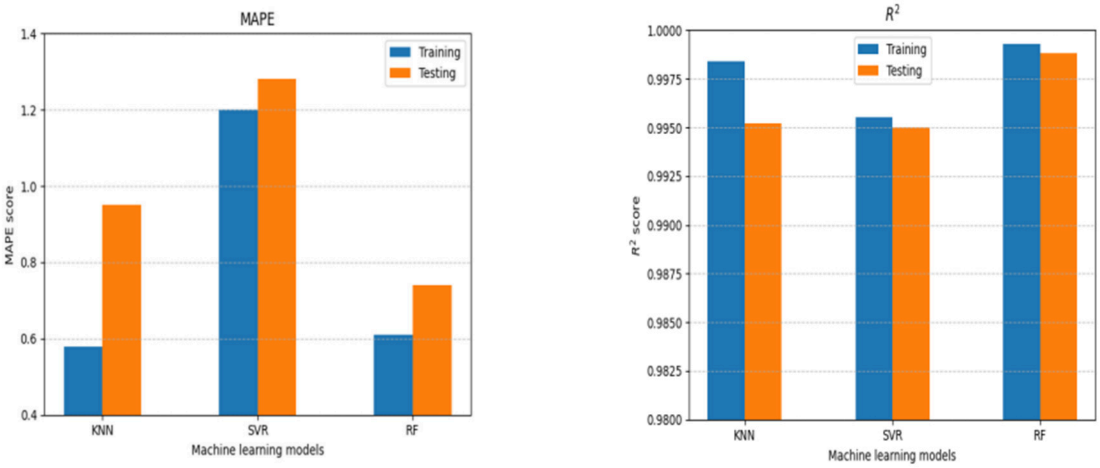
3.5. Machine-Learning Model Performance

Based on the established climate–yield relationships, KNN, SVR, and RF were evaluated using five-fold cross-validation and independent testing (results are illustrated in Table 4 and Figure 5). KNN demonstrated strong predictive capability, with a mean MAE of 0.1211 t ha^{-1} and RMSE of 0.2933 t ha^{-1}. Its performance suggests that observations with similar climatic characteristics tend to exhibit similar yield patterns. However, as an instance-

based method, KNN may have limited ability to extrapolate beyond climatic conditions represented in the historical training data. SVR produced comparatively larger errors, with a mean MAE of 0.3876 t ha^{-1} and RMSE of 0.8415 t ha^{-1}. Its comparatively lower performance may reflect sensitivity to hyperparameter selection and limitations in representing the complex nonlinear interactions contained in the dataset. RF achieved the strongest and most stable predictive performance, with a mean MAE of 0.0803 t ha^{-1} and RMSE of 0.1014 t ha^{-1}. The relatively low variation across validation folds further indicates stable performance within the reported evaluation framework.

Table 4. Models' performance through 5-fold cross-validation.

Fold	KNN		SVR		RF	
	MAE	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
1	0.1282	0.3683	0.4944	1.1339	0.0848	0.1069
2	0.1155	0.1996	0.3211	0.5396	0.0816	0.1025
3	0.128	0.3457	0.3748	0.8793	0.0767	0.0981
4	0.1163	0.2975	0.4207	0.9456	0.0779	0.0983
5	0.1176	0.2556	0.3269	0.7091	0.0807	0.1011
Mean	0.1211	0.2933	0.3876	0.8415	0.0803	0.1014
Sd.	0.0057	0.0610	0.0644	0.2032	0.0029	0.0032



(**a**) (**b**)

Figure 5. Performance comparison of ML models in terms of: (**a**) MAPE and (**b**) R^2.

RF's superior performance is consistent with its ensemble architecture, which can represent nonlinear relationships, interactions among climatic variables, and threshold effects that characterize climate–yield systems. It also provides intrinsic feature-importance information, allowing predictive performance and interpretability to be integrated within the same overall framework.

4. Conclusions

This study developed an interpretable machine-learning framework for early potato-yield prediction and climate-based agricultural decision support using historical agro-meteorological and production data from Bangladesh.

The framework integrates a complete analytical pipeline comprising data extraction and integration, preprocessing, data augmentation, feature engineering, climate-variable

identification, correlation analysis, KDE-based climatic recommendation, ML prediction, hyperparameter optimization, and model evaluation. A key contribution is the integration of two complementary analytical systems: a crop-yield prediction system, which estimates potato yield, and a crop-yield recommendation system, which explains influential climatic variables, their relationships with yield, and the environmental conditions associated with high-yield observations.

The climate analysis demonstrates that potato yield is strongly influenced by stage-specific meteorological conditions. Rainfall variables, particularly those corresponding to December, October, and January, contributed substantially to yield variability, while excessive rainfall was generally negatively associated with potato yield. Temperature effects differed across growth stages, with November temperature showing a positive association with yield and relatively cooler December conditions being associated with favourable productivity.

The KDE analysis further identified monthly climatic ranges associated with high-yield observations. High potato productivity was generally associated with relatively limited rainfall and seasonally appropriate temperatures, providing an interpretable basis for climate-informed recommendations.

Among the three evaluated ML algorithms, RF achieved the strongest performance, with a mean MAE of 0.0803 t ha^{-1} and RMSE of 0.1014 t ha^{-1} across five-fold cross-validation. Its superior performance, stability, nonlinear modelling capability, and intrinsic feature-importance mechanism make it particularly suitable for the proposed framework. Future work can extend the framework through the integration of soil characteristics, crop-management information, remote-sensing indicators, and real-time IoT sensor data to develop continuously updated agricultural decision-support systems.

Author Contributions: Conceptualization, M.J.H. and T.A.T.; methodology, M.J.H., T.A.T. and T.S.; software, T.A.T. and T.S.; validation, M.J.H., T.A.T. and T.S.; formal analysis, M.J.H., T.A.T. and T.S.; investigation, M.J.H. and T.A.T.; resources, M.J.H. and T.S.; data curation, T.A.T. and T.S.; writing—original draft preparation, M.J.H., T.A.T. and T.S.; writing—review and editing, M.J.H.; visualization, M.J.H., T.A.T. and T.S.; supervision, M.J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be available on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sanyaolu, M.; Sadowski, A. The role of precision agriculture technologies in enhancing sustainable agriculture. *Sustainability* **2024**, *16*, 6668. [CrossRef]
2. Hernández-Ochoa, I.M.; Gaiser, T.; Grahmann, K.; Engels, A.M.; Ewert, F. Within-field temporal and spatial variability in crop productivity for diverse crops—A 30-year model-based assessment. *Agronomy* **2025**, *15*, 661. [CrossRef]
3. Abbas, F.; Afzaal, H.; Farooque, A.A.; Tang, S. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy* **2020**, *10*, 1046. [CrossRef]
4. Jannat, A.; Ishikawa-Ishiwata, Y.; Furuya, J. Assessing the impacts of climate variations on the potato production in Bangladesh: A supply and demand model approach. *Sustainability* **2021**, *13*, 5011. [CrossRef]
5. Salan, M.S.A.; Hossain, M.M.; Sumon, I.H.; Rahman, M.M.; Kabir, M.A.; Majumder, A.K. Measuring the impact of climate change on potato production in Bangladesh using Bayesian hierarchical spatial-temporal modeling. *PLoS ONE* **2022**, *17*, e0277933. [CrossRef] [PubMed]

6. Barrile, V.; Maesano, C.; Genovese, E. Optimization of crop yield in precision agriculture using WSNs, remote sensing, and atmospheric simulation models for real-time environmental monitoring. *J. Sens. Actuator Netw.* **2025**, *14*, 14. [CrossRef]
7. Kumar, V.; Sharma, K.V.; Kedam, N.; Patel, A.; Kate, T.R.; Rathnayake, U. A comprehensive review on smart and sustainable agriculture using IoT technologies. *Smart Agric. Technol.* **2024**, *8*, 100487. [CrossRef]
8. Iqbal, N.; Shahzad, M.U.; Sherif, E.-S.M.; Tariq, M.U.; Rashid, J.; Le, T.-V.; Ghani, A. Analysis of wheat-yield prediction using machine learning models under climate change scenarios. *Sustainability* **2024**, *16*, 6976. [CrossRef]
9. Zhang, P.; Lu, B.; Shang, J.; Wang, X.; Hou, Z.; Jin, S.; Yang, Y.; Zang, H.; Ge, J.; Zeng, Z. Ensemble learning for oat yield prediction using multi-growth stage UAV images. *Remote Sens.* **2024**, *16*, 4575. [CrossRef]
10. Kuradusenge, M.; Hitimana, E.; Mtonga, K. SMART-CYPS: An intelligent internet of things and machine learning powered crop yield prediction system for food security. *Discov. Internet Things* **2024**, *4*, 23. [CrossRef]
11. Singh, G.; Sharma, S. Enhancing precision agriculture through cloud based transformative crop recommendation model. *Sci. Rep.* **2025**, *15*, 9138. [CrossRef] [PubMed]
12. Hoque, M.J. Incorporating meteorological data and pesticide information to forecast crop yields using machine learning. *IEEE Access* **2024**, *12*, 47768–47786. [CrossRef]
13. Piekutowska, M.; Niedbała, G. Review of methods and models for potato yield prediction. *Agriculture* **2025**, *15*, 367. [CrossRef]
14. Wang, Y.; Zhang, Q.; Yu, F.; Zhang, N.; Zhang, X.; Li, Y.; Wang, M.; Zhang, J. Progress in research on deep learning-based crop yield prediction. *Agronomy* **2024**, *14*, 2264. [CrossRef]
15. Bhimavarapu, U.; Battineni, G.; Chintalapudi, N. Improved optimization algorithm in LSTM to predict crop yield. *Computers* **2023**, *12*, 10. [CrossRef]
16. Batool, D.; Shahbaz, M.; Asif, H.S.; Shaukat, K.; Alam, T.M.; Hameed, I.A.; Ramzan, Z.; Waheed, A.; Aljuaid, H.; Luo, S. A hybrid approach to tea crop yield prediction using simulation models and machine learning. *Plants* **2022**, *11*, 1925. [CrossRef] [PubMed]
17. Alsaber, A.; Satpathi, A.; Alsabah, M.; Setiya, P. Optimizing potato yield predictions in Uttar Pradesh, India: A comparative analysis of machine learning models. *Sci. Rep.* **2025**, *15*, 12719. [CrossRef] [PubMed]
18. Lin, Y.; Li, S.; Duan, S.; Ye, Y.; Li, B.; Li, G.; Lyv, D.; Jin, L.; Bian, C.; Liu, J. Methodological evolution of potato yield prediction: A comprehensive review. *Front. Plant Sci.* **2023**, *14*, 1214006. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.