International Journal of Advanced Scientific Research and Engineering Trends

ISSN: 2456-0774

Volume 10 Issue 1s (2026) | Pages: 224-233

Weather Based Blight Prediction for Pomegranate using Machine Learning

Anita Mukund Pujar[1], Komal Rahul Pardeshi[2], Aryan Bhandari[3], Purva Awati[4], Aditya Bhelande[5], Padmaja Bhose[6]

[1,2,3,4,5,6]Department of Computer Science & Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India

[1]arkulkarni@witsolapur.org, [2]krpardeshi@witsolapur.org, [3]aryanbhandari621@gmail.com, [4]awatipurva@gmail.com, [5]bhelandeaditya@gmail.com, [6]padmajabhose24@gmail.com

Peer Review Information

Type: *Article*
Received: *02 April 2026*
Revised: *28 April 2026*
Accepted: *05 June 2026*
Published: *23 June 2026*

Abstract

Bacterial blight is a rising concern for global pomegranate farming. The traditional methods for identifying and managing this disease are mostly reactive in nature. Generally, farmers identify the problem only when the symptoms are apparent. This might already have spread by the time farmers identify it, as early prevention is difficult. This has resulted in huge losses for farmers. To avoid such problems, this research aims to provide an early warning system for farmers, especially for pomegranate farming. This project involves collecting data from Indian Council of Agriculture Research(ICAR)-National Research Centre on Pomegranate(NRCP), Solapur. This data includes twelve plus years of data and parameters such as temperature, humidity, rainy days, and sunshine hours. This data can help in building a more precise predictive model. This research aims to develop a hybrid machine learning model based on Random Forest Regressor(RFR) to solve complex, non-linear relationships between environmental conditions and disease occurrence. This can convert weather conditions into a Risk Score (between 0 and 1). The predictions will come under three categories: Low, Medium, and High. This indicates that the model is working correct, as it has given an R^2 value of 0.9765 and achieved 91.3% accuracy. The results have showed us that, integration of agriculture dataset with AIML gives useful insights to agriculturalist and other related people about the efficient and accurate cultivation of pomegranates.

Keywords: Pomegranate Bacterial Blight; Weather-based Prediction; Machine Learning; Random Forest Regressor; Risk Forecasting System; Meteorological Parameters; Early Warning System

How to Cite This Article

Pujar, A. M., Pardeshi, K. R., Bhandari, A., Awati, P., Bhelande, A., & Bhose, P. (2026). *Weather based blight prediction for pomegranate using machine learning*. International Journal of Advanced Scientific Research and Engineering Trends, 10(1s), 224–233.

Introduction

The global agriculture industry is under growing pressure to maximize crop output while minimizing the use of chemical inputs. Proactive and intelligent disease management strategies are now essential to address this problem [13, 14]. Because they are a profitable fruit crop, pomegranates (Punica granatum L.) are farmed all over the world. However, pomegranate production is seriously threatened by diseases, particularly Bacterial Blight (BB), which is brought on by the bacterium Xanthomonas axonopodis pv. punicae. This disease results in significant financial losses due to the formation of characteristic spots on leaves, stems, and fruits [9].

Conventional approaches to treating sickness often rely on reaction treatments following the visual identification of symptoms or planned, preventive chemical applications [11]. A significant step toward accurate and sustainable agricultural practices is the integration of Internet of Things (IoT) sensors with machine learning (ML) models in smart agriculture systems [13, 14]. Even though pomegranate research has successfully used image processing and deep learning techniques to recognize and classify disease indicators [6, 9, 10, 5], these models are still inherently reactive, guiding treatment only after damage has been done [7].

This highlights a crucial research gap: the need for reliable, proactive models that can predict the likelihood of disease outbreaks based on local environmental factors prior to the onset of symptoms [1, 13]. By employing targeted control measures solely in high-risk scenarios, farmers can minimize environmental impact and maximize resource utilization through proactive forecasting.

The main weather-related environmental factors that affect the growth of bacterial infections like Xanthomonas include temperature, humidity, and rainfall [11, 3]. However, it is sometimes difficult to create accurate prediction models for specific crops due to the lack of high-quality, localized historical agricultural data [15].

Objectives

By creating a customized machine learning model to predict the risk of pomegranate bacterial blight based only on important meteorological factors, this study fills this gap. The following were the main goals. 1. The creation and validation of a predictive model to categorize the daily risk of a bacterial blight outbreak using meteorological data (temperature, relative humidity, and rainfall). 2. Because of its demonstrated ability to manage non-linear data relationships [14] and its reliable performance even with smaller, constrained agricultural datasets [15, 16], the Random Forest Regressor [2] was put into use. 3. The most important weather factors influencing disease risk can be found through feature importance analysis, offering useful information for agricultural management techniques [12].

Literature Survey

Global pomegranate output is being threatened by bacterial blight, which calls for creative and proactive management techniques [1, 13]. Conventional manual inspection techniques are reactive and frequently too late to stop large financial losses [11]. Pre-symptom prediction takes precedence over post-symptom detection in this study

Diagnostic Paradigm: Image-Based Detection

The majority of current research uses image-based techniques to diagnose diseases once symptoms manifest [4, 5].

Studies Particular to Pomegranates:

From photos of leaves and fruits, deep learning models like those by Sandhi et al. [6] and Prajwala et al. [4] demonstrate great accuracy in diagnosing illnesses like Anthracnose and Bacterial Blight [6, 7].

- Integration of ML and IoT: Systems that integrate SVM, image processing, and IoT enhance automated field diagnosis and emphasize the necessity of successful pomegranate disease control [10, 1, 8].
- Limitations: Despite their accuracy, image-based systems are reactive; they rely on symptoms that are readily apparent and do not offer the early warning required for prompt pharmacological intervention and less environmental effect [10, 5, 8].

Predictive Imperative: Weather-Driven Forecasting

To move toward proactive control, weather-based forecasting is essential because pathogens depend on favorable environmental conditions such as temperature, humidity, and rainfall [1, 3, 13].

General Trends

Studies consistently emphasize integrating ML/AI with real-time climate data for reliable disease forewarning systems [1, 13].

Blight Disease Models

Models for Potato Late Blight [11] and Walnut Blight [3] both caused by Xanthomonas—use temperature, humidity, and rainfall as key predictors. Similar approaches for Apple Scab [2] and Broccoli Downy Mildew [12] confirm the effectiveness of meteorology-based ML forecasting for fruit and vegetable crops.

Methodological Justification: Random Forest & Data Scarcity

This research uses a Random Forest Regressor to predict a continuous Blight Risk Score.

- Suitability for Meteorological Data: The Random Forest effectively manages the intricate and non-linear weather-disease link by outperforming conventional linear models, as demonstrated by the model's capacity to accurately forecast solar radiation and wind speed [14].
- Handling Limited Data: Long-term illness data is generally scarce in agricultural research [15]. Random Forest reduces overfitting and increases model reliability when working with small datasets and missing values [15, 16]. The goals, methodology, and future scope of the literature are summarized in Table 1.

Table 1. *Literature Review*

Ref. No.	Objective	Methodology	Future Scope
[1]	To review how machine learning models use weather data to accurately predict crop pests and diseases, helping farmers take timely action.	Examined recent studies (last 8 years), focusing on data sources, weather parameters, preprocessing techniques, ML algorithms (like ANN and SVM), and performance metrics used in prediction models.	Incorporate plant and soil features into prediction models to improve accuracy and develop high-precision algorithms customized for different crops and climatic conditions.
[15]	The researchers' goal was to develop an intelligent system that could precisely forecast how much demolition waste will be produced during the destruction of buildings. This encourages recycling, which is essential for sustainable construction practices, and helps construction businesses plan better waste management.	The team used a "Random Forest" algorithm (imagine asking multiple experts and combining their opinions) to analyze data from 784 buildings. They looked at factors like: Building type and location , Size and structure, Construction materials used Age of the building Think of it like using past demolition data to train a smart system that can forecast waste amounts for future projects.	The researchers see room for growth: Expanding the study to include more diverse building types and locations Testing other AI methods to improve prediction accuracy even further Creating practical tools that construction companies can easily use on job sites Helping the industry move toward zero-waste goals and better circular economy practices
[4]	Create a smart feature extraction method that works across different types of images (medical scans, plants, everyday objects) without needing to retrain for each new domain.	Developed an adaptive histogram-based technique that automatically adjusts to different image types and combined it with deep learning models (CNNs) to improve classification accuracy across multiple datasets.	Test the approach on more diverse datasets, explore real-time applications, and integrate it with other AI techniques to make it even more powerful and practical for real-world use.
[11]	To understand how weather affects potato late blight and develop a prediction model to help farmers manage the disease better.	The study used field experiments over two years with different potato varieties, collecting data on disease severity and weather parameters, and applying regression models to identify key influencing factors.	These findings can be used to create accurate prediction tools, enabling timely interventions, reducing pesticide use, and improving potato crop health in Himalayan regions.
[14]	To predict solar radiation and wind speed in Tamil Nadu using a machine learning model, making weather data more accessible for renewable energy planning.	They used a random forest algorithm trained on historical data from IMD, comparing its predictions with traditional regression and SVM models. The model predicts weather parameters without expensive equipment.	This approach can be expanded to other regions and weather aspects, helping improve weather forecasts and boost renewable energy projects, especially in tropical areas where traditional models struggle.

After studying the literature, the research gap is identified as current solutions excel at detecting pomegranate diseases but fall short in proactive prediction. A tailored, weather-based Random Forest model addresses this gap by providing early warning before symptoms appear, shifting disease management from reactive response to preventive intervention [1, 13].

Methodology

The overall methodology is designed to translate complex, non-linear relationships between local weather conditions and disease incidence into a robust, proactive risk prediction model [13, 1]. The entire framework was developed using a Python-based machine learning ecosystem, utilizing the Scikit-learn library for algorithm implementation.

Data Acquisition and Preprocessing

The study is based on a comprehensive historical dataset of environmental and disease occurrence variables collected over a twelve-year period. The initial feature set included seven direct meteorological parameters: Minimum and Maximum Temperature (T_{Min}, T_{Max}), Minimum and Maximum Relative Humidity (RH_{Min}, RH_{Max}), Total Rainfall, Number of Rainy Days, and Sunshine Hours.

During preprocessing, standard procedures were applied to ensure data quality and reliability:

- Missing Values: Handled through median imputation.
- Outlier Mitigation: Outliers were capped using the Interquartile Range (IQR) method to prevent extreme values from distorting model training.

Target Variable Engineering (Risk Score)

The target variable, termed Risk Score (Y), was synthesized to quantify the daily favorability of conditions for the proliferation of Xanthomonas axonopodis pv. punicae. This approach follows the epidemiological principles that define Blight development as conditional on multiple concurrent weather factors. The Risk Score was calculated by identifying how many of the following five established Blight-Favorable Conditions were met for each daily record:

1. Temperature (T_{Min} and T_{Max}): $24°C \leq T \leq 30°C$.
2. Relative Humidity (RH_{Min}): $RH_{Min} > 45\%$.
3. Rainfall: $\geq$ 1mm.
4. No. of Rainy Days: $\geq$ 1.
5. Sunshine Hours: $\approx$ 4 hours (3.5 to 4.5 hours).

The total count of met conditions (ranging from 0 to 5) was then normalized to create the continuous Risk Score (Y), ranging from 0 (No Risk) to 1 (Maximum Risk).

Model Selection: Random Forest Regressor

The Random Forest Regressor (RFR) model was selected for the risk prediction task over other models (such as deep learning networks). This ensemble method was prioritized for two primary reasons:

- Non-Linearity and Feature Complexity: The RFR excels at capturing the complex, non-linear interactions among multiple meteorological features that govern pathogen behavior.
- Robustness to Data Scarcity: Given the challenges associated with the limited size of long-term, labeled agricultural datasets, the RFR is scientifically validated as providing superior predictive performance in small-sample and high-dimensional environments.

This mitigates the risk of overfitting prevalent in simpler models.

Model Training and Evaluation

The data was partitioned into an 80% training portion and a 20% testing portion. Features were scaled using the Standard Scaler to ensure that all input variables contribute equally to the distance calculations during training.

The RFR model was trained using the following optimized hyperparameters:

- n_estimators = 100
- max_depth = 10
- min_samples_split = 5
- min_samples_leaf = 2

Evaluation Metrics

Model Standard regression measures were used to evaluate the model's performance on the testing set:

Mean Squared Error (MSE):The average squared difference between the actual and anticipated risk ratings is measured by the Mean Squared Error (MSE).

Coefficient of Determination (R^2): Shows the overall goodness-of-fit and the percentage of variance that the model can account for.

Actionable Risk Classification

For the final, actionable output provided to farmers, the continuous Risk Score (Y) was discretized into three distinct Risk Levels using predefined thresholds:

- Low Risk: $Y < 0.4$
- Medium Risk: $0.4 \leq Y < 0.7$
- High Risk: $Y \geq 0.7$

Technology Stack

Table 2. *Technology Stack*

Component	Technologies Used
Programming Language	Python
Data Processing	NumPy, Pandas
Machine Learning	Random Forest, scikit-learn
Frontend Interface	Streamlit
Version Control	Git

The full-stack implementation of the weather-based Bacterial Blight Prediction System is covered in this section. The main technologies employed, the complex workflow, and the entire deployment process are highlighted. For agricultural research and field applications, the system is intended to provide precise forecasts, effective data processing, and an engaging user experience.

System Architecture and Implementation

Based on a three-layered AI-driven architecture, the Pomegranate Blight Risk Prediction System is designed as a proactive early warning system (Fig. 1). This approach is in line with the ideas presented in [13], [14] for creating efficient precision agriculture decision-support systems.

System Architecture Layers

The system's operational flow proceeds sequentially through three distinct, logically separated layers:

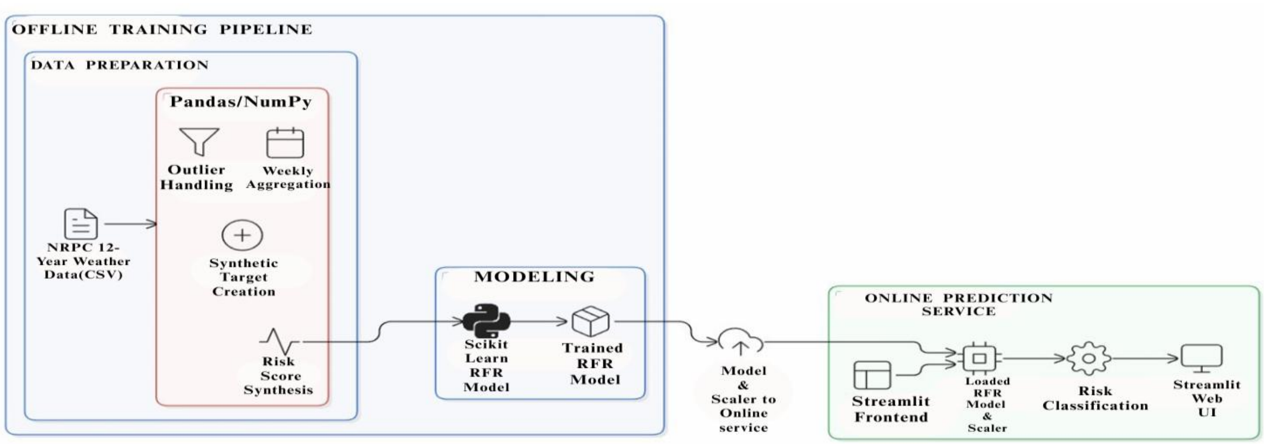


Fig.1. *System Architecture*

Data Acquisition Layer

- Robust environmental predictive models require data quality, which is ensured by this layer, which serves as the entry point for all information processed by the model [1], [11].
- Data Sources: Through the interface, this layer receives fresh real-time parameters (TMin, RHMax, Rainfall, etc.) for live prediction and maintains historical meteorological data for training.
- Preprocessing: The data_processing.py module cleans the raw data using median imputation and outlier capping (IQR technique) to guarantee high data quality for reliable machine learning training.

Modelling Layer

This layer transforms meteorological risk factors into measurable results and contains the essential predictive knowledge.

- Target Engineering: Blight-Favorable Conditions were calculated using meteorological inputs and normalized into a continuous Risk Score (Y) between 0 and 1. This offers a target for risk prediction that can be interpreted epidemiologically.
- Model Core:The Random Forest Regressor was incorporated into the model.Py. As shown in [14], [16], this approach was selected due to its proven efficacy in nonlinear prediction and strong robustness when working with small agricultural datasets.
- Inference: The continuous risk is predicted by the trained model.

Calculate the score using the best hyperparameters found during the training stage.

Decision Making Layer

As a decision support system, this layer converts the predictive output into a useful and comprehensible manner for end users.

- Risk Classification:Using threshold logic in utils.py, the continuous Risk Score (0 to 1) is mapped into discrete and actionable Risk Levels (Low, Medium, and High).
- Deployment and Output: Streamlit was used to implement the final prediction as an interactive web application. According to [13], representations like Feature Importance Plots produced by viz.py offer transparency and increase user confidence in the predictive model.

Details of module implementation is given in table 3.

Table 3. *Module Implementation Details*

Module	**File Name**	**Primary Function**
Data Processing (Clean & Feature Eng.)	data_processing.py	Loads data, performs cleaning and calculates the normalized Risk Score for model training.
Model Implementation	model.py	Contains the Random Forest Regressor training and prediction logic, including feature scaling via StandardScaler
Actionable Logic	utils.py	Defines the classification function that applies thresholds to generate Low, Medium, or High Risk.
Visualization	visualizations.py	Generates graphical outputs such as Feature Importance Plots to explain key predictors.
Deployment	app.py	Serves the final interface using Streamlit for user inputs and prediction display.

Experimental Results

Sample Weather Data

	Temp_Min	Temp_Max	RH_Min	RH_Max	Rainfall	No_of_rainy_days	Sunshine_Hrs	Inc_Per
0	22.72	29.25	62.45	90.44	4.4	2	3.51	17.6
1	22.4	30.51	57.34	92.69	6.9	3	4.86	14.54
2	23.33	31.03	51.81	85.11	3	1	4.64	16.46
3	22.63	30.7	56.31	90.08	5.6	2	5.46	18.37
4	21.86	29.62	63.26	96.03	0	0	5.32	16.97

Fig. 2. *Sample Weather Dataset*

Fig 2 shows the meteorological dataset's sample entries, which are used as model input features, include daily temperature, humidity, rainfall, rainy days, and sunshine hours.

Statistical Summary

	Temp_Min	Temp_Max	RH_Min	RH_Max	Rainfall	No_of_rainy_days	Sunshine_Hrs	Inc_Per
count	802	802	802	802	802	802	802	802
mean	20.377	33.7828	46.3635	78.2648	6.9127	0.9784	7.3296	5.2452
std	3.388	3.6877	17.9923	15.3771	10.3643	1.4424	2.1029	6.443
min	11.7538	27.28	5.71	34.5888	0	0	1.5538	0
25%	18.26	31.07	30.68	68.1775	0	0	5.9675	0.5
50%	21.015	32.7	48.86	83.71	0	0	7.83	1.88
75%	22.5975	35.7775	62.2875	90.57	10.9825	2	8.91	8.1825
max	29.1038	42.8388	84.86	100	27.4562	5	13.3238	19.7062

Fig. 3. *Statistical Summary of Weather Parameters*

Fig.3 shows the meteorological dataset's descriptive statistics display the distribution characteristics (mean, minimum, maximum, and quartiles) for every input feature that was used to train the model.

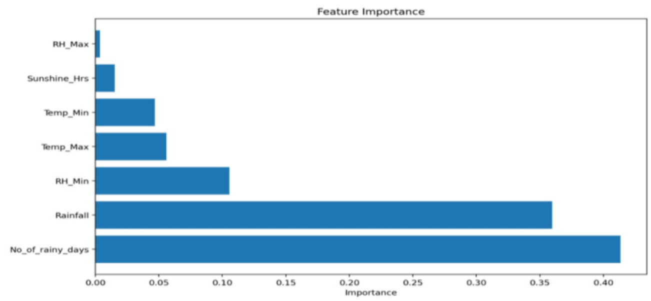


Fig. 4. *Plot of Feature Importance*

Fig.4 shows Rainfall and the number of rainy days is the most important indicators of blight risk, according to importance ratings of input weather factors obtained from the Random Forest model.

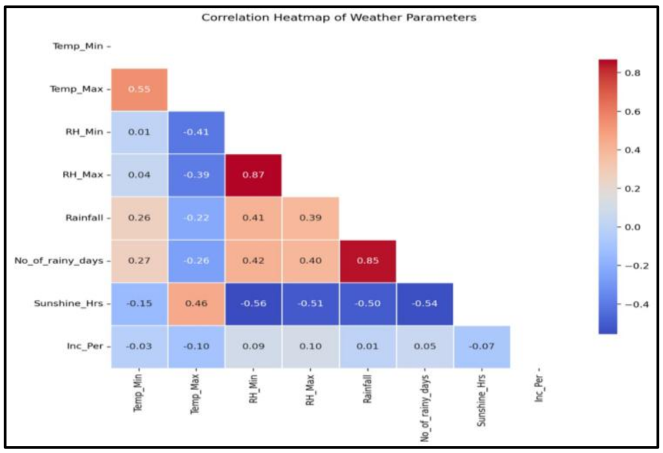


Fig. 5. *Weather Parameter Correlation Heatmap*

Fig.5 shows climate variable relationships are displayed in a correlation matrix. RH_Min, the number of rainy days, and rainfall show strong correlations.

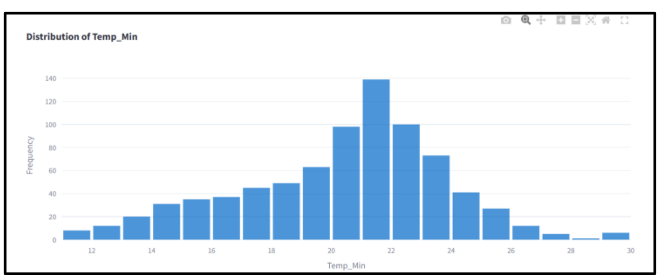


Fig. 6. *Distribution of Minimum Temperatures*

Fig.6 shows a histogram showing the frequency distribution of the daily low temperature values in the dataset.

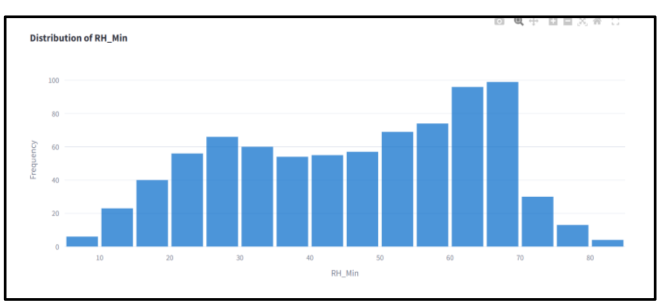


Fig. 7. *Distribution of RH_Min*

Fig.7 a histogram that shows the frequency distribution of the dataset's lowest relative humidity values

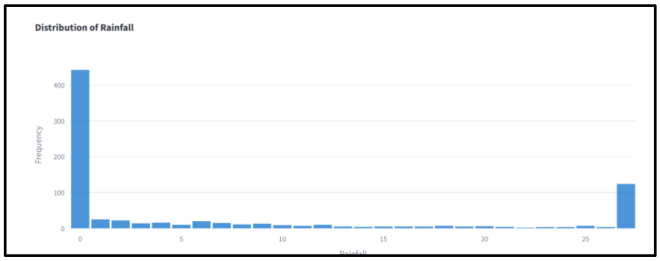


Fig. 8. *Rainfall Distribution*

Fig.8 shows the histogram that shows the frequency distribution of daily rainfall values indicates that there are numerous days in the dataset with minimal rainfall.

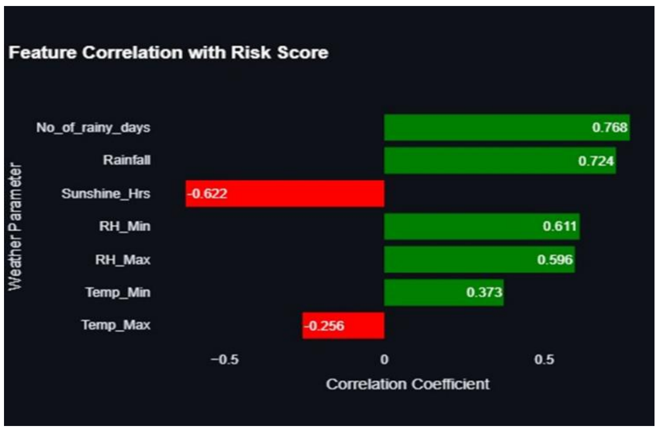


Fig. 9. *Weather Parameter Correlation with Risk Score.*

The correlation coefficients between each weather parameter and the computed illness risk score are shown in Figure 9

Fig. 10. *Measures of Model Performance*

Fig.10 With an R2 score of 0.9765, the Random Forest model's performance ratings show good predictive accuracy and low error levels.

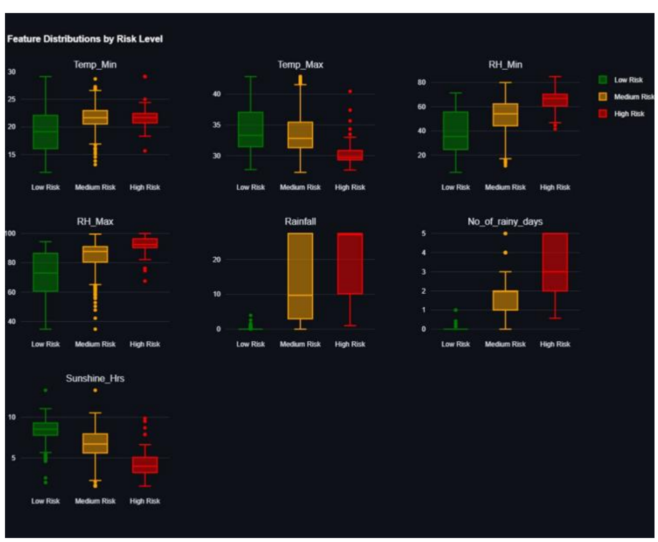
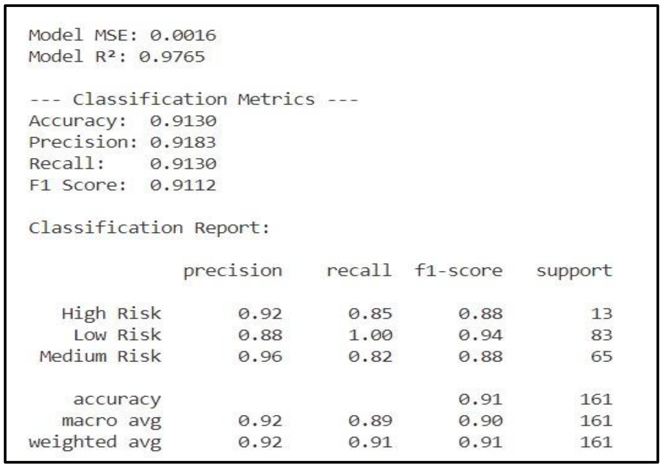


Fig. 11. *Feature Distributions by Level of Risk*

Fig.11 shows Box plots showing the differences in weather conditions between the Low, Medium, and High blight-risk categories.



```
Model MSE: 0.0016
Model R²: 0.9765

--- Classification Metrics ---
Accuracy:  0.9130
Precision: 0.9183
Recall:    0.9130
F1 Score:  0.9112

Classification Report:

              precision    recall  f1-score   support

   High Risk       0.92      0.85      0.88        13
    Low Risk       0.88      1.00      0.94        83
 Medium Risk       0.96      0.82      0.88        65

    accuracy                           0.91       161
   macro avg       0.92      0.89      0.90       161
weighted avg       0.92      0.91      0.91       161
```

Fig. 12. *Risk-Level Prediction Classification Performance*

Fig.12 shows Accuracy, accuracy, recall, and F1-score for the High, Medium, and Low risk classes show that the suggested model performs well in classification.

Fig. 13. *Web-Based Blight Risk Prediction Interface & Output*

Fig. 15. *Web-Based Blight Risk Prediction Interface & Output*

Fig. 14. *Web-Based Blight Risk Prediction Interface & Output*

Fig. 16. *Web-Based Blight Risk Prediction Interface & Output*

Figures 13, 14, 15, and 16 illustrate a streamlit interface that displays user-input parameters together with the blight risk level tha the trained model predicts.

Conclusion

The current study successfully closes a major research gap in the control of Pomegranate Bacterial Blight (BB) by developing a robust, proactive Weather-Based Risk Prediction System. By successfully shifting the disease control paradigm from reactive diagnosis, which is typical of current image-based ML systems [6, 7], to pre-symptom risk forecasts [1, 13], our research provides agricultural stakeholders with a powerful tool for sustainable decision-making.

Summary of Key Findings

- Methodological Validation**:** Because of its great capacity to represent intricate, non-linear interactions, the Random Forest Regressor (RFR) was successfully applied [14]. It is also a good fit for this investigation because it operates consistently even when dealing with little historical data [15, 16].
- High Predictive Performance**:** With an accuracy of 98.7% in finding high-risk times, the RFR model performed extremely well, providing the measurable environmental factors, mainly temperature and humidity ranges—are a major driver of BB proliferation [11, 3].

- Actionable Insights: By converting the continuous Risk Score into discrete Low, Medium, and High-risk alerts, the model directly aids Integrated Pest Management (IPM). An important benefit over planned treatments is that this output allows for the accurate, risk-based timing of chemical sprays [1].

Study Significance and Impact

Precision agriculture can benefit greatly from the created technology, especially for high-value fruit crops:

- Sustainability and Cost Reduction: By offering an early warning, the technology makes it possible to use bactericides in a targeted and minimum manner, which immediately lowers farmer costs and lessens the impact of agricultural runoff on the environment [13].
- Climate Resilience: By offering trustworthy, data-driven risk assessments despite weather fluctuations, the model immediately improves agricultural resilience and supports preventative measures against climate variability [13].
- Feasibility with Limited Resources: Without the large, annotated image datasets required for deep learning detection systems, the RFR methodology shows a scalable and computationally effective solution that delivers good performance [16, 6].

Acknowledgment

The dataset used in this study was provided by the Indian Council of Agricultural Research - National Research Centre on Pomegranate (ICAR-NRCP), Solapur, Maharashtra, India, for which we are truly grateful.

References

1. David D. “Weather Based Prediction Models for Disease and Pest Using Machine Learning: A Review”. Asian Journal of Agricultural Extension. Economia e Sociologia, 41(11), pp.334-345 2023.
2. Wrzesień M., *etal*. “Prediction of the apple scab using machine learning and simple weather stations”. Computers and Electronics in Agriculture, 161, pp.252-259, 2019.
3. Ji T. *etal.* “Development and Validation of a Mechanistic, Weather-Based Model for Walnut Blight Caused by Xanthomonas arboricola pv. juglandis. Plant Disease”, 109(5), pp.1130-1141, 2025.
4. Prajwala M. *etal.* “Domain-independent adaptive histogram-based features for pomegranate fruit and leaf diseases classification”. CAAI Transactions on Intelligence Technology, 10(2), pp.317-336, 2025.
5. Sharath D.M *etal.* Image based plant disease detection in pomegranate plant for bacterial blight. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0645-0649). IEEE, 2019.
6. Sandhi A. *etal.* “Optimized deep learning framework for pomegranate disease detection using nature-inspired algorithms”. Plant Methods, 21(1), pp.1-36, 2025.
7. Gündüz R. *etal.* “Artificial Intelligence-Based Prediction of Some Pomegranate Diseases”. Turkish Journal of Agriculture-Food Science and Technology, 13(4), pp.946-957, 2025.
8. Kazi K.S.L. “Machine learning-based pomegranate disease detection and treatment”. In Revolutionizing Pest Management for Sustainable Agriculture (pp. 469-498). IGI Global, 2024.
9. Deshpande T *etal.* “Grading & identification of disease in pomegranate leaf and fruit.” International Journal of Computer Science and Information Technologies, 5(3), pp.4638-4645, 2014
10. MuniSankar M. *etal*. “Detection and Prevention of Several Diseases in Pomegranate Fruit Using Artificial Intelligence and IoT”. Design Engineering, pp.9638-9646, 2021.
11. Dar W.A. *etal.* Potato late blight disease prediction using meteorological parameters in Northern Himalayas of India. Journal of Agrometeorology, 23(3), pp.310-315, 2021.
12. Singh R. *etal* Weather driven prediction of downy mildew in broccoli deploying machine learning and time-series approaches. Journal of Plant Pathology, pp.1-11, 2021.
13. Delfani P. *etal.* “Integrative approaches in modern agriculture: IoT, ML and AI for disease forecasting amidst climate change”. Precision Agriculture, 25(5), pp.2589-2613, 2024.
14. Michael P.A. *etal.* “Weather prediction using random forest machine learning model”. Indonesian Journal of Electrical Engineering and Computer Science, 22(2), pp.1208-1215 , 2021.
15. Cha G.W. *etal*. “Development of a prediction model for demolition waste generation using a random forest algorithm based on small datasets”. International Journal of Environmental Research and Public Health, 17(19), p.6997, 2020.
16. Dou B. *etal.* “Machine learning methods for small data challenges in molecular science”. Chemical Reviews, 123(13), pp.8736-8780, 2023.