Journal of Computer Science

Research Article

Smart Weather Data Management System for Sustainable Agriculture in Maharashtra Using Machine Learning

Aishwarya V. Kadu and Kuraparthi Tirumala V. Reddy

Department of Artificial Intelligence and Data Science, Faculty of Engineering and Technology, Datta Meghe Institute of Higher Education & Research, Wardha, India

Article history
Received: 27-12-2024
Revised: 08-03-2025
Accepted: 18-03-2025

Corresponding Author:
Aishwarya V. Kadu
Department of Artificial Intelligence and Data Science, Faculty of Engineering and Technology, Datta Meghe Institute of Higher Education & Research, Wardha, India
Email: aishwaryakadu17@gmail.com

Abstract: Agriculture will benefit greatly from efficient weather data sustainable management. Data can be utilized to schedule the prediction yield, crop growth, and irrigation. This paper collects real-time meteorological station data via APIs, primarily focusing on the weather data sustainable management system and data related to Maharashtra Wardha, India, which is involved in a live series period. Agriculture can make potential through the growth of Smart Weather Data Management (SWDM) platforms. Real-time data from India, combined with intelligent insights, can be used to propose innovative and sustainable solutions. Four layers are involved: i. Acquisition ii. Storage iii. Application iv. Processing. ML checks for errors and missing values based on Land-Wardha reanalysis once the information is obtained. To assess the accuracy of the Temperature, Humidity, and Rainfall approximations with the use of some evaluation metrics with coefficients of determination (R2-Scores), (RMSE), and (MSE). The system's services include weather forecast time series, meteorological data visualization and analysis, and ML-based evaluation. Real-time services for temperature, humidity, wind speed, and cloud cover given the prediction for the Previous /Next five days in Maharashtra, Wardha, India. The platform is built with the aim of providing services and solutions that can assist both farmers and representatives.

Keywords: ML-Techniques, Real-Time Prediction, Weather Data, Agriculture

Introduction

By automating laborious processes traditionally performed by hand, advances in technology and industry have improved the quality of life and longevity. However, these advancements come with challenges, particularly the overexploitation and lack of rational use of natural resources on Earth. The effects of climate change, such as global warming and extreme weather events, have been exacerbated by these practices, disrupting the natural balance. As the world's population proliferates, we must reconsider our approach to environmental stewardship, especially in sectors like agriculture, which play a crucial role in feeding the planet (Liu *et al.*, 2020). Water resources exemplify one such challenge, with agriculture being the largest consumer, utilizing approximately 70% of global water resources (Jones *et al.*, 2017). Despite advancements like drip irrigation, inefficiencies persist without proper sustainable management. Achieving an efficient irrigation system necessitates precise monitoring of soil moisture, crop conditions, and weather patterns (Patel *et al.*, 2020).

Crop yield prediction is a crucial agricultural research field that aims to estimate crop harvest potential based on various factors, including crop genotype, environmental conditions, and sustainable management practices. Prediction models that are precise and trustworthy provide helpful information to farmers, policymakers, and stakeholders, supporting them in making educated decisions regarding resource allocation, crop management methods, and sustainable farming practices. In recent years, significant progress has been achieved in creating and enhancing agricultural production prediction systems (Sartore *et al.*, 2022). This progression is affected by incorporating new computing tools, such as machine learning algorithms and data analysis, which allow for more accurate forecasting and a better understanding of intricate interconnections between various components.

Machine Learning (ML) is a promising tool for addressing these agricultural challenges. ML models offer interdisciplinary potential, facilitating knowledge discovery and predictive tasks with performance rivaling traditional methods. ML's ability to analyze vast and diverse datasets enables the creation of comprehensive models by integrating data from various sources, including historical weather patterns and climatic data (Pande *et al.*, 2021). Integrating weather data into ML frameworks aims to develop intelligent agricultural systems that enhance crop productivity and optimize management practices.

Moreover, ML models can adapt and improve by continuously learning from new data, ensuring their relevance in dynamic agricultural environments (Sheth *et al.*, 2022). By leveraging historical data, ML-based weather analysis can enhance forecast accuracy and support sustainable farming practices, ultimately contributing to food security amidst evolving climatic conditions. Efficient weather data sustainable management is crucial for advancing sustainability and precision in agriculture (Kadu & Reddy, 2023). It is vital for various agricultural tasks such as crop growth, development, yield prediction, and irrigation scheduling. In line with technological progress, this study introduces an intelligent weather data sustainable management system designed to seamlessly collect real-time data from meteorological stations using Application Programming Interfaces (APIs) (Kadu & Reddy, 2023). Focused on Maharashtra Wardha, India, the system encompasses data spanning a live series period.

APIs are pivotal in enabling the system to access and retrieve real-time weather data from meteorological stations. They serve as the interface through which the system communicates with these stations, facilitating the efficient and automated collection of crucial weather information. By leveraging APIs, the system ensures timely data acquisition, enabling farmers and stakeholders to make informed decisions based on up-to-date weather conditions (Kadu & Reddy, 2023). Integrating machine learning techniques into the system enhances its capability to derive actionable insights from the collected real-time data (Kadu & Reddy, 2024a). This integration allows for analyzing weather patterns and generating accurate forecasts, contributing to improved agricultural decision-making processes (Kadu & Reddy, 2024b). The system's architectural framework comprises four layers: Data acquisition, data storage, data processing, and application layers, ensuring seamless handling and processing of the collected weather data (Roukh *et al.*, 2020). Evaluation metrics such as the coefficient of determination (R2-Score), Root Mean Squared Error (R.M.S.E.), and Mean Squared Error (M.S.E.) are employed to assess the accuracy of temperature, Rainfall, and humidity estimates generated by the system (Rehman *et al.*, 2022). These metrics measure the system's performance, validating its predictive capabilities and enhancing its reliability. In summary, integrating APIs into the intelligent weather data sustainable management system represents a significant advancement in agricultural technology (Bechar & Vigneault, 2016). By leveraging real-time weather data and machine learning techniques, the system aims to optimize agricultural practices and contribute to the sustainable management of farm resources in Maharashtra, Wardha, India (Popescu *et al.*, 2020).

Literature Review

Natural and human factors threaten agricultural productivity, including adverse weather conditions. These challenges can lead to low crop yields, thereby jeopardizing food security (Sharma *et al.*, 2021). With the increasing deployment of Machine Learning (ML) in intelligent farming, addressing issues related to low productivity becomes increasingly complex (Kadu & Reddy, 2023). In conclusion, given the challenges posed by declining agricultural productivity exacerbated by adverse weather conditions, ML-based solutions for managing agricultural resources must be developed to mitigate the looming global food security crisis.

Selecting appropriate ML models for predicting weather parameters presents a complex challenge due to the diversity of weather features (Benos *et al.*, 2021). This complexity is further compounded by the subjective nature of expert inputs and the need to align ML algorithms with unique data characteristics (Cao *et al.*, 2020). This study aims to review the state of ML research in weather analysis, focusing on identifying essential weather parameters for accurate predictions and recommendations (Zhang *et al.*, 2020). Furthermore, the paper summarizes the performance of ML models in weather prediction, providing valuable insights for farmers and stakeholders (Ramcharan *et al.*, 2017). This review will be instrumental for future research and organizations involved in weather analysis, aiding in selecting fundamental parameters for weather assessment frameworks and analysis tools (Chen *et al.*, 2020). ML has garnered significant attention in weather analysis due to its potential to revolutionize traditional practices (Kadu & Reddy, 2024b). Adopting ML models in weather analysis offers an innovative approach to streamlining processes and enhancing accuracy (Li *et al.*, 2022). ML models can process extensive datasets encompassing weather patterns, climate conditions, and historical data, providing precise predictions and recommendations for agricultural practices (Ramcharan *et al.*, 2017).

The practical advantages of ML extend beyond rapid data processing, with ML models capable of learning and adapting over time (Ayoub Shaikh *et al.*, 2022). This adaptability is crucial in the dynamic agricultural landscape, where seasonal changes and evolving weather conditions demand flexible solutions. By continuously learning from new data, ML models can enhance predictive accuracy and recommendations, ultimately leading to more sustainable and productive farming practices (Shethet al., 2022; Makridakis *et al.*, 2018; Voyant *et al.*, 2017). In the subsequent sections of this study, we will explore specific case studies and applications of ML in weather analysis, providing real-world examples of how these technologies are transforming agriculture and ensuring food security in an ever-changing world (Cedric *et al.*, 2022).

State of Art

Several scholars have explored the complex field of utilizing the big data explosion in agriculture, realizing that it may promote precision farming and sustainability (Paymode & Malode, 2022). Although there have been encouraging developments in other fields, big data analytics is still relatively new in the agricultural industry. This fact is emphasized by (Wang *et al.*, 2021), which provides information on various tools, data sets, algorithms and suggested solutions. The report highlights how big data analytics can revolutionize the agricultural industry and imagines a time when more readily available technologies and a more comprehensive range of data sources would enable wiser farming practices (Fuentes *et al.*, 2017; Sun *et al.*, 2017).

One notable Project described by (Sun *et al.*, 2017) was creating a complex system that included hardware for gathering agricultural data, an online application for data administration, and a mobile application for controlling irrigation. By evaluating crop compatibility in light of climatic variables, including temperature, humidity, and soil moisture, this all-encompassing method attempted to maximize agricultural techniques. In the meantime, (Cai *et al.*, 2019) developed a real-time weather station sustainable management system in response to the urgent demand for weather resilience in Indian agriculture. This program gives farmers accurate weather information to help them make decisions and reduce crop losses. It is motivated by the widespread availability of high-speed internet in rural regions.

Similarly, (Fei *et al.*, 2023) present an intelligent weather station sustainable management system that simplifies meteorological data collection using Internet of Things technologies. The technology provides farmers with actionable insights for improved agricultural Sustainable management by facilitating real-time data transfer and analysis through linked sensors and GSM modules (Jiang *et al.*, 2022).

To tackle the difficulties in managing data that arise from sensor-generated data, (Gomez Selvaraj *et al.*, 2020) suggest WALLeSMART, a platform for cloud-based data architecture. This ground-breaking solution provides a standardized architecture for collecting, analyzing, storing, and displaying large volumes of data (Xu *et al.*, 2022). It was tested on farms in Belgium and paved the way for customized, innovative agricultural services (Tanabe *et al.*, 2023). Our contributions go beyond traditional designs of data platforms and include a whole solution intended to support well-informed agricultural decision-making. Our suggested system integrates parameter estimates, multisource data fusion, and weather forecasting to provide farmers with actionable information from cutting-edge techniques.

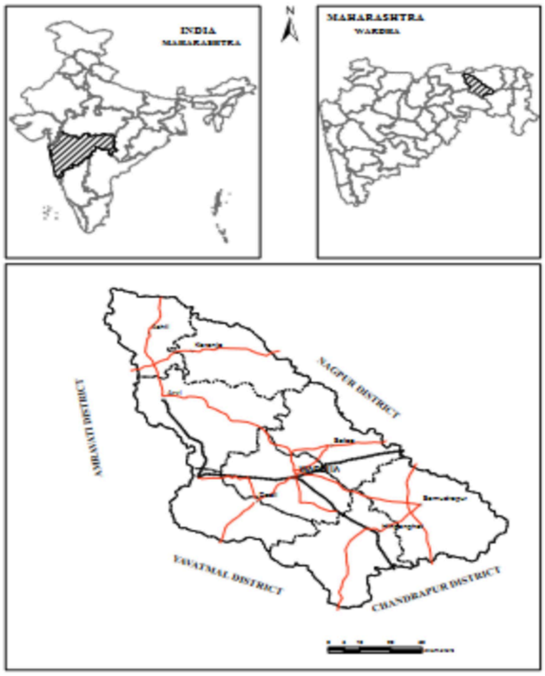


Fig. 2: Location map of Wardha dist

Study Area

In Figure (1), the Location Map of Wardha Dist. The region flourishes in a tropical monsoon climate, a vibrant tapestry woven from distinct seasons, as illustrated in Figure (1) Winters, like a gentle pause, bring comfortable temperatures from December to February, offering a break from the summer's scorching grip. Then, March ushers in the hot season, reaching its peak in May with blistering highs of 43°C (109°F). The harsh sunlight at this time makes shade essential. A dramatic shift arrives with the eagerly awaited southwest monsoon, gracing the region from late June to early October. During this period, they witnessed a downpour, contributing 88% of the annual rainfall. July takes the crown as the wettest month, bringing life-giving showers that rejuvenate the thirsty land and nurture plant life. However, Rainfall can be unpredictable, with historical records showing significant variations. Some years receive an abundance, reaching up to 154% of the average, while others experience a mere 50%. As the monsoon departs, the post-monsoon period covers October and November. This phase witnesses a gradual decline in rainfall and humidity. The once-heavy monsoon clouds disperse, revealing clear skies that dominate the rest of the year. Humidity, which peaks at around 70% during the monsoon, takes a significant plunge during the summer months, dropping to a pleasant 20% throughout the day. Winds play a subtle but essential role in shaping the climate. The area generally experiences gentle to moderate breezes. Warmer seasons see winds blowing

DOI: 10.3844/jcssp.2025.1526.1538

primarily from the northeast, refreshingly contrasting the heat. During the colder months, wind direction becomes more variable, but their strength remains low to moderate, ensuring a consistently comfortable environment.

- Climate Type: Tropical monsoon climate with distinct seasons
- Winter (December – February): Cool and comfortable temperatures
- Summer (March-May): Peak temperatures reach 43°C (109°F) in May. Harsh sunlight makes shade essential
- Monsoon Season (June – October): 88% of annual rainfall occurs during this period. July is the wettest month. Rainfall variability: Some years receive 154% of the average, while others get as low as 50%
- Post-Monsoon (October – November): Gradual decline in rainfall and humidity. Clear skies dominate as monsoon clouds disperse
- Humidity: Peaks at 70% during the monsoon. Drops to 20% in summer.
- Wind Patterns: Summer: Winds mainly from the northeast, providing some relief from the heat
- Winter: Winds become more variable but remain gentle to moderate

Materials and Methods

The proposed system, the Smart Weather Data Management System (SWDM), is crucial for intelligent farming as it optimizes agricultural practices by monitoring and analyzing weather patterns, such as irrigation scheduling and crop selection (Huang *et al.*, 2020). The platform is designed with a service-oriented architecture Figure (2) To provide services addressing the four categories of big data analytics: I. Descriptive ii. Predictive iii. Prescriptive Data Analysis. Descriptive Data Analysis- The system examines past weather data, such as observing a decline and altered frequency of rainfall over the past decade. By analyzing historical data, the system can identify trends and patterns that provide insights into long-term weather changes and temperature shifts. This analysis aims to understand what caused these changes, such as long-term weather patterns and temperature shifts (Bechar & Vigneault, 2016). Predictive Data Analysis- The next stage is predictive data analysis, where the system uses statistical methods and machine learning algorithms to predict future weather conditions based on historical data. This includes services like weather forecasting, which assist in guiding choices to avert catastrophic incidents in the future. Accurate weather predictions can help farmers plan their planting and harvesting schedules more effectively, thereby maximizing crop yields and reducing the risk of crop failure. Prescriptive Data Analysis- Finally, prescriptive data analysis provides actionable recommendations based on the predictions. By analyzing the forecasted weather conditions, the system can suggest specific actions for farmers to take. For example, it can recommend the best times for irrigation, the optimal types of crops to plant, and the necessary precautions to protect crops from adverse weather events. These recommendations aim to enhance decision-making and improve the overall efficiency of agricultural practices.

The SWDM platform is built on a robust service-oriented architecture that includes four primary layers, each playing a crucial role in the overall functionality and effectiveness of the system in managing weather data for SWDM: i. Acquisition: This layer is responsible for collecting data from various sources, such as weather stations, satellites, and sensors. It ensures that the data is accurate, reliable, and up-to-date. ii. Storage: Collected data is stored efficiently in a centralized database, allowing for easy access and analysis. The storage layer uses advanced data sustainable management techniques to handle large volumes of data and ensure data integrity. iii. Processing: Once the data is stored, it undergoes processing to check for errors and missing values. Machine learning algorithms and data validation techniques are applied to clean and preprocess the data. The system uses Land-Wardha reanalysis to ensure that the data is accurate and ready for analysis. Iv. Application: The final layer involves applying the processed data to generate insights and recommendations. This layer includes various analytical tools and visualization techniques to help users understand the data and make informed decisions. The application layer also provides a user-friendly interface for farmers to access the data and receive real-time updates on weather conditions and recommendations.

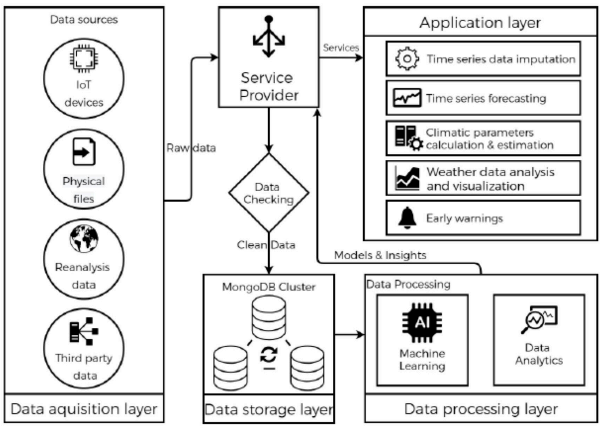


Fig. 2: Architecture of the platform (Huang *et al.*, 2020)

Data Acquisition Layer

This layer unifies heterogeneous data from several sources, including raw files (like CSV and Excel) and IoT weather sensors, reanalysis data, and third-party meteorological services that can provide real-time meteorological data via APIs. The amount, pace, and

DOI: 10.3844/jcssp.2025.1526.1538

diversity of data gathered from these disparate sources introduce the big data concepts.

Weather Station Data

The weather station data for the Wardha district is obtained through an open-source API, ensuring comprehensive and up-to-date meteorological information. The critical parameters measured include temperature (°C, one missing value), which indicates the current warmth or coldness of the atmosphere; humidity (%), reflecting the amount of water vapor present in the air; wind speed (m/s, one missing value), which measures the rate of air movement; wind direction (°), which shows the direction from which the wind is blowing; atmospheric pressure (hPa), monitoring atmospheric pressure, an essential factor in weather forecasting; and cloud cover (%), indicating the extent of the sky covered by clouds. These parameters are crucial for analyzing local weather conditions and trends, contributing to various applications such as agriculture, disaster, and daily weather forecasting. A full description of these data, including statistics for missing values, is shown in Table (1).

Table 1: Meteorological station data description

Variable	Description	Unit	Missing Values
Temperature	Air temperature	°C	1
Humidity	Relative air humidity	%	0
Wind Speed	Wind speed	m/s	1
Wind Direction	Wind direction	°	0
Pressure	Atmospheric pressure	hPa	0
Cloud Cover	Cloud cover	%	0

Data Storage Layer

For the Wardha district weather station data to be handled properly, the data storage layer is critical. It is tasked with efficiently storing, organizing, and ensuring easy access to the large volumes of meteorological data gathered through the open-source API.

Data Ingestion

The data storage layer starts by ingesting weather data and capturing raw data from the API in real-time or at scheduled intervals. Various parameters are collected automatically, including temperature, humidity, wind speed, wind direction, and pressure.

Data Storage Solutions

- Relational databases: Relational databases like PostgreSQL or MySQL are used for structured data. A database like this supports SQL queries and ensures that data integrity and consistency are maintained.
- NoSQL Databases: The flexibility and scalability of NoSQL databases like MongoDB or Cassandra allow large datasets to be stored efficiently
- Data Lakes: The use of data lakes, like those built around Hadoop or Amazon S3, can be very beneficial for historical data analysis and machine learning.

Data Organization

The data is organized into tables or collections based on parameters and time intervals (e.g., hourly, daily). Each record includes fields for Cloud Cover, Humidity, Pressure, Temperature, Wind Direction, and Wind Speed., along with timestamps and geographical metadata.

Data Quality Sustainable Management

- Validation: Incoming data is validated for accuracy and completeness, with any anomalies or errors flagged for further inspection.
- Handling missing values: Strategies are implemented to manage missing values, including imputation techniques and recording missing data statistics, as shown in Table 1.

Data Indexing

Indexes are created on crucial fields such as timestamps and geographical locations to facilitate fast data retrieval and querying.

Data Security and Access Control

Robust security measures, such as encryption, authentication, and authorization, are implemented to protect data from unauthorized access and breaches. Access control mechanisms define user permissions based on roles and responsibilities.

Integration with Analytics and Visualization Tools

The stored data is accessible to analytics and visualization tools like Tableau, Power B.I., or custom dashboards, allowing users to generate insights and make informed decisions based on weather data.

Compliance and Governance

Compliance with data governance policies and regulatory requirements is essential. This includes maintaining data lineage, audit trails, and adhering to standards like GDPR. or CCPA where applicable.

Data Processing Layer

The data processing layer receives input from the data storage layer. It applies statistical, machine learning, and deep learning models to extract data insights and transform them into services Figure (2).

Statistical Models

Statistical models are employed on this platform for forecasting. Initially, the Facebook Prophet model (Xu *et*

DOI: 10.3844/jcssp.2025.1526.1538

al., 2022) was used for long-term weather time series forecasting, as it had been tested on the same dataset in earlier work.

Machine Learning Models

A machine learning model is a mathematical representation derived from data used to make predictions or decisions without being explicitly programmed to perform the task. It is created through a training process where the model learns patterns and relationships in the data (Gasmi *et al.*, 2022). This study employs three machine learning algorithms, Linear Regression, Decision Trees, and Random Forest, to demonstrate their effectiveness across various predictive tasks (Arshaghi *et al.*, 2023).

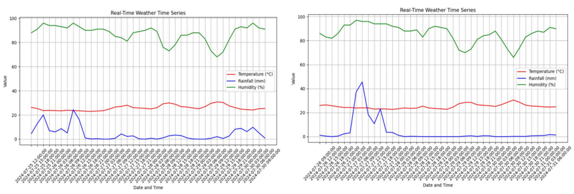
Application Layer

The application layer includes various services associated with real-time weather series data.

Real-Time Weather Series Forecasting

The provided graph, Fig. 3(a-b) illustrates the real-time weather data series forecast for Wardha, Maharashtra, India, including temperature (°C), Rainfall (mm), and humidity (%) over different periods in July 2024. In Figure 3(a), covering the period from 2024-07-25 12:00:00-2024-07-26 09:00:00. Analyzing the graph, the temperature (red line) appears relatively stable around the 25°C mark, with slight fluctuations but overall consistency. The Rainfall (blue line) shows spikes at certain intervals, indicating periods of Rainfall, but the data points appear sporadic, suggesting either intermittent Rainfall or potential missing data points. The humidity (green line) is consistently high, mostly ranging between 80 and 100%, with slight variations throughout the time series.

In Figure 3(b), covering the period from 2024-07-26 09:00:00-2024-07-31 06:00:00, the temperature (red line) remains relatively stable around the 20°C to 25°C mark, with minor fluctuations indicating overall consistency. The Rainfall (blue line) shows significant spikes, particularly around July 26, 2024, reaching up to approximately 40 mm, suggesting intermittent periods of Rainfall. The humidity (green line) is consistently high, ranging from around 80-100%, with slight variations throughout the period.



(a) (b)

Fig. 3: The Real-time weather series forecasting

In real-world meteorological datasets, missing values are expected to be encountered due to network errors or technical problems with measurement sensors. These gaps can adversely impact the performance of various models, including machine learning, numerical and physical models. Thus, it is essential to efficiently identify and manage these gaps during the Exploratory Data Analysis (EDA) and preprocessing phases to maintain the models' accuracy and reliability. The steps to implement the method workflow are shown in Figure (4).

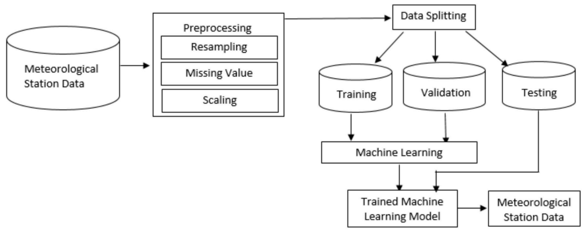


Fig. 4: The flowchart of the ML approach

Exploratory Data Analysis (EDA)

It started with Exploratory Data Analysis (EDA) before implementing the intended machine learning technique. This first stage was essential because it enabled us to make sense of the information we had gathered and create hypotheses for further research and analysis. At this point, we just used observable data and handled the data objectively, making no assumptions about the underlying connections between the variables. We created a correlation matrix Figure (5) as part of our EDA, which helped us find possible predictors for each target variable.

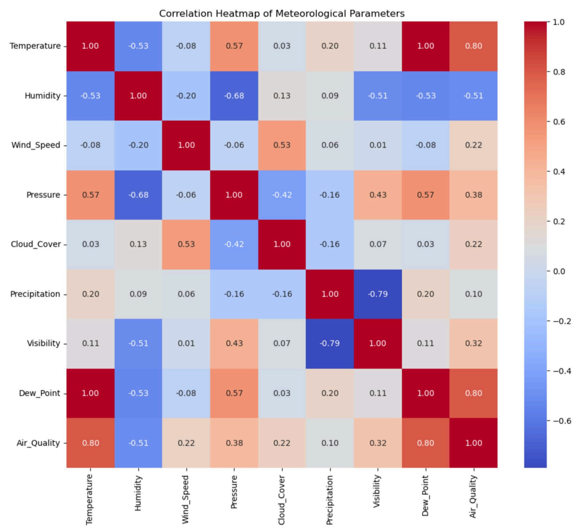


Fig. 5: The correlation heatmap of meteorological parameter

Data Splitting Process

When developing ML models, it is crucial to divide the dataset into training, validation, and test sets, each serving distinct purposes. The training set is utilized to

DOI: 10.3844/jcssp.2025.1526.1538

train the model, allowing it to learn patterns and minimize error. The validation set, which is derived from a portion of the training data, is used to tune hyperparameters and evaluate the model's performance on unseen data during training, thereby preventing overfitting. The test set provides an unbiased evaluation of the final model's performance on new data, ensuring its ability to generalize. In our approach, we allocate 80% of the dataset for training and 20% for testing.

Evaluation Metrics

When assessing the performance of regression models in machine learning, three commonly used metrics are the R^2-Score, Root Mean Squared Error (RMSE), and Mean Squared Error (MSE).

R^2-Score (Coefficient of Determination)

The R^2-Score indicates how well the independent variables explain the variability of the dependent variable. It ranges from 0-1, with 1 signifying a perfect fit and 0 indicating that the model fails to explain the variability in the data:

$$R^2 = 1 - \frac{\sum_1^n (yi - \hat{y}i)^2}{\sum_1^n (yi - \hat{y}i)^2} \tag{1}$$

Mean Squared Error (MSE)

The Mean Squared Error calculates the average of the squares of the differences between actual and predicted values. It provides a sense of the overall magnitude of prediction errors, with lower values indicating better model performance.

$$MSE = \frac{1}{n} \sum_1^n (yi - \hat{y}i)^2 \tag{2}$$

Root Mean Squared Error (RMSE)

The Root Mean Squared Error is the square root of the Mean Squared Error. It represents the average magnitude of prediction errors in the same units as the target variable, making it more interpretable:

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (yi - \hat{y}i)^2} \tag{3}$$

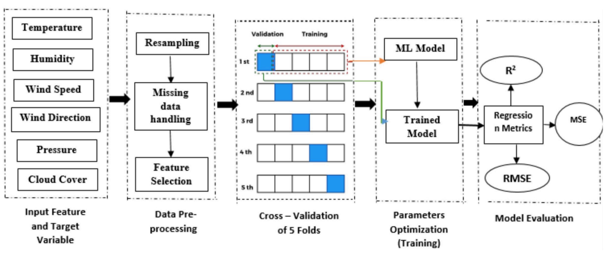


Fig. 6: The flowchart proposed method

Climatic Parameters Calculation and Estimation Service

Irrigation is a critical agricultural technique that farmers use daily. To be efficient, it is essential to accurately predict the water consumption of crops at each stage of their growth throughout the agricultural season. Figure (6) shows the flowchart of the proposed method. One method for achieving this is estimating Evapotranspiration (ET), which indicates the amount of water lost through crop transpiration and soil surface evaporation.

This service aims to give an alternative to the FAO Penman-Monteith calculation process by studying its behavior with constrained climatic variables see Figure (7). It is helpful for stations that lack the requisite hardware and sensors to produce the whole set of meteorological data required for FAO Penman-Monteith or technical issues with sensors, among other reasons.

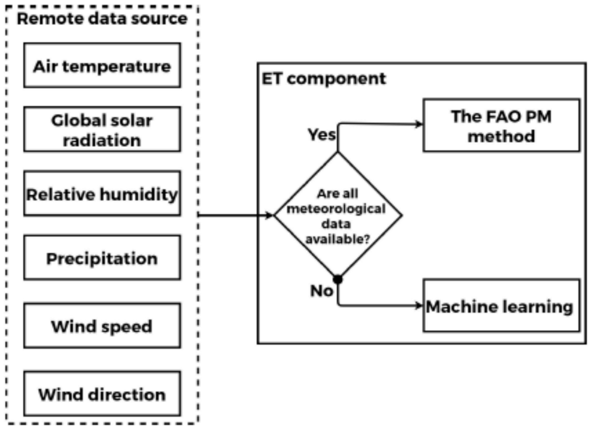


Fig. 7: The clarity of the Evapotranspiration estimation factor (Foucquier *et al.*, 2013)

Weather Data Analysis and Visualization Service

While raw data is meaningless, the information, knowledge, and insights derived from it are precious. Knowledge Discovery in Databases (KDD) is a branch of modern data science focused on transforming data from one state to another. This can be achieved through methodologies like CRISP-DM (Foucquier *et al.*, 2013) or the proposed standard method (Toyama, 2011). Figure (8) provides an example of data visualization and analysis pertinent to our use case. The hybrid process includes gathering, storing, cleaning, visualizing, analyzing, and mining data.

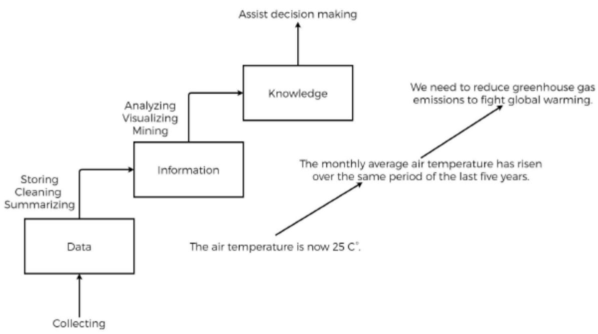


Fig. 8: Example of data visualization and analysis (Baltrušaitis *et al.*, 2019)

DOI: 10.3844/jcssp.2025.1526.1538

The initial three stages are identical to other services available in the marketplace. This service provides different visualization options, including comparison plots (weather time series line charts), relationship plots (scatter plots of weather data) and automatically generated correlation heat maps. These steps are vital in data analysis to comprehend the impact of one variable on another.

Custom Early Warning Alerts Service

This service focuses on outlier or anomaly detection. It identifies unusual data in time series using rule-based approaches, which trigger SMS and email alerts to administrators when specific criteria, like temperature or rainfall thresholds, are met. Another alert method detects sequences that significantly deviate from historical time series data. These anomalies can result from measurement errors or noise, informing managers about sensor statuses at the meteorological station and highlighting events that may require immediate action. Unlike supervised methods, we utilize unsupervised machine learning techniques such as the Local Outlier Factor (LOF) (Kadu & Reddy, 2024c), which operates without needing labeled anomaly data.

Results and Discussions

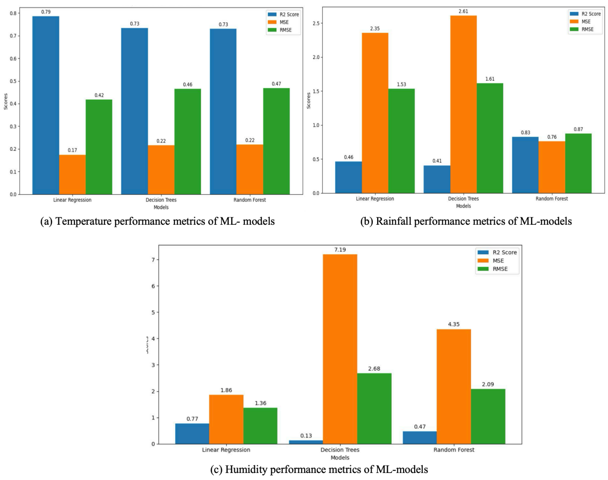
Time Series Data Imputation

Figure (7) illustrates the curves of loss functions MSE, RMSE, and R2 across training and validation sets for each epoch during the application of machine learning for handling missing data.

Table 2: Performance of machine learning models

Model	Temperature			Rainfall			Humidity		
	R2	MSE	RMSE	R2	MSE	RMSE	R2	MSE	RMSE
Linear regression	0.79	0.17	0.42	0.46	2.35	1.53	0.77	1.86	1.36
Decision trees	0.73	0.22	0.46	0.41	2.61	1.61	0.13	7.19	2.68
Random forest	0.73	0.22	0.47	0.83	0.76	0.87	0.47	4.35	2.09

Table (2) illustrates that Once trained, the machine learning model can be deployed in production to analyze data from meteorological stations. Figure (9) illustrates the performance metrics of three machine learning models—Linear Regression, Decision Trees, and Random Forest—across the prediction tasks of temperature, rainfall, and humidity.



(a) Temperature performance metrics of ML- models

(b) Rainfall performance metrics of ML-models

(c) Humidity performance metrics of ML-models

Fig. 9: Monitoring of R2, RMSE, and MSE during training and validation phases: Temperature, Rainfall, and Humidity with LR, DT, and RF

For temperature prediction in Figure 9(a), Linear Regression performs the best among the models, achieving the highest R^2 score of 0.79 along with the lowest Mean Squared Error (MSE) of 0.17 and Root Mean Squared Error (RMSE) of 0.42. This indicates that Linear Regression provides the most accurate predictions with the slightest error. The Decision Trees model shows a slightly lower R^2 score of 0.73 and higher error metrics, with an MSE of 0.22 and an RMSE of 0.46. Similarly, the Random Forest model has an R^2 score of 0.73, an MSE of 0.22, and an RMSE of 0.47. Overall, these results suggest that Linear Regression is the most effective model for temperature prediction, delivering both high accuracy and moderate error values.

In Figure 9(b), for rainfall prediction, the Random Forest model delivers the best performance among the models, achieving a notably higher R^2 score of 0.83 and significantly lower Mean Squared Error (MSE) of 0.76 and Root Mean Squared Error (RMSE) of 0.87. This superior performance highlights Random Forest's effectiveness in predicting rainfall. In comparison, Linear Regression has an R^2 score of 0.46 with an MSE of 2.35 and an RMSE of 1.53, while Decision Trees show slightly lower performance with an R^2 score of 0.41, an MSE of 2.61 and an RMSE of 1.61. Overall, these results demonstrate that Random Forest is the most accurate model for rainfall prediction, outperforming both Linear Regression and Decision Trees.

For humidity prediction depicted in Figure 9(c), the Linear Regression model outperforms the others with an R^2 score of 0.77, coupled with the lowest Mean Squared Error (MSE) of 1.86 and Root Mean Squared Error (RMSE) of 1.36. This indicates that Linear Regression provides the most accurate predictions and the slightest error for humidity. In comparison, the Random Forest model has an R^2 score of 0.47, an MSE of 4.35, and an RMSE of 2.09, which reflects better performance than Decision Trees but falls short of Linear Regression's accuracy. Decision Trees exhibit the lowest performance with an R^2 score of 0.13, an MSE of 7.19, and an RMSE of 2.68. Overall, linear Regression demonstrates the best performance for humidity prediction, while random forests perform better than decision trees but do not reach the accuracy of linear Regression.

Linear Regression provides the best performance for temperature and humidity predictions, with the highest accuracy and lowest error metrics. Random Forest excels in rainfall prediction, showing the highest R^2 score and lowest error rates. Decision Trees exhibit variable performance, with particular challenges in humidity prediction, suggesting potential overfitting issues. Overall, Random Forest and Linear Regression are the most reliable models for their respective tasks.

Climatic Parameters Calculation and Estimation

We divided our dataset into five randomly shuffled parts (five folds). One-fold was used for model evaluation, while the remaining four were used for training. We then assessed the model's performance by calculating regression metrics across the five folds, specifically the Root Mean Squared Error (RMSE), the coefficient of determination (R^2), and the Mean Squared Error (MSE). According to the results in Table (3), the main objective was to evaluate the models' performance using different parameters (Temperature, Rainfall, and Humidity) rather than achieving perfect results.

Table 3: Cross-validation results for the scenarios

Model	Metric	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Linear Regression	Temp R^2	0.74	0.79	0.1	-1.67	0.17
	Rainfall R^2	-83.99	-0.14	-1.7	-1.14	-188.23
	Humidity R^2	0.94	0.84	-0.39	0.03	0.29
	Temp MSE	0.27	0.8	0.11	0.44	-0.36
	Rainfall MSE	3.38	15.78	6.64	2.78	-2.57
	Humidity MSE	0.76	8.81	2.43	4.21	-3.71
	Temp RMSE	0.52	0.89	0.33	0.66	0.6
	Rainfall RMSE	1.84	3.97	2.58	1.67	1.6
	Humidity RMSE	0.87	2.97	1.56	2.05	1.93
Decision Trees	Temp R^2	0.33	0.69	-0.89	-2.14	0.11
	Rainfall R^2	-127.26	-0.02	-1.32	-3.39	-70.33
	Humidity R^2	0.48	0.71	-1.71	-0.78	-1.1
	Temp MSE	0.71	1.19	0.23	0.51	-0.38
	Rainfall MSE	6.01	14.19	5.72	5.72	-0.95
	Humidity MSE	6.62	-16.5	-4.75	-7.75	-11
	Temp RMSE	0.84	1.09	0.48	0.72	0.62
	Rainfall RMSE	2.45	3.77	2.39	2.39	0.97
	Humidity RMSE	2.57	4.06	2.18	2.78	3.32
Random Forest	Temp R^2	0.53	0.68	-0.77	-2.27	0.09
	Rainfall R^2	-68.57	-0.04	0.65	-2.65	-49.97
	Humidity R^2	0.26	0.72	-1.49	-0.09	-0.39
	Temp MSE	-0.63	-1.22	-0.22	-0.59	-0.37
	Rainfall MSE	2.73	13.72	1.54	4.78	-0.76
	Humidity MSE	7.45	15.34	4.6	5.37	-7.98
	Temp RMSE	0.8	1.1	0.47	0.77	0.6
	Rainfall RMSE	1.65	3.7	1.24	2.19	0.87
	Humidity RMSE	2.73	3.92	2.14	2.32	2.82

Based on Table (3), the metrics for temperature, Rainfall, and humidity predictions using Linear Regression, Decision Trees, and Random Forests, here is a detailed and consolidated analysis using statistical methods:

Temperature Prediction

Linear Regression: This is the best model with high R^2 and low MSE and RMSE values in most folds, though it performs poorly in fold 4.

Decision Trees and Random Forest: Show varying performance, with negative R^2 in some folds indicating poor fit.

Temperature Prediction: Linear Regression is the best model.

Rainfall Prediction

None of the models perform well in rainfall prediction. All models have negative R^2 values in multiple folds, indicating poor predictive performance.

Random Forest: Shows a positive R^2 in fold three but performs poorly otherwise.

Rainfall Prediction: None of the models stand out as the best for rainfall prediction based on the provided data. All models show inadequate performance, suggesting a need for different approaches or feature engineering.

Humidity Prediction

Linear Regression: Best model with high R^2 in folds 1, 2, and 3, though it performs less consistently in folds 4 and 5.

Decision Trees and Random Forest: Show varying performance with negative R^2 in some folds, indicating poor fit.

Humidity Prediction: Linear Regression is a robust model, although some inconsistency is noted in specific folds.

These results show that while the models may perform well under certain conditions, their effectiveness can vary significantly depending on the predicted parameter and the fold used for evaluation.

Prototype of the System

The platform, tailored for agriculture in India, offers real-time visualization of day-wise weather data for the Previous /Next five days, including predictions for temperature, wind speed, humidity, and cloud cover. As shown in Figure (10), this data is presented through line charts. Users can customize the displayed period according to their needs. Additionally, the platform can provide predictions for the past day.

Figure 10 (a-b) shows the temperature trends for the previous five days (27-07-2024 to 31-07-2024) and the next five days (31-07-2024 to 04-08-2024), with the y-axis representing degrees Celsius. The dots indicate the

DOI: 10.3844/jcssp.2025.1526.1538

original temperature measurements, while the line represents their linear interpolation, creating a smooth trend. Figure 10 (c) displays the humidity trend over the same period, with the y-axis showing the humidity percentage. The dots represent the original humidity measurements and the line illustrates their linear interpolation, providing a continuous trend. Figure 10 (e) presents the wind speed trend for the Previous /Next five days, with the y-axis indicating wind speed in meters per second. The dots represent the original wind speed measurements, with the line showing their linear interpolation (Bechar & Vigneault, 2016). Figure 10 (g) depicts the cloud cover trend over the Previous /Next five days, with the y-axis representing the cloud cover percentage. The dots indicate the original cloud cover measurements and the line represents their linear interpolation, providing a smooth trend.

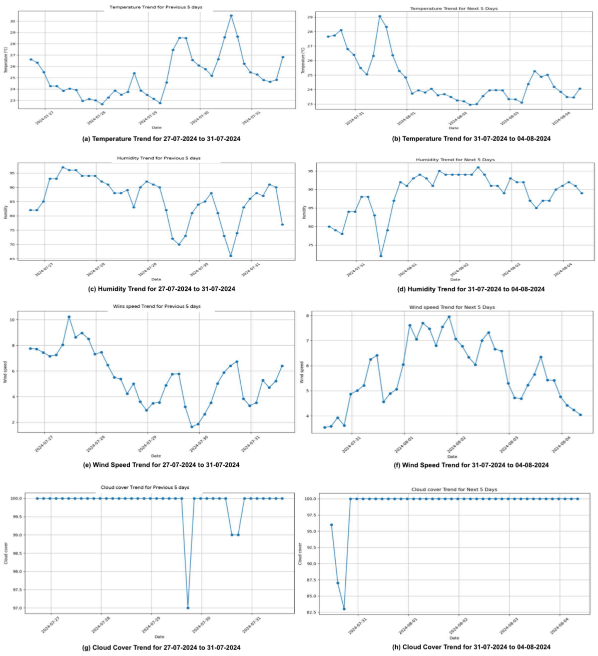


Fig. 10: A screenshot of the real-time weather time series visualization service: (a) Temperature trend for Previous five days (b) Temperature trend for next five days (c) Humidity trend for Previous five days (d) Humidity trend for next five days (e) Wind Speed trend for Previous five days (f) Wind Speed trend for next five days (g) Cloud Cover trend for Previous five days (h) Cloud Cover trend for next five days The dots represent the original measurements, while the line represents the linear interpolation of the dots

Performance Metrics

1. R^2 (coefficient of determination): A statistical measure of how well the regression line approximates the fundamental data points. A value of 1 indicates a perfect fit. Negative values indicate that the model does worse than a horizontal line.
2. RMSE (root mean squared error): Indicates the average magnitude of the difference between the predicted and (Kadu & Reddy, 2024c) actual values. Lower RMSE indicates a better fit and less spread-out residuals.
3. MSE (mean squared error): The average squared difference between the estimated and actual values. Lower MSE indicates a better fit, penalizing more significant errors more severely than RMSE.

Temperature Prediction

For temperature prediction, Linear Regression has an R^2 value of 0.03, indicating a positive but still low fit. It achieves the lowest RMSE of 0.60 and the lowest MSE of 0.40, suggesting it captures the variance relatively well. Decision Trees perform better in terms of R^2 with a value of 0.38 but have a higher RMSE of 0.75 and an MSE of 0.61. Random Forest, despite having a negative R^2 of -0.45, demonstrates a competitive RMSE of 0.59 and an MSE of 0.59. This indicates that while Linear Regression shows the lowest RMSE and MSE, Random Forest provides a more balanced performance despite the negative R^2, indicating it captures the variance relatively well.

Rainfall Prediction

In rainfall prediction, all models perform poorly, reflected by their negative R^2 values. Linear Regression has the most negative R^2 of -55.04, indicating abysmal performance. It achieves an RMSE of 2.33 and an MSE of 6.23. Decision Trees have an R^2 of -44.89, with an RMSE of 2.63 and an MSE of 6.34. Random Forest has the least damaging R^2 of -22.67, indicating relatively better performance compared to the other models. It achieves an RMSE of 1.83 and the lowest MSE of 3.21. This suggests that while none of the models perform well, Random Forest is the least poor performer in terms of RMSE and MSE.

Humidity Prediction

For humidity prediction, Linear Regression shows the best performance with an R^2 of 0.34, the highest among the models. It achieves the lowest RMSE of 1.87 and the lowest MSE of 3.98. Decision Trees have an R^2 of -0.48, an RMSE of 2.98, and an MSE of 9.32. Random Forest has an R^2 of -0.18, indicating relatively better performance compared to Decision Trees. It achieves an RMSE of 2.75 and an MSE of 8.45. These metrics indicate that Linear Regression is the best model for humidity prediction, with the highest R^2 and the lowest RMSE and MSE values.

Based on the provided metrics, Random Forest is the most robust model for predicting temperature, rainfall, and humidity. Temperature prediction offers a balanced performance with competitive RMSE and MSE values compared to linear Regression and decision trees. In rainfall prediction, although all models perform poorly, Random Forest demonstrates the least poor performance with the lowest RMSE and MSE. For humidity prediction, Linear Regression excels with the highest R^2 and the lowest RMSE and MSE, but Random Forest remains consistent across all parameters. Overall, Random Forest provides the most reliable predictions across temperature, rainfall, and humidity. Figure (11) shows these findings visually, underscoring the need for more sophisticated models to improve rainfall forecasting accuracy.

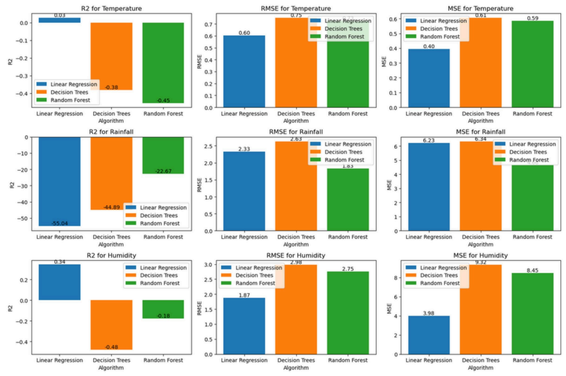


Fig. 11: Displays the results of simulations for three algorithms used to predict temperature, rainfall, and humidity: Linear Regression, Decision Trees, and Random Forests

Conclusion

The work conducted in this study integrates statistical methods (state-of-the-art) and ML techniques to manage and interpret real-time weather data, which is crucial for promoting innovative and sustainable agriculture in India. The platform uses a large amount of data to provide valuable insights and help with decision-making. The proposed platform facilitates weather-data-related services such as handling missing data, analysis, visualization, estimation, and forecasting. The proposed method gives real-time weather time series visualization service for temperature, humidity, wind speed, and cloud cover trends to predict the previous or next five days. Knowing the relationship between the three data sources gave promising results from the performance matrix (R2 = -13.33, RMSE = 1.81, and MSE= 4.35) for all parameters. The same is true for estimating the reference evapotranspiration using the Machine Learning Model with all the algorithms (Linear Regression = 1.08, Decision Trees = 1.78, and Random Forest 1.20). Linear Regression consistently shows the best performance across all metrics and weather parameters (temperature, rainfall, and humidity). The platform is built with the aim of providing services and solutions that can assist both farmers and representatives. All models highlight the challenges inherent in predicting these weather variables using the current dataset and models.

This research uniquely focuses on real-time weather data sustainable management and analysis for Wardha, Maharashtra, a region previously unexplored in such studies. By integrating advanced statistical methods and machine learning techniques, the platform provides comprehensive services like visualization, estimation, and forecasting, addressing local agricultural needs, empowering farmers with actionable insights, and enhancing food security. This pioneering approach serves as a model for extending similar solutions to other underrepresented areas.

Further improvements may require exploring additional features, more complex models, or advanced data preprocessing techniques. The models need to demonstrate more accuracy in rainfall forecasting, highlighting the need for better methods or additional features. Variability in model performance suggests that more refinement is necessary, potentially through exploring more sophisticated techniques or enhancing data preprocessing. Addressing these issues will be crucial for improving the reliability and precision of weather predictions. Future research should focus on addressing current limitations, exploring advanced models, and optimizing the use of both real-time and static data to enhance prediction accuracy and model effectiveness.

Acknowledgment

We appreciate the anonymous reviewers' thorough review and insightful suggestions, which greatly improved our work.

Funding Information

The authors did not receive any grants or funding for this study.

Author’s Contributions

Aishwarya V. Kadu: Participated in all the experiments, including data collection and analysis. Written, reviewed, and edited.

Kuraparthi Tirumala V. Reddy: Participated in data curation, investigation, and validation.

Ethics

The content presented here is the author's original research and has not been published elsewhere.

References

Arshaghi, A., Ashourian, M., & Ghabeli, L. (2023). Potato diseases detection and classification using deep learning methods. *Multimedia Tools and Applications*, *82*(4), 5725-5742. https://doi.org/10.1007/s11042-022-13390-1

Ayoub Shaikh, T., Rasool, T., & Rasheed Lone, F. (2022). Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Computers and Electronics in Agriculture*, *198*, 107119. https://doi.org/10.1016/j.compag.2022.107119

Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(2), 423-443. https://doi.org/10.1109/TPAMI.2018.2798607

Bechar, A., & Vigneault, C. (2016). Agricultural robots for field operations: Concepts and components. *Biosystems Engineering*, *149*, 94-111. https://doi.org/10.1016/j.biosystemseng.2016.06.014

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, *21*(11), 3758. https://doi.org/10.3390/s21113758

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*, 144-159. https://doi.org/10.1016/j.agrformet.2019.03.010

Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Han, J., & Li, Z. (2020). Identifying the Contributions of Multi-Source Data for Winter Wheat Yield Prediction in China. *Remote Sensing*, *12*(5), 750. https://doi.org/10.3390/rs12050750

Cedric, L. S., Adoni, W. Y. H., Aworka, R., Zoueu, J. T., Mutombo, F. K., Krichen, M., & Kimpolo, C. L. M. (2022). Crops yield prediction based on machine learning models: Case of West African countries. *Smart Agricultural Technology*, *2*, 100049. https://doi.org/10.1016/j.atech.2022.100049

Chen, C., Tsai, Y., Chang, F., & Lin, W. (2020). Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Systems*, *37*(5), e12553. https://doi.org/10.1111/exsy.12553

Fei, S., Hassan, M. A., Xiao, Y., Su, X., Chen, Z., Cheng, Q., Duan, F., Chen, R., & Ma, Y. (2023). UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat. *Precision Agriculture*, *24*(1), 187-212. https://doi.org/10.1007/s11119-022-09938-8

Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, 272-288. https://doi.org/10.1016/j.rser.2013.03.004

Fuentes, A., Yoon, S., Kim, S., & Park, D. (2017). A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors*, *17*(9), 2022. https://doi.org/10.3390/s17092022

Gasmi, A., Gomez, C., Chehbouni, A., Dhiba, D., & El Gharous, M. (2022). Using PRISMA Hyperspectral Satellite Imagery and GIS Approaches for Soil Fertility Mapping (FertiMap) in Northern Morocco. *Remote Sensing*, *14*(16), 4080. https://doi.org/10.3390/rs14164080

Gomez Selvaraj, M., Vergara, A., Montenegro, F., Alonso Ruiz, H., Safari, N., Raymaekers, D., Ocimati, W., Ntamwira, J., Tits, L., Omondi, A. B., & Blomme, G. (2020). Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin. *ISPRS Journal of Photogrammetry and Remote Sensing*, *169*, 110-124. https://doi.org/10.1016/j.isprsjprs.2020.08.025

Huang, H., Lan, Y., Yang, A., Zhang, Y., Wen, S., & Deng, J. (2020). Deep learning versus Object-based Image Analysis (OBIA) in weed mapping of UAV imagery. *International Journal of Remote Sensing*, *41*(9), 3446-3479. https://doi.org/10.1080/01431161.2019.1706112

Jiang, J., Atkinson, P. M., Zhang, J., Lu, R., Zhou, Y., Cao, Q., Tian, Y., Zhu, Y., Cao, W., & Liu, X. (2022). Combining fixed-wing UAV multispectral imagery and machine learning to diagnose winter wheat nitrogen status at the farm scale. *European Journal of Agronomy*, *138*, 126537. https://doi.org/10.1016/j.eja.2022.126537

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., & Wheeler, T. R. (2017). Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agricultural Systems*, *155*, 269-288. https://doi.org/10.1016/j.agsy.2016.09.021

Kadu, A., & Reddy, K. T. V. (2023). Enhancing Crop Prediction Accuracy Via Soil and Environmental Feature Selection Techniques in Machine Learning. *2023 1st DMIHER International Conference on Artificial Intelligence in Education and Industry 4.0 (IDICAIEI)*, 1-6. https://doi.org/10.1109/idicaiei58380.2023.10406703

Kadu, A., & Reddy, K. T. V. (2024). Comprehensive analysis of deep and machine learning approaches for predicting crop yields. *AIP Conference Proceedings*, *3188*(1), 080002. https://doi.org/10.1063/5.0240204

Kadu, A. V., & Reddy, K. T. V. (2024a). Agriculture Yield Forecasting via Regression and Deep Learning with Machine Learning Techniques. *Communication and Intelligent Systems*, 219-233. https://doi.org/10.1007/978-981-97-2079-8_18

Kadu, A. V., & Reddy, K. T. V. (2024b). Empowering Agriculture: Smart Crop Recommendation with Machine Learning Challenges and Future Prospects. *Recent Advances in Science, Engineering & Technology*, 176-182. https://doi.org/10.1201/9781003527442-20

Li, D., Miao, Y., Ransom, C. J., Bean, G. M., Kitchen, N. R., Fernández, F. G., Sawyer, J. E., Camberato, J. J., Carter, P. R., Ferguson, R. B., Franzen, D. W., Laboski, C. A. M., Nafziger, E. D., & Shanahan, J. F. (2022). Corn Nitrogen Nutrition Index Prediction Improved by Integrating Genetic, Environmental, and Management Factors with Active Canopy Sensing Using Machine Learning. *Remote Sensing*, *14*(2), 394. https://doi.org/10.3390/rs14020394

Liu, Y., Ma, X., Shu, L., Hancke, G. P., & Abu-Mahfouz, A. M. (2020). From Industry 4.0 to Agriculture 4.0: Current status, enabling technologies, and research challenges. *IEEE Transactions on Industrial Informatics*, *17*(6), 4322-4334. https://doi.org/10.1109/TII.2020.3003910

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, *13*(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Pande, S. M., Ramesh, P. K., Anmol, A., Aishwarya, B. R., Rohilla, K., & Shaurya, K. (2021). Crop Recommender System Using Machine Learning Approach. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 1066-1071. https://doi.org/10.1109/iccmc51019.2021.9418351

Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation system and prospective extensions. *Computers and Electronics in Agriculture*, *178*, 105779. https://doi.org/10.1016/j.compag.2020.105779

Paymode, A. S., & Malode, V. B. (2022). Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG. *Artificial Intelligence in Agriculture*, *6*, 23-33. https://doi.org/10.1016/j.aiia.2021.12.002

Popescu, D., Stoican, F., Stamatescu, G., Ichim, L., & Dragana, C. (2020). Advanced UAV-WSN System for Intelligent Monitoring in Precision Agriculture. *Sensors*, *20*(3), 817. https://doi.org/10.3390/s20030817

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep Learning for Image-Based Cassava Disease Detection. *Frontiers in Plant Science*, *8*, 1852. https://doi.org/10.3389/fpls.2017.01852

Rehman, A., Saba, T., Kashif, M., Fati, S. M., Bahaj, S. A., & Chaudhry, H. (2022). A revisit of internet of things technologies for monitoring and control strategies in smart agriculture. *Agronomy*, *12*(1), 127. https://doi.org/10.3390/agronomy12010127

Roukh, A., Fote, F. N., Mahmoudi, S. A., & Mahmoudi, S. (2020). WALLeSMART: Cloud Platform for Smart Farming. *SSDBM '20: Proceedings of the 32nd International Conference on Scientific and Statistical Database Management*, 1-4. https://doi.org/10.1145/3400903.3401690

Sartore, L., Rosales, A. N., Johnson, D. M., & Spiegelman, C. H. (2022). Assessing machine leaning algorithms on crop yield forecasts using functional covariates derived from remotely sensed data. *Computers and Electronics in Agriculture*, *194*, 106704. https://doi.org/10.1016/j.compag.2022.106704

Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2021). Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access*, *9*, 4843-4873. https://doi.org/10.1109/access.2020.3048415

Sheth, V., Tripathi, U., & Sharma, A. (2022). A Comparative Analysis of Machine Learning Algorithms for Classification Purpose. *Procedia Computer Science*, *215*, 422-431. https://doi.org/10.1016/j.procs.2022.12.044

Sun, L., Yang, Y., Hu, J., Porter, D., Marek, T., & Hillyer, C. (2017). Reinforcement Learning Control for Water-Efficient Agricultural Irrigation. *2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, 1334-1341. https://doi.org/10.1109/ispa/iucc.2017.00203

Tanabe, R., Matsui, T., & Tanaka, T. S. T. (2023). Winter wheat yield prediction using convolutional neural networks and UAV-based multispectral imagery. *Field Crops Research*, *291*, 108786. https://doi.org/10.1016/j.fcr.2022.108786

Toyama, K. (2011). Technology as amplifier in international development. *IConference '11: Proceedings of the 2011 IConference*, 75-82. https://doi.org/10.1145/1940761.1940772

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 569-582. https://doi.org/10.1016/j.renene.2016.12.095

Wang, S., Guan, K., Wang, Z., Ainsworth, E. A., Zheng, T., Townsend, P. A., Liu, N., Nafziger, E., Masters, M. D., Li, K., Wu, G., & Jiang, C. (2021). Airborne hyperspectral imaging of nitrogen deficiency on crop traits and yield of maize by machine learning and radiative transfer modeling. *International Journal of Applied Earth Observation and Geoinformation*, *105*, 102617. https://doi.org/10.1016/j.jag.2021.102617

Xu, P., Tan, Q., Zhang, Y., Zha, X., Yang, S., & Yang, R. (2022). Research on maize seed classification and recognition based on machine vision and deep learning. *Agriculture*, *12*(2), 232. https://doi.org/10.3390/agriculture12020232

Zhang, N., Yang, G., Pan, Y., Yang, X., Chen, L., & Zhao, C. (2020). A Review of Advanced Technologies and Development for Hyperspectral-Based Plant Disease Detection in the Past Three Decades. *Remote Sensing*, *12*(19), 3188. https://doi.org/10.3390/rs12193188