PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 12(1): 1-19
https://doi.org/10.14719/pst.5976

HORIZON
e-Publishing Group
HePG

REVIEW ARTICLE

Rice yield predictions using remote sensing and machine learning algorithms: A review

Ajay prakash M[1], Ragunath KP[2*], Pazhanivelan S[2], Muthumanickam D[1], Sivamurugan AP[2] & Vanitha G[3]

[1]Department of Remote Sensing and Geographic Information System, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

[2]CWGS (Centre for Water and Geospatial Studies), Tamil Nadu Agricultural University 641 003, Tamil Nadu, India

[3]Office of Dean, SPGS (School of Post Graduate Studies), Tamil Nadu Agricultural University, Coimbatore 641 003,Tamil Nadu, India

*Email: ragunathkp@tnau.ac.in

OPEN ACCESS

ARTICLE HISTORY
Received: 20 October 2024
Accepted: 29 October 2024
Available online
Version 1.0 : 06 February 2025
Version 2.0 : 07 February 2025

Check for updates

Additional information
Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE
Ajay PM, Ragunath KP, Pazhanivelan S, Muthumanickam D, Sivamurugan AP, Vanitha G . Rice yield predictions using remote sensing and machine learning algorithms: A review. Plant Science Today. 2025; 12(1): 1-19. https://doi.org/10.14719/pst.5976

Abstract

Crop yield prediction is becoming increasingly crucial due to global food security concerns, as highlighted by recent reports from the World Health Organization. Accurate early predictions can mitigate famine risks by estimating food supply, which is essential for 820 million people facing hunger globally. Rice is the primary staple food consumed worldwide; therefore, global rice yield and rice area are monitored using emerging technologies such as remote sensing (RS) and machine learning (ML). These technologies provide valuable tools for enhancing rice yield predictions. RS includes critical information on crop health, soil conditions and weather patterns. In contrast, ML algorithms analyze these datasets to identify patterns and relationships that affect yield. Integrating these technologies offers promising improvements in yield forecasting accuracy, with applications showing successful yield predictions 1-3 months before harvest. Various ML techniques, including Random Forest, Support Vector Machines and deep learning models such as LSTM (Long-Short Term Memory), have been employed, often in combination with RS data. However, these models face challenges, such as data quality, managing high-dimensional RS data and accounting for spatial and temporal variability. Despite these challenges, integrating RS and ML has significant potential for advancing precision agriculture and achieving sustainable food production. This study explores the advancements, practical applications and challenges associated with using RS and ML for rice yield prediction, emphasizing the importance of these technologies in addressing global food security and promoting sustainable agricultural practices.

Keywords

crop yield; machine learning; remote sensing; rice yield; yield prediction

Introduction

Oryza sativa, commonly known as the cultivated rice plant, is a yearly grass belonging to the Gramineae family. This plant yields rice, a consumable grain rich in starch. Half the global population, particularly in East and Southeast Asia, relies exclusively on rice as their primary food source. Developing countries account for 95 % of the total rice production, with China and India contributing nearly half of the world's (1). Crop yield prediction is becoming increasingly important due to growing concerns about food security, as shown by recent reports from the World Health Organization (2). Early crop yield predictions are crucial in reducing famine by estimating the available food supply for the growing global population (3). This is particularly important because hunger is a significant issue worldwide and increasing crop yield is a viable solution to address this problem and world health organization (WHO) reported that there is still an inadequate food supply for approximately 820 million people globally (1). The United Nations Sustainable Development Goals aim to end hunger, achieve food security and promote sustainable agriculture by 2030 (4). Therefore, crop yield prediction is essential for providing critical information that can help develop practical solutions to achieve these goals and end hunger.

According to FAO 2022, China is the leading producer of rice, with a peak production of 208.49 mt, followed by India, Bangladesh, Indonesia and Africa closely behind, with rice production levels of 196.25 mt, 57.19 mt, 54.75 mt and 39.88 mt respectively (3). On the other hand, Australia recorded the highest rice yield at 11.05 tons per hectare. Egypt came second, with a yield 8.97 tons per hectare, followed by Peru, United States, Morocco, Greece and China, which had yields of 6.64, 5.66, 8.33, 8.28, 7.77, 7.46 and 7.08 tons per hectare, respectively which was represent as shown in Fig. 1. Additionally, based on the global rice production and yield map, a trend chart spanning 1961-2022 for Asia is presented in Fig. 2. In India, rice is a fundamental staple food crop, crucial in the nation's food and livelihood security system. It also contributes significantly to the country's foreign exchange earnings through exports. Over the past 66 years, India has witnessed a remarkable increase in rice production, with output growing fivefold and yield per hectare quadrupling. The cultivated area for *kharif* crops reached 41315 km^2, while *rabi* crops covered 38,360 km^2 (5). The total rice cultivation area is approximately 45152 km^2 (4). Rice production in India for *kharif* crops totalled 11145800 metric tons, with rabi crops yielding 12357 mt. The overall rice production in India amounts to 12381500 metric tons.

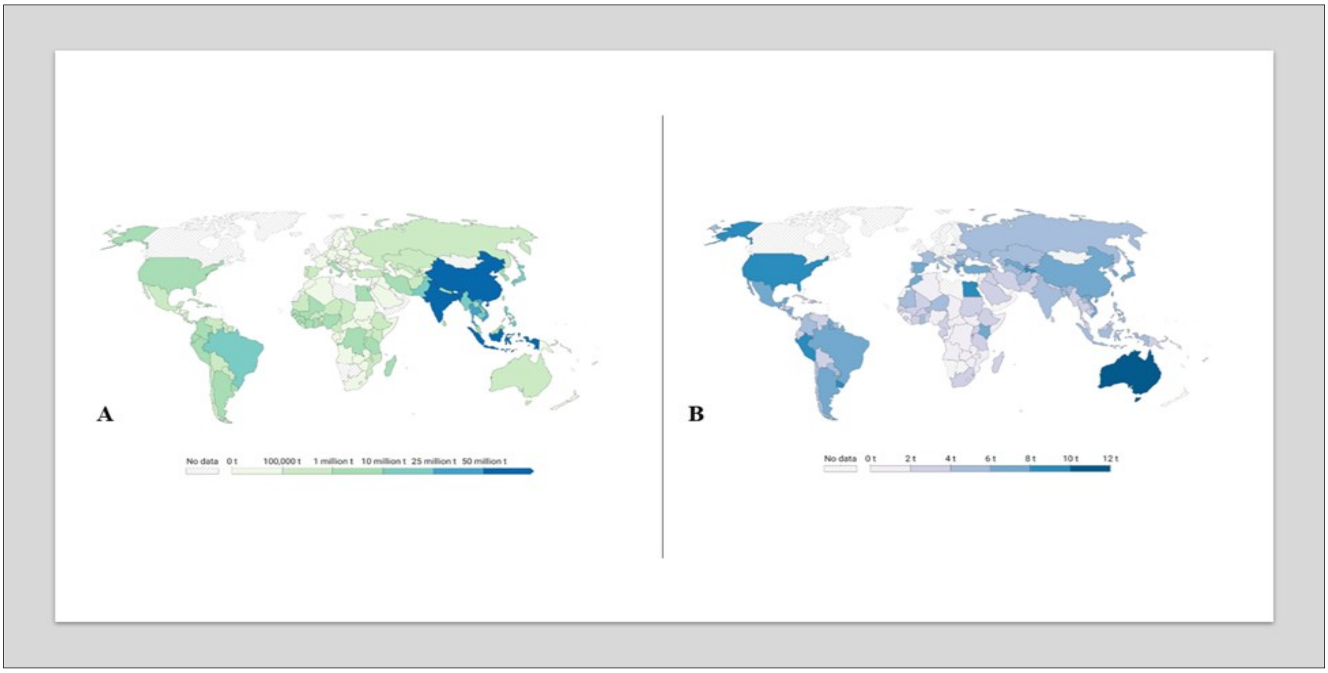

Fig. 1. (A) Rice production (Tons)2022. [Source: Food and Agriculture Organization of the United Nations processed by our world data. (B) Rice yield (t/ha)" [dataset]. Food and Agriculture Organization of the United Nations [original data].

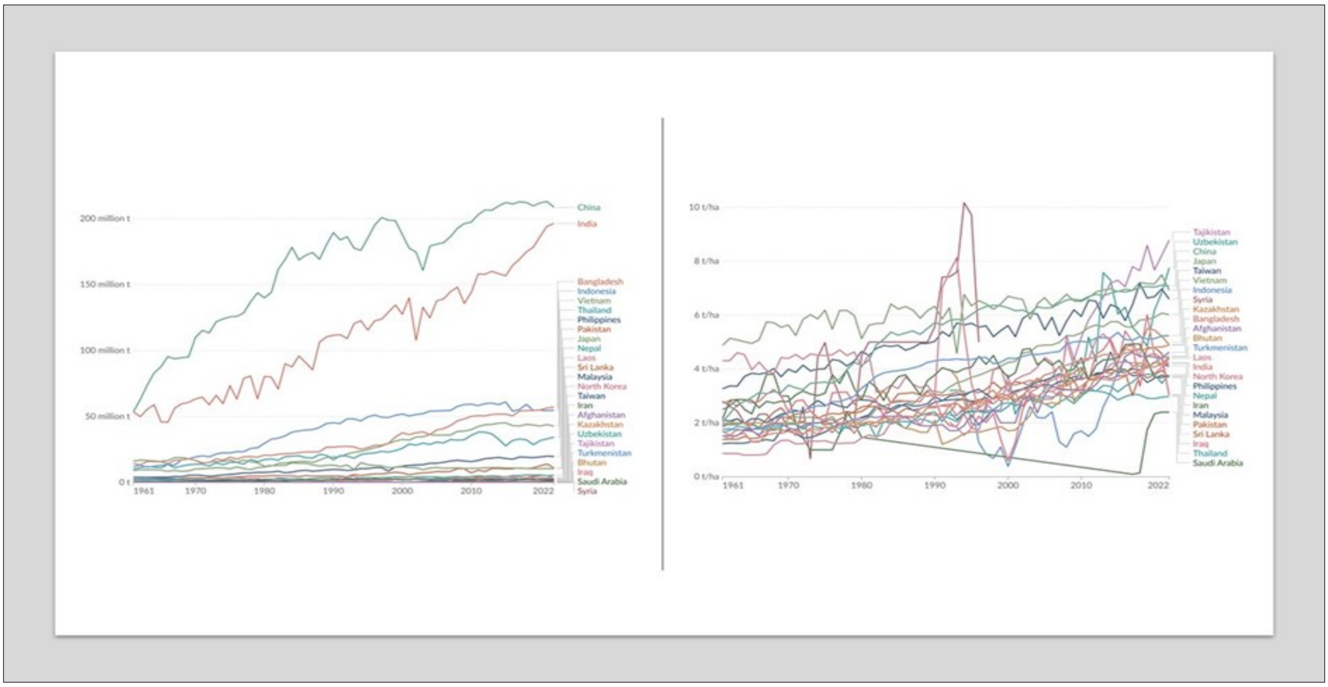

Fig. 2. (A) Trend of rice production (t) (1961-2022). [Source: Food and agriculture organization of the United Nations-processed by our world data. (B) Trend of Rice yield (t/ha) (1961-2022). Food and Agriculture organization of the United Nation [original data].

Furthermore, rice productivity in India for *kharif* crops was 2,698 kg/ha, while *rabi* crops achieved 3,221 kg/ha. The total rice productivity in India stands at 2,742 kg/ha. This remarkable increase in agricultural output can be attributed to the adoption sophisticated and efficient farming technologies. This study aims to consolidate fundamental principles in assessing rice yields using remote sensing (RS) and geospatial methodologies, uncover research limitations and suggest future research directions. Fig. 3 presents a diagrammatic overview of the concepts explored in this study.

Additionally, this review examines the obstacles and constraints of current RS-based yield estimation and ML approaches. One of the primary challenges is the complexity and redundancy of high-dimensional hyperspectral (HS) signals, which can affect the performance of ML models (6). The effects of spectral variabilities in HS imaging further complicate the analysis, limiting the ability of conventional ML tools to handle complex practical problems. Interestingly, while ML techniques have shown promise in various RS applications, there are contradictions in their effectiveness. For instance, in geothermal exploration, ML methods have demonstrated high accuracy (92-95 %) in predicting geothermal potential using surface manifestations. However, when applied to different sites with varying characteristics, the accuracy drops to 72-76 % (7). This highlights the challenge of developing ML models that generalize well across geographical contexts. This highlights the necessity for enhanced spatial and temporal resolution in satellite imagery.

Remote sensing of paddy (Rice) area

Advanced remote sensing techniques have proven effective for identifying and mapping rice-growing areas. Sensors leverage the unique spectral signatures of various surfaces in rice paddies to extract essential information about vegetation health and growth stages, which are key factors in predicting final crop yield. The accuracy of yield predictions is influenced by the type of sensor employed and its spatial and temporal resolution capabilities. This is particularly significant given the diverse nature of paddy ecosystems, which vary in size, management practices and environmental conditions. Remote sensing techniques offer valuable tools for assessing vegetation health and productivity. One key measure derived from these techniques is the Leaf Area Index (LAI), which quantifies the total leaf surface area relative to a specific ground area. LAI is a crucial indicator of crop yield and is frequently incorporated into yield forecasting models. Two important indices that can be calculated using remote sensing data are the Normalized Difference Vegetation Index (NDVI) and the Fraction of Photosynthetically Active Radiation (FPAR) (8). These indices provide insights into various aspects of vegetation health and productivity: NDVI evaluates the contrast between near-infrared light (reflected by healthy vegetation) and red light (absorbed by plants). It helps gauge vegetation health, greenness and vitality, offering a comprehensive view of plant condition.

FPAR indicates the proportion of incoming solar radiation that is utilized in photosynthesis. It provides insights into how efficiently the vegetation canopy converts light into energy for growth. By analyzing these remote sensing-derived indices alongside LAI, researchers and agronomists can better understand crop health, potential yield and overall ecosystem productivity (9). This information is invaluable for agricultural management, environmental monitoring and climate change studies. Lastly, accurate rice area mapping is crucial as yield is an area-dependent quantity. Current trends in mapping rice areas include modern technologies such as deep learning, sensor fusion, machine learning and traditional pixel-based methods. Remote sensing techniques offer significant advantages over conventional agriculture yield estimation methods. Contemporary approaches provide more accurate, timely and cost-effective crop monitoring and yield prediction solutions. There are several key advantages to utilizing remote sensing technologies, such as large-scale coverage, temporal analysis and integration of multiple data sources that provide more comprehensive, accurate and timely information, enabling improved agricultural management decision-making and

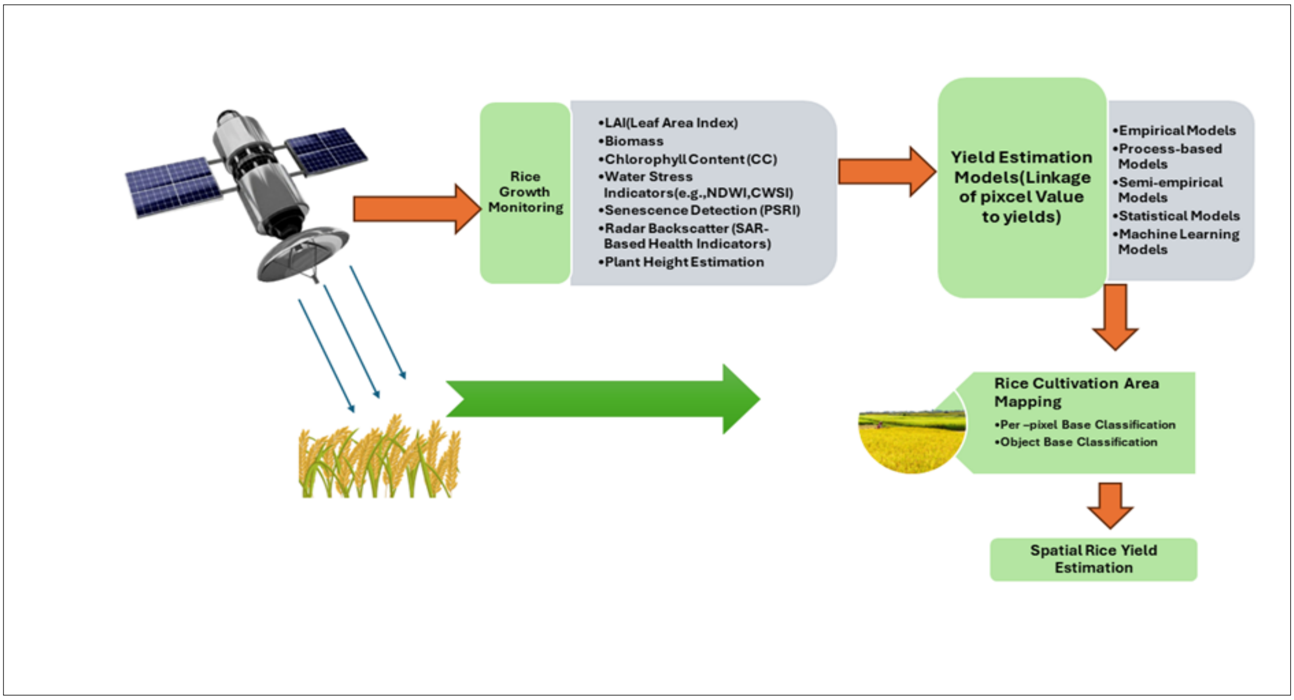

Fig. 3. Remote sensing-based rice yield estimation.

contributing to sustainable farming practices. Machine learning and deep learning algorithms have demonstrated superior accuracy in yield prediction compared to conventional approaches. For instance, a study using LSTM (Long-Short Term Memory) achieved high-performance metrics with 0.989 recall, 0.979 precision and 0.984 F1 score for rice yield estimation (10). These advanced models can capture complex non-linear relationships between various factors affecting crop yields. These advanced techniques offer improved accuracy, spatial coverage, timeliness and the ability to integrate diverse data sources, making them powerful tools for enhancing rice yield estimation and supporting sustainable agricultural practices.

The life cycle of rice

Rice (*Oryza sativa L.)* undergoes several distinct growth stages throughout its lifecycle, each with unique characteristics and requirements. The main phases of rice development include seedling, vegetative and reproductive stages Fig. 4 (6). The life cycle of rice typically spans 100 to 210 days, with most cultivars falling between 110 and 150 days (11). In temperate climates, the average duration from sowing to harvest is 130 to 150 days. The rice plant's development can be divided into three main phases: vegetative, reproductive and ripening, which are influenced by temperature and day length (12). The vegetative stage involves leaf formation and tillering for plant development. During this phase, the plant is particularly susceptible to environmental stresses, including drought and salinity (13, 14). The reproductive stage, which includes panicle initiation, booting and flowering, is critical for determining the number of spikelets per branch and overall yield potential (15). The ripening or grain-filling stage ensues when the plant allocates resources to grain development. Grain-filling constitutes a critical stage for rice yield and quality formation, with nitrogen playing a significant role in this process. Nitrogen applications have been demonstrated to prolong the duration of grain filling for both superior and inferior grains in rice (16). Optimal temperature ranges are crucial for each stage of plant growth and any temperature fluctuations outside these ranges can substantially decrease crop yields, potentially resulting in complete crop loss. Reduced sunlight exposure or insufficient solar radiation can negatively impact crop production. Moreover, inadequate and excessive precipitation, often associated with extreme weather events, can also lead to diminished crop yields (17).

Machine learning (ML) for rice

On the other hand, Machine learning (ML) has been used in

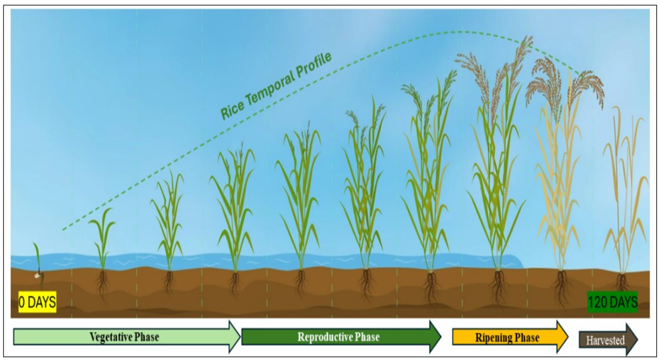

Fig. 4. Agronomic growth stages of rice (*Oryza Sativa*) are divided into three primary phases: vegetative, reproductive and ripening. Each contains distinct sub-phases that characterize specific developmental processes within each stage.

agriculture for several years (18). A primary focus of machine learning is to produce models automatically. A model is a pattern, plan, representation, or description designed to show the working of a system or concept, such as rules to determine mathematical operations and obtain specific results, a function mapping sets of formulae to formulae, or patterns (models) that can be used to generate things or parts of a thing from data (19). This (ML), a subfield of artificial intelligence (AI) dedicated to learning, is a practical method that can effectively predict crop yield based on various features. By analyzing datasets and uncover hidden knowledge, machine learning can identify patterns and relationships.

Several examples of soil characteristics frequently incorporated into machine learning models include macronutrient and micronutrient content, pH levels and moisture content (20). For instance, root-zone soil moisture and soil wetness have been identified as significant factors in oil palm yield prediction (21). Regarding meteorological factors, weather data plays a crucial role in these models, with parameters such as rainfall frequency, temperature, cloud cover, number of precipitation days and wind speed being particularly influential. Some studies have even explored the impact of lagged weather data, or "look-back periods," on model performance for soil temperature estimation (22). In crop yield and variety, machine learning techniques have been successfully applied to classify various crop types and varieties, demonstrating high accuracy in distinguishing different plant species. For example, a study using 3D LIDAR sensors and supervised learning achieved over 98 % accuracy in identifying six other plant species commonly found in nurseries (23). In another study, a multiclass classification model was developed to identify seven varieties of dry beans using machine learning algorithms. The CatBoost algorithm achieved the highest overall mean accuracy of 93.8 % in these classified bean varieties (24). ML models can be either descriptive or predictive, depending on the research problem and research questions (19).

Machine Learning (ML) has emerged as a powerful tool in rice production, offering significant improvements across various cultivation and post-harvest processes. ML algorithms are being applied to analyze data from sensors and IoT devices, transforming traditional rice farming into smart or precision agriculture (25). In rice production, ML is utilized for multiple purposes, including smart irrigation, yield estimation, growth monitoring, disease detection, quality assessment and sample classification (26). For instance, ML models can predict crop yields by evaluating structural characteristics, meteorological data and climatic factors (27). The region of Vietnam demonstrated significant improvements in rice yield forecasting using a framework that employed higher-order spatial independent component analysis and a combination of principal component analysis and ML. This approach improved subregional rice yield forecasting models by an average of 20 % up to 60 % compared to traditional methods, generating predictions 1-2 months ahead of harvest with an average error of only 5 % (28).

Other studies conducted in Michigan's non-irrigated corn, soybean and winter wheat crops showed that advanced deep learning techniques, particularly XGBoost, consistently outperformed other methods in crop yield estimation accuracy.

These models achieved remarkable precision, predicting yields with only a 7.5 % margin of error an entire month before harvest (29). This capability is crucial for addressing the growing food demand and preventing starvation (30). ML techniques are also employed in soil analysis, irrigation control and farm equipment automation, contributing to more efficient and sustainable rice cultivation practices (25).

Interestingly, while ML shows excellent promise in rice farming, its effectiveness heavily depends on the quality of data collected from sensors (14). This highlights the importance of integrating ML with technologies like Big Data and the Internet of Things (IoT) for optimal results. Additionally, adopting ML technologies in agriculture faces data quality issues, infrastructure limitations and high implementation costs, particularly for small-scale farmers in developing regions (31). ML is pivotal in revolutionizing rice production by enabling data-driven decision-making, optimizing resource use and enhancing overall productivity. Integrating ML with other technologies like IoT and autonomous farming equipment is paving the way for more advanced predictive models and the widespread use of smart farming systems in rice cultivation. The IoTML-SIS (Internet of Things Machine Learning enabled smart irrigation system) utilizes various IoT-based sensors to collect data on soil moisture, humidity, temperature and light conditions in the farmland. This data is then transmitted to a cloud server for processing and decision-making. The system employs an artificial algae algorithm (AAA) in conjunction with a least squares-support vector machine (LS-SVM) model to classify when irrigation is necessary. By optimizing the LS-SVM parameters using the AAA, the system achieves a high classification efficiency, with a maximum accuracy of 0.975 (32). However, addressing challenges related to data quality, accessibility and ethical concerns will be crucial for the equitable and sustainable adoption of ML in rice farming across different scales and regions.

Integrating remote sensing and machine learning (ML) for rice

Integrating remote sensing and machine learning (ML) is a growing field of study that aims to enhance the analysis and interpretation of data collected from various sensors to monitor and understand Earth systems, which consist of atmosphere, Hydrosphere, Geosphere, Biosphere, Cryosphere and Anthroposphere (33). This integration leverages the strengths of ML algorithms in handling the high-dimensional data characteristic of remote sensing applications, such as land use/land cover (LULC) classification, environmental monitoring and water quality assessment (34,35). Integrating remote sensing and ML is a dynamic and evolving field with significant promise for advancing environmental research and sustainable management practices.

Remote sensing and ML techniques are increasingly integrated to enhance rice monitoring and management. Satellite imagery combined with UAV data can provide large-scale and precision agricultural monitoring of rice crops, as shown in Fig. 5 (36). ML algorithms like support vector machines (SVM) and neural networks have demonstrated superior performance in estimating key rice parameters such as (LAI) from fused satellite and UAV data (21). Fusing these data sources has improved the inversion accuracy of the Leaf Area Index (LAI) for damaged rice, with R^2 increasing by approximately 0.3 and RMSE decreasing by about 0.1 (21).

These integrated approaches enable more accurate predictions of rice yield and optimal nitrogen management. For example, a variational autoencoder (VAE) model using soil, remote sensing, climate and farming practice data accurately predicted rice yield, with an average yield increase of 4.32 % when applied in practice (37). The VAEs are powerful generative models that encode input data into a latent space and then decode it to reconstruct the original input. The VAE model consists of an encoder network that maps input data to a latent representation and a decoder network that reconstructs the input from the latent space (38). The key feature of VAEs is using variational inference to learn a probabilistic encoding of the input data. The encoder produces a distribution over latent variables, typically modelled as a Gaussian, from which latent codes are sampled. This stochastic encoding allows VAEs to generate diverse outputs and capture uncertainty in the data (39).

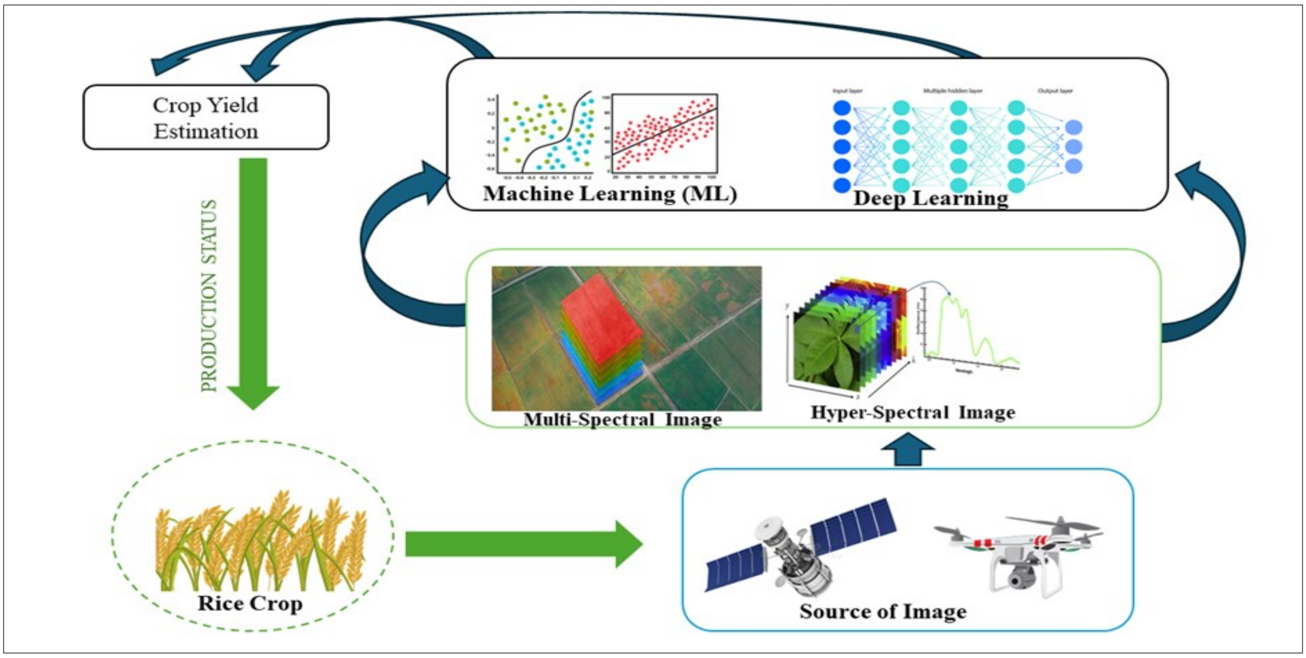

Fig. 5. Integrating remote sensing and machine learning.

Integrating Big Data, ML and IoT technologies transforms traditional rice farming practices into smart or precision agriculture (14). This integration enables various applications such as smart irrigation, yield estimation, growth monitoring, disease detection and quality assessment (14). However, challenges remain in building sufficiently labelled datasets for training ML models using satellite imagery, particularly when utilizing all available Sentinel-2 bands ranging from visible light to short-wave infrared (40). Despite limited field sampling, these approaches provide high-accuracy mapping of rice paddy distribution, growth stages and crop health. The pheno-deep method, which couples phenological methods with deep learning, has shown promising results in achieving high mapping accuracy without the need for extensive field samples as these technologies continue to evolve, they are expected to play a crucial role in optimizing rice production and supporting sustainable agricultural practices.

Materials and Methods

Search Strategies

The current review offers a comprehensive examination of international studies on rice yield prediction, focusing on various aspects. It investigates the application of Machine Learning and Remote Sensing technologies in these fields, highlighting recent advancements, practical uses and advantages and disadvantages of these technologies in rice cultivation. An extensive search through academic databases, including Google Scholar, Scopus, ResearchGate and Web of Science, was conducted for relevant literature published between 1981 and 2024, as shown in Fig. 6. Approximately 150 articles were collected and subsequently filtered. From this collection, approximately 111 articles were utilized.

This process resulted in the collection of 150 publications from multiple countries, as shown in Fig. 7. Among these publications, seventy-six were focused on the latest advancements in rice yield prediction, utilizing Machine Learning or Remote Sensing technologies.

Publications demonstrated the efficient use of ML and/or RS in aspects of rice cultivation, such as crop identification, yield prediction and acreage evaluation. A comprehensive search strategy was employed to gather relevant publications across multiple databases, incorporating specific keywords such as 'rice yield' and 'rice crop. Various tasks can be accomplished using machine learning and remote sensing, such as rice classification and yield prediction. To identify relevant literature on remote sensing and machine learning in rice cultivation, keywords like 'Remote Sensing', 'Yield Prediction' and 'Rice Yield' were used. The word count analysis of the review indicates that the most frequently used terms were "Remote Sensing" (21 %), "Yield Prediction" (17 %), "Rice Yield" (16 %) and "Machine Learning" (15 %). In contrast, less frequently used terms included "Sensing Data" (7 %), "ML Algorithm" (5 %) and "Support Vector," which was not mentioned (0 %). The results were analyzed in a word count analysis presented in Fig. 8 of this review, focusing on ML-based remote sensing applications for managing rice crops. These keywords helped obtain a wide range of research papers from multiple databases, with recent articles providing access to older publications for a deeper understanding of the field.

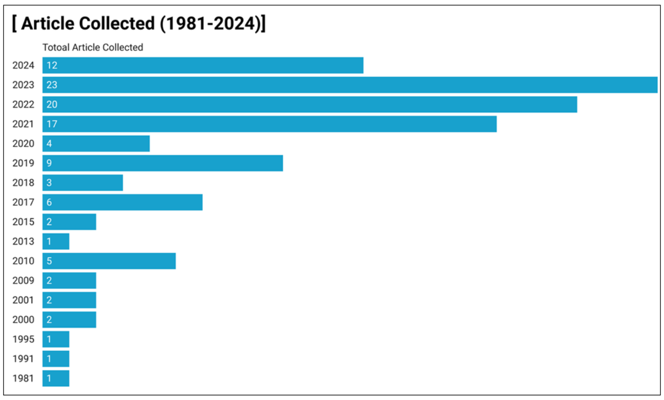

Fig. 6. Article papers collected 1981-24.

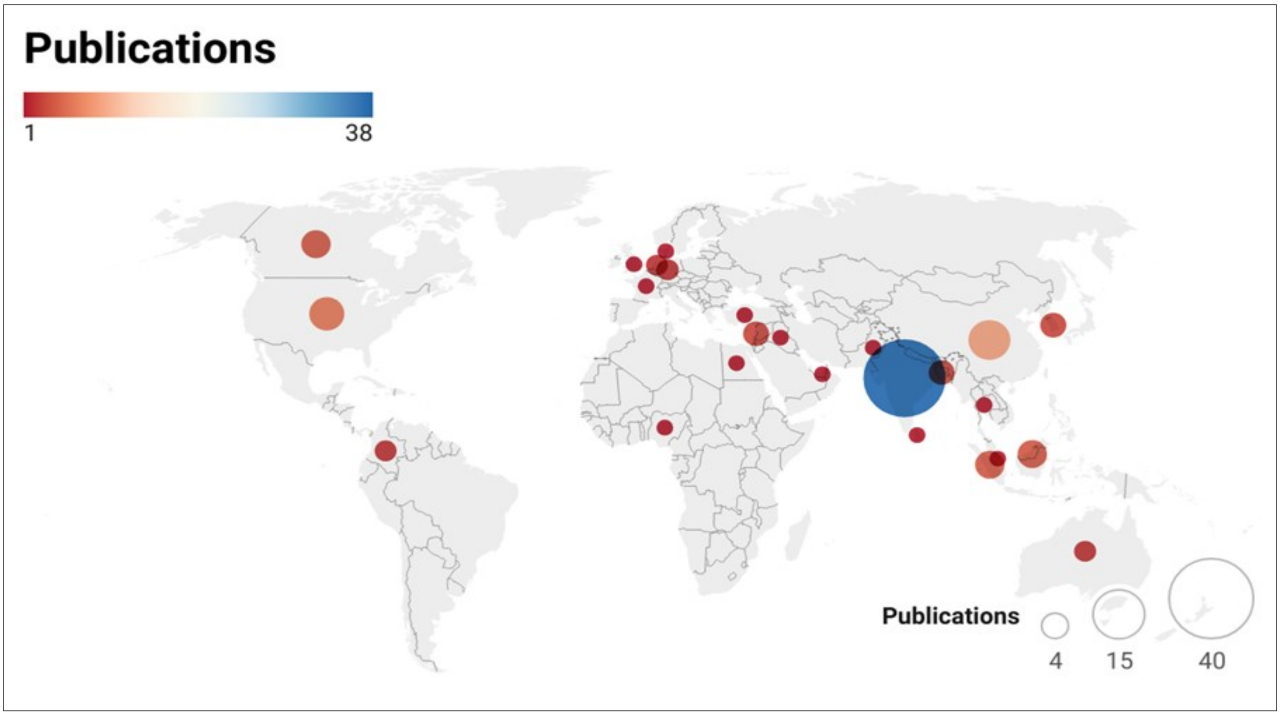

Fig. 7. Publications collected from different parts of the country.

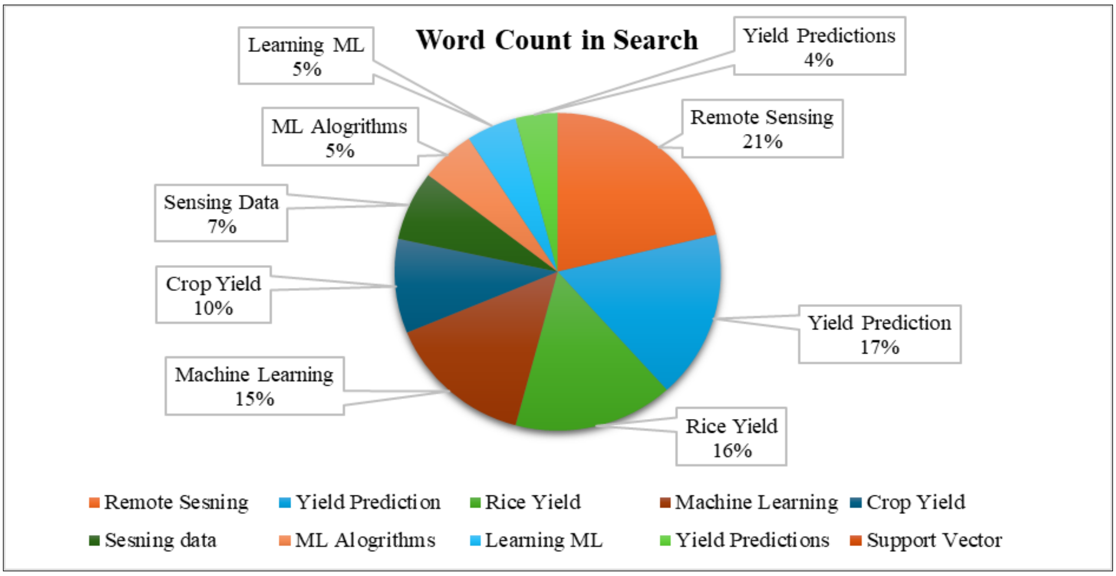

Fig. 8. Keywords or phrases or main words count observed in the references.

Literature Review

Rice yield prediction using optical remote sensing: Optical remote sensing uses visible, near-infrared (NIR) and short-wave infrared sensors to images of the earth surface features by detecting solar radiation reflected from ground targets. This technology enables the identification of vegetation features through their unique spectral signatures by using optical remote sensing to find the unique spectral signature of vegetation. This spectral signature is crucial in agriculture because it provides valuable information about crop health, soil conditions and overall farm management. These signatures, obtained through hyperspectral and multispectral remote sensing technologies, enable the extraction of detailed information about vegetation and soil properties (41). Spectral signatures are significant for precision agriculture applications, as they allow for the detection of plant diseases, insect pests and invasive species, as well as the crop yield estimation and classification of crop distributions (41).

The reflection is low in both the blue and red regions of the spectrum due to the absorption by chlorophyll for photosynthesis. In the near-infrared (NIR) region, the reflection is much higher than that in the visible band due to the cellular structure in the leaves. Hence, the vegetation can be identified by the high NIR but generally low visibility reflectance. Based on the type of sensor and data collected, optical remote sensing is classified into several types: Multispectral, Hyperspectral, Panchromatic, Thermal Infrared and Very high-resolution sensing. Each system offers unique capabilities tailored to specific applications, as shown in Table 1. Researchers have devoted considerable efforts to predicting rice yield using optical remote sensing images. Optical remote sensing data, often susceptible to cloud cover, can be limited during certain seasons, necessitating alternative data such as microwave remote sensing (42). However, when available, optical data, including various forms of the Normalized Difference Vegetation Index (NDVI), have been successfully integrated into crop simulation models like Rice-SRS, which is based on the ORYZA1 model, to provide accurate yield estimations with minimal error (42, 43).

By examining specific spectral bands of these images, pre-harvest yield estimation becomes possible due to the responsiveness of these bands to vegetation conditions. For instance, plants absorb energy in the spectral range of 0.45-0.70 μm and reflect it in the 0.70-0.90 μm range. Employing these spectral ranges, various vegetation indices have been utilized, including NDVI (Normalized Difference Vegetation Index), RVI (Ratio vegetation index), DVI (Difference vegetation index), IPVI (Infrared percentage vegetation index), SAVI (Soil-Adjusted Vegetation Index), VCI (Vegetation Condition Index), VHI (Vegetation health index), TCI (Temperature condition index) and GNDVI (Green Normalized Difference Vegetation Index), to estimate yield which is fully shown in Table 2 before harvesting. Interestingly, while optical remote sensing is a powerful tool for yield prediction, it has challenges. The presence of clouds can significantly reduce the availability of quality optical data, which is crucial for accurate yield estimation (44). However, the limitations posed by cloud cover and the complexity of remote sensing data necessitate advanced processing techniques to ensure the accuracy and reliability of yield predictions (42, 44, 45). The continued development of models and strategies to overcome these challenges is essential for enhancing the predictability of rice yields using optical remote sensing data.

Visualization parameter for optical remote sensing: Remote sensing technologies effectively capture various physical changes associated with rice growth. Paddy rice is a distinct crop variety that requires abundant water throughout its lifecycle, except for the maturation phase. The study area selected for this research was Thiruvarur district of Tamil Nadu. The major agricultural crop was rice, grown throughout the three seasons in a year. Due to the continuous cultivation of rice in this area, TCC(True Colour Composite) represents an image in the visible spectrum that closely resembles what the human eye would see in real-world life, which is created by Red (R), Green (G) and Blue (B) bands of a satellite image to their respective colour in the RGB channel (46). FCC (False Colour Composite) display an image using a non-visible portion of the electromagnetic spectrum, often replacing visible bands with infrared or near-infrared bands. This FCC consisted of three channels, namely NIR, Red and Green. NDVI (Normalized Difference Vegetation Index) consists of two leading spectral bands: red and near-infrared (NIR). NDVI was utilized to identify paddy growth patterns, as illustrated in Fig. 9. During the early growth stages, rice plants have limited canopy coverage, causing the spectral

Table 1. Types of optical remote sensing with satellite products, resolutions and applications in rice yield estimation

Type	Satellite	Spatial resolution	Spectral bands	Temporal resolution	Application in rice yield estimation	Product source
Multispectral	Landsat-8/9 (OLI/TIRS)	30 m (VIS-NIR), 15 m (PAN)	11 bands	16 days	Monitoring crop health (NDVI), evapotranspiration and land-use classification	USGS Earth Explorer
	Sentinel-2A/B	10 m, 20 m, 60 m	13 bands	5 days	Assessing crop phenology, vegetation indices (e.g., NDVI, EVI)	Copernicus Open Access Hub
	MODIS (Terra/Aqua)	250 m - 1 km	36 bands	1-2 days	Large-scale monitoring of rice-growing regions and drought analysis	NASA LP DAAC
Hyperspectral	PRISMA (Italy)	30 m	239 bands	29 days	Detecting crop stress, nitrogen content and soil conditions for rice yield optimization	ASI (Italian Space Agency)
	Hyperion (EO-1)	30 m	220 bands	16 days	Identifying water stress and estimating biochemical parameters	NASA Earth Observing System
Panchromatic	WorldView-3	0.31 m	1 band	<1 day	High-resolution monitoring of rice field boundaries and crop layout	Maxar Technologies
	GeoEye-1	0.41 m	1 band	3 days	Precision agriculture: Identifying irrigation patterns and field-level management	Maxar Technologies
Thermal Infrared	Landsat-8/9 (TIRS)	100 m (resampled to 30 m)	2 bands	16 days	Estimating evapotranspiration (ET) to predict crop water demand	USGS Earth Explorer
	MODIS (Terra/Aqua)	1 km	3 bands	1-2 days	Monitoring surface temperature and drought stress in rice fields	NASA LP DAAC
	ASTER	90 m	5 bands	On-demand	Detecting heat stress in crops and assessing water stress levels	NASA Earth Data
Very high resolution	WorldView-2/3	0.31 m (PAN), 1.24 m (MS)	8 bands	<1 day	Monitoring rice crop growth and yield prediction at field scale	Maxar technologies
	Pleiades-1A/1B	0.5 m	5 bands	2 days	Change detection and disaster impact assessment on rice fields	Airbus defence and space

Table 2. Various vegetation indices used to estimate the yield of rice

Index	Full Form	Formula	Purpose
NDVI	Normalized difference vegetation	(NIR - Red) / (NIR + Red)	Detects live green vegetation.
RVI	Ratio vegetation index	NIR / Red	Highlights vegetation by using a simple ratio between NIR and Red bands.
DVI	Difference vegetation index	NIR - Red	Measures vegetation density by directly subtracting reflectance values.
IPVI	Infrared percentage vegetation index	(NDVI + 1) / 2	Scales NDVI to a 0-1 range for easier interpretation.
SAVI	Soil-adjusted vegetation index	(NIR - Red) / (NIR + Red + L) * (1 + L)	Reduces soil brightness effects for areas with sparse vegetation.
VCI	Vegetation condition index	(NDVI - NDVI_min) / (NDVI_max - NDVI_min) * 100	Monitors drought by comparing current NDVI with historical extremes.
VHI	Vegetation health index	α * VCI + (1 - α) * TCI	Combines VCI and TCI to assess overall vegetation health.
TCI	Temperature condition index	(T_max - T) / (T_max - T_min) * 100	Identifies vegetation stress caused by temperature anomalies.
GNDVI	Green normalized difference vegetation index	(NIR - Green) / (NIR + Green)	More sensitive to chlorophyll concentration than NDVI.

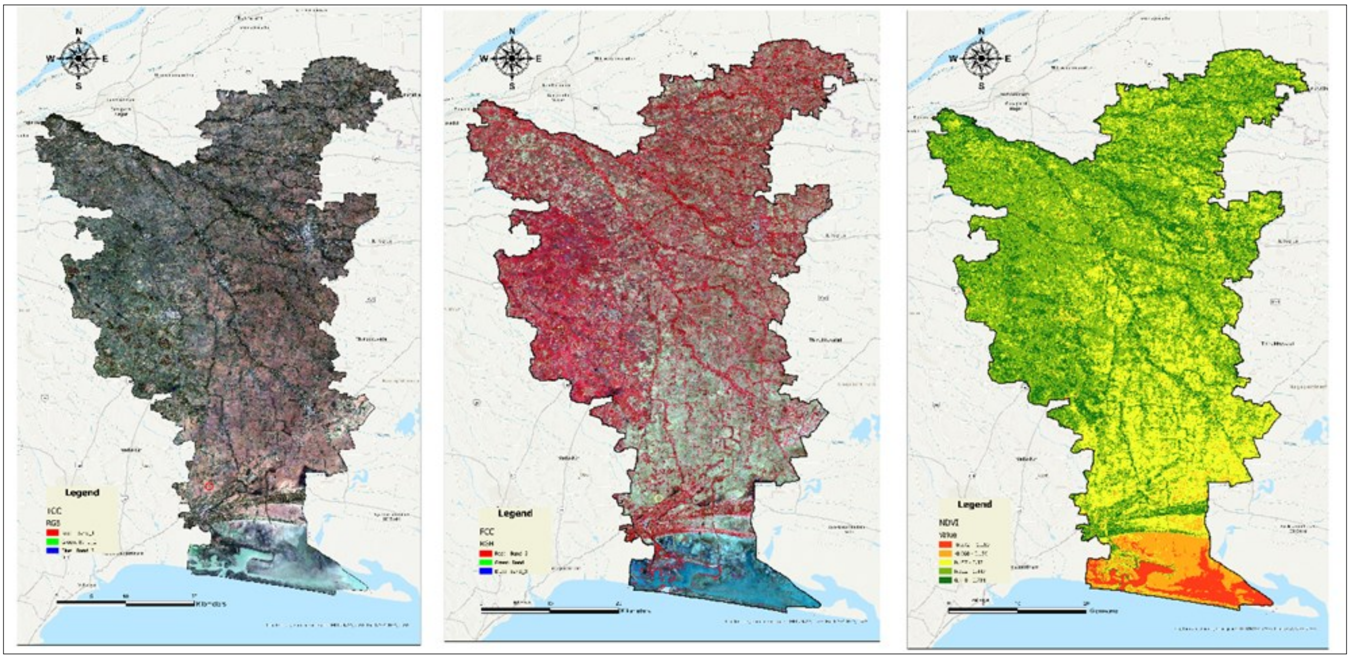

Fig. 9. Sentinel-2 True colour composite (TCC), False colour composite (FCC) and normalized difference vegetation index (NDVI) image in 2023. For Sentinel-2 imagery, different band combinations are used for various representations. The true colour composite (TCC), which mimics natural vision, uses B4 (Red), B3 (Green) and B2 (Blue). The false colour composite (FCC), helpful in highlighting vegetation, combines B8 (NIR), B4 (Red) and B3 (Green), making healthy vegetation appear bright red. The normalized difference vegetation index (NDVI), calculated using B8 (NIR) and B4 (Red), provides a measure of vegetation health with higher values showing healthier vegetation.

response to be dominated by underlying soil and water. As the rice develops, the leaves and stems increasingly influence the signal until reaching a saturation point, typically occurring during the booting phase when photosynthetic activity peaks. The optimal periods for assessing rice yield through remote sensing are during the emergence of the yellowing panicle and when leaves begin to senesce and head (47). Throughout their growth cycle, rice plants and their surroundings exhibit varying spectral responses, providing valuable opportunities for precise area mapping and crop yield prediction using multitemporal remote sensing techniques. When employing single-date imagery, the optimal time for data acquisition is during the booting stage. For multi-date sensing approaches, the most accurate estimations can be achieved by capturing images during the early vegetative and reproductive stages. Rice plants and their environment display distinct spectral signatures throughout their developmental stages, offering significant potential for accurate area mapping and yield forecasting using multi-temporal remote sensing methods.

Rice yield prediction using microwave remote sensing **:** Microwave remote sensing (MRS) presents several advantages. One is its ability to capture images even in adverse weather conditions, such as heavy rain, snow, cloud cover and strong solar irradiance. This makes the imagery obtained from microwave sensors an excellent source for mapping rice regions, as rice cultivation frequently occurs during the rainy season, characterized by persistent cloud cover. Although the sensor has a solid capacity to capture images without sunlight in almost all-weather conditions, high revisit capabilities were not achieved until the launch of the Copernicus Sentinel-1A-satellites in 2014. Carrying C-band SAR sensors that operate at a high-frequency equivalent to a wavelength of 5.5 cm for observation, these satellites have enabled monitoring crops with shorter growth durations. Accordingly, many agricultural studies have employed Sentinel-1A data for agricultural land mapping, yield estimation and crop growth monitoring (48, 49). The Sentinel-1A satellite operates in the interferometric wide (IW) mode for land observation, utilizing VV and VH dual polarizations. In this context, V and H represent the vertical and horizontal polarizations, respectively, while the first and second characters pertain to the transmitted and received polarizations. Interestingly, while MRS techniques are adequate, the synergistic use of microwave and optical remote sensing data has yielded promising results for crop parameter assessment and condition monitoring at a regional level (50). MRS is a robust method for rice yield prediction, offering reliable estimates even under adverse weather conditions. The integration of MRS with crop simulation models enhances the accuracy of yield predictions (43). While the combined use of optical and microwave data can be beneficial, the extent of improvement varies depending on the specific application and data availability Table 3. Interestingly, active microwave sensors like Synthetic Aperture Radar (SAR) can provide higher spatial resolution (e.g., 10 m for Sentinel-1A) but may have lower temporal resolution (51). This creates an opportunity for synergistic use of passive and active microwave and optical data to enhance spatial and temporal resolutions. Overall, the advancements in MRS technology and its application in agriculture promise to improve food security through accurate and timely yield forecasts (52).

Rice yield prediction using machine learning algorithms : Machine learning (ML) algorithms have been increasingly applied to predict rice yields. The literature indicates that various ML techniques and features predict agricultural yields, with temperature, rainfall and soil types being common predictors (53, 54). In particular, Artificial neural networks (ANN), Support vector machines (SVM), linear regression and Long-short term memory (LSTM) networks are commonly utilized (53, 55). However, challenges in ML rice yield prediction include selecting appropriate input variables, handling missing data and capturing non-linear relationships between variables (53). Interestingly, while some studies have achieved high accuracy

Table 3. Microwave remote sensing satellite datasets for rice yield predict

Satellite / Sensor	Frequency Band	Spatial Resolution	Temporal Resolution	Application	Agency
Sentinel-1A (SAR)	C-band (5.405 GHz)	10 m	6-12 days	Crop classification, growth monitoring, yield prediction	ESA (European Space Agency)
RADARSAT-2	C-band (5.405 GHz)	8-100 m	24 days	Rice area mapping, soil moisture, crop yield estimation	Canadian Space Agency
RISAT-1 (SAR)	C-band (5.35 GHz)	3-50 m	25 days	Crop discrimination, soil moisture retrieval	ISRO (Indian Space Research Organization)
ALOS PALSAR-2	L-band (1.27 GHz)	10-100 m	14 days	Rice growth stage monitoring, biomass estimation	JAXA (Japan Aerospace Exploration Agency)
SMAP (Soil Moisture Active Passive)	L-band (1.41 GHz)	36 km	2-3 days	Soil moisture retrieval, yield estimation in large areas	NASA (National Aeronautics and Space Administration)
TerraSAR-X	X-band (9.65 GHz)	1-40 m	11 days	High-resolution crop monitoring, yield prediction	DLR (German Aerospace Centre)
COSMO-SkyMed	X-band (9.6 GHz)	1-100 m	On-demand	Crop classification, yield estimation	ASI (Italian Space Agency)

rates, such as 96.72 % using a classification-based interactive model and 99.82 % with a novel decision support system, others have reported lower accuracy, like 72 % with the Decision Tree Regressor (56, 57, 58). These discrepancies may arise from differences in the datasets, regional focus, or specific ML algorithms applied.

Additionally, the lack of transferability of some models across different crops and locations is noted, highlighting the need for modular and reusable workflows (59). The choice of algorithm, data quality and regional specificity can influence the effectiveness of these algorithms. Further research is warranted to refine these models for broader application and to address challenges such as model interpretability and scalability (60). Techniques that include feature selection, regularization and data preprocessing are recommended to address these challenges. Furthermore, research indicates that prediction accuracy primarily depends on the features considered.

Rice yield prediction using common machine learning algorithms: Machine learning (ML) algorithms are increasingly employed in agricultural contexts to predict crop yields, including rice. The literature shows that Random Forest (RF) is a frequently utilized algorithm for this purpose (54, 61), as shown in Table 4. RF is known for its ability to handle classification and regression tasks, making it suitable for crop yield prediction (CYP) (62). Support Vector Machine (SVM) is also mentioned as a significant tool in predicting rice yield, often in combination with RF for enhanced accuracy (63). Other algorithms such as Decision Trees, Naive Bayes, k-nearest Neighbour (k-NN) and various deep learning techniques like Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) are also applied in the field of agricultural yield prediction (55, 63). Interestingly, while RF is commonly cited, there are instances where other algorithms are more effective in specific contexts. For example, a study in Nigeria identified the Decision Tree Regressor as the most accurate model for predicting crop yields within their dataset (58).

***Integration of Remote Sensing and Machine Learning for Rice Yield Prediction*:** Integrating remote sensing (RS) and machine learning (ML) for rice yield prediction is a multidisciplinary approach that leverages both technologies' strengths to enhance the accuracy of agricultural forecasts, represented in Table 3. For example, a Nepal study developed a rapid rice yield estimation workflow by combining remote sensing-derived NDVI with meteorological variables like rainfall, soil moisture and evapotranspiration (64). This approach, using stacked tree-based regression models, achieved 92 % accuracy in yield estimation.

Low-altitude remote sensing imaging for rice yield estimation : Low-altitude remote sensing imaging has emerged as a promising technique for estimating rice yield with high accuracy and spatial resolution. Multiple studies have demonstrated the effectiveness of this approach using various platforms and sensors. Studies have shown the efficacy of low-altitude remote sensing systems in monitoring rice crop conditions and predicting yield in Table 5. One investigation employing a spectroradiometer-based approach achieved a yield estimation error of 12.78 kg/10 acres and a protein content estimation error of 0.149 % (65). In a separate study, researchers successfully quantified nitrogen levels in a rice canopy using an IMEC hyperspectral snapshot camera with 25 bands (600-1000 nm) mounted on a low-altitude platform (66). Interestingly, different studies have explored various spectral indices and modelling techniques for yield estimation. While some research focused on the importance of ultraviolet and short-wave infrared regions (67), others found success using normalized difference vegetation index (NDVI) derived from multispectral imagery (68). The choice of platform also varied, with studies employing unmanned aerial vehicles (UAVs) (69), radio-controlled unmanned helicopters (69) and other low-altitude systems. Low-altitude remote sensing imaging has proven to be a valuable tool for rice yield estimation, offering high spatial and temporal resolution. Integrating spectral information, vegetation indices and crop growth models has shown promising results in accurately predicting rice yield across different environmental conditions and management practices. This technology provides a rapid, non-destructive approach for site-specific rice nutrient management and yield forecasting.

High-altitude remote sensing imaging for rice yield estimation : Remote sensing techniques have significant potential for estimating rice yields across large areas. Multiple studies have demonstrated the effectiveness of various remote sensing platforms and methods. High-altitude remote sensing, particularly satellite-based imaging, offers regional and national-scale crop monitoring advantages. Satellite data can be

Table 4. Different machine learning algorithms and remote sensing data are used for rice yield prediction

Source	Study Area	Dataset used	Methods used	Results
(95)	Location of Qian Gorlos in Jilin province of China	Sentinel -2	ML (Regression Models) Random Forest, Support vector machine.	The model achieved a higher accuracy with an R2 of 0.87, RMSE of 0.33 and REP of -1.1, significantly outperforming the model that considered the study area as a whole.
(96)	Murray and Murrumbidgee Valleys in New South Wales, Australia	Sentinel-1 and Sentinel-2	Phenological and time-series data were derived from remote sensing and weather datasets to analyze drivers of rice yield variability and develop yield forecast models	with models achieving an RMSE of 1.6 t/ha and Lin's concordance correlation coefficient of 0.67 30 days after flowering at the field level
(97)	Jiangxi is positioned in the southern part of the middle segment of the Yangtze River in south China.	Moderate Resolution Imaging Spectroradiometer (MODIS) has a resolution of 250m and a 16-day interval. Landsat 5 TM and Landsat 8 OLI	Standardized Precipitation Index (SPI) decision-tree algorithm was utilized for rice plantation mapping and yield estimation was conducted using MODIS data from 2000 to 2020.	remote sensing-based rice yield model using the NDVI (normalized difference vegetation index) to estimate rice production, with a specific model formula YD = 4.899 × 10^(-6) × NDVI^2 + 2.891 × NDVI + 98511.218
(98)	Tongxiang County in Zhejiang Province, southeast China	Sentinel -1A	rice green leaf area index (LAI) estimation using four machine learning regression models (SVM, k-NN, RF, GBDT)	The most accurate rice green leaf area index (LAI) estimates with the Gradient Boosting Decision Tree (GBDT) model, achieving an R^2 of 0.82 and RMSE of 0.68 m^2/m^2. The growing season, recording an R^2 of 0.68 and RMSE of 0.98 m^2/m^2 with the k-Nearest Neighbour (k-NN) model
(99)	the study region is situated in western Taiwan.	Sentinel-2 A, B	Three machine learning models, namely random forest (RF), support vector machine (SVM) and artificial neural networks (ANN), were employed to predict rice crop yields.	SVM demonstrated superior performance compared to RF and ANN in predicting rice yields, with RMSPE and MAPE values below 5.5 % and 4.7 % for the 2019 second crop and 2020 first crop, respectively. Conversely, RF and ANN exhibited higher RMSPE and MAPE values for both crops, with 9.4 % and 7.1 % for the 2019 first crop and 11.0 % and 9.4 % for the 2020 second crop, respectively.
(100)	Situated in the central region of Rio Grande do Sul, Brazil,	Remote Piloted Aircraft Systems (RPAS) equipped with a Sequoia® camera on a Phantom 4® Pro platform to acquire multispectral images for monitoring agronomic parameters	The Multi-Layer Perceptron (MLP) algorithm devised predictive models for the N area and grain yield. Performance assessment was conducted through training and testing phases.	MLP demonstrated superior performance in predicting N-area with a strong correlation between predicted and observed values (0.82 and 0.71) and a lower mean absolute error (MAE) of 9.47. It also performed exceptionally well in grain yield models across all stages.
(87)	The six regions of Bangladesh are Sunamganj, Maulvibazar, Sylhet, Habiganj, Kishoreganj and Netrokona.	Sentinel-2	Parametric (simple and multiple) and nonparametric (artificial neural network, ANN) regression analyses were employed to develop and validate the crop yield prediction models. Identify strong correlations, such as NDVI-RGVI, NDWI-MSI and RGVI-LAI.	The artificial neural network (ANN) models, particularly those using NDVI, exhibited higher accuracy for boro rice yield predictions, with R^2 values of 0.84 and 0.91 for simple and multiple regression approaches, respectively. The parametric and nonparametric regression analysis results revealed considerable agreement with the ground reference yield data, as evidenced by the R^2 values, which ranged from 0.44 to 0.91 for various vegetation indices and approaches.
(101)	In the Jiangsu Province, China is positioned within the Middle and Lower Reaches of the Yangtze River.	Sentinel-1, Sentinel-2	meta-learning ensemble regression (MLER) framework for accurate rice yield prediction at field and county levels. (RF, XGBoost, SVR) and Long Short-Term Memory (LSTM) to predict rice yield.	The MLER algorithm accurately predicted rice yield for Xinghua and Suining, outperforming individual models (RF, XGBoost, SVR) and LSTM with an R^2 of 0.89 and RMSE of 0.54 t/ha. The algorithm's performance was consistent across Jiangsu Province, with leave-one-year-out assessments showing R^2 ranging from 0.42 to 0.67 and RMSE from 0.23 to 1.22 t/ha.
(102)	Tiruchirapalli, Thanjavur, Tiruvarur, Nagapattinam, Ariyalur and Perambalur, Ramanathapuram, Sivaganga and Pudukottai districts	Sentinel-1A SAR.	Integrated remote sensing products with the ORYZA2000 crop growth model to estimate rice yield/ Estimated LAI	Mapped over 1.07 million hectares of rice fields with high accuracy, ranging from 90.3 % to 94.2 % and Kappa values between 0.81 and 0.88. Achieved yield simulation accuracies of 86-91 % at the district level and 82-97 % at the block level using SAR products and the ORYZA crop growth model.

(103)	The states are Andhra Pradesh, Assam, Bihar, Chhattisgarh, Karnataka, Haryana, Jharkhand, Madhya Pradesh, Odisha, Punjab, Telangana, Tamil Nadu, Uttar Pradesh and West Bengal.	Radarsat-2, RISAT-1, Sentinel-1	District-level rice crop yield was estimated using three different procedures: i) Agro-meteorological regression models, ii) Crop Simulation Models and iii) Remote sensing index (VCI) based empirical models.	The Root Mean Square Error (RMSE) percentage ranged from 2.3 to 4.3 for area estimation, 0.84 to 1.35 for production estimation and 0.24 to 0.27 for yield estimation. The coefficient of determination (R^2) ranged from 0.62 to 0.92 for area, 0.75 to 0.91 for production and 0.5 to 0.83 for yield estimation.
(104)	The rice-growing areas, 18 states in India	MODIS Leaf Area Index (LAI)	Gradient Boosted Regression (GBR) models	observation 2003 to 2015 (r=0.85 & MAE=0.15t/ ha), correlations up to r=0.93, out of sample validation for the year 2016 and 2017showed result with r=0.86 and r=0.77.
(105)	The study location is Katsina State in northern Nigeria	No satellite data	Logistic Regression, Artificial Neural Network, Random Forest, Random Trees and Naïve Bayes algorithms	Random Forest (RF) and Random Trees (RT) demonstrated superior performance in yield prediction, with an actual positive rate of 1. At the same time, Naive Bayes (NB) had a rate of 0.19 and Neural Network (ANN) and Logistic Regression (LR) achieved a rate of 0.75. RF, RT and NB achieved an ROC Area value of 1, indicating a higher capability to predict rice yield accurately than ANN and LR. NB had the highest prediction accuracy of 91.7, while RF and RT demonstrated a perfect accuracy of 100 in predicting rice yield.
(106)	study area is the Sahibganj district of Jharkhand state (India)	Sentinel-2B multi-spectral sensor (MSI) data and 2 Sentinel-1A (SAR) data	The Random Forest classifier	The estimated paddy acreage was 68.3 to 77.8 thousand hectares and the predicted yield was 1.60 t/ha. (Simple Linear regression and Random Forest)
(107)	Chhattisgarh, three central districts representing three agroclimatic zones	No satellite data used	Stepwise multiple linear regression (SMLR), artificial neural network (ANN), least absolute shrinkage and selection operator (LASSO), elastic net (ELNET) and ridge regression	The ANN model exhibited excellent performance in Raipur and Surguja districts, with high R^2 calibration (1) and validation (1) scores, as well as low RMSE values (0.002 and 0.003, respectively). In contrast, the ELNET and LASSO models performed better in the Bastar district, with R^2 calibration (90 and 93) and validation (0.48 and 0.568) scores. The ensemble models demonstrated improved performance compared to individual models, with random forest (RF) outperforming in the Bastar district, achieving R^2 values of 0.85 and 0.81 for calibration and validation, respectively.
(108)	China	No satellite data used	Support Vector Regression (SVR) models were developed for different rice growth stages and combinations of these stages to predict optimal yields.	The yield prediction models use RMSE, which measures the differences between predicted and observed values. The RMSE values for various stages of rice production varied, with values such as 126.8 kgh m^2, 96.4 kgh m^2 and 109.4 kgh m^2 reported for the tillering, heading and milk stages of middle-season rice, respectively.
(97)	Maha Sarakham, north-eastern Thailand,	No satellite	XGBoost, an ensemble machine-learning algorithm	The study yields promising results with low root mean square error (RMS) values for rice and wheat prediction; the root mean square error (RMS) values for rice and wheat prediction using xgboost are as follows: Rice 0.02538, Wheat 0.02198. Random Forest, another ensemble learning technique, also performs well in predicting crop yield with low RMS values: Rice 0.01942, Wheat 0.01680.
(109)	Study area of Kerala	No satellite data used	algorithms like Decision Tree, Random Forest, Linear Regression, K-nearest neighbour (KNN), xgboost and Support Vector Regression.	The accuracy of the KNN regression model for the rice dataset was reported to be 98.77 %, outperforming other regression models used in the study and results highlight the potential of KNN regression as a reliable method for yield prediction in agriculture.
(110)	Study area Larkana district in Sindh province, Pakistan,	Landsat 7 ETM	relationship between reported rice crop yield and NDVI/RVI values at the peak	indicating a high potential for estimating rice yield using Landsat ETM+ data = positive and strong relationship ((R^2 = 0.875) and root mean square error (RMSE) of 80.726.
(111)	Estimate rice yield in Nepal	MODIS	Popular tree-based regressor models like XGBoost were applied and a customized stack-ensemble model was proposed, combining XGBoost, LightGBM, Gradient Boost and random forest models.	The initial benchmark linear regression model using only NDVI had an average RMSE of 685.17 and MAE of 633.83 for all districts. Other ML models performed slightly better. The proposed stack-ensemble model incorporating NDVI and five auxiliary variables reduced the average RMSE to 328.06 and MAE to 317.21, achieving an average 92 % accuracy in yield estimation.

Table 4. Different machine learning algorithms and remote sensing data are used for rice yield prediction

Crop Varieties	Literature	Year	Task	Network Framework and Algorithms	Result
Rice	(90)	2023	Predict the yield of rice	CNN	RMSPE: 14 %
Rice	(91)	2022	Predict the yield of rice	3D-CNN, 2D-CNN	RMSE: 8.8 %
Rice crop-yield calculation based on High -altitude remote sensing					
Rice	(92)	2022	Predict the yield of rice	SVM, RF, ANN	MAPE: 3.5 %
Rice deep learning in crop-image yield calculation					
Rice	(93)	2021	Calculate Rice Seed Setting Rate (RSSR)	YOLOv4	MAPE: 99.43 %
Rice	(94)	2024	Predict the yield of rice	CNN-LSTM-Attention	R^2 of 0.76, RMSE of 519.07 kg/ha and MAPE of 4.67 %,

integrated with crop growth models to estimate rice yields spatially. For example, a study in the Cauvery Delta of India used Sentinel-1A SAR data with three yield estimation methods - spectral indices regression, a semi-physical approach and integration with the DSSAT crop model (70). The spectral indices regression and DSSAT integration approaches performed well, with $R^2 > 0.80$ and NRMSE < 10 % compared to ground measurements. While high-altitude remote sensing is valuable, some studies have found low-altitude platforms like unmanned aerial vehicles (UAVs) to provide higher spatial and temporal resolution data. In one study, UAV-based hyperspectral imaging at critical growth stages allowed for pixel-scale yield estimation with R2 of 0.74 and RMSE of 248.97 kg/ha (71). This highlights the complementary roles that different remote sensing platforms can play. The high-altitude remote sensing, especially with crop models, shows promise for large-scale rice yield estimation. However, integrating data from multiple platforms and resolutions may provide the most comprehensive yield predictions. Future work should focus on refining models, incorporating dynamic environmental factors and leveraging machine learning approaches to improve accuracy.

Rice deep learning in crop-image yield calculation: Deep learning approaches have shown significant promise in estimating rice crop yields using image data. Convolutional neural networks (CNNs) applied to RGB images of rice canopies at harvest have demonstrated the ability to predict 68-70 % yield variation with relative root mean square errors around 0.22 (72). This low-cost, rapid approach can provide valuable insights for assessing productivity interventions and identifying improvement areas. The deep learning models show robustness across different imaging conditions. Models maintain predictive power even with images taken at angles up to 30° from vertical, in diverse lighting and at reduced resolutions down to 3.2 cm/pixel (48). This suggests the potential for scaling the approach using unmanned aerial vehicles.

Additionally, some studies have found that images taken during the ripening stage, weeks before harvest, can forecast final yields. Deep learning applied to crop imagery offers a promising rice yield estimation and forecasting tool. Various architectures have been explored, including hybrid models combining CNNs with recurrent networks like LSTM, as shown in Table 4. While performance varies, these approaches outperform traditional vegetation index-based methods, especially at later growth stages. The ability to provide rapid, low-cost yield estimates at field and pixel scales could significantly benefit crop management, food security planning and agricultural policy decisions.

Data Preprocessing and Feature Extraction

Data cleaning and normalization: Data preprocessing and feature extraction are critical steps in developing predictive models for rice yield prediction. Data cleaning and normalization ensured that the dataset was free from errors and inconsistencies and that the data were scaled to a uniform range for better comparability and performance of machine learning algorithms (73, 74). Normalization techniques such as min-max scaling and Z-score significantly impact model performance and feature importance, influencing both predictive accuracy and feature selection. The choice of normalization method can affect model performance to varying degrees depending on the dataset and algorithm used. For instance, in radiomics, the z-Score method generally performed best, with an average gain of +0.012 in AUC (Area Under the Curve) compared to no normalization and up to +0.051 on some datasets (75). While (Z-score) normalization techniques generally help reduce bias towards features with larger magnitudes, their impact on model performance can vary significantly. Experimenting with different normalization methods and evaluating their effects on model performance and feature selection for each problem and dataset is advisable.

Feature selection and engineering: Feature selection and engineering involve identifying the most relevant variables that contribute to the rice yield prediction and may include techniques such as the artificial bee colony (ABC) algorithm for feature selection (76).

Handling missing data: Managing missing data is a common challenge in data preprocessing and various imputation methods are employed to address this issue. For instance, Bayesian Principal Component Analysis (BPCA) has been identified as an effective method for imputation in the context of the VASA dataset (77). These methods could be adapted for rice yield prediction, considering the importance of accurate data imputation in predictive modelling. The data preprocessing, including data cleaning, normalization, feature selection and managing missing data, is essential for building robust predictive models for rice yield prediction.

Factors Influencing Prediction Accuracy

Quality and quantity of remote sensing data : The quality and quantity of remote sensing data, particularly spatial and temporal resolution, are critical factors in agricultural applications such as rice yield prediction. The high spatial resolution allows for detailed crop conditions at a fine scale, while high temporal resolution ensures frequent monitoring to capture changes over time (78). The importance of multispectral and hyperspectral data lies in their ability to provide a wealth of spectral information, which is essential for accurately estimating crop yields. With their high spectral resolution, hyperspectral sensors can detect subtle differences in crop vigour and stress, indicating yield potential (79). However, there are challenges associated with remote sensing data. For instance, cloud cover and low temporal resolution can limit the availability of satellite images, affecting the estimation process (78). Cloud cover poses a significant challenge in remote sensing, particularly for urban land cover (ULC) monitoring and flood mapping. Researchers have developed methods combining optical and synthetic aperture radar (SAR) data to address this issue. One innovative approach is the weighted cloud dictionary learning method (WCDL) for fusing optical and SAR data in cloud-prone areas. This method incorporates a cloud probability weighting model and pixel-wise cloud dictionary learning to mitigate cloud interference. Experiments show that the WCDL method improves overall accuracy by more than 6 % compared to single SAR data and 20 % compared to optical data alone (80). Another technique, the global-local fusion-based cloud removal (GLF-CR) algorithm, leverages SAR information to guide the relationship among optical windows and transfer complementary information to generate reliable texture details in cloudy areas (81). Studies have developed optimization models and intelligent systems incorporating additional data, such as energy balance equations, to address these issues and enhance yield predictions (78). Furthermore, integrating spatial and spectral resolutions has improved crop mapping accuracy in heterogeneous areas (82).

Model Training and Validation

Importance of cross-validation and testing: Cross-validation and testing are critical components in developing predictive models for rice yield. These processes are essential for assessing the generalizability and robustness of the models to unseen data, thereby ensuring the reliability of yield predictions (83). While cross-validation is a standard practice in model evaluation, its impact on interpreting machine learning models, particularly in the context of temporal agricultural data, has been highlighted. The choice of cross-validation strategy can significantly affect the interpretation of the model and its performance on held-out data, emphasizing the need for domain-specific best practices in the application of cross-validation (83). Moreover, the studies reveal that while deep learning models, such as LSTM and GRU (Gated Recurrent Unit), show promise in handling complex spatiotemporal data, their performance does not necessarily improve with increased model complexity and simpler models can be equally effective for small-sample data. This suggests that the benefits of complex models must be weighed against their computational costs and convergence rates. Cross-validation and testing are indispensable for validating rice yield prediction models. The findings from the literature underscore the importance of selecting appropriate cross-validation strategies to ensure accurate model interpretation and performance (83). Additionally, the complexity of the model should be carefully considered, as simpler models may offer sufficient predictive power with lower computational demands (84). These insights are crucial for advancing the field of rice yield prediction and supporting decision-making in agriculture.

Techniques for model tuning and optimization: The literature offers various techniques for optimizing and tuning models to predict rice yield. Hybrid deep learning models employ shared layers of classification and regression models along with statistical analyses like PCC, SHAP and RFECV for feature selection (85). This approach yields better results than other deep learning models, with an RMSE of 344.56 and an R-squared of 0.64. However, some studies suggest that regression-based models can outperform ANN models in certain situations. For instance, the importance of meteorological factors like Bright sunshine hours showcases the high accuracy of stepwise multiple linear regression models with an R-squared up to 0.95. (86) propose the SCA-WRELM method, which incorporates min-max data normalization and optimal parameter tuning via the Sine Cosine Algorithm for improved predictive results.

The literature suggests that deep learning and traditional statistical models have their merits in rice yield prediction, with the choice of model depending on specific conditions and data characteristics. Hybrid models and ensemble methods are particularly effective, with feature selection and hyperparameter tuning playing crucial roles in optimizing model performance. Integrating machine learning techniques with meteorological and satellite data is also highlighted as a promising direction for enhancing the accuracy of yield predictions.

Practical Implementation Issues

Scalability and computational requirements: The practical implementation of ML models for rice yield prediction using remote sensing data presents several challenges, particularly regarding scalability and computational requirements. Scalability is a critical issue, as the models must manage large volumes of data and potentially be applied across diverse geographical regions with varying environmental conditions. Computational requirements are also significant, as advanced ML algorithms and high-dimensional remote sensing data processing demand substantial computational resources.

Challenges and future directions: We face several constraints when discussing RS and ML techniques for predicting rice yields. One of the main challenges is the quality and processing of data required for accurate predictions. Integrating remote sensing and meteorological data with ML models requires careful attention to data quality issues, selecting suitable ML models and understanding the complex, non-linear relationships between historical crop yield and various factors (87). Furthermore, it was found that relying solely on remote sensing-derived Normalized Difference Vegetation Index (NDVI) is insufficient for accurate yield estimation and incorporating meteorological variables such as rainfall, soil moisture and evapotranspiration is essential (87). Another challenge is the high dimensionality of remote sensing data, which makes training models infeasible using raw pixels. Methods such as feature selection and dimensionality reduction, such as principal component analysis (PCA), address this issue (45). However, these techniques may discard potentially informative spectral bands. In addition, the spatial and temporal variability of subregional environmental and climatic factors can affect the models' output, requiring tailored approaches for different regions (88). While remote sensing and ML offer promising approaches for predicting rice yields, they are constrained by data quality and processing challenges, suitable ML models and the requirement to account for spatial and temporal variability.

Integration into future agricultural systems is another complex aspect. While ML models offer improved accuracy in yield prediction, their adoption requires compatibility with current farming practices and decision-making processes (57, 87). Moreover, the integration process must consider the distinct physio-geographical settings of different districts, which can affect the accuracy of yield estimations (87). IoT and Big Data technologies can facilitate this integration by enabling the efficient collection and processing of data, but they also introduce additional layers of complexity (89). The practical implementation of ML models for rice yield prediction using remote sensing data involves addressing scalability and computational challenges and ensuring seamless integration into existing agricultural systems. The success of such models depends on their ability to operate efficiently across various scales and to be incorporated into the agricultural workflow without disrupting established practice. Addressing these constraints is essential for improving the accuracy of yield predictions and ensuring their practical applicability for food security and agricultural decision-making.

Conclusion

This paper provides a comprehensive review of the application of RS and ML techniques to predict rice yields, focusing on addressing food security challenges. RS offers insights on crop health through optical and microwave imagery, while ML uses this data to make more accurate predictions. Optical RS effectively predicts yields before harvest using vegetation indices sensitive to plant health. At the same time, microwave RS is applicable in all weather conditions and it is particularly valuable during the rainy seasons, which are common in rice cultivation. Various ML algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Random Forest (RF), are used to predict rice yields, processing high-dimensional data to identify key patterns for accurate forecasts. The integration of RS and ML improves predictive accuracy by utilizing high-resolution spatial and temporal data. Methods such as feature selection and data normalization enhance model performance and hybrid models show promise in optimizing predictions.

Data quality and processing challenges, scalability, computational requirements and environmental adaptability must be addressed to successfully implement precision agriculture and decision-support technologies. These technologies offer numerous benefits for precision agriculture, including improved crop yield prediction, optimized resource management and enhanced environmental sustainability. Processing large amounts of data from various sensors and platforms allows for more accurate decision-making and targeted interventions in farming practices. This can lead to increased productivity, reduced environmental impact and improved food security, aligning with broader policy goals for sustainable agriculture and food production. Future research should focus on improving data integration by incorporating meteorological data and using advanced data assimilation techniques to enhance predictions. Continued development of hybrid models that balance accuracy with computational efficiency is necessary. It is crucial to ensure that predictive models can be scaled across diverse regions and integrated seamlessly into existing agricultural systems for widespread adoption. Integrating IoT and Big Data technologies can facilitate efficient data collection and processing, which is essential for enhancing these models' real-time applicability.

In conclusion, the combination of remote sensing and machine learning presents a potent means of advancing rice yield predictions, thereby contributing to improved food security. These methodologies offer trustworthy, expandable and applicable intelligence for agricultural decision-making by tackling current challenges and capitalizing on technological breakthroughs. Future research and development in this area are essential to attaining the Sustainable Development Goals, which aim to eradicate hunger and foster sustainable agriculture by the year 2030.

Acknowledgements

The authors extend their gratitude to the Department of Remote Sensing and GIS and acknowledge the support provided by the Centre for Water and Geospatial Studies, Tamil Nadu Agricultural University, in preparing and successfully submitting this manuscript. Furthermore, the authors express appreciation to their colleagues and the reviewers who offered constructive feedback and recommendations to enhance the quality of the manuscript.

Authors' contributions

MA conducted the literature review, structured the manuscript, and prepared the initial draft. KPR developed the framework and revised the manuscript. SP, DM, APS and GV contributed to the revision and refinement of the manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Chakraborty N. Rice Proteomics and Beyond. Rice Ress. 2015;03 (02).https://doi.org/10.4172/2375–4338.1000e113
2. FAO, IFAD, UNICEF, WFP and WHO. The State of Food Security and Nutrition in the World [internet].Rome: FAO; 2021[cited 2024 Sept 18]. Available from: https://openknowledge.fao.org
3. FAO. World Food and Agriculture – Statistical Yearbook 2022 [internet]. Rome: FAO; 2022[cited 2024 Sept 18]. Available from: https://openknowledge.fao.org
4. Furey S. Food poverty: zero hunger and the right to food. In: Leal

Filho W, Azul AM, Brandli L, Ozuyar PG, Wall T, editors. Zero hunger. Cham: Springer International Publishing; 2020. p. 329-38. https://doi.org/10.1007/978-3-319-95675-6_115

5. IndiaStat. Agriculture rice statistics and growth figures year-wise of india [internet]. India: Indiastat Data Pvt Ltd.; 2024 [cited 2024 18 Sept]. https://www.indiastat.com/data/agriculture/rice.
6. Hong D, He W, Yokoya N, Yao J, Gao L, Zhang L, et al. Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing. IEEE Geosci Remote Sens Mag. 2021;9(2): 52-87. https://doi.org/10.1109/mgrs.2021.3064051
7. Moraga J, Duzgun HS, Cavur M, Soydan H. The geothermal artificial intelligence for geothermal exploration. Renewable Energy. 2022;192:134-49. https://doi.org/10.1016/j.renene.2022.04.113
8. Bala SK, Islam AS. Correlation between potato yield and MODIS derived vegetation indices. Int J Remote Sens. 2009;30(10):2491-2507. https://doi.org/10.1080/01431160802552744
9. Li T, Angeles O, Marcaida M, Manalo E, Manalili MP, Radanielson A, et al. From ORYZA2000 to ORYZA (v3): An improved simulation model for rice in drought and nitrogen-deficient environments. Agricultural and Forest Meteoro. 2017;237-238:246-56. https://doi.org/10.1016/j.agrformet.2017.02.025
10. Mishra N, Mishra S, Tripathy HK. Rice yield estimation using deep learning. In: Panda M, Dehuri S, Patra MR, Behera PK, Tsihrintzis GA, Cho SB, et al. editors. Innovations in intelligent computing and communication. Cham: Springer International Publishing; 2022. p. https://doi.org/10.1007/978-3-031-23233-6_28
11. Counce PA, Keisling TC, Mitchell AJ. A uniform, objective and adaptive system for expressing rice development. Crop Sci. 2000;40 (2):436-43. https://doi.org/10.2135/cropsci2000.402436x.
12. Vergara BS. Rice plant growth and development. In: Luh BS, editor. Rice. Boston, MA: Springer US; 1991. p. 13-22. https://doi.org/10.1007/978-1-4899-3754-4_2
13. Hussain S, Zhang J Hua, Zhong C, Zhu L Feng, Cao X Chuang, Yu S Miao, et al. Effects of salt stress on rice growth, development characteristics and the regulating ways: A review. Journal of Integrative Agriculture. 2017;16(11): 2357-74. https://doi.org/10.1016/S2095-3119(16)61608-8
14. Sadhukhan D, Mukherjee T, Sarkar A, Devi ND, Bisarya D, Kumar V, et al. A comprehensive analysis of drought stress responses in rice (*Oryza sativa* L.): insights into developmental stage variations from germination to grain filling. Int J Environ Cli Chan. 2024;14(7):141-58. https://doi.org/10.9734/ijecc/2024/v14i74260
15. Linh TM, Huong TT, Viet BT. Study on the flowering in rice plant (*Oryza sativa* cv. OM5451). Sci Technol Develop J. 2023;26(3):2943–9.https://doi.org/10.32508/stdj.v26i3.4088
16. Zhang J, Zhang Y Yan, Song N Yuan, Chen Q Li, Sun H Zheng, Peng T, et al. Response of grain-filling rate and grain quality of mid-season indica rice to nitrogen application. J Int Agric. 2021;20(6):1465-73. https://doi.org/10.1016/S2095-3119(20)63311-1
17. Yoshida S. Fundamental of rice crop science. Los Baños (Philippines) International Rice Research Institute;1981
18. McQueen RJ, Garner SR, Nevill-Manning CG, Witten IH. Applying machine learning to agricultural data. Comput. Electron. Agric 1995;12(4): 275-93. https://doi.org/10.1016/0168-1699(95)98601-9
19. Zhang Y, editor. New advances in machine learning. London: Intechopen; 2010.
20. Parameswari P, Rajathi N, Harshanaa KJ. Machine learning approaches for crop recommendation. In: 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA); 2021 Oct 8–9 Coimbatore, India. New York: IEEE; 2021 [cited 2024 Sep 19]. p. 1-5. https://ieeexplore.ieee.org/document/9675480
21. Khan N, Kamaruddin MA, Ullah Sheikh U, Zawawi MH, Yusup Y, Bakht MP, et al. prediction of oil palm yield using machine learning in the perspective of fluctuating weather and soil moisture conditions: evaluation of a generic workflow. Plants. 2022;11 (13):1697. https://doi.org/10.3390/plants11131697
22. Kovacevic T, Mrcela L, Mercep A, Kostanjcar Z. Impact of look-back period on soil temperature estimation using machine learning models. In: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); 2020 May 25-28, Dubrovnik, Croatia. New York: IEEE; 2020 [cited 2024 Sep 19]. p. 1-5. https://ieeexplore.ieee.org
23. Weiss U, Biber P, Laible S, Bohlmann K, Zell A. Plant species classification using a 3D LIDAR sensor and machine learning. In: 2010 Ninth International Conference on Machine Learning and Applications; 2010 Dec12-14, Washington, DC, USA. New York: IEEE; 2020 [cited 2024 Sep 19]. p. 339-45. https://ieeexplore.ieee.org
24. Krishnan S, Aruna SK, Kanagarathinam K, Venugopal E. Identification of dry bean varieties based on multiple attributes using catboost machine learning algorithm. Scientific Programming. 2023;2023(1):2556066. https://doi.org/10.1155/2023/2556066
25. Mayur Rajaram Salokhe. Machine learning: applications in agriculture (crop yield prediction, disease and pest detection). Int J Adv Res Sci Comm Techno. 2023;592-7. https://doi.org/10.48175/ijarsct-12088
26. Alfred R, Obit JH, Chin CPY, Haviluddin H, Lim Y. Towards paddy rice smart farming: a review on big data, machine learning and rice production tasks. IEEE Access. 2021;9:50358-80. https://doi.org/10.1109/access.2021.3069449
27. Kaur AP, Bhatt DP, Raja L. Applications of deep learning and machine learning in smart agriculture: A Survey. In: Hashmi MF, Kesakr AG, editors. Advances in environmental engineering and green technologies. IGI Global; 2023. p.34-57. https://doi.org/10.4018/978-1-6684-9975-7.ch003
28. Pham HT, Awange J, Kuhn M, Nguyen BV, Bui LK. Enhancing crop yield prediction utilizing machine learning on satellite-based vegetation health indices. Sensors. 2022;22(3):719. https://doi.org/10.3390/s22030719
29. Hashemi MGZ, Tan PN, Jalilvand E, Wilke B, Alemohammad H, Das NN. Yield estimation from SAR data using patch-based deep learning and machine learning techniques. Computers and Electronics in Agriculture. 2024;226. https://doi.org/10.1016/j.compag.2024.109340
30. Murugesan R, Sudarsanam SK, Malathi G, Vijayakumar V, Neelanarayanan V, et al. Artificial intelligence and agriculture 5.0. Int J Rec Technol Eng. 2019;8:1870-7https://doi.org/10.35940/ijrte.B1510.078219.
31. Gul D, Banday RUZ. Transforming Crop Management Through Advanced AI and Machine Learning: Insights into Innovative Strategies for Sustainable Agriculture. London: Intechopen. 2024. https://doi.org/10.5772/acrt.20240030
32. Abuzanouneh K, N. Al-Wesabi F, Abdulrahman Albraikan A, Al Duhayyim M, Al-Shabi M, Mustafa Hilal A, et al. Design of machine learning based smart irrigation system for precision agriculture. Computers, Mat Continua.2022;72(1):109-24. https://doi.org/10.32604/cmc.2022.022648
33. Chang NB, Bai K. Multisensor data fusion and machine learning for environmental remote sensing. Baco Raton: CRC Press; 2018. https://doi.org/10.1201/9781315154602
34. Chang NB, Bai K, Chen CF. Integrating multisensor satellite data merging and image reconstruction in support of machine learning for better water quality management. J Environ Manage. 2017;201:227-240.https://doi.org/10.1016/j.jenvman.2017.06.045
35. Xu X, Gao P, Zhu X, Guo W, Ding J, Li C, et al. Design of an integrated climatic assessment indicator (ICAI) for wheat production: A case study in Jiangsu Province, China. Ecol Indi. 2019;101:943-53. https://doi.org/10.1016/j.ecolind.2019.01.059

36. Chen C, Bao Y, Zhu F, Yang R. Remote sensing monitoring of rice growth under *Cnaphalocrocis medinalis* (Guenée) damage by integrating satellite and UAV remote sensing data. Int J Rem Sens. 2024;45(3):772-90. https://doi.org/10.1080/01431161.2024.2302350

37. Iatrou M, Karydas C, Tseni X, Mourelatos S. Representation learning with a variational autoencoder for predicting nitrogen requirement in rice. Rem Sens. 2022;14(23):5978. https://doi.org/10.3390/rs14235978

38. Yoshimura T, Koike N, Hashimoto K, Oura K, Nankaku Y, Tokuda K. Discriminative feature extraction based on sequential variational autoencoder for speaker recognition. In: 2018 Asia–Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC); 2018 Nov 12-15; Honolulu, HI, USA. New York: IEEE; 2018 [cited 2024 Sep 19]. p. 1742-66 Available from: https://ieeexplore.ieee.org

39. Mak HWL, Han R, Yin HHF. Application of variational autoencoder (VAE) model and image processing approaches in game design. sensors. 2023;23(7):3457. https://doi.org/10.3390/s23073457

40. Langevin S, Bethune C, Horne P, Kramer S, Gleason J, Johnson B, et al. Useable machine learning for Sentinel–2 multispectral satellite imagery. In: Bruzzone L, Bovolo F, Benediktsson JA, editors. Image and signal processing for remote sensing Spain Online Only SPIE; 2021. p. 112. https://doi.org/10.1117/12.2599951

41. Teke M, Deveci HS, Haliloglu O, Gurbuz SZ, Sakarya U. A short survey of hyperspectral remote sensing applications in agriculture. In: 2013 6th International Conference on Recent Advances in Space Technologies (RAST); 2013 12–14 June; Istanbul, Turkey. New York: IEEE; 2013 [cited 2024 Sep 19]. p. 171-176. Available from: https://ieeexplore.ieee.org

42. Dutta S, Patel NK, Srivastava SK. District wise yield models of rice in Bihar based on water requirement and meteorological data. J Ind Soc Rem Sens. 2001;29(3):175-82. https://doi.org/10.1007/BF02989929

43. Huang Jingfeng, Tang Shuchuan, Ousama Abou–Ismail, Wang Renchao. Integration of remote sensing data and simulation model to estimate rice yield. In: 2001 International Conferences on Info–Tech and Info–Net. Proceedings. Beijing, China. New York: IEEE; 2001 [cited 2024 Sep 19]. p. 101-7. Available from: https://ieeexplore.ieee.org

44. Zhuo W, Huang J, Li L, Huang R, Gao X, Zhang X, et al. Assimilating sar and optical remote sensing data into wofost model for improving winter wheat yield estimation. In: 2018 7th International Conference on Agro–geoinformatics (Agro–geoinformatics). Hangzhou. New York: IEEE; 2018 [cited 2024 Sep 19]. p. 101-7. Available from: https://ieeexplore.ieee.org

45. Divakar MS, Elayidom MS, Rajesh R. Design and implementation of an efficient and cost effective deep feature learning model for rice yield mapping. Int J Comp Sci Engin. 2022;25(2):128. https://doi.org/10.1504/ijcse.2022.122205

46. Prakash MA, Pazhanivelan S, Muthumanickam D, Ragunath KP, Sivamurugan AP. Mapping rice area in the cauvery delta zone of Tamil nadu using Sentinel 1a synthetic aperture radar (SAR) Data. Int J Environ Clim Chan. 2023;13(10):195-204. https://doi.org/10.9734/ijecc/2023/v13i102630

47. Chen J, Lin H, Huang C, Fang C. The relationship between the leaf area index (LAI) of rice and the C band SAR vertical/horizontal (VV/HH) polarization ratio. Int J Rem Sens. 2009;30(8):2149-54. https://doi.org/10.1080/01431160802609700

48. Beriaux E, Jago A, Lucau–Danila C, Planchon V, Defourny P. Sentinel -1 Time series for crop identification in the framework of the future cap monitoring. Rem Sens. 2021;13(14):2785. https://doi.org/10.3390/rs13142785

49. Wali E, Tasumi M, Moriyama M. Combination of linear regression lines to understand the response of sentinel–1 dual polarization sar data with crop phenology–case study in Miyazaki, Japan. Rem Sens. 2020;12(1): 189. https://doi.org/10.3390/rs12010189

50. Srikanth P, Chakraborty A, Murthy CS. Crop monitoring using microwave remote sensing. In: Mitran T, Meena RS, Chakraborty A, editors. Geospatial technologies for crops and soils. Singapore: Springer Singapore; 2021. p. 201-28. https://doi.org/10.1007/978–981–15–6864–0_5

51. Zeng Z, Gan Y, Kettner AJ, Yang Q, Zeng C, Brakenridge GR, et al. Towards high resolution flood monitoring: An integrated methodology using passive microwave brightness temperatures and Sentinel synthetic aperture radar imagery. J Hydrol. 2020;582:124377. https://doi.org/10.1016/j.jhydrol.2019.124377

52. Mandal D, Rao YS. SASYA: An integrated framework for crop biophysical parameter retrieval and within–season crop yield prediction with SAR remote sensing data. Remote Sensing Applications. 2020;20:100366. https://doi.org/10.1016/j.rsase.2020.100366.

53. Dissanayake DMPW, Rathnayake RMKT, Chathuranga LLG. Crop yield forecasting using machine learning techniques–a systematic literature review. KDU J Multidis Stud. 2023;5(1):54-65. https://doi.org/10.4038/kjms.v5i1.62

54. Narayan KJ. Review of crop yield prediction using machine learning techniques. International Journal for Res App Sci Engin Techno. 2021;9(VI):4626-28.https://doi.org/10.22214/ijraset.2021.36058

55. Kulyal M, Saxena P. Machine learning approaches for crop yield prediction: A Review. In: 2022 7th International Conference on Computing, Communication and Security (ICCCS); 2022 Nov 3; Seoul, Korea. New York: IEEE; 2022 [cited 2024 Sep 19]. p. 1–7. Available from: https://ieeexplore.ieee.org

56. Chaudhary S, Mongia S, Sharma S, Singh N. Classification based Interactive Model for Crop Yield Prediction: Punjab State. In: 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART); 2022. Moradabad, India. New York: IEEE; 2022 [cited 2024 Sep 19]. p. 1-6. Available from: https://ieeexplore.ieee.org

57. Wijayanto AW, Putri SR. Estimating rice production using machine learning models on multitemporal landsat–8 satellite images (Case Study: Ngawi Regency, East Java, Indonesia). In: 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom). Malang, Indonesia New York: IEEE; 2022 [cited 2024 Sep 19]. p. 1-6. Available from: https://ieeexplore.ieee.org

58. Shuaibu N, Obunadike GN, Jamilu BA. Crop yield prediction using selected machine learning algorithms. Fud J Sci. 2024;8(1):61-8.https://doi.org/10.33003/fjs–2024–0801–2220

59. Paudel D, Boogaard H, De Wit A, Janssen S, Osinga S, Pylianidis C, et al. Machine learning for large–scale crop yield forecasting. Agric Sys. 2021;187:103016. https://doi.org/10.1016/j.agsy.2020.103016

60. GSM, Paudel S, Nakarmi R, Giri P, Karki SB. Prediction of crop yield based–on soil moisture using machine learning algorithms. In: 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS); 2022. Tashkent, Uzbekistan; New York: IEEE; 2022 [cited 2024 Sep 19]. p. 491-5. Available from: https://ieeexplore.ieee.org

61. Thirumal S, Latha R. Automated rice crop yield prediction using sine cosine algorithm with weighted regularized extreme learning machine. In: 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS); 2023. Madurai, New York: IEEE; 2023 [cited 2024 Sep 19]. p. 491-5. Available from: https://ieeexplore.ieee.org

62. Reddy DJ, Kumar MR. . Crop yield prediction using machine learning algorithm. In: 2021 4th International Conference on Computing and Communications Technologies (ICCCT); 2021. Chennai, India. New York: IEEE; 2021 [cited 2024 Sep 19]. p. 611-6. Available from: https://ieeexplore.ieee.org

63. Pavani S, Sophy Beulet P A. Improved precision crop yield prediction using weighted–feature hybrid svm: analysis of ml algorithms. IETE J Res. 2023;1-13. https://

doi.org/10.1080/03772063.2023.2192000

64. Islam MD, Di L, Qamer FM, Shrestha S, Guo L, Lin L, et al. Rapid rice yield estimation using integrated remote sensing and meteorological data and machine learning. Remote Sens. 2023;15 (9):2374. https://doi.org/10.3390/rs15092374

65. Hanya I, Ishii K, Noguchi N. Monitoring rice growth environment by low-altitude remote sensing using spectroradiometer. IFAC Proceedings Volumes. 2010;43(26):184-9. https://doi.org/10.3182/20101206-3-JP-3009.00032

66. Yin W, Cen H, He Y. Quantification of nitrogen status in rice canopy by low-altitude hyperspectral remote sensing.). Am Soc Agric Bio Engin. 2017. https://doi.org/10.13031/aim.201700537

67. Han-Ya I, Ishii K, Noguchi N. Satellite and aerial remote sensing for production estimates and crop assessment. Environ Control Biol. 2010;48(2):51-8. https://doi.org/10.2525/ecb.48.51

68. Bulacan Agricultural State College, Cuevas CM, Saludes R, Dorado M, Sta Cruz P. Assessment of nutrient status of lowland irrigated rice (*Oryza sativa* L.) using low altitude remote sensing. Philippine Journal of Agric Biosys Engineer. 2021;17(1):39-48. https://doi.org/10.48196/017.01.2021.04

69. K. C. Swain, S. J. Thomson, H. P. W. Jayasuriya. Adoption of an unmanned helicopter for low-altitude remote sensing to estimate yield and total biomass of a rice crop. Transactions of the ASABE. 2010;53(1):21-7. https://doi.org/10.13031/2013.29493

70. Pazhanivelan S, Geethalakshmi V, Tamilmounika R, Sudarmanian NS, Kaliaperumal R, Ramalingam K, et al. Spatial rice yield estimation using multiple linear regression analysis, semi-physical approach and assimilating sar satellite derived products with DSSAT crop simulation model. Agronomy. 2022;12(9):2008. https://doi.org/10.3390/agronomy12092008

71. Wang F, Wang F, Zhang Y, Hu J, Huang J, Xie L, et al. Rice yield estimation at pixel scale using relative vegetation indices from unmanned aerial systems. In: 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics); 2019. Istanbul, Turkey. New York: IEEE; 2019 [cited 2024 Sep 19]. p. 1-6. Available from: https://ieeexplore.ieee.org

72. Tanaka Y, Watanabe T, Katsura K, Tsujimoto Y, Takai T, Tanaka TST, et al. Deep learning enables instant and versatile estimation of rice yield using ground-based rgb images. Plant Phenomics. 2023;5:0073. https://doi.org/10.34133/plantphenomics.0073

73. Leelavathi KS, Rajasenathipathi M. A Novel crop yield prediction using deep learning and dimensionality reduction. Int Res J Multidis Scope. 2024;05(01):101-12. https://doi.org/10.47857/irjms.2024.v05i01.0158

74. M. C, Dhanraj RK. Ensemble deep learning algorithm for forecasting of rice crop yield based on soil nutrition levels. ICST Transactions on Scalable Information Systems. 2023;e7. https://doi.org/10.4108/eetsis.v10i3.2610

75. Demircioğlu A. The effect of feature normalization methods in radiomics. Insights into Imaging. 2024;15(1):2. https://doi.org/10.1186/s13244-023-01575-7

76. Prince Pallayan B, Chitradurga Manjunath M. Artificial bee colony algorithm-based feature selection and hybrid ml framework for efficient rice yield prediction. Int J Electr Comput Eng Sys. 2024;15 (3): 235-46. https://doi.org/10.32985/ijeces.15.3.3.

77. Anitha S, Vanitha M. Imputation methods for missing data for a proposed VASA dataset. Int J Innov Technol Explor. Eng. 2019;9(1). https://doi.org/10.35940/ijitee.A5204119119

78. Awad M. Toward precision in crop yield estimation using remote sensing and optimization techniques. Agriculture. 2019;9(3): 54. https://doi.org/10.3390/agriculture9030054

79. Zhang Y, Wang D, Zhou Q. Advances in crop fine classification based on Hyperspectral Remote Sensing. In: 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics); 2019. Istanbul, Turkey. New York: IEEE; 2019 [cited 2024 Sep 19]. p. 1-6. Available from: https://ieeexplore.ieee.org

80. Ling J, Zhang H. WCDL: A Weighted Cloud Dictionary Learning Method for Fusing Cloud-Contaminated Optical and SAR Images. IEEE J Sel Top Appl Earth Obs Remote Sens.2023;16: 2931-41. https://doi.org/10.1109/JSTARS.2023.3259469

81. Xu F, Shi Y, Ebel P, Yu L, Xia GS, Yang W, et al. GLF-CR:SAR-enhanced cloud removal with global-local fusion. ISPRS J Photogramm Remote Sens. 2022;192:268-78. https://doi.org/10.1016/j.isprsjprs.2022.08.002

82. Wu M, Huang W, Niu Z, Wang Y, Wang C, Li W, et al. Fine crop mapping by combining high spectral and high spatial resolution remote sensing data in complex heterogeneous areas. Comp Elect Agric. 2017;139: 1-9. https://doi.org/10.1016/j.compag.2017.05.003

83. Sweet L belle, Müller C, Anand M, Zscheischler J. Cross-validation strategy impacts the performance and interpretation of machine learning models. Artif L Earth Syst. 2023;2(4):e230026. https://doi.org/10.1175/aies-d-23-0026.1

84. Qiu Y. Rice yield prediction based on LSTM and GRU. SPIE 2023; 126041V:1-7. https://doi.org/10.1117/12.2674760

85. Chang CH, Lin J, Chang JW, Huang YS, Lai MH, Chang YJ. hybrid deep neural networks with multi-tasking for rice yield prediction using remote sensing data. agriculture. 2024;14(4):513. https://doi.org/10.3390/agriculture14040513

86. Wangkheimayum N, Paliwal HB. development of rice yield forecasting model using linear regression for imphal west district, manipur, india. Int. J. Environ. Clim. 2023;13(9):485-490. https://doi.org/10.9734/ijecc/2023/v13i92258

87. Islam MdM, Matsushita S, Noguchi R, Ahamed T. Development of remote sensing-based yield prediction models at the maturity stage of boro rice using parametric and nonparametric approaches. Remote Sens App: Soc Environ. 2021;22:100494. https://doi.org/10.1016/j.rsase.2021.100494

88. Pham HT, Awange J, Kuhn M, Nguyen BV, Bui LK. Enhancing crop yield prediction utilizing machine learning on satellite-based vegetation health indices. Sensors. 2022;22(3):719. https://doi.org/10.3390/s22030719

89. Alfred R, Obit JH, Chin CPY, Haviluddin H, Lim Y. Towards paddy rice smart farming: a review on big data, machine learning and rice production tasks. IEEE Access. 2021;9:50358-80. https://doi.org/10.1109/access.2021.3069449

90. Mia MdS, Tanabe R, Habibi LN, Hashimoto N, Homma K, Maki M, et al. multimodal deep learning for rice yield prediction using uav-based multispectral imagery and weather data. Remote Sens. 2023;15(10):2511. https://doi.org/10.3390/rs15102511

91. Bellis ES, Hashem AA, Causey JL, Runkle BRK, Moreno-García B, Burns BW, et al. Detecting Intra-field variation in rice yield with unmanned aerial vehicle imagery and deep learning. Frontiers in Plant Science. 2022;13:716506. https://doi.org/10.3389/fpls.2022.716506

92. Son NT, Chen CF, Cheng YS, Toscano P, Chen CR, Chen SL, et al. Field-scale rice yield prediction from Sentinel-2 monthly image composites using machine learning algorithms. Ecolog Info. 2022;69: 101618. https://doi.org/10.1016/j.ecoinf.2022.101618

93. Guo Y, Li S, Zhang Z, Li Y, Hu Z, Xin D, et al. Automatic and accurate calculation of rice seed setting rate based on image segmentation and deep learning. Fron Pl Sci. 2021;12:770916. https://doi.org/10.3389/fpls.2021.770916

94. Lu J, Li J, Fu H, Tang X, Liu Z, Chen H, et al. Deep learning for multi-source, data-driven crop yield prediction in Northeast China. Agriculture. 2024;14(6):794. https://doi.org/10.3390/agriculture14060794

95. Liu L, Xie Y, Zhu B, Song K. Rice leaf chlorophyll content estimation with different crop coverages based on Sentinel-2. Ecol Inform. 2024;81:102622. https://doi.org/10.1016/j.ecoinf.2024.102622

96. Brinkhoff J, Clarke A, Dunn BW, Groat M. Analysis and forecasting of Australian rice yield using phenology–based aggregation of satellite and weather data. Agric Forest Meteorol. 2024;353:110055. https://doi.org/10.1016/j.agrformet.2024.110055

97. Fu X, Zhao G, Wu W, Xu B, Li J, Zhou X, et al. Assessing the impacts of natural disasters on rice production in Jiangxi, China. International J Remote Sens. 2022;43(5):1919-41. https://doi.org/10.1080/01431161.2022.2049914

98. Mansaray LR, Wang F, Kanu AS, Yang L. Evaluating Sentinel–1A datasets for rice leaf area index estimation based on machine learning regression models. Geocarto Int. 2022;37(5):1225-36. https://doi.org/10.1080/10106049.2020.1773545

99. Son NT, Chen CF, Cheng YS, Toscano P, Chen CR, Chen SL, et al. Field–scale rice yield prediction from Sentinel–2 monthly image composites using machine learning algorithms. Ecol Info. 2022;69: 101618. https://doi.org/10.1016/j.ecoinf.2022.101618

100. Eugenio FC, Grohs M, Schuh M, Venancio LP, Schons C, Badin TL, et al. Estimated flooded rice grain yield and nitrogen content in leaves based on RPAS images and machine learning. Field Crops Res. 2023;292:108823.https://doi.org/10.1016/j.fcr.2023.108823.

101. Yu W, Yang G, Li D, Zheng H, Yao X, Zhu Y, et al. Improved prediction of rice yield at field and county levels by synergistic use of SAR, optical and meteorological data. Agric For Meteoro. 2023;342:109729. https://doi.org/10.1016/j.agrformet.2023.109729

102. Pazhanivelan S, Ragunath KP, Sudarmanian NS, Kumaraperumal R, Setiyono T, Quicho ED. Integrating time–series sar data and oryza crop growth model in rice area mapping and yield estimation for crop insurances. Int Arch Photo, Remote Sens Spa Info Sci. 2019;XLII-3/W6: 239-43. https://doi.org/10.5194/isprs-archives-XLII-3-W6-239-2019

103. Jain V, Saxena S, Dubey S, Choudhary K, Sehgal S, Neetu, et al. Rice (*kharif*) production estimation using sar data of different satellites and yield models: a comparative analysis of the estimates generated under fasal project. Int Arch Photo, Remote Sens Spa Info Sci. 2019;XLII–3/W6:99-107. https://doi.org/10.5194/isprs–archives–XLII–3–W6–99–2019

104. Arumugam P, Chemura A, Schauberger B, Gornott C. Remote Sensing Based Yield Estimation of Rice (Oryza Sativa L.) Using Gradient Boosted Regression in India. Remote Sens. 2021;13 (12):2379. https://doi.org/10.3390/rs13122379

105. Jiya EA, Illiyasu U, Akinyemi M. Rice yield forecasting: a comparative analysis of multiple machine learning algorithms. J Info Sys Infor. 2023;5(2):785-99. https://doi.org/10.51519/journalisi.v5i2.506

106. Ranjan AK, Parida BR. Paddy acreage mapping and yield prediction using Sentinel–based optical and SAR data in Sahibganj district, Jharkhand (India). Spatial Info Res. 2019;27(4):399-410. https://doi.org/10.1007/s41324–019–00246–4

107. Satpathi A, Setiya P, Das B, Nain AS, Jha PK, Singh S, et al. Comparative Analysis of Statistical and Machine Learning Techniques for Rice Yield Forecasting for Chhattisgarh, India. Sustainability. 2023;15(3):2786. https://doi.org/10.3390/su15032786.

108. Su Y xue, Xu H, Yan L jiao. Support vector machine–based open crop model (SBOCM): Case of rice production in China. S J Biol Sci. 2017;24(3):537-47. https://doi.org/10.1016/j.sjbs.2017.01.024

109. Hemalatha N, Akhil W, Vinod R. Computational yield prediction of rice using knn regression. In: Shukla PK, Singh KP, Tripathi AK, Engelbrecht A, editors. Computer vision and robotics. singapore: Singapore: Springer Nature; 2023. p. 295-308. https://doi.org/10.1007/978–981–19–7892–0_23

110. Siyal AA, Dempewolf J, Becker–Reshef I. Rice yield estimation using Landsat ETM + Data. J App Rem Sens. 2015;9(1):095986. https://doi.org/10.1117/1.JRS.9.095986

111. Islam MD, Di L, Qamer FM, Shrestha S, Guo L, Lin L, et al. Rapid rice yield estimation using integrated remote sensing and meteorological data and machine learning. Remote Sensing. 2023;15(9):2374. https://doi.org/10.3390/rs15092374