Innovative: International Multi-disciplinary Journal of Applied Technology
(ISSN 2995-486X) VOLUME 03 ISSUE 5, 2025

Integrating Sentinel _ 2 and Landsat 8 Imagery with Machine Learning Algorithms for Crop Yield Prediction and Agricultural Monitoring

Wassan A. Hassan [1], Ayat Farhan Abed [2], Rafah R. Ismail [3], Noor Zubair Kouder [3], Dina H. Haider [4]

[1, 4] Department of Forensic Chemistry, Higher Institute of Forensic Sciences, Al-Nahrain University, Jadriya, Baghdad, Iraq

[2] Department of Scholarships& Culture Relations, Presidency of Al-Nahrain University, Al-Nahrain University, Baghdad, Iraq

[3] Dept. of Astronomy and Space, College of Science, University of Baghdad, Baghdad, Iraq

Abstract:

This study looks into the integration of Sentinel-2 and Landsat 8 satellite imagery with machine learning algorithms for enhanced crop yield prediction and agricultural monitoring. The use of remote sensing technologies has transformed precision agriculture through real-time assessment of vegetation health, soil conditions, and environmental changes. A good complement to this long-term history and thermal imagery from Landsat 8, Sentinel-2 has high spatial resolution and high revisit cycles to enable a robust dataset for accurate yield estimation. Machine learning models, which include decision trees, random forests, and neural networks, have started processing vast datasets in agriculture that offer predictive insights into crop growth patterns, resource optimization, and risk management. Data pre-processing techniques such as atmospheric correction and cloud removal are very essential in making the satellite imagery reliable, improving the accuracy of vegetation indices and predictive models. Even though data quality, model interpretability, and high implementation costs are still issues, advances in artificial intelligence and deep learning have been refining remote sensing applications. The study highlights the transformative potential of integrating satellite technology and machine learning to enhance food security, optimize resource utilization, and promote sustainable farming practices and pave the way for more precise and data-driven agricultural decision-making.

Keywords: Sentinel-2, Landsat 8, satellite imagery, machine learning, crop yield prediction, agricultural monitoring, precision agriculture, remote sensing, vegetation indices, decision trees, random forests.

1. INTRODUCTION

Crop yield prediction is a prediction of accurate estimates of potential yields for specific crops at specific times during seasons of specific regions [1]. It involves precision predictability in the application that requires science combined with knowledge about local areas of location and types of crops produced therein [2]. Diverse data sets comprising weather, seed genetics, properties of the soil, management factors, statistical databases, etc, are applied for estimations. An experimentally proven multi-level algorithm is implemented in arriving at yield computations [3].

1.1.1. Advantages of Predicting Yield

Precise yield predictability offers a multitude of benefits for agricultural stakeholders [4]:

- **Enhanced food security:** Governments and organisations may anticipate potential food shortages and take precautions with the help of accurate yield projections. To ensure its residents have consistent access to food, the government may prepare ahead of time for imports, oversee export regulations, and allocate resources effectively.
- **Improved farm management**: Businesses may reduce their impact on the environment, save money, and increase efficiency by planning production around anticipated yields and making the most of available resources (such as water, fertiliser, and pesticides).
- **Informed decision-making**: By using data-driven yield estimates, farmers may minimise losses and maximise revenues by strategically planting different crop kinds, obtaining inputs, and harvesting the crop.
- **Risk management in financial:** In order to tailor their services, insurance companies and banks use yield forecast data when evaluating agricultural loans and crop insurance. Both farmers and lenders will feel less danger as a result of this.
- **Better market forecast:** Yield predictability at the regional and national levels leads to better market forecast and stabilizes food prices without any drastic fluctuations that would adversely affect the farmer, businessperson, and the consumer.
- **Supply-demand management has improved:** the predictability of yield means a reduction in having to buffer for seed companies and associated waste. Food producing companies can get better at stock management, avoid wasting space and reduce waste.

1.2. Overview of Sentinel-2 and Landsat 8 imagery

The sentinel-2 and Landsat 8 are among the most applied remote sensing satellite systems that provide vital information for monitoring agriculture, environmental management, and land cover change [5]. These satellites offer high-resolution, multispectral imagery that helps in the evaluation of vegetation health, soil condition, and growth patterns of crops [6]. Their integration with advanced computational techniques, such as machine learning, has significantly improved the accuracy of crop yield prediction models.

Part of the Copernicus program by the European Space Agency, Sentinel-2 offers systematic and continuous Earth observation data. It comprises two satellites: Sentinel-2A and Sentinel-2B. These two satellites work in tandem to create global coverage with a revisit time of five days [7]. Fitted with MSI, Sentinel-2 captures data of 13 spectral bands covering from the visible range to the shortwave infrared range. They include indices such as NDVI, EVI which are used for monitoring vegetation accurately [8]. These, in turn help in the plant health assessment of biomass accumulation besides water stress evaluation. The 10 to 60 meters spatial resolutions of Sentinel-2 bands makes it very much useful for the precision agriculture industry, where exact field-level understanding of the crop under cultivation is quite often required.

Landsat 8 is a project of NASA and the USGS as part of the long-running Landsat program, initiated in 1972 to collect Earth observation data. It captures imagery at 11 spectral bands, including visible, near infrared, shortwave infrared, and thermal infrared wavelengths. Its OLI and TIRS can accurately assess the surface temperature and land cover, which is necessary for detecting crop stress, soil moisture levels, and irrigation efficiency [9]. With a spatial resolution of 30 meters for most bands and a revisit cycle of 16 days, Landsat 8 is ideal for monitoring long-term agricultural trends, tracking deforestation, and analyzing land use changes over vast areas.

These sources provide free, open-access data to researchers, agronomists, and policymakers. When used in tandem, these datasets enhance the temporal resolution, enabling researchers to observe crop dynamics and environmental changes more frequently. For example, Sentinel-2's five-day revisit cycle is complemented by Landsat 8's longer 16-day cycle, so crop development in relation to seasonal changes can be better monitored. With these, the machine learning models to be used in yield estimation, soil fertility analysis, and pest infestation will be improved because the two satellites will improve the spectral information [10].

The fusion of Sentinel-2 and Landsat 8 imagery ensures a comprehensive agricultural monitoring approach with real-time as well as historical information, helping to make better informed decisions by all stakeholders. Evolution in satellite technology will lead the integration of artificial intelligence and cloud computing to add more precision into agriculture to guarantee sustainable food and resource management [11].

1.3. Applications Of Remote Sensing

There has been a lot of effort to use remote sensing to investigate agriculture, and the field is becoming increasingly popular [12]. From basic tasks like field identification to more complex ones like precision farming, remote sensing has many uses in agriculture. First, let's take a brief look at the ways remote sensing has benefited farming [13]:

1. **Land Cover Mapping:** Land cover mapping is a popular use case for remote sensing. It is possible to distinguish between every type of land cover on Earth through the use of land cover mapping. When deciding what crops to grow and how much of each to harvest, land cover is a key factor. In crop management, crops are chosen based on field and soil types, and treatment procedures are developed to increase crop yields while decreasing the likelihood of damage caused by pests or diseases. To accomplish this, several factors such as weather patterns both present and historical, models of crop production, soil characteristics, and market circumstances are integrated with the various crop kinds.

2. **Precision Agriculture:** Precision agriculture, also referred to as Precision farming, describes the set of techniques, instruments, and management methods which help to ensure that treatments given to crops enhance growth and returns of a farm through adapting applications according to changes in biophysical conditions prevailing across the agricultural area instead of an identical application all over the land. It has been feasible, due to advancements in remote sensing and the addition of functionalites in GIS, to characterize and model almost every crop, with a mapping feature-which puts a lot into GIS and remote sensing for a better future of precision agriculture.

3. **Irrigated Land Cover Mapping:** Another important application of remote sensing in agriculture is Irrigated Land Cover Mapping. The observation from space of the Earth's surface offers a reliable, cost-effective, and accurate synoptic information. The data assist in agricultural land cover mapping, especially on the land cover mapping theme. Image classification systems are often used to make current strategies for characterizing agricultural land cover.

4. **Crop health monitoring:** By analysing spectral data collected from satellites, aeroplanes, or instruments on the ground, remote sensing can keep an eye on the growth and well-being of crops. A farmer can use this data to pinpoint how much water, fertiliser, or pesticide their crops need at specific stages.
5. **Yield estimation:** More recently, agricultural yields may be estimated using plant height, biomass, and chlorophyll content, thanks to advancements in remote sensing. satellite imaging. Farmers may make better use of this information when planning harvests and managing to crops.

2. EVOLUTION OF PRECISION AGRICULTURE

The path to precision agriculture is an evolutionary process under the influence of technology and our growing understanding of natural ecosystems. Mechanization, coming at the beginning of the 20th century, brought about a complete revolution in agriculture based on experience and intuition [14]. The true precursor to precision agriculture, however, was the Green Revolution that swept the earth in the 1950s and 1960s, promoting high-yielding crops and new agricultural practices to increase global production by leaps and bounds [15]. In its successes, the approach taken of the Green Revolution was a one-size-fits-all approach as all fields were treated as identical units. But fields are heterogenous, varying in soil properties, moisture, and nutrient content. It is during this time that site-specific crop management emerged in the late 1980s and was said to be the birth of precision agriculture. Precision agriculture, on the other hand, started booming in the 1990s when the Global Positioning System was invented.

GPS technology allowed farmers to map their field variations in an accurate manner. This meant applying inputs in the right amounts and avoiding waste. At the same time, yield monitors were developed to assist farmers measure the yield variations across different fields [16]. The 2000s have brought forth VRT that allows the applications of different input rates in various parts of the field under the data obtained from GPS and yield monitors. With the integration of GIS in precision agriculture, it has enabled the collection, storage, analysis, and display of geographically referenced information that can further improve the precision of agricultural practices [17].

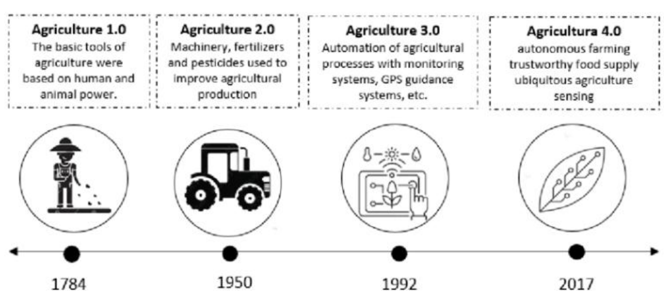

Figure 1: Timelines of agricultural revolutions [18]

2.1. Remote Sensing Technologies

While many technologies have shaped the evolution of precision agriculture, remote sensing has been one of the most influential. Remote sensing is the process of obtaining information about an object or area without physically touching it. It is used in a variety of ways in agriculture, including crop health monitoring, yield prediction, and irrigation management. There are two types of remote sensing technologies active and passive. Active remote sensors emit radiation and measure the reflected signal, whereas passive sensors measure natural radiation emitted or reflected by the object or area of interest. In the first generation of remote sensing technology in agriculture, aerial photography was primarily utilized. Satellites became the dominant remote sensing platform during the last quarter of the 20th century. Large-scale agricultural monitoring through satellite remote sensing enjoys many advantages. For instance, satellites achieve wide coverage and frequent revisits [19]. They may also offer multiple spectral bands or resolution. However, in real applications, there are some limitations, and these might include lower spatial resolution and increased sensitivity to atmospheric conditions. Recently, an alternative tool has gained significant attention: Unmanned Aerial Vehicles, or drones. Drones can capture high-resolution imagery at low altitudes, enabling detailed monitoring of individual plants [20]. With advancements in technology, multispectral and hyperspectral imaging have emerged. Multispectral imaging captures images in specific, broad bands of the electromagnetic spectrum, while hyperspectral imaging captures images in many very narrow bands. These technologies allow for a more detailed and accurate assessment of crop health, nutrient status, and other critical parameters[21].

Active Remote Sensing

Active remote sensing systems produce their energy source. One of the most common forms of active remote sensing is radar, which is short for Radio Detection and Ranging. It emits radio waves and measures the time delay for the signal to bounce back after hitting the target. The lag is then measured in a manner of distance, which will then allow the creation of a three dimensional image of the object or space [22]. LiDAR is another form of active remote sensing technology, in which the light energy takes the form of a pulsed laser, which measures variable distances to Earth. This technology generates highly accurate three-dimensional information regarding the Earth's shape and its surface properties and thus finds a high use in topographic and elevation mapping for agricultural landscapes [23].

Passive Remote Sensing

Passive remote sensing systems are those that rely on the natural radiation emitted or reflected from an observed surface. They are quite common in precision agriculture and come in two categories: multispectral and hyperspectral sensors. Multispectral sensors measure certain bands of the electromagnetic spectrum, typically between the visible and near-infrared regions. The application of these sensors in precision agriculture includes crop health monitoring, yield prediction, and detection of pests and diseases [24]. Hyperspectral sensors measure hundreds of narrow, contiguous spectral bands over the whole visible, near-infrared, and short-wave infrared parts of the electromagnetic spectrum. These sensors make it possible to visualize the scene with greater resolution and identify the same materials by their spectral signatures.

Platforms for Remote Sensing

Remote sensing technologies may be mounted on different platforms that have different advantages and disadvantages. Satellites give a wide area of coverage, and the data are highly repeatable; however, data from satellites could be affected by cloud cover and the temporal resolution may be compromised [25]. Aircraft, on the other hand, can take higher spatial resolutions and are also flexible in operation, but these are costlier and restricted to flight regulations. UAV, or drone, is an emerging tool in precision agriculture; it can fly relatively low and slow compared to aircraft. This enables a higher resolution for images of much higher resolution than can be accomplished by manned aircraft. The low cost and ease of use make drones accessible to individual farmers [26]

Table 1: Remote Sensing Technologies [27]

Technology Type	Active/Passive	Advantages	Limitations
Radar	Active	Can operate in all weather and lighting conditions	Complex data processing
LiDAR	Active	Highly accurate 3D mapping	Costly, complex data analysis
Multispectral	Passive	Broad applications in agriculture	Limited spectral

Sensors			resolution
Hyperspectral Sensors	Passive	High spectral resolution	Complex data analysis, high data volume
Satellites	Platform	Broad, consistent coverage	Affected by cloud cover, limited temporal resolution
Aircraft	Platform	High spatial resolution, flexible use	Higher cost, subject to flight regulations
Drones	Platform	High-resolution imagery, accessible and affordable	Limited flight duration, weather-dependent

2.2. Comparison of Sentinel-2 and Landsat 8 data

Table 2: Comparison of Sentinel-2 and Landsat 8 data [28]

Feature	Sentinel-2	Landsat 8
Launch Date	Both the Sentinel-2A and Sentinel-2B spacecrafts were launched in the years 2015 and 2017, respectively.	Landsat 8 launched on February 11, 2013
Mission	An ESA-led initiative as part of the Copernicus Earth Observation Program	Part of the long-running Landsat Earth Observation Program managed by NASA/USGS
Orbit Type	Sun-synchronous orbit, 786 km above Earth's surface on average	Sun-synchronous orbit with an average altitude of 705 km
Revisit Time	5-day revisit time (with both Sentinel-2A and Sentinel-2B in operation, providing frequent coverage)	16-day revisit cycle (depending on location, can be less frequent in high-latitude regions)
Spatial Resolution	The visibility bands have a range of 10 meters, the red-edge and shortwave infrared bands of 20 meters, and the atmospheric correction bands of 60 meters.	Visual, near-infrared, and shortwave infrared bands typically have a range of 30 meters, whereas thermal infrared bands have a range of 100 meters.
Spectral Bands	13 bands ranging from visible light to shortwave infrared, including unique bands for vegetation monitoring and water content analysis	11 bands, including visible, near-infrared, and thermal infrared bands, suitable for vegetation, land cover, and temperature analysis
Radiometric Resolution	12-bit (providing a greater range of values for pixel intensities, allowing better detection of subtle changes)	12-bit (providing similar detailed information as Sentinel-2 in terms of pixel intensity range)
Data Availability	All data collected by Sentinel-2 are freely available through the Copernicus Open Access Hub.	Data collected by Landsat 8 is freely available online via the USGS Earth Explorer

		website.
Key Instruments	Images are captured over thirteen different spectral bands using the Multi-Spectral Instrument (MSI).	The Operational Land Imager (OLI) uses infrared and visible light, while the Thermal Infrared Sensor (TIRS) collects data on temperatures.
Temporal Coverage	Ongoing data collection since 2015, with continuous operations expected as part of the Copernicus program	Ongoing since 2013, part of the Landsat continuity program; continuous data collection starting from Landsat 1 in 1972
Cloud Cover Sensitivity	Sensitive to cloud cover, with limited data availability in areas with frequent cloud cover	Sensitive to cloud cover, but longer revisit time can mitigate gaps in cloud-covered areas

3. MACHINE LEARNING IN CROP YIELD PREDICTION

The agricultural sector plays a crucial role in meeting the nutritional needs of the world's expanding population [29]. Farmers must maximise their resources to reduce wastage and maximise production if they are to meet the rising demand for food. Machine learning has emerged as a potent tool for modern farmers to use in their pursuit of the modern agricultural goal line of predicting and analysing harvest growth. Precision agriculture, sometimes known as "smart farming," is a relatively new method of farming that makes use of cutting-edge technological tools to maximise harvest yield while minimising input costs. The goal of smart farming is to maximise harvest yield with minimal input of water, fertiliser, and energy [30].

The processes for crop analysis and prediction that are based on machine learning and the IOT are shown in Figure 2. When thinking about smart farming, the Internet of Things is a crucial technology. The optimal times to sow, irrigate, and harvest crops may be defined using data obtained from Internet of Things (IoT) devices that assess soil moisture, temperature, and other environmental parameters. By using Internet of Things (IoT) sensors, we can increase yield quality and quantity by precisely dosing crops with water and fertiliser [31].

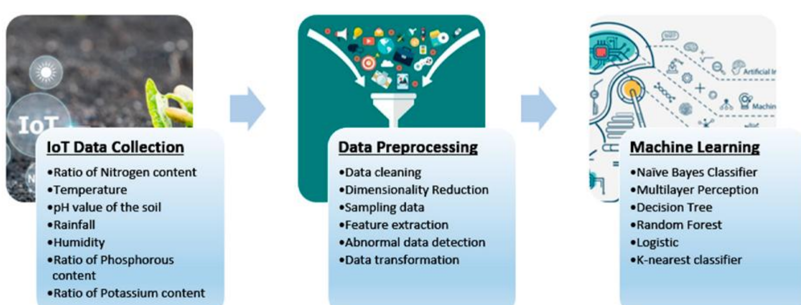

Figure 2: Crop study and forecast using IoT and machine learning [32]

ML applications have taken successful steps in decision-making in recent years, and they have penetrated our lives in many domains, from the health industry to the defence industry, from education to urbanisation [33]. Simultaneously, it started making tech and information solutions by laying the groundwork for the new search engine architecture, which includes ChatGPT, Google Bard, and other tools based on artificial intelligence. Numerous research firms have shown that new trends will continue to expand across different platforms. Regarding this, the implementation of machine learning models will change several domains, such as chip design and traffic predictions, and the influence of these systems and solutions inside technology would make them a big multiplier [34].

In essence, machine learning algorithms are required to gather bulk and quality data. If you want reliable results and precise forecasts, you need high-quality and large-scale data collecting [35]. In a nutshell, "big data" describes data that is large, fast, and diverse. For example, the size of the corpus allows for comprehensive findings to be produced by the data and guarantees a decrease in randomness [36]. Simultaneously, there's need for improvement in the organisation of big data analyses. The success rate of an analysis will be enhanced if data is used from many sources or datasets. Sensors, social media, data networks, physical gadgets, the financial market, and healthcare facilities are just a few of the many places that may provide data. Web collection, access pathways, and application programming interfaces all provide data access. Both static datasets and stream data are possible forms of the data. When processing data, it is common practice to include data from several sources. Machine learning algorithms rely heavily on clean, pre-processed data, which makes data collection and preparation all the more important [37].

For example, this algorithm type can sift through mountains of data collected by Internet of Things (IoT) devices. This area of study has the potential to revolutionise agricultural production and yield forecasting, and it has grown at a rapid pace. In order for computers to learn and, in turn, become better over time, machine learning algorithms examine data using statistical and mathematical models and algorithms to provide predictions [38]. Data acquired from farms, particularly the agricultural production region, may be used to train machine learning algorithms. This data can include weather patterns, soil characteristics, crop growth phases, and pest and disease outbreaks. Machine learning algorithms may utilise this data assessment to provide very accurate predictions about yield, quality, and growth [39].

The use of data and technology to optimise agricultural practices like fertilisation, irrigation, and pest management in order to increase yield and quality is known as precision farming, and it is a significant use of machine learning in agriculture [40]. Machine learning algorithms can sift through mountains of data collected from sources like drone footage, soil sensors, and satellite images to create detailed maps of nutrition and moisture levels, as well as maps of crop growth. In order to maximise harvest yield while minimising water loss, farmers will be able to use these maps to precisely regulate irrigation and fertiliser applications [41]. Machine learning also aids farmers in weighing market demand and environmental factors when deciding which crops to sow. By examining past market data and weather trends, machine learning algorithms are able to forecast the demand for certain crops. As a result, the model can suggest the best times and places to plant [42]. Farmers may increase their profitability and decrease the likelihood of crop failure with this method. Machine learning can assess the quality of harvested crops in addition to forecasting their growth and yield. To guarantee that fruits and vegetables are of high quality and maturity, machine learning algorithms can forecast their form, colour, and texture. Only the highest quality fruit and vegetables are offered to customers in this fashion, thanks to efficient harvesting [43,44].

Deploying machine learning in agriculture is fraught with challenges, such as insufficient data foundation, prohibitive sensor and equipment costs, and the requirement for specialised expertise in solution development and maintenance. Machine learning's potential for profit in agriculture will become clearer as more farms adopt precision agriculture and collect data. It is worth noting that machine learning is still in its early stages in the agricultural sector, and that further study is needed to fully harness the potential of this technology. So far, the outcomes are encouraging, and it's probable that machine learning will become more crucial [45].

Table 3: Research Studies on Machine Learning and Smart Farming

Author Name	**Topic Covered**	**Research Study**	**Title**
Li et al. (2022)	Machine learning	Developed machine learning	" Machine

[46]	models for crop yield prediction	models to predict wheat yield in China using multi-source environmental data (climatic, soil, remote sensing) and tested various machine learning algorithms.	Learning Models for Predicting China's Wheat Yield Using Environmental Data from Multiple Sources"
Van Klompenburg et al. (2020) [47]	Crop yield prediction using machine learning	Conducted a systematic literature review on crop yield prediction models, analyzing machine learning techniques like ANN, SVM, and deep learning for improving prediction accuracy.	"Systematic Review on Predicting Crop Yield using Machine Learning Techniques"
Kuradusenge et al. (2023) [48]	Application of machine learning for Irish potato and maize yield prediction	investigated how well random forests and gradient boosting machines forecast maize and Irish potato yields by using climate and satellite information.	"Application of Machine Learning for Predicting Yield of Irish Potatoes and Maize"
Xu et al. (2021) [49]	Smart farming approach using IoT, cloud computing, and AI	Introduced a six-domain smart farming model incorporating IoT devices, cloud computing, and AI, highlighting the benefits and challenges of digital agriculture.	"Smart Farming Approach Based on IoT, Cloud Computing, and AI"
Moysiadis et al. (2022) [50]	Cloud computing for smart farming	Developed a cloud computing-based web application using microservices architecture for real-time data collection and analysis, facilitating decision-making in crop management.	"Cloud Computing-Based Web Application for Smart Farming Using Microservices Architecture"
Ranjan et al. (2022) [51]	AI in soil and crop management	Investigated AI-powered decision support systems for optimizing irrigation, fertilization, and pest control strategies in precision agriculture.	"Artificial Intelligence in Soil and Crop Management"
Oré et al. (2020) [52]	Drone-borne DInSAR for crop growth monitoring	investigated the application of Differential Interferometric Synthetic Aperture Radar (DInSAR), which is carried by drones, for high-resolution crop yield and health estimates.	"Use of Drone-Borne DInSAR for Crop Growth Monitoring"
Gehlot et al. (2022) [53]	Deep learning in crop production	Conducted technical analysis comparing deep learning	" Technical Evaluation of

	prediction	models (CNNs, LSTMs) and traditional machine learning techniques in predicting crop production.	Machine Learning and Deep Learning for Crop Production Prediction "

3.1. Machine Learning Algorithms In Agriculture

1. Naïve Bayes

The Naïve Bayes model, which is based on Bayes' theorem, is often used for classification problems in numerous applications. Multinomial, Bernoulli, and Gaussian algorithms are the three that make up Naive Bayes [54]. When it comes to classification difficulties, the Naive Bayes Algorithm is absolutely indispensable. The system is based on the premise that all characteristics have the same chance of occurring and that these chances are completely separate from one another. After another event has taken place, the Bayes theorem determines the probability of the third event taking place. Bayes theorem is utilised in multi-class classification. Furthermore, as compared to other ML approaches, it is both faster and easier to develop. Little data is needed for training. It works with both continuous and discrete data. It doesn't care about superfluous features and can scale up or down quite well.

2. Decision Trees

In supervised machine learning, decision trees, which are similar to flowcharts, are very commonly used for classification and prediction. A decision tree (DT) can be decomposed into a set of rules, where each path from the root node to a leaf node represents a different rule. In this tree structure, every leaf node is a class, which is achieved when an attribute satisfies the condition given by the previous branch. The internal nodes in this structure act as decision nodes, which can be interpreted as tests, conditions, or attributes used to make the classification or prediction [55].

3. KNN

Classification and regression issues are both addressed by utilising a kNN-based machine learning approach. Algorithms in supervised ones use labelled data. There are a few of techniques to calculate the distances between locations, which is the basis of the procedure. Always keep in mind that the distance can only be positive or zero [56]. One way to achieve this is by using the absolute values, square root, or raising the distance to a certain power. Prior to using the kNN method, it is imperative that all labelled data undergo pre-processing. The first step is to standardise all of the data. When there are too many features, kNN stops working, hence it's necessary to do feature selection to remove the unimportant ones. Please complete the fields that are blank. In every other case, the specific record must be deleted. Adding more samples to the train set might boost performance. One major drawback of KNN is that computational costs and algorithmic speed both rise as dataset sizes get larger.

4. Random Forests (RF)

As an example of ensemble learning, the RF approach increases a model's efficiency by connecting several classifiers to tackle a difficult task. The resultant "forest" using this technique is really just a collection of decision trees. Random RF characteristics are chosen at each decision split [57]. The connection among trees is reduced by the selection of attributes that aid in prediction and increase efficiency. After dividing the dataset into smaller parts, the The final prediction that a Random Forest machine learning classification method produces is by combining the multiple decision trees.

It comes as a form of ensemble learning under the Bagging technique. Here, using randomly selected subsets of rows and features of the original dataset, decision trees are constructed, which can be suitable for both classification and regression tasks. Also, Random Forest enhances the accuracy of the model, overcomes overfitting, and it performs really well when dealing with large data of high dimensions.

3.2. Advantages and challenges of ML-based yield prediction

Advantages of Machine Learning-Based Yield Prediction [58]

1. **Improved Accuracy:** ML algorithms can research a complex data set to highly accurately predict crop yield, including any variables, be it weather conditions, soil health, or historical.
2. **Real-time Decision Making:** ML algorithms can process real-time data. This makes the farmer have an opportunity to take decisions based on the real-time information about crop and environmental health [59].
3. **Resource Optimization:** Accurate yield prediction helps optimize water, fertilizers, and pesticides resources in that it becomes efficient and does not waste its share of them.
4. Risk Management: **The ML models can predict variations in yield and environmental risks that will enable the farmer to mitigate risks such as adverse weather or pest outbreaks.**
5. Crop Selection and Breeding: **ML can be applied in the selection of crops most suitable for a given region and identifying traits for selective breeding that improve crop performance and resilience.**

Challenges of Machine Learning-Based Yield Prediction

1. **Data Quality and Availability:** In ML models, high-quality and reliable data are expected. Lack of correct data normally makes more errors on the wrong side of prediction, especially for developing countries.
2. **Data Integration and Compatibility:** Weather, soil, and yield data integration from multiple sources might become complex and generate compatibility issues [60].
3. **Complexity and interpretability of ML models:** The complexity of deep learning models often inhibits their interpretability. This tends to lower the confidence level of farmers in such results [61].
4. **Need for expertise:** The successful deployment of ML models requires specialized knowledge in both data science and agriculture, which acts as a barrier for most farmers.
5. **Cost of Technology and Infrastructure:** The initial investment and maintenance costs for ML-based yield prediction, including data collection devices and computing infrastructure, are likely to be too high for small-scale farmers.

4. DATA INTEGRATION AND PRE-PROCESSING

4.1. Fusion of Sentinel-2 and Landsat 8 imagery

Integrating Sentinel-2 and Landsat 8 imagery combines data from both satellite systems to enhance the precision and reliability of remote sensing applications, especially in agriculture and environmental monitoring [62]. Sentinel-2, managed by the European Space Agency (ESA), offers high spatial resolution and frequent revisit intervals, making it well-suited for detailed vegetation assessments [63]. On the other hand, Landsat 8 offers NASA and the USGS with ample historical datasets with a moderate spatial resolution and a greater spectral range. The fusion of the datasets exploits both systems' advantages, thereby facilitating better crop yield forecasts, more accurate land use classification, and more detailed assessments of the environment.

The most important advantage of merging Sentinel-2 and Landsat 8 data is the increase in temporal resolution. Sentinel-2 revisits a location approximately every five days, while Landsat 8 returns every 16 days [64]. By integrating both datasets, the effective revisit frequency increases, thereby allowing for more consistent monitoring of crop conditions and land cover changes. This is particularly valuable in agriculture, where frequent observations are necessary to track vegetation health, soil moisture levels, and phenological variations [65].

Data fusion also improves spectral resolution since Sentinel-2 and Landsat 8 have different spectral bands. Sentinel-2 has a broader band in the visible, near-infrared (NIR), and shortwave infrared (SWIR) spectra, which helps in the detailed analysis of vegetation [66]. On the other hand, Landsat 8 has thermal infrared bands that give critical information about land surface temperatures, which are necessary for drought stress monitoring and irrigation requirements. These datasets, combined together, would help the researchers derive more accurate indices like Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Land Surface Temperature (LST) and better crop health and resource management assessments [67].

Combining these datasets helps to sidestep some of the shortcomings created by cloud cover [68]. Because both Sentinel-2 and Landsat 8 function in the optical spectrum, such data can be obscured by clouds, thereby creating gaps or introducing untrustworthiness into their records [69]. Combining images taken from each satellite means that where clouds obscure one dataset at an image location for a given period, those gaps are likely to be filled by data captured by the other satellite, thereby increasing the probability of cloud-free observations for any given date [70].

An important application of Sentinel-2 and Landsat 8 fusion relates to machine learning and predictive modelling [71]. Information multispectral and temporal extracted from the two sources can improve the performance of algorithms in crop yield prediction, land cover classification, and precision agriculture. Training the machine learning models using both datasets enables the researchers to enhance the accuracy of predictions that relate to soil health, vegetation stress, and productivity in crops. This leads to more efficient decision-making for farmers, policymakers, and agricultural organizations [72].

Integration on such a high level has also some advantages though, but mixing Sentinel-2 and Landsat 8 brings along with this some challenges -dissimilar spatial resolutions, sensor's calibration, also the differences concerning data processing among others [73]. There is a larger spatial resolution, 30 m, 10–20m for the two key spectral bands in Sentinel -2 compared with Landsat. This means alignment would be some what different compared to the existing datasets. In some cases, the radiometric calibration and the techniques used in atmospheric correction would vary to allow for effective integration [74]. Advanced image processing techniques, such as machine learning-based super-resolution and deep learning algorithms, are applied to harmonize the data and make it more usable.

In conclusion, integration with Landsat 8 imagery adds a comprehensive approach toward agricultural monitoring as well as yield prediction through the Sentinel-2 [75]. Complementing strengths in such satellites here improve time, spectral as well as spatial resolution in a remote sensing application, and accordingly, result in more accurate output [76]. Even if data fusion posed some challenges, developing image processing technique and machine learning have further led to better the feasibility of employing these merged data sets more operational [77].

4.1.1. Data pre-processing techniques

Data pre-processing is the process that needs to be performed before any type of analysis on satellite images in order to ensure that the raw data obtained from remote sensing is accurate, standardized, and suitable for further processing [78,79]. Pre-processing involves various techniques applied to correct distortion, remove unwanted noise, and standardize imagery for application with machine learning models or analytical assessments [80]. Atmospheric correction and cloud removal are the most important pre-processing techniques used on Sentinel-2 and Landsat 8 imagery [81].

4.1.2. Atmospheric Correction

Since images of satellites captured from space pass through the Earth's atmosphere before they reach sensors, they introduce distortion in spectral reflectance values [82]. Such distortions arise due to the presence of water vapor, aerosols, dust, and other gases within the atmosphere, which scatter and absorb the sunlight. In the case of such interferences, remote sensing analyses are adversely affected; in applications like vegetation monitoring, land cover classification, and crop yield prediction [83]. Without atmospheric correction, the effects might distort the interpretation of satellite images and make it difficult to compare data collected on different dates or from different sources [84]. Hence, these atmospheric correction techniques are implemented to remove or minimize the distortions in the reflectance values so that the latter better represents the actual conditions of the Earth's surface.

Dark Object Subtraction is a widely used method for atmospheric correction. The basic assumption here is that in certain regions of the image, like deep water bodies or shadowed areas, near-zero reflectance is expected in some spectral bands [85]. Reflectance values in such regions are then regarded as errors introduced by atmospheric scattering. This unwanted signal is subtracted from the entire image [86]. DOS thus helps correct the reflectance values. This way, the data can be trusted further. The technique is simple and computationally efficient and is applied widely in a number of remote sensing applications. However, this method is based on an assumption that there exist truly dark objects within the scene, which might not always be the case [87].

More advanced in terms of atmospheric correction technique is the FLAASH, Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes. The method, grounded on the physical process, makes use of radiative transfer equations that represent how light and particles within the atmosphere interact with one another [88]. Contrasting the DOS empirical approach, FLAASH incorporates the factors of altitude, humidity, and aerosol content in an image for atmospheric corrections, giving much better precision, particularly with hyperspectral and multispectral high-resolution imagery. FLAASH is widely used in scientific researches and environmental monitoring projects, given that it gives the best highly accurate surface reflectance values. However, it has vast computational needs, requiring more complex data inputs in the atmosphere, and is relatively hard to apply [89].

Sen2Cor is the specialized atmospheric correction tool developed for Sentinel-2 imagery. The algorithm converts top-of-atmosphere (TOA) reflectance to bottom-of-atmosphere (BOA) reflectance, thus enhancing the spectral analysis accuracy. Since Sentinel-2 data is predominantly used for monitoring vegetation, agricultural assessments, and land cover mapping, Sen2Cor is the most important one to ensure coherence in different images. It corrects for variations in atmospheric conditions like changes in humidity and aerosol levels, which might otherwise affect the accuracy of NDVI (Normalized Difference Vegetation Index). Therefore, Sen2Cor is commonly used by the researchers and analysts working with Sentinel-2 data because it helps in the efficient automation of atmospheric corrections [90].

Similarly, for the Landsat 8 imagery, LaSRC has been applied for atmospheric correction. It corrects with physics-based corrections that are fine-tuned to be unique sensor characteristics of the Landsat. These contributions include the solar angle, the aerosol thickness, and the water vapor content. Data gathered during radically different periods would then be compared by using this code on the Landsat imagery. This will come in very handy for long-term environmental studies, monitoring deforestation, and analyzing agriculture productivity [91].

Atmospheric correction forms one of the fundamental preprocessing processes in remote sensing. This is because atmospheric correction allows for reliable and consistent analyses of images. It becomes imperative, especially at the integration point of data gathered from various platforms, such as Sentinel-2 and Landsat 8.Because these satellites have different sensor characteristics and revisit frequencies, atmospheric correction standardizes the reflectance values of the data, thus making it easy to fuse and compare the data. Whether using simple empirical methods as DOS or advanced physics-based algorithms like FLAASH, Sen2Cor, and LaSRC, atmospheric correction remains an essential procedure for ensuring high-quality satellite data for various applications in agriculture, forestry, and environmental monitoring.

Cloud Removal

The major challenge in optical satellite remote sensing is cloud cover, which covers the land surface and limits the ability to analyze images. Since optical sensors depend on visible and infrared light to capture surface features, the presence of clouds and their shadows can result in missing or distorted data. This is the most challenging aspect in agricultural monitoring, forest management, and land classification, where the imagery should be clear to accurately assess the issues. Several techniques have been designed to remove the clouds from imagery so that an area covered with clouds can be detected, masked, or even replaced to yield reliable data to be further processed [92].

The most common process used for cloud removal is called cloud masking. It identifies the pixels covered by clouds and removes those pixels from the dataset. There are many algorithms that have been successfully developed over time, including Fmask (Function of Mask) and MAJA (MACCS-ATCOR Joint Algorithm) [93]. These methods use spectral characteristics and thermal bands to distinguish clouds from the land and water surfaces. With reflectance values in certain wavelengths, including SWIR and TIR, these algorithms are able to accurately identify clouds and their shadows. Once masked, cloud-covered pixels do not pollute subsequent analyses of images. It has applications in high-precision remote sensing data where the requirement could be something like crop health assessment and environmental monitoring.

The second method of cloud removal is temporal compositing. In this, a composite image with minimum cloud cover is generated instead of using an individual image. Since satellite sensors return to the same location several times, cloud-free pixels in one image can be replaced by cloud-affected pixels in another image [94]. This technique is widely applied in vegetation monitoring and land cover mapping, where seasonal variations in surface conditions need to be analyzed over time. Temporal compositing helps construct a more complete and accurate representation of the land surface by selecting the clearest pixels from a series of images.

Beyond improving the cloud removal process, the data fusion techniques merge images coming from various satellite resources. For example, the revisit time and spectral characteristics are different for Sentinel-2 and Landsat 8, and thus each becomes a complementary source for cloud-free data. Pixels in the image coming from another satellite resource can fill gaps and complement the cloud-filled area in the other dataset. This cross-sensor approach improves data availability and ensures a more holistic spatial and temporal coverage. Data fusion is the most useful in applications that require constant monitoring such as precision agriculture, land use change detection, and disaster responses.

Advances in AI, such as improved machine learning, have thus enabled improved methods of cloud removal. This now includes reconstruction of covered areas with deep learning models using CNNs [95]. These models are trained on large datasets of cloud-free imagery and learn to predict surface features that are missing in spatial patterns. Machine learning-based approaches can generate realistic representations of cloud-obscured regions by analyzing surrounding pixels and historical data [96 - 98]. This technique is promising for high-resolution remote sensing applications, where even small cloud-covered areas can significantly impact the accuracy of analysis .

Preprocessing steps in remote sensing include atmospheric correction and cloud removal as a way to ensure that satellite imagery is as accurate and reliable as possible [99]. Providing multispectral and hyperspectral image quality is enhanced by these preprocessing methods through the minimization of distortion attributable to the atmosphere and cloud cover, which will significantly help in better decision-making terms in agricultural management, environmental monitoring, and land classification [100]. The further development of satellite-based remote sensing for various applications will be more effective by the integration of traditional methods with advanced machine learning and data fusion approaches as technology continues to evolve.

5. CONCLUSION

The integration of Sentinel-2 and Landsat 8 imagery with machine learning algorithms significantly enhances crop yield predictions and agriculture monitoring based on the complementarity of both satellite systems. The high spatial resolution and high temporal capability of Sentinel-2 complement long-term historical data and thermal imaging capability from Landsat 8, thus allowing a more holistic approach toward agricultural assessment. This integration makes possible the establishment of exact indices about vegetation; the monitoring in real time of the health of the soil; and the prediction of crop yield, upon which it is possible for farmers, agronomists, and policymakers to make data-driven decisions. Remote sensing applications evolved from simple land cover mapping to advanced precision agriculture, optimizing resource use, mitigating risks, and improving farm productivity in general. This further improves these capabilities through processing huge datasets on weather patterns, soil properties, and crop growth, thus leading to smarter and more adaptive agricultural practices. Challenges still abound in the form of data quality, model interpretability, and high costs of implementation, which have so far restricted the full-scale adoption of these technologies, especially for small-scale farmers. Effective data pre-processing techniques like atmospheric correction and cloud removal are a critical step for making satellite images reliable. Accuracy of satellite observations has been further enhanced by various DOS, FLAASH, and Sen2Cor methods in cloud masking and data fusion to be used in crop monitoring applications. Continued development in artificial intelligence and deep learning continues to take agricultural analytics into uncharted territories, that is, making remote sensing more accurate and efficient. Indeed, these make the resultant transformation of satellite technology and AI-driven analytics a potent opportunity for the promotion of sustainable farming practices, enhancing food security, and mitigating agriculture-related climate risks- an eventual case toward the widespread global adoption of these transformative technologies.

REFERENCES

1. Vashisht, S.; Kumar, P.; Trivedi, M.C. Improvised Extreme Learning Machine for Crop Yield Prediction. In Proceedings of the 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 27–29 April 2022; pp. 754–757. [**Google Scholar**]

2. OpenAI. New and Improved Content Moderation Tooling. OpenAI. 2022. Available online: **https://openai.com/blog/new-and-improved-content-moderation-tooling/** (accessed on 1 April 2023).

3. Google. Bard Chatbox. Google. Available online: **https://bard.google.com** (accessed on 2 April 2023).

4. Dean, J. The deep learning revolution and its implications for computer architecture and chip design. In Proceedings of the IEEE International Solid-State Circuits Conference-(ISSCC), San Francisco, CA, USA, 16–20 February 2020. [**Google Scholar**]

5. Cui, Y.W.; Henrickson, K.; Ke, R.; Pu, Z.; Wang, Y. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 4883–4894. [**Google Scholar**] [**CrossRef**]

6. Shahrin, F.; Zahin, L.; Rahman, R.; Hossain, A.J.; Kaf, A.H.; Abdul Malek Azad, A.K.M. Agricultural Analysis and Crop Yield Prediction of Habiganj using Multispectral Bands of Satellite Imagery with Machine Learning. In Proceedings of the 11th International Conference on Electrical and Computer Engineering (ICECE), Dhaka, Bangladesh, 17–19 December 2020; pp. 21–24. [**Google Scholar**]
7. Tawseef, A.S.; Tabasum, R.; Faisal, R.L. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Comput. Electron. Agric.* **2022**, *198*, 107119. [**Google Scholar**]
8. Senthil KS, D.; Mary, D.S. Smart farming using Machine Learning and Deep Learning techniques. *Decis. Anal. J.* **2022**, *3*, 100041. [**Google Scholar**]
9. Senthil, K.M.; Akshaya, R.; Sreejith, K. An Internet of Things-based Efficient Solution for Smart Farming. *Procedia Comput. Sci.* **2023**, *218*, 2806–2819. [**Google Scholar**]
10. Vivek, S.; Ashish, K.T.; Himanshu, M. Technological revolutions in smart farming: Current trends, challenges & future directions. *Comput. Electron. Agric.* **2022**, *201*, 107217. [**Google Scholar**]
11. Mamatha, J.C.K. Machine learning based crop growth management in greenhouse environment using hydroponics farming techniques. *Meas. Sens.* **2023**, *25*, 100665. [**Google Scholar**] [**CrossRef**]
12. Rashid, M.; Bari, B.S.; Yusup, Y.; Kamaruddin, M.A.; Khan, N. A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches with Special Emphasis on Palm Oil Yield Prediction. *IEEE Access* **2021**, *9*, 63406–63439. [**Google Scholar**] [**CrossRef**]
13. Alpaydın, “Introduction to machine learning, second edition.” MIT Press, 2010. ISBN: 978-0-262-01243-0.
14. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D., “machine learning in agriculture: a review”, Sensors, vol. 18, no. 8, pp. 2674, August 2018. https://doi.org/10.3390/s18082674.
CrossRef
15. Sabitha, “A study on sectorial contribution of gdp in india from 2010 to 2019”, AJEBA, 19, no. 1, pp. 18-31, January 2020. Article no. AJEBA. 62227. CrossRef
16. Jain A., “Analysis of growth and instability in the area, production, yield, and price of rice in India”, Journal of Social Change and Development, vol. 2, pp. 46-66, N/A,
17. Wolfert S, Ge L, Verdouw C, Bogaardt MJ, “Big data in smart farming– a review. Agricultural Systems”, 153, pp. 69-80, May 2017. CrossRef
18. Sangeeta, Shruthi G. “Design and implementation of crop yield prediction model in ” International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 4, pp. 305-310, Apr. 2020.
19. Johnson LK, Bloom JD, Dunning RD, Gunter CC, Boyette MD, Creamer NG, “Farmer harvest decisions and vegetable loss in primary production. Agricultural Systems”, 176, pp. 102672, November 2019. CrossRef
20. Sharma A, Jain A, Gupta P, Chowdary V. Machine learning applications for precision agriculture: A comprehensive review. IEEE Access. 2020 Dec 31;9:4843-73. CrossRef
21. Meshram V, Patil K, Meshram V, Hanchate D, Ramkteke SD. Machine learning in agriculture domain: A state-of-art survey. Artificial Intelligence in the Life Sciences. 2021 Dec

1;1:100010.
CrossRef

22. Reddy, D. J., & Kumar, M. R. (2021). Crop Yield Prediction using Machine Learning Algorithm. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). doi:10.1109/iciccs51141.2021.9432236 CrossRef
23. Ranjini B Guruprasad, Kumar Saurav, Sukanya Randhawa,"Machine learning methodologies for paddy yield Estimation in India: a case study", CrossRef
24. S. P. Raja, B. Sawicka, Z. Stamenkovic and G. Mariammal, "Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers," IEEE Access, vol. 10, pp. 23625-23641, 2022, doi: 10.1109/ACCESS.2022.3154350.
CrossRef
25. Venugopal, Anakha, S, Aparna, Mani, Jinsu, Mathew, Rima, Williams, Vinu. "Crop yield prediction using machine learning algorithms." International Journal of Engineering Research & Technology (IJERT) NCREIS – 2021, vol. 09, no. 13, pp. 1-6, 2021.
26. S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and SHAURYA, "Crop recommender system using machine learning approach," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351. CrossRef
27. Suresh, N., et al. "Crop yield prediction using random forest algorithm." 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 279-282, 2021, doi: 10.1109/ICACCS51430.2021.9441871. CrossRef
28. E. Manjula and S. Djodiltachoumy, "A model for prediction of crop yield," Int. J. Comput. Intell. Inform., vol. 6, no. 4, pp. 298–305, 2017.
29. van Klompenburg, , Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review in computers and electronics in Agriculture, 177, pp. 105709. doi: 10.1016/j.compag.2020.105709. CrossRef
30. M. Liu, T. Wang, A. K. Skidmore, and X. Liu, "Heavy metal-induced stress in rice crops detected using multi-temporal Sentinel-2 satellite images," Sci. Total Environ., vol. 637-638, pp. 18-29, Oct. 2018. CrossRef
31. K. E. Eswari and L. Vinitha, "Crop yield prediction in tamil nadu using bayesian network," Int. J. Intell. Adv. Res. Eng. Comput., vol. 6, no. 2, pp. 1571-1576, 2018.
32. D. A. Reddy, B. Dadore, and A. Watekar, "Crop recommendation system to maximize crop yield in ramtek region using machine learning," Int. J. Sci. Res. Sci. Technol., vol. 6, no. 1, pp. 485-489, Feb. 2019. CrossRef
33. Priya, P., Muthaiah, U., Balamurugan, M. "Predicting yield of the crop using machine learning algorithm." International Journal of Computer Science and Mobile Computing, 4, no. 5, pp. 1-7, May 2015.
34. Medar, Ramesh, S, Vijay, Shweta. "Crop yield prediction using machine learning " International Journal of Advanced Research in Computer Science and Software Engineering, vol. 9, no. 5, pp. 1-6, May 2019. CrossRef
35. Basso, B., Cammarano, D., & Carfagna, E. (2013). Review of Crop Yield Forecasting Methods and Early Warning Systems. In First Meeting of the Scientific Advisory Committee of the Global Strategy to improve Agricultural and Rural Statistics, FAO Headquarters.

36. Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. Precision Agriculture, 5(4), 359-387. https://doi.org/10.1023/B:PRAG.0000040806.3 9604.aa

37. Gebbers, R., & Adamchuk, V. I. (2010). Precision Agriculture and Food Security. Science, 327(5967), 828-831. https://doi.org/10.1126/science.1183899

38. Chen, H., Lan, Y., Fritz, B. K., Hoffmann, W. C., & Liu, S. (2021). Review of agricultural spraying technologies for plant protection using unmanned aerial vehicle (UAV). International Journal of Agricultural and Biological Engineering, 14(1), 38-49. DOI: 10.25165/j.ijabe.20211401.5714

39. Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture, 143, 23-37. https://doi.org/10.1016/j.compag.2017.09.037

40. Khan, S. I., Khaliq, A., & Prabhakar, M. (2015). Remote Sensing and Geographical Information System Application in Irrigation Water Management. Journal of Applied and Natural Science, 7(2), 658-666. https://doi.org/10.3390/w10050608

41. Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. IEEE Geoscience and Remote Sensing Letters, 14(5), 778-782. https://doi.org/10.1109/LGRS.2017.2681128

42. Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. Sensors, 14(11), 20078-20111. https://doi.org/10.3390/s141120078

43. Lobell, D. B., Asner, G. P., Ortiz-Monasterio, J. I., & Benning, T. L. (2003). Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agriculture, Ecosystems & Environment, 94(2), 205-220. https://doi.org/10.1016/S0167- 8809(02)00021-X

44. Mahlein, A. K. (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. Plant Disease, 100(2), 241-251. https://doi.org/10.1094/PDIS-03-15-0340-FE

45. Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies. Managing Forest Ecosystems, 27. https://doi.org/10.1007/978-94-017-8663-8

46. Li, L.; Wang, B.; Feng, P.; Liu, D.L.; He, Q.; Zhang, Y.; Wang, Y.; Li, S.; Lu, X.; Yue, C.; et al. Developing machine learning models with multi-source environmental data to predict wheat yield in China. *Comput. Electron. Agric.* **2022**, *194*, 106790. [**Google Scholar**] [**CrossRef**]

47. van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [**Google Scholar**] [**CrossRef**]

48. Kuradusenge, M.; Hitimana, E.; Hanyurwimfura, D.; Rukundo, P.; Mtonga, K.; Mukasine, A.; Uwitonze, C.; Ngabonziza, J.; Uwamahoro, A. Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize. *Agriculture* **2023**, *13*, 225. [**Google Scholar**] [**CrossRef**]

49. Xu, W.; Kaili, Z.; Tianlei, W. Smart Farm Based on Six-Domain Model. In Proceedings of the IEEE 4th International Conference on Electronics Technology (ICET), Chengdu, China, 7–10 May 2021; pp. 417–421. [**Google Scholar**]

50. Moysiadis, V.; Tsakos, K.; Sarigiannidis, P.; Petrakis, E.G.M.; Boursianis, A.D.; Goudos, S.K. A Cloud Computing web-based application for Smart Farming based on microservices architecture. In Proceedings of the 11th International Conference on Modern Circuits and

Systems Technologies (MOCAST), Bremen, Germany, 8–10 June 2022; pp. 1–5. [**Google Scholar**]

51. Ranjan, P.; Garg, R.; Rai, J.K. Artificial Intelligence Applications in Soil & Crop Management. In Proceedings of the IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, India, 21–23 December 2022; pp. 1–5. [**Google Scholar**]
52. Oré, G.; Alcântara, M.S.; Góes, J.A.; Oliveira, L.P.; Yepes, J.; Teruel, B.; Castro, V. Crop Growth Monitoring with Drone-Borne DInSAR. *Remote Sens.* **2020**, *12*, 615. [**Google Scholar**] [**CrossRef**]
53. Gehlot, A.; Sidana, N.; Jawale, D.; Jain, N.; Singh, B.P.; Singh, B. Technical analysis of crop production prediction using Machine Learning and Deep Learning Algorithms. In Proceedings of the International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 24–25 September 2022; pp. 1–5. [**Google Scholar**]
54. England, J. R., & Viscarra Rossel, R. A. (2018). Proximal sensing for soil carbon accounting. Soil, 4(2), 101-122. https://doi.org/10.5194/soil-4-101-2018
55. Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems Engineering, 114(4), 358-371. https://doi.org/10.1016/j.biosystemseng.2012.0 8.009
56. Pôças, I., Gonçalves, P., & Pereira, L. S. (2019). NDVI from Landsat 8 Vegetation Continuous Fields: A New Approach to Normalize NDVI for the Estimation of Biophysical Parameters in Mediterranean Pear Orchards. Remote Sensing, 11(23), 2777.
57. Schimmelpfennig, D. (2016). Farm Profits and Adoption of Precision Agriculture. Economic Research Report, (217), 1-40. http://dx.doi.org/10.22004/ag.econ.249773 Schimmelpfennig, D., & Ebel, R. (2011). On the doorstep of the information age: Recent adoption of precision agriculture. USDA-ERS Economic Information Bulletin, 80. https://ssrn.com/abstract=2692052
58. Babber, J.; Malik, P.; Mittal, V.; Purohit, K.C. Analyzing Supervised Learning Algorithms for Crop Prediction and Soil Quality. In Proceedings of the 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 29–31 March 2022; pp. 969–973. [**Google Scholar**]
59. Ishak, M.; Rahaman, M.S.; Mahmud, T. FarmEasy: An Intelligent Platform to Empower Crops Prediction and Crops Marketing. In Proceedings of the 13th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 17–20 April 2021; pp. 224–229. [**Google Scholar**]
60. Patel, K.; Patel, H.B. A Comparative Analysis of Supervised Machine Learning Algorithm for Agriculture Crop Prediction. In Proceedings of the Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India, 15–17 September 2022; pp. 1–5. [**Google Scholar**]
61. Memon, R.; Memon, M.; Malioto, N.; Raza, M.O. Identification of growth stages of crops using mobile phone images and machine learning. In Proceedings of the International Conference on Computing, Electronic and Electrical Engineering (ICE Cube), Quetta, Pakistan, 26–27 October 2021; pp. 1–6. [**Google Scholar**]
62. Chandraprabha, M.; Dhanaraj, R.K. Soil Based Prediction for Crop Yield using Predictive Analytics. In Proceedings of the 3rd International Conference on Advances in Computing,

Communication Control and Networking (ICAC3N), Greater Noida, India, 17–18 December 2021; pp. 265–270. [**Google Scholar**]

63. Ray, R.K.; Das, S.K.; Chakravarty, S. Smart Crop Recommender System-A Machine Learning Approach. In Proceedings of the 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 27–28 January 2022; pp. 494–499. [**Google Scholar**]
64. Priyadharshini, K.; Prabavathi, R.; Devi, V.B.; Subha, P.; Saranya, S.M.; Kiruthika, K. An Enhanced Approach for Crop Yield Prediction System Using Linear Support Vector Machine Model. In Proceedings of the International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 10–11 March 2022; pp. 1–5. [**Google Scholar**]
65. Malathy, S.; Vanitha, C.N.; Kotteswari, S.; Mohankkanth, E. Rainfall Prediction for Enhancing Crop-Yield based on Machine Learning Techniques. In Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 9–11 May 2022; pp. 437–442. [**Google Scholar**]
66. Chowdary, V.T.; Robinson Joel, M.; Ebenezer, V.; Edwin, B.; Thanka, R.; Jeyaraj, A. A Novel Approach for Effective Crop Production Using Machine Learning. In Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 16–18 March 2022; pp. 1143–1147. [**Google Scholar**]
67. Yamparla, R.; Shaik, H.S.; Guntaka, N.; Marri, P.; Nallamothu, S. Crop Yield Prediction using Random Forest Algorithm. In Proceedings of the 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 22–24 June 2022; pp. 1538–1543. [**Google Scholar**]
68. Apeksha, R.G.; Swati, S.S. A brief study on the prediction of crop disease using machine learning approaches. In Proceedings of the 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), Nagpur, India, 18–19 June 2021; pp. 1–6. [**Google Scholar**]
69. Kumar, R.; Shukla, N.; Princee. Plant Disease Detection and Crop Recommendation Using CNN and Machine Learning. In Proceedings of the International Mobile and Embedded Technology Conference (MECON), Noida, India, 10–11 March 2022; pp. 168–172. [**Google Scholar**]
70. Bhosale, S.V.; Thombare, R.A.; Dhemey, P.G.; Chaudhari, A.N. Crop Yield Prediction Using Data Analytics and Hybrid Approach. In Proceedings of the Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 16–18 August 2018. [**Google Scholar**]
71. Alwis, S.D.; Hou, Z.; Zhang, Y.; Na, M.H.; Ofoghi, B.; Sajjanhar, A. A survey on smart farming data, applications and techniques. *Comput. Ind.* **2022**, *138*, 103624. [**Google Scholar**] [**CrossRef**]
72. Lyu, Y.; Li, J.; Hou, R.; Zhang, Y.; Hang, S.; Zhu, W.; Zhu, H.; Ouyang, Z. Precision Feeding in Ecological Pig-Raising Systems with Maize Silage. *Animals* **2022**, *12*, 11. [**Google Scholar**] [**CrossRef**] [**PubMed**]
73. Ghobadi, F.; Kang, D. Application of Machine Learning in Water Resources Management: A Systematic Literature Review. *Water* **2023**, *15*, 4. [**Google Scholar**] [**CrossRef**]
74. Padarian, J.; Minasny, B.; McBratney, A.B. Machine learning and soil sciences: A review aided by machine learning tools. *SOIL* **2020**, *6*, 35–52. [**Google Scholar**] [**CrossRef**]

75. Ramos, P.J.; Prieto, F.A.; Montoya, E.C.; Oliveros, C.E. Automatic fruit count on coffee branches using computer vision. *Comput. Electron. Agric.* **2017**, *137*, 9–22. [**Google Scholar**] [**CrossRef**]

76. Sengupta, S.; Lee, W.S. Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions. *Biosyst. Eng.* **2014**, *117*, 51–61. [**Google Scholar**] [**CrossRef**]

77. Su, Y.; Xu, H.; Yan, L. Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi J. Biol. Sci.* **2017**, *24*, 537–547. [**Google Scholar**] [**CrossRef**]

78. Sullivan, D. G. (2010). Hyperspectral Imaging with a Helicopter Platform: Early Detection of Plant Stress. In P. Thenkabail, J. G. Lyon, & A. Huete(Eds.), Hyperspectral Remote Sensing of Vegetation (pp. 541-562). CRC Press.

79. Taghvaeian, S. (2015). Remote Sensing of Evapotranspiration: Theories, Models, and Applications. In Water Conservation in the 21st Century (pp. 71-98). Springer.

80. Tey, Y. S., & Brindal, M. (2012). Factors influencing the adoption of precision agricultural technologies: a review for policy implications. Precision agriculture, 13(6), 713-730. https://doi.org/10.1007/s11119-012-9273-6

81. Thenkabail, P. S. (2015). Remote Sensing Handbook - Three Volume Set: Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press. https://scholar.google.com/scholar_lookup?title =Remote+Sensing+HandbookThree+Volume+Set&author=Thenkabail,+P.&p ublication_year=2018

82. Vasques, G. M., Grunwald, S., & Sickman, J. O. (2008). Comparison of Multivariate Methods for Inferential Modeling of Soil Carbon Using Visible/Near-Infrared Spectra. Geoderma, 146(1-2), 14-25. https://doi.org/10.1016/j.geoderma.2008.04.007

83. Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., & Veroustraete, F. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS Journal of Photogrammetry and Remote Sensing, 108, 273-290. https://doi.org/10.1016/j.isprsjprs.2015.05.005

84. Whelan, B., & Taylor, J. (2013). Precision Agriculture for Grain Production Systems. CSIRO Publishing. https://doi.org/10.1071/9780643107489

85. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming – A review. Agricultural Systems, 153, 69-80. https://doi.org/10.1016/j.agsy.2017.01.023

86. Zhang, N., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. Precision agriculture, 13(6), 693-712. https://doi.org/10.1007/s11119- 012-9274-5

87. Zhang, Z., Jayachandran, K., & Grunwald, S. (2019). Soil Information Derived from Visible/Infrared and Passive Microwave Remote Sensing: Status and Perspectives. Earth Science Reviews, 190, 420-436.

88. Rahman, M. M., & Robson, A. (2020). Integrating landsat-8 and sentinel-2 time series data for yield prediction of sugarcane crops at the block level. *Remote Sensing*, *12*(8), 1313.

89. Zhang, H., Zhang, Y., Liu, K., Lan, S., Gao, T., & Li, M. (2023). Winter wheat yield prediction using integrated Landsat 8 and Sentinel-2 vegetation index time-series data and machine learning algorithms. *Computers and Electronics in Agriculture*, *213*, 108250.

90. Řezník, T., Pavelka, T., Herman, L., Lukas, V., Širůček, P., Leitgeb, Š., & Leitner, F. (2020). Prediction of yield productivity zones from Landsat 8 and Sentinel-2A/B and their evaluation using farm machinery measurements. *Remote Sensing*, *12*(12), 1917.

91. Wolanin, A., Camps-Valls, G., Gómez-Chova, L., Mateo-García, G., van der Tol, C., Zhang, Y., & Guanter, L. (2019). Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. *Remote sensing of environment*, *225*, 441-457.

92. Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., & Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote sensing of environment*, *269*, 112831.

93. Ghayour, L., Neshat, A., Paryani, S., Shahabi, H., Shirzadi, A., Chen, W., ... & Ahmad, A. (2021). Performance evaluation of sentinel-2 and landsat 8 OLI data for land cover/use classification using a comparison between machine learning algorithms. *Remote Sensing*, *13*(7), 1349.

94. Mishra, B., & Shahi, T. B. (2021). Deep learning-based framework for spatiotemporal data fusion: an instance of landsat 8 and sentinel 2 NDVI. *Journal of Applied Remote Sensing*, *15*(3), 034520-034520.

95. Saadi, T. D. T., & Wijayanto, A. W. (2021). Machine learning applied to sentinel-2 and landsat-8 multispectral and medium-resolution satellite imagery for the detection of rice production areas in Nganjuk, East Java, Indonesia. *International Journal of Remote Sensing and Earth Sciences*, *18*(1), 19-32.

96. Skakun, S., Vermote, E., Roger, J. C., & Franch, B. (2017). Combined use of Landsat-8 and Sentinel-2A images for winter crop mapping and winter wheat yield assessment at regional scale. *AIMS geosciences*, *3*(2), 163.

97. Chakhar, A., Ortega-Terol, D., Hernández-López, D., Ballesteros, R., Ortega, J. F., & Moreno, M. A. (2020). Assessing the accuracy of multiple classification algorithms for crop classification using Landsat-8 and Sentinel-2 data. *Remote sensing*, *12*(11), 1735.

98. Dhillon, M. S., Kübert-Flock, C., Dahms, T., Rummler, T., Arnault, J., Steffan-Dewenter, I., & Ullmann, T. (2023). Evaluation of MODIS, Landsat 8 and Sentinel-2 data for accurate crop yield predictions: A case study using STARFM NDVI in Bavaria, Germany. *Remote Sensing*, *15*(7), 1830.

99. Abebe, G., Tadesse, T., & Gessesse, B. (2022). Combined use of Landsat 8 and Sentinel 2A imagery for improved sugarcane yield estimation in Wonji-Shoa, Ethiopia. *Journal of the Indian Society of Remote Sensing*, *50*(1), 143-157.

100. Chen, J., & Zhang, Z. (2023). An improved fusion of Landsat-7/8, Sentinel-2, and Sentinel-1 data for monitoring alfalfa: Implications for crop remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *124*, 103533.