AN OVERVIEW OF DATA ANALYTICS IN SMART AGRICULTURE

Chapter · December 2023

5 authors, including:

Pankaj Das
Indian Agricultural Statistics Research Institute
75 PUBLICATIONS **108** CITATIONS

Rahul Banerjee
Indian Agricultural Statistics Research Institute
48 PUBLICATIONS **43** CITATIONS

Bharti Juneja
Indian Agricultural Statistics Research Institute
31 PUBLICATIONS **102** CITATIONS

Abhilasha Rangi
Central Silk Board
8 PUBLICATIONS **150** CITATIONS

AN OVERVIEW OF DATA ANALYTICS IN SMART AGRICULTURE

Pankaj Das[1], Rahul Banerjee*[1], Bharti[1], Abhilasha Rangi[2] and Nitin Varshney[3]

[1]ICAR-Indian Agricultural Statistics Research Institute,

Library Avenue, New Delhi-110 012, India

[2]Central Silk Technological Research Institute, Central Silk Board, Bengaluru-560 068, India

[3]Navsari Agricultural University, Dandi Road, Erugam, Navsari-396 450, India

*Corresponding author E-mail: rahuliasri@gmail.com

Abstract:

Data analytics is becoming immensely important in the field of smart agriculture to improve farming efficiency, provide accurate real time estimates and thereby increase yields. The application of modern technologies such as sensors, internet of things (IoT) devices, and other technologies like unmanned aerial vehicle (UAV), blockchain methodology, etc. generate large amounts of data, which needa to be analysed accurately to provide scientific insights about crop growth, soil health, weather patterns, and other factors affecting agricultural production. With this data, farmers can confidently determine the optimal timing for planting, fertilizing, and harvesting their crops, and identify the specific areas of their land that need focused attention. Data analytics can also be used to optimize resource allocation, such as water usage, to reduce waste and minimize environmental impact. Overall, the use of data analytics in smart agriculture has the potential to revolutionize the way we produce food, making it more sustainable, efficient, and resilient in the face of climate change and other challenges. This chapter will provide an overview of data analytics in smart agriculture, focusing on how data can be collected, analysed, and used to improve crop yields, reduce costs, and increase sustainability. The challenges associated with data collection and analysis in agriculture including issues related to data quality, privacy, and security will be discussed. Some of the most promising applications of data analytics in smart agriculture, including precision agriculture, predictive analytics and remote monitoring will also be elucidated. Finally, the chapter will conclude with discussions on some of the key opportunities and challenges faced in the adoption of data analytics in smart agriculture and with an outline of a roadmap for future research in the area.

Keywords: Data analytics, Smart agriculture, Precision agriculture, Internet of things (IoT), Unmanned Aerial Vehicle (UAV)

Introduction:

Smart agriculture refers to the use of technology and data analytics to improve crop yield, reduce input costs, and increase profitability. Data analytics plays a crucial role in smart agriculture as it helps farmers to make informed decisions based on real-time data collected from sensors, drones, satellites, and other sources. Data analytics can be used to monitor crop health, predict weather patterns, optimize irrigation systems, and identify pest and disease outbreaks. With the help of data analytics, farmers can also optimize fertilizer and pesticide applications, reduce energy consumption, and minimize waste. Lastly, these practices also aim to ensure agricultural practices are sustainable and environmentally friendly. This is achieved by reducing the use of chemicals, optimizing the use of water and other resources, and limiting pollution (Agbona *et al.,* 2022). These measures will ensure that we are sustainable and able to continue to produce food in the future. Finally, the measures we take now will determine our long-term food security. Therefore, it is our responsibility to ensure the survival of our future generations by taking action now to secure our food supplies. Therefore, in order to secure our food supply, we must examine agricultural practices, increase efficiency, and promote sustainable food production. In this regard, governments should support research and development in agriculture, as well as encourage farmers to adopt new technologies and practices (Chergui and Kechadi, 2022).

Additionally, they should also invest in infrastructure, such as modern storage facilities and means of transportation, to ensure the availability of food throughout the year. The prevalence of data and data analytics in agriculture presents an exciting opportunity for farmers to increase efficiency and productivity while reducing waste and environmental impact. Inefficient farming practices are often the result of a lack of data and insights, leading to ill-informed decision-making and suboptimal outcomes. However, not all farmers have access to the same resources or training necessary to make the most of these opportunities. In order to close this gap, many universities are now offering courses in data analytics and smart agriculture, providing students with the skills and knowledge they need to drive innovation in the industry. These programs enable students to gain a comprehensive understanding of the technology and tools available, as well as develop the skills they need to apply these tools to real-world situations. By learning how data analytics can improve decision-making and increase efficiency, students are helping to drive the industry forward and improve the lives of farmers and consumers alike (Elijah *et al.*, 2018).

Smart agriculture:

Smart agriculture refers to the use of advanced technologies, such as sensors, drones, and big data analytics, to improve crop yield, reduce input costs, and minimize environmental impacts. Smart farming refers to the use of information and data management technologies in agriculture. These technologies, including software, sensors, and robotics, increase productivity and efficiency of agricultural systems. This method offers farmers tools and tactics to increase yields and the sustainability of agricultural output. Smart agricultural technology, which is spreading from other sectors (including autonomous driving), is at the core of the innovation. A farm followed the major procedures *viz*. data collection, diagnostics, decision making and actions to become a smart farm (in Figure 1) (Goel *et al.,* 2021). At the same time, people need to understand the distinction between precision agriculture and smart farming. While the first strategy makes use of cutting-edge instruments to make farming a more predictable and effective process, the technologies used in the second approach are concentrated on getting the most precise measurements.

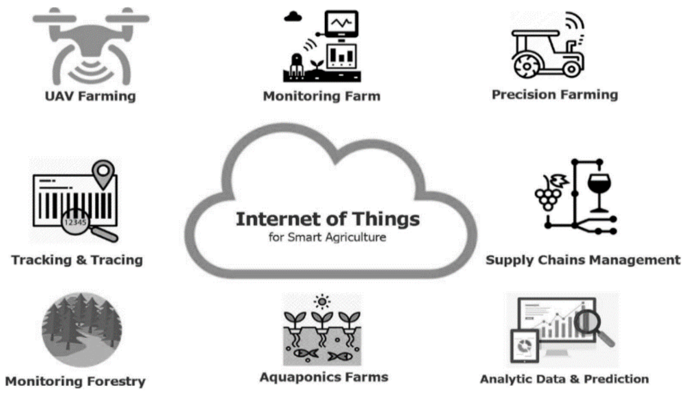

Figure 1: **An illustration of IoT applications for smart agriculture (Source: Quy *et al.*, 2022)**

1. Smart Farming Technologies:

The phrase "smart farming" refers to a variety of techniques and technology that are used to optimise agricultural processes. The following are some of the most practical and efficient smart farming tools:

- **Machine learning:** Data-driven self-adaptive technologies possess the capability to anticipate or predict variations in climate conditions, soil and water characteristics, carbon levels, the spread of diseases and pests, and other relevant factors.
- **Smart farming sensors:** Farmers can track even the smallest changes in the environment and their farms in real time through sensitive sensors.
- **Drones and satellites with cameras:**
- With the aid of these technologies, farmers can generate frequently updated maps and remotely monitor their fields without the need for physical visits.
- **Big data**: Accurate forecasting, activity planning, and the creation of more effective business models are inconceivable without big data. Making long-term judgements and acting immediately are both made possible by smart farming and big data (Quy *et al.*, 2022).
- **Internet of Things (IoT):** With IoT, you may create a single system out of all the available tools and solutions. All hardware and software have the ability to share data and take precise actions based upon patterns.

2. Advantages of Smart Agriculture:

The advantages of smart agriculture can be summarised as follows:

Better crop health monitoring	**Lesser impact on environment**	**Ensure food security in case of climate change**	**Reduce operating expenses while increasing yields**
• Utilising cutting-edge farming tools, farmers may detect crop diseases and other issues earlier.	• Systems for precision agriculture can lower carbon emissions and the usage of toxic chemicals	• More effective smart farming aids in maintaining production levels while adjusting to changing climatic conditions	• Increased yields (by 1.75%), 8% less water use, and lower energy expenditures are all achieved by smart farms.

In short, more extensive adoption of intelligent farm technologies is essential for more reasons than just increased profitability. The demands of the expanding population must also be met. The current technological advancements in the agricultural industry have the potential to drastically alter both small farms and large businesses' work processes (Ramya *et al.,* 2015). Modern farming technologies are being introduced at such a rapid rate that they are now accessible not just in developed regions but also in underdeveloped ones. The detail procedure of data analysis is summarized in Fig. 2.

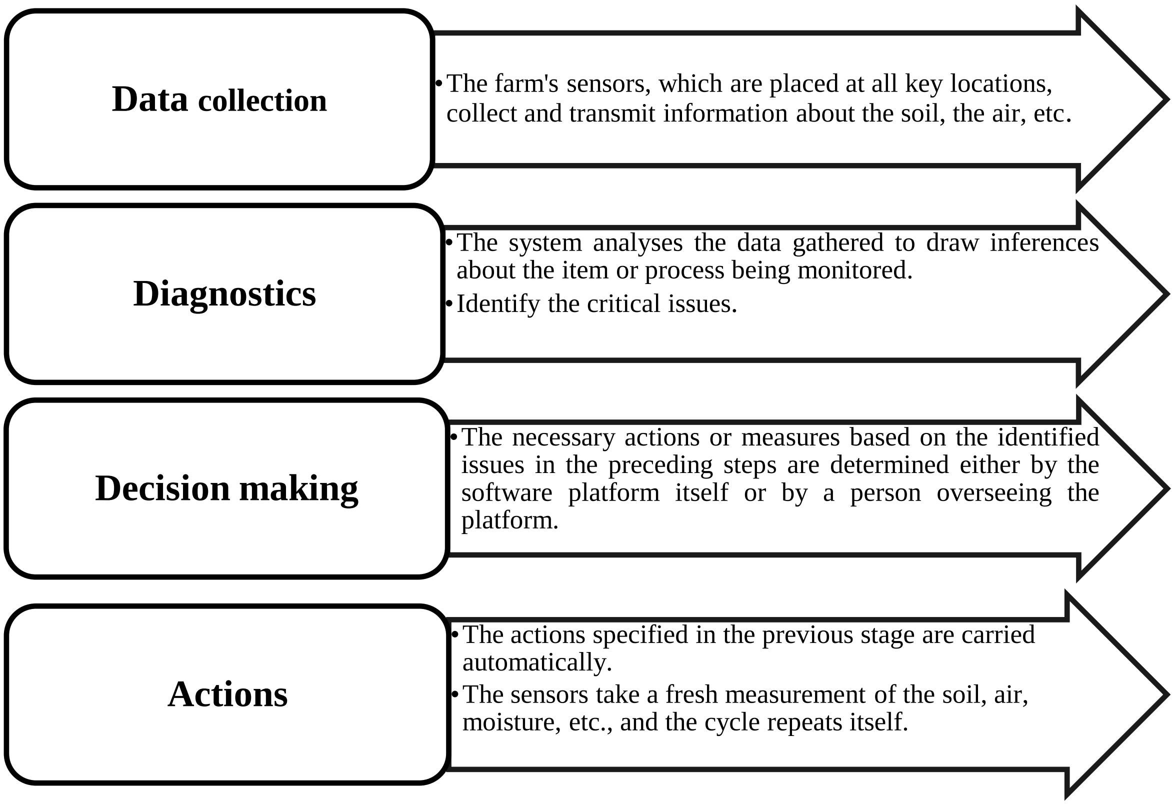

Figure 2: **The procedures used in smart agriculture on a farm**

Some examples of smart farming applications in agriculture:

Farmers have already started utilising a variety of smart agricultural technology, and their results have been amazing like

- **Livestock management and monitoring:** This technology, also known as Precision Cattle Farming (PLF), tracks multiple traits of cattle using IoT sensors and predictive analytics software. These traits are Behaviour, Feeding patterns and Health & well-being. The operation makes use of smart agricultural machinery that includes cameras, sensors, and software to monitor animal health, improve feeding regimens, and carry out real-time monitoring.
- **Smart crop management:** Farmers have typically visited the fields to assess the condition of their crops. However, this strategy is simply unworkable for major agricultural operations if you have to spend hours each day inspecting your field. IoT sensors and drones are used to monitor different factors related to crop health and soil health. AI-powered software is the innovation that enables these processes. Drones

are guided by GPS, image recognition, and numerous IoT devices in agricultural applications so they can carry out their tasks (Sinha and Dhanalakshmi, 2022).

- **Greenhouse monitoring system:** A greenhouse with a monitoring system, sometimes known as a "smart greenhouse," automatically regulates environmental conditions. Sensors, Wireless connectivity (infrastructure + cloud), and User software (a desktop and/or mobile app) are the key elements of the system. A smart greenhouse maintains an ideal micro climate and manages fertilization and irrigation. Besides this, it provides data on potential plant diseases.
- **Smart farm management software:** The goal of smart farming software is to assist with numerous facets of farm management. In essence, it serves as a hub for historical and current data and information about the facility that farmers use to organise, track, and evaluate their operations. The software may include information like weather records, animal monitoring data, farming equipment condition status, ideal and historical planting schedule *etc.* Together, this data enables farmer to handle everything with a single app. They are capable of managing not just the farm but also the finances and human resources (Mohapatra and Rath, 2022; Mohapatra *et al.*, 2023).
- **Autonomous ground vehicles:** The position and speed of the vehicle are programmed by AI and location tracking software with map data. Therefore, all that is required is GPS and equipment for managing machinery (cameras and sensors). The technology for autonomous driving is now available in tractors and other agricultural equipment. It has enabled autonomous farming and decreased the demand for human labour.

Overview of data analytics:

Data analytics is the process of gathering data from various sources, securitizing the raw data, analyzing it, and drawing inferences that help in decision-making. The data is collected based on the hypothesis and objective of the study i.e. experimental data, observational data, reference data etc. Experimental data is collected by researchers through proper planning and conducting an experiment based on the objective of study. Generally, there are two type of experiments, one is absolute experiment in which researcher is interested to know the value of a particular characteristics and another is comparative experiments in which researcher is interested to compare the characteristics under study. Comparison between the yields of two varieties of a crop in an experiment is an example of comparative experiments. However, observational data is collected by making observations. In other words, everything that can be heard or seen is noted. The majority of the data in surveys is almost certainly observational. The observational data can be obtained through remote sensing and conventional data collection methods like simple random sampling, stratified random sampling, multistage sampling, judgmental sampling etc. In sample surveys, data can be collected through various methods depending on the cost of survey and precision required i.e. direct personal interview, telephonic surveys, mail based surveys, questionnaire and schedule based surveys etc. While collecting data through surveys, a proper plan of work with supervision is crucial for getting quality data. Any carelessness at any stage will aid in more sampling error.

In conventional data collection methods, it is impractical to thoroughly survey the vast area for a number of different variables, whereas remote sensing (RS) technologies are perfectly capable of doing so. Remote sensing data has drawn the attention of researchers over the years for the acquisition of information about an object without making physical contact with the object. Precision agriculture and remote sensing are thus developing day by day as they are more efficient and eco-friendly. Advancements in remote sensing data acquisition, processing, and interpretation of ground-based, airborne, and satellite observations has made it feasible to integrate remote sensing technologies with precision crop management systems (Waheed *et al.*, 2006). There are numerous types of satellite data depending upon spatial resolution (ranging from meter to kilometres), technique (active/passive, radiometer/scatterometer), spectral range, and viewing geometry (Oza *et al.*, 2008). Once data is collected, it is cleaned to ensure that there is no mistake and redundancy in data. It will help to fix the problem before data is processed for data analysis. The basic steps in cleaning of data includes removal of duplicate observations, fix structural errors, outlier detection, missing data handling and validation. However, the steps in this process cannot be outlined in a single, unambiguous manner because the procedures will differ from dataset to dataset (Fig. 3).

Figure 3: **Flowchart of Data Analytics**

The cleaned data is then subjected to appropriate statistical analysis to draw valid conclusion about the population. The broad objective of data analysis is to look into variation in the population and the factors that contribute to it. The analysis and simplification of remote sensing data frequently involves the use of specialized methods like geostatistics, image analysis, artificial intelligence (Liaghat and Balasundram, 2010). However, Remote Sensing cannot capture all types of agricultural information. Furthermore, the information gleaned from remote sensing data is more insightful when combined with ground data. The final step in data analysis is to interpret the findings to determine how well the data answered the research question and what recommendations are made in light of the data in decision-making.

Data analytics in smart agriculture:

Data analytics plays a crucial role in smart agriculture, which refers to the use of technology to enhance and optimize agricultural practices. By leveraging data analytics tools and techniques, farmers and agribusinesses can gain valuable insights into their operations and make data-driven decisions to improve yields, reduce costs, and increase efficiency. Here are some ways data analytics can be applied in smart agriculture:

- **Predictive analytics:** Farmers can use data analytics to forecast crop yields, weather patterns, and pest infestations. By analysing historical data and real-time sensor data, farmers can make more accurate predictions and take pre-emptive actions to prevent crop loss.
- **Precision farming:** Precision farming involves using data analytics to optimize crop production by tailoring farming practices to specific areas of the farm. By analysing soil and moisture data, farmers can adjust irrigation, fertilizer application, and seed planting to maximize crop yields and minimize waste.
- **Livestock management:** Livestock farmers can use data analytics to monitor animal health and behaviour, as well as track feed consumption and growth rates. This information can help farmers make more informed decisions about breeding, feeding, and veterinary care.
- **Supply chain management:** Data analytics can be used to optimize the supply chain in agriculture, from seed production to distribution. By analysing data on demand, transportation, and inventory levels, agribusinesses can minimize waste and ensure timely delivery of products.

Overall, data analytics can help farmers and agribusinesses make more informed decisions, optimize production, and reduce costs. As technology continues to advance, the role of data analytics in smart agriculture is likely to become even more important. Data analytics is a rapidly evolving field that encompasses a wide range of methods and techniques for processing, analysing, and interpreting data. In recent years, there have been several trends that have emerged in data analytics, each of which has the potential to transform the way organizations approach data-driven decision-making. One major trend in data analytics is the increasing use of machine learning and artificial intelligence (AI) techniques. These methods enable organizations to analyse vast amounts of data and extract insights that might otherwise be difficult or impossible to identify. Machine learning algorithms can be trained to recognize patterns and anomalies in data, making it possible to automate many aspects of data analysis and decision-making. Another trend in data analytics is the growing importance of data visualization. Data visualization tools enable organizations to present complex data sets in a more intuitive and easily digestible format, making it easier for decision-makers to identify trends and patterns. Visualization tools can also be used to explore data in new ways, uncovering insights that might have been missed using traditional methods. Another emerging trend in data analytics is the use of big data technologies. Big data technologies are designed to handle large volumes of data, typically measured in petabytes or even exabytes. These technologies include distributed file systems, such as Hadoop and Spark, as well as data processing and storage tools such as NoSQL databases. Finally, there is an increasing focus on the ethical use of data analytics.

As organizations collect and analyse more data, there are growing concerns about privacy, security, and the potential misuse of data. As a result, many organizations are developing ethical frameworks and guidelines for the use of data analytics, with the goal of ensuring that data is used in a responsible and transparent manner. Overall, these trends are transforming the field of data analytics, enabling organizations to make more informed and data-driven decisions than ever before. As data analytics continues to evolve, it is likely that we will see further innovations and developments that will further enhance our ability to extract insights from data.

The escalation in global population has heightened the imperative for the adoption of intelligent agricultural practices. This makes food security a primary concern for most countries, along with the loss of natural resources, the lack of arable land, and the growth in weather unpredictability. The Internet of Things (IoT) and data analytics (DA) are employed in the agriculture business to increase operational effectiveness and production. The usage of IoT and DA is replacing the use of wireless sensor networks (WSN) as one of the main drivers of smart agriculture. The Internet of Things (IoT) merges a number of already-existing technologies, including cloud computing, radio frequency identification, middleware systems, and end-user applications. Elijah *et al.* (2018) have highlighted several benefits and challenges of IoT. They have presented the IoT ecosystem and depicted the of IoT and DA enabling smart agriculture. They have also highlighted future trends and the opportunities of IoT and DA. The most current developments are covered in-depth in a report by Sinha and Dhanalakshmi (2022). Their survey's objective is to assist prospective researchers in identifying pertinent IoT issues and choosing the best solutions depending on the application needs. They have also emphasised the importance of data analytics and the Internet of Things.

The three primary pillars of smart agriculture consist of research, innovation, and space technologies, which are considered fundamental for national development. Space technologies play a crucial role in enhancing soil quality, reducing water waste in irrigation, and providing farmers with valuable agricultural knowledge. Geospatial data collected from various sources, including terrestrial, aquatic, and aerial sensors, satellites, and surveillance tools, is extensively analyzed and utilized for smart farming and crop protection. Innovations in technology include the integration of drones in agriculture, precise gene editing in plants, epigenetics, big data, the Internet of Things (IoT), efficient utilization of renewable energy sources like smart wind and solar energy, as well as the application of artificial intelligence in robotics and large-scale desalination technology. Many of these innovations are already being implemented in industrialized countries. However, the adoption of digital farming in rural areas will be particularly beneficial for emerging economies, as agriculture plays a significant role in their development. With an estimated 85% of the world's population projected to reside in developing nations by 2030, these countries require data-driven technological advancements to boost their GDP and ensure food security for their populations. Researchers, such as Goel *et al.* (2021) and Ramya *et al.* (2015), have explored the use of data and analytics to predict and mitigate the impacts of extreme weather events on global finance and address the economic dimensions of climate change, including the challenge of global food insecurity caused by climate change. They have also developed improved automated prediction methods and scalable meteorological data analysis and forecasting systems using frameworks such as Hadoop.

Real-time estimates:

Real-time estimates refer to estimates or predictions that are updated and provided in real-time, meaning they are available instantly or with very little delay. Real-time estimates are useful in many contexts, such as in finance, weather forecasting, traffic management, and industrial processes. Real-time estimates are made possible through the use of advanced technologies such as sensors, machine learning algorithms, and data analytics. These technologies enable the collection and analysis of large amounts of data in real-time, allowing for quick and accurate predictions. Real-time estimates are particularly valuable in situations where timely decision-making is critical, such as in emergency response scenarios or in financial trading. They can also help businesses and organizations optimize their operations, reduce costs, and improve efficiency. Overall, real-time estimates are an important tool for decision-makers in a wide range of industries, enabling them to make informed decisions based on up-to-date information. Real-time estimates in agriculture can refer to the use of various technologies and techniques to gather and analyse data in real-time to improve farming efficiency and productivity. Some examples of real-time estimates in agriculture include:

- **Weather and climate monitoring:** Real-time weather and climate data can be used to help farmers make decisions about planting, irrigation, and harvesting. This can be done through various sensors and weather stations that collect data and provide real-time updates to farmers.
- **Soil monitoring:** Real-time soil monitoring can help farmers understand soil health and nutrient levels, allowing them to make informed decisions about fertilization and other farming practices.
- **Crop monitoring:** Crop monitoring can involve the use of drones, satellites, or other sensors to provide real-time information about crop health, growth, and yield. This can help farmers identify problems early on and take corrective action before it's too late.
- **Livestock monitoring:** Real-time monitoring of livestock can help farmers keep track of animal health and behaviour, allowing them to identify potential problems and take action before they become more serious.

Overall, real-time estimates in agriculture can help farmers make more informed decisions, reduce waste, and increase productivity, ultimately leading to more sustainable and profitable farming practices.

1. Generation of real time crop yield estimates:

Real-time crop yield estimates in agriculture can be achieved through the use of various technologies such as satellite imagery, drones, and sensors. These technologies can collect data on various factors that affect crop growth and development, such as soil moisture, temperature, and nutrient levels. This data can be used to estimate crop yields in real-time. One approach to real-time crop yield estimation is to use machine learning algorithms to analysed the data collected by these technologies. Machine learning algorithms can use historical data to make predictions about future crop yields based on the current environmental conditions. For example, if the soil moisture and nutrient levels are optimal, and the temperature is within a certain range, the algorithm may predict a higher crop yield. Another approach is to use sensors and other devices to directly measure the growth and development of crops. For example, a sensor may be used to measure the height of corn plants, which can be used to estimate the yield of the crop. This data can be collected and analysed in real-time to provide an estimate of the crop yield. Real-time crop yield estimation has the potential to revolutionize agriculture by allowing farmers to make more informed decisions about planting, fertilizing, and harvesting crops. By providing accurate and timely information on crop yields, farmers can optimize their production and reduce waste, ultimately leading to higher profits and more sustainable farming practices. A hypothetical work flow process involved in crop yield monitoring process in smart agriculture has been elucidated in Fig. 4.

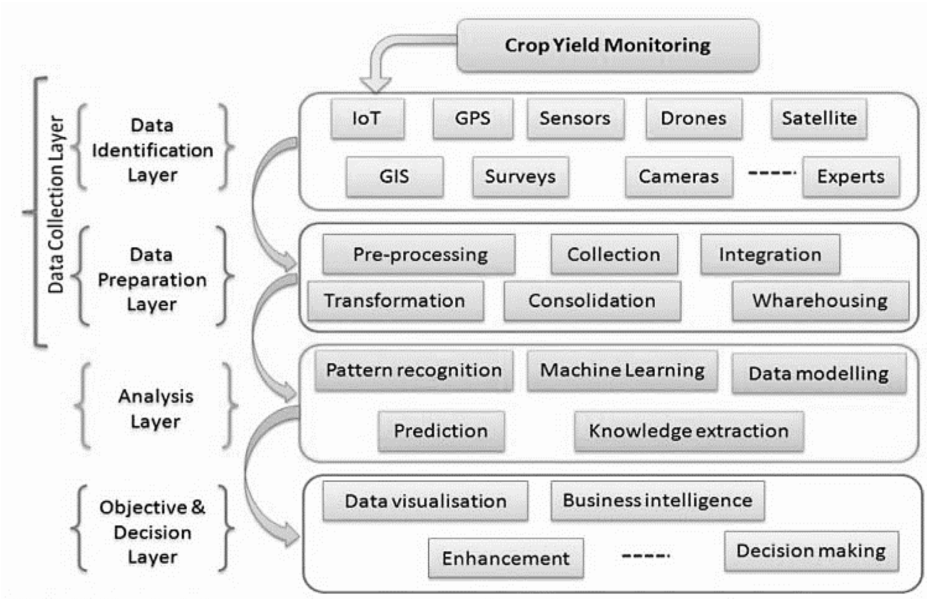

Figure 4: Process Flow of Crop Yield Monitoring in Smart Agriculture (Source: Chergui and Kechadi, 2022)

2. Real time soil monitoring:

Real-time soil monitoring is a process of collecting and analysing soil data in real-time using various technologies and sensors. Real-time soil monitoring systems can provide farmers with up-to-the-minute information about soil health, moisture levels, nutrient levels, and other soil characteristics, allowing them to make informed decisions about fertilization, irrigation, and other farming practices.

Some examples of technologies used for real-time soil monitoring include:

- **Soil sensors**: Soil sensors can be placed in the field to measure soil moisture, temperature, and other characteristics. This information can be transmitted in real-time to a central system for analysis.
- **Wireless networks:** Wireless networks can be used to transmit data from soil sensors and other monitoring systems to a central system for analysis.
- **Satellite imagery:** Satellite imagery can be used to monitor changes in soil moisture levels and other characteristics over time, providing farmers with valuable insights into soil health and fertility.
- **Automated irrigation systems:** Automated irrigation systems can be programmed to deliver water based on real-time soil moisture data, ensuring that crops receive the optimal amount of water without wasting resources.

Overall, real-time soil monitoring can help farmers optimize their use of resources, improve crop yields, and reduce environmental impact. By providing farmers with real-time data and insights, real-time soil monitoring can help farmers make more informed decisions about fertilization, irrigation, and other practices, leading to more sustainable and profitable farming practices.

3. Real time livestock monitoring:

Real-time livestock monitoring is a process of collecting and analyzing data on livestock behavior and health in real-time using various technologies and sensors. Real-time monitoring can help farmers identify potential health problems, track animal behavior, and improve overall livestock management.

Some examples of technologies used for real-time livestock monitoring include:

- **Sensors:** Sensors can be attached to livestock to monitor vital signs such as heart rate, body temperature, and activity levels. This information can be transmitted wirelessly to a central system for analysis.
- **RFID tags:** RFID tags can be used to track livestock movements and monitor feed intake. This information can be used to identify potential health problems and improve feeding and management practices.
- **Cameras:** Cameras can be used to monitor livestock behaviour and detect signs of distress or illness. This information can be transmitted to a central system for analysis.

- **Wearable technology:** Wearable technology such as GPS trackers can be used to monitor livestock movements and detect unusual behaviour or signs of stress.

Real-time livestock monitoring can help farmers detect potential health problems early, improve feeding and management practices, and optimize livestock production. By providing farmers with real-time data and insights, real-time livestock monitoring can help farmers make more informed decisions about animal health and welfare, leading to more sustainable and profitable farming practices.

4. Real time weather monitoring:

Real-time weather and climate monitoring is a process of collecting and analysing weather and climate data in real-time using various technologies and sensors. Real-time monitoring can help farmers make informed decisions about planting, irrigation, and harvesting, as well as help scientists study weather and climate patterns over time.

Some examples of technologies used for real-time weather and climate monitoring include:

- **Weather stations:** Weather stations can be used to collect data on temperature, humidity, wind speed, and other weather parameters. This information can be transmitted wirelessly to a central system for analysis.
- **Satellites:** Satellites can be used to monitor weather patterns over large areas, providing real-time data on cloud cover, precipitation, and other weather parameters.
- **Doppler radar:** Doppler radar can be used to detect and track severe weather events such as thunderstorms and tornadoes, allowing farmers and emergency responders to take timely action.
- **Climate models:** Climate models can be used to simulate and predict future weather patterns based on historical data and current trends.

Real-time weather and climate monitoring can help farmers make more informed decisions about planting, irrigation, and other practices, leading to more sustainable and efficient farming practices. By providing scientists with real-time data and insights, real-time weather and climate monitoring can also help improve our understanding of weather and climate patterns over time, leading to better predictions and preparation for future weather events

Challenges of data collection & analysis:

Smart Agriculture arose as a result of the democratization of digital devices, as well as advancements in artificial intelligence and data science. Smart agriculture pioneered innovative methods for increasing farm productivity and efficiency while respecting the environment. Recent and sophisticated digital technologies and data science have enabled the collection and analysis of huge amounts of agricultural datasets to assist farmers, agronomists, and professionals in understanding and making better judgments about farming activities.

1. Data collection:

Data collection is the process of gathering data for decision-making, strategic planning, research, and other purposes. It's a crucial part of data analytics and research projects. Effective data collection provides the information that's needed to answer questions, analyse performance or other outcomes, and predict future trends, actions, and scenarios. For research in science, Agricultural science, medicine, higher education and other fields, data collection is often a more specialized process, in which researchers create and implement measures to collect specific sets of data. In both the business and research contexts, though, the collected data must be accurate to ensure that analytics findings and research results are valid.

2. Data analysis:

Data analysis refers to the process of transforming, cleaning, and modelling raw data in order to extract useful insights and information from it. The goal of data analysis is to make sense of complex data sets, identify patterns and trends, and inform decision-making. Any organization may face a number of challenges in collecting consistent and quality data. To develop methods to improve data collection practices, it is necessary to first identify barriers to consistent data collection. Government departments, agencies and service providers with responsibility for data collection should consider these challenges and improvement opportunities as part of implementation planning.

3. Common challenges in data collection and analysis:

Some of the challenges often faced when collecting data include the following:

- **Poor data quality:** Raw data often contains errors, inconsistencies, and other issues that can impact its reliability. To address this, data profiling is commonly employed to identify and rectify problems in the acquired data. Without ensuring the quality and accuracy of the input data, the resulting analysis may not be trustworthy. This is particularly critical when the analysis is used to inform decision-making, as relying on flawed data can lead to serious negative consequences.
- **Collecting meaningful and real-time data:** It might be challenging to sift through all the data and find the insights that are most required. Decision-making can be significantly harmed by outdated data. Decision-makers can be sure that any decisions they make are based on complete and accurate information by using real-time reports and notifications.
- **Deciding what data to collect:** This is a crucial decision that must be made both for the initial collection of raw data and when users gather data for analytics apps. Unneeded data collection increases process time, expense, and complexity. The usefulness of a data set can be limited and analytics outcomes might be impacted by leaving out useful data.
- **Dealing with big data:** Big data settings frequently contain a mix of structured, unstructured, and semi-structured data in significant quantities. Because of this, the earliest stages of data collecting and processing are more difficult. Additionally, for particular

analytics applications, data scientists frequently need to filter collections of raw data stored in a data lake.

- **Low response:** Lack of responses or willing participants in research investigations calls into question the reliability of the information gathered.
- **Data volume:** It can be intimidating and challenging to manage a lot of data. To do this, it may be necessary to use specialised software and methods for handling, analysing, and interpreting data.
- **Data complexity:** Data can come in a variety of sorts, formats, and sources and be complicated and heterogeneous. Such data analysis may call for a variety of knowledge and abilities in data administration, statistical analysis, and machine learning.
- **Data privacy and security:** Data security and privacy are crucial considerations in data analysis. Access to sensitive information must be restricted to authorised individuals in order to maintain security.
- **Interpretation and communication:** Data analysis alone won't help; stakeholders also need to be adequately informed of insights. This necessitates the capacity to recognise essential facts and convey them in a direct and succinct manner.
- **Time and resource constraints**: It can take a lot of time and resources to analyse data. Organisations may experience staffing or financial limitations that prevent them from doing thorough analysis.
- **Bias and ethics:** Biases and ethical issues may affect data analysis. It is critical to be aware of these problems and take action to address them, such as making sure diverse representation is included in data gathering and analysis.
- **Data quantity:** It can be difficult to gather enough information to draw conclusions that are statistically significant. Sometimes there might not be enough data available, or gathering the data might be expensive or complicated.
- **Confusion or anxiety:** Even if they are aware of the advantages of automation, users may experience anxiety or confusion when migrating from conventional data processing techniques. Nobody likes change, especially when the current method of doing things is convenient and familiar to them.

Applications of data analytics in smart agriculture:

Smart farming encompasses the utilization of advanced agricultural technology to enhance and sustain production. It encompasses a range of technologies, including high-resolution satellite data, Global Positioning Systems (GPS), Geographic Information Systems (GIS), field sensors, Artificial Intelligence (AI), and automated machinery. Additionally, emerging technologies such as Unmanned Aerial Vehicles (UAVs), the Internet of Things (IoT), and cloud computing are anticipated to contribute to this progress, introducing a fusion of robotics, computer science, and IoT in farming practices. Big data holds promise as a potential technology for assessing and managing farm-level decisions, enabling the identification and rectification of inefficient practices to boost farm productivity. To address the challenges posed by increasing food demand and climate change, policymakers and industry leaders are turning to technological advancements such as IoT, big data, analytics, and cloud computing for support. Data analytics can be applied to many aspects of smart agriculture to improve the efficiency, productivity, and sustainability of farming practices. Here are some applications of data analytics in smart agriculture

- **Crop management:** Data analytics can be used to monitor crop growth, identify potential issues such as nutrient deficiencies or disease outbreaks, and optimize irrigation and fertilization practices.
- **Precision agriculture:** Precision agriculture involves using data analytics to tailor farming practices to individual plants or small areas of land, based on data such as soil moisture, temperature, and nutrient levels. This can help optimize yields while minimizing waste.
- **Livestock management:** Data analytics can be used to monitor the health and behaviour of livestock, predict potential health issues, and optimize feeding and breeding practices.
- **Supply chain management:** Data analytics can be used to track and optimize the supply chain for agricultural products, from farm to table. This can help improve efficiency, reduce waste, and ensure product safety and quality.
- **Weather forecasting:** Accurate weather forecasting is essential for effective agricultural planning and management. Data analytics can be used to analyze weather data and provide more accurate and timely forecasts.
- **Soil analysis:** Data analytics can be used to analyze soil samples and provide insights into soil quality, nutrient levels, and potential contaminants. This can help farmers make more informed decisions about fertilization and land management practices.
- **Pest control:** Data analytics can be used to monitor pest populations and predict potential outbreaks, allowing farmers to take proactive measures to prevent crop damage.

Overall, data analytics has the potential to revolutionize agriculture by providing farmers with the tools they need to make informed decisions, optimize their practices, and improve their yields and profitability.

Opportunities and future prospects of data analytics in smart agriculture:

In the agriculture, the impact of data analytics has been so extensive that it is challenging to identify all of its implications and even more challenging to predict what changes this could bring. With the advancement in smart measurements and data analysis, it has become reliable to assess the trend in agricultural related data and to forecast production, demand/supply in agricultural systems. Incorporating data analytics is instrumental in estimating production levels, resource usage, and the environmental impact within agricultural systems. It also plays a pivotal role in gauging the disparity between agricultural production and consumption, a critical factor in making decisions to ensure food security. It gives real-time insights into production, demand and supply of agricultural commodities by integrating information and technologies. As a result, data science, remote sensing data, machine learning and big data analytics are becoming increasingly promising in agriculture. Remote sensing technology, which involves using satellites and other sensors to gather information about the Earth's surface, can provide valuable data on crop health and growth patterns. This data can be combined with machine learning algorithms to predict crop yields, detect nutrient deficiencies or disease outbreaks, and make more informed decisions about crop management. A representation of the Data management workflow is mentioned in Figure 5.

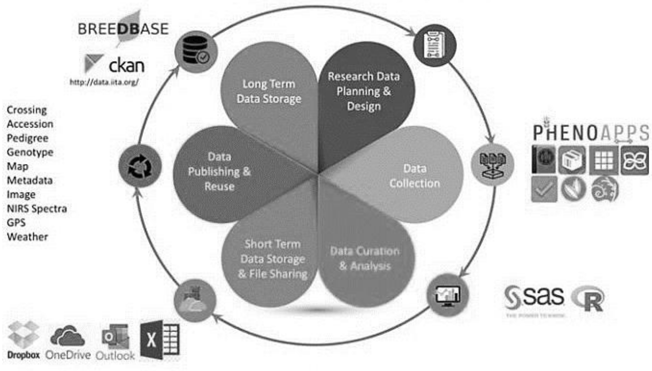

Figure 5: **Data management Workflow in Smart Agriculture**
(Source: Agbona *et al.*, 2022)

Conclusion:

Data analytics can help in increasing the production in the future by combining the time series data of previous and current years and can extract insights of data using appropriate algorithms. Data analytics also aids farmers in better comprehending environmental trends and consequently they can better plan for challenges and embrace opportunities without wasting resources. Before the advent of data analytics, farmers were unable to predict the early warning signs of failing crops; as a result, it was too late to take any action. Scientists can use data analytics to look for disease indicators in plants and predict future harvests to determine if crops are at risk from illness. Future agriculture will be heavily reliant on technologies like robotics, sensors, aerial photography, and GPS. Robotic systems, modern machinery and precision agriculture will make agriculture more effective, safe, and environmentally friendly.

References:

Agbona, A., Peteti, P., Teeken, B., Olaosebikan, O., Bello, A., Parkes, E., Mueller, I. R. L., Egesi, C., & Kulakow, P. (2022). Data Management in Multi-disciplinary African RTB Crop Breeding Programs. In *Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development* (pp. 85-103). Springer Publication. https://doi.org/10.1007/978-3-031-13276-6

Chergui, N., & Kechadi, M. (2022). Data analytics for crop management: a big data view. *Journal of Big Data*, *9*(123), 1-37. https://doi.org/10.1186/s40537-022-00668-2

Elijah, O., Rahman, T. A., Orikumhi, I., Leow, C. Y., Hindia, M. H. D. N. (2018). An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges. *IEEE Internet of Things Journal*, *5*(5), 3758-3773. https://doi.org/10.1109/JIOT.2018.2844296

Goel, R. K., Yadav, C. S., Vishnoi, S., & Rastogi, R. (2021). Smart agriculture-Urgent need of the day in developing countries. *Sustainable Computing: Informatics and Systems, 30*. https://doi.org/10.1016/j.suscom.2021.100512

Mohapatra, H., & Rath, A. K. (2022). IoE based framework for smart agriculture. *J Ambient Intell Human Comput*, *13*, 407–424. https://doi.org/10.1007/s12652-021-02908-4

Mohapatra, H., Dehury, M. K., Guru, A., & Rath, A. K. (2023). IoT-Enabled Zero Water Wastage Smart Garden. In *IoT Enabled Computer-Aided Systems for Smart Buildings* (pp. 407–424). EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-031-26685-0_4

Quy, V. K., Hau, N. V., Anh, D. V., Quy, N. M., Ban, N. T., Lanza, S., Randazzo, G., & Muzirafuti, A. (2022). IoT-Enabled Smart Agriculture: Architecture, Applications, and Challenges. *Applied Sciences 2022*, *12*(3396), 1-19. https://doi.org/10.3390/app12073396

Ramya, M. G., Balaji, C., & Girish, L. (2015). Environment Change Prediction to Adapt ClimateSmart Agriculture Using Big Data Analytics. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 4*(5), 1995-2000.

Sinha, B. B., & Dhanalakshmi, R. (2022). Recent advancements and challenges of Internet of Things in smart agriculture: A survey. *Future Generation Computer Systems, 126*, 169-184. https://doi.org/10.1016/j.future.2021.08.006