

UDC 621.01:631.31

DOI: 10.31548/machinery/4.2025.43

Stanislav Nikolayenko

Doctor of Pedagogical Sciences, Professor, Academician of the National Academy of Agricultural Sciences of Ukraine
National University of Life and Environmental Sciences of Ukraine
03041, 15 Heroiv Oborony Str., Kyiv, Ukraine
<https://orcid.org/0000-0001-5271-1777>

Mykola Volokha*

Doctor of Technical Sciences, Professor
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
03056, 37 Beresteisky Ave., Kyiv, Ukraine
<https://orcid.org/0000-0002-0112-7324>

Digitalisation of field research data – the basis of farm management

Abstract. The efficiency of modern agricultural production management was shown to depend largely on the implementation of digitalisation processes, which transformed the paradigm of production process management and created conditions for increasing agricultural productivity through the use of innovative digital technologies. The purpose of the article was to review scientific publications devoted to the digitisation of field experiment databases in order to enhance the profitability of agricultural production through the development of effective management decisions in crop production based on data obtained from numerous sources of digital transformation. It was demonstrated that the diversity of technological processes, as well as input and output data involved in agricultural production, resulted in significant complexity and integrative challenges in assessing production sustainability. It was identified that the majority of farmers were not specialists in information technologies and were therefore unable to fully comprehend the complexity of the algorithms underlying modern digital solutions. Approaches to the use of big data in agriculture were systematised with regard to the evolution of the 4V concept towards the 5V framework, in which data value was emphasised as a critical requirement. A scientific and methodological approach aimed at improving agricultural production efficiency through the application of machine learning methods was proposed, including automatic crop recognition, disease and weed detection, and the forecasting of crop yield and product quality. It was established that, alongside material and technical support, the intellectualisation of production and management activities based on digitalisation constituted a priority factor in the development of the agro-industrial sector, being considered a highly effective means of increasing the return on investment in agriculture by 10-15%

Keywords: agricultural production efficiency; digitalisation; yield; soil fertility; moisture reserves

INTRODUCTION

Digitalisation and the integration of advanced technologies, including data analytics, machine learning (ML), and automation, offer significant potential to transform agricultural management, increase productivity, and ensure long-term sustainability. These innovations are becoming essential as part of efforts to increase agricultural production while minimising environmental impact. In review of the concepts and methods of sustainable agricultural

development, conducted by a team of authors from different countries it is argued that the way to solve the above problems, which put enormous pressure on the level of agricultural production, is to increase the efficiency of agricultural technologies while reducing the burden on the environment. These two important factors form the basis for transforming production in the agricultural sector based on the principles and characteristics of precision farming.

Article's History: Received: 07.07.2025; Revised: 17.10.2025; Accepted: 27.11.2025.

Suggested Citation:

Nikolayenko, S., & Volokha, M. (2025). Digitalisation of field research data – the basis of farm management. *Machinery & Energetics*, 16(4), 43-52. doi: 10.31548/machinery/4.2025.43.

*Corresponding author



Copyright © The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (<https://creativecommons.org/licenses/by/4.0/>)

Such modernisation has great potential for ensuring the sustainable development of agricultural production while maximising its productivity and environmental safety.

Researchers A. Cravero & S. Sepúlveda (2021) explored the use of ML in big data applications within agriculture. They focused on the practical adaptations of these technologies to improve efficiency in agricultural processes. By leveraging advanced data analysis techniques, they demonstrated how automation can enhance decision-making and optimize agricultural operations. In general, according to F. Zecca (2019), smart agriculture is based on four basic principles that make it possible to meet the growing needs of the population: a) optimal management of natural resources; b) use of advanced technologies; c) preservation of the ecosystem; d) development of relevant services.

An essential feature of modern agriculture is the use of information and communication technologies (ICTs), which are encouraged by policy makers around the world. ICTs may, for example, include farm management information systems, air and soil moisture sensors, accelerometers, wireless sensor networks, cameras, drones, low-cost satellites, online services, automated guided vehicles, etc. The disadvantage and a serious obstacle to obtaining valuable information from field research data is that traditional data processing methods are unable to meet the ever-growing needs in the new era of “smart” agriculture based on the algorithmic and software being created everywhere (Sørensen *et al.*, 2021). That is why ML has emerged, taking advantage of the rapid exponential growth of computing power and the tools of artificial intelligence and the Internet of Things (IoT) (Evstatiev & Gabrovska-Evstatieva, 2020). Researchers from Brazil G.M.D. Almeida *et al.* (2021) have demonstrated that logistics is a particularly important challenge in agriculture. As a cutting-edge technology, ML can solve significant problems, particularly in sugarcane transport chains, thereby significantly improving the overall economics of production. Ukrainian researchers also raise issues related to the digitisation process in Ukraine’s agro-industrial complex. M. Rudenko (2019), revealing digitisation as a component of modern management methods in agriculture, notes the lack of a unified approach to interpreting its essence. The author considered the multifaceted definition of this concept from different points of view: the state; scientists; practitioners (entrepreneurs); society.

In the article by Y. Voloshchuk (2019) the processes of formation and development of the digital economy of Ukraine are considered, the differences and peculiarities of its functioning and the impact on the competitiveness of the national economy are being studied. It is determined that digitalisation of the country is not a simple transition to e-government, but a full-scale introduction of digital technologies in all spheres of society: from education, skills, production, business to bridging the digital divide and creating a complete digital infrastructure. The basic moments of the digital strategy of the country are defined, problems of digitisation of industries, in particular agro-industrial complex, and ways of their solution are

outlined. M. Volokha *et al.* (2022) concluded that at the level of agricultural enterprises or farms, the digitisation of the data obtained allows monitoring production processes and processes of plant growth and development; assessing the condition of crops and forecasting crop yields; identifying areas requiring fertilisers and pesticides; controlling crop rotations and the quality of agro-technical measures; determining soil moisture, assessing snow cover, freezing areas, etc.

M. Lampridi *et al.* (2019) noted that one of the main problems is that the vast majority of farmers around the world are not experts in this field and therefore cannot fully understand the basic patterns obtained using ML algorithms. Despite this, a group of scientists from Italy, Tunisia, Malaysia, and the United Kingdom (Rejeb *et al.*, 2022) argues that the growing body of knowledge related to artificial intelligence and ML requires a bibliometric approach to understand the structure of knowledge. As research in this field develops and becomes more complex, it is extremely important to conduct a retrospective assessment of accumulated publications in order to identify new contributions, reflect research trends, and identify promising directions for future research.

The purpose of this article was to review publications devoted to the digitisation of field experiment databases in order to increase the profitability of agricultural production by developing effective management decisions in crop production based on the use of data from sources of digital transformation. To achieve this goal, the following tasks were set:

- systematise existing methods of digitising field research data based on the 5V requirements: volume, variety, velocity, reliability, and value;
- define the fundamental principles for developing realistic solutions that integrate the expertise of specialists in both computing and agriculture;
- develop a scientific and methodological approach to enhancing the efficiency of ML for automatic crop recognition, disease and weed detection, and yield forecasting.

To identify relevant studies related to the objective, literature published mainly between 2019 and 2024 was filtered, and the bibliometric method was used, which allows collecting information about publications for further generalisation. Scientometric databases and search platforms, including Scopus, Web of Science, Google Scholar, and ScienceDirect, were used. To study the aspects of agricultural management, combinations of the keywords “crop management”, “water management” and “soil management” in combination with the word “machine learning” were used. The texts and references in the named articles were scanned, and a critical analysis of the publications was conducted. Journal articles, conference papers, monograph, PhD dissertations were considered appropriate.

The literature review process was carried out on the methodological basis of synthesising all available criteria and a two-level analysis of relevant publications, which facilitates systematisation. Thus, a systematic literature

review of more than three hundred studies on ML methods for crop management and decision-making, and the assessment of precision crop production as a component of smart agriculture over the past five years was implemented. The study of the methodologies used is of particular importance, as there are no recommended standards or norms for assessing the sustainability of methods in the crop production industry. The known methods most often include tools based on indicators, indices, and multicriteria. They are often used in combination, so a clear distinction is not always possible.

Intelligent data processing and ML in agricultural logistics and supply chains

ML algorithms are fundamental to artificial intelligence and data science, enabling computers to learn and improve from data without explicit programming. They have a wide range of applications, including image and speech recognition, recommendation systems, and data analysis. In supervised learning, algorithms are trained on labelled data to make predictions or classifications. For classification tasks, common algorithms include Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks. Regression algorithms, such as Linear Regression and Polynomial Regression, predict continuous outcomes like prices or scores. Unsupervised learning involves finding patterns or structures in unlabelled data. Clustering algorithms like K-Means, Hierarchical Clustering, and DBSCAN group similar data points, aiding market segmentation and anomaly detection. Dimensionality reduction algorithms like Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbour Embedding (t-SNE) help in reducing the number of features in a dataset while preserving key information. This approach combines labelled and unlabelled data, useful when acquiring fully labelled data is costly or time-consuming. It's used in various applications, including text classification and image recognition. In reinforcement learning, agents learn to make sequences of decisions by receiving feedback in the form of rewards. It's crucial in robotics, game playing, and autonomous systems. These ML algorithms play a pivotal role in solving complex problems and making data-driven decisions across diverse domains (Jiang *et al.*, 2019).

To eliminate many of the shortcomings inherent in supply chains, ML can analyse large amounts of data, identify patterns and generate effective solutions. X. Fan *et al.* (2021) defined the role of ML in supply chain optimisation as follows. Learning plays a transformative role in optimising supply chains, revolutionising the way businesses manage their operations. Its impact is multi-faceted and extends across various aspects of supply chain management. One of the key areas where ML excels is demand forecasting. By analysing historical data, market trends, and external factors, ML algorithms can predict future demand more accurately than traditional methods. This enables businesses to adjust their inventory levels, production schedules, and distribution strategies to meet customer demand

while minimising overstock or stockouts. Route optimisation is another crucial aspect of supply chain management where ML shines. ML algorithms can analyse real-time data on factors like traffic, weather, and road conditions to determine the most efficient routes for transportation. This not only reduces transportation costs but also minimises delivery times, enhancing customer satisfaction. Inventory management is also greatly enhanced by ML. These algorithms can identify optimal inventory levels, taking into account lead times, demand variability, and costs associated with holding excess inventory. This prevents overstocking and understocking, leading to cost savings and improved efficiency. Quality control and maintenance in supply chains benefit from ML as well. Sensors and data analytics allow for real-time monitoring of equipment and product quality. Predictive maintenance models can anticipate when machinery or vehicles are likely to fail, enabling proactive maintenance to prevent costly breakdowns. Risk management is an essential aspect of supply chain optimisation. ML algorithms can analyse a multitude of data sources to identify potential risks, whether they are related to geopolitical issues, supplier reliability, or market fluctuations. This early warning system enables supply chain managers to mitigate risks and develop contingency plans.

Solving logistics problems requires a combination of technology, logistics innovation, and cooperation between supply chain participants. According to V. Borodin *et al.* (2016), developing strategies to manage seasonal fluctuations, optimise transportation, forecast demand fluctuations, and ensure consistent quality control are vital to the successful and sustainable operation of agricultural supply chains. ML can optimise transport routes, reducing waste and ensuring the timely delivery of perishable goods for example for the purpose of extending and regulating the shelf life of fruits and vegetables as in China (Li *et al.*, 2024). New artificial neural network models are being developed to accurately assess the temperature of food products stored in multi-temperature refrigerated vehicles (Zou *et al.*, 2023).

The American Chemical Society offers a technology for large-scale food storage by analysing images using a combination of intelligent sensor film and QR codes. Such rapid, real-time monitoring of food quality will ultimately lead to a reduction in food waste and losses (Pounds *et al.*, 2022). Another advantage of ML is that quality control systems based on it can automate the verification process, ensuring that only the highest quality products reach consumers. A. Rejeb *et al.* (2022) examined the interaction between artificial intelligence technologies and the agri-food industry, with particular emphasis on their role in improving efficiency, quality control, and sustainability across the entire food supply chain. The authors analysed how ML-based systems enable automated monitoring, traceability, and decision support, thereby enhancing product quality and reducing operational risks. It was concluded that the integration of artificial intelligence represents a key driver of digital transformation in the agri-food sector. By

reducing its environmental impact, ML can contribute to sustainable agriculture by supporting the efficient use of water, fertilisers, and pesticides.

Given that farmers are becoming increasingly accustomed to smartphones, scientists, particularly Indian (Mayuri, 2018) and Chinese (Yuan *et al.*, 2022), have proposed special smartphone applications as a possible solution to the above problem. These researchers believe that more user-friendly systems should be developed. In particular, devices that are easy to understand and operate, such as a visualisation tool with a user-friendly interface for the correct presentation and processing of data, would be valuable, as explored in detail in the work of Australian scientists (Ang & Seng, 2021).

In Ukraine, many works by international and Ukrainian scientists and experts are devoted mainly to the formation of digitalisation as a tool for the development of informatisation. Researchers emphasise that digitalisation requires a systematic approach to digital transformation and human capital development, as well as effective state policy and institutional support to ensure sustainable economic development. A similar view is shared by T. Lazorenko & I. Sholom (2018), who underlined the importance of coordinated digital reforms and the development of human potential. In addition, A. Trushlyakova (2018) pointed out the need for thorough planning of a digital strategy and the selection of an effective model to achieve the expected outcomes of digitalisation.

Thus, most authors whose research has focused on the essence of the concept of digitalisation limit themselves to the processes of national economic development in the context of globalisation. Instead, V. Varga (2020) conducted a content analysis of the basic term “digitalisation” and presented his own vision of the concept of digitalisation as one of the factors of enterprise competitiveness, when, thanks to the more convenient use of digital technologies, work with a significant database of information for economic growth is accelerated. The term “digitalisation” is understood as the process of using, applying, transferring, and translating information into digital format. The author notes that economic growth and transformation occur through the expansion and more convenient use of digital technologies. It is emphasised that innovations are designed to accelerate and facilitate work with a large database, use a unified system of cooperation with all departments of the company, and establish interaction with customers and suppliers. All changes that occur during the implementation of innovative innovations lead to the improvement and optimisation of the efficiency of the enterprise as a whole. It has been determined that digitalisation is currently leading not only to the emergence of new specialties, but also to changes in the labour market as a whole. It has been found that in order to increase competitiveness, economic entities must create a strategy that focuses on the use of innovative solutions, while introducing scientific and technical innovations, applying intellectualisation, and effectively implementing digital innovations.

All these actions are a driving force for development, which implements and accelerates the profitable activities of the enterprise. It has been analysed that digitalisation provides enterprises, including agricultural ones, with competitive advantages in all areas of activity, and thanks to this, the business entity will ensure confident stability in the future.

Big Data and data fusion technologies for smart and precision agriculture

The processes of digitising large amounts of field data, i.e., collecting field research data, processing or storing it, are inextricably linked to the concept of “Big Data”, as farmers, researchers and other agricultural professionals are now accumulating large amounts of data (Sonka, 2016). According to the United Nations, the exponentially growing amount of data generated by digital technologies requires greater capacity to store it on electronic computing applications for editing, analysis, and interpretation. At the same time, it should be noted that since the second half of the 1980s, the volume of data storage has doubled. The latter has significant potential and motivates decision makers to improve their intellectual level (The Guardian, 2017). If while in 2018 the US National Institute of Standards and Technology (NIST) noted that big data consisted of sets with characteristics of large volume (1), variety (2), velocity (3), and variability (4), i.e., 4V (Chang & Grady, 2018), then in 2020-2022, in addition to these characteristics, emphasising the scale and complexity of big data, scientists R.Y. Choi *et al.* (2020) added the descriptors of high dimension, high complexity, and high uncertainty, i.e., 5V, to the definition. Characterising the requirements of 5V it was highlighted the following (Table 1). The rationale for increasing the number of V units is presented in the review by A.Y. Sun & B.R. Scanlon (2019), which highlights the impact of big data and ML technologies on various aspects of environmental and water resource management, as advances in remote sensing methods have contributed to the spread of big data in many areas of environmental and water management, including weather forecasting and natural disaster prevention. Big data opens up new opportunities, although it requires, first and foremost, new forms of information analytics, provided that it is properly integrated with ML, which will help to fully unlock its potential. The review presents actual applications of big data ML in environmental and water resource management, where the authors focus on deep learning, especially since the illustrated data on environmental and water resource management require the use of analytical tools for big data analysis, ranging from virtualisation and edge computing to high-performance processing. It is noted that one of the most challenging applications of big data and deep interdisciplinary learning is precisely data from environmental and water resource management. The scope of the applications cited in this review includes more than 1,000 articles, which prove that: 1) researchers of environmental and water resource management data will significantly change the planning, conduct, and analysis of experiments; 2) the

most effective use of big data advantages is possible with low data cleaning costs; 3) thanks to a more reliable and faster flow of information, the significant effectiveness of deep learning methods has already been demonstrated in solving many problems of improving disaster relief efforts. Finally, big data and artificial intelligence will be most

valuable when they improve the level of reasoning and discussion. For big data in machine learning applied to environmental and water resource management, it is expected that deep learning will significantly improve the ability to understand the conditions and changes the planet is experiencing as a result of human and environmental impacts.

Table 1. Requirements for components in the 5V kit

5V component		Characteristic
No.	Name	
1	Volume	Refers to the size of big data. In the last decade, the growth of big data has coincided with the rapid development of ICT and the Internet, and the volume of its deployment is often measured in terabytes, petabytes, and sometimes exabytes (the amount of information in 1 exabyte is equal to a quintillion (1,018 bytes). To process data at scale, it is necessary to switch to a distributed computing platform. At the same time, the size of a data set is relative and may have different values in different applications.
2	Variety	This applies to data collection methods. Depending on the sources of observation, big data can be classified as passive or active, with the former originating from derivative digital media (e.g., GPS data from mobile phones) and the latter from planned data collection activities (field observations or targeted remote sensing). Passive data is collected without the active participation of the user and, as a result, does not explain the reasons for events. Depending on the format, data can be structured (sensor readings and satellite data) or unstructured (e-mail, video, audio, or text), with the main difference being that unstructured data cannot be easily described using a predefined model due to the limited information resources in the case of unstructured data. In other words, diversity creates an additional level of complexity when working with big data. Two methods are commonly used to reduce data diversity: virtualisation or data fusion.
3	Velocity	Speed attribute. This may refer not only to the speed of data generation, but also to the speed of analysis required to receive it. Data generation speed is a direct consequence of network connectivity, widespread use of smart devices, and real-time monitoring networks. High-speed data sets that are continuously generated by various sources require special processing. A data stream is defined as an unlimited sequence of events that requires continuous, incremental processing.
4	Variability	Variability refers to changes in all other attributes of big data, such as changes in data flow rates or changes in data values, with the latter often occurring in data obtained through crowdsourcing, i.e., a method of finding solutions to problems without significant external assistance. This is a serious problem for data processing, as many ML algorithms operate on the principle of stationarity.
5	Value	This refers to potential issues with data inconsistency and quality, such as partial missing data, anomalies, and input errors. High-quality products created by data centres involve a significant number of processing and integration steps. However, as data centres constantly update their algorithms, managing their versions becomes a serious challenge when accepting data from different sources. Therefore, when developing platforms for data collection and analysis, it is necessary and important to protect the reliability of information.

Source: developed by the authors based on W.L. Chang & N. Grady (2018)

Chinese electrical engineers from Xinjiang University X. Sibó *et al.* (2022) developed a data fusion algorithm based on the idea of parallel computing to improve the performance of the system for monitoring unevenly distributed temperature in smart greenhouses, which optimises the quality of a hierarchical network of wireless sensors and increases computing efficiency using a real-time data fusion strategy. In the study by R.Y. Choi *et al.* (2020), the application of ML methods and neural networks to real-time intelligent data processing tasks was analysed. The authors examined a multi-level system architecture in which, at the lower level, temperature data were collected and pre-processed using an improved Kalman filter, traditionally applied in applied mechanics and control theory to reduce measurement noise. At the intermediate level, characteristic of automation systems and computer-integrated technologies, a parallel intersection data fusion algorithm was implemented, enabling the aggregation of information from multiple sensors and increasing the reliability of the estimates. The upper level of the system was based on the use of an extreme ML algorithm with a global fusion centre, reflecting modern approaches in electrical engineering and power engineering to intelligent data analysis and the optimisation of operating modes of

complex technical objects. The simulation results reported by R.Y. Choi *et al.* demonstrated that the integration of classical control theory methods with ML tools ensured a significant improvement in the accuracy and responsiveness of the temperature control system. Thus, the study synthesised and systematised contemporary scientific developments at the intersection of mechanics, electrical engineering, automation, and computer-integrated technologies, and confirmed the feasibility of applying multi-level intelligent algorithms for the design of efficient, safe, and energy-efficient engineering systems.

Automation of crop production management using ML

According to the works reviewed in this study with regard to crop production, soil fertility, and water regime, sections devoted to water and soil resource management were less extensively investigated and accounted for approximately 18-20% of the total number of articles (up to 10% per section). In contrast, crop production management represented the majority of publications (up to 80% of the total) and was divided into four main categories in which ML methods were applied: crop recognition, crop quality

assessment, disease and weed detection, and profitability forecasting. Automatic crop recognition based on leaves is the most common approach for determining specific leaf characteristics such as colour, shape and texture. S. Virnodkar *et al.* (2020) systematically analysed remote sensing data and ML algorithms as tools for automated assessment of crop water stress, demonstrating the potential of image-based methods to replace manual field inspections and support automated decision-making in precision agriculture. The proposed approaches emphasised the integration of sensor data, image processing, and algorithmic analysis, which directly corresponds to modern trends in automation and computer-integrated agricultural technologies. With the expansion of the use of satellites and aircraft, crop classification using remote sensing has become particularly popular. E.I. Papageorgiou *et al.* (2018) focused on the development of intelligent computational models for crop and product classification, highlighting the role of automated data acquisition systems, signal processing, and soft computing methods in reducing human subjectivity and increasing the reliability of quality assessment. Their work illustrates how automated classification systems can be embedded into larger cyber-physical and electromechanical control frameworks for agricultural production. In other words, the development of computer software and image processing devices has led to the automatic recognition and classification of crops, in particular based on yield. Such solutions reflect a broader shift towards automation of analytical and control functions in agricultural systems, where software-driven image analysis replaces traditional manual monitoring and enables real-time feedback for management and optimisation processes.

At the same time, farmers in countries such as Bangladesh choose crops for planting based on their own knowledge. M. Hasan *et al.* (2023) addressed this limitation by proposing an ensemble ML-based recommendation system that formalises empirical farmer knowledge into an automated decision-support tool. The study demonstrated how data-driven algorithms can enhance traditional practices by providing objective, scalable, and reproducible recommendations, thereby supporting the transition from experience-based decision-making to automated and intelligent crop management systems.

High-quality agricultural products are usually sold at more favourable prices, which gives farmers higher incomes. For example, fruit quality, flesh firmness, soluble solids content, and skin colour are some of the simplest maturity indicators used for apple harvesting (Zhang *et al.*, 2021). Harvesting dates have a significant impact on the quality characteristics of harvested products of both high-value crops (grapes, vegetables, woody crops, herbs, etc.) and row crops. Thus, the development of decision support systems can help farmers and managers make appropriate management decisions to improve product quality. For example, selective harvesting is a management technique that can significantly improve quality. In addition, crop quality is closely related to food waste, which is

an additional problem that modern agriculture has to face, as if the crop deviates from the desired shape, colour, or size, it may simply be thrown away as substandard. However, by applying image processing technologies for such discarded raw materials, developed on the basis of ML algorithms, it is possible to reduce the amount of food waste for processing companies.

Diseases were identified by experienced agronomists through field reconnaissance. However, this process is time-consuming and based solely on visual inspection. Technological advances have made sensor systems commercially available that can identify diseased plants before symptoms appear. In addition, computer vision, especially using deep learning, has made significant progress over the past few years, as noted in the article by J. Zhang *et al.* (2021). The authors demonstrated that convolutional neural networks can be effectively applied to automated disease recognition using image data collected within the Agricultural IoT framework, reducing dependence on expert judgement and enabling real-time diagnostic decision-making. Their work highlighted the role of software-based vision systems as functional elements of automated control loops in crop production. A review by S.R. Upadhyaya *et al.* (2024) as a scientific achievement highlights the potential of ML as an additional tool for predicting genetic loci that contribute to pathogen resistance. This study extended automation concepts beyond operational monitoring by introducing data-driven predictive models at the genetic level, illustrating how ML can support automated breeding strategies and long-term optimisation of system resilience in agricultural production.

Weeds typically germinate and develop very quickly due to their abundant seed production and longevity and spread over large parts of the field, competing with crops for resources including space, sunlight, nutrients and water availability. In addition, weeds often emerge earlier than crops and do not have to face natural enemies that negatively affect cultivated plants. To prevent a decrease in yields, weed control can be done either by mechanical cultivation, herbicide application, or a combination of both. However, mechanical tillage is sometimes difficult to implement and ineffective if not done properly, making herbicide application the most common practice. The use of large quantities of herbicides is both expensive and harmful to the environment, especially if applied uniformly without taking into account differences in weed density across the field. Prolonged use of herbicides is likely to make weeds more resistant, leading to more difficult and costly weed control. Significant progress has been made in solving the problem of weed differentiation in crops based on smart agriculture. This differentiation can be achieved by remote or proximity sensing with sensors mounted on satellites, airborne and ground vehicles, and drones, both ground and airborne. However, converting data collected by UAVs into “usable” information is a challenging task, as their collection and classification require considerable effort. ML algorithms combined with imaging or non-imaging

spectroscopy technologies can provide real-time differentiation and localisation of certain weed species, enabling precise herbicide application and planning of the shortest weeding path in predefined areas instead of spraying the entire field (Ofosu *et al.*, 2023).

According to X. Gao *et al.* (2025), ML-based approaches represent an innovative solution to the persistent challenge of weed management in agriculture by enabling automated detection and control. The system integrates advanced ML algorithms with robotics to accurately find and cut weeds in real-time. A convolutional neural network (CNN) is trained on a diverse dataset comprising images of various crops and weed species, allowing the model to develop a robust understanding of visual cues associated with both. This trained model is deployed on a robotic platform equipped with cameras and actuators for real-time decision-making and precise weed removal. The integration of deep learning models with control systems that include laser modules, robotic manipulators, and precision spray nozzles facilitates the development of intelligent robotic mowing systems that significantly reduce chemical herbicide consumption and improve operational efficiency compared to traditional approaches. Significant obstacles remain, including limited adaptability to the environment, real-time processing limitations, and insufficient model generalisation. Future directions include integrating diverse data sources, developing optimised models, and improving intelligent decision-making systems, creating a foundation for the development of sustainable agricultural production.

Yield forecasting and resource management based on IoT and ML

Yield forecasting is one of the most important and complex topics in modern agriculture. An accurate model can help farm owners, for example, make informed management decisions about which crops to grow to ensure that the yield meets current market demands. However, this is not an easy task, it consists of different steps and directions of digitalisation, because the yield forecast can be determined by several factors, such as the technical level of the used machine complexes and the sophistication of the adapted technologies, soil and weather conditions, genotypic and phenotypic characteristics of crops and their interaction. Thus, it requires a fundamental understanding of the relationship between these interacting factors. In their study, M.R. Islam *et al.* (2023) proposed a ML-enabled IoT system for automated soil nutrient monitoring and crop recommendation, which directly supports yield forecasting through continuous data acquisition and real-time analysis. The authors integrated distributed soil sensors with data transmission modules and decision-support algorithms, forming a cyber-physical system capable of transforming raw sensor measurements into actionable management recommendations. Such an approach reflects the principles of automation and computer-integrated

technologies, as it reduces manual intervention, enables closed-loop decision-making, and enhances the reliability of predictive models by incorporating dynamic soil and environmental parameters. The proposed system demonstrated how the integration of IoT-based measurement infrastructure with ML algorithms can improve the robustness and practical applicability of yield forecasting models in modern agricultural production systems.

The agricultural sector is the main consumer of available fresh water on a global scale, as plant growth and development largely depend on the availability of water. Given the rapid depletion of many aquifers, the use of digital technologies provides more efficient management of the processes of accumulating and preserving productive moisture to ensure sustainable crop production, which is also confirmed by research by scientists at Texas Tech University J. Neupane & W. Guo (2019). As a result of the research by W. Zhao *et al.* (2018) in the field of precision agriculture, variable rate irrigation is proposed to save water. This can be achieved by applying irrigation rates that vary with field variability based on the specific water needs of individual management zones instead of using a single rate for the entire field. The level of digitisation of agronomic factors, including terrain slopes, affects the effectiveness and feasibility of variable irrigation rates, the amount of productive moisture in the soil. S.A. Mauget *et al.* (2017) noted that careful monitoring of soil moisture and other conditions for crop growth and development, combined with weather forecasting, can help to develop digital irrigation programs and efficient water management, especially where groundwater sources are used for irrigation and precipitation provides only a portion of the total evapotranspiration of crops, as was studied in cotton cultivation in arid areas of West Texas. Accurate soil information by agro-climatic zonation is important because it facilitates better soil management according to the land's potential. Effective soil management is also of great interest in relation to problems such as land degradation (loss of biological productivity), soil nutrient imbalances (due to excessive fertiliser use), and soil erosion (due to excessive vegetation cutting, improper crop rotation, and destruction by energy-driven vehicles). Useful soil properties can include texture, organic matter, and nutrient content, among others. Traditional methods of soil assessment include sampling and laboratory analysis, which are usually expensive and require significant time and effort. At the same time, remote sensing and soil mapping sensors can provide an inexpensive and simple solution to studying soil spatial variability. However, merging and processing such heterogeneous big data can be a serious drawback when using traditional analysis methods.

A serious obstacle to improving the economics of modern agriculture by obtaining valuable information is the limited processing of field data using traditional methods, especially against the backdrop of the ever-growing needs of precision (smart) agriculture, which requires the use of information and communication technologies. The basis for this is agricultural software, farm management

information systems, various sensors, cameras, drones, low-cost satellites, etc. As a result, ML has emerged, where artificial intelligence and IoT tools are widely used. The way to solve the existing problems that put enormous pressure on the level of agricultural development is to increase the efficiency of agricultural technologies while reducing the environmental impact by transforming production in the agricultural sector into precision agriculture. Such a significant and timely modernisation has a great potential to ensure the sustainability of agricultural production while increasing the return on investment in its development, profitability and a safe environment.

The efficiency of complex processes of agricultural production is significantly affected by the introduction of digitalisation, which is a way to accumulate large amounts of field data and a tool for the development of new, modern areas of research. Based on a systematic analysis of the methods and features of digitising field research data for use as a basis for crop management and decision-making based on the development of realistic solutions that combine the expert knowledge of specialists in the field of computing and agriculture, a scientific and methodological approach to improving the efficiency of ML for automatic crop recognition, detection of diseases and weeds, and prediction of crop yield and quality is proposed.

CONCLUSIONS

A serious obstacle to improving the economy of modern agriculture by obtaining valuable information is the limited processing of field data using traditional methods, especially against the backdrop of the ever-growing needs of precision (smart) agriculture, which requires the use of information and communication technologies. The basis for this is agricultural software, farm management information systems, various sensors, cameras, drones, inexpensive satellites, etc. As a result, ML has emerged, where artificial intelligence and IoT tools are widely used.

The way to solve existing problems that put enormous pressure on the level of agricultural development is to

increase the efficiency of agricultural technologies while reducing the impact on the environment by transforming production in the agricultural sector into smart agriculture and precision farming. Such a significant and timely modernisation has great potential to ensure the sustainability of agricultural production while increasing the return on investment in its development by 10-15%, profitability, and environmental safety. The efficiency of complex agricultural production processes is significantly influenced by the introduction of digitalisation, which is a way of accumulating large amounts of field data and a tool for developing new, modern areas of research. Based on the conducted systematic analysis, it was established that the integration of ML, sensor networks, and computer-integrated technologies forms a methodological basis for automating key decision-making processes in crop production.

Based on a systematic analysis of the methods and characteristics of digitisation of field research data for use as a basis for crop management and decision-making based on the development of realistic solutions that combine the expert knowledge of specialists in the field of computer technology and agriculture, a scientific and methodological approach is proposed to improve the effectiveness of ML for automatic crop recognition, disease and weed detection, and yield and crop quality prediction. Further research should focus on the development and validation of integrated cyber-physical systems that combine real-time sensor data, advanced ML models, and automated control algorithms to enhance the reliability and scalability of precision agriculture solutions.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

- [1] Almeida, G.M.D., Pereira, G.T., Bahia, A.S.R.D.S., Fernandes, K., & Marques, J.J. (2021). Machine learning in the prediction of sugarcane production environments. *Computers and Electronics in Agriculture*, 190, article number 106452. [doi: 10.1016/j.compag.2021.106452](https://doi.org/10.1016/j.compag.2021.106452).
- [2] Ang, K.L.-M., & Seng, J.K.P. (2021). Big data and machine learning with hyperspectral information in agriculture. *IEEE Access*, 9, 36699-36718. [doi: 10.1109/ACCESS.2021.3051196](https://doi.org/10.1109/ACCESS.2021.3051196).
- [3] Borodin, V., Bourtembourg, J., Hnaien, F., & Labadie, N. (2016). Handling uncertainty in agricultural supply chain management: A state of the art. *European Journal of Operational Research*, 254(2), 348-359. [doi: 10.1016/j.ejor.2016.03.057](https://doi.org/10.1016/j.ejor.2016.03.057).
- [4] Chang, W.L., & Grady, N. (2018). *NIST big data interoperability framework: Volume 1, big data definitions* (NIST special publication 1500-1). Gaithersburg, MD: National Institute of Standards and Technology. [doi: 10.6028/NIST.SP.1500-1](https://doi.org/10.6028/NIST.SP.1500-1).
- [5] Choi, R.Y., Coyner, A.S., Kalpathy-Cramer, J., Chiang, M.F., & Campbell, J.P. (2020). Introduction to machine learning, neural networks, and deep learning. *Translational Vision Science & Technology*, 9(2), article number 14. [doi: 10.1167/tvst.9.2.14](https://doi.org/10.1167/tvst.9.2.14).
- [6] Cravero, A., & Sepúlveda, S. (2021). Use and adaptations of machine learning in Big Data-applications in real cases in agriculture. *Electronics*, 10, article number 552. [doi: 10.3390/electronics10050552](https://doi.org/10.3390/electronics10050552).

- [7] Evstatiev, B.I., & Gabrovska-Evstatieva, K.G. (2020). Review on the methods for big data analysis in agriculture. *IOP Conference Series: Materials Science and Engineering*, 1032, article number 012053. doi: [10.1088/1757-899X/1032/1/012053](https://doi.org/10.1088/1757-899X/1032/1/012053).
- [8] Fan, X., Nan, Z., Ma, Y., Zhang, Y., & Han, F. (2021). Research on the spatio-temporal impacts of environmental factors on the fresh agricultural product supply chain and the spatial differentiation issue – an empirical research on 31 Chinese provinces. *International Journal of Environmental Research and Public Health*, 18(22), article number 12141. doi: [10.3390/ijerph182212141](https://doi.org/10.3390/ijerph182212141).
- [9] Gao, X., Gao, J., & Qureshi, W.A. (2025). Applications, trends, and challenges of precision weed control technologies based on deep learning and machine vision. *Agronomy*, 15(8), article number 1954. doi: [10.3390/agronomy15081954](https://doi.org/10.3390/agronomy15081954).
- [10] Hasan, M., Marjan, M.A., Uddin, M.P., Afjal, M.I., Kardiy, S., Ma, S., & Nam, Y. (2023). Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation. *Frontiers in Plant Science*, 14, article number 1234555. doi: [10.3389/fpls.2023.1234555](https://doi.org/10.3389/fpls.2023.1234555).
- [11] Islam, M.R., Oliullah, K., Kabir, M.M., Alom, M., & Mridha, M.F. (2023). Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation. *Journal of Agriculture and Food Research*, 14, article number 100880. doi: [10.1016/j.jafr.2023.100880](https://doi.org/10.1016/j.jafr.2023.100880).
- [12] Jiang, L., Zeng, G., Wan, C., & Liu, Z. (2019). Research on the countermeasures for rural revitalization of the ethnic minorities in border areas: A case study of Heihe City in Heilongjiang Province. In *Proceedings of the 3rd international conference on culture, education and economic development of modern society (ICCESE 2019)* (pp. 657-660). Paris: Atlantis Press. doi: [10.2991/iccese-19.2019.144](https://doi.org/10.2991/iccese-19.2019.144).
- [13] Lampridi, M., Sørensen, C., & Bochtis, D. (2019). Agricultural sustainability: A review of concepts and methods. *Sustainability*, 11, article number 5120. doi: [10.3390/su11185120](https://doi.org/10.3390/su11185120).
- [14] Lazorenko, T., & Sholom, I. (2018). [Digitalization as the main factor of business development](https://doi.org/10.3390/agronomy9020087). *Proceedings of the international scientific and practical conference* (pp. 50-51). Kyiv: Kyiv Polytechnic Institute.
- [15] Li, D., Bai, L., Wang, R., & Ying, S. (2024). Research progress of machine learning in extending and regulating the shelf life of fruits and vegetables. *Foods*, 13(19), article number 3025. doi: [10.3390/foods13193025](https://doi.org/10.3390/foods13193025).
- [16] Mauget, S.A., Adhikari, P., Leiker, G., Baumhardt, R.L., Thorp, K.R., & Ale, S. (2017). Modeling the impact of management and increase on cotton production in arid areas of West Texas. *Agricultural and Forest Meteorology*, 247, 385-398. doi: [10.1016/j.agrformet.2017.07.009](https://doi.org/10.1016/j.agrformet.2017.07.009).
- [17] Mayuri, K.P. (2018). Role of image processing and machine learning techniques in disease recognition, diagnosis and yield prediction of crops: A review. *International Journal of Advanced Research in Computer Science*, 9(2), 788-795. doi: [10.26483/ijarcs.v9i2.5793](https://doi.org/10.26483/ijarcs.v9i2.5793).
- [18] Neupane, J., & Guo, W. (2019). Agronomic bases and strategies of accurate water resources management: A review. *Agronomy*, 9(2), article number 87. doi: [10.3390/agronomy9020087](https://doi.org/10.3390/agronomy9020087).
- [19] Ofosu, R., Agyemang, E.D., Márton, A., Pásztor, G., Taller, J., & Kazinczi, G. (2023). Herbicide resistance: Managing weeds in a changing world. *Agronomy*, 13, article number 1595. doi: [10.3390/agronomy13061595](https://doi.org/10.3390/agronomy13061595).
- [20] Papageorgiou, E.I., Angelopoulou, K., Gemtos, T.A., & Nanos, G.D. (2018). Development and evaluation of a fuzzy inference system and a neuro-fuzzy inference system for grading apple quality. *Applied Artificial Intelligence*, 32(3), 253-280. doi: [10.1080/08839514.2018.1448072](https://doi.org/10.1080/08839514.2018.1448072).
- [21] Pounds, K., Bao, H., Luo, Y., De, J., Schneider, K., Correll, M., & Tong, Z. (2022). Real-time and rapid food quality monitoring using smart sensory films with image analysis and machine learning. *ACS Food Science & Technology*, 2(7), 1123-1134. doi: [10.1021/acfoodscitech.2c00124](https://doi.org/10.1021/acfoodscitech.2c00124).
- [22] Rejeb, A., Rejeb, K., Zailani, S., Keogh, J.G., & Appolloni, A. (2022). Examining the interplay between artificial intelligence and the agri-food industry. *Artificial Intelligence in Agriculture*, 6, 111-128. doi: [10.1016/j.aiaa.2022.08.002](https://doi.org/10.1016/j.aiaa.2022.08.002).
- [23] Rudenko, M. (2019). The features of digitization influence on the functioning of agricultural enterprises. *Bulletin of the Kharkiv National Agricultural University of V.V. Dokuchaeva*, 1, 202-212. doi: [10.31359/2312-3427-2019-1-202](https://doi.org/10.31359/2312-3427-2019-1-202).
- [24] Sibó, X., Xinyuan, N., Xin, C., & Xumeng, L. (2022). Data fusion based wireless temperature monitoring system applied to intelligent greenhouse. *Computers and Electronics in Agriculture*, 192, article number 106576. doi: [10.1016/j.compag.2021.106576](https://doi.org/10.1016/j.compag.2021.106576).
- [25] Sonka, S. (2016). Big Data: Fueling the next evolution of agricultural innovation. *Journal of Innovation Management*, 4(1), 114-136. doi: [10.24840/2183-0606_004.001_0008](https://doi.org/10.24840/2183-0606_004.001_0008).
- [26] Sørensen, C.A.G., Kateris, D., & Bochtis, D. (2021). ICT innovation and smart agriculture. In M. Salamasis & T. Bournaris (Eds.), *Information and communication technologies in modern agricultural development* (vol. 953, pp. 1-19). Cham: Springer. doi: [10.1007/978-3-030-12998-9_1](https://doi.org/10.1007/978-3-030-12998-9_1).
- [27] Sun, A.Y., & Scanlon, B.R. (2019). How can Big Data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. *Environmental Research Letters*, 14(7), article number 073001. doi: [10.1088/1748-9326/ab1b7d](https://doi.org/10.1088/1748-9326/ab1b7d).
- [28] The Guardian. (2017). *Tsunami of data could consume one fifth of global electricity by 2025*. Retrieved from <https://www.theguardian.com/environment/2017/dec/11/tsunami-of-data-could-consume-fifth-global-electricity-by-2025>.
- [29] Trushlyakova, A.B. (2018). Development of digitalization in Ukraine: Factors of influence, advantages and challenges of today. *Economies' Horizons*, 4(7), 186-191. doi: [10.31499/2616-5236.4\(7\).2018.212762](https://doi.org/10.31499/2616-5236.4(7).2018.212762).
- [30] Upadhyaya, S.R., Danilevicz, M.F., Dolatabadian, A., Neik, T.X., Zhang, F., Al-Mamun, H.A., Bennamoun, M., Batley, J., & Edwards, D. (2024). Genomics-based plant disease resistance prediction using machine learning. *Plant Pathology*, 73(9), 2298-2309. doi: [10.1111/ppa.13988](https://doi.org/10.1111/ppa.13988).

- [31] Varga, V. (2020). Digitalization as one of the factors of enterprise competitiveness. *Efektivna Ekonomika*, 8. doi: [10.32702/2307-2105-2020.8.154](https://doi.org/10.32702/2307-2105-2020.8.154).
- [32] Virnodkar, S., Pachghare, V., Patil, V., & Jha, S. (2020). Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precision Agriculture*, 21, 1121-1155. doi: [10.1007/s11119-020-09711-9](https://doi.org/10.1007/s11119-020-09711-9).
- [33] Volokha, M., Bondar, S., Boldyrieva, L., & Kisilevskaya, M. (2022). [Advantages of digitalization in field data processing and crop management](#). In *Modern research in world science* (pp. 414-420). Lviv: Scientific Publishing Center "Sci-conf. com.ua".
- [34] Voloshchuk, Y. (2019). Digital directions of agrarian enterprises. *Efektivna Ekonomika*, 2. doi: [10.32702/2307-2105-2019.2.66](https://doi.org/10.32702/2307-2105-2019.2.66).
- [35] Yuan, Y., Chen, L., Wu, H., & Li, L. (2022). Advanced agricultural disease image recognition technologies: A review. *Information Processing in Agriculture*, 9(1), 48-59. doi: [10.1016/j.inpa.2021.01.003](https://doi.org/10.1016/j.inpa.2021.01.003).
- [36] Zecca, F. (2019). [The use of Internet of Things for the sustainability of the agricultural sector: The case of climate smart agriculture](#). *International Journal of Civil Engineering and Technology*, 10, 494-501.
- [37] Zhang, J., Rao, Y., Husband, K., Jiang, Z., & Lee, S. (2021). Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things. *International Journal of Distributed Sensor Networks*, 17(4). doi: [10.1177/15501477211007407](https://doi.org/10.1177/15501477211007407).
- [38] Zhao, W., Li, J., Yang, R., & Li, Y. (2018). Determining placement criteria of moisture sensors through temporal stability analysis of soil water contents for a variable rate irrigation system. *Precision Agriculture*, 19(4), 648-665. doi: [10.1007/s11119-017-9545-2](https://doi.org/10.1007/s11119-017-9545-2).
- [39] Zou, Y., Wu, J., Wang, X., Morales, K., Liu, G., & Manzardo, A. (2023). An improved artificial neural network using multi-source data to estimate food temperature during multi-temperature delivery. *Journal of Food Engineering*, 351, article number 111518. doi: [10.1016/j.jfoodeng.2023.111518](https://doi.org/10.1016/j.jfoodeng.2023.111518).

Станіслав Ніколаєнко

Доктор педагогічних наук, професор, академік Національної академії аграрних наук України
Національний університет біоресурсів і природокористування України
03041, вул. Героїв Оборони, 15, м. Київ, Україна
<https://orcid.org/0000-0001-5271-1777>

Микола Волоха

Доктор технічних наук, професор
Національний технічний університет України «Київський політехнічний інститут
імені Ігоря Сікорського»
03056, просп. Берестейський, 37, м. Київ, Україна
<https://orcid.org/0000-0002-0112-7324>

Оцифрування даних польових досліджень – основа управління фермерським господарством

Анотація. Ефективність управління сучасним сільськогосподарським виробництвом була визначена як така, що значною мірою залежить від впровадження процесів діджиталізації, які змінюють парадигму управління виробничими процесами та створюють умови для підвищення продуктивності аграрного виробництва шляхом використання інноваційних цифрових технологій. Метою статті було здійснення огляду наукових публікацій, присвячених оцифруванню баз даних польових досліджень, з метою підвищення прибутковості аграрного виробництва шляхом розроблення ефективних управлінських рішень у рослинництві на основі використання даних, що надходять із численних джерел цифрової трансформації. Було показано, що різноманіття технологічних процесів, а також вхідних і вихідних даних, задіяних у сільськогосподарському виробництві, зумовлює складність та комплексність узагальнення процесів оцінювання його сталості. Визначено, що значна частина фермерів не є фахівцями з інформаційних технологій і не має можливості повною мірою розуміти алгоритмічну складність сучасних цифрових рішень. У роботі було систематизовано підходи до використання великих даних у сільському господарстві з урахуванням еволюції концепції 4V до 5V, де додатково враховується цінність даних. Запропоновано науково-методологічний підхід, спрямований на підвищення ефективності сільськогосподарського виробництва шляхом застосування методів машинного навчання для автоматичного розпізнавання культур, виявлення хвороб і бур'янів, прогнозування врожайності та якості сільськогосподарської продукції. Встановлено, що поряд із матеріально-технічним забезпеченням першочерговим чинником розвитку агропромислового комплексу є інтелектуалізація виробничої та управлінської діяльності на основі цифровізації, що розглядалась як високоефективний інструмент підвищення рентабельності інвестицій у сільське господарство на 10-15 %

Ключові слова: ефективність сільськогосподарського виробництва; діджиталізація; врожайність; родючість ґрунту; запаси вологи