https://doi.org/10.31891/csit-2024-4-10
UDC 004.4

Dmytro OKRUSHKO, Olga PAVLOVA
Khmelnytskyi National University

DECISION-MAKING SUPPORT SYSTEM REGARDING THE OPTIMIZAION PROCESS OF CROP CULTIVATION USING REMOTE SENSING DATA

This paper explores the development of an information system to support decision-making in agriculture, specifically focusing on optimizing crop production. This system leverages the power of remote sensing (RS) data, which offers valuable insights into crop health and environmental conditions from a bird's-eye view. By analyzing this data, the proposed system empowers farmers with the information they need to make informed choices throughout the agricultural season, ultimately leading to increased yields and improved resource management.

Following the introduction, the paper delves into the core functionalities of the information system. It details the process of acquiring RS data from various platforms, such as Landsat, Sentinel-2, or PlanetScope satellites. Here, the discussion emphasizes the importance of selecting data with appropriate spatial and temporal resolution to capture the most relevant details for specific agricultural applications. Pre-processing techniques for handling the raw RS data are then discussed, outlining methods for removing noise and errors to ensure the accuracy of subsequent analyses. The paper then details the implementation of various algorithms for data analysis. These algorithms extract meaningful features from the pre-processed RS data, such as vegetation indices that provide insights into plant health and biomass, or other indicators that can detect potential crop stress due to nutrient deficiencies or water scarcity.

In conclusion, this crop monitoring and decision-support system equips farmers with the tools necessary to make informed, data-driven decisions, which ultimately increases yields, improves resource efficiency, and reduces the environmental footprint of agriculture. The adoption of such systems ensures that modern agriculture can meet the demands of a growing population while preserving the ecological balance.

Keywords: remote sensing, agriculture, decision-making, information system, optimization, yield, resource management, disease outbreak, sustainability.

Дмитро ОКРУШКО, Ольга ПАВЛОВА
Хмельницький національний університет

СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ЩОДО ОПТИМІЗАЦІЇ ПРОЦЕСУ ВИРОЩУВАННЯ УРОЖАЮ ЗА ДАНИМИ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ

У цій статті досліджується розробка інформаційної системи для підтримки прийняття рішень у сільському господарстві, особливо зосереджуючись на оптимізації виробництва сільськогосподарських культур. Ця система використовує потужність даних дистанційного зондування (RS), яка пропонує цінну інформацію про здоров'я врожаю та стан навколишнього середовища з висоти пташиного польоту. Аналізуючи ці дані, запропонована система надає фермерам інформацію, необхідну для прийняття обґрунтованого вибору протягом сільськогосподарського сезону, що зрештою призводить до підвищення врожайності та покращення управління ресурсами.

Також розглянуто основні функції інформаційної системи. Детально описується процес отримання даних RS з різних платформ, таких як супутники Landsat, Sentinel-2 або PlanetScope. Тут обговорення наголошує на важливості вибору даних з відповідною просторовою та часовою роздільною здатністю для захоплення найбільш релевантних деталей для конкретних сільськогосподарських застосувань. Потім обговорюються методи попередньої обробки для обробки необроблених даних RS, окреслюються методи видалення шуму та помилок для забезпечення точності наступних аналізів. Далі в статті детально описується реалізація різних алгоритмів для аналізу даних. Ці алгоритми витягують значущі характеристики з попередньо оброблених даних RS, наприклад індекси рослинності, які дають уявлення про стан рослин і біомасу, або інші показники, які можуть виявити потенційний стрес урожаю через дефіцит поживних речовин або дефіцит води..

Таким чином, ця система моніторингу посівів і підтримки прийняття рішень надає фермерам інструменти, необхідні для прийняття обґрунтованих рішень на основі даних, що в кінцевому підсумку підвищує врожайність, покращує ефективність використання ресурсів і зменшує вплив сільського господарства на навколишнє середовище. Впровадження таких систем гарантує, що сучасне сільське господарство може задовольнити потреби зростаючого населення, зберігаючи при цьому екологічний баланс.

Ключові слова: дистанційне зондування, сільське господарство, прийняття рішень, інформаційна система, оптимізація, врожайність, управління ресурсами, спалах захворювань, стійкість.

Introduction

Food security, the ability to consistently access enough safe and nutritious food, is a fundamental human right facing increasing pressure on a global scale. A confluence of factors is driving this challenge[1]:

Population Growth**.** The global population is projected to reach nearly 10 billion by 2050, placing a significant strain on our ability to produce enough food to meet this growing demand.

Climate Change**.** Rising temperatures, changing precipitation patterns, and extreme weather events driven by climate change are disrupting agricultural production and threatening crop yields. Additionally, climate change is contributing to land degradation, desertification, and salinization, further reducing arable land.

Resource Scarcity**.** Freshwater availability for irrigation is a growing concern, particularly in arid and semi-arid regions. Additionally, unsustainable agricultural practices have led to soil erosion and nutrient depletion, requiring more resources to maintain productivity.

Traditional agricultural practices, while providing sustenance for centuries, are no longer sufficient to meet the demands of the 21st century. The agricultural sector needs a paradigm shift towards **sustainable intensification**, focusing on producing more food with fewer resources and minimizing environmental impact.

This urgency is driving a wave of innovation in agriculture, often referred to as **precision agriculture**. Precision agriculture leverages advanced technologies like remote sensing, data analytics, and automation to gain deeper insights into crops and their environment. This allows farmers to make data-driven decisions that optimize resource use, improve crop health, and ultimately enhance yields.

While precision agriculture holds immense potential, implementing it effectively requires access to real-time, actionable data. Farmers need to move beyond traditional broad-acre management practices and gain a granular understanding of the health and status of their crops across their entire fields.

Existing methods of data collection in agriculture often rely on manual labor or limited field sensors, which can be time-consuming, expensive, and fail to provide a comprehensive picture. Therefore, a critical challenge lies in developing efficient and cost-effective methods for gathering and analyzing data that can guide informed decision-making throughout the agricultural season.

This paper proposes a novel information system that addresses this challenge by leveraging the power of remote sensing (RS) data. By analyzing RS data collected from satellites and aerial platforms, the system can provide farmers with valuable insights into various aspects of their fields, empowering them to make data-driven decisions for optimal crop production and resource management.

The Purpose of the Research. This research aims to develop a comprehensive information system that utilizes remote sensing (RS) data to support decision-making for optimizing crop cultivation. The system will empower farmers with valuable insights derived from RS data analysis, ultimately leading to increased yields, improved resource management, and sustainable agricultural practices.

Object of Research**.** The object of this research is the creation of an information system specifically designed for the agricultural sector. This system will bridge the gap between raw RS data and actionable insights for farmers, focusing on optimizing crop production throughout the agricultural season.

Subject of Research. The subject of this research encompasses two key areas:

Methods for processing and analyzing RS data**.** This involves developing algorithms and techniques to extract meaningful information from raw RS data relevant to crop health, environmental conditions, and potential stress factors.

Decision-support methodologies. This focuses on translating the analyzed RS data into actionable recommendations for farmers, such as optimizing irrigation practices, targeted nutrient application, and early detection of pest or disease outbreaks.

Research Methods. The research will employ a combination of methods to achieve its objectives: Remote Sensing Data Acquisition**:** Strategies for acquiring appropriate RS data from various platforms like satellites (Landsat, Sentinel-2) or aerial imagery will be explored, Data Pre-processing**:** Techniques to clean and prepare the raw RS data for analysis, ensuring accuracy and removing noise or errors, will be implemented, Data Analysis**:** Algorithms for extracting meaningful features from the pre-processed data will be developed. These features may include vegetation indices, crop health indicators, or stress detection based on nutrient deficiencies or water scarcity, Data Visualization**:** User-friendly methods for presenting the analyzed RS data will be established, such as generating informative maps, graphs, and charts, to facilitate comprehension for farmers, Decision-Support Model Development: This involves designing a system that translates the analyzed data into actionable recommendations tailored to specific crop types and environmental conditions.

Scientific Novelty of Results. This research strives to deliver novel contributions through the development of the information system. These contributions may include: advanced algorithms for efficient and accurate RS data analysis specific to agricultural applications, innovative visualization techniques that effectively communicate complex data to farmers with varying technical backgrounds, a robust decision-support model that generates precise and actionable recommendations based on real-time RS data.

Practical Significance of the Obtained Results. The information system developed in this research will offer significant practical benefits for farmers and the agricultural sector as a whole. These benefits include:increased crop yield and improved resource efficiency through data-driven decision making, enhanced farm management practices through real-time insights into crop health and environmental conditions,early detection and mitigation of pest and disease outbreaks, minimizing crop losses, promotion of sustainable agricultural practices by optimizing resource use and minimizing environmental impact.

Analysis of existing tools for optimization process of crop cultivation using remote sensing data

While advancements in remote sensing technology have opened doors for data-driven agriculture, a variety of existing tools and software solutions address crop optimization using this technology. However, translating this data into actionable decisions necessitates robust information systems. This analysis examines five existing tools designed to bridge this gap, highlighting their strengths, weaknesses, and potential areas for improvement. The key findings are summarized in a table (Table 1.1 and Table 1.2) and a chart (Figure 1.1) for clear comparison.

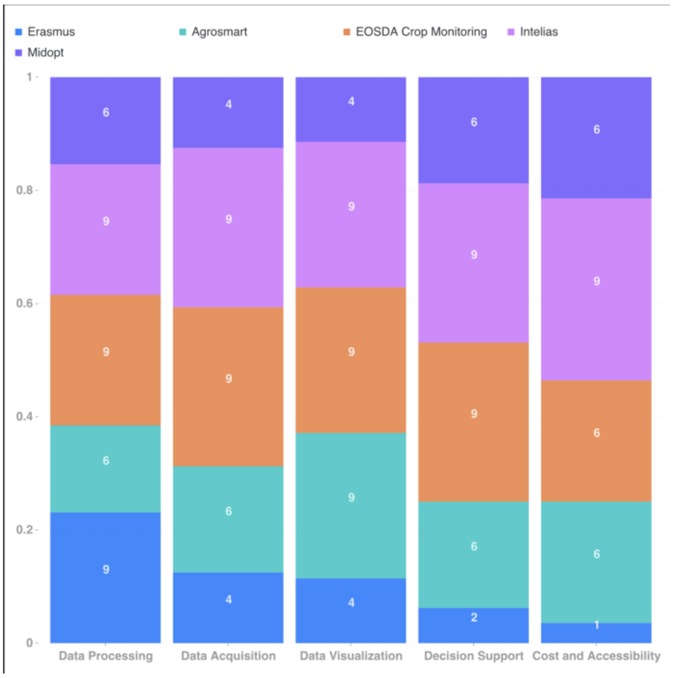

Fig. 1. Tool Comparison chart

Erasmus creation [12]**.** This is a project funded by the European Union's Horizon 2020 research and innovation program. While an exact launch date is difficult to pinpoint, project information suggests it started around 2016-2017. Erasmus primarily caters to researchers, agricultural scientists, and developers familiar with remote sensing data analysis. A research team studying drought tolerance in wheat crops might leverage Erasmus to analyze multi-temporal satellite imagery and assess crop health variations across a test field. They can utilize Erasmus' modular framework to integrate specific algorithms for vegetation index calculation and drought stress detection. Erasmus prioritizes data processing and analysis, catering to researchers and developers.

Agrosmart [10]. While a precise launch date isn't readily available, Agrosmart seems to be a well-established company with a presence in South America and potentially other regions. Agrosmart is designed for farmers of all technical backgrounds, with its user-friendly interface and mobile app integration promoting accessibility. A soybean farmer in Brazil concerned about potential nutrient deficiencies in their fields can use Agrosmart. The farmer uploads high-resolution satellite imagery of their fields, and Agrosmart analyzes the data to generate insights on crop health and potential nutrient deficiencies. Based on these insights, the farmer can make informed decisions about targeted fertilizer application to optimize yield. Agrosmart prioritizes user-friendliness and readily available recommendations for farmers.

EOSDA Crop Monitoring [14]. EOSDA is a geospatial analytics company founded in 2015. EOSDA Crop Monitoring is one of their agricultural product offerings. EOSDA Crop Monitoring targets farms of various sizes, with their tiered subscription plans catering to different data needs and farm acreage. A large-scale corn producer in the United States can utilize EOSDA Crop Monitoring. The producer subscribes to a plan that offers high-resolution satellite imagery and advanced AI-powered yield prediction. EOSDA's system analyzes data throughout the growing season, providing insights on potential yield variations across different zones within the vast cornfields. This allows the producer to strategically allocate resources like irrigation and fertilizer to maximize overall yield. EOSDA focuses on advanced analytics and AI-powered insights, potentially requiring more technical knowledge for interpretation.

Intelias [8]. Intelias is a software development company founded in 2009. While an exact launch date for their agricultural solutions is unavailable, their experience suggests a strong presence in the geospatial and agricultural technology sectors. Intelias targets large-scale farms, agricultural enterprises, and government agencies seeking highly customized solutions for crop optimization. Their services likely cater to users with a designated IT team or technical personnel to manage the custom system. A government agency in charge of managing agricultural land for smallholder farmers in Africa might partner with Intelias. Intelias would develop a custom solution that integrates with various data sources, including satellite imagery, weather data, and soil maps. This system would provide smallholder farmers with targeted recommendations on crop selection, planting times, and resource allocation based on their specific land characteristics and local weather conditions. Intelias prioritizes custom-built solutions for individual farm needs, offering comprehensive control over functionalities.

Midopt [9]. Similar to Intelias, specific details about Midopt's founding date are limited. Their focus on optimization algorithms suggests expertise in the agricultural technology and decision-support systems domain. Midopt likely targets medium to large-scale farms seeking to optimize resource allocation for maximizing yield. A dairy farm facing rising costs for water and fertilizer might utilize Midopt. The farm uploads data on their past crop yields, resource allocation history, and soil characteristics. Midopt's system analyzes this data and suggests optimized resource allocation strategies for the upcoming growing season. This could involve recommendations on water usage per field zone or fertilizer application rates based on predicted crop needs. Midopt prioritizes resource optimization using advanced algorithms, potentially requiring training for effective utilization.

The extended comparison of the highlighted tools is presented in the Tables 1. and 2.

Table 1.

Comparison of Decision-Support Tools for Crop Optimization (features)

Tool name	Focus	Cost and accessibility	Technical expertise
Erasmus	Open-source data processing and analysis	Free and open source	Requires advanced technical knowledge
Agrosmart	User-friendly recommendations	Subscription based model	User-friendly with for farmers
EOSDA	AI powered insight and yield prediction	Subscription based with various tiers.	Requires some technical skills for advanced features
Intelais	Customizable solutions	Custom development cost (pay as you go).	High level of technical skills needed for customization
Midopt	Resource optimization	Subscription based with training courses.	Required training for effective use.

Table 2

Comparison of Decision-Support Tools for Crop Optimization (work with data)

Tool name	Data acquisition	Data processing	Data visualization	Decision support
Erasmus	Limited to compatible open-source data sources	User manages pre-processing data and analysis algorithms	Basic visualization tools provided	Requires expertise to interpret analysis results
Agrosmart	Integrates with various platform (may require additional setup	Pre-processing included, limited customization for analysis algorithms	User-friendly interface with interactive maps, graphs and charts.	Offers actionable recommendations for various farm practices
EOSDA	Offers data acquisition as part of the service	Advanced AI and machine learning for data analysis.	Advanced visualization options with customization capabilities.	AI powered insights with yield prediction and targeted recommendations
Intelais	Custom integrations with desired platforms	Extensive customization for pre-processing data and analysis algorithms	Highly customizable dashboards and reports	Comprehensive recommendations tailored to specific farm needs
Midopt	Limited data source options	Focuses on optimization algorithms, limited crop health analysis	Limited visualization options	Focuses on resource allocation recommendations for optimal yield

Crop monitoring workflow

The paper [3] explores the use of remote sensing and Geographic Information Systems (GIS) in agriculture. The use of remote sensing and Geographic Information Systems (GIS) in agriculture has revolutionized the way farmers manage crops. As explored in the article [23], remote sensing platforms such as satellites, drones, and unmanned aerial vehicles (UAVs) provide multi-spectral and hyperspectral imagery that can be used to monitor crop health over time. Below is an extended workflow for crop monitoring using these technologies.

Key Stages of Crop Monitoring Workflow (Figure 2): Data Collection, Data Pre-Processing, Feature Extraction and Indices Calculation, Data Integration with GIS, Crop Health Assessment and Analytics, Decision Support Systems (DSS), Visualization and Reporting, Feedback Loop.

Data Collection with Plain map. This stage focuses on collecting raw data from various sources. The plain map acts as the foundation for geographic reference. It provides boundaries for fields, spatial markers, and helps geo reference other datasets such as sensor readings or remote sensing imagery. Remote sensing data, collected from satellites (e.g., NASA's Landsat, MODIS) or drones, provides real-time information on crops, soil, and environmental conditions.

Using Google Earth API and NASA data (Landsat, MODIS) from 2012 to today. Field Sensors: soil compaction sensors, moisture sensors, nutrient sensors. Historical Data: Previous crop yields, weather data. We can collect from remote sensing data such as Normalized Difference Vegetation Index (NDVI), moisture content, and soil compaction indicators.

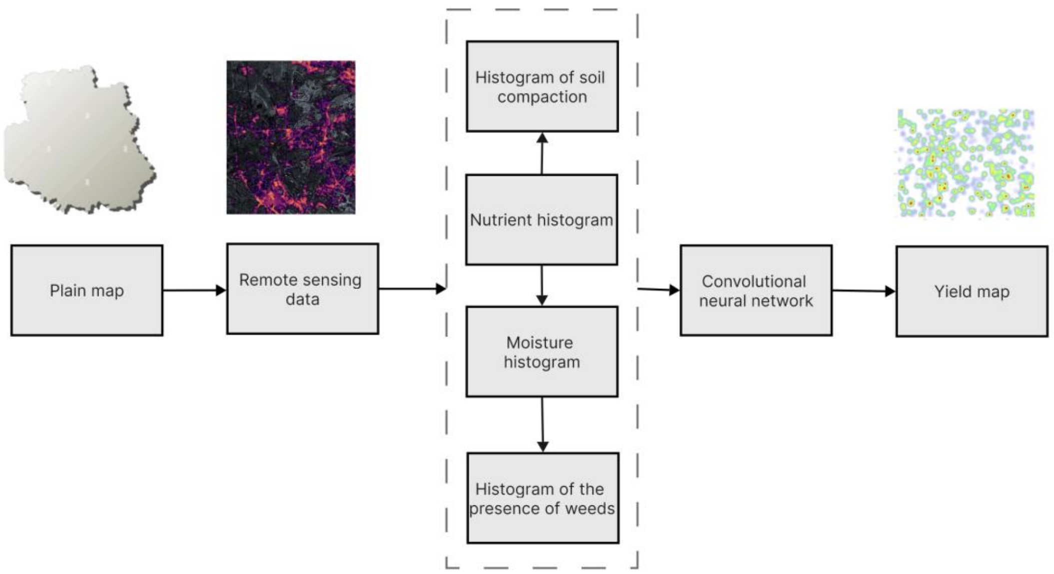

Fig. 2. A list of necessary skills describing a person

NDVI (Normalized Difference Vegetation Index) is presented in the formula 1:

$$NDVI = \frac{(NIR + RED)}{(NIR - RED)} \tag{1}$$

where:

- NIR is the reflectance in the near-infrared spectrum (which is strongly reflected by healthy vegetation).
- RED is the reflectance in the red spectrum (which is absorbed by chlorophyll).

Also, we can use SAVI. It improves vegetation detection in semi-arid regions where the soil has significant reflectance. SAVI adjusts NDVI to account for the influence of soil reflectance, which can impact vegetation indices in areas with sparse vegetation.

SAVI (Soil Adjusted Vegetation Index) is presented in the Formula 2:

$$SAVI = \frac{(NIR - RED)}{(NIR + RED + L)} \times (1 + L) \tag{2}$$

where:

- NIR is the reflectance in the near-infrared spectrum.
- RED is the reflectance in the red spectrum.
- L is a soil brightness correction factor (usually set to 0.5).

For a areas with high biomass we can ise EVI. EVI is an improvement over NDVI, providing better sensitivity in areas with high biomass and correcting for atmospheric influences and background noise from soil.

EVI (Enhanced Vegetation Index) is presented in the Formula 3:

$$EVI = G \times \frac{(NIR - RED)}{(NIR + C_1 \times RED - C_2 \times BLUE + L)} \tag{3}$$

where**:**

- NIR, RD, and BLUE are the reflectance values in the respective bands.
- G is a gain factor (usually set to 2.5).
- C_1 and C_2 are coefficients to correct for atmospheric effects (usually 6 and 7.5).
- L is a canopy background adjustment (typically set to 1).

Data Pre-Processing. The raw data often contains noise and errors caused by atmospheric interference, sensor limitations, or inconsistencies in data acquisition. Data cleaning and preparation are essential to remove noise and outliers from raw data. Techniques such as median filtering, interpolation, and smoothing are applied to sensor data and satellite imagery. The preprocessing stage involves: Radiometric calibration, Geometric correction, Cloud masking.

We can use median filter to smooth NDVI, temperature and moisture. The median filter is a non-linear filter used to reduce noise by replacing each value with the median of its neighboring values. This is particularly effective for removing outliers or spikes from sensor data like NDVI, soil moisture, and temperature.

Median Filter (for noise reduction in sensor data) is presented in the Formula 4:

$$y[i] = median(x[i-k], \dots, x[i+k]) \quad (4)$$

where**:**

- y[i] is the filtered value at index i.
- x[i−k], ..., x[i+k] are the neighboring data points within a window size k.

The simple moving average smooths time-series data by averaging the last nnn values. This can help to reduce short-term fluctuations, making long-term trends more visible.

Simple Moving Average (for smoothing data) is presented in the Formula 5:

$$SMA_n = \frac{1}{n}\sum_{i=0}^{n-1} x_{t-i} \quad (5)$$

where:

- SMA_n is the simple moving average over the last n observations.
- $x_{t-i}$the individual data points at t−i.

Feature Extraction and build histograms. List of histograms: Histogram of Soil Compaction, Nutrient Histogram, Moisture Histogram, Histogram of the Presence of Weeds. After preprocessing, the data is analyzed to extract meaningful features about crop health The next step involves creating a histogram of soil compaction based on remote sensing data or field data from compaction sensors. Compaction influences the ability of plant roots to absorb water and nutrients. We need to calculate the compaction index based on soil penetration resistance measurements. This formula helps quantify soil compaction at various depths.

Histogram of Soil Compaction. Soil Compaction Index (SCI) is presented in the Formula 6:

$$SCI = \frac{Force\ applied\ (N)}{Area\ of\ Penetrometer\ tip\ (cm^2)} \quad (6)$$

Since sensor data is often sparse, we can use Kriging to interpolate soil compaction measurements across the field. The formula is displayed in the formula 7:

$$\hat{Z}\ (s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \quad (7)$$

where:

- $\hat{Z}(s_0)$is the estimated value at the unknown location s0s_0s0.
- $Z(s_i)$are the known values at the sampled locations.
- λ_i are the Kriging weights, which are determined based on the spatial correlation between the known data points.

Nutrient Histogram. Nutrient availability (such as nitrogen, phosphorus, and potassium) is crucial for crop health. Data can be collected via soil sensors or remote sensing data with spectral analysis. Nutrient levels will be visualized as histograms.

Nutrient Index Calculation. If spectral data is available, indices such as the Nitrogen Reflectance Index (NRI) can be calculated by the formula 8:

$$NRI = \frac{Reflectance_R - Reflecyance_G}{Reflectance_R - Reflectance_G} \quad (8)$$

where R and G represent red and green spectral bands.

To model spatial variability in nutrient levels, Gaussian Process Regression (GPR) is a suitable tool. It can account for spatial dependencies and provide a probabilistic estimate of nutrient levels across a field.

$$f\ (x) \sim GP(m(x), k(x, x^*))$$

where m(x) is the mean function and $k(x, x^*)$ is the covariance function between any two input pointsx and x^*.

In case of outlier nutrient readings, statistical outlier detection like Z-score or IQR methods can be used to clean data. A data point is considered an outlier if its Z-score is greater than a certain threshold, typically Z > 3.

Outlier Detection:

$$Z = \frac{x - \mu}{\sigma} \tag{9}$$

where:

- x is the data point.
- μ is the mean of the dataset.
- σ is the standard deviation of the dataset.

Moisture Histogram. Soil moisture data is key to determining water availability for crops. We will collect this data using soil moisture sensors or satellite remote sensing platforms such as SMAP.

Soil Moisture Index (SMI). This index normalizes soil moisture levels between minimum and maximum values.

$$SMI = \frac{Moisture\ Content - Min\ Moisture}{Max\ Moisture - Min\ Mousture} \tag{10}$$

Also, we can apply here Savitzky-Golay Filter. The Savitzky-Golay filter works by fitting a polynomial to a moving window of data points and then evaluating this polynomial at a single point to estimate the smoothed value (Formula 11).

$$\hat{y}(n) = \sum_{k=-m}^{m} c_k\, x\,(n + k) \tag{11}$$

where:

- $\hat{y}(n)$ is the smoothed value at time point n.
- x(n+k) are the original data points within the window, centered at nnn.
- m is half the window size (so the total window is 2m+1).
- c_k are the convolution coefficients derived from a least-squares fit of a polynomial to the data.
-

Histogram of the Presence of Weeds. Remote sensing data, particularly from multi-spectral or hyperspectral imagery, can be used to detect weeds. The vegetation indices differ between crops and weeds due to their different reflectance characteristics. The presence of weeds is represented as a histogram. We can use thresholding techniques on the NDVI or Green Vegetation Index (GVI) to identify areas with weeds.

$$NDVI_{threshold} = 0.2$$

NDVI values lower than the threshold can indicate weed presence, as weeds often have lower NDVI values compared to healthy crops.

For a more robust weed detection approach, Random Forest can be trained using labeled remote sensing data to distinguish between crops and weeds (Formula 12).

$$f(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \tag{13}$$

Where:

- $T_i(x)$is the i-th decision tree
- N is the total number of trees in the forest.

Data Integration with GIS. GIS (Geographic Information Systems) allow for spatial mapping of the extracted features. All data from different sources (remote sensing, field sensors) are georeferenced to ensure proper spatial alignment. Ensuring that sensor data and satellite imagery align with geographic coordinates of the fields. Combine NDVI maps, soil moisture data, nutrient levels, and other features to build a complete field model. Use GIS tools helps to identify spatial patterns in data (e.g., compaction areas, moisture gradients, etc.). GIS helps in tracking crop health at different spatial scales, allowing farmers to visualize problem areas within a field, compare temporal changes, and make decisions on a localized level.

Crop Health Assessment and Analytics. Crop health is evaluated using vegetation indices, moisture levels, and nutrient availability. This step focuses on diagnosing areas of the field that may need attention, such as nutrient deficiencies or high soil compaction.

At this stage we can apply Linear Regression to predict future yield based on extracted features.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

where**:**

- Y is the predicted yield.

- β_0 is the intercept (constant term).
- $\beta_1,\beta_2,\dots\ \beta_n$ are the coefficients corresponding to the features $X_1,X_2,\dots\ X_n$.
- $X_1,X_2,\dots\ X_n$ are the input variables, such as NDVI, soil moisture, temperature, etc.

Linear regression is used to model the relationship between multiple features and the predicted yield. The model assigns weights (coefficients) to each feature to predict the outcome (yield) based on input data.

For example yield preduction due to water stress.

$$Y = Y_{max} \times\ (1 - K_s \times D)$$

where**:**

- Y_{max} is the potential yield without water stress.
- K_s is a crop-specific sensitivity coefficient to water stress.
- D is the water deficit ratio.

This formula models the reduction in crop yield due to water stress. The potential yield is reduced proportionally to the crop's sensitivity to water stress and the severity of the water deficit.

After generating the histograms, we feed them into a convolutional neural network to predict the final yield map. CNNs can automatically learn relevant features from the input histograms (such as moisture, soil compaction, nutrient levels, etc.). The CNN will have multiple layers, including convolutional layers, pooling layers, and fully connected layers. The key idea is to allow the network to learn spatial hierarchies in the input data.

Convolution Operation:

$$y[i,j] = \sum_m \sum_n x[i+m, j+n] \cdot w[m,n]$$

where :

- $x[i,j]$ is the input data
- $w[m,n]$ is the filter (or kernel) applied to the input data.

When training the CNN use backpropagation with the Adam optimizer to minimize the Mean Squared Error (MSE) between predicted yield and actual yield.

MSE Formula (14):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (14)$$

where:

- y_i is the actual yield
- $\hat{y_i}$is the predicted yield.

Evaluation and Testing. To assess the accuracy of the entire system, we will evaluate it using Root-Mean-Square Error (RMSE), cross-validation.

RMSE is presented in the Formula 15:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (15)$$

This metric helps measure the error between predicted and actual yield.

Also we will use k-fold cross-validation and test the model multiple times to ensure robustness and avoid overfitting.

Decision Support Systems (DSS). DSS tools take processed data and provide farmers with actionable insights, such as:

Optimized irrigation schedules based on soil moisture data and weather forecast, Fertilizer application rates, tailored to soil nutrient deficiencies detected through remote sensing, Pest or disease outbreak warnings, allowing farmers to take preventive action before damage occurs.

DSS also provides variable rate application (VRA) maps, which help farmers apply inputs (e.g., fertilizers, pesticides) in variable amounts across different sections of the field based on specific needs.

Visualization and Reporting. Once the analysis is complete, visualization tools are used to present data in an intuitive and actionable format. Heatmaps, yield maps, and histograms provide a visual representation of crop health, soil conditions, and predicted yield. The analyzed data and recommendations are visualized in user-friendly formats, such as yield maps, heat maps, or 3D field models. These visualizations help farmers understand the condition of their fields at a glance. Data is shared with farmers via mobile applications or cloud platforms, allowing for real-time monitoring and decision-making. Farmers can use this information to implement precision agriculture practices, such as targeted irrigation or nutrient management, which optimize input use and maximize yields.

At this stage we can calculate yield Increase for farmers and visualise it.

Formula for Yield Increase Based on Crop Health Indicators:

$$\Delta Y = a_0 + a_1 \times \Delta NDVI + a_2 \times \Delta Moisture + a_3 \times \Delta Nutrients + a_4 \times \Delta Temperature + \cdots + a_n \times \Delta Feature_n$$

where**:**

- ΔY is the predicted change in yield (yield increase or decrease).

-a_0 is the intercept term or baseline yield (without any improvements).

-$a_1, a_2, \ldots, a_n$ are the coefficients (weights) for each crop health indicator, representing how strongly each indicator affects the yield.

-$\Delta NDVI$ is the change in the Normalized Difference Vegetation Index.

-$\Delta Moisture$ is the change in soil moisture levels (e.g., from soil moisture sensors).

-$\Delta Nutrients$ is the change in nutrient levels (e.g., nitrogen, phosphorus, potassium).

-$\Delta Temperature$ is the change in temperature (e.g., average or extreme temperatures affecting crop growth).

-$\Delta Feature$ are other factors that can affect yield, such as soil pH, humidity, crop-specific stress factors, weed presence, or any additional data from sensors or indices.

Feedback Loop. The final step is closing the feedback loop. After each growing season, the actual outcomes (yields, crop health) are compared to the model's predictions. This helps improve the system over time through machine learning techniques like CNN and parameter tuning.

Experiments and Results

For expected results we need to build charts , that will help demonstrate how the system can enhance decision-making by providing insights into crop health, yield predictions, resource management, and more. Below are some suggested charts, along with detailed descriptions of their purpose and design.

The first chart will be - Yield Prediction vs. Actual Yield (Figure 3).

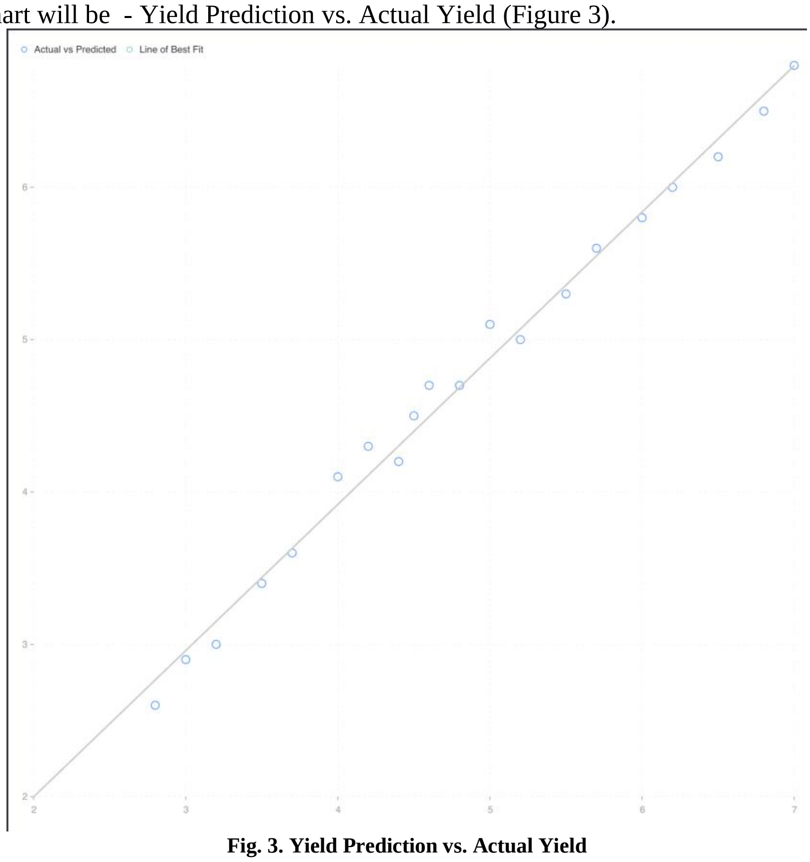

Fig. 3. Yield Prediction vs. Actual Yield

This chart shows the predicted yield from your system compared to the actual yield collected from the field. A scatter plot is used to represent each field or data point, with a line of best fit showing the overall accuracy of the predictions.

The purpose is: to visually demonstrate the accuracy of the system's yield prediction model, it highlights the system's ability to predict crop yield based on factors like NDVI, soil moisture, nutrients, etc, a well-fitting line shows that the model's predictions are closely aligned with actual yield, thus validating its use for decision-making.

The next chart will be Crop Health Heatmap (Figure 4).

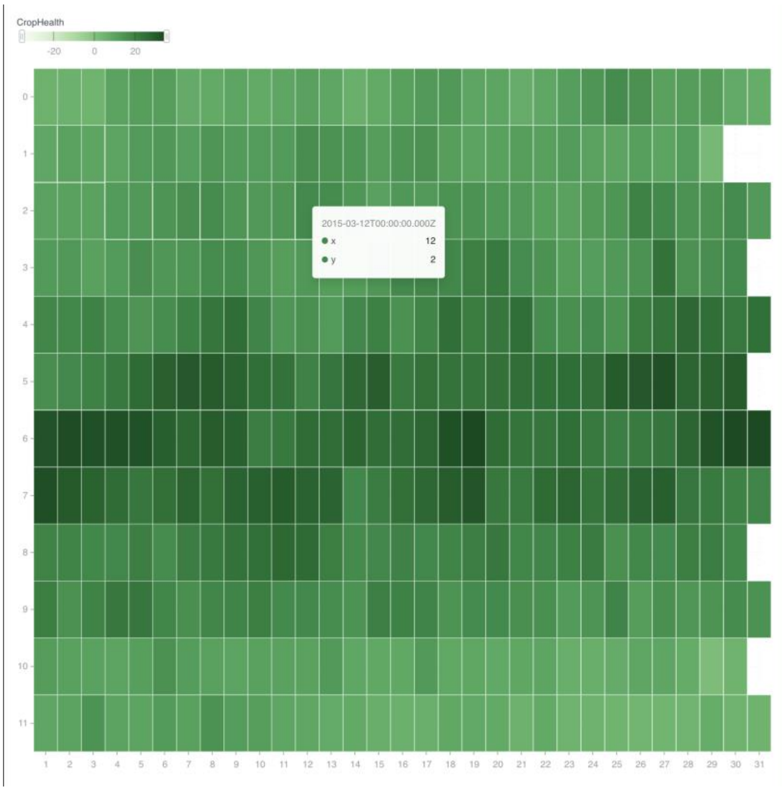

Fig. 4. Crop Health Heatmap (NDVI Distribution Over Time)

A heatmap that visualizes the changes in NDVI (Normalized Difference Vegetation Index) across time, providing a clear view of crop health. The color gradient will indicate areas of the field with varying levels of vegetation health (from poor to excellent). The purpose is: to show how crop health evolves over time based on NDVI readings, helps farmers identify specific areas of the field that may require attention (e.g., irrigation, fertilization).
This chart is useful for decision-making related to resource allocation (e.g., identifying high-stress zones).

Water Use Efficiency (WUE) is presented in the Figure 5.

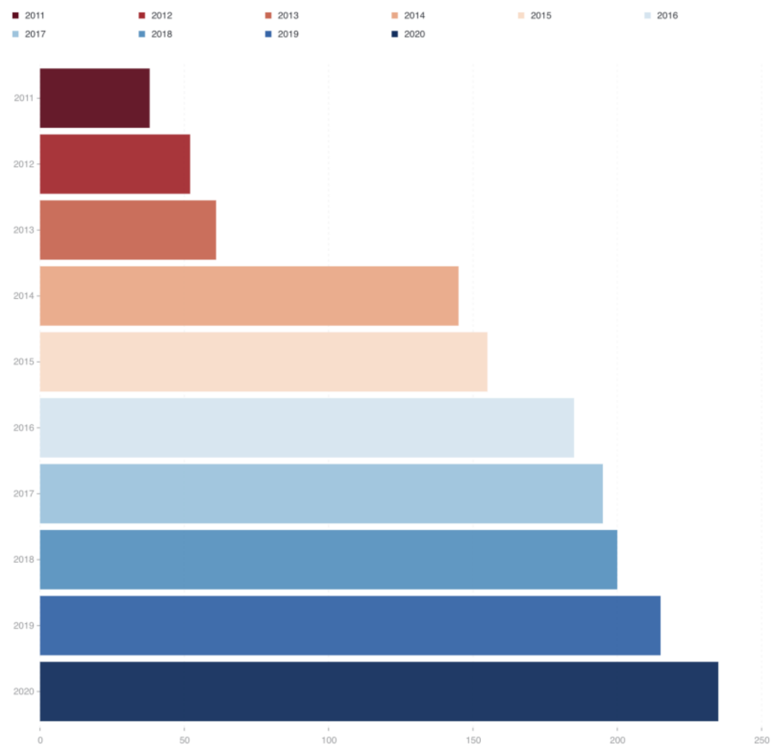

Fig. 5. Water Use Efficiency

A bar chart comparing the Water Use Efficiency (WUE) before and after implementing the decision-support system. WUE is calculated as the yield per unit of water used.

The Purpose is: To visually highlight how the system optimizes water usage, Demonstrates that farmers can achieve higher yields with less water after using the system for irrigation decision-making, Helps in promoting sustainable farming practices by showing improvements in resource efficiency.

The next chart will be Fertilizer Use Efficiency (FUE) (Figure 6)

Fig. 6. Fertilizer Use Efficiency

This chart that shows how Fertilizer Use Efficiency (FUE) changes over time. FUE measures the yield increase per unit of fertilizer applied and helps monitor how efficiently fertilizer inputs are being used.

The purpose is: To show how the decision-making system optimizes fertilizer use over time,Demonstrates that farmers can achieve higher yields with targeted and efficient fertilizer application based on sensor and remote sensing data,A rising trend would indicate improved fertilizer efficiency.

The next chart is Soil Moisture Histogram **(**Figure 7, 8)

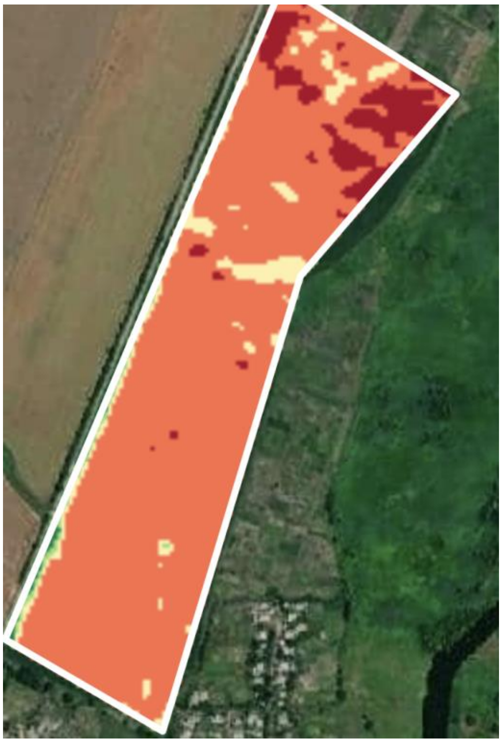

Fig. 7. Soil Moisture Histogram (Before System Use)

Fig. 8. Soil Moisture Histogram (After System Use)

A chart showing the distribution of soil moisture levels before and after using the decision-support system. The chart will visualize how soil moisture is more evenly distributed after the system's recommendations for irrigation.

The Purpose is:To show how the system optimizes soil moisture distribution across a field,Demonstrates the system's impact on balancing irrigation to avoid over- or under-watering,Can help justify the system's value in improving water management for farmers.

The next chart is Resource Allocation Map (Figure 9)

Fig. 9. Resource Allocation Map (GIS Map with Variable Rate Application Zones)

A GIS map showing different zones within the field that require variable amounts of resources (e.g., water, fertilizer) based on the system's decision-making outputs. Each zone is color-coded according to the recommended input level (low, medium, high)

The purpose is :To visually demonstrate how the system allocates resources efficiently based on crop health and soil conditions, Helps in planning variable rate application (VRA) of inputs such as water and fertilizer, reducing waste and increasing yield, Provides farmers with a clear visual guide for applying resources in different areas of their field.

The next chart is Yield Increase Indicator (Figure 10)

Fig. 10. Yield Increase (Yield Improvement)

This chart comparing the crop yield before and after implementing the decision-support system. Each bar represents the yield for a given field or crop type, showing the percentage increase after using the system.

The purpose is: To clearly demonstrate the yield increase that farmers can expect after using the system, Highlights the impact of the system on improving overall crop productivity, Useful for convincing stakeholders of the system's effectiveness.

These charts will help visually demonstrate how the decision-support system benefits farmers by optimizing resource use, improving crop health, increasing yield, and enhancing predictive accuracy. Each chart provides a clear view of the system's impact and helps in justifying its adoption for better decision-making in agriculture.

Conclusions

In summary, the crop monitoring workflow begins with remote sensing and GIS integration, offering invaluable insights into crop health and field variability. By utilizing vegetation indices such as NDVI, SAVI, and EVI, the system provides an accurate assessment of crop vigor, enabling farmers to monitor changes in real time. The integration of machine learning models further enhances this process, allowing for the precise prediction of crop yields, the detection of pest infestations, and the optimization of resource allocation based on data-driven insights.

The workflow's decision-making capabilities extend beyond monitoring, by offering actionable recommendations through Decision Support Systems (DSS). These systems advise on optimal irrigation schedules, fertilizer application rates, and pest control measures, leading to increased productivity while minimizing water and fertilizer waste. The improvements in Water Use Efficiency (WUE) and Fertilizer Use Efficiency (FUE) clearly demonstrate the system's ability to enhance resource utilization, leading to higher yields with less environmental impact.

The integration of feedback loops into the system allows for continuous refinement of management practices. By incorporating real-time data and analyzing historical performance, farmers can iteratively improve crop management strategies, ensuring that precision agriculture is not only effective but also sustainable in the long term. This approach aligns with the growing need for sustainable farming practices that address the challenges of climate change, resource scarcity, and food security.

In conclusion, this crop monitoring and decision-support system equips farmers with the tools necessary to make informed, data-driven decisions, which ultimately increases yields, improves resource efficiency, and reduces the environmental footprint of agriculture. The adoption of such systems ensures that modern agriculture can meet the demands of a growing population while preserving the ecological balance.

References

1. How 'Glowing' plants could help scientists predict flash drough. URL: https://www.nasa.gov/earth/how-glowing-plants-could-help-scientists-predict-flash-drought/ (Last accessed October 17, 2024).
2. Food and Agriculture Organization of the United Nations. Crops and livestock products. URL: https://www.fao.org/faostat/en/#data/QCL (Last accessed October 17, 2024).
3. Food and Agriculture Organization of the United Nations. Food lose and waste database. URL: https://www.fao.org/platform-food-loss-waste/flw-data/en/ (Last accessed October 17, 2024).
4. Food and Agriculture Organization of the United Nations. Land use. URL: https://www.fao.org/faostat/en/#data/RL (Last accessed October 17, 2024).
5. Combining remote sensing data and machine learning to predict crop yield. Stanford. URL: https://sustain.stanford.edu/crop-yield-analysis (Last accessed October 17, 2024).
6. Google Earths engine. URL: https://code.earthengine.google.com/ (Last accessed October 17, 2024).
7. National Agricultural Statistics Service. URL: https://quickstats.nass.usda.gov/ (Last accessed October 17, 2024).
8. Applications of GIS in Agriculture for Predictability & Sustainable Growth. URL: https://intellias.com/gis-in-agriculture/ (Last accessed October 17, 2024).
9. MIDOPT TRAINING. Optical system. URL: https://midopt.com/training/ (Last accessed October 17, 2024).
10. Food systems transition: the key to solve the Climate crises. Agrosmart. URL: https://agrosmart.com.br/eng/ (Last accessed October 17, 2024).
11. Digital farming. URL: https://agriculture.basf.com/global/en/business-areas/digital-farming.html (Last accessed October 17, 2024).
12. Public documents, results and outcomes of the FIELDS project. URL: https://www.erasmus-fields.eu/documents/ (Last accessed October 17, 2024).
13. Crop yield Prediction with Deep Learning. URL: https://github.com/JiaxuanYou/crop_yield_prediction?tab=readme-ov-file#crop-yield-prediction-with-deep-learning (Last accessed October 17, 2024).
14. Vegetation Indices to Drive Digital Agri Solutions. URL: https://eos.com/blog/vegetation-indices/ (Last accessed October 17, 2024).
15. KOMPSAT 3/3A Satellite Images. URL: https://eos.com/find-satellite/kompsat-3-3a/ (Last accessed October 17, 2024).
16. The Platform for 3D Geospatial. URL: https://cesium.com/platform/cesiumjs/ (Last accessed October 17, 2024).
17. Landviewer. URL: https://eos.com/landviewer/ (Last accessed October 17, 2024).
18. Martin Kleppmann.Designing Data-Intensive Applications/ Martin Kleppmann, 2017. – 35 p.
19. Climate change and global food security. - Resource: https://science.org/content/article/food-and-farming-could-stymie-climate-efforts-researchers-say
20. P.P. Lopatin, A.O. Babenko, V.I. Shevchenko, O.V. Moroz, V.M. Glukhov, O.M. Kovalchuk. Use of Earth remote sensing data to assess the condition of winter rapeseed crops in the conditions of the Forest Steppe of Ukraine. Biological systems.
21. Mark Sabini, Gili Rusak and Brad Ross. Understanding Satellite-Imagery-Based Crop Yield Predictions. Stanford University, 2017. 4p.
22. Filed analytics. URL: https://www0.efcsystems.com/cgi-bin/login.cgi (Last accessed October 17, 2024).
23. ARSET - Satellite Remote Sensing for Agricultural Applications. URL: https://appliedsciences.nasa.gov/get-involved/training/english/arset-satellite-remote-sensing-agricultural-applications (Last accessed October 17, 2024).
24. Steklov V.K. Information System: Textbook for Students of Higher Educational Institutions in the Direction of "Telecommunications". V.K. Steklov, L. Berckman. K. Tekhnika, 2014. 792 p.
25. Erasmus Guide. URL: https://erasmus-plus.ec.europa.eu/erasmus-programme-guide (Last accessed October 17, 2024).

Dmytro Okrushko	Master degree student of Computer Engineering & Information Systems Department, Khmelnytskyi National University, Khmelnytskyi, Ukraine https://orcid.org/0009-0007-6496-6436 e-mail: okrdima@gmail.com	
Olga Pavlova	PhD, Associate professor of Computer Engineering & Information Systems Department, Khmelnytskyi National University, Khmelnytskyi, Ukraine https://orcid.org/0000-0001-7019-0354 e-mail: pavlovao@khmnu.edu.ua	