2021 IEEE Smart Information Systems and Technologies (SIST)
28-30 April, 2021, Nur-Sultan, Kazakhstan

A Conceptual Research Model of Development of the Geographic Information System for Agriculture

Mingxin Huang[1*], Andrii Biloshchytskyi[2], Yurii Andrashko[3], Serik Omirbayev[4]
[1*] Taras Shevchenko National University of Kyiv, Ukraine, bao1978@gmail.com
[2]Astana IT University, Taras Shevchenko National University of Kyiv, Kazakhstan, Andrii.Biloshchytskyi@astanait.edu.kz
[3]Uzhhorod National University, Ukraine, yurii.andrashko@uzhnu.edu.ua
[4]Astana IT University, Kazakhstan, serik.omirbayev@astanait.edu.kz

***Abstract* — The development of geographic information systems is an important prerequisite for implementing information systems in agriculture. The work is devoted to the Conceptual Research Model of the Development of the Geographic Information System for Agriculture. The model reflects four stages: goal setting, functionality description, implementation, and diagnosis. The paper describes the objects and subjects of the environment in which the geo-information system operates. A mathematical model for describing the problem of crop yield monitoring is also proposed. The model is based on the analysis of the cropped image and the use of time series analysis.**

***Keywords* — building enterprise; multiobjective optimization**

I. INTRODUCTION

In recent decades, the use of digital images of geographical areas for agricultural purposes has become relevant. This area's rapid development is facilitated by the development of wireless technology, digital photography technology, image display devices, operating systems, and services for storing and processing these images. These opportunities have become a prerequisite for the intensive development of geographic information systems and technologies in general.

The development of geographic information systems and technologies is associated with the improvement of imaging and image processing technologies, the emergence of high-quality maps, such as Google Maps, increasing data rates, and developing a new direction - the Internet of Things, which has direct application in agriculture. A description of the evolution of geoinformation technologies was made in [1].

Paper [2] describes the hypothesis that inhomogeneities can directly determine crop yields in field images. The processing of digital images of fields provides valuable information about the state of crops, allows you to assess plant health, and predict the yield, timing, quantity, and quality of products in the future.

The image of sown areas can be presented in the form of time series, allowing them to apply appropriate forecasting methods [3-5]. The works [6-8] describe the methods and technologies of processing time series of digital images for agriculture decision-making.

For effective management to obtain a higher yield, forecast the harvest price, prompt elimination of plant diseases, information is needed: on soil quality, the need to apply fertilizers to certain agricultural areas, plant morbidity, analysis of flowering, etc.

Today, the popular concept is based on SSCM (Site Specific Crop Management). An important task of SSCM is to increase yields, i.e., to ensure the quality and quantity of agricultural products. SSCM is implemented based on GPS technology, which determines the location of objects and the condition of plants on agricultural land through a network of satellites. RTK (Real-Time Kinematic) technology allows you to identify objects in the area with great accuracy. After collecting information on the state of the soil, SSCM technology allows you to control the use of fertilizers, chemicals, herbicides, pesticides, and other substances. This technology is relevant not only for Ukraine but also for China, wherein some regions and farms produce crops several times a year. All this requires prompt management of the ripening process of crops. The equipment that uses the GPS module makes it possible to cultivate large areas quickly and minimize human impact.

Consider the leading geographic information systems that exist today: CMaps Analytics by Centigon Solutions, Maptitude by Caliper, Map Business Online by MapBusinessOnline.com, Rosmiman IWMS by Rosmiman Software Corporation, and special noteworthy ArcGIS by ESRI,

CMaps Analytics [9] is a system that positions itself as a convenient visualization tool based on electronic Google Maps. The system is a web service, so it is accessible through the browser of any device. The system can work with built maps in other systems, such as ArcGIS. The software package's main feature is a large number of applications and extensions for different products, depending on needs. Access to the resource is paid, the price is determined individually during negotiations with the client.

Quantum GIS [10] is an open-source graphical information system distributed under the General Public License. Quantum GIS is an Open-Source Geospatial Foundation project that runs on Linux, Unix, Mac OSX, Windows, Android, supports many vector and raster formats, databases, and has extensive analysis and presentation capabilities geospatial data. There is some set of raster image analysis. However, the main emphasis is on working with raster and vector layers, databases, and various types of data, such as Oracle GeoRaster, Oracle Spatial, WMS \ WMTS (Web Map Service), WCS (Web Coverage Service), WFS (Web Feature Service), CSV (Comma-Separated Values), etc.

ArcGIS [11] is the most popular geographic information system. Forms a whole set of software applications allows you to convert raw data using advanced analysis and highly specialized applications. Helps to integrate all types of environmental data and apply analytical methods. Presents results in the form of maps and graphs. The system supports the following data sets: spatial data, bitmaps, and tables. There are specialized methods for analyzing sown areas: soil analysis, collection of information on yield, construction of the history of the yield of the zone. Based on many factors, recommendations for soil improvement are determined. The software is paid.

II. A CONCEPTUAL

We will consider the conceptual research model of a geographic information system for agriculture as a way to describe it. The main task of this model is to display the features and key characteristics of the system. The model also answers the question of developing a geographic information system for the agriculture system and the expected results of its operation.

In developing the research model, the authors followed the sequence of application of scientific approaches. According to the conceptual approach, the statement is formulated in the first step, the main direction, course, and architecture of the research are determined. The primary reference points of information system development are defined. As a result of the conceptual approach, the conceptual model [12] is received.

Next, it is applied to the aspect approach. As a result, the most critical aspects of the system are highlighted. In [13], the peculiarities of using image analysis to determine crop yields are highlighted.

Further research is proposed to continue using a systems approach. The main purpose is to determine the nature of the links between the elements of the system.

The process of developing a geographic information system for agriculture is proposed to be divided into four stages.

1. The stage of determining the goal. At this stage, it is necessary to explore the environment where the information system will operate. Select all objects and entities in this environment. Formulate the requirements of the final product. Determine the results expected from the launch of the information system and its benefits.

Crop yield management is part of the management tasks of an agricultural enterprise, which requires the application of new concepts of project management and programs

2. Stage of the description of functionality. This stage is to formalize the research task. Development of a mathematical model to describe the problem. They are choosing methods that best solve the research problem.

3. Implementation stage. It is necessary to implement the methods and models obtained at the previous step in the form of software. Also, the features of the state task require the use of specialized hardware. Therefore, at the implementation stage, it is necessary to configure it.

4. Stage of diagnosis. At this stage, it is necessary to identify criteria and quantitative indicators by which to evaluate the results of the geographic information system for agriculture. Also, in the fourth stage, the system is tested.

They considered four stages are presented as four components of the conceptual research model (See Pic. 1). Consider each element in the model in more detail.

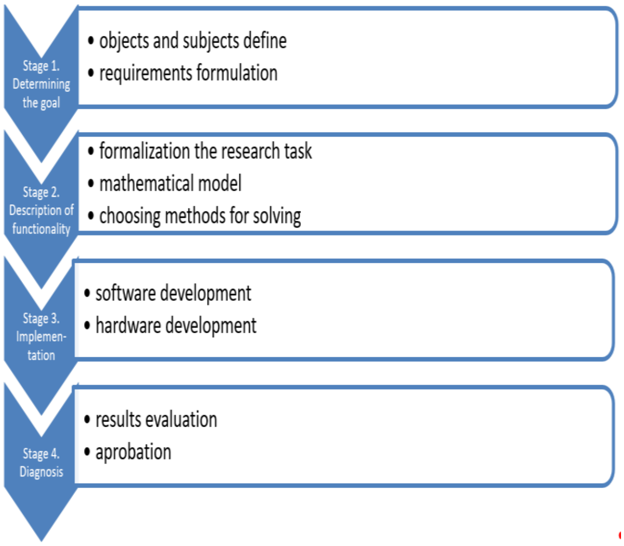


Figure 2. Conceptual research model of development of the geographic information system for agriculture

At the first stage, it is proposed to formulate the purpose of the geographic information system for agriculture as a collection and presentation of information necessary for management decisions. In particular, for the management of sown areas, this may be information about:

1. plant health;
2. plant needs for fertilizers, watering, etc .;
3. forecast of crop yields that can be obtained in the current year;
4. readiness of the crop for harvesting, digging, and other activities;
5. signs of pest spread;
6. damage from flooding, hail, wind, or other emergencies.

Essential requirements that geographic information systems must meet are:

1. the ability to use new maps of territories and update old maps;
2. the ability to save and access digital images of the region and their comparison with plans;
3. integrated with data collection, storage, processing, and visualization systems.
4. cross-platform system.

Also, at the first stage, the following subjects and objects of the system environment were identified:

1. The territory on the Earth's surface. The area is the main object of observation. The region is divided into a finite set of sections. Each is considered a valuable resource for user activity. The site has a specific set of characteristics. The result of action on the specified site depends on the concrete value of these characteristics at each moment. The features of the site are dynamic values, i.e., they change over time. Its location on the territory determines the features of each section. Influence on the significance of the land plot's characteristics, ecological, biological factors, and results of activity on the site previous periods. Estimating land characteristics values is usually known as a priori and can be obtained from the cadastral register, agricultural databases, etc.
2. The user of the land plot. It can be an agricultural enterprise, an individual, or the state. This facility uses land and other resources to generate food.
3. Environmental monitoring systems. For these purposes, you can use drones, which are equipped with a specialized camera to create images of maps of the state of crops, considering the infrared reflection. Remote sensing using aircraft is also used.

At the second stage of information system development, it is necessary to define mathematical models and methods that ensure the information system's functioning. Consider a mathematical model for monitoring crop yields.

Yield monitoring will be understood as a system of monitoring and measuring crops' growth, considering meteorological, agrometeorological, phenological, and other indicators based on the analysis of time series images obtained by photographing sown areas to assess and predict potential crop yields. It is also envisaged to consider the uneven yield on different areas of sown areas, which can be visualized with the help of yield maps. Different colors indicate the volume of production after harvest. It should also be borne in mind that yield monitoring should involve a two-dimensional assessment: the number of products that will be potentially obtained and the quality of these products on an appropriate scale. To ensure effective monitoring at all levels, it is necessary to provide high-quality aerospace images for analysis, stored and analyzed in time series.

The main phenological indicator in the analysis of images of sown areas is the normalized differential vegetation index. The vegetation index is a quantitative indicator of active photosynthetic biomass used to quantify vegetation, which is calculated by the formula:

$$NDVI = \frac{N - R}{N + R},$$

R is the intensity of the light reflected from the spectrum's surface in the visible red region (630-690 nm), N is the intensity of the light reflected from the surface of the spectrum in the infrared region of the spectrum (760-900 nm).

The input data for the information system is the image of the site. The image can be represented as some size matrix. The image shows culture on certain pixels. Others show a soil that is not culture. Each image's intensities have an intensity distribution close to the normal distribution (See fig.2). Therefore, a threshold function can be introduced

$$\delta(x, y) = \begin{cases} 1, if \ \beta(x, y) \geq \mathrm{B} \\ 0, if \ \beta(x, y) < \mathrm{B} \end{cases},$$

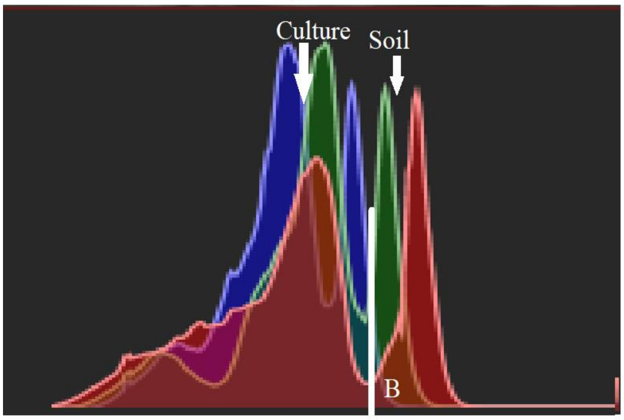
where x, y are the coordinates of the pixel in the image, β is the value of the NDVI in a particular pixel, and B is some threshold value. Marking

Then the yield of the site is described by the following model:

$$\beta(x, y) = \delta(x, y)\, p(x, y) + \left(1 - \delta(x, y)\right) \overline{p}(x, y), \quad (1)$$

where o the distribution density of the culture in the image, r the distribution density of the culture. Yield can be estimated by finding a two-dimensional integral of crop density.

a)



b)

Figure 3. a)The image of the field with growing corp retrieved from Google maps b) Color histogram

The forecasting model based on phenological indicators is based on the definition of critical points in the NDVI trajectory. Sequential measurements of NDVI for one site at specified intervals (day, week, two weeks, etc.) form a discrete-time series of estimates

$$\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$$

Using prediction methods for time series [14-19], predict NDVI values for subsequent periods. And using the model of yield forecasting (1) - to estimate future yields.

In the third stage, specialized mathematical and statistical packages and tools MatLab [20], R [21] environment, as well as software for image processing GRASS [22], IDRIDI [23], ENVI can be used to analyze the images of agricultural lands for yield monitoring [24]. Also, it is not enough to use traditional databases such as MySQL or PostgreSQL to store information, so you need to use special spatial databases such as MLPQ / GIS [25-29].

At the fourth stage, it is necessary to determine the criteria and indicators of the functioning of the geographic information system. These indicators should make it possible to quantify the performance. It is a diagnostic tool that allows you to evaluate the usefulness of GIS in decision-making and improve it in future seasons. This issue needs a more in-depth study.

III. CONCLUSIONS

The article describes the known methods of multi, as a result of research built Conceptual Research Model of developing the geographic information system for agriculture. This model includes four stages that completely determine the outline of the study. In the first stage, the objects and subjects in the environment where the geographic information system operates are identified. The requirements for the final product are formulated, and the expected results are determined. At the second stage, the mathematical model of estimation and forecasting of productivity based on images' analysis is constructed. According to the third stage, the software is being developed. Determining the criteria and indicators of the geographic information system's functioning is the subject of further research.

REFERENCES

[1] F. Rembold, C. Atzberger, I. Savin, and O. Rojas, "Using low resolution satellite imagery for yield prediction and yield anomaly detection," Remote Sens, 5 (4), pp. 1704-1733, 2013.

[2] A. DeJoia, M. Duncan, "What is «Precision Agriculture and why is itimportant," 2015, [Online]. Available: https://soilsmatter. Wordpress.com/2015/02/27/what-is-precision-agriculture-and-why- is-it-important/.

[3] A. Kuchansky, A. Biloshchytskyi, "Forecasting time series by selective comparison with a sample," Eastern-European Journal of Enterprise Technologies, 6/4 (78), pp. 13-18, 2015. [ukrainian]

[4] A. Kuchansky, and V. Nikolenko, "Forecasting time series by the method of comparison with the sample," management of development of complex systems, 22, pp. 101-106, 2015. [ukrainian]

[5] A. Kuchansky, A. Biloshchytskyi, Yu. Andrashko, S. Biloshchytska, Ye. Shabala, and O. Myronov, "Development of adaptive combined models for predicting time series based on similarity identification," Eastern-European Journal of Enterprise Technologies, ¼ (91), pp. 32–42, 2018. Doi: 10.15587/1729-4061.2018.121620.

[6] F. Petitjean, J. Inglada, and P. Gançarski, "Satellite image time series analysis under time warping," IEEE Trans. Geosci. Remote Sens, 50(8), pp. 3081-3095, 2012.

[7] H. Eerens, D. Haesen, F. Rembold, F. Urbano, C. Tote, and L. Bydekerke, "Image time series processing for agriculture monitoring," Environmental Modelling and Software, 53, pp. 154–162, 2014.

[8] Drone data collection and analytics for agriculture. Quantify plant and soil health, improve productivity and maximize field output. [Online]. Available: https://www.precisionhawk.com/agriculture.

[9] Centigon Solutions Cmaps Analytics. (2016) [Online]. Available: http://cmapsanalytics.com/

[10] QGIS The Leading Open Source Desktop GIS. [Online]. Available: http://www.qgis.org/

[11] ArcGIS Online. [Online]. Available:https://www.arcgis.com/index.html

[12] M. Huang , Y. Y. Shabala, “Conceptual model of geogerphic information system for agriculture,” Scientific Bulletin of Uzhhorod University. Series of Mathematics and Informatics, 2(35), pp. 149-155, 2019. Doi: https://doi.org/10.24144/2616-7700.2019.2(35).149-155

[13] M. Huang, and V. Vatskel, “Digital image analysis technologies for decision support systems in agricultural,” management of development of complex systems, 37, pp.164 – 167, 2019. https://doi.org/10.6084/m9.figshare.9783227

[14] A. Kuchansky, V. Nikolenko, and A. Rachenko, “Methods for identifying trends in financial time series based on trend forecasting models,” management of development of complex systems, 24, pp. 84-89, 2015. [ukrainian]

[15] A. Kuchansky, and A. Biloshchytskyi, “Selective pattern matching method for time-series forecasting,” Eastern-European Journal of Enterprise Technologies, 6(4(78)), pp. 13-18, 2015. Doi: 10.15587/1729-4061.2015.54812

[16] A. Kuchansky, A. Biloshchytskyi, Y. Andrashko, S. Biloshchytska, T. Honcharenko and V. Nikolenko, “Fractal Time Series Analysis in Non-Stationary Environment,” 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 2019, pp. 236-240, doi: 10.1109/PICST47496.2019.9061554.

[17] O. Berzlev, “Methods of pre-forecast fractal analysis of time series,” Management of development of complex systems, 16, pp. 76-81, 2013. [ukrainian]

[18] O. Berzlev, “The current state of information systems for forecasting time series,” Management of development of complex systems, pp. 78-82, 2013. [ukrainian]

[19] A. Biloshchytskyi, “Vector method of goal setting of projects in the design vector space,” management of development of complex systems, 11, pp. 110 – 114, 2012. [russian]

[20] MATLAB per l’intelligenza artificiale Progettare modelli e sistemi guidati dall’AI. [Online]. Available: http://www.mathworks.it/

[21] The R Project for Statistical Computing. [Online]. Available: http://www.r-project

[22] GRASS GIS. Bringing advanced geospatial technologies to the world. [Online]. Available: http://grass.osgeo.org/

[23] Analytics. Insights. Answers you can trust. [Online]. Available: http://www.exelisvis.com/

[24] Features TerrSet 2020 Software. [Online]. Available: http://clarklabs.org/

[25] JRC MARS Explorer. Crop and weather monitoring in Europe. [Online]. Available: http://www.marsop.info/marsop3/

[26] P. Revesz, R.Chen, P. Kanjamala, Y. Li, Y. Liu, and Y. Wang, “The MLPQ/GIS constraint database system,” In Proceedings of the 2000 ACM SIGMOD international conference on management of data, 601 p., 2000.

[27] V. Morozov, O. Kalnichenko, M. Proskurin, and M. Mezentseva, “Investigation of Forecasting Methods of the State of Complex IT-Projects with the Use of Deep Learning Neural Networks,” International Scientific Conference “Intellectual Systems of Decision Making and Problem of Computational Intelligence”, pp. 261-280, 2019.

[28] V. Morozov, O. Kalnichenko, and I. Liubyma, “Projects change management in based on the projects configuration management for developing complex projects,” 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), pp. 939–942, 2017.

[29] Kuchansky et al., “Combined Models for Forecasting the Air Pollution Level in Infocommunication Systems for the Environment State Monitoring,” 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), Lviv, 2018, pp. 125-130, doi: 10.1109/IDAACS-SWS.2018.8525608.