Development of a software module for the analysis of the state of the forest using remote sensing data of the Earth

T Tatarnikova, S Stepanov,Y Petrov* and A Sidorenko

Department of applied computer science, *Russian State Hydrometeorological University*, 79 Voronezhskaya Street, St. Petersburg 192007, Russian Federation

*Corresponding email:**yaroslav.petrov025@gmail.com**

Abstract. The scientific work examines the process of designing and implementing the software module of the QGIS geographic information system for analyzing the state of the forest on the basis of a series of index images of different times. The module is successfully used as a separate tool and can be integrated into automated models of a more complex level. The scientific results serve as an example of applying an integrated approach to the processing of remote sensing data to detect violations of the integrity of the forest cover, with the aim of rapid response and strategic management.

1. Introduction

The study of the dynamics of the variability of natural phenomena and processes, operational monitoring and satellite monitoring of natural resources, analysis and forecasting of the likely consequences, determining ways to prevent emergency situations, all this is an integral feature of the methodology for collecting and processing geospatial information about the state of the study area, necessary for making timely and strategically correct management decisions [1].

When using hardware and software systems for monitoring, in combination with remote sensing data, it is possible to identify areas of soil and water pollution, negative changes in forestry and agriculture, illegal logging and waste storage sites, and unauthorized places for the extraction of natural resources. Space monitoring allows a retrospective analysis of the study area, aimed at photographing the beginning and end of various events.

Modern geographic information systems (GRASS, QGIS, ArcGIS, SAGA, ENVI, Erdas Imagine, gvSIG, etc.) have great functionality [2] for processing satellite data from remote sensing and allow solving the tasks of regular space-based environmental monitoring.

The main problem when using the built-in GIS functions is in certain scenarios when working with data, each geographic information system has its own set of built-in procedures for processing and analyzing data. To solve specific, not standard tasks, the built-in GIS functionality is often not enough. The functionality of most open geographic information systems is significantly inferior to proprietary products. Using our own application modules for analysis and calculations, modeling, planning and other procedures, in accordance with the specific requirements of the application, it becomes possible to significantly refine the GIS functionality, which allows for the efficient processing of geodata according to a given algorithm.

During of the scientific work, a module of the geographic information system QGIS was developed to support decision-making using geospatial data to conduct an operational analysis of the state of the territories. The practical value of the work is to increase the efficiency of remote sensing data

management due to the comprehensive optimization of data processing methods in the geographic information system. The results of scientific work provide an opportunity to improve the processing of remote sensing data.

2. Methods and Materials

The theoretical basis of the study is the work of domestic and foreign scientists on the issue under consideration, such as N V Zhukovskaya, A G Karmanov, A I Knyshev, V V Eliseeva, A V Koshkarev, M Yu Kataev and others. The results of scientific research, regulatory documents related to the operational management of territories, the development of regions and the problems of using remote sensing data of the Earth in geographical systems for solving problems of operational monitoring of the state of the territories are analyzed.

The methodological basis of the scientific work is constituted by the methods of system analysis, algorithmization, modeling, programming and preliminary processing of satellite images - conversion of brightness values and atmospheric data correction, and channel processing methods for creating "index" images.

In scientific work, data processing from the Landsat-8 [3] satellite was carried out. The spacecraft collects [3] valuable geospatial data and images used in agriculture, education, business, science and government. The following satellite spectral channels are used in scientific work: B2 (Blue), B3 (Green), B4 (Red), B5 (NIR), B6 (SWIR2), B7 (SWIR3), B9 (SWIR1).

In most cases, Landsat images are delivered with geometric and radiometric corrections; the archive name indicates the L1TP data processing level, an example of the name of the downloaded geodata archive is: LC08_L1TP_xxxxxx_xxxxxxxx_xxxxxxxx_01_T1_B4. If Landsat images are delivered in raw form - "raw", the processing level L1GT is indicated in the archive name, and therefore, before working with the data, it is necessary to correct them [4].

Landsat L1GT data can be recalculated to reflect the upper atmospheric reflectance and / or radiation (TOA) using the radiometric scaling factors provided in the metadata file that comes with the Level 1 product. The metadata file also contains the thermal constants needed to convert the data thermal range to brightness temperature TOA. Formulas for transformations (1.1 and 1.2) [5, 6]:

$$p\lambda' = M_p * Q_{\text{cal}} + A_p \tag{1.1}$$

where: $p\lambda'$ – the value of the upper atmospheric planetary reflection of radiation (TOA reflectance), excluding correction for the angle of incidence and reflection of sunlight; M — channel-specific multiplicative calculation factor to enhance the reflection value; A_p — channel-specific multiplicative calculation factor for biasing the reflection value; Q_{cal} — values of the brightness of the pixel of the "raw" geo-image, nominated in conventional relative units of DN (Digital Numbers) [5, 6].

$$P\lambda = \frac{p\lambda'}{\sin(\theta SE)}, \tag{1.2}$$

where, $P\lambda$ — the value of the upper atmospheric planetary reflection of radiation (TOA reflectance), taking into account correction for the angle of incidence and reflection of sunlight; $p\lambda'$ — spectral radiation arriving at the satellite sensor; θSE — the height of the sun above the horizon.

To analyze the study area on the basis of combinations of brightness values in certain channels and calculate the values of the "spectral index" of the object, an image is constructed corresponding to the index value in each pixel, which allows a detailed study of the object and an assessment of its condition.

The approach used to process remote sensing data using spectral indices used to study and assess the state of vegetation has received the generally accepted name - vegetation indices. Currently, there are more than 160 options for vegetation indices. The following most promising indexes are used in the scientific work for the analysis of the forest massif: NDVI, EVI, VARI, BAI, NBR, PSRI, NDWI, DMCI [7].

NDVI normalized vegetation index, shows the presence and state of vegetation (relative biomass).

$$DVI = \frac{NIR - Red}{NIR + Red}, \tag{1.3}$$

EVI modification of the normalized relative index of vegetation NDVI, the influence of soil and atmosphere in the values of this index is minimized.

$$EVI = \frac{NIR - Red}{(NIR + 0.6*Red - 7.5*Blue + 1)} \times 2.5, \tag{1.4}$$

VARI index is designed to highlight vegetation in the visible part of the spectrum, while attenuating the effects of differences in lighting and atmospheric phenomena.

$$VARI = \frac{Green - Red}{Green + Red - Blue}, \tag{1.5}$$

BAI index of burnt areas, used to identify areas of the surface exposed to fire [8].

$$BAI = \frac{1}{(0.1 - Red)^2 + (0.06 - NIR)^2}, \tag{1.6}$$

NBR Burn Index, designed to highlight burnt areas while dimming the differences in lighting and atmospheric conditions.

$$NBR = \frac{NIR - SWIR}{NIR + SWIR}, \tag{1.7}$$

PSRT index of coarsening carbon of plant tissues, used to estimate the total number of dry and dead plants.

$$PSRT = \frac{Red - Green}{NIR}, \tag{1.8}$$

NDWI normalized differential water index, designed to assess the moisture content of forest combustible materials.

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR}, \tag{1.9}$$

DMCI dryness index, used to assess dryness of vegetation.

$$DMCI = \frac{SWIR_3 - SWIR_2}{SWIR_3 + SWIR_2}, \tag{1.10}$$

2.1. Data processing conceptual model

In the scientific work, a systematic analysis of the processing of satellite images from the moment of their receipt to the interpretation of the final results was carried out in order to support decision-making on the management of the territory. The conceptual representation of any system under study is based on the totality of interacting and interconnected elements (processes and resources), the functioning of which is aimed at obtaining a specific result [9].

A generalized contextual diagram of the process of processing Earth remote sensing data for analyzing the state of the forest and developing design solutions for managing the territory [10] is presented in figure 1.

The final function of the system is as follows (1.11):

$$D = F(d, M, U), \tag{1.11}$$

where: d — input data; D — output (result: layers); M — process resources; U — process control resources.

The context diagram provides visual information about the necessary resources involved in the processing. The process of obtaining resources has a multi-stage structure before the information will be used by the consumer.
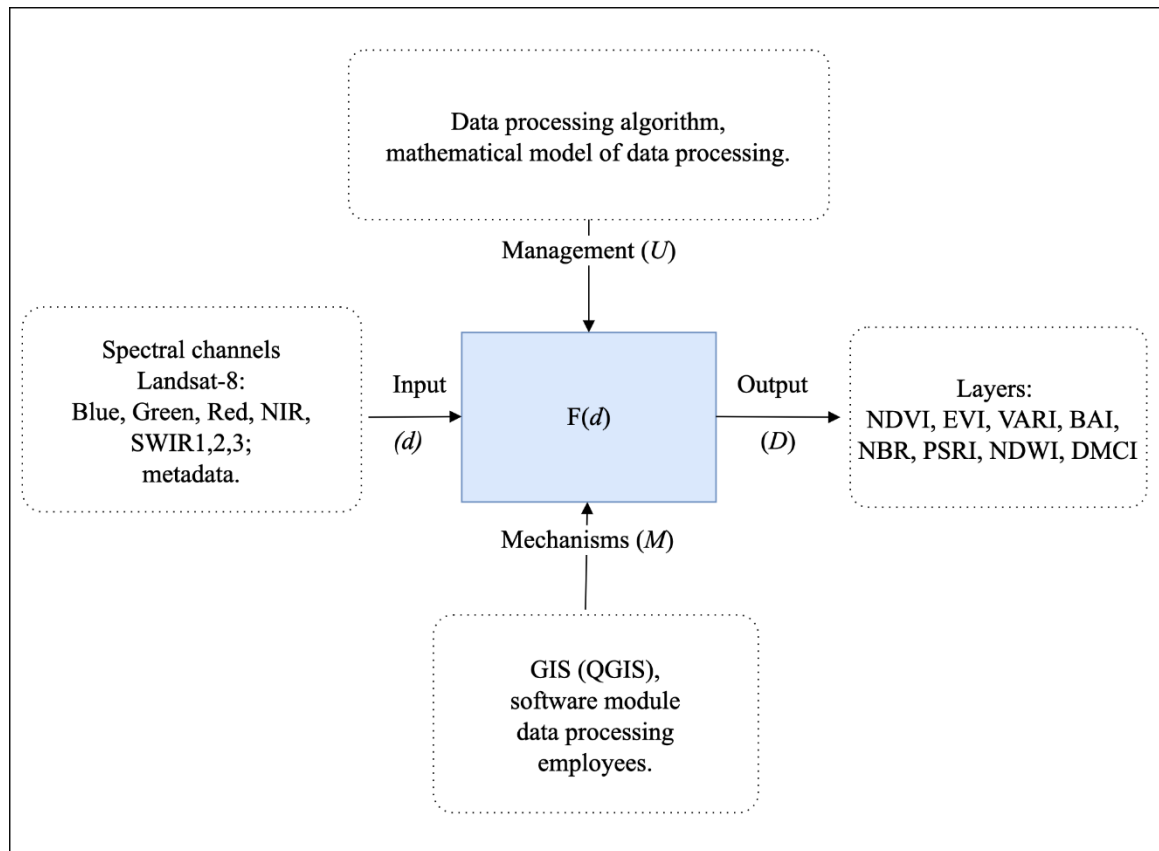
Figure 1. Context diagram of the process of processing remote sensing data.

2.2. Remote sensing data processing algorithm for images Landsat-8

Based on the conceptual model, an algorithm for processing remote sensing data from the Landsat-8 satellite has been developed (figure 2).

The processing begins with obtaining geodata, Landsat-8 satellite images (spectral channels: Blue, Green, Red, NIR, SWIR1, 2, 3) coming from open access with a frequency of 16 days are the input data. The spectral channels from the satellite go to the data preparation unit, in which the channel images are loaded into memory, and the numerical values of the reflection of each pixel are extracted in one or another range (depending on the spectral channel).

Before the data is processed, satellite images are pre-processed. It consists of radiometric and atmospheric correction, which minimizes the influence of the atmosphere on the recorded brightness valuesof objects on the earth's surface. All thematic products are calculated using adjusted data, which allows to obtain a significant increase in the reliability of information extracted from satellite images.

Then the data goes to the "Data Processing" block, where the calculations associated with the immediate task are performed. In this block, the user is given the opportunity to choose a data processing method. For comparative analysis, image processing data for the previous period of time must be received in the processing unit. After the processing of spectral channels, the geodata is subject to manual display settings. At the output, the processed data is presented as a geoinformation raster layer.
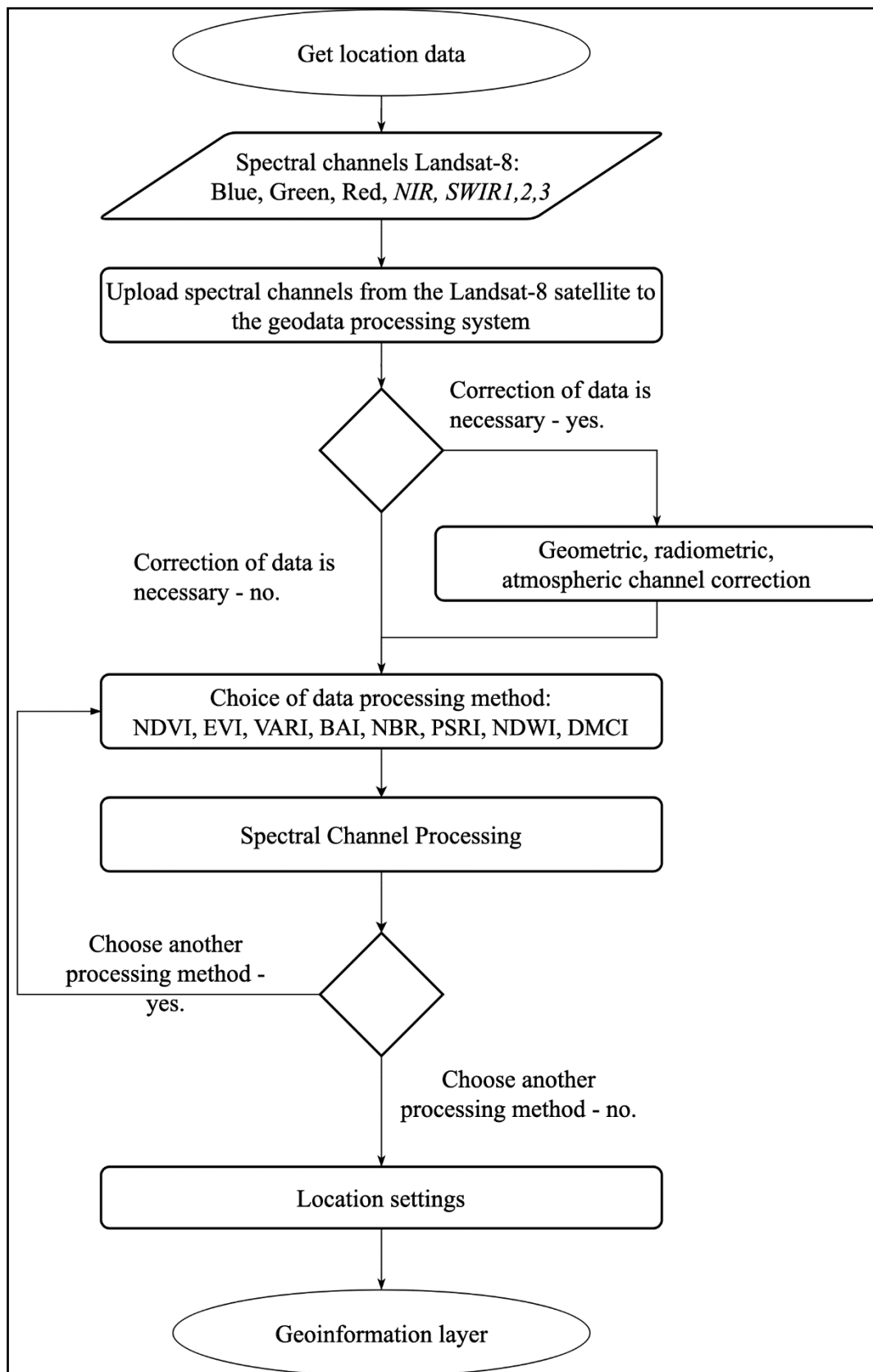
Figure 2. Flowchart for geodata processing.

2.3. Conceptual model of a business system

To formulate technical specifications for developing a program based on the geodata processing algorithm (figure 2), diagrams were modeled in the object modeling language UML. On the basis of the geodata processing flowchart the activity chart of the data processing precedent was developed (figure 3).
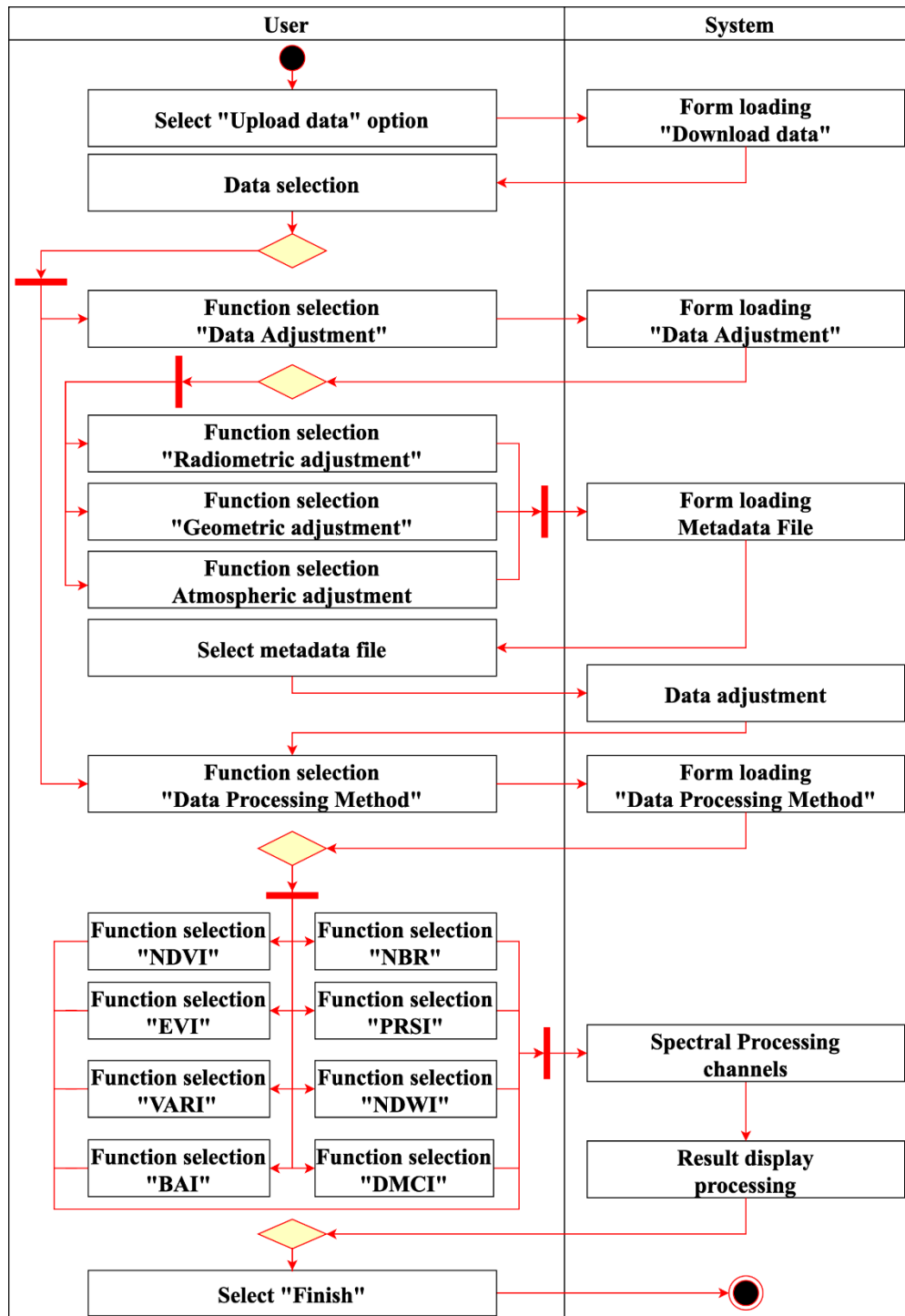

Figure 3. Activity chart of the data processing precedent.

The activity diagram describes the specification of the executable behavior in the form of coordinated sequential and parallel execution of subordinate elements interconnected by flows that go from the outputs of one node to the inputs of another.

Figure 4 shows a diagram of options for using the system (use case diagram) of remote sensing data processing.
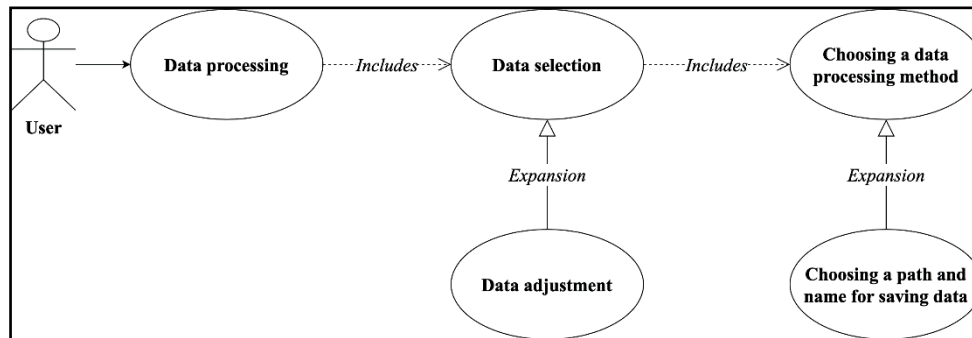

Figure 4. Diagram of system use cases.

Each use of the system determines the sequence of actions that must be performed by the developed system in its interaction with the corresponding user. Using the system, the user gets access to the direct selection of data, with the possibility of their correction, the system also provides the user with the opportunity to choose the data processing method from the proposed: NDVI, EVI, VARI, BAI, NBR, PSRI, NDWI, DMCI.

3. Results and Discussion
Based on the geodata processing algorithm (figure 2) and UML diagrams (figures 3, 4), the «Vegetation Indices» program has been developed. The graphical interface of the system is implemented in Python in the cross-platform free environment Qt Designer (figure 5).
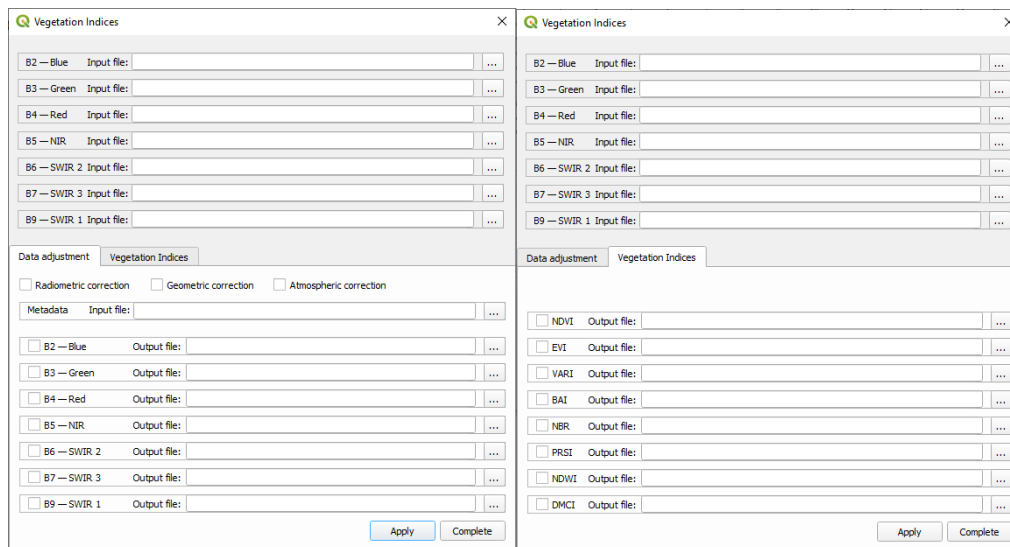

Figure 5. User GUI of the program «VegetationIndices».

To process raster channels according to formula (1.3-1.10), the QGIS Raster Calculator extension functionality was added to the program code. As an approbation of the work of the "Vegetation Indices" module the NDVI indices were calculated using the data from the Landsat-8 satellite covering part of the territory of the Ust-Kutsky district of the Irkutsk region (57 ° 34'45.0 "N, 106 ° 27'03.4" E), shooting periods 2019/07/04 and 2019/08/05.

Figure 6 shows the results of processing raster data using the Vegetation Indices module. A non-standardized gradient scale with various color palettes is used to display spectral channels.
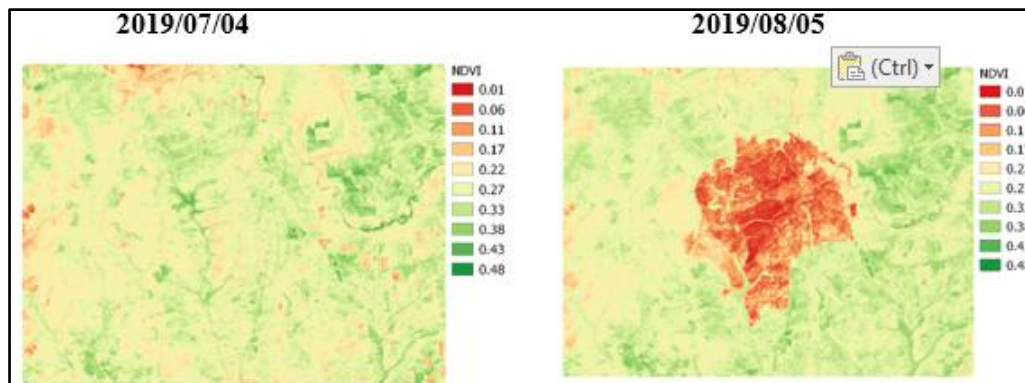

Figure 6. Example of satellite data processing results (NDVI).

The obtained cartographic geodata helps to identify problem areas of oppressed vegetation - areas noticeably differ in color during visualization. The vegetation index can serve as a tool for managing the bioproduction process in the forest management system.

The main advantage of indexes is the ease of obtaining them and a wide range of tasks to be solved with their help. It should be noted that the indices do not give absolute indicators of the studied property, and their values depend on the characteristics of the sensor (spectral channel width, resolution), shooting conditions, lighting, and atmospheric conditions. The indices provide relative estimates of the properties of the vegetation cover, which can be interpreted and converted to absolute using field data.

4. Conclusion

As a result of the study, an object decomposition of the problem was carried out, an algorithm for processing remote sensing data was developed, diagrams of the use case and system activity were modeled, software was developed that allowed the satellite images to be read and spatial data from remote sensing of the Earth to be processed. The developed software allows the analysis of the surface conditions of the territories under consideration on the basis of a series of different data. The created plugin can be successfully applied both as a separate tool, and be integrated into automated models [11] of a more complex level as a separate element.

Analysis of Earth remote sensing data is a progressive scientific method for obtaining information about the state of the territory of interest, necessary for making strategically correct management decisions [12].

The results can serve as an example of applying an integrated approach to the processing of remote sensing data to detect violations of the integrity of the forest [13] cover or seasonal negative changes in order to attract operational management and strategic response.

In turn, satellite imagery has a number of significant drawbacks when analyzing Earth remote sensing data:

1) Loss of information and its distortion due to atmospheric phenomena - images taken in cloudy weather do not have the necessary information.

2) Low resolution of images located in the public domain 30–250 m / pixel; higher resolution images (1–6 m / pixel) are acquired commercially.

3) Interval for repeating images - 16 days. There is no possibility of daily data analysis.

Further scientific work is aimed at designing and developing a geographic information web service to support decision making, providing access to an array of diverse, constantly updated information. Using a web service will provide an opportunity to quickly evaluate the large amount of available space images at different times in the area of interest.

The further development of the project is related to the introduction of a mathematical model [14] for processing and analyzing spatio-temporal data into the module to build forecast maps of the development of processes on the earth's surface based on satellite monitoring data of different times using probabilistic and spatial characteristics.

References

[1]   Vagizov M R, Ustyugov V A and Kvochkin D O 2017 Determination of the forest inventory indicators according to the photographs of the unmanned aerial vehicles *Ecology, Environment and Conservation* vol 23 **1** pp 582-586

[2]   Vagizov M R, Mihailova A A, Fetisova A A, Habirova A I and Vaisero O S Study of reforestation after cuttings based on materials of open web-mapping service *IOP Conf. Ser.: Earth Environ. Sci* **316** 012075DOI:10.1088/1755-1315/316/1/012075

[3]   Landsat 8 (L8) Data Users Handbook *NASA.GOV* available at: https://www.usgs.gov/media/files/landsat-8-data-users-handbook (accessed: 02.02.2020)

[4]   Neil Flood 2014 Continuity of Reflectance Data between Landsat-7 ETM+ and Landsat-8 OLI, for Both Top-of-Atmosphere and Surface Reflectance: A Study in the Australian Landscape. Remote Sens. 2014, **6** 7952-70

[5]   Atmospheric correction by the DOS method [in Russian – Atmosfernaya korrekciya po metodu DOS] *GIS Laboratoriya* available at: https://gis-lab.info/qa/atcorr-dos.html (accessed: 02.02.2020)

[6]   Lu D, Mausel P, Brondizio E and Moran E 2002 Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. International Journal of Remote Sensing, **23 (13)** 2651-71

[7]   Index Gallery [in Russian – Galereya indeksov] *Environmental Systems Research Institute* available at: https://pro.arcgis.com/ru/pro-app/help/data/imagery/indices-gallery.htm (accessed: 02.02.2020)

[8]   Martyn I, Petrov Y, Stepanov S, Sidorenko A and Vagizov M 2020 Monitoring forest fires and their consequences using MODIS spectroradiometer data. *IOP Conf. Ser.: Earth Environ. Sci.* **507** 012019 DOI:10.1088/1755-1315/507/1/012019

[9]   Tatarnikova T, Stepanov S, Petrov Y, Sidorenko A and Martyn I 2020 A conceptual model for geodata processing for sustainable forest management. *IOP Conf. Ser.: Earth Environ. Sci.* **507** 012029 DOI:10.1088/1755-1315/507/1/012029

[10]   Istomin E, Petrov Y, Stepanov S, Sidorenko A and Martyn I 2020 About the methodology of geo-risk management in forestry. *IOP Conf. Ser.: Earth Environ. Sci.* **507** 012006 DOI:10.1088/1755-1315/507/1/012006

[11]   Istomin E, Petrov Y, Stepanov S, Kolbina O and Sidorenko A 2019 Model of optimum integration of diverse geodata for the benefit of management of forestry. *IOP Conf. Ser.: Earth Environ. Sci.* **316** 012013 DOI:10.1088/1755-1315/316/1/01201

[12]   Istomin E, Martyn I, Petrov Y, Stepanov S and Sidorenko A 2020 Study of intra-day dynamics of currents in the area of the navigable strait of Baltiysk to adjust the movement of water transport. *IOP Conf. Ser.: Mater. Sci. Eng*. **817** 012013 DOI:10.1088/1757-899X/817/1/012013

[13]   Istomin E, Petrov Y, Stepanov S, Kolbina O and Sidorenko A 2019 About technology of risk management in forestry. *IOP Conf. Ser.: Earth Environ. Sci.* **316** 012011 DOI:10.1088/1755-1315/316/1/012011

[14]   Istomin E, Sidorenko A, Stepanov S, Petrov Y and Martyn I 2020 Application of Kalman-Bucy filter for vessel traffic control systems in the northern sea route. *IOP Conf. Ser.: Mater. Sci. Eng.* **817** 012012 DOI:10.1088/1757-899X/817/1/012012