System for automated environmental monitoring using remote sensing data of the Earth from open data sources

Konstantin Maslov
Division for Information Technology
Tomsk Polytechnic University
Tomsk, Russian Federation
kam20@tpu.ru

Olga Tokareva
Division for Information Technology
Tomsk Polytechnic University
Tomsk, Russian Federation
ost@tpu.ru

***Abstract*—Environmental monitoring using remote sensing data requires an analyst to perform a large amount of routine work related to downloading, processing and analyzing data, especially in cases when the study area is covered with a large number of satellite imagery. The paper presents the results of the design and software implementation of the system that automates downloading and processing of remotely sensed data according to developed scenarios and, thus, greatly simplifies the processing of satellite imagery. It provides the description of tools for accessing data from the archive of the United States Geological Survey (USGS) and describes the data flow in the system. The paper gives an analysis of results obtained using the developed system on the example of monitoring the state of Siberian pine forests of the Tomsk region.**

Keywords—remote sensing, Siberian pine forests, land-cover change, automation, geographic information system, environmental monitoring, NDVI, NDWI

I. Introduction

Monitoring of land-cover change using remotely sensed data helps to make an objective assessment of the state of environmental objects and to identify changes occurring because of natural processes or anthropogenic impact.

Methods of remote sensing of the Earth are widely used in solving problems of environmental monitoring [1]–[4]. However, processing of Earth remote sensing data by an analyst is slow and can lead to errors due to inattention or loss of researcher concentration, and it consists of steps that can be easily automated, therefore, the issue of automating this process is relevant [5].

To solve problems of monitoring of land-cover change, index images based on combination of pixel values from different electromagnetic spectrum regions are widely used [6]–[8]. Examples of such index images are vegetation index maps, e.g. normalized difference vegetation index (NDVI) or normalized difference water index (NDWI), standing for the amount of photosynthetically active phytomass and water content in leaves and needles of vegetation, respectively [9], [10].

To provide continuous monitoring of the land-cover using remotely sensed data of the Earth, it is necessary to have a system that simultaneously automates both the processes of downloading data from external sources and processing of the fetched data according to developed scenarios.

The research is carried out at Tomsk Polytechnic University within the framework of Tomsk Polytechnic University Competitiveness Enhancement Program.

II. System design

A. General architecture of the system

Fig. 1 demonstrates the diagram showing the general architecture of the designed system for automated environmental monitoring using remote sensing data of the Earth from open data sources. The user controls the system by interacting with the console interface of the monitoring system. Through network requests, the system interacts with data from the external sources, downloads them onto local storage and processes them.

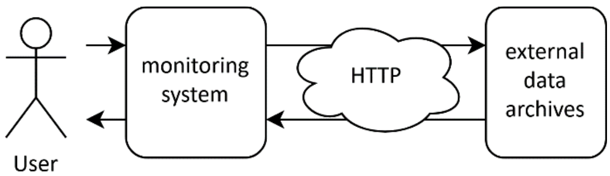

Fig. 1. Architecture of the system

The system provides a framework focused on developing and executing scenarios, which are sets of instructions that automate data acquisition and processing. Scenarios are run at regular intervals according to a predefined schedule.

Note that the architecture of the designed system does not depend on terms of the remote sensing domain, so the system can be used with any kind of data obtained from external sources, e.g. financial, meteorological, etc.

B. Data access tools

Nowadays network resources allow end-users to get access to remotely sensed imagery via public application program interfaces (APIs). These tools automate data searching and downloading processes.

In this paper, the United States Geological Survey (USGS) data archive was used as a source of external data. To automate the process of obtaining the correct identifiers of satellite images EarthExplorer API was used [11]. The EarthExplorer API provides a flexible and powerful mechanism for searching remotely sensed data and their products in the USGS data archive, allowing to search across hundreds of data collections from various spacecraft and sensors [12], e.g. Landsat, MODIS, ASTER, VIIRS, etc., giving the opportunity to search using complex queries. The ESPA ordering system API [13] was used to order scenes and get the URLs to download them.

C. Data flow in the system

The designed architecture and selected data access tools allow to automate the process of acquisition and processing

978-1-7281-2896-2/19/$31.00 ©2019 IEEE

scenes of satellite imagery. Fig. 2 shows the corresponding data flow diagram.

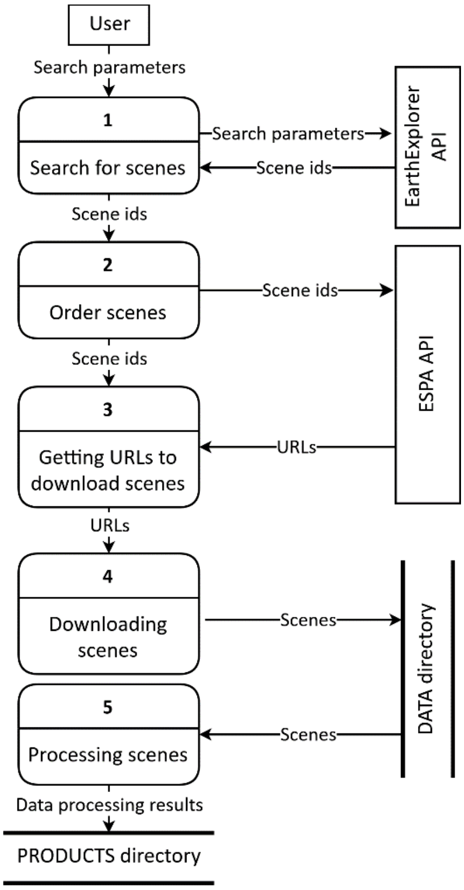

Fig. 2. Data flow diagram for a scenario of downloading and processing satellite imagery from the USGS data archive

The user specifies search parameters for the particular area of interest. Then the system obtains the appropriate scenes identifiers via the EarthExplorer API, after that it orders scenes and receives URLs to download them via the ESPA ordering system API. Finally, the system downloads the scenes and processes them, it puts the scenes in the DATA directory and places the results of the processing in the PRODUCTS directory.

D. Scenario algorithm

Below is an example of a scenario for monitoring of vegetation, which allows to trace changes in its condition and locate areas of its degradation and restoration.

At each iteration, the scenario waits for its next run, and then it acquires new satellite imagery and processes each image. For each image it performs the calculation of open area mask, cloud mask and maps of vegetation indices and calculates the zonal statistics (mean values of vegetation indices and percentage of cloud pixels) for the study area, which is represented in a vector format as a set of polygons.

The following is the algorithm in pseudo-code that demonstrates the steps performed by the designed scenario:

```
BEGIN
  WHILE running
    Wait for the next run
    Search for scenes
    Order new scenes
    Get URLs to download available scenes
    FOR each URL in URLs
      Download scene archive by URL
      Unzip the archive
      Calculate open area mask, cloud
   mask, NDVI map and NDWI map
      Calculate zonal statistics for the
   studied polygons
      Write zonal statistics to a .csv
   file
      IF no clouds within the polygons
      THEN
        Write zonal statistics to a
   separate .csv file
        Calculate new dNDVI and dNDWI
   maps
      END IF
    END FOR
  END WHILE
END
```

Vegetation indices NDVI and NDWI are calculated by the formulas [9], [10]:

$$NDVI = (NIR - RED) / (NIR + RED), \quad (1)$$

$$NDWI = (NIR - SWIR) / (NIR + SWIR), \quad (2)$$

where NIR, RED and SWIR are spectral reflectance measurements acquired in the near-infrared, red and short-wave infrared regions of the electromagnetic spectrum, respectively.

To assess changes in the state of vegetation, the differenced indices dNDVI and dNDWI are calculated by the formulas:

$$dNDVI = NDVI_{post} - NDVI_{pre}, \quad (3)$$

$$dNDWI = NDWI_{post} - NDWI_{pre}, \quad (4)$$

where $NDVI_{post}$ and $NDWI_{post}$ are vegetation index values for imagery with a later acquisition date, $NDVI_{pre}$ and $NDWI_{post}$ are vegetation index values for imagery with an earlier acquisition date.

The .csv extension files obtained during the execution of the scenario can be opened in Microsoft Excel or other software for further analysis, including, for example, plotting the dynamics of the values of vegetation indices. Calculated open area masks and cloud masks, as well as maps of NDVI, NDWI, dNDVI, dNDWI indices are raster images with spatial reference and, accordingly, can be opened and visualized in geographic information system (GIS), used for more complicated geoanalysis, etc.

III. System implementation

To implement the system, an object-oriented approach was chosen. Python was used as a programming language; rasterio and rasterstats libraries were used to deal with raster files and to calculate zonal statistics.

The system and the scenario have been implemented. Fig 3 shows the corresponding class diagram.

The *Manager* class implements the console user interface and is responsible for the interpretation of user commands. The *AbstractScenario* and *AbstractDataManager* classes are abstract and represent, respectively, a generalized scenarios classes and classes for receiving data from external sources.

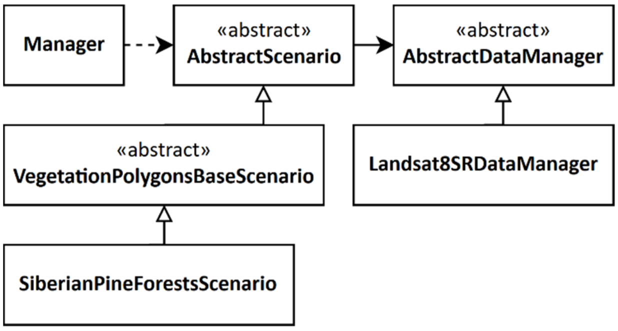

Fig. 3. Classes of the developed system and the scenario (only essential classes and their relationships are shown)

The *Landsat8SRDataManager* class implements a mechanism for searching, ordering and downloading satellite images acquired by the Landsat 8 satellite of the L2 processing level (surface reflectance) from the USGS archive [14]. This class can be reused by other scenarios to access data from the same collection of USGS data archive, but for other spatial and temporal parameters.

The *VegetationPolygonsBaseScenario* class implements the algorithm described previously, and the *SiberianPineForestsScenario* class implements the scenario for monitoring the state of the Siberian pine forests, which are specially protected natural areas.

IV. Results

This section gives the description of the results obtained with the *SiberianPineForestsScenario* scenario for monitoring of the state of the Siberian pine forests of the Tomsk Region, which are specially protected natural areas (Aksenovskiy, Belousovskiy, Bogashevskiy, Loskutovskiy, Luchanovo-Ipatovskiy, Magadaevskiy, Nizhne-Sechenovskiy, Petrovskiy, Petukhovskiy, Plotnikovskiy, Protopopovskiy, Trubachevskiy, Voronovskiy Siberian pine forests and forest park near the Yar settlement). Reference [1] gives a more detailed description of the Siberian pine forests.

The mean values of the NDVI and NDWI indices were calculated within the boundaries of the Siberian pine forests of the Tomsk region. We compared the values of vegetation indices calculated automatically with values obtained with GIS, in this way the correctness of the scenario implementation was shown.

The scenario automatically determined cloudless scenes within the boundaries of all Siberian pine forests. Table I contains the description of these scenes. Using data from the file with zonal statistics for these scenes, plots were built showing the changes of mean values of the NDVI and NDWI indices (Fig. 4) by year for mid-July.

TABLE I. List of used Landsat 8 satellite imagery

WRS2 path/row	Acquisition date	Identifier
148/21	14.07.2013	LC08_L1TP_148021_20130714_20170503_01_T1
148/21	22.07.2016	LC08_L1TP_148021_20160722_20180525_01_T1
148/21	26.08.2017	LC08_L1TP_148021_20170826_20170913_01_T1
148/21	12.07.2018	LC08_L1TP_148021_20180712_20180717_01_T1

As depicted in Fig. 4, the Luchanovo-Ipatovskiy Siberian pine forest showed significant negative state dynamics, not correlating with other Siberian pine forests. In July 2018, the mean NDVI and NDWI indices of this pine forest decreased.

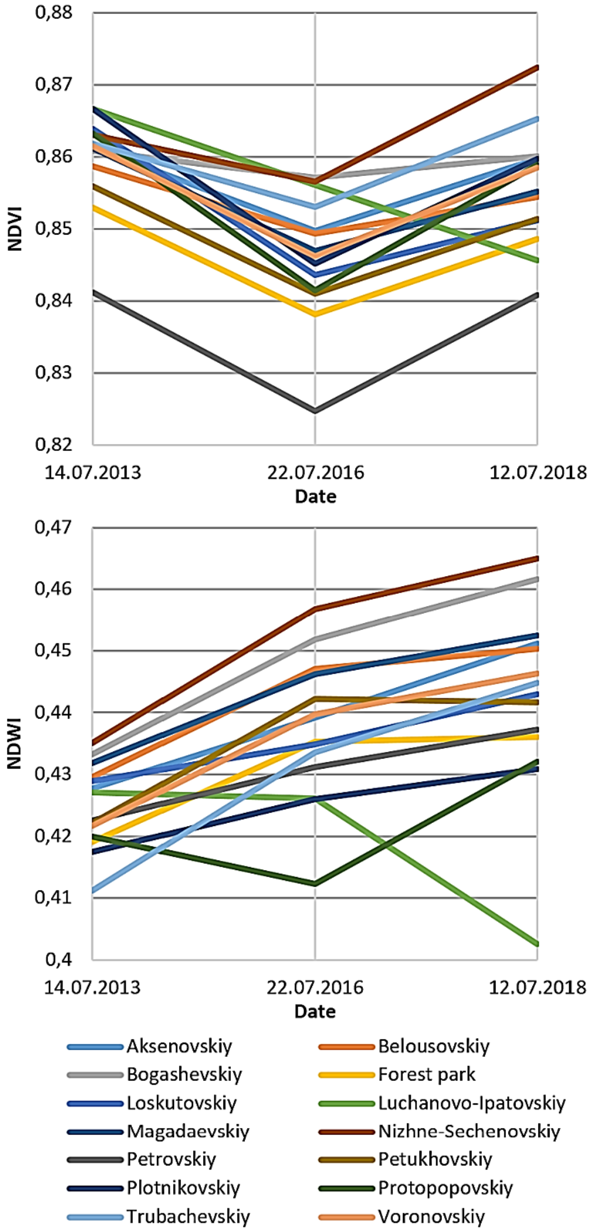

Fig. 4. Changes of the mean values of the vegetation indices NDVI (top) and NDWI (bottom) by years for the Siberian pine forests

According to the scenario dNDVI and dNDWI maps were calculated for the scenes listed in Table I. Fig. 5 demonstrates fragments of these maps for the Luchanovo-Ipatovskiy Siberian pine forest, boundaries of the Siberian pine forest are shown with black line. According to these fragments, areas of the most negative state dynamics were established: those are forest areas near the Luchanovo settlement and the Ipatovo settlement, these areas are shown in red shades; and the period of negative impact is the period from 2016 to 2017.

The analysis of high spatial resolution imagery and ground surveys of the territory showed that these areas were affected by Siberian silkworm larvae. It is worth noting, that the NDWI index is more sensitive to this kind of forest damage, as Fig. 5 shows.

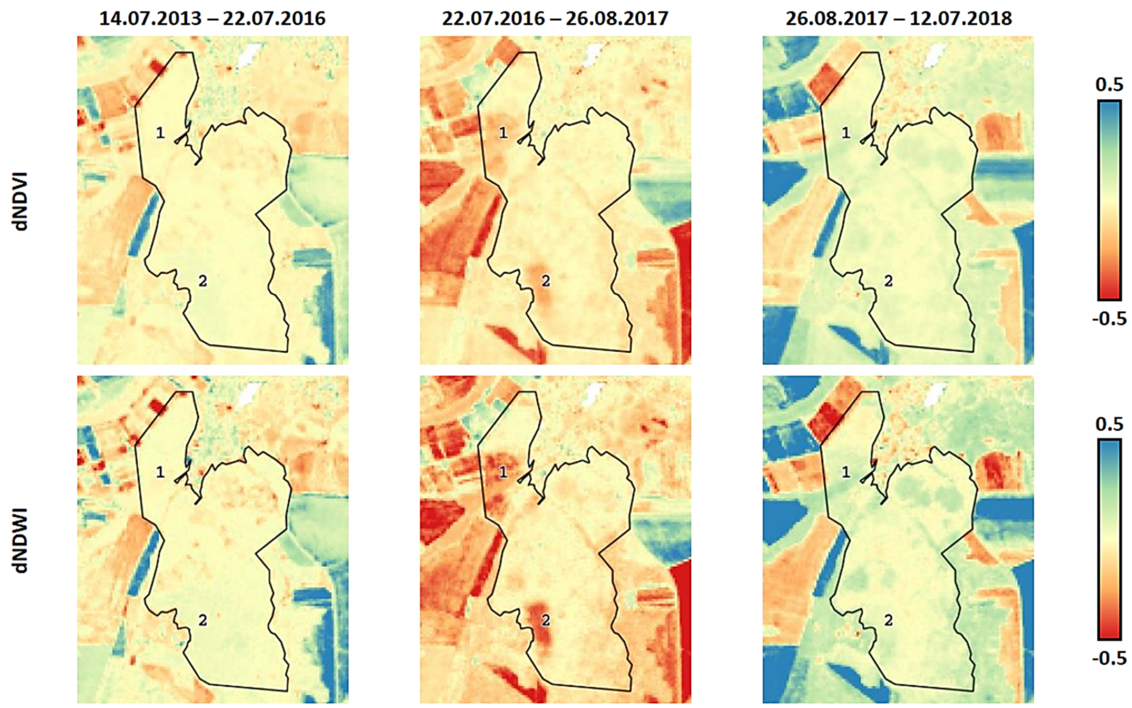

Fig. 5. Fragments of the obtained dNDVI and dNDWI maps for Luchanovo-Ipatovskiy Siberian pine forest. Location of the damaged areas: 1 – the area near the Luchanovo settlement, 2 – the area near the Ipatovo settlement.

V. Conclusions

This paper presented the results of the design and implementation of the automated system for monitoring of land-cover change using Earth remote sensing data obtained via public APIs. A scenario was designed and implemented to monitor the state of the Siberian pine forests of the Tomsk region. The functionality of the developed system was successfully tested. The results of the testing showed the negative dynamics of the state of the Luchanovo-Ipatovskiy Siberian pine forest and allowed to locate its damaged areas.

In the future, the developed system may be extended for environmental monitoring of other objects or phenomena using a wider range of Earth remote sensing data.

Acknowledgment

The research is carried out at Tomsk Polytechnic University within the framework of Tomsk Polytechnic University Competitiveness Enhancement Program.

References

[1] O. A. Pasko, O. S. Tokareva, A. Alshaibi, T. Y. Chernikova, and P. Cabral, "Assessment of state of cedar forests in Tomsk region using remote sensing data of the Earth", Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering, vol. 330, no. 1, pp. 98–109, 2019. (in Russian)

[2] K. A. Maslov, "Analysis of the dynamics of the state of coniferous forests damaged by the Dendrolimus superans using remotely sensed data", Proc. XVI International scientific and technical conference of students, graduates and young scientists «Youth and modern information technology», pp. 98–99, Apr. 2019. (in Russian)

[3] K. V. Mjachina, C. W. Baynard, A. A. Chibilyev, and R. D. Richardson, "Landscape disturbance caused by non-renewable energy production in a semi-arid region: a case study on the Russian steppe", International journal of sustainable development & world ecology, vol. 25, no. 6, pp. 541–553, 2018.

[4] A. A. Lamqadem, G. M. Afrasinei, and H. Saber, "Analysis of Landsat-derived multitemporal vegetation cover to understand drivers of oasis agroecosystems change", Journal of applied remote sensing, vol. 13, no. 1, February 2019, [online] Available: https://www.researchgate.net/publication/331199740_Analysis_of_Landsat-derived_multitemporal_vegetation_cover_to_understand_drivers_of_oasis_agroecosystems_change.

[5] I. V. Balashov, O. A. Khalikova, M. A. Burtsev, E. A. Loupian, and A. M. Matveev, "Organization of automatic data acquisition from satellite and meteorological data archiving and distribution centers", Current problems in remote sensing of the Earth from space, vol. 10, no. 3, pp. 9–20, 2013. (in Russian)

[6] L. Liu, Q. Niu, J. Heng, H. Li, and Z. Xu, "Transition Characteristics of the Dry-Wet Regime and Vegetation Dynamic Responses over the Yarlung Zangbo River Basin, Southeast Qinghai-Tibet Plateau", Remote sensing, vol. 11, no. 10, May 2019, [online] Available: https://www.mdpi.com/2072-4292/11/10/1254.

[7] R. Albarakat and V. Lakshmi, "Comparison of Normalized Difference Vegetation Index Derived from Landsat, MODIS, and AVHRR for the Mesopotamian Marshes Between 2002 and 2018", Remote sensing, vol. 11, no. 10, May 2019, [online] Available: https://www.mdpi.com/2072-4292/11/10/1245.

[8] N. Quintero, O. Viedma, I. R. Urbieta, and J. M. Moreno, "Comparison of Normalized Difference Vegetation Index Derived from Landsat, MODIS, and AVHRR for the Mesopotamian Marshes Between 2002 and 2018", Forests, vol. 10, no. 6, June 2019, [online] Available: https://www.mdpi.com/1999-4907/10/6/518.

[9] A. S. Cherepanov and E. G. Gruzhinina, "Spectral properties of vegetation and vegetation indices", Geomatics, no. 3, pp. 28–32, 2009. (in Russian)

[10] A. Bannari, D. Morin, F. Bonn, and A. R. Huete, "A review of vegetation indices:, Remote sensing reviews, vol. 13, no. 1, pp. 95–120, 1995.

[11] United States Geological Survey, "EarthExplorer – Service Documentation", United States Geological Survey, May 2019, [online]. Available: https://earthexplorer.usgs.gov/inventory/documentation.

[12] United States Geological Survey, "EarthExplorer – Home", United States Geological Survey, April 2019, [online] Available: https://earthexplorer.usgs.gov/.

[13] United States Geological Survey, "API for the ESPA ordering system", Github, May 2019, [online] Available: https://github.com/USGS-EROS/espa-api.

[14] United States Geological Survey, "Landsat Surface Reflectance", United States Geological Survey, April 2019, [online] Available: https://www.usgs.gov/land-resources/nli/landsat/landsat-surface-reflectance.