IOP Conf. Series: Materials Science and Engineering 1019 (2021) 012100 doi:10.1088/1757-899X/1019/1/012100

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

K A Maslov and O S Tokareva
Division for Information Technology, Tomsk Polytechnic University, Tomsk, Russia
E-mail: kam20@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.

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

To solve the 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 specific scenarios.

2. System design

2.1. General architecture of the system
Figure 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 the local storage and processes them.

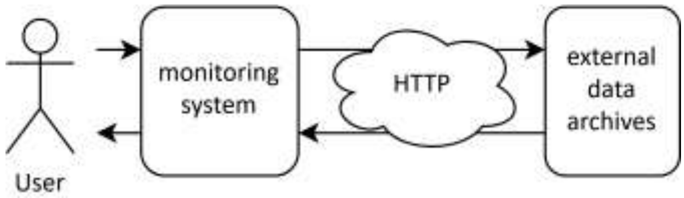

Figure 1. Architecture of the system.

The system provides a framework focused on development and execution of the scenarios, which are sequences of instructions that automate data acquisition and processing. The scenarios are run at regular intervals according to a predefined schedule.

Note that the architecture of the 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.

2.2. 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. It also gives an opportunity to search satellite imagery using complex queries. Earth resources observation and science (EROS) center science processing architecture (ESPA) on demand interface API [13] was used to order scenes and get the URLs to download the scenes.

2.3. Data flow in the system

The designed architecture and selected data access tools allow to automate the processes of acquisition and processing of the satellite imagery scenes. Figure 2 shows the corresponding data flow diagram.

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.

2.4. 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.

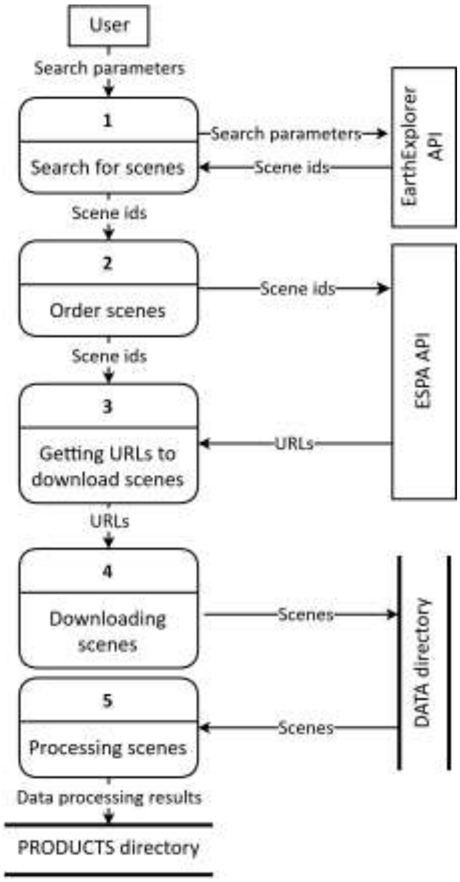

Figure 2. Data flow diagram for a scenario of downloading and processing satellite imagery from the USGS data archive.

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]:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}}, \tag{1}$$

$$\mathrm{NDWI} = \frac{\mathrm{NIR} - \mathrm{SWIR}}{\mathrm{NIR} + \mathrm{SWIR}}, \tag{2}$$

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:

$$\mathrm{dNDVI} = \mathrm{NDVI}_{\mathrm{post}} - \mathrm{NDVI}_{\mathrm{pre}}, \tag{3}$$

$$\mathrm{dNDWI} = \mathrm{NDWI}_{\mathrm{post}} - \mathrm{NDWI}_{\mathrm{pre}}, \tag{4}$$

where $\mathrm{NDVI}_{\mathrm{post}}$ and $\mathrm{NDWI}_{\mathrm{post}}$ are vegetation index values for imagery with a later acquisition date, $\mathrm{NDVI}_{\mathrm{pre}}$ and $\mathrm{NDWI}_{\mathrm{pre}}$ are vegetation index values for imagery with an earlier acquisition date.

The csv 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.

3. System implementation

To implement the system, the 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. Figure 3 shows the corresponding class diagram.

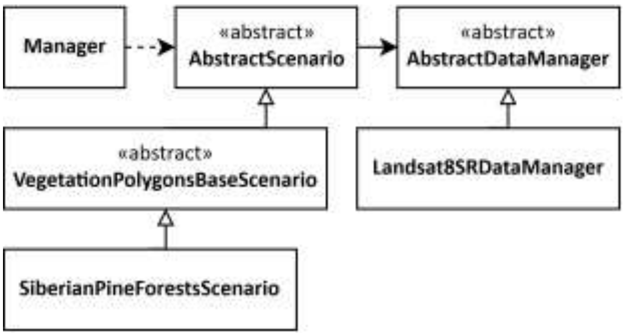

Figure 3. Classes of the developed system and scenario (only essential classes and their relationships are shown).

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.

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 the 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.

4. 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 village Yar). 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 1 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 (Figure 4) by year for mid-July.

Table 1. List of the used cloudless 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 Figure 4, the Luchanovo-Ipatovskiy Siberian pine forest has shown significant negative state dynamics, not correlating with other Siberian pine forests. In July 2018, the mean NDVI and NDWI indices of this forest decreased.

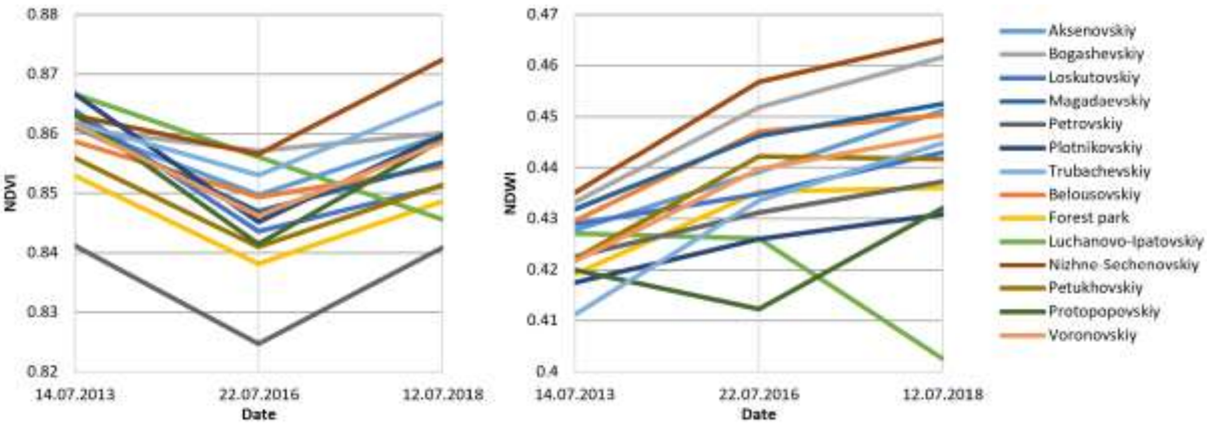

Figure 4. Mean values of the vegetation indices NDVI (left) and NDWI (right) by years for the Siberian pine forests.

According to the scenario dNDVI and dNDWI maps were calculated for the scenes listed in table 1. Figure 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 village Luchanovo and the village Ipatovo, these areas are shown in red shades; and the period of negative impact is the period from 2016 to 2017.

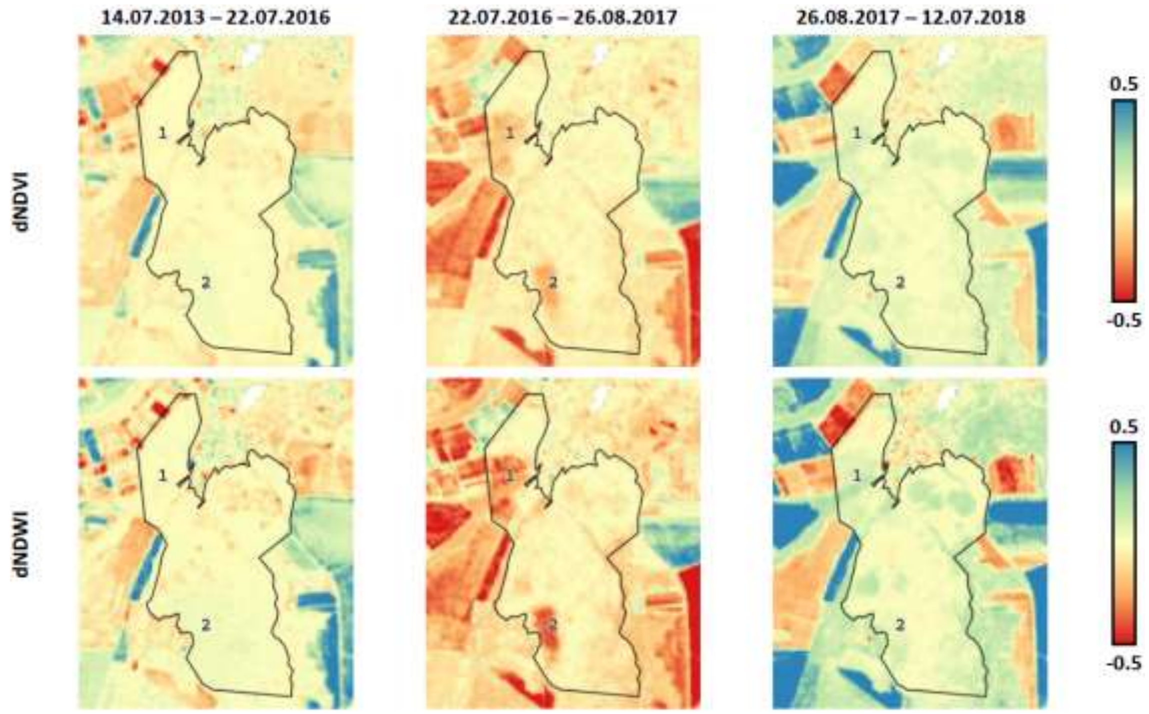

Figure 5. Fragments of the obtained dNDVI and dNDWI maps for Luchanovo-Ipatovskiy Siberian pine forest. Location of the damaged areas: 1 – the area near the village Luchanovo, 2 – the area near the village Ipatovo.

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 Figure 5 shows.

5. Conclusion

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 modified and extended for environmental monitoring of other objects or phenomena using a wider range of Earth remote sensing data.

Acknowledgements

The research has been carried out with the financial support of the RFBR (Project 18-47-700010).

References

[1] Pasko O A, Tokareva O S, Alshaibi A, Chernikova T Y and Cabral P 2019 *Bull. TPU Geo Assets Eng.* **330** 98–109

[2] Maslov K A 2019 *Proc. XVI Int. Scientific and Technical Conf. of students, graduates and young scientists «Youth and modern information technology»* 98–99

[3] Mjachina K V, Baynard C W, Chibilyev A A and Richardson R D 2018 *Int. J. Sust. Dev. World Ec.* **25** 6 541–553

[4] Lamqadem A A, Afrasinei G M and Saber H 2019 *J. App. Rem. Sens.* **13** 1 180590 doi: 10.1117/1.JRS.13.014517,

[5] Balashov I V, Khalikova O A, Burtsev M A, Loupian E A and Matveev A M 2013 *Curr. Prob. Rem. Sens. Earth Space* **10** 3 9–20
[6] Liu L, Niu Q, Heng J, Li H and Xu Z 2019 *Rem. Sens.* **11** 10 1254 doi: 10.3390/rs11101254
[7] Albarakat R and Lakshmi V 2019 *Rem. Sens.* **11** 10 1245 doi: 10.3390/rs11101245
[8] Quintero N, Viedma O, Urbieta I R and Moreno J M 2019 *Forests* **10** 518 doi: 10.3390/f10060518
[9] Cherepanov A S and Gruzhinina E G 2009 *Geomatics* 3 28–32
[10] Bannari A, Morin D, Bonn F and Huete A R 2009 *Rem. Sens. Rev.* **13** 1 95–120
[11] United States Geological Survey, EarthExplorer – Service Documentation. Available at: https://earthexplorer.usgs.gov/inventory/documentation
[12] United States Geological Survey, EarthExplorer – Home. Available at: https://earthexplorer.usgs.gov/
[13] United States Geological Survey, API for the ESPA ordering system, Github. Available at: https://github.com/USGS-EROS/espa-api
[14] United States Geological Survey, Landsat Surface Reflectance. Available at: https://www.usgs.gov/land-resources/nli/landsat/landsat-surface-reflectance