

Technology Automation of Satellite Data Processing for Operational Mapping of Water and Ice Surfaces

Anastasiya Voronova, Sergey Kuzminykh

Siberian Center of SRC “Planeta”, Novosibirsk, Russia
35voran@gmail.com

Abstract

The process of technology automation of satellite data thematic processing presented in the article includes two main stages: the development of classification algorithm and the creation of a software module based on them.

At the first stage, algorithm (decision tree) was developed to isolate water surfaces in conditions of destroyed ice cover using medium resolution satellite images (Landsat-8, OLI). The article presents the final dendrogram, examples of the thematic maps and the results of accuracy assessment in comparison with standard masks and supervised classification.

At the second stage, a software module based on the IDL language was created. This module allows to perform the following tasks automatically: primary data processing, classification, “cleaning” the results, exporting to a vector format.

Keywords: decision tree, classification, indices, ice, NDVI, NDII, OLI, Landsat-8

АВТОМАТИЗАЦИЯ ТЕХНОЛОГИИ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ ДЛЯ ОПЕРАТИВНОГО КАРТОГРАФИРОВАНИЯ ВОДНОЙ И ЛЕДОВОЙ ПОВЕРХНОСТЕЙ

А.Е. Воронова, С.А. Кузьминых

Сибирский центр НИЦ «Планета», Новосибирск, Россия
35voran@gmail.com

Представленный в статье процесс автоматизации технологии тематической обработки спутниковых данных включает в себя разработку порогового алгоритма классификации и создание на его основе программного модуля.

На первом этапе был разработан пороговый алгоритм (дерево решений) для выделения водных поверхностей в условиях разрушенного ледового покрова по спутниковым изображениям среднего пространственного разрешения (Landsat-8, OLI, TIRS). В статье приведена итоговая дендрограмма алгоритма, полученные тематические карты и результаты оценки точности разработанного алгоритма в сравнении со стандартными масками и обучаемой классификацией.

На втором этапе был создан программный модуль на базе языка IDL. Данный модуль позволяет выполнять в автоматическом режиме следующие задачи: первичную обработку данных, классификацию с помощью выбранного дерева решений, «очистку» результатов классификации, экспорт целевого класса в векторный формат.

Ключевые слова: дерево решений, классификация, индексы, лед, NDVI, NDII, OLI, Landsat-8

Введение

Оперативный спутниковый мониторинг состояния водных объектов во время весенних ледовых явлений играет важную роль в предупреждении и ликвидации чрезвычайных ситуаций. При этом время, затраченное на подготовку тематических карт, становится ключевым фактором обеспечения ежедневного мониторинга. В общем виде технология подготовки тематической продукции включает в себя три основных этапа: подготовка данных, классификация, экспорт данных. В зависимости от исходных данных, эти этапы могут включать в себя различные операции. Самым времязатратным этапом является процесс классификации изображения с целью выделения классов «воды», «льда» и «облачности».

В мировой практике широкое применение для создания масок воды, снега и облачности получили пороговые алгоритмы. Таким образом, например, Европейское космическое агентство (ESA) обеспечивает обработку изображений со спутника Sentinel-2 до уровня «2A». В таких алгоритмах зачастую используются различные спектральные индексы для задания пороговых значений. В настоящий момент существует ряд работ, посвященных применению спектральных индексов для идентификации водной поверхности, как у отечественных авторов [1], так и у зарубежных [2, 3]. Однако в данных статьях рассматриваются водные объекты, не покрытые льдом, и поэтому вопрос, насколько точно по данным индексам можно отделить ледовую поверхность от водной, остается открытым. Кроме того, как показывает практика, многие алгоритмы выдают различные результаты в зависимости от геоклиматических условий исследуемого региона.

Целью данной работы является автоматизация технологии тематической обработки спутниковых изображений. Основными задачами являются разработка порогового алгоритма

идентификации целевых классов («вода», «лед», «снег») в условиях геоклиматических особенностей сибирского региона и создание программного модуля для автоматизации выполнения классификации и других рутинных процедур (первичная обработка, экспорт данных в векторный формат).

Исследуемый регион

Сибирский центр ФГБУ «НИЦ «Планета» обеспечивает оперативный мониторинг обширных территорий Уральского и Сибирского федеральных округов. Наиболее пристального внимания во время весенних ледовых явлений требуют пойменные территории реки Оби и ее притоков. Вскрытие реки происходит в период со второй половины марта до начала июня и совпадает с волной половодья. Весенний ледоход часто сопровождается заторами и выходами воды на пойму. При максимальных уровнях половодья разливы на пойме в нижнем и среднем течении Оби достигают нескольких десятков километров, что может приводить к затоплению населенных пунктов и сельскохозяйственных угодий [4].

Данные

В качестве данных использовались снимки среднего разрешения с космического аппарата Landsat-8 (OLI, TIRS). Характеристики каналов радиометров OLI и TIRS приведены в таблице 1 [5].

Таблица 1. Характеристики каналов OLI и TIRS (Landsat-8)

<i>Каналы</i>	<i>Длина волны (мкм)</i>	<i>Пространственное разрешение (м)</i>
Coastal/Aerosol	0.43 – 0.45	30
Blue	0.45 – 0.51	30
Green	0.53 – 0.59	30
Red	0.64 – 0.67	30
NIR	0.85 – 0.88	30
SWIR1	1.57 – 1.65	30
SWIR2	2.11 – 2.29	30
panchromatic	0.50 – 0.68	15
Cirrus	1.36 – 1.38	30
Thermal Infrared 1	10.60 – 11.19	100
Thermal Infrared 2	11.50 – 12.51	100

Исходные изображения были преобразованы в TOA Reflectance стандартными методами обработки программного комплекса ENVI.

Алгоритм

В ходе исследования стало ясно, что среди специалистов нет единой терминологии относительно спектральных индексов. Некоторые из них имеют одинаковые названия, а некоторые и вовсе «изобретены» дважды. Например, индекс MNDWI, предложенный Hanqiu Xu [6], является, по сути, разработанным Dozier индексом снежного покрова NDSI [7]. При этом в 1996 году было предложено сразу два водных индекса NDWI [8].

В статье [8] также приведены другие примеры одинаковых индексов, основанных на соотношении ближнего (NIR) и коротковолнового (SWIR) инфракрасных каналов, и предложена единая терминология, учитывающая различный диапазон волн ИК-каналов.

В нашей работе были использованы спектральные индексы NDVI [9] и NDII [8], выбранные с учетом доступных спектральных диапазонов сенсора OLI. Подробное описание индексов приведено в таблице 2.

Таблица 2. Описание спектральных индексов

Название индекса	Формула	Спектральные диапазоны (мкм)	Автор
Normalized Difference Vegetation Index (NDVI)	$\frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}}$	ρ_{Red} : 0,60–0,68 ρ_{NIR} : 0,85–0,88	Tucker C.J. (1979)
Normalized Difference Infrared Index (NDII)	$\frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}$	ρ_{NIR} : 0,85–0,88 ρ_{SWIR} : 1,55–1,75	Kimes et al. (1981) Hardisky et al. (1983)

Для проведения исследования были выбраны три сцены Landsat-8 на территорию Томской области в период с мая по июнь 2018 года. Данные сцены характеризуются наличием всех интересующих классов: вода, лед, снег, облачность. Так же стоит отметить, что пойма Оби на данном участке имеет довольно сложный характер подстилающей поверхности, что приводит к ложным результатам применения стандартных масок для Landsat-8 (BQA) и необучаемых классификаций.

В ходе эмпирического исследования был построен пороговый алгоритм (рисунок 1) для выделения классов облачности (clouds), воды (water), льда (ice), снега (snow). Неклассифицированные пиксели обозначаются классом «back». Класс «снег» имеет условное название, и к нему также могут относиться участки светлого льда. Правые ответвления узлов дендрограммы означают выполнение условия, левые – невыполнение. Переменные с литерой «b» соответствуют номерам каналов сенсора OLI, с литерой «t» – сенсора TIRS. Индексы NDVI и NDII рассчитываются согласно формулам в таблице 2.

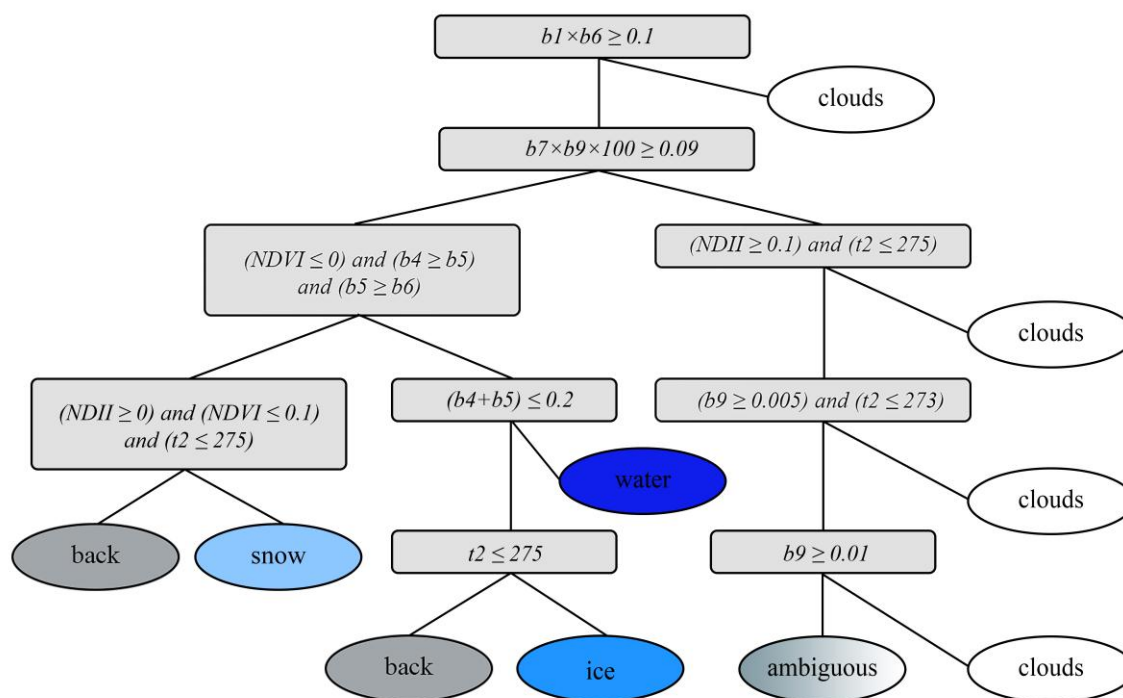


Рис. 1. Пороговый алгоритм для выделения классов облачности, воды, льда, снега по данным Landsat-8 (OLI, TIRS)

Помимо целевых классов в алгоритме также присутствует класс «спорных пикселей» (ambiguous), который может быть определен либо как «облачность», либо как «земля» в зависимости от температуры и рельефа подстилающей поверхности. Данный класс занимает

незначительную долю среди классифицированных пикселей, а окончательное решение о его принадлежности принимает оператор.

Результаты

На рисунке 2 представлены фрагмент сцены Landsat-8 на территорию Томской области в псевдоцветах (05.05.2018) и результат классификации, наложенный на карту.

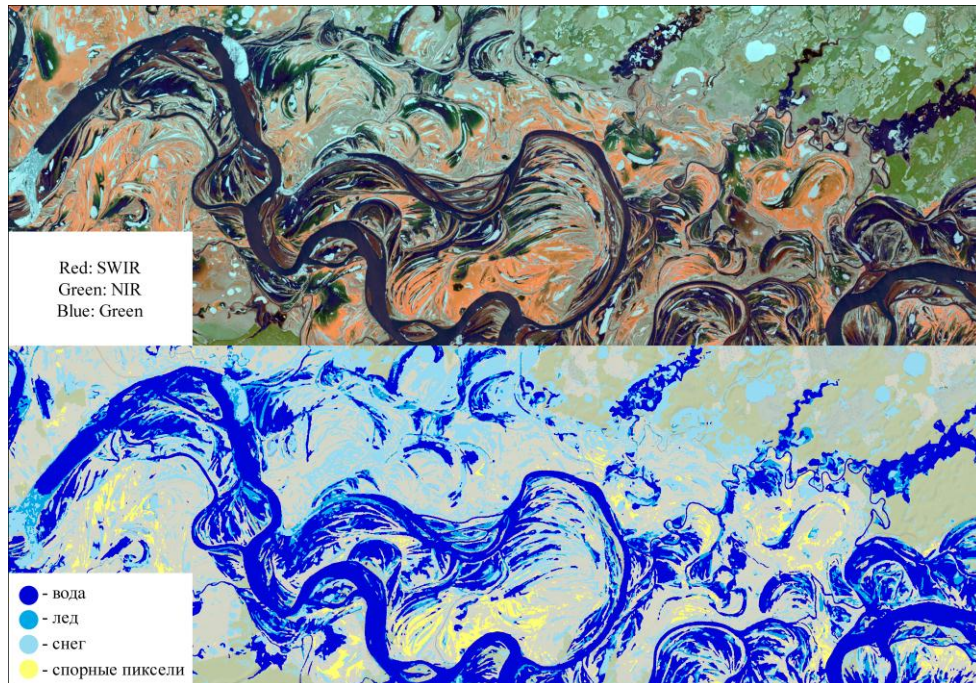


Рис. 2. Фрагмент снимка Landsat-8 (05.05.2018) и результат классификации

Хорошо видно, что в класс «спорных пикселей» попали участки поймы, поэтому при дальнейшей работе над тематической картой указанный класс будет исключен. На рисунке 3 приведены результаты классификации снимка за 08.06.2018.

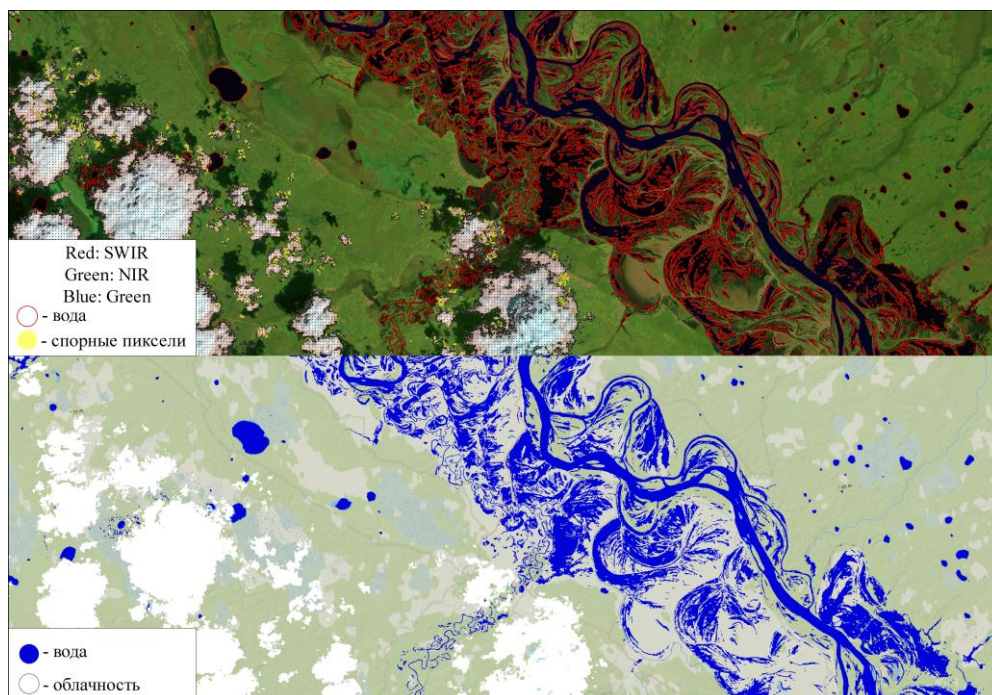


Рис. 3. Фрагмент снимка Landsat-8 (08.06.2018) и результат классификации

В данном случае к классу «спорных пикселей» относится небольшая часть облачности, при последующей обработке эти пиксели будут присоединены к классу облаков.

Оценка точности

Для оценки точности классификации был выбран снимок другой территории (озеро Байкал) за 18 мая 2018 года. Из-за отсутствия наземных данных и принципиальной невозможности их получения, за фактические данные были приняты 1118 точек, выбранных опытным дешифровщиком и отнесенных к четырем классам: «облака», «вода», «лед/снег», «прочее». Для сравнительного анализа была проведена оценка точности результатов обучаемой классификации по методу спектрального угла и использования стандартных масок Landsat-8 (BQA). Ввиду отсутствия класса «воды» в BQA, оценка точности для него проводилась по трем классам с использованием 851 точки. На рисунке 4 представлены результаты классификаций и фрагмент снимка с набранными точками.

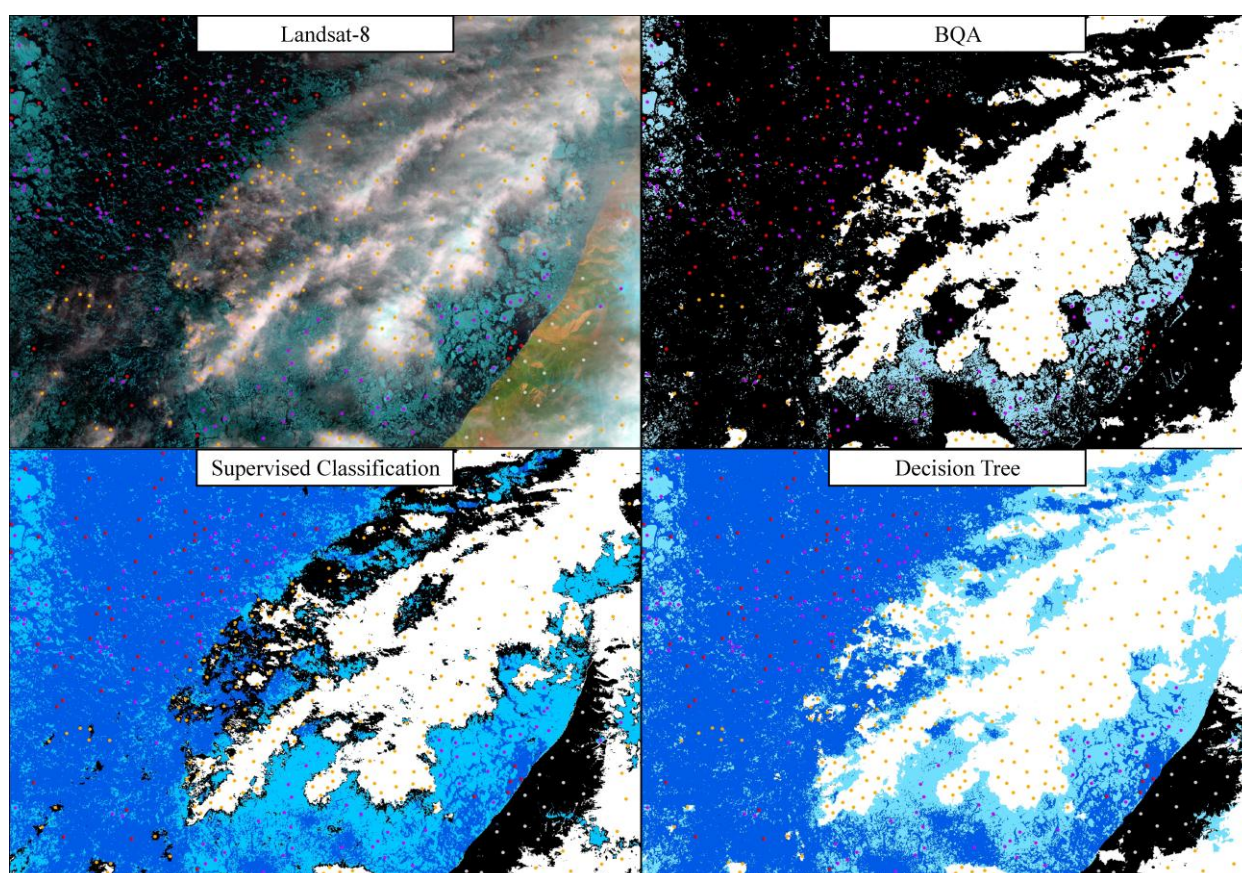


Рис. 4. Фрагмент снимка Landsat-8 (18.05.2018) и результаты классификации

В таблице 3 приведены результаты оценки точности.

Таблица 3. Результаты оценки точности

Показатель	Стандартные маски	Обучаемая классификация	Дерево решений
Количество классов	3	4	4
Количество точек	851	1118	1118
Общая точность (ТА)	0,76	0,81	0,86
Коэффициент Каппа (К)	0,65	0,75	0,81

Результаты классификации с помощью разработанного порогового алгоритма (дерева решений) показали наилучшую общую точность ($TA=0,86$) и степень согласованности ($K=0,81$). Обучаемая классификация лишь в малой степени уступает по точности, однако значительно проигрывает по временным затратам (приблизительно в 5 раз). Применение стандартных масок показало наихудшие результаты: $TA=0,76$; $K=0,65$.

Программный модуль

На втором этапе автоматизации тематической обработки спутниковых данных был создан программный модуль на базе языка IDL. В него интегрирован представленный алгоритм, а также другие специализированные деревья решений, разработанные в Сибирском центре. Алгоритм работы модуля включает в себя выполнение следующих задач:

- определение исходных данных (тип сенсора);
- первичная обработка данных (радиометрическая калибровка);
- классификация с помощью выбранного дерева решений;
- «очистка» результатов классификации;
- экспорт целевых классов в векторный формат.

В процессе опытной эксплуатации модуля были получены средние значения времени, затрачиваемого на разные этапы обработки. Данные значения сравнивались с временными затратами при работе по традиционной схеме в ручном режиме (обучаемая классификация). Полученные результаты и разница между ними (дельта в процентах) указаны в таблице 4.

Таблица 4. Время, затрачиваемое на различные этапы обработки (в минутах)

	<i>Традиционная схема</i>	<i>Автоматическая обработка</i>	<i>Дельта</i>
Первичная подготовка	4	1	75%
Классификация	25	3	88%
Экспорт данных	2	1	50%
Суммарно	31	5	84%

Таким образом, разработанный модуль позволяет сократить общее время обработки приблизительно на 80% благодаря трем факторам: автоматизации рутинных действий, выполняемых оператором, автоматическому выполнению классификации и отсутствию необходимости загрузки (визуализации) изображений, что обычно занимает значительное время.

Заключение

Разработанная технология позволила значительно сократить время подготовки тематических карт паводковой и ледовой обстановки в оперативном режиме, не жертвуя при этом точностью классификации. Тем не менее, представленный алгоритм требует дальнейшей апробации на более масштабных наборах данных и, возможно, доработки пороговых условий с целью повышения точности. Кроме того, предполагается дальнейшее расширение функционала программного модуля для обеспечения возможности обработки данных с других космических аппаратов.

References

- [1] Kataev M.Yu., Bekerev A.A., Methodology of the water objects detection from multi-spectrum satellite measurements, *Proceedings of TUSUR journal*, 2017, part 20, №4. DOI: 10.21293/1818-0442-2017-20-4-105-108. (In Russian).

- [2] Moradi M., Sahebi M., Shokri M., Modified optimization water index (MOWI) for Landsat-8 OLI/TIRS, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2017, XLII-4/W4, pp. 185-190. DOI: 10.5194/isprs-archives-XLII-4-W4-185-2017.
- [3] Zhou Y., Dong J., Xiao X., Xiao T., Yang Z., Zhao G., Zou Z., Qin Y., Open surface water mapping algorithms: a comparison of water-related spectral indices and sensors, *Water*, 2017, 9, p. 256. DOI: 10.3390/w9040256
- [4] Berkovich K.M., Ruleva S.N., Chalov R.S., River Ob, *Popular science encyclopedia «Water of Russia»*, <http://water-rf.ru/a1754>. (In Russian).
- [5] Loveland T.R., Irons J.R., Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 2016, 185, 1–6. DOI: 10.1016/j.rse.2016.07.033
- [6] Xu H., Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, *International Journal of Remote Sensing*, 2006, Vol. 27, No. 14, pp. 3025–3033. DOI: 10.1080/01431160600589179
- [7] Hall D.K., Riggs G.A., Normalized-difference snow index (NDSI), *Encyclopedia of Snow, Ice and Glaciers*, 2010, pp. 779–780. DOI: 10.1007/978-90-481-2642-2_376
- [8] Ji L., Zhang L., Wylie B.K., Rover J.R., On the terminology of the spectral vegetation index (NIR – SWIR)/(NIR + SWIR), *International Journal of Remote Sensing*, 2011, Vol. 32, No. 21, pp. 6901–6909, DOI: 10.1080/01431161.2010.510811.
- [9] Tucker C.J., Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sensing of Environment*, 1979, 8, pp. 127–150. DOI: 10.1016/0034-4257(79)90013-0