ACCURACY ANALYSIS OF THE INLAND WATERS DETECTION

Marina Gudelj[1]**, MSc**

Assist. Prof. Dr. Mateo Gašparović[1]

Assoc. Prof. Dr. Mladen Zrinjski[1]

[1]University of Zagreb, Faculty of Geodesy, **Croatia**

ABSTRACT

Climate changes and human activities on Earth's surface effect on natural Earth's resources as well as on inland waters. Any changes in the Earth's atmosphere, such as changes in temperature, humidity, precipitations and their intensity affect the inland waters area and level. In order to carry out the monitoring of the areas under the inland waters economically, the satellite imageries collected by remote sensing sensors are used. Satellite imagery used for this research were cloud-free Landsat-8 imagery at study area Zagreb. The classification of satellite imagery gives an insight into the state of the Earth's surface at a given moment. Today, many classification methods of satellite imagery have been developed, but in order to find the most accurate classification method for the extraction of inland waters, this research compare the accuracy assessment of the classification methods. For purposes of this research, the satellite imageries are classified into two classes, inland waters and others. Firstly, the DOS1 atmospheric correction was performed in QGIS software. Then study area Zagreb was classified using supervised classification methods Maximum Likelihood Classification (MLC) and Random Forests (RF). Both supervised classification methods were performed on six Landsat-8 30 m spatial resolution bands and same training polygons. For this research, normalised difference water index (NDWI), modified normalised difference water index (MNDWI) and automated water extraction index (AWEI) were created and used for classifying satellite imagery scene in two classes. Spectral indices, supervised classification, as well as unsupervised classification, were made in open software SAGA GIS. In order to compare all the methods of extraction inland waters, visual inspection and objective analysis were carried out. The objective analysis was performed by determining a figure of merit, overall agreement, omission and commission. By objective and subjective analysis, the best method for extraction of inland waters is the Random Forests method of supervised classification whose overall agreement 99.78%. The second method, whose accuracy is slightly lower than the RF method, is a method based on MNDWI and unsupervised classification whose overall agreement 99.71%. Accuracy assessment of the MLC method is lower than the previous two methods whose overall agreement 99.41%. The accuracy of methods based on NDWI and AWEI are worse. Overall agreement, for a method based on AWEI and k-means unsupervised classification, is 83.23%, while overall agreement for a method based on NDWI and k-means unsupervised classification is 79.05%.

Keywords: inland water monitoring, inland water detect, unsupervised classification, supervised classification, spectral indices

INTRODUCTION

Water is a substance composed of the chemical elements hydrogen and oxygen which is necessary for life on Earth. Water is necessary for natural processes, the life of plants, animals and humans. On inland water level affecting climate change such as seasonal changes. Furthermore, the amount of exhaust gases produced by human activity affect climate changes and, consequently, on the water level. Scientists predict that water will be more valuable in the future and will be the source of many conflicts. Since water is of great importance for maintaining life on Earth, it is necessary to monitor its quality, water level and the area it occupies. Monitoring of the inland waters areas was carried out using remote sensing methods. Land-cover classification (LCC) methods for satellite imagery have been developed and tested in various remote sensing research [1], [2], [3]. For the purposes of this research, the satellite imagery scene is classified by different classification methods, Maximum Likelihood Classification (MLC), Random Forests (RF) and K-means unsupervised classification. K-means unsupervised classification is performed on the spectral indices. This research aims to compare the accuracy of water extraction that can be obtained by different classification methods.

STUDY AREA AND DATA

For this research, the study area covered the area surrounding the city of Zagreb was taken (Figure 1). Zagreb is the capital city of Croatia with about 790000 inhabitants. Study area Zagreb was chosen because it contains a highly diverse landscape inland water, vegetation, bare land and urban areas.

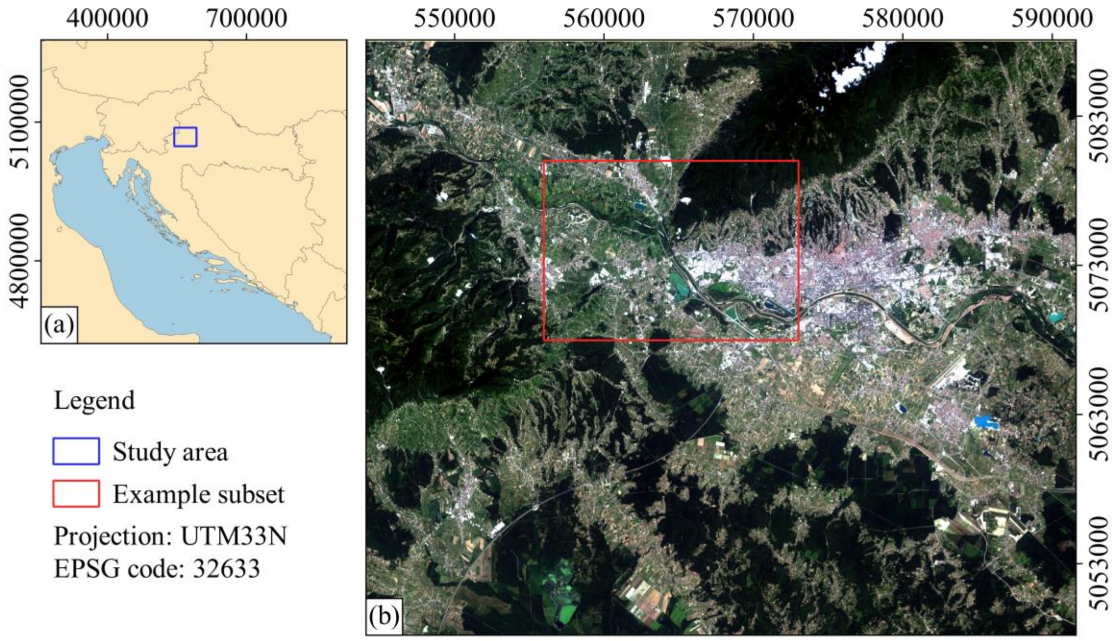

Figure 1. (a) Geographic location of the study area; (b) the study area Zagreb.

Satellite imagery used for this research were cloud-free Landsat-8 imagery (Table 1) in the summer period. Landsat-8 satellite imagery was downloaded from the USGS Earth Explorer service [4]. Landsat-8, launched in 2013, has two sensors for data collection Operational Land Imager and Thermal Infrared Sensor. These two sensors provide seasonal coverage of the global landmass at a spatial resolution of 30 meters (visible, near infrared, shortwave infrared); 100 meters (thermal); and 15 meters (panchromatic). The Landsat-8 scene size is 185 km × 180 km [5].

Table 1. Landsat-8 satellite imagery used in this research.

Study area	Sensing date	Satellite imagery ID
Zagreb	19 July 2015	LC08_L1TP_189028_20150719_20170407_01_T1

METHODS

In order to monitor the inland water surface more accurate, the accuracy of inland waters extraction has been made with various classification methods of satellite imageries. First, the pre-processing of satellite imageries was made. DOS1 atmospheric correction was performed in Quantum GIS software (QGIS, version 2.18.19). This method is found to be data dependent and well accepted by the geospatial community to correct light scattering in remote sensing data [6]. Satellite imageries of study area Zagreb were georeferenced in the WGS 84 UTM 33N coordinate system. Then the extraction of the inland water areas was done by supervised classification methods Maximum Likelihood Classifier (MLC) [7] and Random Forests (RF) [8]. Supervised methods require training polygons of every class to classify the satellite imagery. Satellite imagery scene was divided into two classes inland waters (rivers, lakes, morass and ponds) and the others (forests, woody vegetation, cropland, grassland, rocks, soil, built-up). The same training polygons were used for both classification methods, while the classification was performed on six Landsat-8 30 m spatial resolution bands. On the other hand, unsupervised classification methods do not require knowledge of the terrain and taking training polygons. Unsupervised classification methods require the prior definition of final classes number. The drawback of such methods is the classification of the scene in an unknown, random, classes. For purposes of this research k-means unsupervised classification method was used [9]. To improve unsupervised classification methods, k-means unsupervised classification is based on spectral indices [10]. Spectral indices are rasters that are obtained by the various combinations of satellite imageries bands, and they have been proposed to emphasise areas where certain classes are dominant. Nowadays, there are a lot of spectral indices that emphasise water. For the purpose of this research, normalised difference water index (NDWI), modified normalised difference water index (MNDWI) and automated water extraction index (AWEI) were used. NDWI is a measure of liquid water molecules in vegetation canopies that interacted with the incoming solar radiation. NDWI was first proposed to detect surface waters in wetland environments and to allow for the measurement of surface water extent. Although the index was created for use with Landsat Multispectral Scanner (MSS) image data, it has been successfully used with other sensor systems in applications where the measurement of the extent of open water is needed. The NDWI is calculated as follows [11], [12]:

$$NDWI = \frac{NIR - SWIR1}{NIR + SWIR1}, \quad (1)$$

where *NIR* is a near-infrared band such as Landsat-8 band 5, where *SWIR1* is a Short Wavelength Infrared band such as Landsat-8 band 6. The MNDWI is one of the most widely used water indices for various applications, including surface water mapping, land-cover change analyses and ecological research. The MNDWI can be expressed as follows [13]:

$$MNDWI = \frac{Green - MIR}{Green + MIR}, \quad (2)$$

where *Green* is a green band such as Landsat-8 band 3, *MIR* is a middle infrared band such as Landsat-8 band 6. The AWEI is an index formulated to effectively eliminate non-water pixels, including dark built surfaces in areas with urban background. AWEI is obtained as [14]:

$$AWEI = 4 \cdot (Green - NIR) - 0.25 \cdot Red + 2.75 \cdot SWIR2, \quad (3)$$

where *Green* is a green band such as Landsat-8 band 3, *NIR* is a near-infrared band such as Landsat-8 band 5, *Red* is a red band such as Landsat-8 band 4 and *SWIR2* is a Short Wavelength Infrared band such as Landsat-8 band 7. Figure 2 shows the workflow of this research.

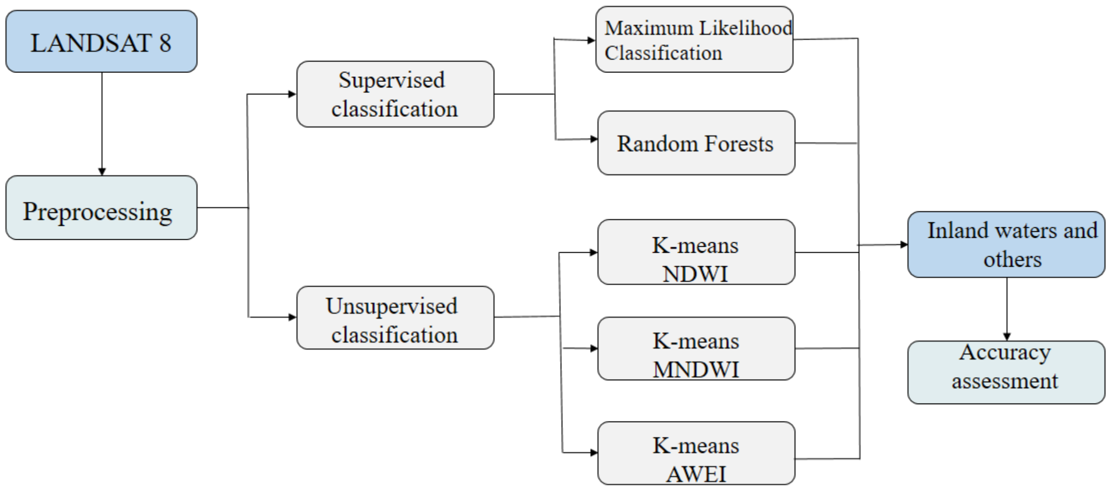

Figure 2. The workflow of this research.

In order to determine which method is the best for extract inland waters areas, the classification accuracy was done. Accordingly to Pontius and Millones [15] the figure of merit (F), overall agreement (A), omission (o) and commission (c) were used for quantitative accuracy assessment. All classification methods were made in SAGA GIS (version 6.2.0). Since SAGA GIS is open software, available to everyone, it allows the application of inland waters detection and monitoring to all those interested.

RESULTS

The satellite scene is classified into two classes, inland waters and others, with different classification methods. In order to determine the best classification method by visual inspection, Figure 3 was made.

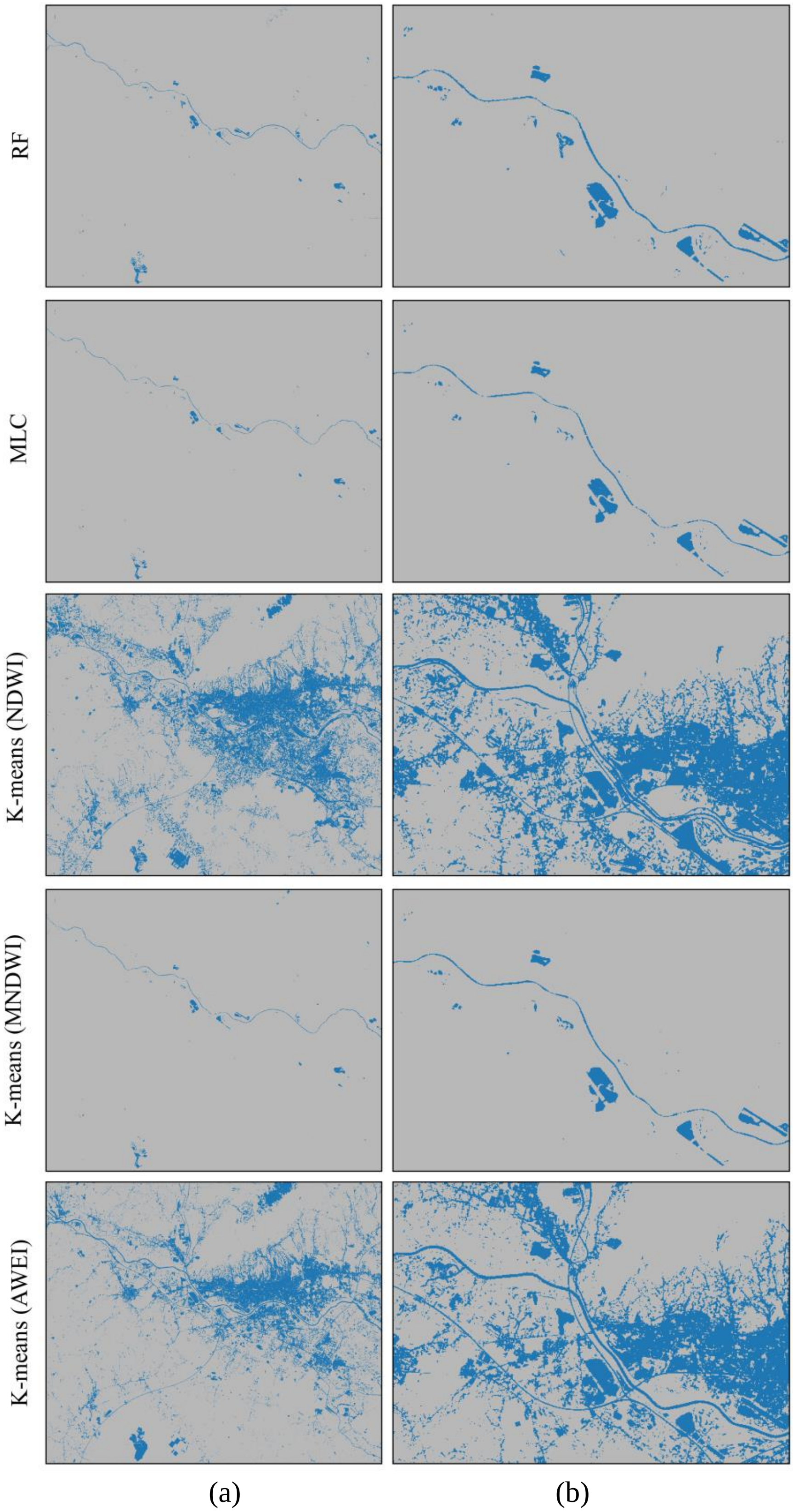

Figure 3. The results of applied classification methods for extracting inland waters on (a) entire study area, and (b) example subset. Blue colour represents class inland waters, and grey colour represents class others.

Figure 3 shows how different classification methods distinguish water surfaces. By visual inspection, it can be seen that classification methods based on k-means unsupervised classification and spectral indices AWEI and NDWI have been the worst for the extraction of inland waters areas. Furthermore, Figure 3 shows that supervised classification methods MLC and RF, as well as unsupervised k-means classification method based on MNDWI spectral index, similarly distinguish inland waters. By visual inspection, the RF classification method of satellite imagery scene showed slightly better than the MLC classification method and the k-means classification method which is based on MNDWI raster. How the accuracy analysis would not only be done by a subjective experience, objective accuracy was made, for all classification methods, on the same polygon samples. Table 2 shows the results of the objective analysis.

Table 2. Classification accuracies for five classification methods.

Study area	Classification method	Inland waters			
		F [%]	o [%]	c [%]	A [%]
Zagreb	RF	96.72	0.15	0.07	99.78
	MLC	91.05	0.55	0.04	99.41
	K-means (NDWI)	23.73	0.02	20.93	79.05
	K-means (MNDWI)	95.53	0.25	0.05	99.71
	K-means (AWEI)	28.03	0.00	16.77	83.23

The results shown in Table 2 are objective confirmation of the visual inspection. According to Table 2, the worst classification method for extracting inland waters areas is an unsupervised k-means method, which is based on the NDWI spectral index whose figure of merit, overall agreement, omission and commission are 23.73%, 79.05%, 0.02% and 20.93%. Furthermore, a method which has also shown as bad for extracting inland waters, but better than the method based on k-means and the NDWI spectral index, is a method based on the k-means classification and the AWEI spectral index raster. The figure of merit, overall agreement, omission and commission for a method based on AWEI and k-means unsupervised classification are 28.03%, 83.23%, 0.00%, 16.77%, respectively. Classification methods MLC, RF and k-means based on MNDWI spectral index have shown as good for extracting inland waters areas, and the difference in their overall agreement is small. According to Table 2, the best method for extraction of inland waters is the Random Forests method of supervised classification. Accuracy assessment indicators figure of merit, overall agreement, omission and commission for the RF method are 96.72%, 99.78%, 0.15%, 0.07%, respectively. The second method, whose accuracy is slightly lower than RF method, is a method based on MNDWI and unsupervised classification whose figure of merit, overall agreement, omission and commission are 95.53%, 99.71%, 0.25% and 0.05%. Accuracy of the MLC method is lower than the previous two methods whose figure of merit, overall agreement, omission and commission are 91.05%, 99.41%, 0.55% and 0.04%.

DISCUSSION

The main problem of mankind's progress and the increase in living standards is that progress is not being made in harmony with nature. Environmental pollution is one of the byproducts of progress that lead to climate change that further reflects on the resources that are necessary for the life we know. Today, the quality of air, water and food is damaging. In order to sustain the resources we need, it is necessary to detect and monitor the changes that occur with them timely. This research provides a comparison of the accuracy of inland waters separation with different classification methods to make monitoring as good as possible. A comparison of accuracy was performed on MLC, RF, k-means based on NDWI spectral index, k-means based on MNDWI spectral index and k-means based on AWEI spectral index classification methods. The results have shown that inland waters separation is possible to quality perform by RF, MLC and k-means based on the MNDWI spectral index classification methods. Although the results of the three above-mentioned classification methods are similar, the RF method proved to be the best. This research was done on Landsat-8 30 m spectral resolution satellite imagery. Future research could inspect the accuracy of the inland waters extraction by various classification methods on satellite imageries with higher spatial resolution.

CONCLUSIONS

This study deals with the comparison of various methods of separating inland waters. Water, a molecule of two elements, hydrogen and oxygen, is essential for life on Earth and its quality detection on the Earth’s surface is of great importance. Water extraction was done using remote sensing, on Landsat satellite imagery, more specifically on Landsat-8 satellite imagery. The Landsat mission was chosen because forty decades of continuous data collection has made the invaluable archive, which allows monitoring and comparison of various phenomena. For the study area, the area of town Zagreb was chosen because it contains variety of landscapes and quite of inland waters. In the research, the comparison of the accuracy of the classification was performed subjectively (visual inspection) and objectively (determination figure of merit, overall agreement, omission and commission). The best classification method for extracting inland waters is RF. MLC and k-means based on MNDWI spectral index classification methods can be used to extract the inland waters. Classification methods k-means based on NDWI raster and k-means based on AWEI raster should be avoided for inland waters extraction. Although this research provides the best method of extracting inland waters used for monitoring, it should not be the only one to be used for the keeping and protection of water. This study detects changes in the areas covered by inland waters. The importance of water is so invaluable that its keeping cannot rely only on one research, this research should only be one part of a comprehensive study. Furthermore, water quality control, diversity of flora and fauna, the amount of individual gases, etc. should be conducted and inspected. Water keeping requires association and contribution of different professions.

REFERENCES

[1] Foody, G. M. Status of land cover classification accuracy assessment, Remote Sensnsing of Environment, 80, 185-201, 2002.

[2] Gašparović, M., Jogun, T. The effect of fusing Sentinel-2 bands on land-cover classification, International Journal of Remote Sensing, 39, 822-841, 2018.

[3] Gašparović, M., Zrinjski, M., Gudelj, M. Analysis of Urbanization of Split, Geodetski list, 71(3), 189-202, 2017.

[4] USGS – EarthExplorer, https://earthexplorer.usgs.gov/, (1.9.2018).

[5] Jia, K., Wei, X., Gu, X., Yao, Y., Xie, X., Li, B. Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China, Geocarto International, 29, 941-951, 2014.

[6] Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., Macomber, S. A. Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects, Remote Sensing of Environment, 75, 230-244, 2001.

[7] Otukei, J. R., Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms, International Journal of Applied Earth Observation and Geoinformation, 12, 27-31, 2010.

[8] Breiman, L. Random Forests, Machine learning, 45, 5-32, 2001.

[9] Hartigan, J. A., Wong, M. A. Algorithm AS 136: A K-Means Clustering Algorithm, Journal of the Royal Statistical Society, C 28, 100, 1979.

[10] Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., Liu, J., Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., Liu, J. Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index, Remote Sensing, 9, 249, 2017.

[11] McFeeters, S. K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, International journal of remote sensing, 17(7), 1425-1432, 1996.

[12] Gao, B. C., NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space, Remote sensing of environment, 58(3), 257-266, 1996.

[13] Xu, H., Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery, International Journal of Remote Sensing, 27, 3025–3033, 2006.

[14] Feyisa, G. L., Meilby, H., Fensholt, R., Proud, S. R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery, Remote Sensing of Environment, 140, 23-35, 2014.

[15] Pontius Jr, R. G., & Millones, M., Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment, International Journal of Remote Sensing, 32(15), 4407-4429, 2011.