

DETECTION OF SHALLOW WATER AREA WITH MACHINE LEARNING ALGORITHMS

N. Yagmur^{1*}, N. Musaoglu¹, G. Taskin²

¹ ITU, Civil Engineering Faculty, Department of Geomatics Engineering 34469 Maslak Istanbul, Turkey
(yagmurn, musaoglune)@itu.edu.tr

² ITU, Institute of Earthquake Engineering and Disaster Management, 34469 Maslak Istanbul, Turkey
gulsen.taskin@itu.edu.tr

KEY WORDS: Shallow water, Remote sensing, Machine learning, Water indices, SVM, Feature selection.

ABSTRACT:

Remote sensing techniques has been widely used for detecting water bodies in especially wetlands. Different classification methods and water indices has used for this purpose and there are numerous studies for detecting water bodies. However, detecting shallow water area is difficult comparing with deep water bodies because of the mixed pixels. Akgol Wetland is chosen as study area to detect shallow water. For this purpose, Sentinel 2 satellite image, which gives more accurate results thanks to higher spatial resolution than the images having medium spatial resolution, is used. In this study, two classification approaches were applied on Sentinel 2 image to detect shallow water area. In the first approach, effectiveness of indices was determined and classification of spectral bands with indices shows higher accuracy than classification of only spectral bands by using support vector machine classification method. In the second approach, support vector machine recursive feature elimination method used for the most effective features in the first approach. Besides overall accuracy of only spectral bands is obtained as 88.10%, spectral bands and indices' accuracy was obtained as 91.84%.

1. INTRODUCTION

Water has significant role in ecological cycle. The surface area of the water bodies diminishes because of climate and environmental change. Six percent of the world land surfaces are wetlands consisting of shallow water area in general (Erwin, 2019). As a terrestrial resource wetlands have numerous functions in terms of hydrological such as storage of water, prevention of floods, coastal stabilization (Bergkamp and Orlando, 1999) and ecological such as storing carbon, preserving biological diversity and critical habitat for many migrating birds (Coban, 2017).

Remote sensing technologies have widely been used in detection of water body. There are numerous studies which focused on wetlands or water bodies using remote sensing images and techniques. Halabisky et al. (2016) used spectral mixture analysis to classify surface-water extent of wetlands at scales below 30 m. Niemuth et al. (2010) used a combination of supervised classifications and photo interpretation techniques to determine wetlands in the Prairie Pothole Region.

For a more accurate estimation of wetlands, several water indices have been developed for detecting water area and jointly been used with spectral bands of the satellite images (Ma et al, 2019). Acharya et al. (2017) applied four different indices to Landsat 8 OLI satellite images of to Phewa Lake's. The classification with the combination of these indices, including Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), Water Ratio Index (WRI), and Normalized Difference Vegetation Index (NDVI) shows better results than those of spectral bands. Fisher et al. (2016) utilized multiple water indices which are Automated Water Extraction Index with shadow (AWEIsh) and with no shadow (AWEInsh), Tasseled Cap Wetnes (TCW_{Crist}) NDWI and Water Index (WI₂₀₁₅) for multi-temporal Landsat (TM, ETM and OLI) satellite images and analysed performance and accuracy results

of afformentioned indices on pure and mixed water pixels. Zhou et al. (2017) applied multiple water indices on Landsat 7 ETM+, Landsat 8 OLI, and Sentinel-2 MSI satellite images to show the effectiveness of the indices in classification performance. Acharya et al. (2016) performed different classification methods using water indices with the aim of analysing performance of the indices in detecting water area, and the results showed that support vector machine achieves higher classification accuracy than the other methods.

Machine learning algorithms have gained popularity (Maxwell et al., 2018) for classification of land use/cover such as wetland classification. Image classification methods automatically categorize all of the pixels in an image into land use/cover classes. Franklin et al. (2018) used object and pixel based machine learning algorithms to classify Hudson Bay Lowlands Ecoregion wetlands area using optical and radar data. Abdel-Hamid et al. (2018) used different machine learning classification algorithms for mapping mangroves Extents on the Red Sea coastline in Egypt using radar and high resolution optical data.

One of the most basic classification application is to distinguish water bodies from dry land surfaces. Spectral signature of water is changeable depending on the height of the water, especially in the shallow water. The shallow water area cannot be easily detected because the spectral reflectance of the shallow water might include vegetation or soil land types. Such mixing areas especially exists in the wetlands, and their detection determines the coastline changes in the wetlands. In this study, two different classification approaches were applied on Sentinel 2 image to detect the shallow water area. In the first approach, only six spectral bands were used as the features, while nine water indices were used in conjunction with spectral bands. In classification support vector machine (SVM) with radial basis kernel was conducted. The results showed that feeding water indices to classifier improves the classification accuracy. Besides, the

* Corresponding author

feature selection method, so called SVM recursive feature elimination (SVM-RFE), was utilized to determine which water index has more impact on the classification performance.

Deep water bodies can be detected easily with classification and water indices which has been quite often used in literature. However, in the literature there has not been encountered any study about detecting shallow water using combined different methods. In this study, Akgol is a wetland which has shallow water as mentioned before in definition of wetland. Water depth in Akgol Wetland has not exceed to 1 m and detecting shallow water are more difficult than detecting deep water bodies. This study offers two different classification approaches for detecting shallow water area with successful results.

2. STUDY AREA

The study area is chosen as Akgol Wetland where is located in Konya Closed Basin (KCB). The KCB is located in Central Anatolia of Turkey, which is one of the seven basins in Turkey covering 7% of Turkey's surface area. Akgol Wetland is nature conservation area and therefore has an ecological importance. The study area is shown in the Figure 1.

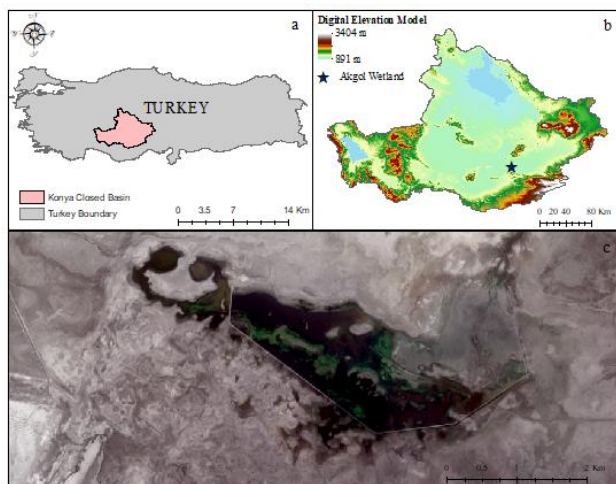


Figure 1. Geographical location of (a) Turkey, (b) Konya Closed Basin (KCB), (c) Akgol Wetland.

3. DATASET AND METHODOLOGY

In this study, Sentinel 2 satellite image, which is freely available since 2015, was used as the dataset. Sentinel 2 satellite have 13 spectral bands which change between 0.665 and 2.190 μm wavelengths in electromagnetic spectrum. Sentinel 2 satellite images have 10 m. spatial resolution, Level 2 top of atmosphere (TOA), and Level 1 bottom of atmosphere (BOA) reflectance and BOA image was used in this study (URL – 1). Sentinel 2 satellite images have high spatial resolution and higher accuracy compared to the images having medium high spatial resolution such as Landsat (Zhou et al., 2017). Moreover, in situ data was collected at the same time with the satellite image. 167 sample points were used for classification of satellite image and collected from 4 classes which are 1 (deep water comparing with another classes), 2 (shallow water), 3 (more shallow water) and 4 (other). Figure 2 shows the sample points.

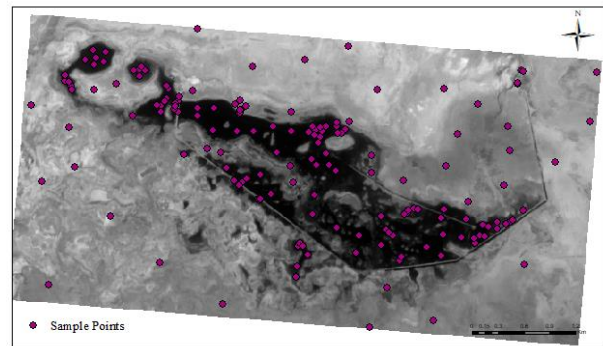


Figure 2. Sample points on Sentinel 2 NIR band.

The water indices have been widely used for detection of water bodies, in this way, different water indices have been developed based on visible (VIS), near infrared (NIR), and shortwave infrared (SWIR) bands of satellite images and separates the water and background according to some predefined threshold values. In this study, nine indices which are Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), Automated Water Extraction Index with shadow (AWEIsh) and with no shadow (AWEInsh), Normalized Difference Pond Index (NDPI), Water Ratio Index (WRI), two types of Tasseled Cap Wetness (TCW) and Water Index (WI) are used for determining effect of the detecting shallow water area. The definition of the indices is given in Table 1.

Index	Formula	Reference
NDWI - 1	$(B_{\text{green}} - B_{\text{Nir}}) / (B_{\text{green}} + B_{\text{Nir}})$	Mcfeeters (1996)
MNDWI	$(B_{\text{green}} - B_{\text{swir}}) / (B_{\text{green}} + B_{\text{swir}})$	Xu (2006)
NDPI	$(B_{\text{swir}} - B_{\text{green}}) / (B_{\text{swir}} + B_{\text{green}})$	Lacaux et al. (2007); Ji et al., (2009); Feyisa et al. (2014)
AWEIsh	$B_{\text{blue}} + 2.5 \times B_{\text{green}} - 1.5 \times (B_{\text{Nir}} + B_{\text{swir}1}) - 0.25 \times B_{\text{swir}2}$	Feyisa et al. (2014)
AWEInsh	$4 \times (B_{\text{green}} - B_{\text{swir}1}) - (0.25 \times B_{\text{Nir}} + 2.75 \times B_{\text{swir}2})$	Feyisa et al. (2014)
TCW _{cris}	$0.0315 \times B_{\text{blue}} + 0.1973 \times B_{\text{green}} + 0.3279 \times B_{\text{red}} + 0.3406 \times B_{\text{Nir}} - 0.7112 \times B_{\text{swir}1} - 0.4572 \times B_{\text{swir}2}$	Crist (1985)
WRI	$(B_{\text{green}} + B_{\text{red}}) / (B_{\text{Nir}} + B_{\text{swir}1})$	Shen & Li (2010)
TCW	$0.1509 \times B_{\text{blue}} + 0.2021 \times B_{\text{green}} + 0.3102 \times B_{\text{red}} + 0.1594 \times B_{\text{Nir}} - 0.6806 \times B_{\text{swir}1} - 0.6109 \times B_{\text{swir}2}$	Crist (1984; 1985)
WI ₂₀₁₅	$1.7204 + 171 \times B_{\text{green}} + 3 \times B_{\text{red}} - 70 \times B_{\text{Nir}} - 45 \times B_{\text{swir}1} - 71 \times B_{\text{swir}2}$	Fisher et al. (2016)

Table 1. Definition of spectral water indices used in this study.

SVM was used to determine the shallow water. SVM was created as a statistical learning theory (Vapnik et al., 1997), finds an

optimum hyperplane by maximizing the margin between the classes. SVMs is quite useful in classification problems. In this study, SVM was used with radial basis function (RBF) kernel.

SVM has also been used for feature selection, and in this study SVM-RFE method was used for choosing the most effective water index in detection of the shallow water. The SVM-RFE algorithm has been proposed by Guyon et al. (2002). Based on the SVM classification, SVM-RFE ranks the features based on the criterions using the weight magnitude (Guyon, et al., 2002).

4. EVALUATION AND RESULTS

The sample points were randomly divided into two sets: training and test, and nine indices were used as features with six multispectral bands (Red, Green, Blue, NIR, SWIR, and SWIR2). In total, the samples with 15 features was generated for a possible classification model.

Firstly, the effects of the indices were determined with classification. Training data was classified with two different types: with indices and with no indices for the purpose of determining indices effect of the detecting shallow water area and classification results are shown in Table 2. According to the results, utilizing water indices with spectral features dramatically increases the overall accuracy. Accuracy of the first class is seen as lowest in classification; but considering with one hundred trials, it can reach 85.71 because of the training set used during each trial. However, mean of these trial values shown in the following table.

	Classes' Accuracy (%)				Overall Accuracy (%)
	1	2	3	4	
Spectral bands	52.28	92	88.8	98.26	88.10
Spectral bands + water indices	100	84.62	90	94.74	91.84

Table 2. Accuracy assessment of two classification approaches.

Figure 3 shows the averaged overall classification accuracy with respect to the selected features obtained by SVM-RFE based on the one hundred trials.

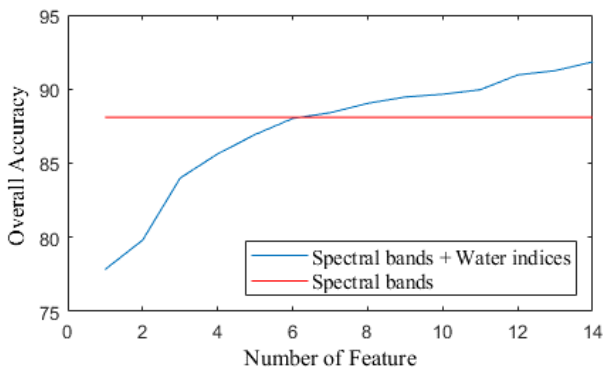


Figure 3. Trend of overall accuracy when indices were added to classification.

Red line shows the overall accuracy of classification with six spectral bands. According to Figure. 3, when number of feature reach to 6, overall accuracies of these two classifications are equal each other. Note that ranking of the most important features might change depending on the training set used during each trial. Figure 4 shows the ranked features obtained with hundred trials. This figure infers that in overall the most effective features are Blue, Red, Green, NDWI, WRI, and SWIR2. X axis of the figure shows the hundred trials and y axis of the figure shows the first six features ranking according to significance.

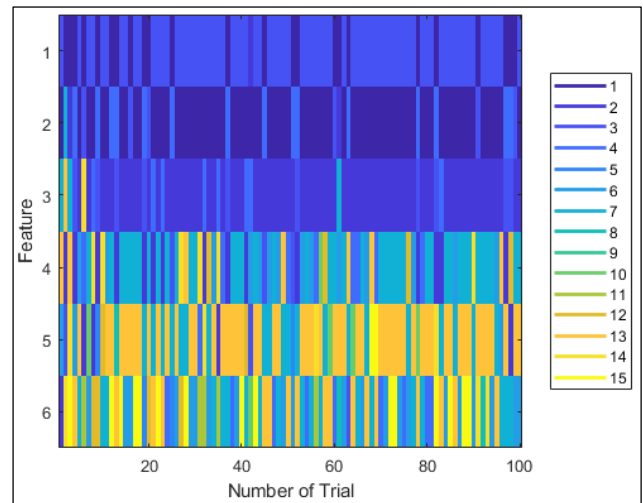


Figure 4. Ranked features according to hundred trial.

Positive effects of water indices in terms of accuracy were proved with Figure 3. According to these results, Sentinel 2 image was classified with SVM with RBF kernel. Figure 5 and 6 show the classification with only using spectral bands and the classification with spectral bands and water indices, respectively.

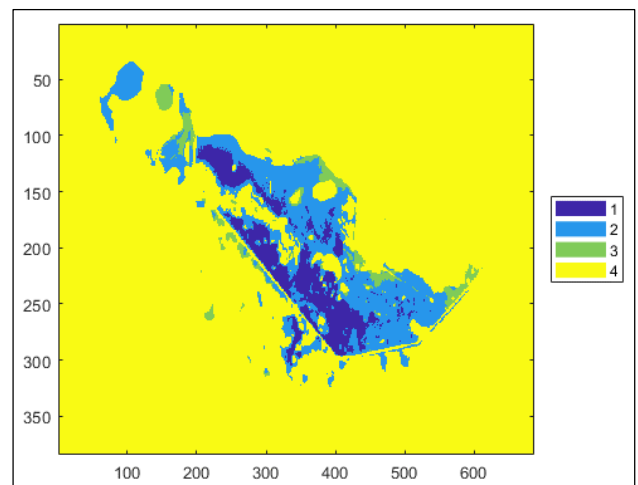


Figure 5. Classification result with only spectral bands.

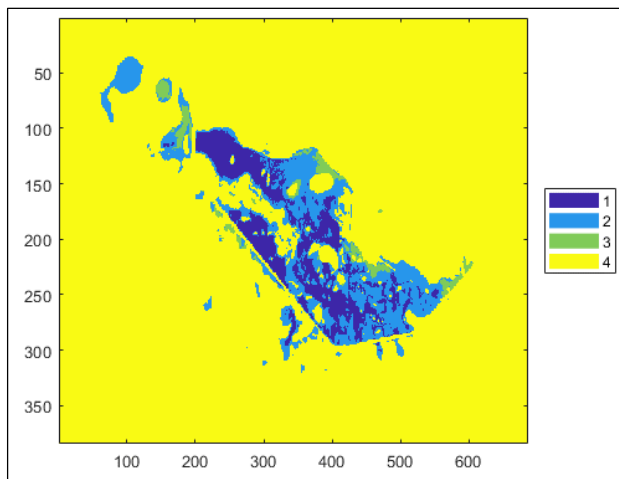


Figure 6. Classification result with spectral bands and water indices.

The differences between thematic maps are seen in the Figure 5 and 6. In this study, the first class namely as deep water is a slightly higher than other classes. The deep water, shows as dark blue colour and the maximum change is seen in this class. According to overall accuracies, thematic map of the classification with spectral bands and indices has more accurate in this study.

5. CONCLUSION

According to results, it is determined that using indices in classification problems, it can be very usable and increases the overall accuracy. Feature selection method is important in terms of detecting the most significance features (spectral bands or indices) in classification. This study is expected to be guide to future studies about water body and shallow water area detection. Based on the further studies, in situ data has significant role in especially shallow water related studies.

ACKNOWLEDGEMENTS

The authors would like to express their thanks to the National Scientific and Technological Research Council of Turkey (TUBITAK) for their financial support to Project number 116Y142.

REFERENCES

Abdel-Hamid, A., Dubovyk, O., Abou El-Magd, I., & Menz, G., 2018. Mapping Mangroves Extents on the Red Sea Coastline in Egypt using Polarimetric SAR and High Resolution Optical Remote Sensing Data. *Sustainability*, 10(3), 646.

Acharya, T. D., Subedi, A., Yang, I. T., & Lee, D. H., 2017. Combining Water Indices for Water and Background Threshold in Landsat Image. In *Multidisciplinary Digital Publishing Institute Proceedings*, 2(3), 143.

Acharya, T., Lee, D., Yang, I., & Lee, J., 2016. Identification of water bodies in a Landsat 8 OLI image using a J48 decision tree. *Sensors*, 16(7), 1075.

Bergkamp, G. and Orlando, B. 1999. Exploring Collaboration between the Convention on Wetlands (Ramsar, Iran 1971) and

the UN Framework Convention on Climate Change, Climate Initiative, IUCN, Washington, USA.

Coban, A. 2017. The effect of climate change on wetlands and the situation of wetlands in Turkey, *The Turkish Journal of Occupational/Environmental Medicine and Safety*, 2.1 (3): 3-3.

Crist, E.P., 1985. A TM Tasseled Cap Equivalent Transformation for Reflectance Factor Data. *Remote Sensing of Environment*, 17, 301–306.

Crist, E.P.; Cicone, R.C., 1984. A Physically-Based Transformation of Thematic Mapper Data-The TM Tasseled Cap. *IEEE Trans. Geoscience Remote Sens.*, 22, 256–263.

Erwin, K. L., 2009. Wetlands and global climate change: the role of wetland restoration in a changing world. *Wetlands Ecology and management*, 17(1), 71.

Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014. Automated water extraction index: a new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, 140, 23–35.

Fisher, A., Flood, N., & Danaher, T., 2016. Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sensing of Environment*, 175, 167-182.

Franklin, S. E., Skeries, E. M., Stefanuk, M. A., & Ahmed, O. S., 2018. Wetland classification using Radarsat-2 SAR quad-polarization and Landsat-8 OLI spectral response data: a case study in the Hudson Bay Lowlands Ecoregion. *International Journal of Remote Sensing*, 39(6), 1615-1627.

Gao, B.-C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3), 389-422.

Halabisky, M., Moskal, L. M., Gillespie, A., & Hannam, M., 2016. Reconstructing semi-arid wetland surface water dynamics through spectral mixture analysis of a time series of Landsat satellite images (1984–2011). *Remote Sensing of Environment*, 177, 171-183.

Ji, L., Zhang, L., Wylie, B., 2009. Analysis of dynamic thresholds for the normalized difference water index. *Photogrammetric Engineering & Remote Sensing*, 75(11), 1307–1317.

Lacaux, J.P., Tourre, Y.M., Vignolles, C., Ndione, J.A., Lafaye, M., 2007. Classification of ponds from high-spatial resolution remote sensing: application to Rift valley fever epidemics in Senegal. *Remote Sensing of Environment*, 106, 66–74.

Ma, S., Zhou, Y., Gowda, P. H., Dong, J., Zhang, G., Kakani, V. G., ... & Jiang, W., 2019. Application of the water-related spectral reflectance indices: A review. *Ecological Indicators*, 98, 68-79.

Maxwell, A. E., Warner, T. A., & Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784-2817.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17, 1425–1432.

Niemuth, N. D., Wangler, B., & Reynolds, R. E., 2010. Spatial and temporal variation in wet area of wetlands in the Prairie Pothole Region of North Dakota and South Dakota. *Wetlands*, 30(6), 1053-1064.

Scholkopf, B., Sung, K. K., Burges, C. J., Giroso, F., Niyogi, P., Poggio, T., & Vapnik, V., 1997. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. *IEEE transactions on Signal Processing*, 45(11), 2758-2765.

Shen, L.; Li, C., 2010. Water Body Extraction from Landsat ETM Imagery using Adaboost Algorithm. *In Proceedings of the 18th International Conference on Geoinformatics*, Beijing, China, 1–4.

URL – 1: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>

Vapnik, V., Golowich, S. E., & Smola, A. J., 1997. Support vector method for function approximation, regression estimation and signal processing. *In Advances in neural information processing systems*, 281-287.

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

Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., ... & Qin, Y., 2017. Open surface water mapping algorithms: a comparison of water-related spectral indices and sensors. *Water*, 9(4), 256.