DETERMINATION OF SEA SURFACE MUCILAGE FORMATIONS USING MULTITEMPORAL SENTINEL-2 IMAGERY

Ismail Colkesen[1], Muhammed Yusuf Ozturk[1], Taskin Kavzoglu[1] and Umut Gunes Sefercik[1]

[1]Department of Geomatics Engineering, Technical University, Gebze, Kocaeli, Turkey

Email: icolkesen@gtu.edu.tr; m.ozturk2020@gtu.edu.tr; kavzoglu@gtu.edu.tr; sefercik@gtu.edu.tr

KEY WORDS: Mucilage, Sentinel-2A, Random Forest, XGBoost, Marmara Sea

ABSTRACT: In recent years, there has been a markedly increase in temperatures of the Earth's surface including oceans and seas due to global warming. Mucilage, or sea saliva observed in the inland seas and bays is one of the destructive results of this phenomenon. Mucilage formations have been widely observed in May 2021 in the Marmara Sea, especially in the Izmit bay. The main purpose of this study is to determine mucilage formations observed in the Izmit bay using Sentinel-2A (S2A) imagery captured on May 19 and 24. For this purpose, three datasets were formed and utilized for classification of mucilage formations. While the first dataset is consisted of only the 10m spatial resolution bands of the S2A imagery, the second dataset is included 6 bands of the S2A imagery at 20m resolution. The third dataset includes 10 bands of the S2A imagery resampled at 10m spatial resolution. Random forest (RF) and XGBoost ensemble learning algorithms were utilized to classification of these three datasets created separately for the 19 and 24 May imagery. To assess the classification performances, overall accuracy and Kappa coefficient measures calculated from confusion matrix were utilized. Results showed that the highest overall classification accuracies were estimated as 98.9% and 98.5% with RF and XGBoost, respectively for 19 May 10-band S2A dataset. On the other hand, the lowest overall accuracies were obtained with the use of 4-bands at 10m spatial resolution for both 19 and 24 May S2A imagery with RF and XGBoost algorithms. This result confirms that the use of the 10m and 20m bands of S2A images together can have a positive effect on the classification accuracy. Whilst the mucilage covered area on Izmit bay was calculated as 15.3 km² on May 19, it was calculated as 20.13 km² on May 24. In other words, it was observed that mucilage formations increased by about 31% in the 5-day time period in the study area. This clearly indicating the harmful effect of the mucilage phenomenon on the Izmit bay.

1. INTRODUCTION

Marine mucilage, also named as sea snot, marine snow, or sea saliva, is defined as a collection of mass of microorganisms or mucus-like organic matter observed in the sea (Precali et al., 2005). Mucilage events frequently observed after sudden changes of the physical and biogeochemical conditions in the marine environment. For instance, significant decreases in the number of phytoplankton species and diversity index just before the mucilage event were observed (Tas et al., 2020). In recent years, many studies have been conducted to investigate the effects of climate change and related global warming and eutrophication in marine environments, especially on mucilage events (Laffoley and Baxter, 2016). The mucilage phenomenon has been experienced frequently in the Adriatic and Tyrrhenian Seas for over 200 years. Apart from the Adriatic and Tyrrhenian Seas (Rinaldi et al., 1995; Innamorati et al., 2001; Mecozzi et al., 2001), mucilage events have been observed in various geographies of the world, such as the Eastern Mediterranean (Gotsis-Skretas 1995; Scoullos et al., 2006; Nikolaidis et al., 2008), North Sea (Lancelot, 1995), New Zealand (MacKenzie et al., 2002) and East China Sea (Fukao et al., 2009).

The first massive mucilage phenomenon in Turkish seas was observed in the Sea of Marmara between October 2007 and February 2008 and affected Straits of Bosphorus and Dardanelles (Aktan et al., 2008; Tüfekçi et al., 2010; İşinibilir Okyar et al., 2015). Before this case, another event was observed in 1992 in the western part of the sea, around Erdek Bay, and was recorded by underwater cameras only by athletes (Tüfekçi et al., 2010). Mucilage, which was not observed in the Marmara Sea for a long time after the 2008 event, was reported again in the deep waters of the Dardanelles during regular research on coral monitoring in December 2020 (Özalp, 2021). Finally, in March 2021, a more severe case of mucilage was observed in the Eceabat region of Dardanelles (Özalp, 2021).

The main goal of this study is to determine spatial allocations of mucilage formations observed in the Izmit bay using Sentinel-2A (S2A) imagery captured on May 19 and 24. In order to evaluate the effects of spectral bands of S2A on the classification of surface mucilage formations, three datasets including 4 bands of S2A at 10m spatial resolution (Dataset-1), 6 bands of S2A at 20m spatial resolution (Dataset-2) and whole 10 bands of the S2A imagery resampled at 10m spatial resolution (Dataset-3) were formed and classified using RF and XGBoost ensemble learning algorithms.

2. STUDY AREA AND DATASET

The main aim of this study is to monitor and detect mucilage formations from multi-temporal satellite images using various spectral band configurations via pixel-based image classification. For this purpose, Izmit bay covering the eastern part of Marmara Sea and intense mucilage formation observed due to lack of oxygen level, was chosen as the study area. The study area is located in the southeastern part of Istanbul and northern part of Yalova cities. In addition, the longest shoreline part of study area is belonged to boundaries of Kocaeli province as shown in Figure 1 with red rectangle. It covers approximately 300 km^2 sea surface region and its shoreline perimeter correspond to 260 km. The region having ports that contribute to national and international maritime trade, makes significant contributions to the Turkey's economy in terms of seafood, especially fisheries. Moreover, recreation areas and beaches located in the coastal areas of the study area are important attraction centers for the tourism sector.

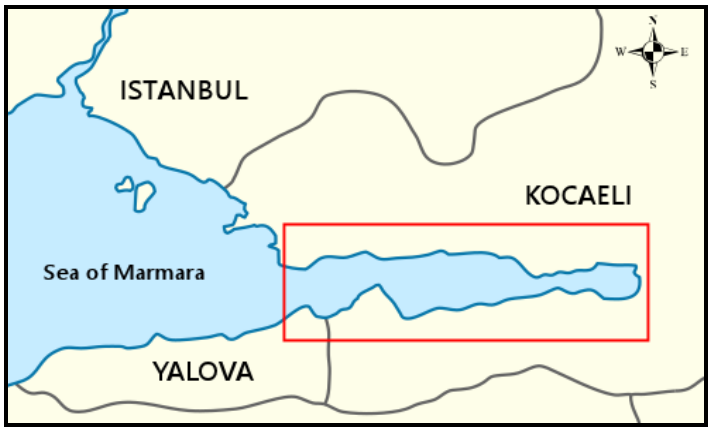

Figure 1. Location of the study area

The atmospherically corrected and georeferenced level 2A Sentinel-2 (S2A) images served by ESA were used as main data product for monitoring and detection of mucilage masses appeared on the study area. Sentinel-2A imagery provides 4-bands at 10m spatial resolution, 6-bands at 20m spatial resolution and 3-bands at 60m spatial resolution, in total 13 spectral bands. The 3-bands at 60m resolution were excluded because they are utilized for atmospheric correction and scanning of cirrus clouds in the literature. In addition to the existing spatial and spectral features, the presentation of images to users in 5-day periods at the point of temporal resolution has revealed the use of Sentinel-2A images as a main data source in many studies (Kavzoglu et al., 2021). In order to accurately determine mucilage formations available in the study area and to obtain reliable results, the cloud-free images were investigated. In this context, two imagery acquired on 19 and 24 May 2021 by Sentinel-2A were used as main data source. In addition, a level-2 Sentinel-2 image captured on 16 October 2020 was utilized to draw the shoreline boundaries of study area.

3. METODOLOGY

The main goal of this study is to detect of mucilage formations observed on Izmit bay and to monitor these formations in 5-day period using different band combinations of multi-temporal S2A images. In this context, various band combinations were composed for image classification process and images were classified by robust ensemble methods, RF and XGBoost. The workflow followed in this study was explained in detail in the following sections.

3.1. Data Acquisition and Pre-Processing

The mucilage formations were first observed in the eastern part of the Marmara Sea in the beginning of April, their maximum concentration levels were monitored between middle of May and June. On the other hand, satellite images taken within the specified time period are exposed to the cloud effect due to the spring season in the study area. In this context, S2A images, captured over study area with minimum cloud effect, were scanned from ESA's web page to obtain reliable results about mucilage formations observed on the Izmit bay. The cloud-free S2A products acquired on 19 and 24 May were selected and utilized in this study. In addition, S2A data, non-cloud with clean water surface and taken on 16 October 2020, was utilized to extract the shoreline boundaries of the study area.

Before image classification process, three bands at 60m spatial resolution were excluded due to their limited spatial

resolution and used for cloud masking. In order to evaluate the effects of spectral bands of S2A on classification of mucilage formations, the remaining 10-bands were divided into three group as 4-bands at 10m resolution (Dataset-1), 6-bands at 20m resolution (Datase-2) and whole bands resampled at 10m (Dataset-3). Based on the spectral properties of water surface that are showing limited reflectance in the SWIR region, the shoreline of study area was extracted using histogram distribution of SWIR-1 band of October 2020 dataset. The shoreline created as vector was overlapped with multi-spectral images and observed errors occurred during shoreline extraction were manually corrected. Shoreline covers about 300 km^2 water surface area and coastline with 260 km perimeter. In the last step of the image processing, sea areas were masked on the raw images dated 19 and 24 May 2021 with the help of the October 2020 dated shoreline.

3.2. Image Classification

Extraction of LULC information from remotely sensed imagery, which results in a thematic map of study area, has traditionally been carried out through pixel-based image classification (Kavzoglu, 2017). The pixel-based image classification technique was employed on the S2A datasets obtained from image processing for both dates to monitor and to detect mucilage formations. For this purpose, the sample pixels representing mucilage formations, sea and ships were collected on the S2A images for each date. The 70% of sample pixels were labelled as training data to construct classification models and the remaining %30 were used as test dataset. Overall accuracy (OA) and kappa coefficient metrics derived from confusion matrix were used to compare performance of classifiers in the detecting mucilage formations.

Two popular ensemble methods namely random forest (RF) and extreme gradient boosting (XGBoost) were employed for the classification purpose. The RF is one of the commonly used tree-based algorithm on the classification and regression problems due to its robustness, efficiency and speed (Immitzer et al., 2012; Fu et al., 2017; Sheykhmousa et al., 2020). RF utilizes multiple decision trees to construct a classification model known as ensemble model to make a prediction of a unknown sample. (Breiman, 2001). The RF classifier has lately become popular in the remote sensing community, particularly for the classification of remotely sensed imagery (Kavzoglu et al., 2018; Tonbul and Kavzoglu, 2020). The RF is a classifier that includes a large number of decision tree classifiers (Kavzoglu, 2017). Before training of each tree in the decision forest, original dataset was divided into subsamples. The 2/3 of subsamples are used for the construction of a tree and the remaining 1/3 of the subsamples are used to test the accuracy of the classification model. The majority voting rule is applied to predict the class label of the unknown sample (Sonobe et al., 2018). In recent years, various ensemble learning algorithms were developed to obtain satisfactory, reliable and fast results. XGBoost, one of the advanced decision tree-based gradient boosting algorithm, was developed by Chen and Guestrin (2016). The main principle of this model is that iteratively construct decision trees in the forest by modifying labels of misclassified samples in each iteration by giving them more weight. XGBoost classifier uses loss function and regularization parameter values to accurate the most classification model and to limit overfitting problems, respectively (Hamedianfar et al., 2020).

4. RESULTS

In this study, detection and statistical analysis of mucilage formations observed on the Izmit bay were carried out utilizing S2A imagery. For this purpose, the S2A images captured on the 19 and 24 May were used as main dataset for pixel-based image classification process. To assess the effect of different band combinations on the separation of mucilage and water surface, S2A bands were divided into three datasets. The first dataset includes original 10m bands of S2A (Dataset-1), the second dataset consists of original 20m bands of S2A (Dataset-2) and the last dataset was formed using 10 spectral bands resampled at 10m spatial resolution (Dataset-3). To classify each dataset, RF and XGBoost ensemble learning algorithms were used. To calculate the OA and Kappa coefficient values of each data set for different dates, constructed classification models were applied to the test dataset. It should be noted that all classification process was conducted with open-source R software. Classification results of each dataset captured on 19 and 24 May were given in Table 1.

It was clear that the highest classification accuracies were estimated as 98.9% (Kappa coefficient value of 0.97) by RF and 98.5% (Kappa coefficient value of 0.97) by XGBoost with the use of 10-band dataset (i.e., Dataset-3). On the other hand, the lowest OA values were calculated with the use of Dataset-1 including only 4-bands at 10m spatial resolution for both classifiers. In addition, the OA values obtained with the use of Dataset-2 having 6-bands and using a 10-bands data set were quite similar, while the estimated OA values were 9% higher compared to using a 4-band data set. In other words, it was observed that the classification accuracies produced from both ensemble models increased with the use of 20m spectral bands of S2A. In addition, XGBoost showed superior classification performance than RF in the limited spectral information available for the 24 May dataset. Furthermore, the mucilage

formations were detected with about 97.5% OA value for 19 May and about 96% OA value for 24 May images using 6-band at 20m spatial resolution by two classifiers. These results showed that spectral ranges of 6-bands with 20m spatial resolution were more effective in separating mucilage from the water surface compared to spectral range of 4-bands at 10m resolution. In addition, RF showed better classification performance in the more spectral bands available.

Table 1. Classification results of each dataset for 19 and 24 May. Note: Calculated highest classification accuracy and corresponding metric about mucilage covered area are represented in bold.

		Dataset-1		Dataset-2		Dataset-3	
Date	**Classifiers**	**RF**	**XGBoost**	**RF**	**XGBoost**	**RF**	**XGBoost**
19.05.2021	**OA (%)**	89.5	89.8	98.1	97.3	**98.9**	**98.5**
	Kappa	0.79	0.80	0.97	0.96	0.97	0.97
	Area (km^2)	21.17	21.34	17.70	19.94	**15.13**	15.03
24.05.2021	**OA (%)**	89.8	91.7	96.1	95.8	**98.1**	**97.1**
	Kappa	0.79	0.81	0.92	0.92	0.95	0.94
	Area (km^2)	24.79	21.56	21.96	19.63	**20.13**	17.41

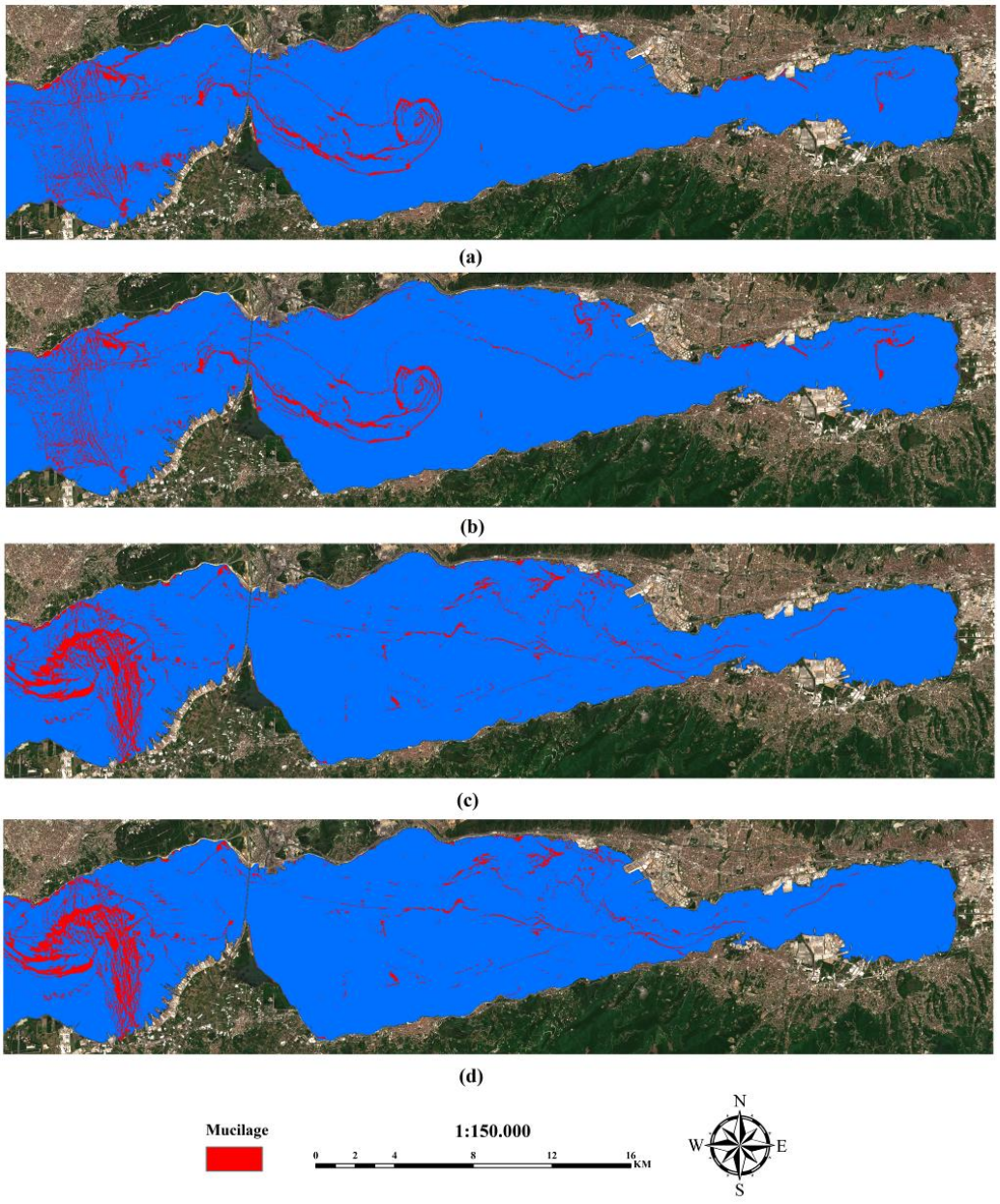

Figure 2. Distribution of mucilage formation on the Izmit bay produced by (a) RF for 19 May; (b) XGBoost for 19 May; (c) RF for 24 May; (d) XGBoost for 24 May

As a result of the analysis conducted using all datasets of both dates, the metric information of the areas covered by mucilage formations on the sea surface were calculated and given in Table 1. There is an inversely relationship between the mucilage areas detected on the sea surface and the classification accuracies. The main reason may be related to misclassification of mucilage areas in images with limited spectral information (i.e., 4-band at 10m resolution dataset). Thus, the mucilage areas determined with the use of Dataset-3 were considered as the the most accurate result. On 19 May, 15.13 km^2 and 15.03 km^2 mucilage formations were detected in the study area by RF and XGBoost, respectively. These mucilage formations correspond to approximately 5% of the whole study area. On the other hand, the classification results of RF and XGBoost showed that mucilage formations on the Izmit bay covered 20.13 km^2 and 17.41 km^2, respectively on 24 May. These results showed that mucilage formations increased by about 31% in the 5-day time period in the study area due to harmful effect of the mucilage phenomenon.

To visually analyze the spatial locations of the mucilage formations on the study area, the thematic maps produced with the use of 10-band dataset by RF and XGBoost were presented in Figure 2. It was found that RF outperformed the XGBoost algorithm in detecting low-density mucilage formations. As can be seen from the figure, mucilage formations were more observed on the coasts, middle and west of the study area on 19 May. On the other hand, it was observed that mucilage formations were less monitored in the inner parts of the Izmit bay (i.e., part of east). Only 5 days later, on May 24, mucilage formations were concentrated in the west of the gulf of Izmit.

REFERENCES

Aktan, Y., Dede, A., Ciftci, P.S., 2008. Mucilage event associated with diatoms and dinoflagellates in Sea of Marmara, Turkey. Harmful Algae News 36, pp. 1-3.

Breiman, L., 2001. Random Forests. In: Machine Learning, Chapman and Hall/CRC, pp. 5–32.

Chen, T., Guestrin, C., 2016. XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17 August, pp. 785–794.

Fu, B., Wang, Y., Campbell, A., Li, Y., Zhang, B., Yin, S., Xing, Z., Jin, X., 2017. Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. Ecological Indicators, 73, pp. 105–117.

Fukao, T., Kimoto, K., Yamatogi, T., Yamamoto, K., Yoshida, Y., Kotani, Y., 2009. Marine mucilage in Ariake Sound, Japan, is composed of transparent exopolymer particles produced by the diatom Coscinodiscus granii. Fisheries Science, 75, pp. 1007–1014.

Gotsis-Skretas, O., 1995. Mucilage appearances in Greek waters during 1982-1994. The Science of the Total Environment. 165, pp. 229–230.

Hamedianfar, A., Gibril, M. B. A., Hosseinpoor, M., Pellikka, P. K. E., 2020. Synergistic use of particle swarm optimization, artificial neural network, and extreme gradient boosting algorithms for urban LULC mapping from WorldView-3 images. Geocarto International, pp. 1–19.

Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. Remote Sensing, 4(9), pp. 2661–2693.

Innamorati, M., Nuccio, C., Massi, L., Mori, G., Melley, A., 2001. Mucilages and climatic changes in the Tyrrhenian Sea. Aquatic Conservation, Marine and Freshwater Ecosystems 11, pp. 289-298.

Isinibilir Okyar, M., Üstün, F., Orun, D.A., 2015. Changes in abundance and community structure of the zooplankton population during the 2008 mucilage event in the northeastern Marmara Sea. Turkish Journal of Zoology 39, pp. 28-38.

Kavzoglu, T., 2017. Object-Oriented Random Forest for High Resolution Land Cover Mapping Using Quickbird-2 Imagery. In: Handbook of Neural Computation, edited by P. Samui, S. S. Roy, and V. E. Balas. Amsterdam: Elsevier. Chapter 33, pp. 607-619.

Kavzoglu, T., Tonbul, H., Yildiz Erdemir, M., Colkesen, I., 2018. Dimensionality Reduction and Classification of Hyperspectral Images Using Object-Based Image Analysis. Journal of Indian Society of Remote Sensing, 46, pp. 1297–1306.

Kavzoğlu, T., Çölkesen, İ., Sefercik, U. G., Öztürk, M. Y., 2021. Detection and Analysis of Marine Mucilage Bloom in the Sea of Marmara by a Machine Learning Algorithm from Multi-Temporal Optical and Thermal Satellite Images. Harita Dergisi (Map Journal), 166(2), pp. 1-9.

Laffoley, D., Baxter, J.M., 2016. Explaining Ocean Warming: Causes, Scale, Effects and Consequences. Full Report. Gland, Switzerland: IUCN, 456 pp.

Lancelot, C., 1995. The mucilage phenomenon in the continental coastal waters of the North Sea. The Science of the Total Environment, 165, pp. 83–102.

MacKenzie, L., Sims, I., Beuzenberg, V. Gillespie, P., 2002. Mass accumulation of mucilage caused by dinoflagellate polysaccharide exudates in Tasman Bay, New Zealand. Harmful Algae, 1(1), pp. 69-83.

Mecozzi, M., Acquistucci, R., Noto, V., Pietrantonio, E., Amici, M. Cardarilli, D., 2001. Characterization of mucilage aggregates in Adriatic and Tyrrhenian Sea: Structure similarities between mucilage samples and the insoluble fractions of marine humic substance. Chemosphere, 44, pp. 709-20.

Nikolaidis, G., Aligizaki, K., Koukaras, K., Moschandreou, K., 2008. Mucilage phenomena in North Aegean Sea, Greece: another harmful effect of dinoflagellates. Proceedings of the 12th International Conference on Harmful Algae. 4–8 September 2006, pp. 219–222

Ozalp, H.B. 2021. First massive mucilage event observed in deep waters of Çanakkale Strait (Dardanelles), Turkey. Journal of the Black Sea/Mediterranean Environment, 27(1), pp. 49-66

Precali, R., Giani, M., Marini, M., Grilli, F., Ferrari, C. R., Pečar, O., Paschini, E., 2005. Mucilaginous aggregates in the northern Adriatic in the period 1999–2002: Typology and distribution, Science of the Total Environment, 353(1–3), pp. 10–23.

Rinaldi, A., Vollenweider, R. A., Montanari, G., Ferrari, C. R. & Ghetti, A., 1995. Mucilages in Italian seas: the Adriatic and Tyrrhenian Seas, 1988–1991, Science of the Total Environment, 165(1–3), pp. 165–183.

Scoullos, M., Mata Plavšić, M., Karavoltsos, S., Sakellari, A., 2006. Partitioning and distribution of dissolved copper, cadmium and organic matter in Mediterranean marine coastal areas: The case of a mucilage event, Estuarine, Coastal and Shelf Science, 67(3), pp. 484-490.

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S., 2020. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, pp. 6308–6325.

Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K., 2018. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. Journal of Applied Remote Sensing, 12, 026019.

Tas, S., Kus, D., Yilmaz, I. N., 2020. Temporal variations in phytoplankton composition in the north-eastern Sea of Marmara: potentially toxic species and mucilage event. Mediterranean Marine Science, 21(3), pp. 668-683.

Tonbul, H., Kavzoglu, T., 2020. Semi-automatic building extraction from worldview-2 imagery using Taguchi optimization. Photogrammetric Engineering & Remote Sensing, 86(9), pp. 547-555.

Tufekci, V., Balkıs, N., Beken Polat, Ç., Ediger, D., Mantıkcı, M., 2010. Phytoplankton composition and environmental conditions of a mucilage event in the Sea of Marmara. Turkish Journal of Biology, 34, pp. 199-210.