ML for Mapping Ocean Colours

Sarah Parsons

Dr. Meisam Amani

Armin Moghimi

Authors Parsons, Amani, and Moghimi use satellite images to map ocean colour.

Who should read this paper?

This research study is of benefit to fishers and researchers interested in theory and applications of ocean colour remote sensing, the marine ecosystem and pollution, water quality and eutrophication, and inland and coastal water quality. Fishers can use ocean colour data to identify potential fishing zones. Managers and researchers can use remote sensing derived ocean colour data for monitoring the negative impacts of harmful algal blooms and aquaculture. Moreover, such data is very efficient for ecosystem modelling.

Why is it important?

Numerous machine learning (ML)-based methods have been proposed to estimate the ocean colour from remote sensing data that have yielded promising results. However, most of them were non-automatic or semi-automatic approaches and required in-situ samples. The framework presented here is designed to automatically estimate the level of ocean colour by applying an unsupervised ML classification method on a remote sensing image. The ocean colour map generated by the proposed method can be used as an input data for navigation and saving fuel in commercial shipping. The results can also be utilized in fish finding analyses to predict places where the possibility of fish is high. Additionally, the results can be helpful in coastal monitoring and measuring water quality.

About the authors

Sarah Parsons is enrolled in her third year of the Ocean Mapping program at the Fisheries and Marine Institute of Memorial University. **Dr. Meisam Amani** is currently a senior remote sensing engineer and the key specialty leader of data analytics at Wood PLC, a global consulting and engineering company, where he manages and leads various industrial, governmental, and academic remote sensing projects worldwide. Over the past 11 years, he has worked on different applications of remote sensing, including but not limited to land cover/land use classification, soil moisture estimation, drought monitoring, water quality assessment, watershed management, power/transmission line monitoring, fog detection and nowcasting, and ocean wind estimation. To do these, Dr. Amani has utilized various remote sensing datasets (e.g., UAV, optical, LiDAR, SAR, scatterometer, radiometer, and altimeter) along with different machine learning and big data processing algorithms. A list of his research works, including over 50 peer-reviewed journal and conference papers, can be found at www.researchgate.net/profile/Meisam_Amani3. **Armin Moghimi** received his B.Sc. in geomatics engineering from the Shahid Beheshti University (formerly Geomatics College of National Cartographic Center), Tehran, Iran, in 2013, and a M.Sc. degree in photogrammetry engineering in 2015 from the K.N. Toosi University of Technology, Tehran, Iran, where he is currently working toward his PhD in photogrammetry and remote sensing. His research interests include change detection techniques, image preprocessing, image registration, and machine learning.

OCEAN COLOUR MAPPING USING REMOTE SENSING TECHNOLOGY AND AN UNSUPERVISED MACHINE LEARNING ALGORITHM

Sarah Parsons[1], Meisam Amani[2], and Armin Moghimi[3]
[1]Ocean Mapping Program, Fisheries and Marine Institute of Memorial University of Newfoundland, St. John's, N.L., Canada
[2]Wood Environment and Infrastructure Solutions, Ottawa, O.N., Canada
[3]Department of Photogrammetry and Remote Sensing, Geomatics Engineering Faculty, K.N. Toosi University of Technology, Tehran, Iran

ABSTRACT

Satellites allow users to observe ocean colour in a way that is not possible from a ship or the shore. Ocean colour depends on interactions of incident light with particles or substances in the water. These light interactions cause the ocean to be a variety of shades depending on what the water is composed of and how these materials change the reflections of the light. The ocean colour fluctuation can be caused by different compositions, such as the biomass of phytoplankton or zooplankton, and can lead to a change in ocean colour, for example, from normal clear blue into a variety of shades of green. Satellites take measurements that can be used to calculate ocean colour and concentrations of materials in the ocean. This study focused on ocean colour mapping using satellite images captured from the Mediterranean Sea. The Iterative Self Organizing Data Analysis Techniques Algorithm (ISODATA) unsupervised machine learning (ML) algorithm was employed to determine ocean colour. The produced map is a basic way of displaying ocean colour and is easy for users of any skill level to produce. Finally, it was observed that having more information about phytoplankton and applying it to the algorithm could improve the results.

KEYWORDS

Ocean colour; Remote sensing; Phytoplankton; ISODATA; Machine learning

INTRODUCTION

The ocean gets its colour from interactions of incident light with substances or particles that are in the water [Aiken, 1994; Blondeau-Patissier, 2014; Groom et al., 2019]. Among these substances, phytoplanktons have a significant effect on the seawater colour changes through the process of its photosynthetic organisms that contain chlorophyll [Blondeau-Patissier, 2014]. This is mainly because green pigment – chlorophyll – transmits light spectrum at green wavelengths while absorbs it at blue and red wavelengths [Blondeau-Patissier, 2014; Groom et al., 2019]. Consequently, the ocean colour will change from blue to green over regions with high phytoplankton densities [Aiken, 1994; Blondeau-Patissier, 2014; Groom et al., 2019]. Depending on the phytoplankton abundance and its species and weather conditions, the level of this discolouration is usually different [Aiken, 1994]. Ocean colour data is typically obtained by measuring the light intensity at different wavelengths [Aiken, 1994; Blondeau-Patissier, 2014; Groom et al., 2019; Abbas et al., 2019]. Such data can then be utilized to estimate material concentrations in surface ocean waters and determine the amount of biological activity [Cullen, 1982; Dore et al., 2008].

Ocean colour can be mainly estimated using direct or indirect measurements. Direct methods use field data derived from water sampling by vessels or a more permanent method, such as a land-based observation site [Joint and Groom, 2000]. These measurements are not ideal for monitoring the ocean colour in large spatio-temporal scales and typically include gaps and patchy observations [Abbas et al., 2019]. Moreover, they cannot collect observations fast enough to provide a global synoptic view of phytoplankton abundance [Joint and Groom, 2000]. However, indirect ocean colour measurements would come from data collected by airborne or spaceborne remote sensing systems [Abbas et al., 2019]. Regarding ocean colour estimation, passive remote sensing systems provide valuable information [Amani et al., 2020] affordably. These satellite systems record the ocean colour by measuring the sunlight reflected from the water's surface in different wavelengths spectrum. Data collected from passive satellites can also be used to determine the concentrations of particles, such as phytoplankton and dissolved materials in shallow ocean water [Abbas et al., 2019].

There are many passive spaceborne remote sensing systems, the data of which can be used to estimate ocean colour [Aiken, 1994; Blondeau-Patissier, 2014; Groom et al., 2019]. In 1978, the National Aeronautics and Space Administration (NASA) launched the Nimbus-7 rocket equipped by the Coastal Zone Colour Scanner for ocean colour measurements [Hovis et al., 1980]. Before this mission, different methods were used to observe ocean colour, but this new mission specifically focused on the parameters that would affect the ocean colour [Hovis et al., 1980]. This satellite had four spectral bands primarily used for ocean colour [Hovis et al., 1980]. The Sea-viewing Wide Field-of-view Sensor was also the first sensor dedicated to global ocean colour observation [Unninayar and Olsen, 2015]. This sensor had eight spectral bands, and its main objective was to determine the bio-optical reflectance characteristics of the upper region

of the ocean and understand the methods used for observed variations [Unninayar and Olsen, 2015]. Additionally, the data derived from NASA's Aqua Moderate Resolution Imaging Spectroradiometer (MODIS), launched in 2002, provide a wide range of essential data in nine spectral bands (in the visible-NIR) for daily ocean colour estimation [Guo et al., 2019]. The data acquired by the Ocean and Land Colour Imager carried by Sentinel 3 satellite (launched in February 2016) with 21 spectral bands from 400 to 1,020 nm is another valuable source for ocean colour determination [Nieke et al., 2012]. The ocean colour can also be achieved from higher-resolution optical satellites, such as Landsat 8 at 30 m and Sentinel 2 at 10-60 m [Blondeau-Patissier et al., 2014; Groom et al., 2019].

So far, many remote sensing and ML algorithms have been developed to reconstruct the ocean colour from remote sensing images. For example, Franz et al. [2014] measured the ocean colour from Landsat 8 images based on the remote sensing reflectance data acquired by SeaDAS. Moreover, Zheng and DiGiacomo [2017] introduced the generalized stacked-constraints model for measuring Chl-a from remote sensing images in the coastal regions. Abbas et al. [2019] also proposed the Ocean Colour 3M algorithm to estimate Chl-a from MODIS data in the coastal area. This method was based on the blue-green spectral band ratios and colour indices which are not suited for cloudy regions. Finally, Park et al. [2019] employed the random forest algorithm and extremely randomized tree method for the reconstruction of the ocean colour data in the Polar Regions using microwave-based remote sensing data.

Although many research studies have been developed for ocean colour estimation from remote sensing data, a relatively low number of studies have utilized automatic ML algorithms to determine ocean colour. Thus, this study aims to automatically determine the ocean colour levels from Landsat 8 satellite images using an unsupervised ML classification algorithm.

STUDY AREA AND DATA

Study Area

This paper focused on an area of the Mediterranean Sea which marked around the coordinates 40° 19' 58.48" N and 25° 34' 29.68" E (Figure 1). The reason behind choosing this area was its clear water in which the plankton or other causes of ocean colour changes can be better distinguished from satellite images. Moreover, visual analysis of satellite images between 2013 to 2020 in Google Earth confirms that the plankton concentrations in the study area were increased from 2018 up to now.

Remote Sensing Data and Preprocessing

A multispectral optical image acquired by Landsat 8 (i.e., downloaded from the United States Geological Survey web site: https://earthexplorer.usgs.gov/) was employed to determine ocean colour of the study area. The detail of the satellite data used in this study is listed in Table 1. Figure 2 also shows the true colour composite of the Landsat-8 image.

Before any processing with optical satellite images, they should be radiometrically corrected with a reliable atmospheric correction. To this end, the FLAASH®

Figure 1: Locations of the study area and time-series satellite images from 2013-2020.

	Sensor	Spectral Band	Wavelength	Resolution	Size	Date
Landsat 8	OLI	Band 1 - Coastal / Aerosol	0.433 – 0.453 µm	30 m	7821x7941	2019/08/30
		Band 2 - Blue	0.450 – 0.515 µm	30 m		
		Band 3 - Green	0.525 – 0.600 µm	30 m		
		Band 4 - Red	0.630 – 0.680 µm	30 m		
		Band 5 - Near Infrared	0.845 – 0.885 µm	30 m		
		Band 6 - Short Wavelength Infrared	1.560 – 1.660 µm	30 m		
		Band 7 - Short Wavelength Infrared	2.100 – 2.300 µm	30 m		
		Band 8 - Panchromatic	0.500 – 0.680 µm	15 m		
		Band 9 - Cirrus	1.360 – 1.390 µm	30 m		
	TIRS	Band 10 - Long Wavelength Infrared	10.30 – 11.30 µm	100 m		
		Band 11 - Long Wavelength Infrared	11.50 – 12.50 µm	100 m		

Table 1: Characteristics of the Landsat 8 image used in this study.

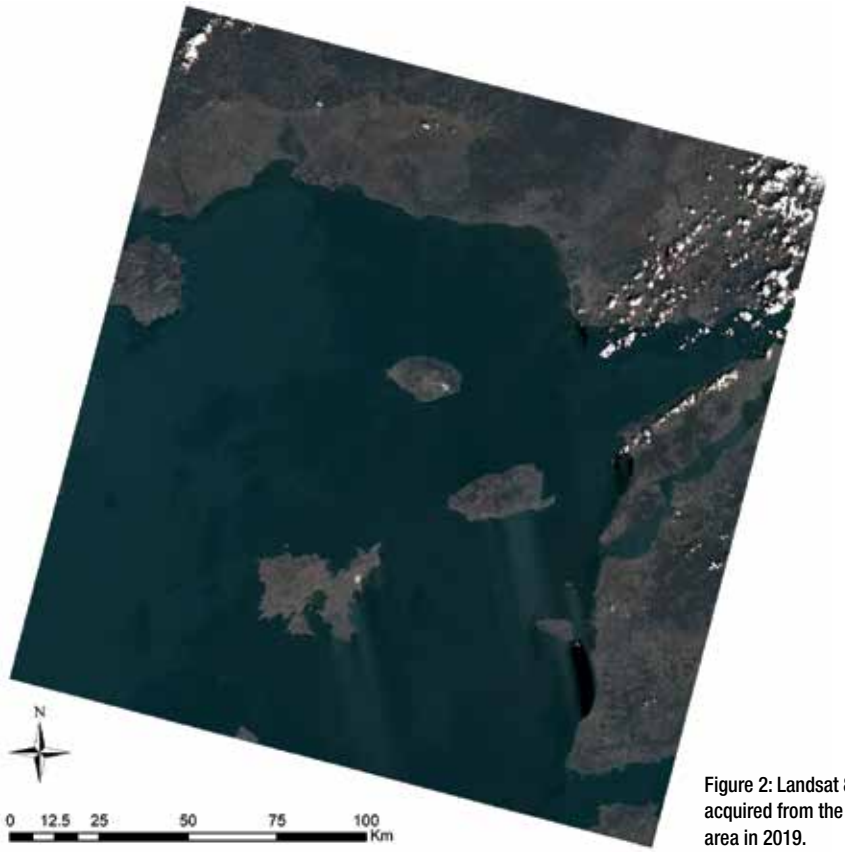


Figure 2: Landsat 8 image acquired from the study area in 2019.

atmospheric correction tool in the ENVI software package was employed to generate surface reflectance data by removing the effects of clouds and aerosols from the radiance image.

METHOD

Along with the main spectral bands of the Landsat 8 image (Table 1), we also used two widely used remote sensing indices of the normalized difference vegetation index (NDVI) [Jensen, 1996] (Equation 1) and normalized difference water index (NDWI) [Gao, 1996] (Equation 2) to improve the ocean colour classification.

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (1)$$

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (2)$$

where ρ_{Red}, ρ_{NIR}, and ρ_{SWIR} are the reflectance values in the red, near infrared (NIR), and Shortwave Infrared (SWIR) spectral bands (i.e., bands 4, 5, and 6), respectively.

Higher plankton levels produce different shades of blue and green in water bodies, which can easily be detected using a ML algorithm. There are two main groups of ML classification algorithms: supervised and unsupervised [Chaovalit and Zhou, 2005; Sathya and Abraham, 2013]. The supervised ML classifications are human-guided algorithms that assign a class label to input data based on a set of training data [Chaovalit and Zhou, 2005]. In contrast, unsupervised ML classifications automatically analyze and learn image pixels from unlabelled data and assign pixels to spectral clusters [Chaovalit and Zhou, 2005; Sathya and Abraham, 2013]. In many remote sensing applications where a high level of automation and low-cost computational expenses are required, unsupervised ML classification algorithms are of interest [Chaovalit and Zhou, 2005; Moghimi et al., 2017].

In this study, we employed the ISODATA [Ball and Hall, 1965] unsupervised algorithm for clustering the study area into different ocean colour levels due to its high efficacy and low computational complexity [Jensen, 1996; Memarsadeghi et al., 2007]. The ISODATA is a developed version of k-means clustering in which splitting and merging of clusters are embedded [Navulur, 2006]. The ISODATA algorithm is also more flexible than the k-means algorithm because it allocates a dynamically different number of clusters. However, the number of clusters is fixed in the k-means clustering algorithm [Navulur, 2006]. ISODATA initially computes the means of evenly distributed clusters as initial centroids within the feature space and then iteratively assigns the remaining pixels to the closest centroid using minimum distance techniques [Jensen, 1996; Memarsadeghi et al., 2007; Ma et al., 2020]. The new centroids are recalculated using the current cluster memberships and pixels re-clustered in each iteration [Jensen, 1996; Ma et al., 2020]. The number of clusters is decreased/increased during each iteration by merging similar clusters or splitting clusters with high standard deviations using actual thresholds [Jensen, 1996; Ma et al., 2020]. This process is terminated when a convergence criterion is

Parameters	Default Value
Range of cluster numbers	[2,20]
Minimum number of pixels per cluster	20
Maximum number of iterations	40
Maximum number of clusters that can be merged at one time	2
Change threshold to end the iterative process	0.05

Table 2: The tuning parameters of the ISODATA algorithm.

met or the maximum number of iterations is reached [Jensen, 1996; Ma et al., 2020].

The results of ISODATA do not represent ocean levels, and labels need to be allocated to them based on the spectral properties. To this end, the water body was first extracted by thresholding of NDWI with the Otsu algorithm [Otsu, 1979]. The mean of the NDVI for each cluster in the detected water body area was then considered as levels of planktons in the study area, where the cluster with the low mean value was assigned to a lower level and vice versa.

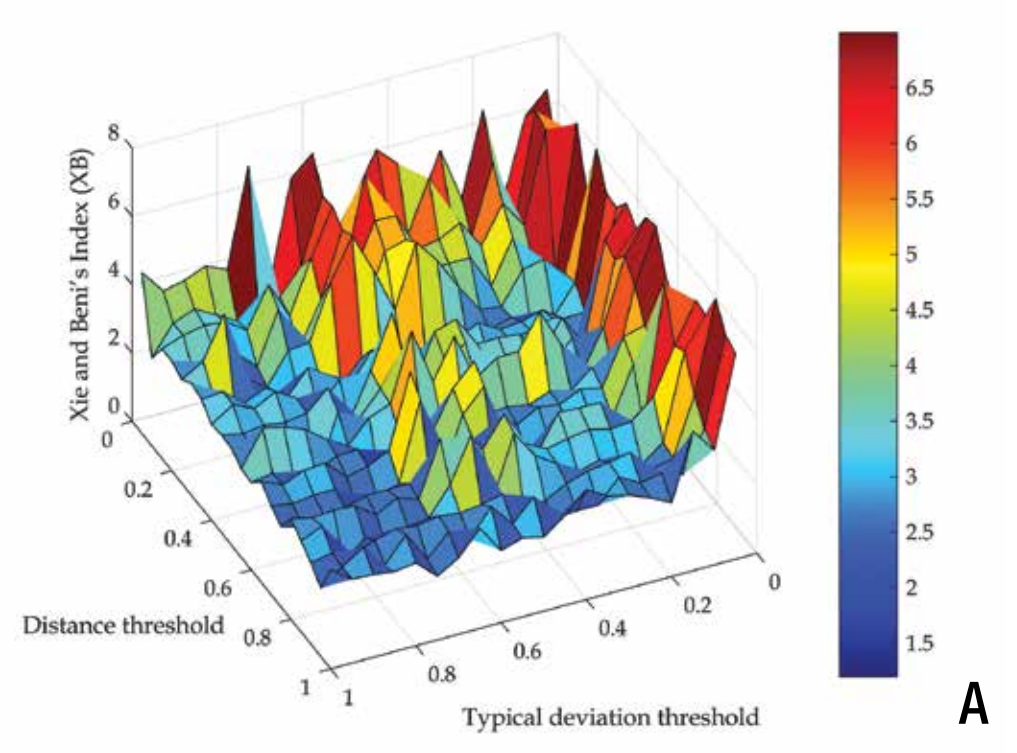
RESULTS AND DISCUSSION

Even though the ISODATA algorithm can regulate the number of clusters, a set of parameters directly affect its results. Among these parameters, the distance and typical deviation thresholds are the most critical parameters because they control the ISODATA convergence by merging clusters and/or dividing clusters. Although default settings of the ISODATA parameters may produce a good performance, we employed the widely used grid search method for hyperparameter tuning. In this way, the Xie and Beni's Index (XB) [Xie and Beni, 1991] was used as a parameter selection measure to search for optimum distance and typical deviation thresholds. The XB can be represented based on the ratio of the compactness and separability of clusters as following [Xie and Beni, 1991]:

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{N}\left\|x_j - v_i\right\|^2}{N\, min_{i,j}\left\|x_j - v_i\right\|^2} \quad (3)$$

where c is the number of clusters, the set of N observations is represented by the $X=\{x_j| j=1,2,...,N\}$, and v_i is the i-th cluster centre. The lower XB value provides better clustering results in terms of the compactness and separability of clusters [Xie and Beni, 1991]. It is worth noting that the other parameters of the ISODATA algorithm were set as defaults (Table 2) when the distance and typical deviation thresholds were obtained by the grid search method. The results from the grid search are shown in Figure 3.

As shown in Figure 3A, the variation of the distance and typical deviation thresholds significantly affected the ISODATA performance, indicating the high sensitivity to these threshold vaues. Therefore, a combination of these thresholds in the optimum space can satisfy the requirements of XB. Finally, the optimal thresholds meet in a minimum space with blue colour in the grid, where the optimal combination of the distance and typical deviation thresholds, respectively, were 0.5 and 0.8. Figure 3B indicated that the credible XB value was attained for 16 clusters



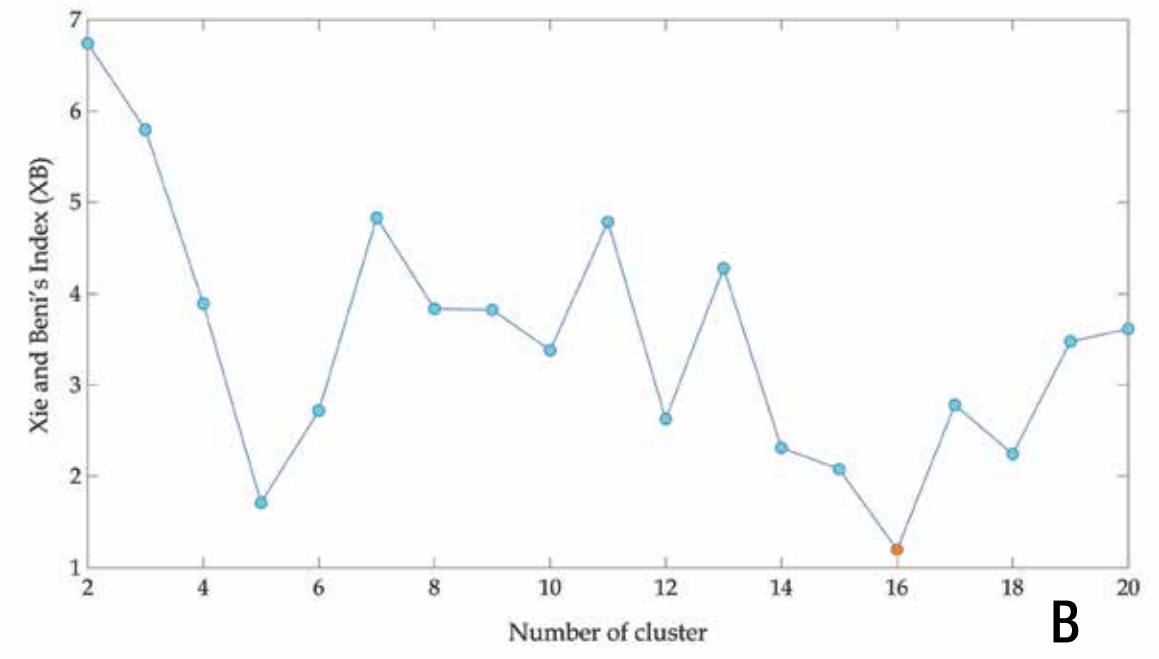


Figure 3: (A) The grid search for the distance and typical deviation thresholds in the ISODATA clustering based on the Xie and Beni's Index (XB), and (B) the final cluster number resulted from the selected optimum thresholds.

within range of [2,20] after the clustering process with optimum values obtained for the desired thresholds.

As discussed, the ISODATA algorithm was applied to the input data (i.e., spectral bands of Landsat 8 as well as the NDVI and NDWI) by assigning the optimum distance and typical deviation thresholds and initial cluster range [Park et al., 2019; Moghimi et al., 2017]. When the unsupervised classification was applied, the map with 16 clusters was initially produced, including different ocean colour levels. To assign labels to each cluster, water bodies and clouds were first separated from land by thresholding the NDWI index. Then, by masking the land using the NDVI index, clouds with positive values in NDVI were separated from water bodies (i.e., negative values of NDVI). Finally, the rest of the clusters in the water body were combined based on the NDVI values and were assigned to five ocean colour levels. The final unsupervised classification is provided in Figure 4.

As shown in Figure 4, five levels of ocean colours were determined for the study area using the ISODATA algorithm. Different ocean colour levels indicate various plankton levels. Level 1 is the areas with a lower concentration of plankton, while level 5 shows the highest concentrations of plankton. In fact, level 1 has the most apparent ocean colour, while level 5 has the deepest green shade. Moreover, lighter green areas were found in level 1, while the deeper green shades were found in level 5.

The ISODATA unsupervised classification algorithm is an easy but effective algorithm to produce ocean colour maps. Moreover, since unsupervised classifications do not require previous knowledge and in-situ data, it significantly reduces cost and human error in producing the results.

With all advantages of unsupervised classification algorithms for ocean colour mapping, there are still multiple disadvantages. For example, unsupervised classification algorithms usually have lower accuracies compared to supervised machine learning methods. However, it should be noted supervised methods need in-situ data, which were not available in this study.

CONCLUSION

The changes in ocean colour can reveal valuable information about the ocean and what is happening below its surface. Remote sensing methods are practical tools for ocean colour estimation. In this study, we presented a simple but efficient method to determine ocean colour levels using the ISODATA unsupervised algorithm. The technique used in this study demonstrates a simple approach to determine the levels of phytoplankton concentrations. Moreover, this is a straightforward approach for users with any skill level to analyze remote sensing data for different oceanographic applications. One of the limitations of this study was the lack of the ground truth. In future studies, more advanced ML algorithms along with suitable number of field data should be employed to determine the amount of ocean colour change and the reasons behind these changes.

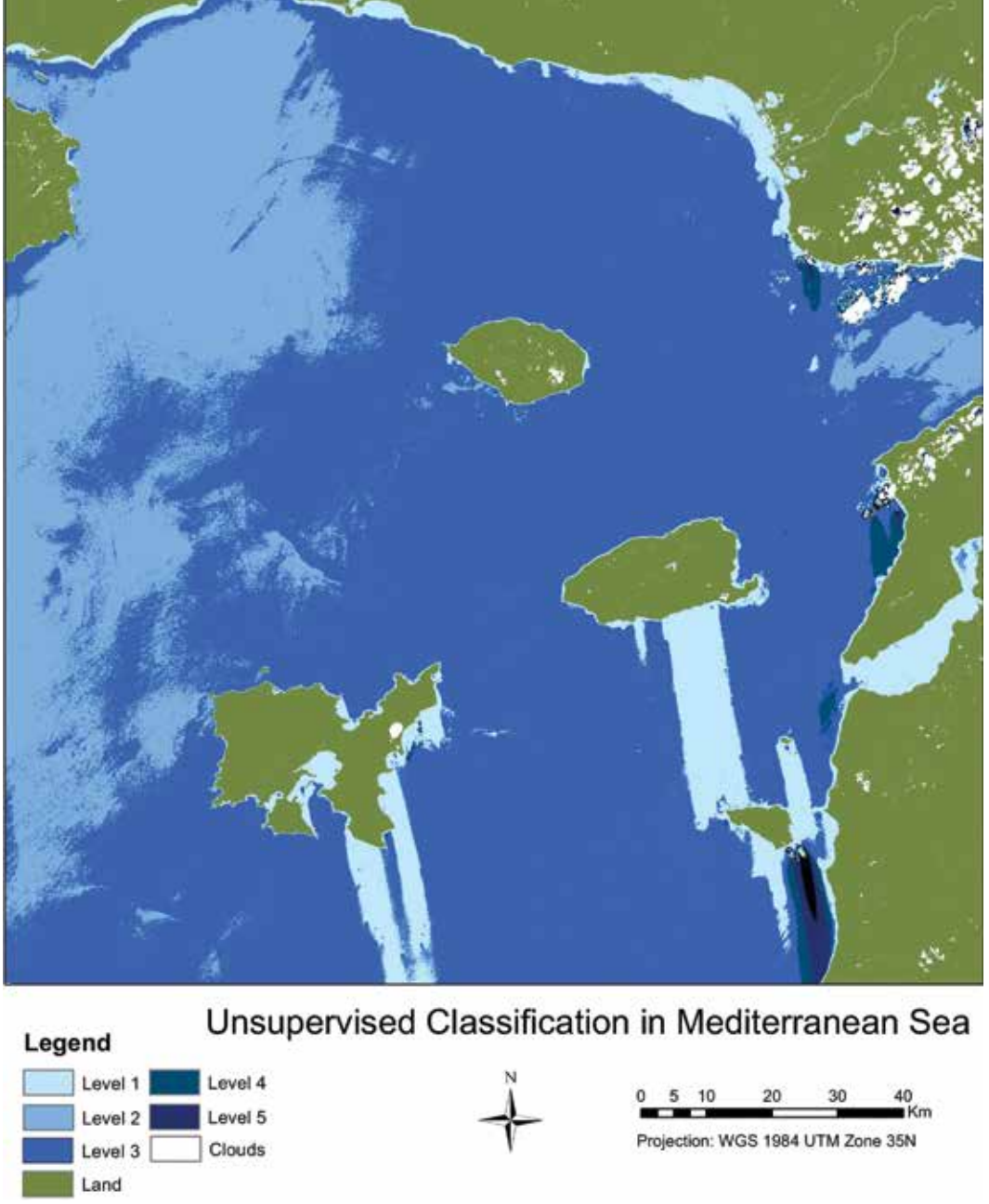


Figure 4: Unsupervised classification of ocean colour in the Mediterranean Sea.

REFERENCES

Abbas, M.M.; Melesse, A.M.; Scinto, L.J.; and Rehage, J.S. [2019]. *Satellite estimation of chlorophyll-a using moderate resolution imaging spectroradiometer (MODIS) sensor in shallow coastal water bodies: validation and improvement*. Water (Switzerland), doi: 10.3390/w11081621.

Aiken, H. [1994]. *Ocean colour: theory and applications in a decade of CZCS experience*. Journal of Experimental Marine Biology and Ecology, doi: 10.1016/0022-0981(94)90205-4.

Amani, M.; Ghorbanian, A.; Ahmadi, S.; Kakooei, M.; Moghimi, A.; Mirmazloumi, S.M.; Moghaddam, S.H.A.; Mahdavi, S.; Ghahremanloo, M.; Parsian, S.; Wu, Q.;

and Brisco, B. [2020]. *Google Earth Engine cloud computing platform for remote sensing big data applications: a comprehensive review*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, doi: 10.1109/JSTARS.2020.3021052.

Ball, G.H. and Hall, D.J. [1965]. *ISODATA, a novel method of data analysis and pattern classification*. Stanford Research Institute, Menlo Park, C.A.

Blondeau-Patissier, D.; Gower, J.F.R.; Dekker, A.G.; Phinn, S.R.; and Brando, V.E. [2014]. *A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans*. Progress in Oceanography, doi: 10.1016/j.pocean.2013.12.008.

Chaovalit, P. and Zhou, L. [2005]. *Movie review mining: a comparison between supervised and unsupervised classification approaches*. Proceedings: 38th annual Hawaii international conference on system sciences, pp. 112c-112c.

Cullen, J.J. [1982]. *The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a*. Canadian Journal of Fisheries and Aquatic Sciences, doi: 10.1139/f82-108.

Dore, J.E.; Letelier, R.M.; Church, M.J.; Lukas, R.; and Karl, D.M. [2008]. *Summer phytoplankton blooms in the oligotrophic North Pacific Subtropical Gyre: historical perspective and recent observations*. Progress in Oceanography, doi: 10.1016/j.pocean.2007.10.002.

Franz, B.A.; Bailey, S.W.; Kuring, N.; and Werdell, P.J. [2014]. *Ocean color measurements from Landsat-8 OLI using SeaDAS*. Proceedings: Ocean Optics XX11, DOI:10.13140/2.3686.4965.

Gao, B.C. [1996]. NDWI - *a normalized difference water index for remote sensing of vegetation liquid water from space*. Remote Sensing of Environment, doi: 10.1016/S0034-4257(96)00067-3.

Groom, S.B.; Sathyendranath, S.; Ban, Y.; Bernard, S.; Brewin, R.; Brotas, V.; Brockmann, C.; Chauhan, P.; Choi, J-K.; Chuprin, A.; Ciavatta, S.; Cipollini, P.; Donlon, C.; Franz, B.; He, X.; Hirata, T.; Jackson, T.; Kampel, M.; Krasemann, H.; Lavender, S.; Pardo-Martinez, S.; Mélin, F.; Platt, T.; Santoleri, R.; Skakala, J.; Schaeffer, B.; Smith, M.; Steinmetz, F.; Valente, A.; and Wang, M. [2019]. *Satellite ocean colour: current status and future perspective*. Frontiers in Marine Science, doi: 10.3389/fmars.2019.00485.

Guo, H.; Bao, A.; Liu, T.; Ndayisaba, F.; Jianag, L.; Zheng, G.; Chen, T.; and DeMaeyer, P. [2019]. *Determining variable weights for an Optimal Scaled Drought Condition Index (OSDCI): evaluation in Central Asia*. Remote Sensing of Environment, Vol. 231, p. 111220, doi: https://doi.org/10.1016/j.rse.2019.111220.

Hovis, W.A.; Clark, D.K.; Anderson, F.; Austin, R.W.; Wilson, W.H.; Baker, E.T.; Ball, D.; Gordon, H.R.; Mueller, J.L. El-Sayed, S.Z.; Sturm, B.; Wrigley, R.C.; and Yentsch, C.S. [1980]. *Nimbus-7 coastal zone color scanner: system description and initial imagery*. Science, Vol. 210, Issue, 4465. pp. 60-63, doi: 10.1126/ science. 210.4465.60.

Jensen, J.R. [1996]. *Introductory digital image processing: a remote sensing perspective*. Second edition.

Joint, I. and Groom, S.B. [2000]. *Estimation of phytoplankton production from space:*

current status and future potential of satellite remote sensing. Journal of Experimental Marine Biology and Ecology, doi: 10.1016/S0022-0981(00)00199-4.

Ma, Z.; Liu, Z.; Zhao, Y.; Zhang, L.; Liu, D.; Ren, T.; Zhang, X.; and Li. S. [2020]. *An unsupervised crop classification method based on principal components isometric binning*. ISPRS International Journal of Geo-Information, doi: 10.3390/ijgi9110648.

Memarsadeghi, N.; Mount, D.M.; Netanyahu, N.S.; and Le Moigne, J. [2007]. *A fast implementation of the isodata clustering algorithm*. International Journal of Computation Geometry and Applications, Vol. 17, No. 1, pp. 71-103, doi: 10.1142/S0218195907002252.

Moghimi, A.; Khazai, S.; and Mohammadzadeh, A. [2017]. *An improved fast level set method initialized with a combination of k-means clustering and Otsu thresholding for unsupervised change detection from SAR images*. Arabian Journal of Geosciences, Vol. 10, No. 13, pp. 1-18, doi: 10.1007/s12517-017-3072-3.

Navulur, K. [2006]. *Multispectral image analysis using the object-oriented paradigm*. CRC press.

Nieke, J.; Borde, F.; Mavrocordatos, C.; Berruti, B.; Delclaud, Y.; Riti, J.B.; and Garnier, T. [2012]. *The Ocean and Land Colour Imager (OLCI) for the Sentinel 3 GMES mission: status and first test results*. Proceedings: SPIE 8528, Earth Observing Missions and Sensors: Development, Implementation, and Characterization II, 85280C. doi: 10.1117/12.977247.

Otsu, N. [1979]. *Threshold selection method from gray-level histograms*. IEEE Transactions on Systems, Man, and Cybernetics, doi: 10.1109/tsmc.1979.4310076.

Park, J.; Kim, J-H.; Kim, H-C; Kim, B-K.; Bae, D; Jo, Y-H.; Jo, N.; and Lee, S.H. [2019]. *Reconstruction of ocean color data using machine learning techniques in polar regions: focusing on off Cape Hallett, Ross Sea*. Remote Sensing, doi: 10.3390/rs11111366.

Sathya, R. and Abraham, A. [2013]. *Comparison of supervised and unsupervised learning algorithms for pattern classification*. International Journal of Advanced Research Artificial Intelligence, Vol. 2, No. 2, pp. 34-38.

Unninayar, S. and Olsen, L.M. [2015]. *Monitoring, observations, and remote sensing – global dimensions* ★ . In: Reference Module in Earth Systems and Environmental Sciences.

Xie, X.L. and Beni, G. [1991]. *A validity measure for fuzzy clustering*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 8, pp. 841-847.

Zheng, G. and DiGiacomo, P.M. [2017]. *Remote sensing of chlorophyll-a in coastal waters based on the light absorption coefficient of phytoplankton*. Remote Sensing of Environment, doi: 10.1016/j.rse.2017.09.008.