

Remote sensing data driven bathing water quality assessment using sentinel-3

Antonia Senta, Ljiljana Šerić

Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, R. Boškovića 32, Split 21000, Croatia

Article Info

Article history:

Received Oct 2, 2020

Revised Dec 2, 2020

Accepted Dec 23, 2020

Keywords:

Bathing water quality

DT

KNN

Machine learning

Remote sensing

Sentinel-3

ABSTRACT

In this paper we are investigating the possibility of usage of remote sensing satellite data, more precisely sentinel-3 OLCI and SLSTR data, for assessment of bathing water quality. In this research we used data driven approach and analysis of data in order to pinpoint aspects of remote sensing data that can be useful for bathing water quality assessment. For this purpose we collected satellite images for period from start of June till end of September of 2019 and results of in-situ measurement for the same period. Results of in-situ measurement were correlated with satellite images bands and analyzed. We propose a simple method for rapid assessment of possible deterioration of bathing water quality to be used by public health authorities for better planning of *in situ* measurements. Results of implementation of predictive models based on k-nearest neighbour (KNN) and decision tree (DT) are described.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Antonia Senta

Faculty of Electrical Engineering

Mechanical Engineering and Naval Architecture

University of Split, R. Boskovicica 32, Split 21000, Croatia

Email: Antonia.Senta.00@fesb.hr

1. INTRODUCTION

Bathing water quality has many health implications [1, 2]. Assessments on sea bathing water quality are determined on the basis of criteria defined by regulation on sea bathing water quality (OG 73/08) and the EU directive on management of bathing water quality No 2006/7/EC. Traditionally, bathing water quality is measured and published by public health authorities. However, although effort is made to provide timely information, information about bathing water quality is often not published on time and does not reach the public on time to prevent bathing in polluted water. As an example, if a measurement for one location shows good water quality the next measurement is scheduled in 15 days. State of water quality in a given moment is result of previous nearby activities and events and any deterioration of water quality that happens in that period will not be detected.

In this paper we are investigating if it is possible to make assessment of the bathing water quality based on values of individual bands or combination of bands from collected images. Collected images we used in this study are provided by satellite sentinel-3 using ocean and land colour instrument (OLCI) and sea and land surface temperature instrument (SLSTR). These images are available for scientific and commercial exploitation. In order to evaluate and compare this collected data, we implemented k-nearest neighbors (KNN) and decision tree algorithms (DT) and applied them on *in situ* measurements data on sea bathing water quality

at the beaches in the Republic of Croatia. This measurements data is public and published by Institute of Oceanography and Fisheries from Split, Croatia [3].

Responding quickly to events that pose a risk to human health is a remote sensing mission, so we propose a method and a software system to be used by authorities to make rapid assessment of bathing water quality and detect possible deterioration of bathing water quality to schedule additional *in situ* measurements when necessary. This tool would be beneficial for public health, since deterioration of bathing water quality would be detected more quickly than having only 15 days scheduled measurements. All of these research and implemented algorithms lead us to the first prototype of software system which will be used for classifying bathing water quality in the area in the Bay of Kaštela and the Brač channel.

2. RELATED WORK

Based on directive 2006/7/EC of the European parliament and of the Council, each member state of the European Union need to control coastal bathing water quality by measuring the concentration of Escherichia Coli, Intestinal Enterococci, Total Coliforms and Faecal Coliforms. The results of the monitoring of bathing water quality are disseminated to the public via web page [3, 4]. In recent years *data driven prediction* has found its application in many domains. *Model driven approach* requires hypothesis of physical background of phenomenon and modeled event, while *data driven approach* uses data and mathematical operations on large set of data to find laws and correlations between variables using statistics and filtering. Aside from *in situ* measurements, there is increasing number of the studies using remote sensing data and techniques for assessment of bathing water quality [5]. Their focus is on optical properties as well as chlorophyll [6], turbidity, total suspended matters [7] and dissolved organic matters [8]. Different sensors onboard satellites or aircraft platforms provide an opportunity to choose the most adequate one for the analyzing bathing water quality. In [9] authors used Landsat-8 thermal images for determination of bathing water quality.

Beside remote sensing techniques, there is a wide use of hyperspectral images from drone based coastal water quality monitoring. Specifically, hyperspectral monitoring can be used for measuring turbidity, colored dissolved organic matter (CDOM) and phytoplankton used for characterisation of the spectra of bacteria responsible for water quality [10]. In this paper we decided to use Sentinel-3 data as input data for assessment of bathing water quality on a case study area in the Bay of Kaštela and the Brač channel.

3. MATERIALS AND METHODS

Thanks to the governments programs, such as Copernicus and NASA, there is a vast amount of freely available data which originates from a remote sensing. Remote sensing data comes from various sensors mounted on satellites. The data provided by sensors is characterized by wavelengths (bands or channels), time resolution and spatial resolution. Usually, a sensors measures several bands of wavelength in a single capturing. Specific bands or a combination of bands values can be used for detecting different phenomenon on the Earth surface. As shown in previous analysis, efforts have been made to find correct equations to measure features such as surface temperature, salinity and chlorophyll.

In this paper we investigate if we can identify features of various satellite bands that can indicate changes in bathing water quality. The whole work is strictly data driven - we examined changes in features of remote sensing signals in case of bad bathing water quality. We use data obtained from satellite sensors monitoring both land and sea and associate their values to measured water quality in search for indicators that can detect or predict incidents of lower bathing water quality.

The research question of this study can be formulated as. Is it possible to pinpoint relevant remote sensing bands for assessment of bathing water quality using data driven approach? To answer this question we selected a study area and appropriate data and performed a case study to demonstrate the data driven assessment of bathing water quality using data obtained from Sentinel-3 satellite and discussed results.

3.1. Study area

The study area is located in the Bay of Kaštela and the Brač Channel, as shown in Figure 1. Both locations are located near the town of Split, which is the second largest city in the Republic of Croatia. The main motivation for using this study area was the exceedance of the limit values in summer of 2019 defined by Regulation on Sea bathing water quality (OG73/08) for the microbiological indicator Escherichia Coli and Enterococci. Having in mind these *in situ* results, we were encouraged to use remote sensing techniques,

machine learning algorithms and data mining, in effort to relate *in situ* data and data obtained from satellite image processing.



Figure 1. The satellite image (Sentinel-2 L2A) of study area

3.2. Data sources

In this study we used images related to Sentinel-3 SLSTR and OLCI instrument. The argumentation for using this specific satellite is that it provides the most suitable spatial and temporal resolution for the selected area and the time period of the study. Apart from spatial and temporal resolution, the availability of images for all bands via the API is another reason why we used the Sentinel-3 satellite for this study. For comparison, data originating from moderate resolution imaging spectroradiometer (MODIS) consists of two satellites, terra and aqua, has good time resolution, same as Sentinel-3 data, but worse spatial resolution for majority of bands so it wouldn't be a good choice for collecting and analyzing data for a short time period. Table 1 shows a comparison of the spatial and temporal resolution for the Sentinel-2, Sentinel-3, Landsat-8 and MODIS satellites data and shows that Sentinel-3 is the most suitable for our research.

Table 1. Comparison of satellite spatial and temporal resolution

Satellite	Temporal Resolution	Spatial Resolution	Reference
Sentinel-2	<= 5 days	10m, 20m, and 60m, depending on the wavelength	[11]
Sentinel-3	<= 2 days	SLSTR: 500m for bands S1 to S6, 1km for bands S7 to S9, F1, and F2 OLCI: 300m	[12] [13]
Landsat-8	<= 8 days	130m, thermal (100m resampled to 30m) and panchromatic (15m) bands	[14]
MODIS	<= 2 days	250m (bands 1-2), 500m (bands 3-7), 1000m (bands 8-36)	[15]

Satellite Sentinel-3 is developed by European Space Agency (ESA) as a part of Copernicus programme. Sentinel-3 is a satellite mission consisting of two satellites, Sentinel-3A and Sentinel-3B, developed to provide accurate and timely information for better measurements of sea surface topography, sea and land surface temperature and ocean and land surface colour. Four main instruments on Sentinel-3 spacecraft are Ocean and Land Colour Instrument (OLCI), Sea and Land Surface Temperature Instrument (SLSTR), SAR Radar Altimeter (SRAL) and Microwave Radiometer (MWR) [16]. The motivation to use sensors related both to the sea and land remote sensing is found in the fact that the locations of measurements where we inspect bathing water quality are located on waters in near vicinity of land. Bathing water can be considered as an interface between land and sea and its current state is influenced by both sea and land features.

Images related to SLSTR instrument have 11 bands from which nine are spectral bands S1-S9 and two additional bands are optimised for fire monitoring F1 and F2. Spatial resolution of images obtained by this instrument is 500 m for bands S1 to S6, 1km for bands S7 to S9, F1 and F2. Revisit time for SLSTR is less than or equal to one day using both satellites. OLCI instrument has 21 spectral bands, ranging from the visible to the near-infra-red (400 nm to 1020 nm). Spatial resolution for this instrument is 300 m, while revisit time is two days using both satellites [16].

All images collected from these satellites are taken in period of June 1st 2019 to October 1st 2019, with cloud cover percentage less than 20% for SLSTR images. Data acquired by OLCI and SLSTR sensors are available on Sentinel Hub EO Browser [17]. Measurements for Sentinel-3 SLSTR are processed at Level

1B and represent top of atmosphere reflectance or brightness temperature, while data acquired with the OLCI sensor are at full resolution and processed at Level 1 reflectance. Locations of interest are related to geographical coordinates of *in situ* measurements locations for which we have available data published by Institute of Oceanography and Fisheries from Split, Republic of Croatia [18]. The band values for each location of the study area were extracted from each image. Web scrapping, data filtering, consolidation and all data manipulation was automated in Python [19] scripts.

3.3. Methodology

The methodology consists of several phases required to obtain results, and in this section we will describe: data collection and consolidation, implementation of machine learning techniques for assessment and prediction of bathing water quality.

3.3.1. Data collection and consolidation

The Sentinel-3 satellite, due to its short revisit time, is ideal for extracting data that is time-continuous. This satellite sensors also have good spatial resolution suitable for a small study area. Acquisition of bathing water quality data is done by web scrapping from the web site of Institute of Oceanography and Fisheries from Split, Croatia. The quality of bathing water at a particular location can have one of four ratings: *excellent*, *good*, *satisfactory*, and *unsatisfactory*. *In situ* measurement is performed every 15 days unless the measurement shows pollution. In case of lower quality, measurement is done every day until the results are improved. Due to the very small number of different ratings of samples for individual locations, we replaced these four classes of ratings with two ratings, in a manner that excellent rating is denoted as *good*, and all other ratings are considered suboptimal and denoted as *notGood*.

Remote sensing data was also retrieved using a web API from the Sentinel Hub EO Browser [17]. Images were collected in .tiff format georeferenced in the WGS 84 coordinate system (EPSG: 4326). Collected images coverage corresponds to study area and selected time period - from June 1st 2019 till October 1st 2019. The values of each band for the observed time period were extracted from image related to the location (latitude and longitude) of the *in situ* measurement.

We constructed data set where each measurement represent a sample of data. Sample is described by features of *in situ* measurement and values of each band for that day for two instruments of Sentinel-3 satellite. Each sample is described with a set of features: the measurement date, the bathing water quality score, and the band value for the observed measurement. The Table 2 shows the number of samples with respect to bathing water quality evaluation and used instruments of Sentinel-3 satellite. In this way we associated each measurement with a series of values of satellite sensor bands [20].

Table 2. The number of consolidated data samples in period of 1st June 2019 to 1st October 2019

Evaluation	SLSTR	OLCI
good	865	475
notGood	119	32
Total	984	507

3.3.2. Data analysis and preprocessing

In this section we will present behaviour of all bands, separately for SLSTR and OLCI instruments. We used all available data collected from satellite images consolidated within *in situ* measurements, so we could get two classes of bathing water quality, where one is classified as *good* and other is classified as *notGood* bathing water quality. Figures 2 and 3 present graphs of data distribution for instruments SLSTR and OLCI, corresponding bands values and bathing water quality scores. SLSTR data distribution shows that distribution of values for *good* and *notGood* bathing water quality are most distinctive for bands S5 and S6. Similarly OLCI instruments bands value distribution graph shows the most distinctive values for bands B7, B8, B9, B10 and B11.

After analysis of the distribution of all bands values in order to select bands appropriate for classification problem, we decided to analyse bands behaviour in time domain for two locations, for which we knew that overall rating was *notGood* opposed to those whose overall rating was *good*. To perform further analysis we singled out two locations: one located at Kamp in Kaštel Gomilica that was rated with an overall rating as *notGood* and the other located at Zvončac in Split with overall rating as *good*. We collected values from

SLSTR and OLCI satellite instruments bands that correspond to selected two location for the continuous time period from July 1st 2019 to August 31st 2019. Based on the extracted data, we created time graphs of band values for each location and tried to isolate bands whose values deviated the most for locations with *good* and *notGood* bathing water quality. The time series graphs for all bands for both locations are shown below, one for SLSTR instrument, as shown in Figure 4 and the other for OLCI instrument, as shown in Figure 5.

The SLSTR graph in Figure 4 shows the inconsistency in the bands values during 13 July 2019 to 16 July 2019 and 23 August 2019 to 27 August 2019, especially for bands S5 and S6. Based on the band value graphs and *in situ* data measurements of bathing water quality, we can notice that inconsistency in the data of band values usually occurred during the time, before, or after the determined bathing water pollution.

In Figure 5 we can notice the change of OLCI bands values for the period from August 20, 2019 till September 7, 2019. A sudden change in bands values can be read for August 30, 2019, where the peak of the line representing *notGood* bathing water quality is higher than the line representing *good* bathing water quality. If we pay attention to the previous values of these two lines before the isolated peak, we can conclude that up to that point the band values were following the same trend in periods when both locations had satisfactory measurements.

Based on these observations, we can conclude that our analysis for all locations and for two singled out locations picks out the same bands that have deviation in values between *good* and *notGood* measurements. We choose bands S5 and S6 for SLSTR and B7, B8, B9, B10, B11 for OLCI instrument for further modeling of more accurate algorithms for prediction of bathing water quality. This approach aims toward development of a software system for remote sensing screening of bathing water quality prior to *in situ* measurements. Such a system can be used as a tool for scheduling *in situ* measurements, focused on the locations where deterioration of bathing water quality is suspected.

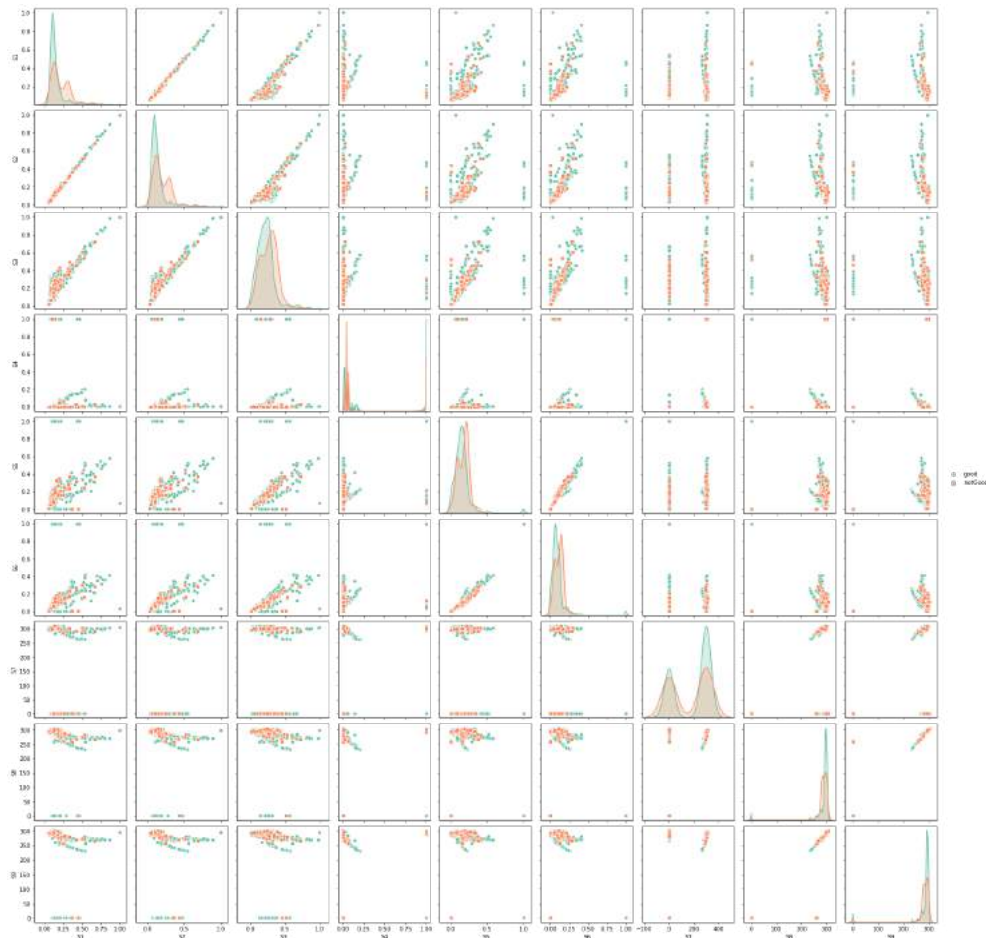


Figure 2. SLSTR-data distribution of all bands

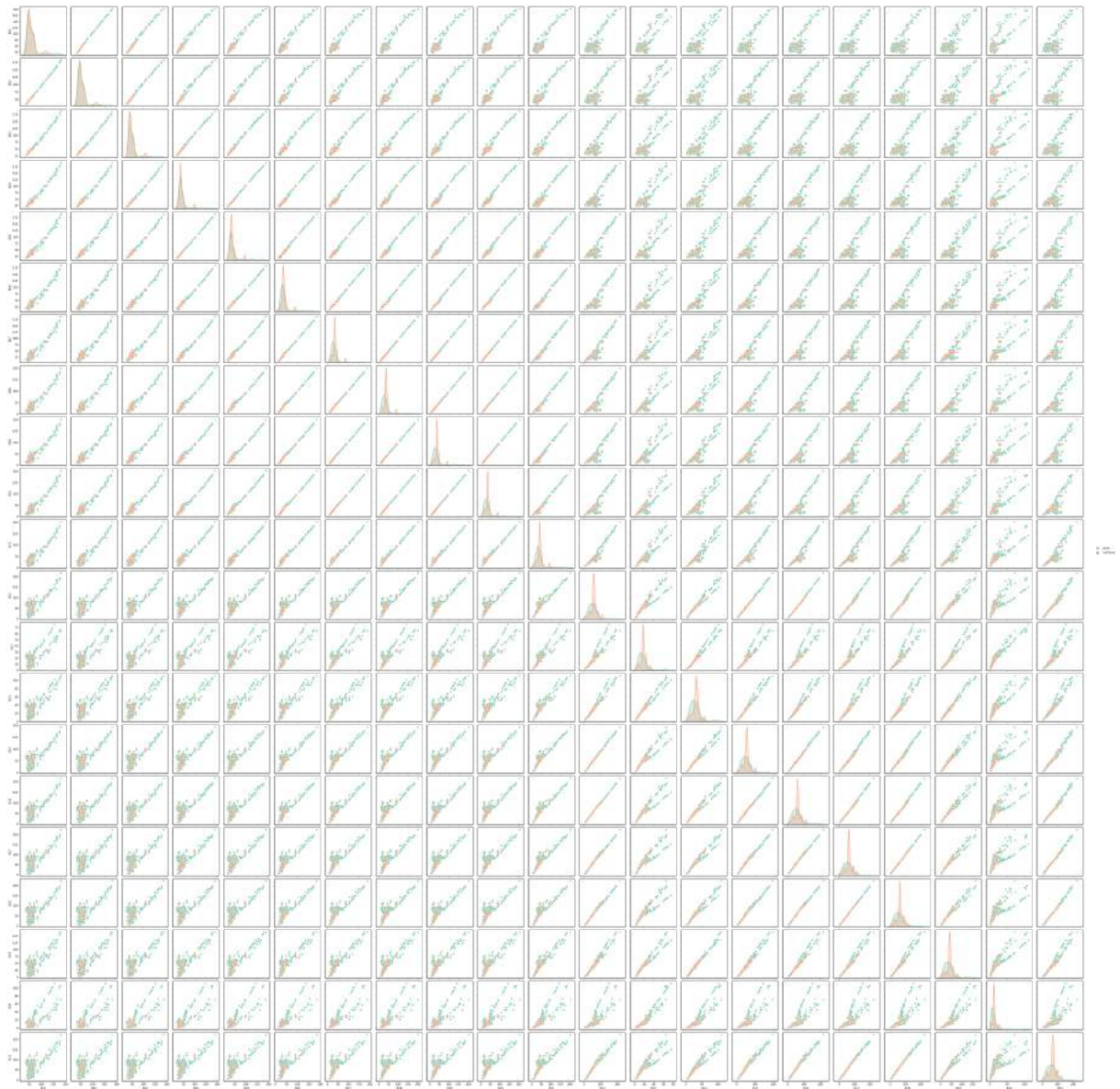


Figure 3. OLCI-data distribution of all band

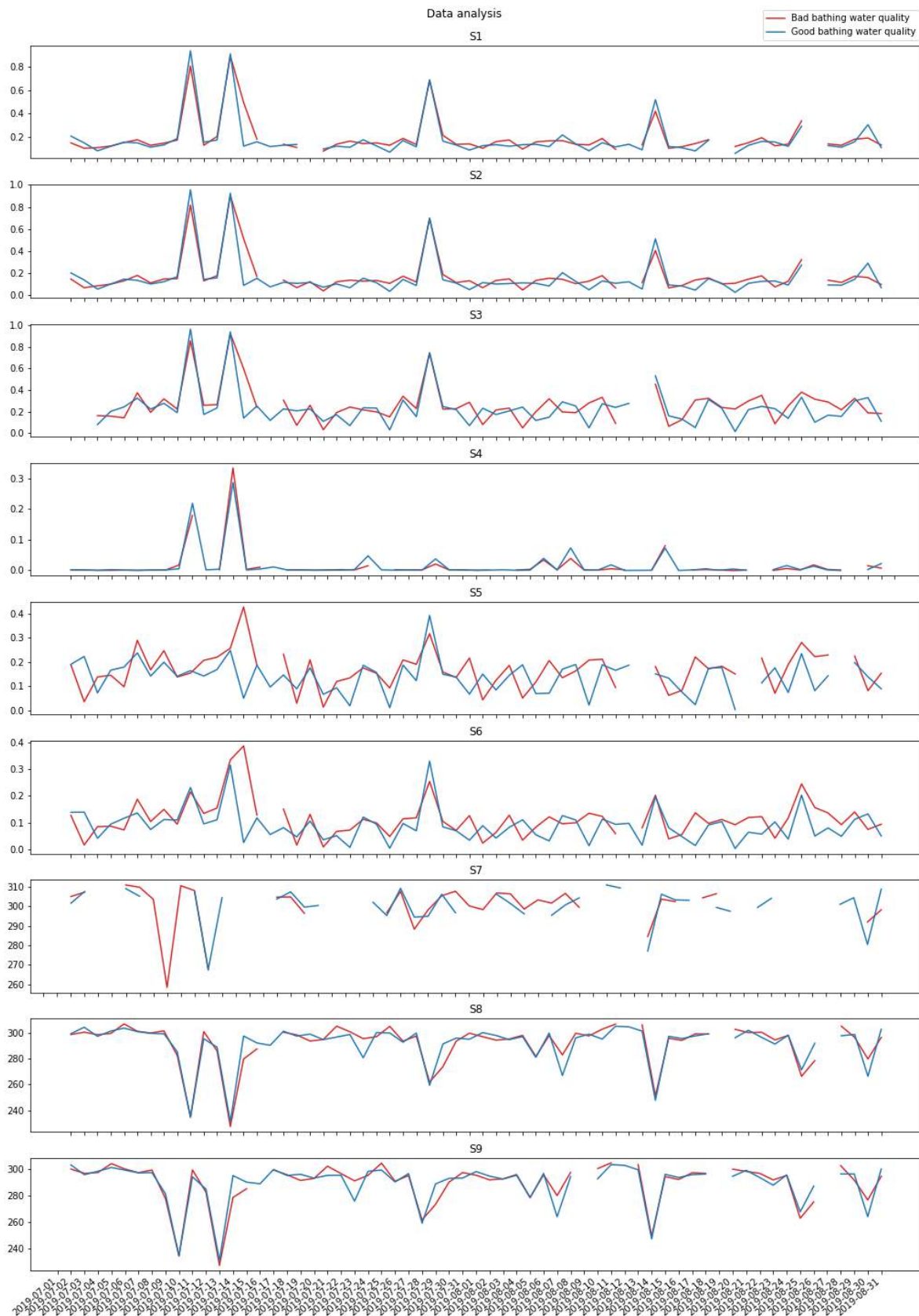


Figure 4. Time series graph of all bands of SLSTR instrument for two selected locations

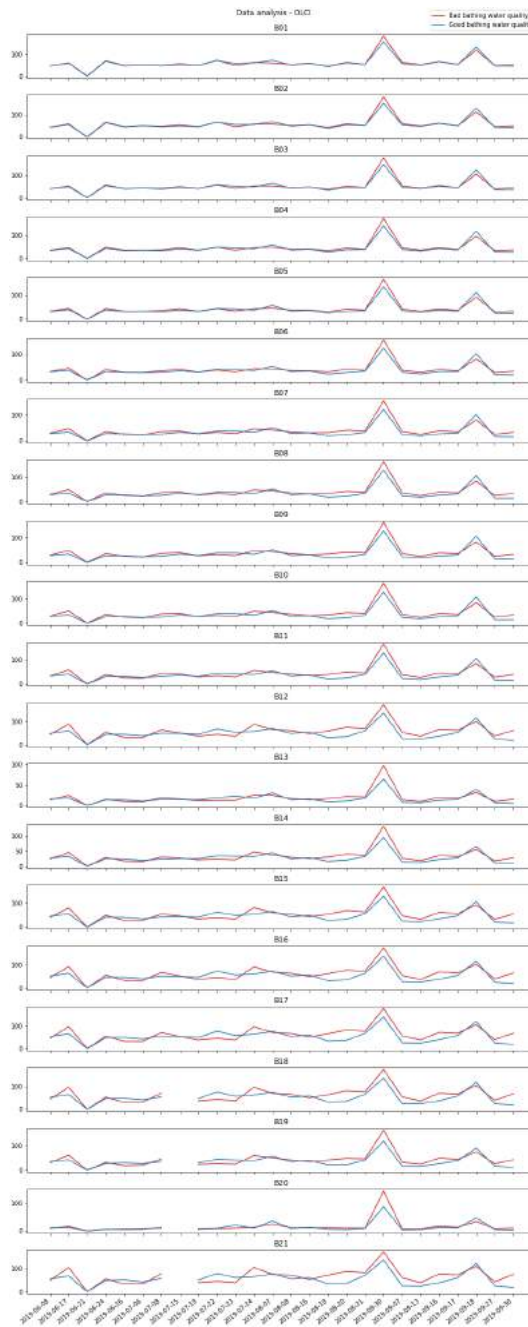


Figure 5. Time series graph of all bands of OLCI instrument for two selected locations

3.3.3. Implementation of assessment and prediction

In this paper we implemented two machine learning algorithms. K-nearest neighbors and decision tree classification as a tools for predictive modeling. Both of these algorithms are good recognition techniques and are widely used for analysis and predictions in remote sensing data sets [21].

K-nearest neighbors (*KNN*) algorithm [22] is an efficient lazy learning algorithm widely applied in pattern recognition and data mining in machine learning classification. It is analytically tractable and highly adaptive to local information. *KNN* algorithm uses the closest data points for estimation. For each data point the algorithm checks against the training table for the *KNN*. Since each data point is independent of the others, the execution of search and score can be conducted in parallel. Samples for training are stored in an n-dimensional space. When an unknown test label is given, the *KNN* classifier searches these samples which are closest to

the unknown sample. Closeness is in this paper defined in terms of Euclidean distance. Euclidean distance closeness was calculated on the basis of the band values and not geographical closeness.

Decision tree (DT) classification algorithm [23] creates a tree model from the data. Model is created by using values of only one attribute at a time. Tree structure is made of root (formed from all of the data), internal (splits) and leaf nodes. Each node in a decision tree has only one parent node and two or more descendant nodes. The algorithm sorts the data set on the attribute's value and then looks for regions in the data set that clearly contain only one class and mark those regions as leaves. For the remaining regions that have more than one classes, the algorithm chooses another attribute and continues the branching process with only the number of instances in those regions, until it produces all leaves or there is no attribute that can be used to produce one or more leaves in the conflicted regions.

These two algorithms were first applied to data set consisting of all bands of OLCI and SLSTR instrument and results are discussed. After the discussion, we selected distinguished bands values based on data analysis of all bands for all locations and two singled out locations described in section 3.3.2. The results of the applied algorithms will be depicted below.

4. RESULTS AND DISCUSSION

In this section the analysis performed on prepared data set will be presented and the results discussed. The discussion is aimed towards argumentation whether selection of smaller subset of bands can be more useful for assessment and prediction of bathing water quality.

4.1. Predictive models used on all SLSTR bands

The data set constructed from the SLSTR instrument, raw bands images are used as an input set for K-nearest neighbor and Decision tree algorithms. For each algorithm, obtained results show influence of the individual bands of the SLSTR instrument in decision if sample will be classified as *good* or *notGood*. Since the data set we used for training the models is imbalanced-selected case study area has much larger number of *good* measurements than *notGood* we trained models with 50% of our data set, and used remaining 50% for evaluation of results. We created a confusion matrix for each of the algorithms (KNN and DT) which shows the performance of the classification model on a set of test data, for which the true values are known. The results of the confusion matrix of both algorithms are shown on Figure 6 and 7. Both of these algorithms are highly accurate, but since data set is imbalanced, model can fail to identify negative labels [24]. In our study, Decision tree algorithm produced better and more precise results than K-nearest neighbor algorithm. The reason for this is that decision tree algorithm identifies many more true positive values than the K-nearest neighbor algorithm does, so even if accuracy of decision tree algorithm is smaller than accuracy of K-nearest neighbor algorithm, true positive and false positive values are in better ratio and make accuracy more reliable.

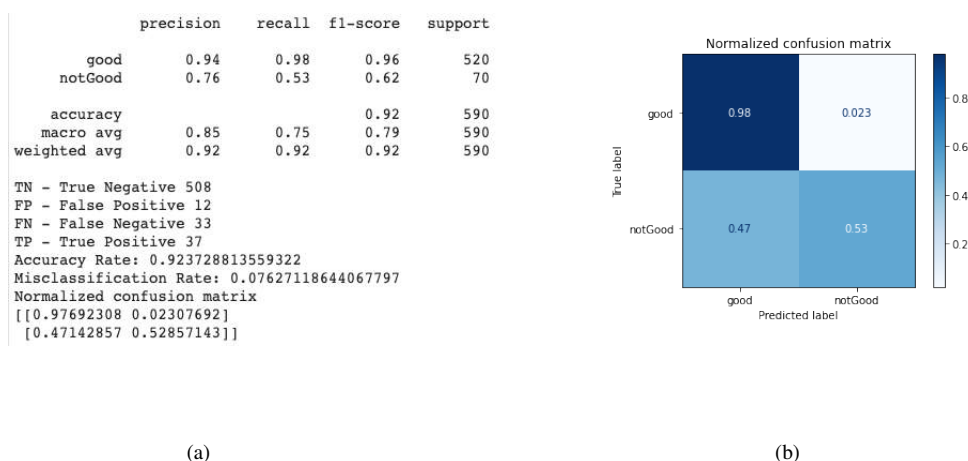


Figure 6. KNN algorithm for SLSTR data, (a) Classification report, (b) Normalized confusion matrix

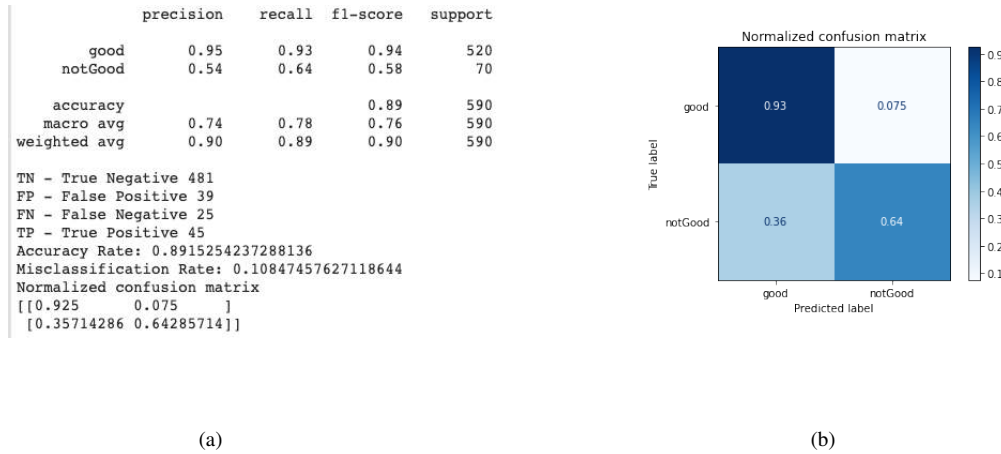


Figure 7. DT algorithm for SLSTR data, (a) Classification report, (b) Normalized confusion matrix

4.2. Predictive models used on all OLCI bands

We performed the same methodology on data from OLCI instrument and implemented K-nearest neighbor and Decision tree algorithms. Amount of SLSTR data is almost double in volume compared to OLCI data because of the OLCI instrument's larger revisit time. Also, OLCI data is even more imbalanced, where data classified as having *good* bathing water quality is in the majority compared to data classified as having *notGood* bathing water quality. The results of the applied K-nearest neighbor and Decision tree algorithms are shown on Figure 8 and 9.

Confusion matrix for both algorithms applied to OLCI data shows high accuracy. But if we look closely, we will notice that because of imbalanced data set, neither of these two models are suitable for predicting deterioration of bathing water quality based only on OLCI data. True positive values in regards to true negative values are in the minority, so our accuracy is not reliable parameter, although it has a high value. Our intuition was that both instruments - OLCI typically used for ocean monitoring and SLSTR typically used for land monitoring should be used for bathing water quality assessment since specific area of interest is located near interface between land and sea.

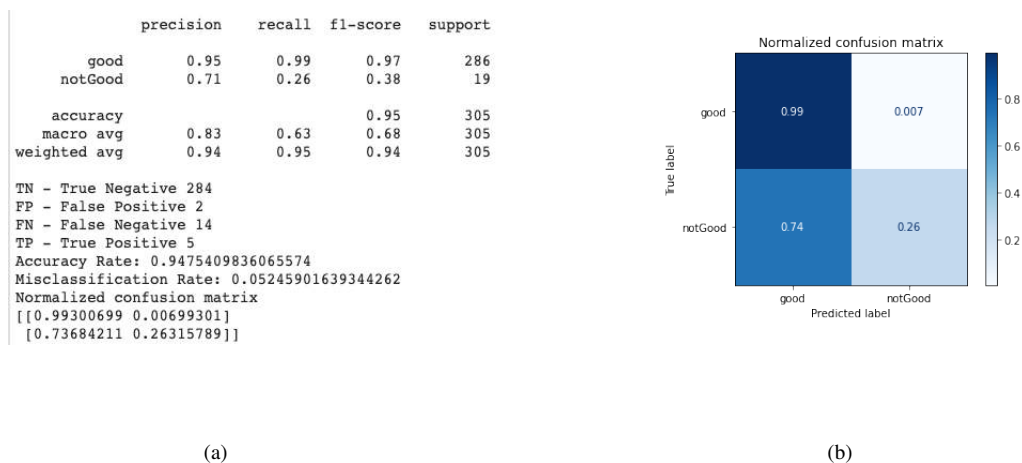


Figure 8. KNN algorithm for OLCI data, (a) Classification report, (b) Normalized confusion matrix

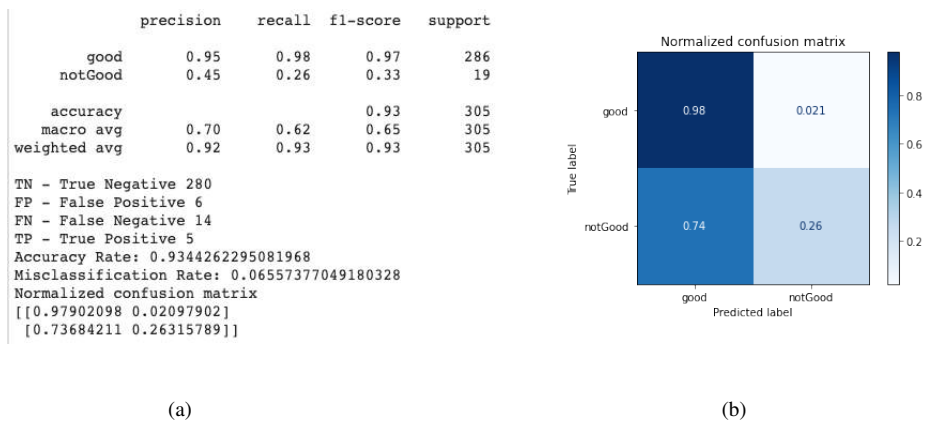


Figure 9. DT algorithm for OLCI data, (a) Classification report, (b) Normalized confusion matrix

4.3. Predictive models used on selection of bands

Based on the analysis of individual bands from section 3.3.2., we decided to extract only the bands where the largest differences in relation to *good* and *notGood* bathing water quality were observed. Similar to previous analysis of data using algorithms, data used for training K-nearest neighbor and decision tree algorithms using only selected bands is also imbalanced in a way that there is much more data with *good* evaluation of bathing water quality. Data were obtained by merging data from the SLSTR and OLCI instruments based on the date and location of the measured satellite values of the selected bands. The bands taken for processing from the SLSTR instrument are S5 and S6, and from the OLCI instrument are B7, B8, B9, B10 and B11. Figure 10 shows difference of distribution of values for selected bands for *good* and *notGood* measurements.

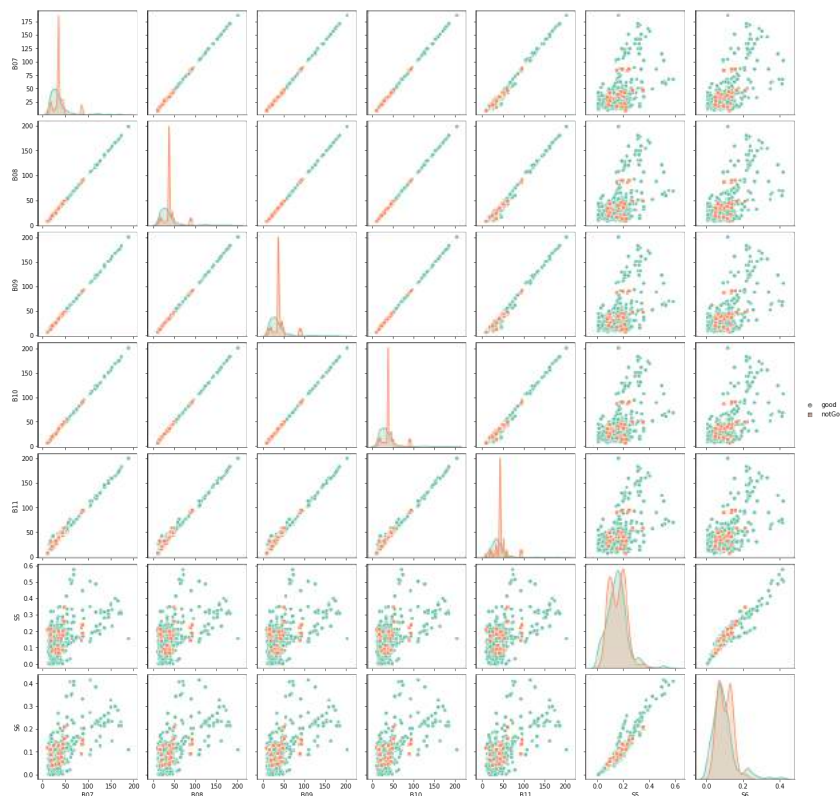


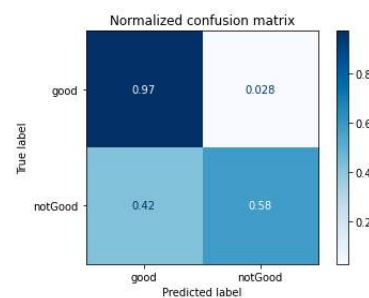
Figure 10. Data distribution of specific bands

Another interesting observation is that bands we selected based on data analysis are the same bands that are used in calculation of indices that are somehow related to bathing water quality. For example, bands B8 and B11 are used in calculation of chlorophyll index *a* (*chl-a*) [25], total suspended matter (TSM), coloured dissolved organic matter (CDOM) and secchi [26]. So, these bands are already proven to have influence on water quality parameters [27]. One more reason why are this indices are good parameters for assessment bathing water quality is that they are optical properties of the bathing water and affect the rapid reaction if bathing water contamination occurs. These properties varies over a quite wide range on interface between land and sea. As we expected, the results shown by confusion matrix for evaluation of KNN and DT based models on selected bands is more promising for intended use than previous models Figure 11 and 12.

	precision	recall	f1-score	support
good	0.97	0.97	0.97	426
notGood	0.61	0.58	0.59	33
accuracy			0.94	459
macro avg	0.79	0.77	0.78	459
weighted avg	0.94	0.94	0.94	459

TN - True Negative 414
 FP - False Positive 12
 FN - False Negative 14
 TP - True Positive 19
 Accuracy Rate: 0.9433551198257081
 Misclassification Rate: 0.05664488017429194
 Normalized confusion matrix
 [[0.97183099 0.02816901]
 [0.42424242 0.57575758]]

(a)



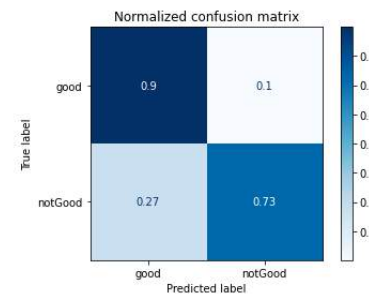
(b)

Figure 11. KNN algorithm results for specific bands, (a) Classification report, (b) Normalized confusion matrix

	precision	recall	f1-score	support
good	0.98	0.90	0.94	426
notGood	0.36	0.73	0.48	33
accuracy			0.89	459
macro avg	0.67	0.81	0.71	459
weighted avg	0.93	0.89	0.90	459

TN - True Negative 383
 FP - False Positive 43
 FN - False Negative 9
 TP - True Positive 24
 Accuracy Rate: 0.8867102396514162
 Misclassification Rate: 0.11328976034858387
 Normalized confusion matrix
 [[0.89906103 0.10093897]
 [0.27272727 0.72727273]]

(a)



(b)

Figure 12. DT algorithm results for specific bands, (a) Classification report, (b) Normalized confusion matrix

We consider that the prediction of bathing water quality using both algorithms on selected bands is more accurate than the prediction based on all bands, because it has significantly higher effect for determining the value of bands which are sensitive and have small difference between values of bands representing *good* and *notGood* bathing water quality. The nature of the task of predicting deterioration of bathing water quality is such that we prefer to have higher rate of false positive as opposed to false negative errors. Evaluation of a model trained on selected bands with DT algorithm shows best value for recall of *notGood* measurement and is the best candidate for implementing a software tool for early screening of bathing water quality.

5. CONCLUSION

This paper presents preliminary results of remote sensing data analysis aimed towards assessing bathing water quality from remote sensing data. The method provided analysis of bands values behaviour for different aspects of bathing water quality. Since we used only raw bands values, the lack of this method can be various adverse weather events and cloud coverage, which are optical parameters that affect the bands values of the individual image. Because of this parameters we decided to use only data set of images with cloud coverage less than 20%. The results of this study indicate that it could be possible to construct a predictive model based on remote sensing data for assessment of bathing water quality since remote sensing data covers aspects of land and sea parameters. Due to small amount of data caused by the small number of *in situ* measurements, results of K-nearest neighbor and Decision tree algorithms constructed in this study can not be generalized. Based on these results we certainly got a broader picture of bands values behavior and limitations, resulting in applying the algorithms to a smaller, but more reliable data set. K-nearest neighbor and Decision tree algorithms display reliable confusion matrix for selected bands, even though confusion matrix has accuracy close to 95% for K-nearest neighbor algorithm and close to 90% for Decision tree algorithm. The method provided in this paper uses many simplifications in time and space dimensions, because only point of data associated with measurement point was taken into account. In future work this will be extended by approximation of band values based on image before and after *in situ* measurement point, when image for the date of measurement is not available. Also, we could increase volume of data using bands values extracted from images taken from different satellites, which will complete time flow and spatial coverage of data. It is reasonable to conclude that higher volume of data would provide more precise and applicable algorithms results for better prediction of bathing water quality. More sophisticated algorithms for model training will also be investigated in future work.

6. ACKNOWLEDGEMENTS

This research was supported through project CAAT (coastal auto-purification assessment technology), funded by the European Union from European Structural and Investment Funds 2014.–2020., Contract Number: KK.01.1.1.04.0064.

REFERENCES

- [1] C Nelson and AT Williams. Bathing water quality and health implications. *WIT Transactions on Ecology and the Environment*, 20, 1997.
- [2] V. J. Cabelli, A. P Dufour, L. J. McCabe, and M. A Levin. Swimming-Associated Gastroenteritis and Water Quality. *American Journal of Epidemiology*, 115(4):606–616, 04 1982.
- [3] Institute of Oceanography and Fisheries. Assessment of sea bathing water quality. (accessed April 7, 2020). <http://baltazar.izor.hr/plazepub/kakvoacaetalji10>.
- [4] European Union. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. *Office Journal of the European Union*, (64/37), 2006.
- [5] Abdiel E. Laureano-Rosario; Andrew P. Duncan; Erin M. Symonds; Dragan A. Savic; Frank E. Muller-Karger. Predicting culturable enterococci exceedances at Escambron Beach, San Juan, Puerto Rico using satellite remote sensing and artificial neural networks. *Journal of Water and Health*, 17(1):137–148, 2019.
- [6] Watanabe Fernanda; Alcântara Enner; Rodrigues Thanan; Rotta Luiz; Bernardo Nariane; Imai Nilton. Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil). *Anais da Academia Brasileira de Ciências*, 90(2, Suppl. 1), 2018.
- [7] A.G.Dekker; R.J.Vos; S.W.M.Peters. Comparison of remote sensing data, model results and in situ data for total suspended matter (TSM) in the southern Frisian lakes. *Science of The Total Environment*, 268, Issues 1–3:197–214, 2001.
- [8] Saleh T Daqamseh, Biswajeet Pradhan, Anas Al-Oraiqat, Maan Habib, et al. Modis derived sea surface salinity, temperature, and chlorophyll-a data for potential fish zone mapping: West red sea coastal areas, saudi arabia. *Sensors*, 19(9):2069, 2019.
- [9] El Khalil Cherif, Farida Salmoun, Francisco Javier Mesas-Carrascosa, et al. Determination of Bathing Water Quality Using Thermal Images Landsat 8 on the West Coast of Tangier: Preliminary Results. *Remote Sensing*, 11(8):972, 2019.

- [10] J Martin, A Medina, A De Luque, O Bergasa, and J Coca. Drone hyperspectral system for monitoring coastal bathing water quality. *YP Remote S*, 7:32, 2018.
- [11] CREODIAS. Sentinel-2 L1 collection. (accessed September 16, 2020). <https://discovery.creodias.eu/dataset/sentinel-2-l1-collection>.
- [12] European Space Agency (ESA). SLSTR. (accessed September 16, 2020). <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-3/data-products/slstr>.
- [13] European Space Agency (ESA). OLCI. (accessed September 16, 2020). <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-3/data-products/olci>.
- [14] Landsat 8. (accessed September 16, 2020). <https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/>.
- [15] European Space Agency (ESA). Mission Facts and Figures. (accessed September 16, 2020). <https://earth.esa.int/web/guest/missions/3rd-party-missions/current-missions/terraaqua-modis>.
- [16] European Space Agency (ESA). User Guides Introduction. (accessed April 7, 2020). <https://sentinel.esa.int/web/sentinel/user-guides>.
- [17] Sinergise Laboratory for geographical information systems, Ltd. Sentinel Hub. (accessed September 15, 2020). <https://apps.sentinel-hub.com/eo-browser/?zoom=10lat=43.42lng=16.56themeId=DEFAULT-THEME>.
- [18] Institute of Oceanography and Fisheries. Official web page. (accessed September 16, 2020). <http://www.izor.hr/web/guest>.
- [19] G. van Rossum. Python tutorial. Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.
- [20] Wang X.; Ling F.; Yao H.; Liu Y.; Xu S. Unsupervised Sub-Pixel Water Body Mapping with Sentinel-3 OLCI Image. *Remote Sensing*, 11(3):327, 2019.
- [21] F. Pirotti; F. Sunar; and M. Piragnolo. Benchmark of machine learning methods for classification of a Sentinel-2 image. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7:335–340, 2016.
- [22] Veena Vijayan V. and Aswathy Ravikumar. Study of Data Mining Algorithms for Prediction and Diagnosis of Diabetes Mellitus. *International Journal of Computer Applications*, 95(17):12–16, 2017.
- [23] Masud Karim and Rashedur M. Rahman. Decision Tree and Naïve Bayes Algorithm for Classification and Generation of Actionable Knowledge for Direct Marketing. *Journal of Software Engineering and Applications*, 6(4):11, 2013.
- [24] M. Brown H. G. Lewis. A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, 22:16:3223–3235, 2001.
- [25] José Luiz Stech Fernanda Sayuri Yoshino Watanabe, Enner Alcântara. High performance of chlorophyll-a prediction algorithms based on simulated OLCI Sentinel-3A bands in cyanobacteria-dominated inland waters. *Advances in Space Research*, 62:265–273, 2018.
- [26] Toming K.; Kutser T.; Uiboupin R.; Arikas A.; Vahter K.; Paavel B. Mapping Water Quality Parameters with Sentinel-3 Ocean and Land Colour Instrument Imagery in the Baltic Sea. *Remote Sens*, 2017; 9(10):1070.
- [27] Blix K; Pálffy K; R. Tóth V; Eltoft T. Remote Sensing of Water Quality Parameters over Lake Balaton by Using Sentinel-3 OLCI. *Water*, 10(10), 1428, 2018.