Saberioon, M., Brom, J., Nedbal, V., Souček, P., Císař, P. (2020): Chlorophyll-a and total suspended solids retrieval and mapping using Sentinel-2A and machine learning for inland waters. - Ecological Indicators, 113, 106236.

https://doi.org/10.1016/j.ecolind.2020.106236

Final version can be accessed here: https://doi.org/10.1016/j.ecolind.2020.106236

Chlorophyll-a and Total Suspended Solids Retrieval and Mapping Using Sentinel-2A and Machine Learning for Inland Waters

Mohammadmehdi Saberioon[a,b,*], Jakub Brom[c], Václav Nedbal[c], Pavel Souček[a], Petr Císař[a]

[a] *University of South Bohemia in České Budějovice, FFPW, South Bohemian Research Centre of Aquaculture and Biodiversity of Hydrocenoses, Institute of Complex Systems, Zámek 136, Nové Hrady 373 33, Czech Republic*
[b] *Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam 14473, Germany*
[c] *University of South Bohemia in České Budějovice, Faculty of Agriculture, Department of Landscape Management, Studentská 1668, České Budějovice 370 05, Czech Republic*

Abstract

Chlorophyll-a (Chl-a) and Total Suspended Solids (TSS) are both key indicators of the biophysical status of inland waters, and their continued monitoring is essential. Existing conventional methods (e.g., in situ monitoring) have shown that they are impractical due to their time and space limitations. The recently operated Sentinel-2A satellite offers the potential to have higher temporal, spatial, and spectral resolution images with no cost for monitoring water quality parameters of inland waters. The main aim of this study was to develop a semi-empirical model for predicting water quality parameters by combining Sentinel-2A data and machine learning methods using samples collected from several water reservoirs within the southern part of the Czech Republic, Central Europe. A combination of 10 spectral bands of the Sentinel-2A and 19 spectral indices, as independent variables, were used to train prediction models (i.e., Cubist) and then produce spatial distribution maps for both Chl-a and TSS. The results showed that the prediction accuracy based on Sentinel-2A was adequate for both Chl-a ($R^2 = 0.85$, $RMSE_p = 48.57$) and TSS ($R^2 = 0.80$, $RMSE_p = 19.55$). The spatial distribution maps derived from Sentinel-2A performed well where Chl-a and TSS were relatively high. The temporal changes in both Chl-a and TSS could be seen in the distribution maps. The temporal changes are showing that The values of TSS dramatically changed in fishponds compared to sand lakes over time which might be due to indifferent management practices. Overall, it can be concluded that Sentinel-2A, when coupled with machine learning algorithms, could be employed as a reliable, inexpensive, and accurate instrument for monitoring the biophysical status of small inland waters like fishponds and sandpit lakes.

Keywords: Water quality, Small inland waters, Cubist modelling, Remote sensing, Monitoring, Fish ponds

---

1. Introduction

Inland waters are the primary source of drinking water and irrigation and are critical to recreational and industrial needs such as energy production, transportation, and fisheries (Carvalho

---

[*] Corresponding author; email: msaberioon@frov.jcu.cz, saberioon@gfz-potsdam.de

et al., 2013). Additionally, they not only provide the habitat for fauna and flora but are also very crucial in the global carbon cycle and climate change (Tranvik et al., 2009; Moss, 2012). Over recent decades, the freshwater quality has been threatened by many human and environmental stressors, posing a significant threat to not only water security but also to the entire ecology system. Therefore, with respect to the above-mentioned dynamical effects, it is essential to have a comprehensive, accurate, fast, and inexpensive monitoring system to observe the biophysical and biochemical conditions of these water bodies to prevent severe damage occurring by applying on-time treatments. An existing conventional in situ monitoring system coupled with geostatistical methods was shown to be impractical due to its time and space limitations (Philipson et al., 2016). On the contrary, Earth observation (EO) techniques have been used by many researchers as efficient methods for retrieving and mapping some water quality parameters due to their micro-dynamic characteristics.

Generally, optical remote sensors from different platforms record radiation from the water's surface to derive information about water properties such as physiochemical properties (e.g. turbidity, total suspended solids (TSS)), organic properties (e.g., total organic carbon (TOC), tentatively identified compounds (TICs)), and microbiological properties (Chlorophyll-a (Chl-a)) (Dörnhöfer et al., 2018; Matsushita et al., 2016; Tyler et al., 2016). Researchers used different remote sensing platforms to quantify and map different water properties for inland waters, for instance, unmanned aerial vehicles (Guimar£es et al., 2019), airborne platforms such as CASI, AISA, and APEX (Hunter et al., 2010; Rößler et al., 2013), and satellites like MERIS (Bresciani et al., 2011),MODIS (Koponen et al., 2004; McCullough et al., 2012), SeaWIFS (Gohin et al., 2019), Landsat (Boucher et al., 2018), and Quickbird (Heblinski et al., 2011). Recently, Dörnhöfer & Oppelt (2016) listed different remote sensing platforms and sensors used for monitoring lake water properties.

Since the late 1970s, satellite remote sensing for monitoring water quality for inland water was set back due to lack of appropriate sensors such as a lack of a sufficient number of spectral bands as well as relatively low radiometric sensitivity and low spatial and temporal resolution (Matsushita et al., 2016; Mouw et al., 2015). For instance, Landsat 1-7 has limited radiometric resolution, and the spatial resolution of the moderate resolution imaging spectroradiometer is not suitable for inland water. However, with the availability of new satellites with a higher spatial, spectral, and temporal resolution, like Landsat-8 and Sentinel-2, water quality retrieval and mapping from the orbit have become more reachable.

The multispectral imager (MSI) aboard Sentinel-2, which was launched on 23 June 2015 with a combination of wide coverage (swath width of 290 km), spatial resolution (10–60 m), and a minimum of five days temporal resolution, provides an exceptional perspective on inland water remote sensing (Drusch et al., 2012). Researchers showed that Sentinel-2 not only can improve global inland water mapping (Du et al., 2016) but can offer a useful range of information for monitoring certain water quality indicators (Toming et al., 2016; Pahlevan et al., 2017). For instance, Toming et al. (2016) showed the suitability of Sentinel-2 data to map different water quality parameters, namely Chl-a, water color, CDOM, and DOC for small inland waters. In Grendait et al. (2018), Sentinel-2 images were used to predict the Chl-a concentration in eutrophic lakes in Lithuania. Chl-a was predicted with an accuracy range between 0.45 and 0.76. In Ansper & Alikas (2018), the suitability of Sentinel-2 A for retrieving Chl-a from water bodies was evaluated. Kutser et al. (2018) also utilized Sentinel-2 data for mapping several water quality parameters such as Chl-a, TSM and CDOM for shallow waters in Baltic sea. Additionally, Pahlevan et al. (2019) evaluated and compared the Landsat 8 and Sentinel-2A/B top of atmospheric, reflectance, and remote sensing reflectance to estimate TSS. Giardino et al. (2019) used

Sentinel-2A to determine the color of water of 170 Italian lakes as a water quality attribute. In other words, MSI has four visible bands, three near-infrared (NIR) bands which certainly makes MSI more potent for the retrievals of concentrations of Chl-a or other pigments in severe bloom conditions (Gower et al., 2005; Moses et al., 2009). Additionally, for accurate measurement of TSS, sensors are needed with a red and NIR band and sufficiently high signal-to-noise ratio (Ruddick et al., 2016; Caballero et al., 2018) which MSI has both bands at high spatial resolution.

One of the widely used methods for the retrieval of water quality properties from remote sensing data for optically complex waters (i.e., inland waters) is band ratio based algorithms. Commonly, band ratio based algorithms can be expressed as the band ratio of surface reflectance (ρ_w) at two, three, or four bands. Usually, these bands are a combination of one (ρ_w) in the red spectrum and two or three (ρ_w) in the near-infrared (NIR) spectrum (Matsushita et al., 2016). For instance, Le et al. (2009) proposed that the combination of four bands at 662, 693, 740, and 705 nm could be used to predict Chl-a in highly turbid waters. In Moses et al. (2009), two (i.e., 665 and 708 nm) and three band (i.e., 665, 708, and 753 nm) models from MERIS were used to predict the Chl-a concentration in inland and turbid coastal waters. The model was shown to predict Chl-a with 96% and 94% accuracy when two and three band models were used, respectively. In Gilerson et al. (2010), a ratio of (ρ_w) at 709 nm was shown, and 665 nm was used to predict Chl-a in moderately turbid waters. These wavelengths correspond with the maximum spectral reflectance of cell tissue and Chl-a of green algae, respectively (Moses et al., 2009; Gilerson et al., 2010). In Ansper & Alikas (2018), it was shown that three and four band ratio models can estimate the Chl-a at levels close to in situ measurements.

Since water properties have complex optical characteristics that strongly affect the performance of the different prediction approaches, different studies provide different results (Kallio et al., 2001; Pepe et al., 2001). Therefore, introducing more efficient approaches is greatly demanded. Consequently, the primary objective of the current study is to introduce a novel approach to use Sentinel-2 water surface reflectance (ρ_w) for retrieving and mapping selected water quality parameters such as Chl-a and TSS for small inland bodies of water. Water quality properties with high dimensional spectral data require intelligent feature extraction, which can be acquired by using machine learning algorithms including the support vector machine (Matarrese et al., 2008), neural network (Sudheer et al., 2006; Mas & Flores, 2008; Chebud et al., 2012), and extreme machine learning (Peterson et al., 2018). Despite physical models, machine learning algorithms are a better approach for handling complex problems without prior knowledge (Chang et al., 2013; Keller et al., 2018) where the limited assumption is required. Additionally, they are less affected by the atmospheric and other background factors under non-ideal contexts (Chebud et al., 2012). Therefore, another objective of this experiment was to develop a semi-empirical model based on the machine learning algorithm for predicting and mapping Chl-a and TSS by considering ten spectral bands and the most available water indices derived from Sentinel-2A images. The introduced method is a completely data-driven approach and does not rely on any prior knowledge. It also not only can be used as an efficient approach to other small inland waters with similar conditions, but also it provides vital information about water quality parameters in a manner that is faster, more accurate, and computationally cheaper than other methods.

2. Materials and Methods

2.1. Study area

Samples of water were collected from water reservoirs within the southern part of the Czech Republic, Central Europe and analyzed for their quality (for more detail, see Fig. 1). To gain a large spectra of various water quality levels, fishponds and sandpit lakes were selected for water sampling. The spatial extent of observed reservoirs varied in the order of tens to hundreds of hectares.

Both fishponds and sandpit lakes were observed in the area of the Biosphere Reserve and Landscape Protection Area Třeboňsko between the towns Třeboň and Veselí nad Lužnicí (South Bohemia). The territory is very flat with an elevation of approximately 420 m a.s.l. The mean annual temperature varies by about 7.8 °C, and the annual sum of precipitation is circa 650 mm. Fishponds are shallow artificial lakes with a depth of up to 2 m that were developed between the 15th and 19th Centuries. The usage of the fishponds is mainly for fish production, mostly common carp (*Cyprinus carpio* L.). The fishponds are supplied by water using a system of ditches and channels. The fishponds are usually very turbid and hypertrophic, typically with very low transparency (tens of centimeters). Sandpit lakes are water reservoirs created in pits after sand mining. The lakes are currently used mostly for recreation and partly for mining (sandpit lake Horusice). The depth of observed sandpit lakes was up to 10 m. The sandpit lakes are predominantly supplied by underground water. The water is relatively clear, oligotrophic to eutrophic (depends on the age of lake) with transparency of about 1 m. Water samples taken from sandpit lakes were used as a reference for the water samples from fishponds. An overview of the essential characteristics of observed reservoirs is shown in Table 1.

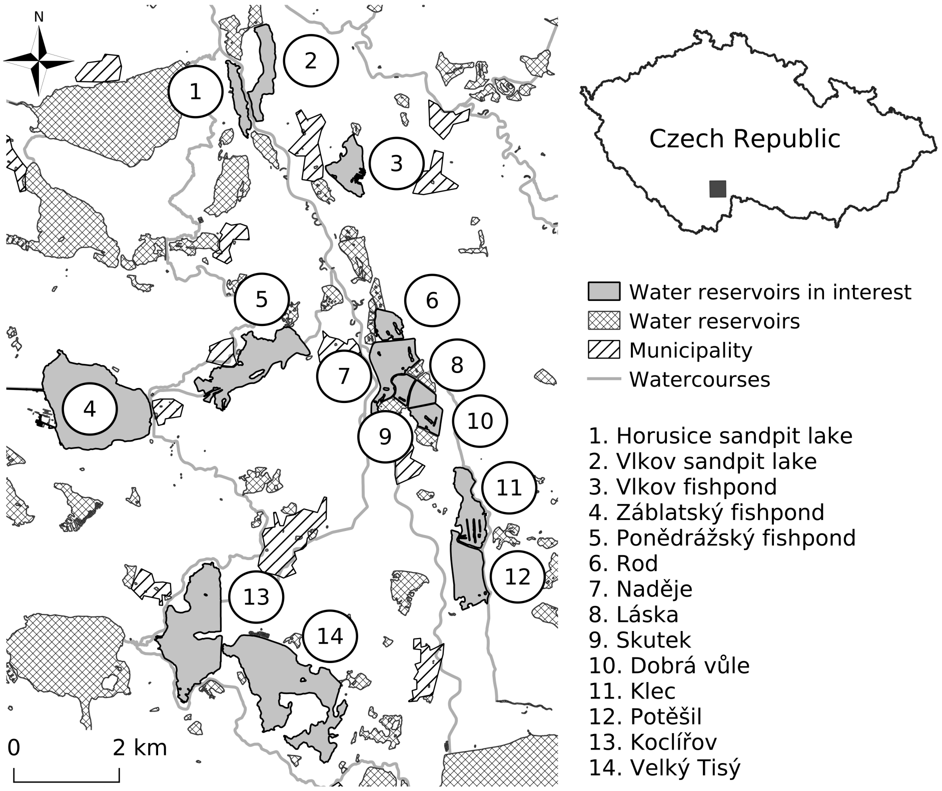


Figure 1: Map of water reservoirs used for water sampling in this study.

Table 1: Basic characteristics of water reservoirs used for water sampling.

Reservoir name	**Type**[a]	**Trophy**[b]	**Area (ha)**	**Altitude (m a.s.l.)**	**Depth (m)**	**Catchment**
Koclířov	F	H	184.6	425	<2	Lužnice
Velký Tisý	F	H	224.8	425	<2	Lužnice
Záblatský	F	H	261.3	430	<2	Lužnice
Ponědražský	F	H	117.5	420	<2	Lužnice
Vlkovský	F	H	44.4	415	<2	Lužnice
Rod	F	H	23.5	415	<2	Lužnice
Naděje	F	H	65.7	415	<2	Lužnice
Láska	F	H	15.6	415	<2	Lužnice
Skutek	F	H	17.9	415	<2	Lužnice
Dobrá vůle	F	H	23.5	415	<2	Lužnice
Klec	F	H	55.2	415	<4	Lužnice
Potěšil	F	H	67.3	415	<4	Lužnice
Sandpit lake Horusice	S	O, M	27.2	410	<10	Lužnice
Sandpit lake Vlkov	S	E	46.0	410	<6	Lužnice

[a] Main type of reservoir: F—fishpond, S—sandpit lake.

[b] Trophy of reservoirs: O—oligotrophic, M—mesotrophic, E—eutrophic, H—hypertrophic.

2.2. Ground sampling and water quality variable measurements

Water samples were collected from water reservoirs during the summer seasons in the years 2017 and 2018. Data were collected from May to October because this time period is the most important in fishponds management point of view. Furthermore, development of algae communities is the most intensive during this period.

Terms of water sample collection were synchronized with Sentinel 2 satellite data acquisition in the area of interest. The reason for the synchronization of data collection was to ensure the comparability of satellite and ground data. Data collection details are shown in Table 2. Water from fishponds and sandpit lakes was sampled at noon, with one or two samples collected from each reservoir.

Water samples were collected from the surface layer in the column of approx. 0.2 to 0.3 m to polyethylene bottles and transported to the laboratory within 4 h. Each sampling point was recorded using a GPS tracker. The distance of sampling points from a bank was greater than 100 m.

Chl-a values were estimated by the reading of absorbance with a double beam UVVis spectrophotometer Heλios Alpha (Unicam, GB) at 664 nm after extraction with a mixture of 90 % acetone:methanol (Pechar, 1987). TSS was determined as the dry weight of seston captured on pre-weighed Whatman GF/C filters and dried to a constant weight at 105 °C.

Table 2: Details of data collected from the study area.

Samples (no.)	***in situ* sampling**	**Superspectral Sentinel-2**	**Properties**
16	11.05.2017	11.05.2017	Chl-a , TSS
16	13.06.2017	13.06.2017	Chl-a, TSS
7	20.06.2017	20.06.2017	Chl-a, TSS
11	03.08.2017	3.08.2017	Chl-a, TSS
12	30.08.2017	30.08.2017	Chl-a, TSS
11	17.10.2017	17.10.2017	Chl-a
19	07.08.2018	07.08.2018	Chl-a
21	17.08.2018	17.08.2018	Chl-a
11	27.08.2018	27.08.2018	Chl-a, TSS
7	16.10.2018	16.10.2018	Chl-a, TSS

2.3. Superspectral satellite data pre-processing and indices retrieval

Ten cloud-free Sentinel-2 images (Level 1C processing) were downloaded from the ESA Sentinels Scientific Data hub according to the closest dates to field sampling (Table 2). All Sentinel-2 level-1C data were atmospherically corrected with ACOLITE software, which is completely image-based. ACOLITE uses the Dark Spectrum Fitting (DSF) algorithm to convert ToA data to water surface reflectance data (ρ_w). The DFS algorithm initially corrected images for atmospheric gas transmittance and sky reflectance. DFS is based on the application of Lookup tables (LUTs) constructed automatically using standard 6SV continental and maritime models (i.e., based on the lowest aerosol optical thickness (τ_a)), except pre-defined dark bands (e.g., NIR and SWIR) were not used; rather, the best model was selected based on the lowest dark spectrum for each band ($\rho_{path}(\lambda)$) (Vanhellemont & Ruddick, 2018). This approach prevents unrealistic negative (i.e., over-corrected) reflectances after atmospheric correction (Kuhn & Quinlan, 2018). Additionally, along with the atmospheric correction, the vicarious calibration gains provided by Pahlevan et al. (2017, 2019) were applied in this study to improve some of the existing biases in MSI-derived products.

Nearest neighbor resampling was used from the original 20 m spatial resolution to the 10 m resolution of the Sentinel-2 bands. This method was chosen, because it is computationally efficient and preserves the input image pixel values (Roy et al., 2016).

The analysis was performed using two sets of remote sensing variables including the water surface reflectance ρ of 10 extracted bands (Table 4) from the Sentinel-2 and 19 calculated spectral indices (Table 3) as co-variances, which was expected to improve the prediction capability. Two different groups of spectral indices including vegetation indices (which are sensitive to Chl-a) and water indices (which are sensitive to TSS) were calculated to indirectly retrieve variables through inter-correlation between target traits. The employed spectral indices were the Normalized Differences Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), Normalized Difference Turbidity Index (NDTI), Water Ratio Index (WRI), Automated Water Extraction Index (AWEI), Simple Ratio (SR), and Simple Ratio Water Color (SRWC). To the best of our knowledge, no studies have been used the proposed methodology for predicting water quality traits.

Table 3: Derived indices details.

Index	**Definition**	**Definition based on Sentinel-2**	**Reference**
NDVI	$(\rho_{NIR} - \rho_{Red})/(\rho_{NIR} + \rho_{Red})$	$B8 - B4/B8 + B4$	(Rouse et al., 1974)
NDWI1	$(\rho_{Green} - \rho_{NIR_1})/(\rho_{Green} + \rho_{NIR_1})$	$B3 - B8/B3 + B8$	(McFeeters, 1996)
NDWI2	$(\rho_{NIR_1} - \rho_{SWIR_1})/(\rho_{NIR_1} + \rho_{SWIR_1})$	$B8 - B11/B8 + B11$	(Gao, 1996)
NDWI3	$(\rho_{NIR_1} - \rho_{SWIR_2})/(\rho_{NIR_1} + \rho_{SWIR_2})$	$B8 - B12/B8 + B12$	(Gao, 1996)
NDWI4	$((\rho_{NIR_2} - \rho_{SWIR_1})/(\rho_{NIR_2} + \rho_{SWIR_2})$	$B8 - B11/B8 + B12$	(Gao, 1996)
NDWI5	$((\rho_{NIR_2} - \rho_{SWIR_2})/(\rho_{NIR_2} + \rho_{SWIR_1})$	$B8 - B12/B8 + B11$	(Gao, 1996)
MNDWI1	$(\rho_{Green} - \rho_{SWIR_1})/(\rho_{Green} + \rho_{SWIR_1})$	$B3 - B11/B3 + B11$	(Xu, 2006)
MNDWI2	$((\rho_{Green} - \rho_{SWIR_2})/(\rho_{Green} + \rho_{SWIR_2})$	$B3 - B12/B3 + B12$	(Xu, 2006)
MNDWI3	$((\rho_{Green} - \rho_{SWIR_1})/(\rho_{Green} + \rho_{SWIR_2})$	$B3 - B11/B3 + B12$	(Xu, 2006)
MNDWI4	$((\rho_{Green} - \rho_{SWIR_2})/(\rho_{Green} + \rho_{SWIR_1})$	$B3 - B12/B3 + B11$	(Xu, 2006)
NDTI	$((\rho_{Red} - \rho_{Green})/(\rho_{Red} + \rho_{Green})$	$B4 - B3/B4 + B3$	(Lacaux et al., 2007)
WRI1	$((\rho_{Green} + \rho_{Red})/(\rho_{NIR} + \rho_{SWIR_1})$	$B3 + B4/B8 + B11$	(Mukherjee & Samuel, 2016)
WRI2	$((\rho_{Green} + \rho_{Red})/(\rho_{NIR} + \rho_{SWIR_2})$	$B3 + B4/B8 + B12$	(Mukherjee & Samuel, 2016)
AWEI1	$4 \times (\rho_{Green} - \rho_{SWIR_1}) - (0.25 \times \rho_{NIR} + 2.75 \times \rho_{SWIR_1}$	$4 \times (B3 - B11) - (0.25 \times B8 + 2.75 \times B11)$	(Feyisa et al., 2014)
AWEI2	$4 \times (\rho_{Green} - \rho_{SWIR_2}) - (0.25 \times \rho_{NIR} + 2.75 \times \rho_{SWIR_2}$	$4 \times (B3 - B12) - (0.25 \times B8 + 2.75 \times B12$	(Feyisa et al., 2014)
AWEI3	$4 \times (\rho_{Green} - \rho_{SWIR_1}) - (0.25 \times \rho_{NIR} + 2.75 \times \rho_{SWIR_2}$	$4 \times (B3 - B11) - (0.25 \times B8 + 2.75 \times B12$	(Feyisa et al., 2014)
AWEI4	$4 \times (\rho_{Green} - \rho_{SWIR_2}) - (0.25 \times \rho_{NIR} + 2.75 \times \rho_{SWIR_1}$	$4 \times (B3 - B12) - (0.25 \times B8 + 2.75 \times B11$	(Feyisa et al., 2014)
SR	ρ_{Red}/ρ_{NIR}	$B4/B8$	(Birth & McVey, 1968)
SRWC	ρ_{Red}/ρ_{Blue}	$B4/B2$	(Zarco-Tejada & Ustin, 2001)

Table 4: Technical details of Sentinel-2 bands used in this study.

Band	Spectral Range (nm)	Central Wavelength (nm)	Bandwidth (nm)	Spatial Resolution (m)	SNR
B2	458–523	492	65	10	154
B3	543–578	560	35	10	168
B4	650–680	665	30	10	142
B5	698–713	704	15	20	117
B6	733–748	740	15	20	89
B7	773–793	783	20	20	105
B8	785–900	833	115	10	174
B8a	855–875	865	20	20	72
B11	1565–1655	1641	90	20	100
B12	2100–2280	2202	180	20	100

2.4. Modeling and prediction performance assessment

The dataset was divided into training and validation sets using random stratified sampling.The training set (70% of total samples) was used for the fitting model, and the testing set (30% of total samples) was used to assess the prediction accuracy of models. To develop the prediction model, Cubist, which is an extension of the M5 model trees (Quinlan, 1992), was used. Cubist is a form of rule-based regression which initially partitions the response data into subsets within which their characteristics are similar concerning the predictors (i.e., Sentinel2A bands and spectral indices) based on a series of hierarchically arranged rules. Additionally, the ensemble of the rule-based model, called the committee, and the number of neighboring observations were adjusted to improve the predictability and stability of the models (Rossel & Webster, 2012). In other words, the Cubist permits to add multiple training committees and reinforcement to make the weights more balanced in comparison to other similar algorithms such as random forest (Kuhn & Quinlan, 2018; Zhou et al., 2019). Cubist has several advantages including (a) it requires the relatively small number of effective tuning hyperparameters, (b) it minimized the risk of overfitting, and (c) it easily can be interpreted due to availability of variable importance in the final predictor model (Zhou et al., 2019).

The error of the prediction model was evaluated by repeated 10-fold cross-validation of the training set (70% of samples) and by using the root-mean-square error (RMSE). The coefficient of determination (R^2) was also measured to show how well the variation of one variable explains the variation in the other. Generally, the largest R^2 and smallest $RMSE_p$ values give the best prediction model. R package Caret (Kuhn, 2018) and Cubist (Kuhn & Quinlan, 2018) were used together for the Cubist regression model.

2.5. Distribution mapping

Once the model was validated, it applied to all spatial data (i.e., Sentinel-2 images from water bodies) to predict the spatial variability of both Chl-a and TSS and create the geospatial raster dataset. The final maps of water properties were produced using R software (R Development Core Team, Vienna, Austria).

3. Results

3.1. Water quality descriptive statistics and correlations

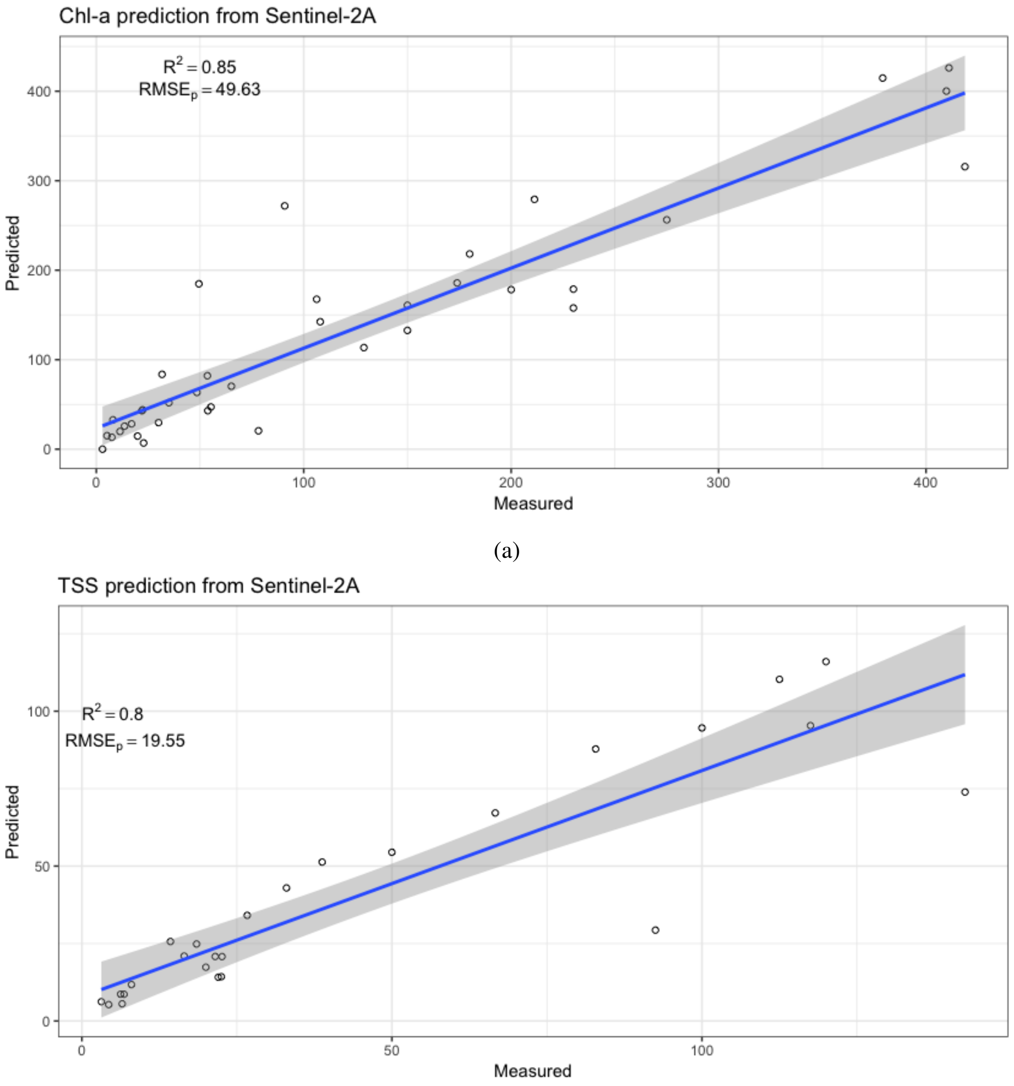
Descriptive statistical results of both Chl-a and TSS from all water bodies including the mean, minimum, maximum, SD, and Coefficient of Variation (CV) are shown in Table 5. Generally, Chl-a increased with the start of algae growth in June, reached its maximum in August, and declined in September. This trend was seen for TSS as well. In other words, both Chl-a and TSS followed the same trend. A comparison of attributes' CV values showed that during June, both Chl-a and TSS had the highest CV values, 192.64% and 85.60%, respectively. In contrast, Chl-a and TSS had the lowest CV values during October, which shows that their distributions are more homogeneous during October than on other dates.

Table 5: Statistical description of water properties

Sampling date	Chl-a					TSS				
	Mean	Min	Max	SD	CV(%)	Mean	Min	Max	SD	CV(%)
11.05.2017	26.29	2.142	111.55	28.57	108.67	7.44	2.00	22.60	6.19	82.79
13.06.2017	78.49	3.94	379.27	85.50	108.92	23.36	7.8	65.00	14.95	63.99
20.06.2017	74.18	2.99	397.12	142.91	192.64	17.91	6.2	51.0	15.33	85.60
03.08.2017	214.09	9.93	430.84	157.68	73.65	67.22	5.8	195.0	54.14	80.54
30.08.2017	245.70	8.22	509.79	149.10	60.43	61.28	5.8	90.0	26.21	42.78
17.10.2017	96.70	5.71	423.26	114.25	118.14	-	-	-	-	-
07.08.2018	36.83	17	150	33.150	90.00	-	-	-	-	-
17.08.2018	110.76	15	240	86.99	78.54	-	-	-	-	-
27.08.2018	283.70	8.05	672.58	235.45	83.34	69.62	4.4	120.0	44.89	64.48
16.10.2018	219.43	128.52	355.57	82.64	37.66	72.57	33.0	142.42	34.59	47.67

3.2. Water variable prediction using Sentinel-2A data

Figure 2 presents the results of the Cubic modeling of water quality traits using superspectral Sentinel-2A data. The estimation of water quality properties provided rather good results for Chl-a, which was predicted with $R_p^2 = 0.85$ and $RMSE_p = 49.64$. Although, the obtained accuracy for TSS was satisfactory ($R_p^2 = 0.80$ and $RMSE_p = 19.55$), it was a bit lower than that of Chl-a. The above-mentioned results highlight the fact that data from Sentinel-2A are suitable for predicting both Chl-a and TSS in this study area.



(a)

(b)

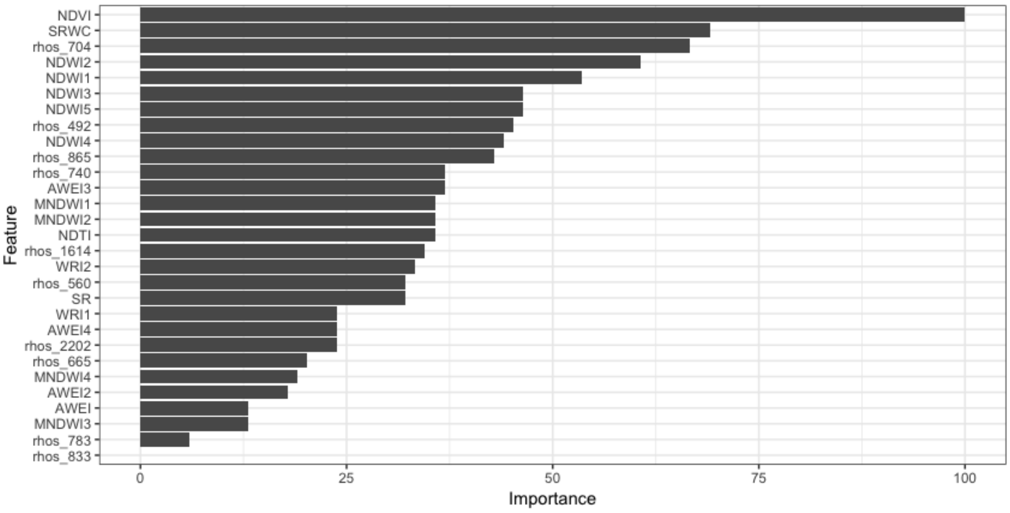
Figure 2: The measured versus predicted values of Chl-a (a) and TSS (b) with Sentinel-2.

The performance of the Cubist model, listed in Table 6, shows good results. The performance in the training dataset is slightly better than on the validation, which can be evidence that the model does not overfit.

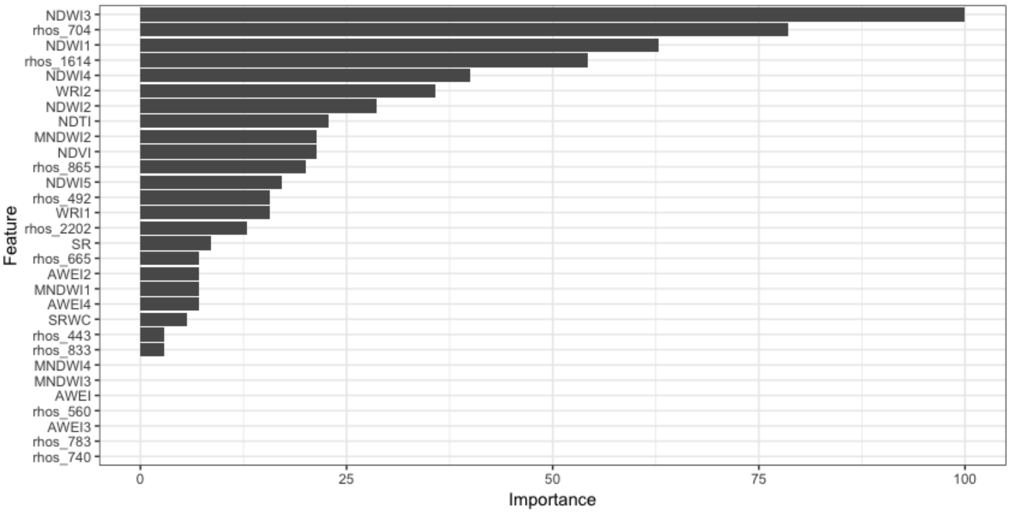
Table 6: Training statistics using Cubist for Chl-a and TSS

	training set		validation set		testing set	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
Chl-a	0.92	35.53	0.89	55.88	0.85	45.63
TSS	0.96	10.05	0.89	18.40	0.80	19.55

As mentioned earlier, Cubist easily can be interpreted due to the availability of variable importance in the final predictor model (Zhou et al., 2019). Therefore, the variable importance for all variable and co-variables showed in figure 3. It indicates that NDVI, SWRI and B5 are the top three most important variables in the dataset and B7 and B8 are the least essential variables Cubist utilized for predicting Chl-a. It also indicates that NDWI3, B5 and NDWI1 are the most variables Cubist algorithm used for predicting TSS.



(a)



(b)

Figure 3: Rank of features by importance for Chl-a (a) and TSS (b) based on Cubist algorithm.

Consequently, to better understand which spectral bands and spectral indices are the most significant drivers in the prediction of Chla and TSS using Sentinel-2A data, correlograms between variables and co-variables were built (Figure 4). It can be seen that the most correlated features with Chl-a were NDWI2, NDWI4, NDWI5, and NDVI, followed by B5 and SR. For TSS, which was successfully predicted using Sentinel-2A data, the highest correlation among the Sentinel-2A bands was B5, regarding the correlation between water spectral indices and TSS. The most correlated indices were NDWI2, NDWI4, NDWI5, NDWI3, and MNDWI4, followed by MNDWI1, MNDWI2, and MNDWI3.

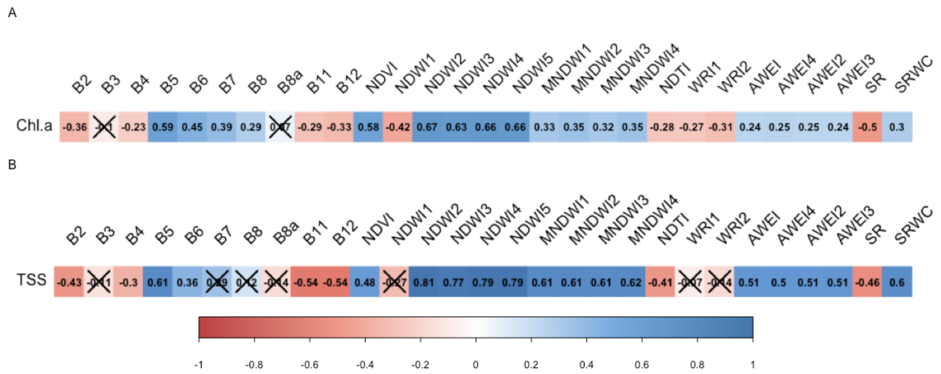


Figure 4: The correlograms of Chl-a (A) and TSS (B) at Sentinel-2 bands and calculated water indices (values in cells show correlation coefficients and crossed out cells indicate insignificant values at the 0.01 level).

3.3. *Spatial distribution of Chl-a and TSS and time series analysis*

The resulting spatial distribution maps of Chl-a and TSS developed through time derived observations from the Sentinel-2A are illustrated in Figure 5 and Figure 6 respectively.
Figure 5 shows that the maps displayed high and very high classes of Chl-a with higher mean values (Table 5), but Sentinel-2 failed to characterize the low level of Chl-a content in the study area. In general, according to the Chl-a map, Chl-a increased in August but decreased by the end of October. This trend is similar to all fishponds; however, for sand lakes, Chl-a did not change. This trend relatively was similar in both data collection years (i.e., 2017 and 2018).
According to the TSS spatial distribution maps (Figure 6), TSS reached its highest value by the end of August for fish ponds, but it decreased until the end of October. However, TSS remained relatively stable for sand lakes over time.

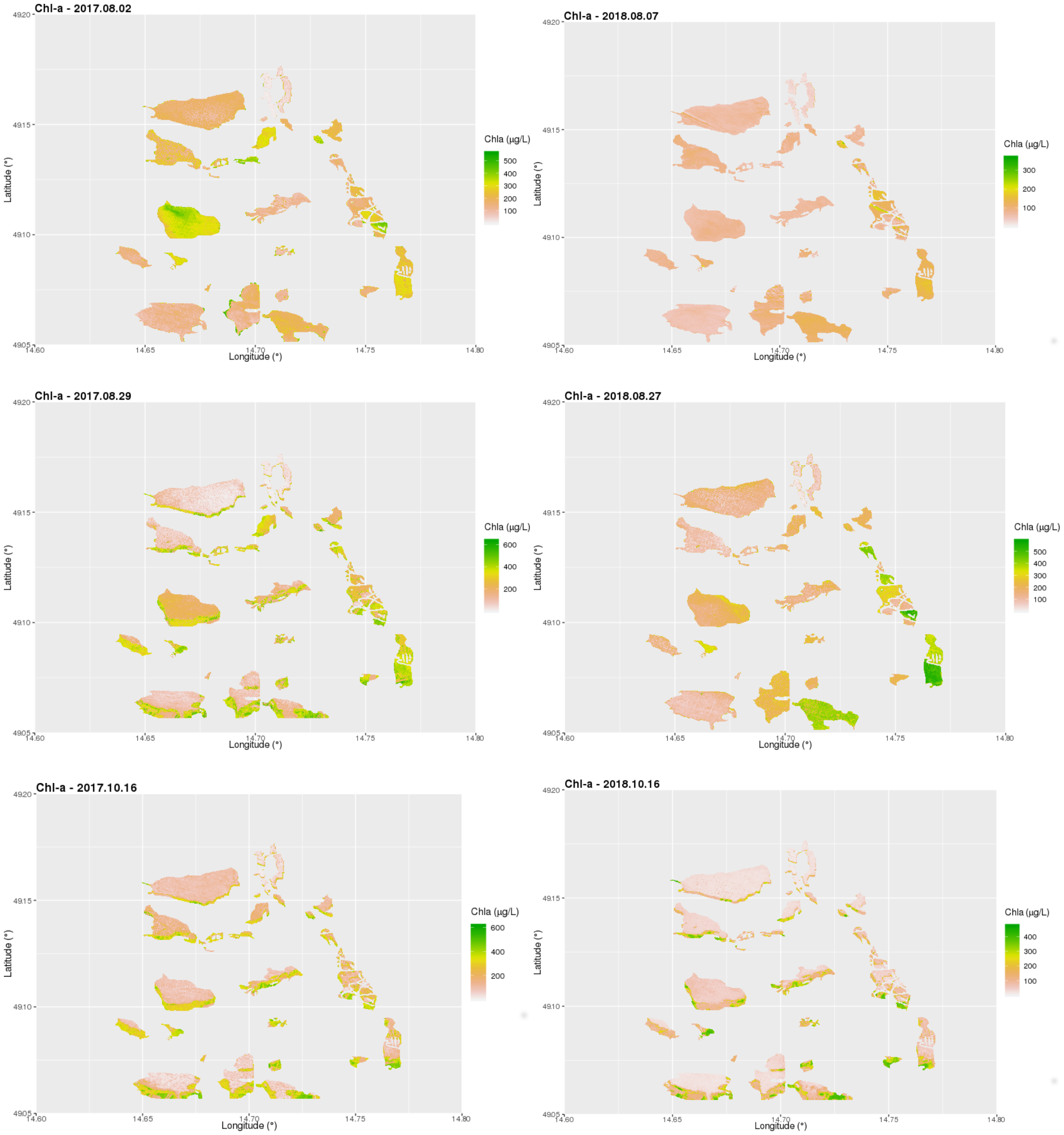
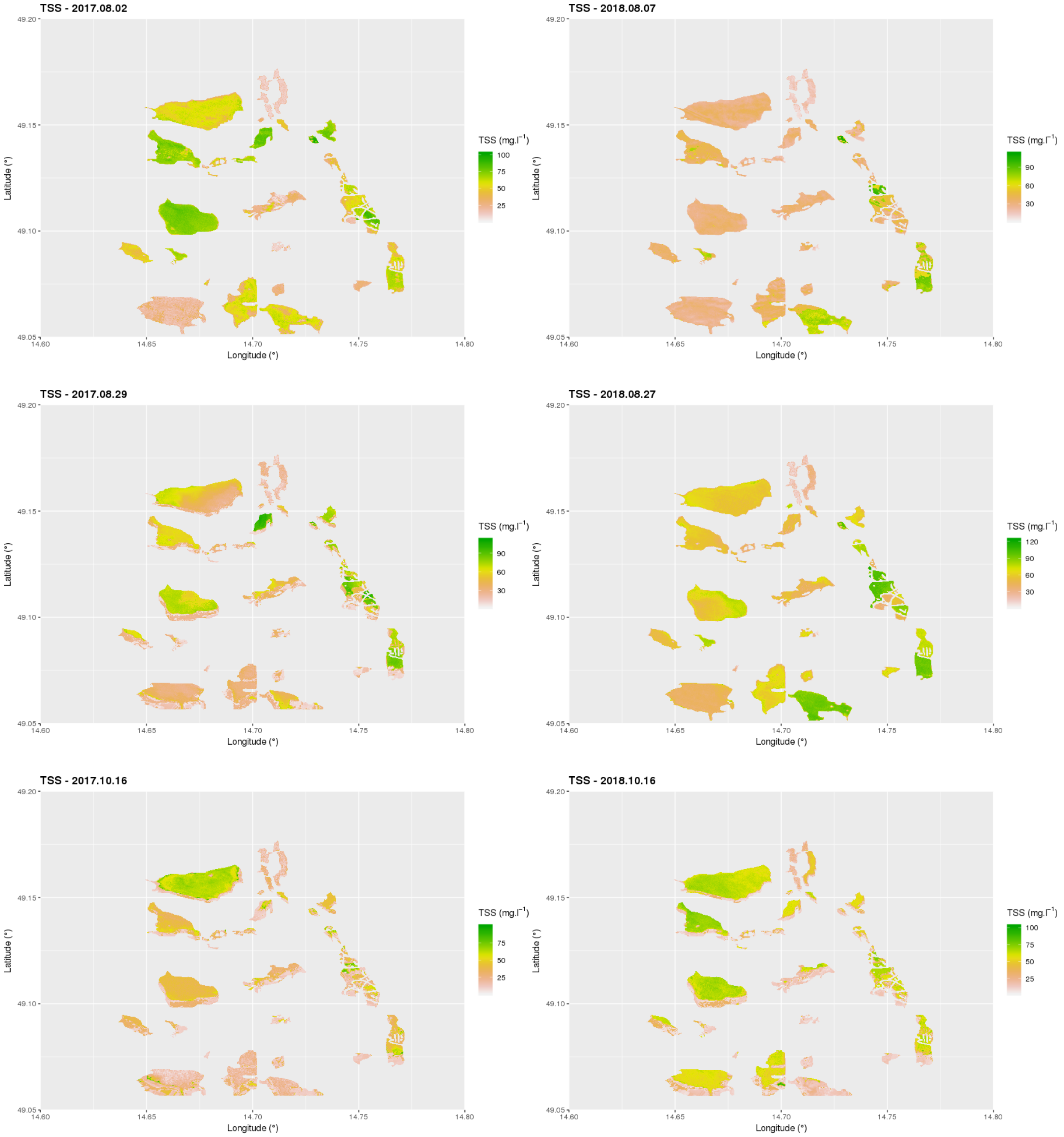


Figure 5: Distribution map of Chl-a over time.


TSS - 2017.08.02

TSS - 2018.08.07

TSS - 2017.08.29

TSS - 2018.08.27

TSS - 2017.10.16

TSS - 2018.10.16

Figure 6: Distribution map of TSS over time.

4. Discussion

The results of this study show that Sentinel-2A products can provide enough data to efficiently predict and visualize temporal and spatial Chl-a and TSS trends in small water bodies. Additionally, it showed that machine learning permits the prediction of Chl-a and TSS with significant accuracy based only on interactions between optical and water properties. In comparison to other studies, such as Toming et al. (2016), which predicted Chl-a for small bodies of inland water based on the band ratio calculated from BoA with 80% accuracy, machine learning managed to improve the accuracy of prediction. Machine learning generates a universal prediction algorithm, which allows better generalization due to utilizing all spectral bands and any number of band ratios. Although most previous studies (Song et al., 2012; Moridnejad et al., 2015; Chang et al., 2017) that used artificial neural network (ANN) to retrieve water quality parameters reported significant results, ANN requires a large dataset for training. It also necessitates an exceedingly long computation time; however, other machine learning methods such as Cubist could train the model with the smaller dataset and lower computation costs.

Considering the correlogram and performance of the extracted bands of Sentinel-2A and the calculated water spectral indices, the specific spectral band of B5 (698–713nm) provided the strongest correlations with both Chl-a and TSS. These results can be attributed mainly to the absorbance of red edge characteristics of vegetation (Gitelson et al., 1996).

The results in Figure 4 also indicate that the highest correlations for both Chl-a and TSS were provided from NDWI2, NDWI4, NDWI5, and NDVI, which represent a combination of Vis, NIR, and SWIR. Similar to what Grebdaute et al. (2018) reported, water indices, which are based on the combination of B4, B5, and B8A, can provide better results for retrieving Chl-a in inland waters using Sentinel-2A water surface reflectance. Similar to Chla, water surface reflectance in the NIR and Vis are sensitive to the TSS concentration. Furthermore, as Novoa et al. (2017) and Din et al. (2017) pointed out, water indices which have the SWIR partially contribute to successful TSS retrieval in high turbidity waters because they have been proven to be reliable for atmospheric correction of ACOLITE in SWIR bands.

Based on the Cubist model, we established that the spatial distribution of the concentration of both Chl-a and TSS for small water bodies can be easily generated. As expected, Chl-a and TSS were relatively higher during summer due to the growth of algal bloom cells. This trend was seen in fishponds; however, Chl-a and TSS concentrations remained low in sandpit lakes.

Regarding the spatial distribution of both Chl-a and TSS over time (Figure 5 and 6), the values of TSS dramatically changed in fishponds compared to in sandpit lakes, where the TSS values were more stable than the former. The reason for this may have arisen due to indifferent fishery management practices. While sandpit lakes are not managed, fishponds are controlled extensively. The mean fish stock is approximately 500 $kg.ha^{-1}$ (Pechar, 2000, 2015) in fishponds in the area of interest. The dominant fish is the benthivorous common carp (*Cyprinus carpio L.*). Carp digs in the bottom sediment while searching for food. As shown by Huser et al. (2016), common carp can disturb the bottom sediment at depths of up to 0.15 m. The result of the intensive bioturbation of the sediment by common carp is high water turbidity with a large amount of TSS in the water with enormous consequences to the water reservoir ecosystem (see, e.g., Zambrano et al. (2001)). In the case of sandpit lakes, the amount of TSS in water can be increased artificially by mining activities (sandpit lake Horusice) as well as recreation activities.

5. Limitations and Perspectives

Although the prediction accuracy of the introduced method is significant; it still needs to be improved. Besides, knowledge of the associated uncertainties related to water quality traits measurements and how to control the sources of errors are also crucial for small inland waters where bio-optical parameters are complex. To overcome these and similar uncertainties, a number of strategies can be recommended; For instance, the accuracy of the model can be improved *i.* by establishing a benchmark between field and satellite measurements in order to avoid mismatch in time scales between in situ and sensor overpass schedules,*ii.* by implying the spectral unmixing to decompose the optical water components which seems crucial for small inland waters (Alcantara et al., 2009), *iii.* by utilizing the super-resolution images in order to minimize the introduced bias due to conventional spatial resampling methods (Lanaras et al., 2018), and *iv.* by optimizing and applying other machine learning algorithms to reach better prediction accuracy. Additionally, as Pahlevan et al. (2019) demonstrated, there is consistency between Landsat-8 and Sentinel-2A/B for retrieving water biogeochemical properties. Thus, further studies should focus on investigating the application of machine learning methods for predicting water properties based on multi-mission surface reflectance.

As previously mentioned, machine learning algorithms are the better approach for handling the complex problems without prior knowledge, and they are less affected by the atmospheric and other background factors under non-ideal contexts. Therefore it can be assumed that the developed approach can be applied to other inland water within the same terrestrial and atmospheric condition; However, it still needs to be validated.

Conclusion

This study used a machine learning approach (i.e., Cubic) to retrieve two influential water quality properties for inland waters, i.e., Chl-a and TSS. To this end, concurrently to Sentinel-2A, several field campaigns were conducted to collect in situ data at several lakes in the south of the Czech Republic. As demonstrated, the enhanced spatial, spectral and temporal capabilities of Sentinel-2A permitted the prediction of biogeochemical properties accurately and inexpensively. Additionally, the machine learning algorithm was able to predict both Chl-a and TSS with significant accuracy in small lakes and ponds over time. This could be used as an alternative approach to commonly used methods such as physical models for predicting and mapping water quality parameters. The results of this study will support the trending idea that implementing data-driven methods (i.e., machine learning algorithms) for predicting water quality parameters improves the overall pipeline for predictive accuracy for complex spectral relationships and interactions. Nevertheless, future works are still essential to expand the knowledge on the other factors affecting the bio-optical parameters, efficient machine learning algorithms for retrieving the water quality parameters, and the associated uncertainties related to remote sensing of water quality traits.

Acknowledgements

This work was funded by the Ministry of Education, Youth and Sports of the Czech Republic project CENAKVA [LM2018099]; The CENAKVA Centre Development [No. CZ.1.05/2.1.00/19.0380]. This work was also supported by the grant of Technological Agency of the Czech Republic no. TG03010027 The Strengthening of Activities Proof-of-Concept at the University of South Bohemia.

Thanks belong to the staff of Laboratory of Applied Ecology, namely Martina Kobesová, Blanka Tesařová, Martin Musil, Libor Pechar and Iva Šímová and furthermore to Povodí Vltavy, State Enterprise for their help with water sampling and analysis of water samples.

References

Alcantara, E., Barbosa, C., Stech, J., Novo, E., & Shimabukuro, Y. (2009). Improving the spectral unmixing algorithm to map water turbidity distributions. *Environmental Modelling Software*, *24*, 1051 – 1061. doi:10.1016/j.envsoft.2009.02.013.

Ansper, A., & Alikas, K. (2018). Retrieval of Chlorophyll a from Sentinel-2 MSI Data for the European Union Water Framework Directive Reporting Purposes. *Remote Sensing*, *11*, 64. doi:10.3390/rs11010064.

Birth, G. S., & McVey, G. R. (1968). Measuring the Color of Growing Turf with a Reflectance Spectrophotometer. *Agronomy Journal*, *60*, 640–643. doi:10.2134/agronj1968.00021962006000060016x.

Boucher, J., Weathers, K. C., Norouzi, H., & Steele, B. (2018). Assessing the effectiveness of Landsat 8 chlorophyll a retrieval algorithms for regional freshwater monitoring. *Ecological Applications*, *28*, 1044–1054. doi:10.1002/eap.1708.

Bresciani, M., Stroppiana, D., Odermatt, D., Morabito, G., & Giardino, C. (2011). Assessing remotely sensed chlorophyll-a for the implementation of the Water Framework Directive in European perialpine lakes. *Science of The Total Environment*, *409*, 3083–3091. doi:10.1016/j.scitotenv.2011.05.001.

Caballero, I., Steinmetz, F., & Navarro, G. (2018). Evaluation of the first year of operational sentinel-2a data for retrieval of suspended solids in medium- to high-turbidity waters. *Remote Sensing*, *10*. doi:10.3390/rs10070982.

Carvalho, L., Poikane, S., Solheim, L. A., Phillips, G., Borics, G., Catalan, J., Hoyos, D. C., Drakare, S., Dudley, B., Jrvinen, M., Laplace-Treyture, C., Maileht, K., McDonald, C., Mischke, U., Moe, J., Morabito, G., Nges, P., Nges, T., Ott, I., Pasztaleniec, A., Skjelbred, B., & Thackeray, S. (2013). Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. *Hydrobiologia*, *704*, 127–140. doi:10.1007/s10750-012-1344-1.

Chang, N.-B., Bai, K., & Chen, C.-F. (2017). Integrating multisensor satellite data merging and image reconstruction in support of machine learning for better water quality management. *Journal of Environmental Management*, *201*, 227–240. doi:10.1016/j.jenvman.2017.06.045.

Chang, N.-B., Xuan, Z., & Yang, J. Y. (2013). Exploring spatiotemporal patterns of phosphorus concentrations in a coastal bay with MODIS images and machine learning models. *Remote Sensing of Environment*, *134*, 100–110. doi:10.1016/j.rse.2013.03.002.

Chebud, Y., Naja, G. M., Rivero, R. G., & Melesse, A. M. (2012). Water Quality Monitoring Using Remote Sensing and an Artificial Neural Network. *Water, Air, & Soil Pollution*, *223*, 4875–4887. doi:10.1007/s11270-012-1243-0.

Din, E. S. E., Zhang, Y., & Suliman, A. (2017). Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. *International Journal of Remote Sensing*, *38*, 1023–1042. doi:10.1080/01431161.2016.1275056.

Dörnhöfer, K., Klinger, P., Heege, T., & Oppelt, N. (2018). Multi-sensor satellite and in situ monitoring of phytoplankton development in a eutrophic-mesotrophic lake. *Science of The Total Environment*, *612*, 1200–1214. doi:10.1016/j.scitotenv.2017.08.219.

Dörnhöfer, K., & Oppelt, N. (2016). Remote sensing for lake research and monitoring Recent advances. *Ecological Indicators*, *64*, 105–122. doi:10.1016/j.ecolind.2015.12.009.

Drusch, M., Bello, D. U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. doi:10.1016/j.rse.2011.11.026.

Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water Bodies Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band. *Remote Sensing*, *8*, 354. doi:10.3390/rs8040354.

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. doi:10.1016/j.rse.2013.08.029.

Gao, B.-c. (1996). NDWIA normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*, 257–266. doi:10.1016/s0034-4257(96)00067-3.

Giardino, C., Kks, K.-L., Bolpagni, R., Luciani, G., Candiani, G., Lehmann, M. K., der Woerd, H. J. V., & Bresciani, M. (2019). The color of water from space: A case study for italian lakes from sentinel-2. In A. Pepe, & Q. Zhao (Eds.), *Geospatial Analyses of Earth Observation (EO) data* chapter 5. Rijeka: IntechOpen. URL: `https://doi.org/10.5772/intechopen.86596`. doi:10.5772/intechopen.86596.

Gilerson, A. A., Gitelson, A. A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., & Ahmed, S. A. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express*, *18*, 24109. doi:10.1364/oe.18.024109.

Gitelson, A. A., Merzlyak, M. N., & Lichtenthaler, H. K. (1996). Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. *Journal of Plant Physiology*, *148*, 501–508. doi:10.1016/s0176-1617(96)80285-9.

Gohin, F., der Zande, D. V., Tilstone, G., Eleveld, M. A., Lefebvre, A., Andrieux-Loyer, F., Blauw, A. N., Bryre, P., Devreker, D., Garnesson, P., Farias, T. H., Lamaury, Y., Lampert, L., Lavigne, H., Menet-Nedelec, F., Pardo, S., & Saulquin, B. (2019). Twenty years of satellite and in situ observations of surface chlorophyll-a from the northern bay of biscay to the eastern english channel. is the water quality improving? *Remote Sensing of Environment*, *233*, 111343. doi:https://doi.org/10.1016/j.rse.2019.111343.

Gower, J., King, S., Borstad, G., & Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the meris imaging spectrometer. *International Journal of Remote Sensing*, *26*, 2005–2012. doi:10.1080/01431160500075857.

Grendait, D., Stoneviius, E., Karosien, J., Savadova, K., & Kasperoviien, J. (2018). Chlorophyll-a concentration retrieval in eutrophic lakes in Lithuania from Sentinel-2 data. *Geologija. Geografija*, *4*. doi:10.6001/geol-geogr.v4i1.3720.

Guimar£es, T. T., Veronez, M. R., Koste, E. C., Souza, E. M., Brum, D., Gonzaga, L., & Mauad, F. F. (2019). Evaluation of Regression Analysis and Neural Networks to Predict Total Suspended Solids in Water Bodies from Unmanned Aerial Vehicle Images. *Sustainability*, *11*, 2580. doi:10.3390/su11092580.

Heblinski, J., Schmieder, K., Heege, T., Agyemang, T., Sayadyan, H., & Vardanyan, L. (2011). High-resolution satellite remote sensing of littoral vegetation of Lake Sevan (Armenia) as a basis for monitoring and assessment. *Hydrobiologia*, *661*, 97–111. doi:10.1007/s10750-010-0466-6.

Hunter, P. D., Tyler, A. N., Carvalho, L., Codd, G. A., & Maberly, S. C. (2010). Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sensing of Environment*, *114*, 2705–2718. doi:10.1016/j.rse.2010.06.006.

Huser, B. J., Bajer, P. G., Chizinski, C. J., & Sorensen, P. W. (2016). Effects of common carp (Cyprinus carpio) on sediment mixing depth and mobile phosphorus mass in the active sediment layer of a shallow lake. *Hydrobiologia*, *763*, 23–33. URL: `http://link.springer.com/10.1007/s10750-015-2356-4`. doi:10.1007/s10750-015-2356-4.

Kallio, K., Kutser, T., Hannonen, T., Koponen, S., Pulliainen, J., Vepslinen, J., & Pyhlahti, T. (2001). Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons. *Science of The Total Environment*, *268*, 59–77. doi:10.1016/s0048-9697(00)00685-9.

Keller, S., Maier, P. M., Riese, F. M., Norra, S., Holbach, A., Brsig, N., Wilhelms, A., Moldaenke, C., Zaake, A., & Hinz, S. (2018). Hyperspectral Data and Machine Learning for Estimating CDOM, Chlorophyll a, Diatoms, Green Algae and Turbidity. *International Journal of Environmental Research and Public Health*, *15*, 1881. doi:10.3390/ijerph15091881.

Koponen, S., Kallio, K., Pulliainen, J., Vepsalainen, J., Pyhalahti, T., & Hallikainen, M. (2004). Water quality classification of lakes using 250-m modis data. *IEEE Geoscience and Remote Sensing Letters*, *1*, 287–291. doi:10.1109/LGRS.2004.836786.

Kuhn, M. (2018). *caret: Classification and Regression Training*. URL: `https://CRAN.R-project.org/package=caret` r package version 6.0-81.

Kuhn, M., & Quinlan, R. (2018). *Cubist: Rule- And Instance-Based Regression Modeling*. URL: `https://CRAN.R-project.org/package=Cubist` r package version 0.2.2.

Kutser, T., Paavel, B., Kaljurand, K., Ligi, M., & Randla, M. (2018). Mapping shallow waters of the baltic sea with sentinel-2 imagery. In *2018 IEEE/OES Baltic International Symposium (BALTIC)* (pp. 1–6). doi:10.1109/BALTIC.2018.8634850.

Lacaux, J., Tourre, Y., Vignolles, C., Ndione, J., & 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. doi:10.1016/j.rse.2006.07.012.

Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., & Schindler, K. (2018). Super-resolution of sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 305 – 319. doi:10.1016/j.isprsjprs.2018.09.018.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., & Lu, H. (2009). A four-band semi-analytical model for estimating chlorophyll a in highly turbid lakes: The case of Taihu Lake, China. *Remote Sensing of Environment*, *113*, 1175–1182. doi:10.1016/j.rse.2009.02.005.

Mas, J. F., & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, *29*, 617–663. doi:10.1080/01431160701352154.

Matarrese, R., Morea, A., Tijani, K., De Pasquale, V., Chiaradia, M. T., & Pasquariello, G. (2008). A specialized support vector machine for coastal water chlorophyll retrieval from water leaving reflectances. In *IGARSS 2008 - IEEE International Geoscience and Remote Sensing Symposium* (pp. IV – 910–IV – 913). volume 4. doi:10.1109/IGARSS.2008.4779871.

Matsushita, B., Yang, W., Jaelani, L., F, & Fukushima, T. (2016). Chapter 9 Monitoring water quality with remote sensing image data. *Boca Raton: CRC Press*, . doi:10.1201/9781315371931-10.

McCullough, I. M., Loftin, C. S., & Sader, S. A. (2012). High-frequency remote monitoring of large lakes with modis 500m imagery. *Remote Sensing of Environment*, *124*, 234 – 241. doi:https://doi.org/10.1016/j.rse.2012.05.018.

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. doi:10.1080/01431169608948714.

Moridnejad, A., Abdollahi, H., Alavipanah, S. K., Samani, J. M. V., Moridnejad, O., & Karimi, N. (2015). Applying artificial neural networks to estimate suspended sediment concentrations along the southern coast of the Caspian Sea using MODIS images. *Arabian Journal of Geosciences*, *8*, 891–901. doi:10.1007/s12517-013-1171-3.

Moses, W. J., Gitelson, A. A., Berdnikov, S., & Povazhnyy, V. (2009). Satellite Estimation of Chlorophyll-a Concentration Using the Red and NIR Bands of MERISThe Azov Sea Case Study. *IEEE Geoscience and Remote Sensing Letters*, *6*, 845–849. doi:10.1109/lgrs.2009.2026657.

Moss, T. (2012). Spatial fit, from panacea to practice: implementing the eu water framework directive. *Ecology and Society*, *17*, 2. doi:http://dx.doi.org/10.5751/ES-04821-170302.

Mouw, C. B., Greb, S., Aurin, D., DiGiacomo, P. M., Lee, Z., Twardowski, M., Binding, C., Hu, C., Ma, R., Moore, T., Moses, W., & Craig, S. E. (2015). Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. *Remote Sensing of Environment*, *160*, 15–30. doi:10.1016/j.rse.2015.02.001.

Mukherjee, N., & Samuel, C. (2016). Assessment of the Temporal Variations of Surface Water Bodies in and around Chennai using Landsat Imagery. *Indian Journal of Science and Technology*, *9*. doi:10.17485/ijst/2016/v9i18/92089.

Novoa, S., Doxaran, D., Ody, A., Vanhellemont, Q., Lafon, V., Lubac, B., & Gernez, P. (2017). Atmospheric Corrections and Multi-Conditional Algorithm for Multi-Sensor Remote Sensing of Suspended Particulate Matter in Low-to-High Turbidity Levels Coastal Waters. *Remote Sensing*, *9*, 61. doi:10.3390/rs9010061.

Pahlevan, N., Chittimalli, S. K., Balasubramanian, S. V., & Vellucci, V. (2019). Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems. *Remote Sensing of Environment*, *220*, 19–29. doi:10.1016/j.rse.2018.10.027.

Pahlevan, N., Sarkar, S., Franz, B., Balasubramanian, S., & He, J. (2017). Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. *Remote Sensing of Environment*, *201*. doi:10.1016/j.rse.2017.08.033.

Pechar, L. (1987). Use of an acetone: methanol mixture for the extraction and spectrophotometric determination of chlorophyll-a in phytoplankton. *Algological Studies/Archiv für Hydrobiologie, Supplement Volumes*, *46*, 99–117.

Pechar, L. (2000). Impacts of long-term changes in fishery management on the trophic level water quality in Czech fish ponds. *Fisheries Management and Ecology*, *7*, 23–31. doi:10.1046/j.1365-2400.2000.00193.x.

Pechar, L. (2015). Stolet eutrofizace rybnkÅsynergick efekt zvyÅovn ztÅe Åivinami (fosforem a duskem) a nrÅstu rybch obsdek. *Vodn hospodÅstv*, *65*, 1–6.

Pepe, M., Giardino, C., Borsani, G., Cardoso, A., Chiaudani, G., Premazzi, G., Rodari, E., & Zilioli, E. (2001). Relationship between apparent optical properties and photosynthetic pigments in the sub-alpine Lake Iseo. *Science of The Total Environment*, *268*, 31–45. doi:10.1016/s0048-9697(00)00691-4.

Peterson, K. T., Sagan, V., Sidike, P., Cox, A. L., & Martinez, M. (2018). Suspended Sediment Concentration Estimation from Landsat Imagery along the Lower Missouri and Middle Mississippi Rivers Using an Extreme Learning Machine. *Remote Sensing*, *10*, 1503. doi:10.3390/rs10101503.

Philipson, P., Kratzer, S., Mustapha, S. B., Strmbeck, N., & Stelzer, K. (2016). Satellite-based water quality monitoring in Lake Vnern, Sweden. *International Journal of Remote Sensing*, *37*, 3938–3960. doi:10.1080/01431161.2016.1204480.

Quinlan, J. R. (1992). Learning with continuous classes. In *Proceedings of Australian Joint Conference on Artificial Intelligence* (pp. 343–348). World Scientific.

Rossel, V. R., & Webster, R. (2012). Predicting soil properties from the Australian soil visiblenear infrared spectroscopic database. *European Journal of Soil Science*, *63*, 848–860. doi:10.1111/j.1365-2389.2012.01495.x.

Rößler, S., Wolf, P., Schneider, T., Zimmermann, S., & Melzer, A. (2013). Water constituent retrieval and littoral bottom mapping using hyperspectral apex imagery and submersed artificial surfaces. *EARseL EProceedings*, *12*, 44–57.

Rouse, J. H., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the great plains with erts. *Proceedings, Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, 1974*, (pp. 3010–3017).

Roy, D. P., Li, J., Zhang, H. K., & Yan, L. (2016). Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data. *Remote Sensing Letters*, *7*, 1023–1032. doi:10.1080/2150704x.2016.1212419.

Ruddick, K., Vanhellemont, Q., Dogliotti, A., Nechad, B., Pringle, N., & Van der Zande, D. (2016). New opportunities and challenges for high resolution remote sensing of water colour. In *The Ocean Optics XXIII* (pp. 23–28).

Song, K., Li, L., Wang, Z., Liu, D., Zhang, B., Xu, J., Du, J., Li, L., Li, S., & Wang, Y. (2012). Retrieval of total suspended matter (TSM) and chlorophyll-a (Chl-a) concentration from remote-sensing data for drinking water resources. *Environmental Monitoring and Assessment*, *184*, 1449–1470. doi:10.1007/s10661-011-2053-3.

Sudheer, K., Chaubey, I., & Garg, V. (2006). Lake water quality assessment from landsat thematic mapper data Using neural network: an approach to optimal band Combination selection1. *Journal of the American Water Resources*

Association, *42*, 1683–1695. doi:10.1111/j.1752-1688.2006.tb06029.x.

Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., & Nges, T. (2016). First Experiences in Mapping Lake Water Quality Parameters with Sentinel-2 MSI Imagery. *Remote Sensing*, *8*. doi:10.3390/rs8080640.

Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., Dillon, P., Finlay, K., Fortino, K., Knoll, L. B., Kortelainen, P. L., Kutser, T., Larsen, S., Laurion, I., Leech, D. M., McCallister, L. S., McKnight, D. M., Melack, J. M., Overholt, E., Porter, J. A., Prairie, Y., Renwick, W. H., Roland, F., Sherman, B. S., Schindler, D. W., Sobek, S., Tremblay, A., Vanni, M. J., Verschoor, A. M., Wachenfeldt, E. v., & Weyhenmeyer, G. A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, *54*, 2298–2314. doi:10.4319/lo.2009.54.6_part_2.2298.

Tyler, A. N., Hunter, P. D., Spyrakos, E., Groom, S., Constantinescu, A. M., & Kitchen, J. (2016). Developments in Earth observation for the assessment and monitoring of inland, transitional, coastal and shelf-sea waters. *Science of The Total Environment*, *572*, 1307–1321. doi:10.1016/j.scitotenv.2016.01.020.

Vanhellemont, Q., & Ruddick, K. (2018). Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. *Remote Sensing of Environment*, *216*, 586–597. doi:10.1016/j.rse.2018.07.015.

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. doi:10.1080/01431160600589179.

Zambrano, L., Scheffer, M., & Martinez-Ramos, M. (2001). Catastrophic response of lakes to benthivorous fish introduction. *Oikos*, *94*, 344–350. doi:10.1034/j.1600-0706.2001.940215.x.

Zarco-Tejada, P. J., & Ustin, S. L. (2001). Modeling canopy water content for carbon estimates from modis data at land eos validation sites. In *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)* (pp. 342–344 vol.1). volume 1. doi:10.1109/IGARSS.2001.976152.

Zhou, J., Li, E., Wei, H., Li, C., Qiao, Q., & Armaghani, D. J. (2019). Random forests and cubist algorithms for predicting shear strengths of rockfill materials. *Applied Sciences*, *9*. doi:10.3390/app9081621.