International Journal of Applied Earth Observation and Geoinformation 51 (2016) 28–36

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Application of Landsat 8 imagery to regional-scale assessment of lake water quality

Jacek Andrzej Urbanski[a,*], Agnieszka Wochna[a], Iwona Bubak[b], Waldemar Grzybowski[b], Katarzyna Lukawska-Matuszewska[b], Magda Łącka[d], Sylwia Śliwińska[b], Bożena Wojtasiewicz[b,c], Marek Zajączkowski[d]

[a] GIS Centre, University of Gdansk, 80-952 Gdansk, Poland
[b] Institute of Oceanography, University of Gdansk, 81-378 Gdynia, Poland
[c] CSIRO Oceans & Atmosphere, Floreat, WA, Australia
[d] Institute of Oceanology Polish Academy of Sciences, 81-712 Sopot, Poland

ARTICLE INFO

Article history:
Received 29 November 2015
Received in revised form 6 April 2016
Accepted 18 April 2016
Available online 4 May 2016

Keywords:
Landsat 8
Lakes
Trophic status
GIS
Water quality

ABSTRACT

The aim of the project was to create a tool with which to support regional lake quality assessment using Landsat 8 imagery data. The model of assigning the ecological status was implemented in GIS for the northern part of Poland and classifies lake quality for several classes according to classification of WFD using two basic assumptions. The first is that there exists a combination of OLI bands (OLI2/OLI4 was used) which correlates well with the trophic state of the lakes; the second assumption is that the reference trophic state depends on the mean depth of the lake. The model uses a lake geodatabase which contains lakes outlines, raster masks of lakes and attribute information about their mean depth. There is no need to provide any field data when using this tool, as calibration of the model is done using subsets of lakes which were classified using legally defined methods. The tool allows fast classification of 2800 lakes from the area of interest. The results show good agreement between satellite and expert based methods.

1. Introduction

Lake water quality is of a global interest. Indeed, the high or bad ecological state of rivers, lakes and coastal waters can strongly influence the values of properties along their coasts (McCullough et al., 2012). In nearly every country it is the responsibility of local administration to care for the health of the water bodies. In the EU, the WFD (Water Framework Directive) obligates all member countries to achieve good ecological quality of all inland water bodies until 2015 (Peeters et al., 2009). These demands require the establishment of complex monitoring and assessment systems. They are traditionally based on water sampling at fixed stations and laboratory analyses. This method may be difficult to implement when there are hundreds of separate water bodies to monitor (Kalio, 2000). There are several spots in the northern hemisphere where the lakes are clustered and their number exceeds the thousands. Such places may be found in Canada, Northern US, Sweden, Finland, Poland and Baltic Countries. They have a common origin connected with the end of the last glaciation. They are surrounded by a mixture of land cover and use. Some lake watersheds constitute a nearly unchanged environment, while others are strongly affected by agricultural runoff, erosion, urban development and loss of wetland, all of which have a strong impact on lake health (Torbick et al., 2013). This is the main reason why the trophic status of lakes incorporating relationships between nutrients (mainly phosphorus), phytoplankton and transparency may vary from ultra-oligotrophic to hyper-eutrophic (Solimini et al., 2006). On the other hand however, the trophic status of lakes with minimal human influence may also vary in a natural way depending on the lake morphometry, and especially with regard the mean depth of the water body (Scheffer, 1998).

There are many parameters used in lake monitoring which typically include measures of chlorophyll-a (Chl-a), suspended material (SPM), light attenuation for which the common proxy is Secchi depth (SD) and colored dissolved organic matter (CDOM). The WFD requirements are extremely demanding and aim to assess five qual-

* Corresponding author.
E-mail address: cgisju@ug.edu.pl (J. Andrzej Urbanski).

http://dx.doi.org/10.1016/j.jag.2016.04.004
0303-2434/

ity classes (Bad, Poor, Moderate, Good, High) by comparing the current status to reference state using biological, physical as well chemical indicators supplemented by expert knowledge. However, some publications suggest that this may be predicted quite well using water transparency only expressed by Secchi depth (Peeters et al., 2009).

The remote sensing monitoring of lakes primarily uses satellite data, mainly from Landsat TM and TM+ sensors which have the perfect spatial and time resolution for such an aim. Indeed, there are many examples of when this tool has been used successfully, and these can be divided into two groups of applications. The first group uses a singular or multiple images of the area of interest and field data obtained from several lakes (Torbick et al., 2013; Tebbs et al., 2013; Duan et al., 2008; Sass et al., 2007; Brezonik et al., 2005). It has been suggested in several publications that the time gap between lake data collection and satellite data should be no longer then 3–10 days (Olmanson et al., 2008; McCullough et al., 2012). In most cases the band ratio multi regression approach is used to obtain the relation between satellite signal and in-situ measurements. Both raw signal (DN), radiance and reflectance have been used. Several methods of less or more advanced atmospheric corrections were applied and discussed, although there are suggestions that simple dark object subtraction is recommended for classification and change detection applications (Song et al., 2001). The results of these projects are, general speaking, similar. They have proved that it is possible to obtain good results for the mapping of SD and chlorophyll-a, as well as weaker results for SPM and CDOM, cyanobacteria or diatoms. In order to obtain acceptable results, the regression equations have to be calibrated separately for a particular image. The second group used monitoring programs and the community approach to collect Secchi depth data from many lakes over several years (McCullough et al., 2012; Olmanson et al., 2008; Kloiber et al., 2002). They also used the regression approach and conducted additional analyses including change detection.

The use of satellite images for lake monitoring, especially in places with hundreds or thousands of lakes, is very tempting. Indeed, recent years have seen the emergence of two important facts. The first is there are now new satellites platforms whose resolution is suitable for lake monitoring, and the second is that their usage is free of charge. Using both Landsat 8 and, in the near future, Sentinel 2, it will be possible to obtain at least one uncloudy image nearly every month. This will give environmental managers a chance to use all benefits of RS technology such as simultaneously monitoring many lakes.

The aim of the project was to create a GIS tool to support regional lake quality assessment. The tool should classify lake quality for classes according to classification used by WFD. We assumed that the tool will require only Landsat 8 image which covers a few dozen of lakes, the water mask of the lakes and its mean depth. The tool will not require any field data which will make them more flexible and operational. To fulfill these aims three research problems have to be solved. First, the best method has to be found out for estimation of crucial parameters used in remote sensing lake monitoring. Then, such a model has to be created which allows to compare the current ecological status of the lakes using the results of satellite image analyses. Finally it has to be found out how to calibrate the results of this comparison using the lakes with known current state or defined as reference state lakes. The mean depth of the lake plays an important role. The postglacial lakes may have different depths on which their ecological state in natural conditions is heavily dependent. There is a well-established difference in the structure and functioning of the lakes which depends on their mean depth (Scheffer, 1998; Poikane et al., 2014). The natural ecological state is different in polymictic shallow systems than in deeper seasonally stratified systems. As a result, and in order to minimize natural biological variations, lakes are often divided into arbitrary depth related classes (Phillips et al., 2008). The tool should make it possible to calibrate the classification process in an iterative way so as the subset of lakes whose quality state is known may be used to calibrate the complete set of lakes.

2. Data and methods

The development of the system was carried out in two steps. First the two field experiments were conducted using two Landsat 8 scenes and field observations carried out at nearly the same time as the images were taken. The aim of these experiments was to find out the best combination of satellite image bands for describing the trophic state of the lakes in the area of interest and to prove that our assumption about the role of mean depth is true. On the basis of the results from these two experiments the model of assigning the ecological status using only satellite data was developed and verified using independent expert assessments.

2.1. Description of study area

The study area covers the northern part of Poland, where there are many densely distributed clusters of postglacial lakes. There are approximately 8000 lakes with a total surface of around 3000 km^2 across, and a total area of approximately 116000 km^2 (Fig. 1B). The mean depth of the lakes varies from 0.2–38.7 m. The land cover of this area is a mosaic of agricultural areas (61%), forests and semi natural areas (34%) as well as urban and suburban fabric (2%). The field experiment area (Fig. 1) is placed in the central area of study which is representative of climate and land cover. In Poland, the ecological state of the lakes (required by WFD) is determined by the regional inspectorates of environmental protection which are the branches of the government administration. They use several legally defined procedures in order to divide the ecological status of lakes into five classes using mainly biological based methods supplemented by some abiotic measurements. These methods are expensive and need a lot of laboratory work. This means that a limited number of lakes may be classified each year. For example, only 22 lakes were assessed in 2014 in the Pomeranian District (Fig. 1).

2.2. Lake geodatabase

The geodatabase of lakes for the northern part of Poland was created using three sources: A vector layer from polish hydrographic geodatabase MPHP in the form of polygons (SHP) containing lake names; an analogous atlas of lakes (Janczak, 1996) which contains tables with positions of the central points of the lakes, and their mean and maximum depth; and Landsat 8 images. First the maximum depth from atlas tables was spatially joined with shapes of the MPHP lakes. Following this, the number of lakes was updated as the MPHP geodatabase contains many small lakes which are not currently in existence and which are now covered by vegetation. This was done using Landsat 8 images for the whole area of northern Poland. The area (Fig. 1B) was covered by twelve uncloudy images (paths: 187–192 and rows: 22–24) from the years 2013–2015. The shortwave infrared band (6th band DN) of Landsat 8, which gives the strongest and most stable difference between water and land, and the median statistics were used to find a threshold to select existing lakes only. The same set of images was used to create a raster of lake zones for extracting pixel values for statistics calculation. It was decided that for each lake only pixels with values lower than the median of all lake pixels in band 6 would create a zone with unique lake identifiers. As a result, our geodatabase contains 2800 lakes with the same mean and maximum depth as a vector

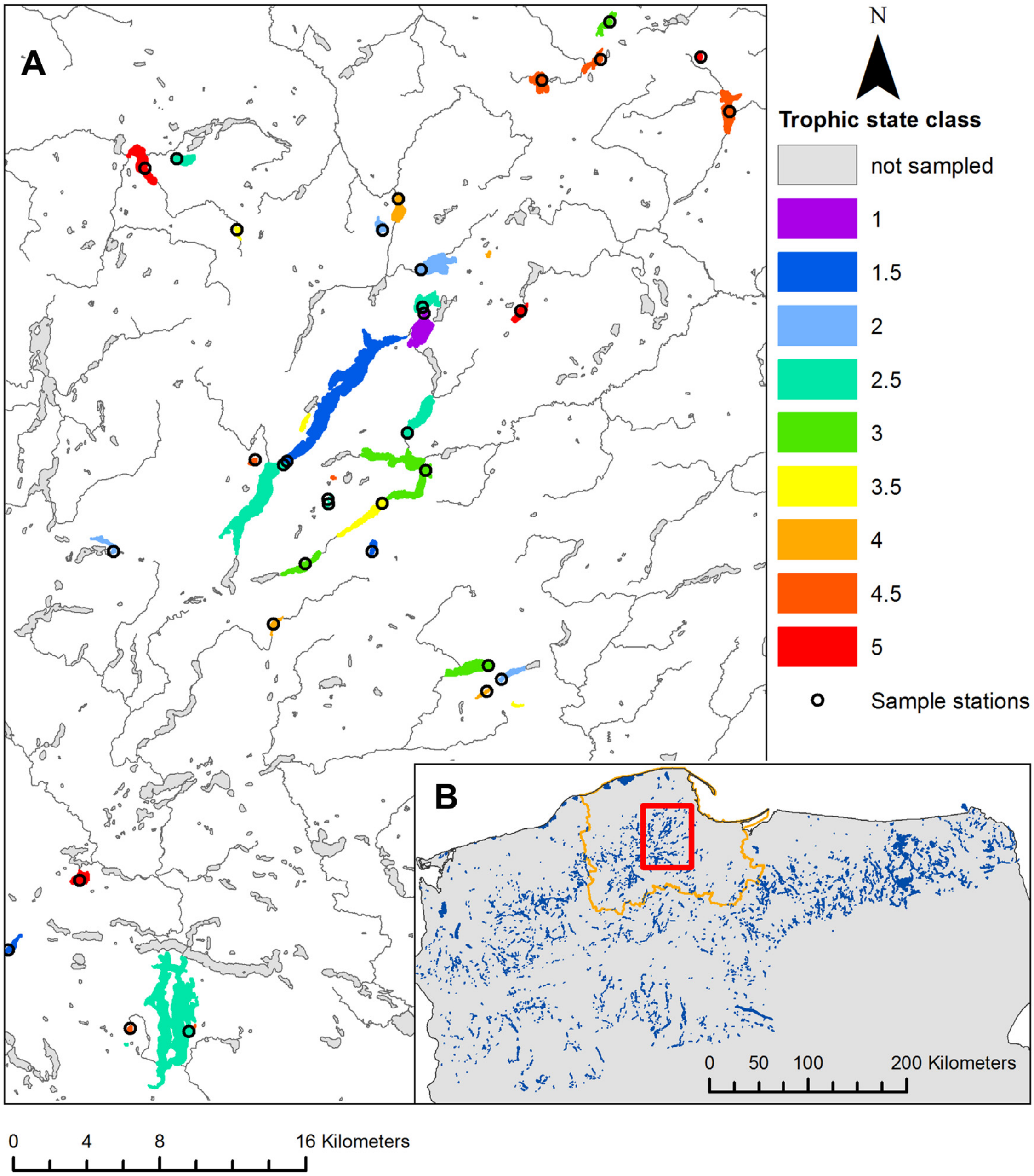

Fig. 1. Study area. A—Landsat 8–field experiment area with trophic state classes of sampled lakes determined during experiment and sample stations; B—area covered by Lake Geodatabase and model of assessments of lake water quality (red rectangle shows position of field experiment area, while orange outline is a border of Pomerania District).

polygon layer and a lake sampling zone raster layer; both have a unique lake identifier (LID).

2.3. Field experiment

The satellite part of the field experiments was carried out on 5 June (Scene 1) and 7 July (Scene 2) 2014. The images were partly cloudy but free from haze and cirrus contamination. First, on the same day as the Landsat 8 images were taken, the scenes were converted to natural color composite and visually inspected to select lakes free of clouds with colors representing the whole spectrum of lake surface colors in the image. For the sake of quality regression analyses it was important to obtain points with values of bands equally distributed across the whole range of extremal values. The sampled lakes with trophic state classes were determined through analysis, as described in Section 2.5 (please see Fig. 1A). The field measurements were carried out on 6–8 June and 11–13 July, meaning that all samples were collected no later than 6 days after the satellite images were taken. The field measurements included: DGPS position, Secchi depth (SD), chlorophyll-a (Chl-a), suspended particulate matter (SPM), colored dissolved organic matter (CDOM) and dissolved organic carbon (DOC). The composition of the phytoplankton species and its biomass were also determined in order to assess the seasonal difference in the development of various planktonic algal groups between both scenes. Secchi depth was measured using a standard procedure with a 20 cm diameter white disk attached to a calibrated depth line. In order to measure chlorophyll *a* concentration, the water samples from each lake were passed through Whatman GF/F filters (25 mm diameter) under low suction pressure in order to measure chlorophyll *a* concen-

Table 1
Modified lake classification scheme (with ranks 1–5) based on trophic state. Ranks 1–5 correspond to Oligotrophic – Eutrophic state. In two bottom rows a number of lakes in each class is presented for field experiment.

	Chl-a [μgl^{-1}]				
SD [m]	<=2.2	>2.2 and <=6.0	>6.0 and <= 22.0	>22.0 and <= 40.0	>40.0
>=4.6	1	1.5			
<4.6 and >=2.3	1.5	2	2.5		
<2.3 and >=0.9		2.5	3	3.5	
<0.9 and >=0.4			3.5	4	4.5
<0.4				4.5	5

Class	1	1.5	2	2.5	3	3.5	4	4.5	5
Number of lakes	1	3	4	8	4	4	4	7	4

tration [$\mu g\,l^{-1}$] no more than a few hours after collecting. Filters were kept in a liquid nitrogen until analysis. After thawing, pigments were extracted with the use of 96% ethanol in darkness at room temperature for 24 h. Then samples were centrifuged for 15 min at 4000 rpm. The absorbance spectra of the extracts were determined using a Perkin Elmer Lambda 850 dual-beam spectrophotometer in a 1-cm cuvette against a 96% ethanol blank. The chlorophyll *a* concentration was calculated through the use of the HELCOM formula (HELCOM, 1988). The chlorophyll *a* concentration taken into further analysis was calculated as the average value from three aliquots. Suspended particulate matter (SPM; mg/dm^3) was assessed through vacuum-filtering of 1.8 dm^2 water samples onto pre-combusted and pre-weighted fibre glass MN GF-5 filters (mesh size 0.4 μm). Large organisms visible to the naked eye were removed from the filters. Each filter was then air dried at 60 °C for 24 h and weighed to determine total suspension dry mass collected on filters (Zajączkowski et al., 2010). For concentrations of DOC and CDOM, water from lakes was collected in acid-cleaned plastic bottles before being transferred to the laboratory within 8 h and filtered at low pressure through pre-combusted Whatman GF/F filters. The concentration of dissolved organic carbon (DOC) was measured with the high temperature catalytic oxidation (HTCO) technique using a Vario TOC Cube analyzer (Elementar Analysensysteme GmbH).). After filtration, samples were acidified to pH < 2 with concentrated HCl and stored in a refrigerator until analysis. 50 ml glass stopper bottles for sample storage were combusted for 5 h at 450 °C. The precision (RSD) of DOC analyses was not worse than 3.5%. Absorbance of the filtered water samples was used as a proxy for CDOM concentration. Absorption spectra were recorded between 250 nm and 700 nm at 1 nm intervals using a Jasco V-630 dual-beam spectrophotometer with matching 100 mm quartz cells (10 mm long cell were used when absorbance values exceeded 2.5). Milli-Q water was used in the reference cell.

Water samples for phytoplankton species and composition were immediately preserved in Lugol's solution. The taxonomic composition and number of phytoplankton were analyzed under a NIKON inverted microscope. Phytoplankton organisms were identified at the species level or, if not possible, assigned to a genus only. Taxons were identified using keys and world literature. For each sample, at least 200 cells were counted with a subset measured to characterize the size range. Based on cell or colony measurements, algal biovolume was estimated. After calculating the average volume of every species, total volume was calculated by multiplying biovolume by the number of species. All measurements were saved as attributes of vector point feature class in the GIS geodatabase (ArcMap 10.2). Each point (one sampling point per lake) was surrounded by polygons for extraction of statistics of band values from the raster image. The surfaces of the polygons varied from 0.28 to 3.6 ha (mean value = 0.92 ha) depending on the size of the lake and the location of the sampling point.

2.4. Preprocessing of satellite data

Both Landsat 8 scenes (path 190, row 22) were obtained from USGS Glovis and processed in the same way. The DN was first converted to TOA reflectance with correction for the sun angle according to USGS Landsat 8 product instruction (USGS, 2015) using Formula (1). Following this, the atmospheric correction using the dark object subtraction (DOS) method was carried out assuming 1% surface reflectance from the dark objects (Chavez, 1988) according to Formulae (2) and (3). The dark objects threshold value is defined as minimum value in particular band. The minimum values are most often localized in places with dark land cover (forest) covered by cloud shadows or on the surface of small dystrophic or humic lakes in the forest in cloudless situation. The atmospheric correction depends on the band. It changes the blue band (OLI2) ten times more than red band (OLI4), for which the correction is usually insignificant. In result the classification has relative character defined by mean and standard deviation of index value. The role of some uncertainty of atmospheric correction is not crucial. Indeed lack of the correction changes only about 30% of assigning classes by ± 1 class and only 1% by ± 2 classes. No bias in assigning is observed.

$$R_{TOA} = (DN \times 0.00002 - 0.1)\,/\,\cos\theta_s \tag{1}$$

$$R_{scatter} = (DN_{minimum} \times 0.00002 - 0.1)\,/\,\cos\theta_s \tag{2}$$

$$R_{earth} = R_{TOA} - (R_{scatter} - 0.01) \tag{3}$$

where:R_{TOA} – reflectance at the top of atmosphere DN – digital number. θ_s – solar zenith. $R_{scatter}$ – reflectance scatter in atmosphere. $DN_{minimum}$ – minimum digital number. R_{earth} – reflectance at earth surface.

For bands 1–5, the statistics of corrected reflectance were calculated from sets of pixels extracted by polygons surrounding the sample points for each lake

2.5. Lake classification

It was essential to test the relation between spectral signal of a lake surface and trophic state of lake. In order to achieve this, a modified OECD (Vollenweider and Kerekes, 1982) classification depending on chl-a and Secchi disk was proposed. Each lake included in the field experiment was assigned a rank of 1–5, corresponding with the oligo-to eutrophic state respectively (Fig. 1A). Ranges of Secchi disk and chlorophyl-a concentration together with assigning classes are presented in Table 1. Lakes selected for sam-

Table 2
OLI variables candidates for building regression equations with field measurements.

Variable	OLI band combination	OLI band combination—description
VAR1	OLI2	Blue
VAR2	OLI3	Green
VAR3	OLI4	Red
VAR4	OLI5	NIR
VAR5	OLI2/OLI3	Blue/Green
VAR6	OLI2/OLI4	Blue/Red
VAR7	OLI3/OLI4	Green/Red
VAR8	OLI2/OLI5	Blue/NIR
VAR9	OLI3/OLI5	Green/NIR
VAR10	(OLI2–OLI4)/(OLI2 + OLI4)	(Blue–Red)/(Blue + Red)
VAR11	(OLI2–OLI3)/(OLI2 + OLI3)	(Blue–Green)/(Blue + Green)
VAR12	OLI3–OLI2	Green–Blue
VAR13	OLI3–OLI4	Green–Red

pling are shown in Fig. 1A, with trophic state assigned on the basis of Secchi disk and Chl-a concentration according to this table.

There is a common methodology used when looking at the best relation between satellite image information and measured values of parameters of water, including trophic state class. Indeed, the aforementioned methodology includes testing different variables (simple function of satellite image bands) using Ordinary Least Squares (OLS) and independent variable transformation. Based on previous research (McCullough et al., 2012; Duan et al., 2008; Sass et al., 2007; Kloiber et al., 2002) several potentially useful Landsat OLI sensor wavelengths were chosen. All tabled (Table 2) variables were used with natural logarithms from all measured water parameters in the OLS model and judged with adjusted R-squared values, the basic metric of OLS performance. Additionally, to improve the method of determining the best combination of bands, we used the significance of each variable defined as the number of times it was statistically significant with its stability. For all models in the table, the variable VAR6 (Blue/Red) represents the greatest significance, with extremely good stability compared to other variables; as such, it was used to calculate the trophic state index.

2.6. Model development

The model (and its implementation as a GIS tool) was created using two basic assumptions. The first is that there is a variable (combination of OLI bands) which correlates well with the trophic state of the lakes, while the second is that the reference trophic state depends on the mean depth of the lake. The lake topology trophic state index was introduced by Carlson in 1977 (Carlson, 1977) on the basis of lake topology (Elster, 1974). The trophic state of freshwater and coastal marine water may be described as oligotrophic, mesotrophic, eutrophic and hypertrophic and is related to amount of total nitrogen, total phosphorus, chlorophyll a and Secchi disk transparency (Smith, 1988; Chin, 2006). There were many projects which proved the relation between remote sensing reflectance and Chlorophyll a and Secchi disk transparency (Chang et al., 2015) which correlate well with trophic state. The relationship between the depth and size of lakes and its natural trophic class is well known in limnology (Lampert and Sommer, 2007). This relation is explained by several factors. The deep lakes have usually higher ratio of lake volume to watershed area and in result smaller flux of nutrients per unit lake volume. In deep lakes only a small part of epilimnion is in contact with the sediments and in result the return of nutrients to the euphotic zones is limited. Large and deep lakes have a longer retension time and a large part of nutrients is lost through sedimentation.

There are two inputs into the model, namely the lake geodatabase and a raster catalog with Landsat 8 image to be used (Fig. 2). As previously described, the lake geodatabase must contain lake vector layers with mean depths and lake raster zones. A raster catalog should include all unzipped Landsat 8 bands and a raster mask of scene with NoData value in areas of clouds and clouds shadows and the value 1 elsewhere. Such a mask must be previously prepared for each Landsat 8 scene we want to analyze. The first step in the model is to calculate reflectance of the first five bands from the satellite image using the DOS method (Chavez, 1988); we must also assume that there is 1% of reflectance from the darkest object, as descried earlier. Following this, all reflectance bands (Ref1,. . .Ref5) are masked using the raster mask. Once this has been achieved, for each lake raster zone in the area of the satellite image, the mean values of reflectance of the analyzed bands are calculated using the zonal statistics geoprocessing tool. In addition new raster Ref2/Ref4 is created for which the mean and standard deviation values of reflectance are calculated using also zonal statistics tool. This step creates a table with lake ID (LID) for each lake and mean values of reflectance for two OLI bands (2,4), Ref2/Ref4 and standard deviation for Ref2/Ref4. The process is controlled by minimum pixel number, which is required to calculate statistics. In the next step lakes are divided into subsets on the basis of mean depth. The initial width of subsets is two meters, with one-meter overlap: 0–2, 1–3, 2–4. . . etc. The one meter overlap is to make classification more realistic and smooth. If there are not enough lakes in the subset, everything which is controlled by a parameter – minimum number of lakes in a subset, consecutive subset is merged. The minimum number of lakes in lakes subsets has to be defined because statistics are calculated from each of the subsets. For each subset, mean value and standard deviation of trophic state index I is calculated to show variability of the ecological state of lakes grouped into one subset. Following this, thresholds of trophic state index are used to assign a class to lakes in each subset.These thresholds are defined as the number of standard deviations of trophic state index below and above mean index I in each subset. These thresholds are common for all subsets. As a result, lakes are divided into one of five classes: Bad, Poor, Moderate, Good, High. In the case where a lake occurs in two subsets, it is possible that depth interval overlap will occur, potentially resulting in two different classes being assigned to one lake. If this happens, the better class is chosen. Defining proper thresholds should be supported with expert knowledge of the area and may be verified by comparing results with other classification methods even on the few lakes that have a diverse ecological state and depth. By comparing the results of different threshold values, models may by calibrated to get results which best fit the expected classes. It is also possible to test lakes for uniformity using coefficient of variation of Ref2/Ref4. For this reason standard deviation and mean value is added to results of modeling. The output of the model is a vector layer of lakes with an assigned class describing ecological state at the time of a particular satellite image and mean and standard deviation values of Ref2/Ref4.

3. Results

3.1. Field measurements analyses

In situ measurements were used to find the best combination of spectral bands of Landsat 8 satellite images describing the trophic state of a lake. We used OLS to obtain the linear equation and to predict parameters such as SD or Chl-a from space using a combination of DN, radiance or reflectance. We tested the relation between the natural logarithm of all measured parameters: Secchi depth (SD), chlorophyll-a (Chl-a), suspended particulate matter (SPM), colored dissolved organic matter (CDOM) and dissolved organic carbon (DOC) with different variables of spectral band combinations, as shown in Table 2. We used adjusted R Squared (above 0.5 to accept the model) as well as several diagnostic criteria to evaluate

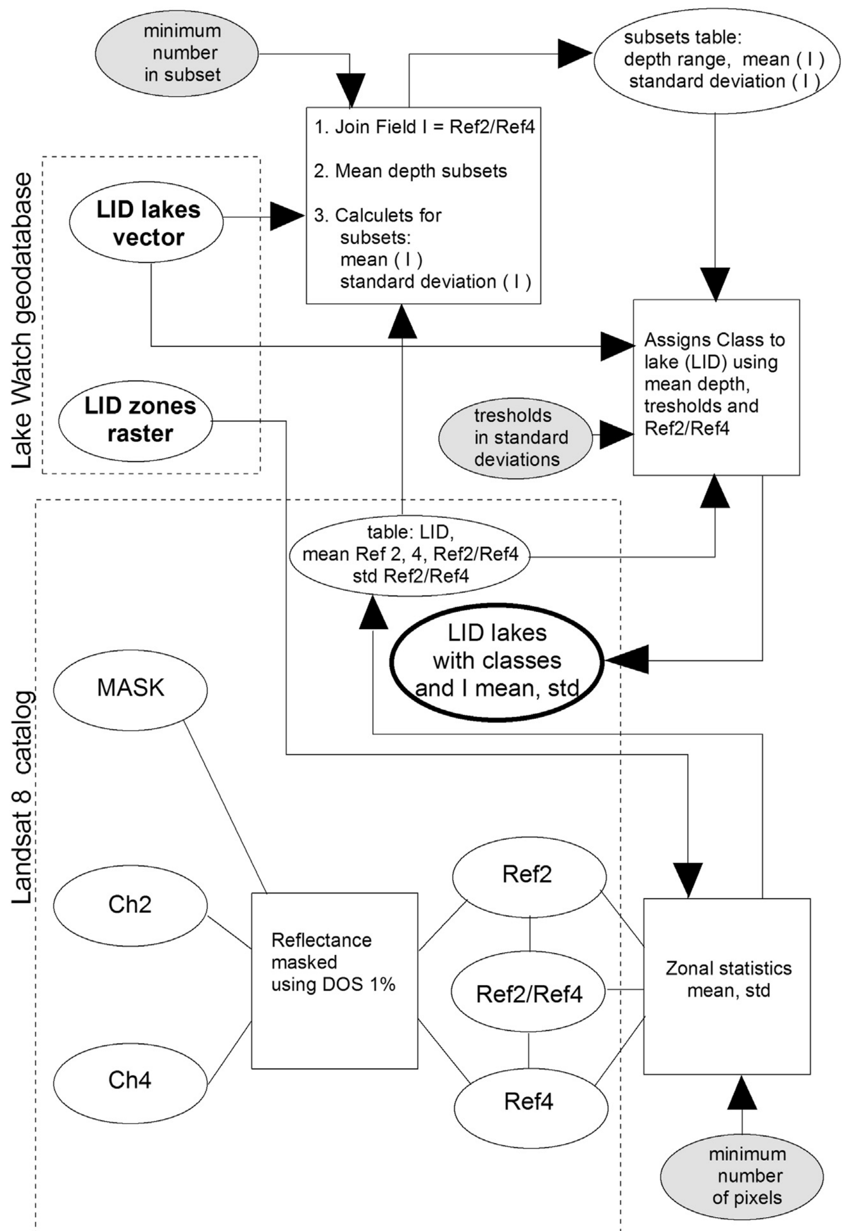

Fig. 2. The model of classification of lakes to five classes of the trophic state and its implementation in GIS.

Table 3
Results for accepted models. Ln(SPM) – June and Ln(DOC) – June models do not pass tests.

Parameter	Scene	Adjusted R^2	p	VAR6 significance [%]	VAR6 stability [%]
Ln(SD)	June	0.76	<0.05	62	+100
Ln(SD)	July	0.81	<0.05	94	+100
Ln(Chl-a)	June	0.52	<0.05	31	−100
Ln(Chl-a)	July	0.76	<0.05	69	−100
Ln(SPM)	June	–	–	28	−98
Ln(SPM)	July	0.77	<0.01	89	−100
Ln(DOC)	June	–	–	42	−96
Ln(DOC)	July	0.70	<0.05	67	−100
Ln(CDOM)	June	0.60	<0.05	50	−97
Ln(CDOM)	July	0.53	<0.05	87	−100

acceptance of the model. Every model was checked for: maximum coefficient p-value cutoff (0.05 to accept), maximum VIF value cutoff (7.5 to accept), minimum acceptable Jarque Bera p-value (0.1 to accept) and minimum acceptable spatial autocorrelation p-value (0.1 to accept). Results of the accepted models are presented in Table 3. In total there were two sampling campaigns following two satellite images. However, the attempt to merge data and thus create a more general model yielded poorer results, and so models were presented only for separate scenes. For all analyzed variables VAR6 (Blue/Red) represents the greatest significance, with very

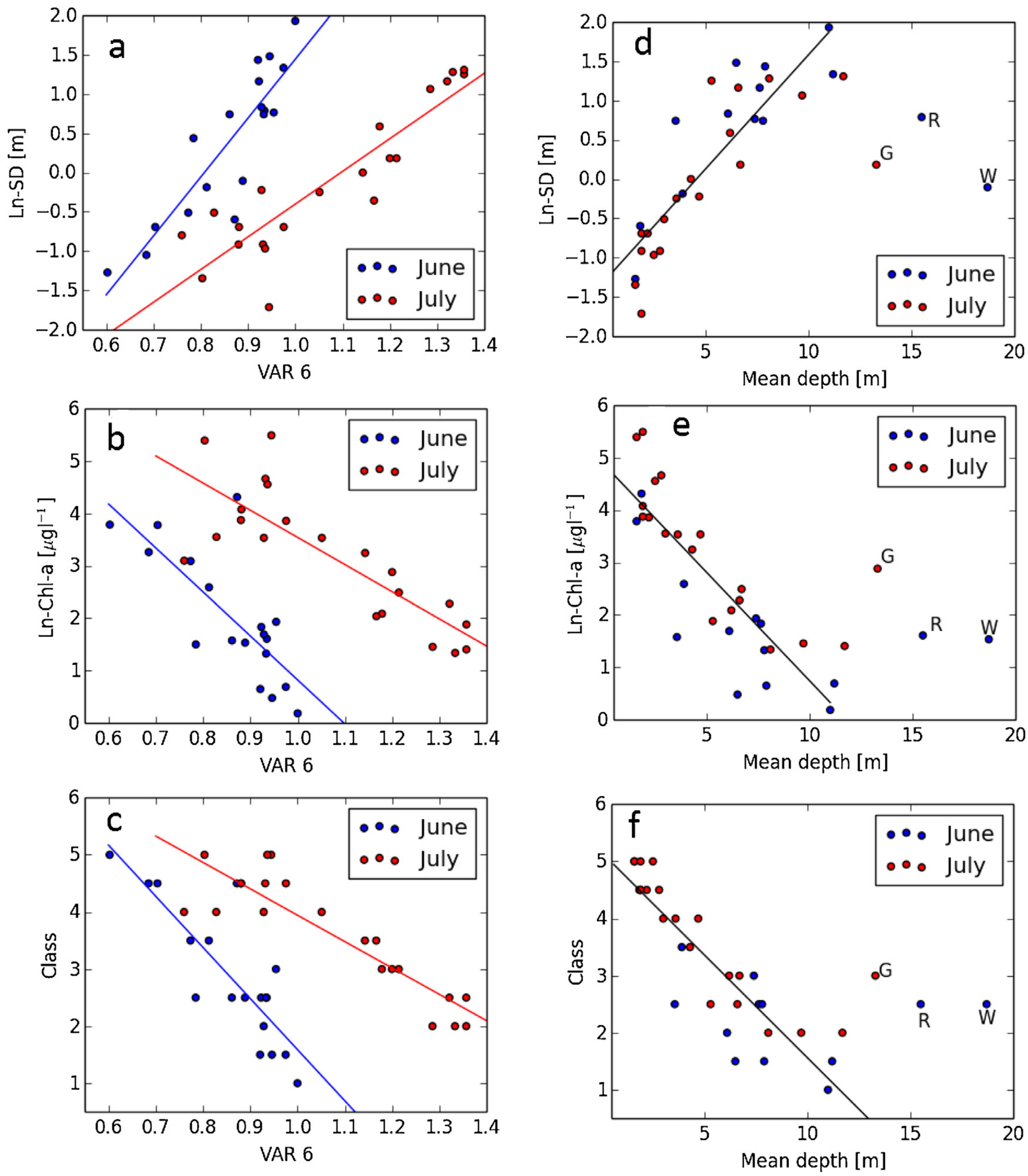

Fig. 3. The relation between measured parameters and VAR 6 (Blue/Red Landsat 8 bands) and mean depth of sampled lakes: a, b,c—SD, Chl-a, tropic class vs VAR 6; d,e,f—SD, Chl-a, tropic class vs mean depth.

good stability compared to other variables. The significance of this variable was noticed by several authors (Kloiber et al., 2002; Sass et al., 2007; Torbick et al., 2013).

The scatterplot analysis (Fig. 3a–c) shows relations described by a model between VAR6 and SD, Chl-a and a trophic state rank defined in Table 1. There is an essential difference in relation for both scenes. The reason for this may be a result not of the perfect atmospheric corrections or the relatively different composition of the phytoplankton in June and July. In June diatomphyceae were present in several lakes while in July they were absent with strong dominance of chlorophyceae. As a trophic state rank is a function of SD and Chl-a it correlates well with VAR6 (AdjR^2 = 0.68 for June and AdjR^2 = 0.79 for July with significant p value < 0.01). In conclusion, the VAR6 may be treated as a good index of trophic state for single scene.

The second part of Fig. 3(d–f) describes the relation of SD, Chl-a and trophic state to mean depth of a lake. Three lakes with mean depth above 12 m did not fit the general trend and were excluded during the creation of the model. Indeed, they have higher Chl-a and lower SD than other lakes with similar mean depth. In lake Radunskie Gorne (G), for example, the bloom of diatomophyceae was observed, with biomass volume of 87 000 mm^3/m^3. Both SD and Chl-a are strongly correlated with mean depth (AdjR^2 = 0.76 for SD and AdjR^2 = 0.73 for Chl-a) creating a clearly visible trend for the relation between trophic state and mean depth (Fig. 3 f) with AdjR^2 = 0.78. This also explains the improvement in results of linear regression achieved by some researchers by including mean depth in band variables (McCullough et al., 2012). In our research, adding mean depth as an additional is not intendent to improve models for separate scenes. Instead it significantly improves the general model, making the mean depth the most significant vari-

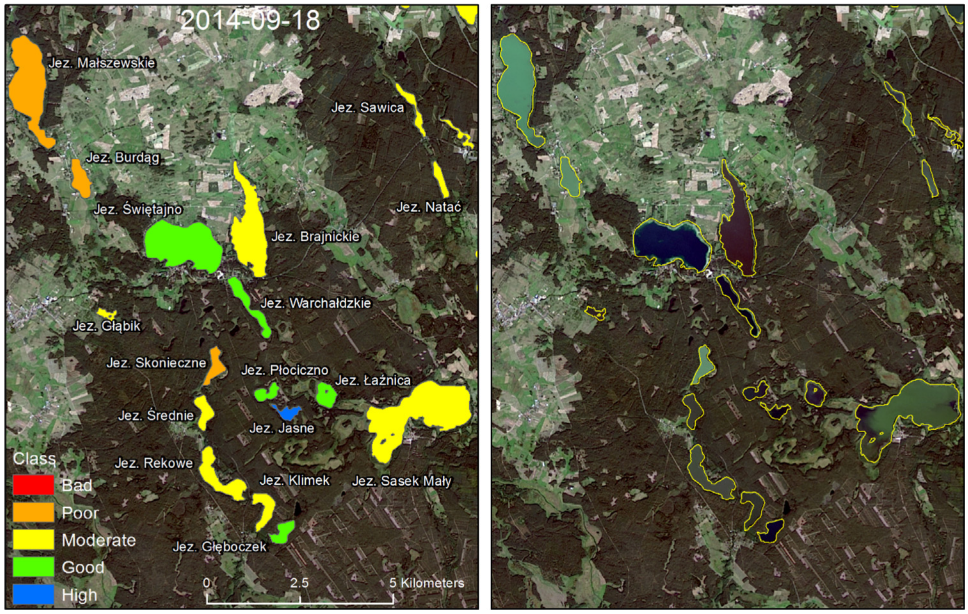

Fig. 4. The product created by the tool for assigning ecological classes for small part of image area. On the right side the part of image with outlines of lakes and on the left side the same image with overlaid classified lakes with labels.

Table 4
Verification of the model. Satellite S – September image; Satellite A – April image; best fit – the most fitting choice from Satellite S and A; worst fit – the least fitting choice from Satellite S and A.

Number of lakes	Satellite S/Satellite A	Expert/Satellite (best fit)	Expert/Satellite (worst fit)
With the same ecological state	9	9	2
With the one class difference	6	9	13
With the two class difference	3	0	3

able. For particular mean depth value, lakes below a regression line for trophic class represent a better ecological state than lakes above the line. We used this property to relative validate of ecological state of the subset of lakes with similar mean depth.

3.2. Model implementation

The implementation of the tool to assign ecological classes consists of Python Toolbox with two tools. The first tool has three parameters: Landsat 8 catalog, Landsat 8 MTL text file and the geodatabase Lake Watch localization. The second tool parameters are: Landsat 8 catalog, localization of Lake Watch geodatabase and the name of the lake polygon shapefile to be created. In addition, the user has to input the value of thresholds between classes Bad and Poor, Poor and Moderate, Moderate and Good and Good and High. The values are the numbers of standard deviations from the mean values (with minus and plus sign), and may be used for model calibration. Fig. 4 presents the example of products created using this tool. This is only small part of the whole image which was analyzed to fulfill the requirements of at least 30 lakes in each mean depth subset.

To verify satellite based methods, we compared results from the legal classification of 18 lakes made by experts in 2014 (however some used data were a few years old) with classification of two Landsat 8 images (6 September 2014 (S) and 21 April 2015 (A)). The results are shown in Table 4. Parameters used in the model: minimum number of pixels – 15, minimum number of lakes in subset – 30, thresholds to differentiate classes: 1.5,0.5,−1,−2 (from High to Bad).). The best fit means that the best fitting result from classification of both images has been chosen and the worst fit means that the worst result was the choice.

(Table 4) shows a good agreement between the satellite and the expert based method. Half of the lakes have an identical state, as indicated by using best fit from both satellite methods and expert judgment. The same number of lakes have identical ecological states, as indicated from both images. There were three cases of two-class difference noted when comparing both satellite methods and satellite and expert judgment worst scenario. The difference was never more than two classes in any comparison and only one class in the best fit case. The time factor may be critical or at least very important when assigning a class to the lake. This, from one side, may be treated as a problem, but from the other, may present the chance to monitor changes in the lakes' ecological state at the time, by comparing results of models using satellite images in time series. This requires further evaluation in the future. We also analyzed spatial variation of both bands (OLI2 and OLI4) and calculates the coefficient of variation (standard deviation/mean) for the sample of about 1000 lakes from several images. The mean value of coefficient of variation was about 0.06 (with standard deviation about 0.05). This shows that reflectance in both bands (as a trophic state index) has in general low level of spatial variation. However some lakes show higher level of spatial variation. There were two reasons for this. The first one is that some shallow lakes may be partly covered by reeds during the summer and the second is the situation when occasional intensive bloom of phytoplankton occurs. As a result the coefficient of variation of band 2 or 4 may be a good additional index describing the state of the surface of particular lake. The values of the coefficient of variation may be calculated for each lake from mean and standard deviation fields of model output layer. It is possible to define a threshold value to inspect some lakes for their uniformity. Also some group analyses may be carried out which will result in dividing lakes to separate parts.

4. Conclusions

The main limitation in using satellite data to predict SD and Chl-a,e which are considered good indicators of a lake's trophic state, is that each scene needs its own formula. There are several possible reasons for this, but it seems as though the main problem is with the determination of true reflectance (or radiance) due to the uncertainty with a proper atmospheric correction of the generally low signal. The other reasons are due to changing environmental conditions: abiotic like lake surface dynamic state or biotic like phytoplankton composition. It is very important to have representative samples for OLS covering the whole range of trophic states in the area of research.

There is strong evidence regarding the statistical relation between mean depth and expected trophic state of the lake which is justified by the functioning of the lake ecosystem. The shallow lakes do not have constant stratification during the phytoplankton growth season, and thus may be effected by resuspension. In addition to this, the shallow water column may be easily contaminated as a result of nutrients delivered by runoff from the land. On the other side, deep lakes are free from resuspension and the long lasting stratification and high water column creates a buffer which influences the nutrient download caused by runoff. They are also colder, which slows down biological processes. As all these processes are natural, it may be expected that reference lakes will have a trophic state depending on their depth. The epilimnion of the lake is an occasionally highly dynamic environment influenced mainly by wind mixing which due to the limited size of the lakes can make their surface nearly uniform both physically and biologically. However in stable conditions this surface may be less uniform due to blooms or local upwelling events. The mean statistic is the simple generalization of the lake surface state in time scale of weeks. This assumption may not be however true for very large lakes or coastal waters when their trophic state may differ in various parts. One of the solutions is the subdivision of such waters body. This has been proposed for some lakes in our project.

The proposed method is an attempt to make use of the increasing amount of high quality satellite data for the monitoring of vast numbers of lakes with different trophic states which may change in the short and long term. As the number of images suitable for such analyses will increase with the launching of new satellite systems, the statistics of trophic classes for particular lakes may reveal information if its state is stable or changes comparing to other lakes. This may be a valuable tool for the lakes' environmental managers.

Acknowledgments

The study was financed by the University of Gdansk (530-G300-D433-14). We also thank Adam Inglot and Anna Piszewska from the GIS Center of University of Gdansk for their contribution to the field measurements.

References

Brezonik, P., Menken, K.D., Bauer, M., 2005. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). Lake Reservoir Manage. 21 (4), 373–382.

Carlson, R.E., 1977. A trophic state index for lakes. Limnol. Oceanogr. vol. 22 (2), 361–369.

Chang, N., Imen, S., Vannah, B., 2015. Remote sensing for monitoring surface water quality status and ecosystem state in relation tot he nutrient cycle: a 40-year perspective. Crit. Rev. Environ. Sci. Technol. 45, 101–166.

Chavez, P.S., 1988. An improved dark-object substraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24, 459–479.

Chin, D.A., 2006. Water-Quality Engineering in Natural Systems. John Wiley & Sons, pp. 609.

Duan, H., Zhang, Y., Zhang, B., Song, K., Wang, Z., Liu, D., Li, F., 2008. Estimation of chlorophyll-a concentration and trophic states for inland lakes in Northeast China from Landsat TM data and field spectral measurements. Int. J. Remote Sens. 29 (3), 761–786.

Elster, H.J., 1974. History of limnology. Mitteilung Internationale Vereinigung fuer Theoretische unde Amgewandte Limnologie 20, 7–30.

HELCOM, 1988. Guidelines for the baltic monitoring programme for the third stage part D. biological determinands. Baltic Sea Environ. Proc. 27 (D), 153.

Janczak, J., 1996–1999. Atlas jezior polski. parts 1,2,3. Poznan, Bogucki Wydawnictwo Naukowe, 778.

Kalio, K., 2000. Remote sensing as a tool for monitoring lake water monitoring. In: Heinonen, P., Ziglio, G., Van der Beken, A. (Eds.), Hydrological and Limnological Aspects of Lake Monitoring. John Wiley & Sons.

Kloiber, S.M., Brezonik, P.L., Bauer, M.E., 2002. Application of Landsat imagery to regional-scale assessments of lake clarity. Water Res. 36, 4330–4340.

Lampert, W., Sommer, U., 2007. Limnoecology: The Ecology of Lakes and Streams. Oxford University Press, Oxford, pp. 324.

McCullough, I.M., Loftin, C.S., Sader, S.A., 2012. Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity. Remote Sens. Environ. 123, 109–115.

Olmanson, L.G., Bauer, M.E., Brezonik, P.L., 2008. A 20-year landsat water clarity census of minnesota's 10000 lakes. Remote Sens. Environ. 112 (11), 4086–4097.

Peeters, E.T., Franken, R., Jeppesen, E., Moss, B., Becares, E., Hansson, L., Romo, S., kairesalo, T., Gross, E., Donk, E., Noges, T., Irvine, K., Kornijow, R., Scheffer, M., 2009. Assessing ecological quality of shallow lakes: does knowledge of transparency suffice? Basic Appl. Ecol. 10, 89–96.

Phillips, G., Pietilainen, O.P., Carvalho, L., Solimini, A., Solheim, L., Cardoso, A., 2008. Chlorophyll-nutrient relationships of different lake types using a large European dataset. Aquat. Ecol. 42, 213–226.

Poikane, S., Portielje, R., Berg, M., Phillips, G., Brucet, S., Carvalho, L., Mischke, U., Ott, I., Soszka, H., Wichelen, J., 2014. Defining ecologically relevant water quality targets for lakes in Europe. J. Appl. Ecol. 51, 592–602.

Sass, G.Z., Creed, I.F., Bayley, S.E., Devito, K.J., 2007. Understanding variation in trophic status of lakes on the Boreal Plain: a 20 year retrospective using Landsat TM imagery. Remote Sens. Environ. 109, 127–141.

Scheffer, M., 1998. Ecology of Shallow Lakes. Chapman & Hall, London.

Smith, V.H., 1988. Cultural eutrophication of inland, estuarine and coastal waters. In: Pace, M.L., Groffman, P.M. (Eds.), Success, Limitations and Frontiers in Ecosystem Science. Springer, New York, pp. 7–49.

Song, C., Woodcock, C.E., Karen, C.S., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. 75, 230–244.

Tebbs, E.J., Remedios, J.j., Harper, D.M., 2013. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline-alkaline, flamingo lake, using Landsat ETM+. Remote Sens. Environ. 135, 92–106.

Torbick, N., Hession, S., Hagen, S., Wiangwang, N., Becker, B., Qi, J., 2013. Mapping inland lake water quality across the Lower Peninsula of Michigan using Landsat TM imagery. Int. J. Remote Sens. 34, 7607–7624, 21–22.

USGS, 2015. Landsat 8 (L8), Data Users Handbook, Version 1.0, June 2015, EROS, Sioux Falls, South Dakota, pp. 98.

Vollenweider, R.A., Kerekes, J.J., 1982. Eutrophication of waters Monitoring, assessment and control. In: Final Report. OECD, Paris, pp. 154.

Zajączkowski, M., Nygård, H., Hegseth, E.N., Berge, J., 2010. Vertical flux of particulate matter in an Arctic fjord: the case of lack of the sea-ice cover in Adventfjorden 2006–2007. Polar Biol. 33, 223–239.