

# Chlorophyll-a concentration retrieval in eutrophic lakes in Lithuania from Sentinel-2 data

Dalia Grendaitė<sup>1</sup>,

Edvinas Stonevičius<sup>1</sup>,

Jūratė Karosienė<sup>2</sup>,

Ksenija Savadova<sup>2</sup>,

Jūratė Kasperovičienė<sup>2</sup>

<sup>1</sup> Vilnius University,  
M. K. Čiurlionio St. 21,  
03101 Vilnius

<sup>2</sup> Nature Research Centre,  
Akademijos St. 2, 08412 Vilnius  
Email: dalia.grendaite@gmail.com

Grendaitė D., Stonevičius E., Karosienė J., Savadova K., Kasperovičienė J. Chlorophyll-a concentration retrieval in eutrophic lakes in Lithuania from Sentinel-2 data. *Geologija. Geografija*. 2018. T. 4(1). ISSN 2351-7549.

Inland waters are an important habitat for flora and fauna and are also used for aesthetic, recreational, and industrial needs; therefore, monitoring the current state of freshwaters and applying measures to improve water quality are of high importance. To have an efficient monitoring system that could cover large areas, the use of remote sensing data is crucial. In this study the suitability of the Sentinel-2 Multispectral Imager data is tested for observing cyanobacteria bloom events in the eutrophic lakes and retrieving the chlorophyll-a concentration – an indicator of phytoplankton biomass. The analysis is carried out using data from four lakes in Lithuania – two eutrophic blooming lakes and two oligo-mesotrophic non-blooming lakes. The results showed that reflectances are higher in the eutrophic lakes than in the oligo-mesotrophic lakes due to the presence of an optically active constituent, namely, chlorophyll-a pigment. We tested empirical equations for chlorophyll-a concentration retrieval in eutrophic lakes derived in other studies to check whether they could be used without adaptation to local conditions. Most of the equations performed well ( $R^2 = 0.5–0.8$ ); however, they had high RMSEs = 17–53  $\mu\text{g L}^{-1}$ . The equation used with the bottom of atmosphere data CHL8\_L2A ( $R^2 = 0.76$ ) had the lowest RMSE = 9  $\mu\text{g L}^{-1}$ . In addition, we derived empirical equations for eutrophic lakes in Lithuania. The equations that were based on the Sentinel-2 band ratio B5/B4 and the three band (B4, B5, and B8A) expression performed the best ( $R^2 = 0.77–0.79$ ) and had lower RMSE = 7  $\mu\text{g L}^{-1}$  than empirical equations from other studies. A larger *in situ* dataset could improve the algorithm performance in retrieving the chlorophyll-a concentration. The first attempts to map water quality parameters in eutrophic lakes in Lithuania using the data received from the Sentinel-2 MSI sensor show good results, as the changes in reflectance, caused by the changes in chlorophyll-a concentration, can be seen from satellite images.

**Keywords:** lakes, water quality, chlorophyll-a, remote sensing, Sentinel-2

## INTRODUCTION

Inland waters provide society with water that is used for aesthetic, recreational and industrial needs. In addition, lakes create habitat for flora and fauna; however, lake water quality is affected by nutrient loads from the catchment basin that cause eu-

trophication and algal and cyanobacteria blooms that pose problems both for biota in the lakes and people. Harmful cyanobacteria blooms can cause adverse impacts on health due to the toxins released by cyanobacteria species (Koreivienė et al., 2012). In addition, recently, inland waters gained more attention since they were shown to be very

important in the global carbon cycle and climate change (Tranvik et al., 2009). However, inland water monitoring systems cannot ensure monitoring of a large number of lakes since those systems use laboratory analyses and, therefore, require considerable human and technical resources. Thus, *in situ* measurements are very limited in space and time (Philipson et al., 2016) and cannot provide continuous operational water quality monitoring, especially for a large number of lakes and large lakes.

The first aquatic remote sensing applications were largely focused on the ocean (Mishra et al., 2017). The first space-borne sensor specifically designed to study living marine resources was the Coastal Zone Colour Scanner (CZCS) that flew 1978–1986 on board Nimbus-7 satellite and for the first time provided the opportunity to observe the oceans on a global scale (Hovis et al., 1980). Further missions that have largely contributed to ocean observation: Sea-viewing Wide Field-of-view Sensor, SeaWiFS (operated by NASA from 1997 to 2010), the Moderate Imaging Spectroradiometer, MODIS (on-board Terra and Aqua satellites, operated by NASA since 1999), and ESA's Medium Resolution Imaging Spectrometer, MERIS (on-board ENVISAT, 2002–2012) (McClain, 2009).

Over these few decades since the first Earth Observation (EO) satellites were launched, there have been attempts to map water quality of inland water bodies using data from these satellites (Giardino et al., 2014). Extracting data from satellite images provides an opportunity to have data coverage for large areas, and overcome the paucity of *in situ* measurements. Using satellite imagery data to derive water quality parameters was challenging as these first satellites had relatively low radiometric, spectral and spatial resolution and, therefore, had difficulty making use of the weak signal from the water surface, given that most of the incoming radiation is absorbed by water rather than reflected. The recently launched optical high spatial resolution (10–20 m) Multispectral Imager (MSI) on board Sentinel-2 (operated by the European Space Agency (ESA) as part of the European Commission Copernicus project), designed for land monitoring studies (Drusch et al., 2012), provide the aquatic scientists community with new opportunities to investigate water quality in smaller water bodies and find solutions for detection of contamination in them from space. The delivery of the Sentinel-2

MSI data is planned to be continuous up to 2030 (ESA, 2014); therefore, it is important to find the algorithms for chlorophyll-a retrieval as they could be used operationally for many years.

In Lithuania there are around 6000 natural lakes that make up about 1.4% of the country's area (Kilkus, Stonevičius, 2012). Most of the lakes are small – there are just 150 lakes larger than 1 km<sup>2</sup> (Kilkus, Stonevičius, 2012), and, hitherto, it has been difficult to observe them from space due to low spatial resolutions of satellite sensors. A large number of lakes have suffered adverse anthropogenic effects. The contamination with nutrients from close proximity agricultural fields, irrigation engineering, induce eutrophication processes that cause overgrowth of aquatic vegetation of lakes, accelerated sedimentation rates, common algal and cyanobacteria blooms, depletion of oxygen; and therefore, lakes grow old and are neither a suitable habitat for many flora and fauna species nor can be used for human needs (Balevičienė et al., 2009).

The main biological parameter describing the eutrophication of lakes and lake water quality is phytoplankton biomass. The photosynthetic phytoplankton pigment chlorophyll-a is often used as it is a main indicator of phytoplankton biomass (Mishra et al., 2017) that also makes water look green since chlorophyll-a absorbs in the blue and red regions of the visible spectrum and reflects in the green (Hovis et al., 1980). The chlorophyll-a concentration is routinely estimated by the environmental protection agencies using laboratory analyses and, therefore, usually data are available only for a small number of lakes and with a low temporal frequency. The use of remote sensing data can help to observe lakes in a large area with a high temporal frequency as the Sentinel-2 imagery is acquired every 3–5 days over the same point (Drusch et al., 2012).

Some attempts were made to use remote sensing data for analysing the temporal variability of chlorophyll-a concentrations (Giardino et al., 2010; Bresciani et al., 2012) and for distinguishing surface accumulations (scum) and dense cyanobacteria blooms (Bresciani et al., 2014) in the Curonian Lagoon in the west of Lithuania. Hitherto, no published analysis is available for inland water bodies in Lithuania that uses remote sensing techniques.

The aim of this study is to evaluate the suitability of the optical Sentinel-2 MSI sensor data

for retrieval of water quality parameters, such as the concentration of chlorophyll-a in eutrophic lakes in Lithuania. We search for an algorithm that could perform well in retrieving the chlorophyll-a concentration. The following tasks are addressed in this study:

1. Retrieve the spectral profile for two eutrophic lakes and two lakes without algal blooms and observe the temporal pattern.

2. Test the empirical equations from other studies for the calculation of chlorophyll-a concentration in selected eutrophic lakes.

3. Derive an empirical chlorophyll-a concentration equation based on the band ratios and test it with available *in situ* measurements.

## DATA AND METHODS

### Study area

The analysis was carried out using data for two eutrophic blooming lakes (Jieznas and Širvys) and two

oligo-mesotrophic non-blooming lakes (Guostus and Šventas). We chose these eutrophic lakes due to their high recreational potential (Balevičienė et al., 2009) and since we had the *in situ* data available for these lakes. Both eutrophic lakes are shallow (average depth is 2.74 and 2.47 m, respectively), and lake basin is mainly covered by agriculture fields (Table 1). The eutrophic lakes suffer from recurrent cyanobacteria blooms that have negative effects on lake ecosystems (Savadova et al., 2016). For comparison and identification of blooming lakes, two oligo-mesotrophic lakes, Guostus and Šventas, in which no cyanobacteria blooms occur, were selected (Fig. 1). Lake Guostus is characterized by surface area similar to Lake Jieznas and agriculture fields dominate in its basin area. The lake is situated close to Lake Jieznas (~12 km). The other oligo-mesotrophic Lake Šventas was selected for the analysis to ensure reliable reflectance from the lake since the lake is situated in a natural environment (70% of lake basin area is forest).

Table 1. Lake characteristics (Balevičienė et al., 2009)

Lake	Coordinates (Longitude, Latitude)	Lake area, ha	Average depth/ maximum depth, m	Land cover (CORINE LC) in lake basin: urban/agriculture fields/natural	Water renewal time, years
Jieznas	54.5945, 24.17536	74.4	2.74/4.44	12.4/83.6/4	1.254
Širvys	54.98771, 25.21384	85.1	2.47/4.2	4.3/48.3/47.4	0.197
Guostus	54.60496, 24.37268	74.7	13/33.7	2.2/57.9/39.9	1.456
Šventas	55.61519, 26.30079	425.9	8.38/18.2	0/29.7/70.3	7.076

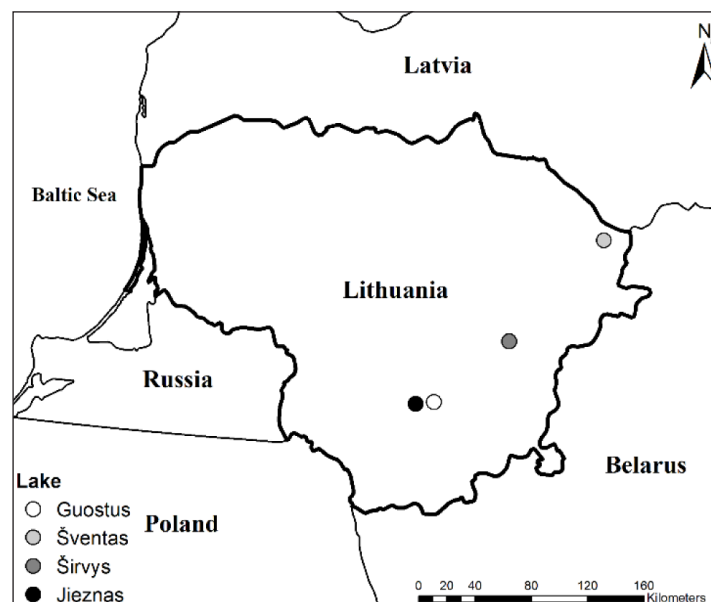


Fig. 1. Studied lakes and their location

In Lake Jieznas *in situ* measurements were carried out once a month, and in Širvys twice per month during the period of July–October 2015 (Table 2). The *in situ* measurements were carried out in the deepest part of the lake, taking sample from the surface layer (0.5 m) using a Ruttner sampler (Wetzel et al., 2001). Water transparency was estimated using the Secchi disk (Wetzel, 2001). The chlorophyll-a concentration was estimated at the laboratory using the fluorometer AlgaeLabAnalyser (Bbe moldanke, 2018). For lakes Guostus and Šventas only yearly averages are available from the Lithuanian Environmental Protection Agency (EPA) (Table 2). The Lithuanian EPA estimates chlorophyll-a concentration spectrophotometrically after extraction with boiling 90% ethanol according to the ISO 10260:1992 standard (LAND 69-2005 ..., 2005).

### Satellite data

Satellite images in 2015 are available only from July since the Sentinel-2A was launched in June 2015. The Multispectral Imager (MSI) sensor observes Earth at 13 spectral bands in 10 m, 20 m, and 60 m spatial resolutions, and spectral resolution spans from the visible (VIS) to the short wave infrared (SWIR) wavelengths of the electromagnetic spectrum (443–2190 nm). The radiometric resolution is 12-bit. The images that

had lower than 25% cloud cover were selected for analysis and downloaded from the Copernicus Open Access Hub web platform (<https://sci-hub.copernicus.eu/>). We got six images from July to September (Table 3). In addition, sometimes there were some scattered small cumulus clouds in satellite images; therefore, it was also important to check visually every image to make sure that the studied lakes were not covered by cloud or cloud shadow.

The downloaded images are of level 1C – radiometrically and geometrically corrected. The images are in JPEG 2000 format and provide top of Atmosphere (toA) reflectances that are dimensionless and vary between 0 and 1 (Drusch et al., 2012). The layer of atmosphere influences toA reflectance and the atmospheric influence is removed by applying an atmospheric correction (AC) to generate the level 2A data – bottom of atmosphere (boA) reflectance.

### Workflow

The atmospheric correction was performed with the Sen2cor (version 2.4.0) (ESA, 2017) tool available in the ESA SNAP toolbox (ESA, 2018). It uses the ATCOR algorithm (Müller-Wilm, 2017). The images were resampled to 20 m resolution and the pixels of central areas of lakes were exported using the ESA SNAP Sentinel-2 toolbox. Only

Table 2. Water transparency and chlorophyll-a concentration in the studied lakes: <sup>1</sup> *in situ* measurements, average of July–October data. <sup>2</sup> annual average from the Lithuanian Environmental Protection Agency

Lake	Secchi disk depth, m	Standard deviation of the Secchi disk depth, m	Chlorophyll-a concentration, $\mu\text{g L}^{-1}$	Standard deviation of the chlorophyll-a concentration, $\mu\text{g L}^{-1}$
Jieznas <sup>1</sup>	0.6	0.1	64.7	8.2
Širvys <sup>1</sup>	0.8	0.1	43.3	8.5
Guostus <sup>2</sup>	5.6	–	2.45	–
Šventas <sup>2</sup>	7.1	–	2.12	–

Table 3. The Sentinel-2 image dates for analysed lakes: \* indicates that a smaller lake area was selected due to the presence of clouds on a part of the lake, – indicate missing satellite data for the lake

Lake	25.07.2015	04.08.2015	11.08.2015	14.08.2015	24.08.2015	30.09.2015
Jieznas	+	+	+	+	+	+
Širvys	+	+	+	+	+	+
Guostus	+	+	+	+	+	–
Šventas	+	+	+	–	+	+

the central area of the lake was chosen to avoid the influence on reflectance of macrophytes and the lake bottom. After the atmospheric correction, data were filtered using the output quality scene classification (generated in the process of the atmospheric correction with Sen2Cor tool), only retaining pixels that were classified as water. Mostly, just several pixels were removed for a lake; however, for Lake Guostus for 11.08.2015 all the pixels were classified not as water but as dark-area pixels (Müller-Wilm, 2016) and, therefore, were removed. The pixels were averaged for each day for each lake and reflectance graphs were drawn using both level 1C and level 2A data. Analysis was done using *Rstudio* and graphs were drawn using the dedicated data visualization *ggplot2* package (Wickham, 2016). We illustrate the satellite data processing scheme in Fig. 2.

### Chlorophyll-a concentration retrieval methods

The chlorophyll-a concentration was calculated using several available band ratio algorithms from other scientific studies (Table 4). Band ratio algorithms are empirical algorithms that are based on statistical relationship between a colour index, i.e. band ratio, and a water quality parameter (Mishra et al., 2017). The algorithms tested were primarily used for other satellites that had different band wavelengths than Sentinel-2. The closest wavelength band from Sentinel-2 bands was then selected for calculation. The bands B4, B5, and B6 were used for calculation in algorithms CHL1, CHL5, CHL6. The bands B4 and B5 were used for algorithms CHL2a, CHL2b, CHL3, CHL4, CHL7 and CHL9. For algorithm CHL8 – B4, B5, and B8A were used and 885 nm wavelength was changed to the 865 nm in the expression

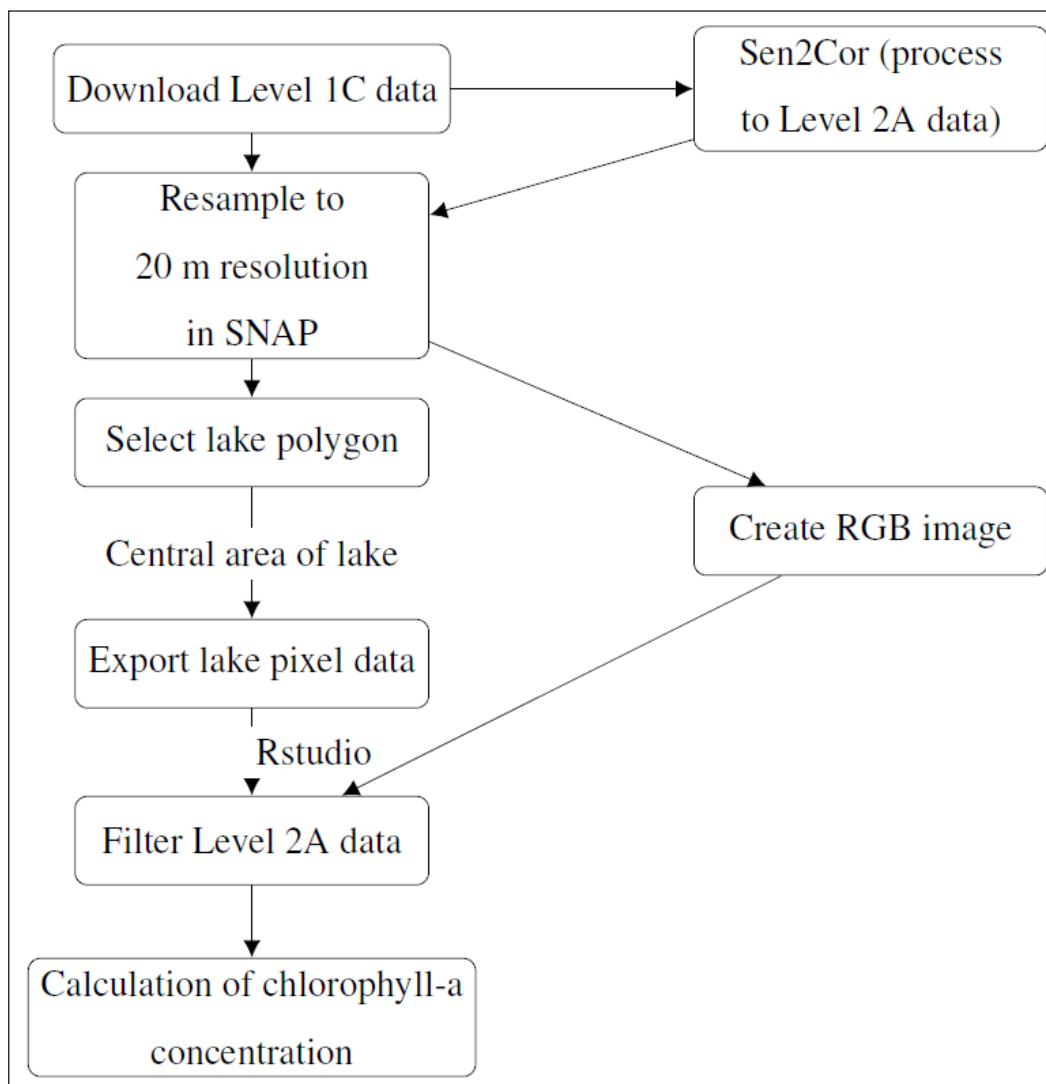


Fig. 2. Satellite data processing scheme

Table 4. The algorithms used for calculation of chlorophyll-a concentration found in the literature

Code	Reference	Algorithm	Sentinel-2 bands used	Empirical equations to calculate the concentration of chlorophyll-a
CHL1a CHL1b	Toming et al. (2016)	$P_1 = R_{703.9} - \frac{R_{664.5} + R_{740.2}}{2}$	B5, B4, B6	For Level1C data: $Chla = 169 \times P_1 + 19.5$ For Level2A data: $Chla = 2231 \times P_1 + 12.7$
CHL2a CHL2b	Duan et al. (2007)	$P_2 = \frac{R_{710}}{R_{670}}, P_2 = \frac{R_{700}}{R_{670}}$	B5, B4	For : $\frac{R_{710}}{R_{670}}$ $Chla = 78.018 \times P_2 + 65.88$ For : $\frac{R_{700}}{R_{670}}$ $Chla = 93.67 \times P_2 - 90.4$
CHL3	Hunter et al. (2008)	$P_2 = \frac{R_{710}}{R_{670}}$	B5, B4	$\log_{10}Chla = 1.33 + 2.44 \times \log_{10}P_3$
CHL4a CHL4b	Jiao et al. (2006)	$P_4 = \frac{R_{719}}{R_{667}}$	B5, B4	For band ratio: $Chla = 0.0282 \times P_4^{3.0769}$ For $R_{719}$ : $Chla = 0.0298 \times R_{719}^{4.03}$
CHL5	Kutser et al. (2016)	$P_5 = R_{700-720} - \frac{R_{665} + R_{770}}{2}$	B5, B4, B6	$Chla = 0.0071 \times P_5 - 0.11$
CHL6	Moses et al. (2009)	$P_6 = (R_{665}^{-1} - R_{702}^{-1}) \times R_{752}$	B4, B5, B6	$Chla = 232.29 \times P_6 + 23.174$
CHL7	Moses et al. (2009)	$P_7 = \frac{R_{708}}{R_{665}}$	B5, B4	$Chla = 61.324 \times P_7 - 37.94$
CHL8	Matthews et al. (2012)	$P_8 = R_{max} - R_{664} - ((R_{885} - R_{664}) \times (\lambda_{max} - 664) / (885 - 664))$	B5, B8A, B4	$Chla = 6903.13 \times P_8 + 2.72$
CHL9	Bresciani et al. (2014)	$P_9 = \frac{R_{708}}{R_{664}}$	B5, B4	$Chla = 86.96 \times P_9 - 67.8$

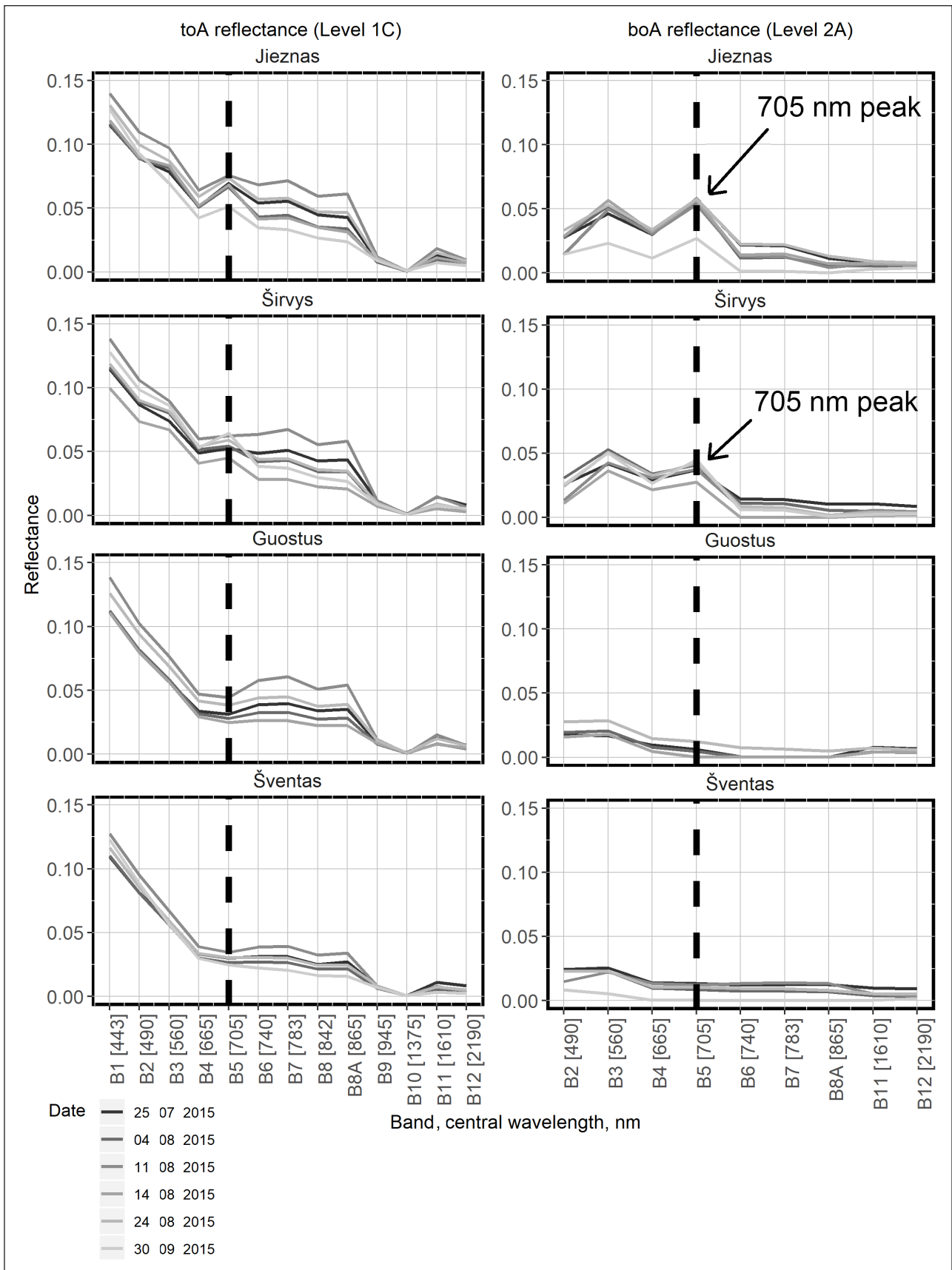
(Table 4). The band ratio expression was the same for  $P_2, P_3, P_4, P_7$  and  $P_9$  algorithms when calculated using Sentinel-2 MSI bands.

The coefficient of determination ( $R^2$ ) and the root mean squared error (RMSE) for chlorophyll-a concentration were calculated to quantify the performance of the algorithms. Calculations were done for both Level 1C data and Level 2A data. The slopes, intercepts and  $R^2$  of the linear relationship were calculated in Rstudio using the  $lm()$  function, and RMSEs were calculated using the *hydroGOF* package (Zambrano-Bigiarini, 2017). The algorithms that had a coefficient of determination equal to or higher than 0.5 were considered to perform well. We also derived the empirical equations for eutrophic lakes in Lithuania using the linear regression method.

## RESULTS

### Top of atmosphere and bottom of atmosphere reflectances of lakes

Reflectances from lakes were very low – across the electromagnetic spectrum top of atmosphere (toA) reflectance varied between 0 in the band B10 and 0.155 in the blue band B1, and bottom of atmosphere reflectance varied from 0 in the bands B6-B8A to 0.074 in the band B5 (Fig. 3). Reflectances were higher in eutrophic lakes (Jieznas and Širvyys) than in oligo-mesotrophic lakes (Guostus and Šventas) in all the bands. The toA reflectance was higher at the short wavelengths, especially in the blue bands (B1 and B2) due to gaseous absorption and the Rayleigh scattering by air molecules in the atmosphere. Hence, it is important to apply an atmospheric correction (AC) that accounts for these



**Fig. 3.** toA and boA reflectances of the eutrophic lakes (Jieznas, Širvys) and oligo-mezotrophic lakes (Guostus, Šventas). The reflectance in band B5 (thick dashed line) is related to chlorophyll-a concentration in water (Mishra et al., 2017)

processes (Mishra et al., 2017). The differences between the atmospherically uncorrected (toA data) and atmospherically corrected data were the highest at these wavelengths (Fig. 3).

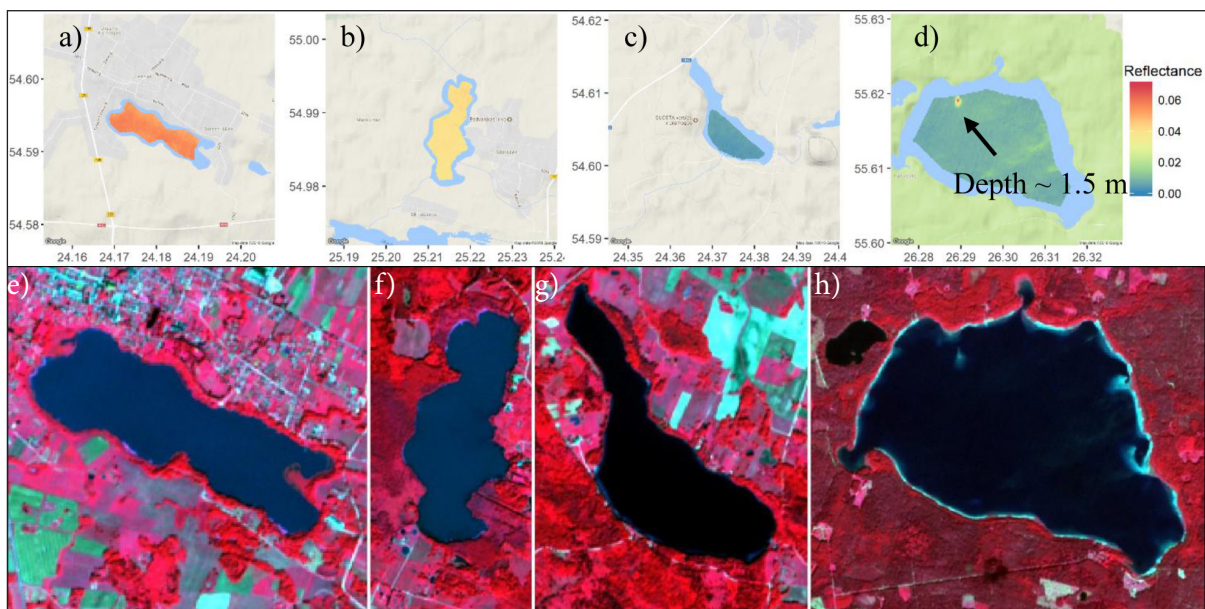
Reflectance in the green band (B3) was still relatively high in comparison to reflectances at the longer wavelengths, partly due to atmospheric processes and partly due to presence of phytoplankton (Mishra et al., 2017). This explains the peak of the B3 band seen in boA data (Fig. 3, boA reflectance).

In the red-edge wavelengths (bands B5-B7) a peak around the 705 nm wavelength in eutrophic lakes is visible caused by the presence of phytoplankton. The peak was more pronounced in the boA reflectance. In oligo-mesotrophic lakes toA reflectance at the 705 nm wavelength was low and boA reflectance was close to zero. The peak height of band B5 (against the baseline of B4 and B6) was used in calculating chlorophyll-a concentration (Kutser et al., 2016; Toming et al., 2016).

The temporal variation was higher in toA than boA reflectance due to the changes of atmospheric conditions (Fig. 3). The variation left after the application of the AC can be accounted for the variation in phytoplankton biomass and other optically active water constituents such as total suspended solids and coloured dissolved organic matter (Mishra et al., 2017). The reflectances on the first image

acquisition date (25 July 2015) were relatively high and later varied over a narrow range in August. In Lake Jieznas a clear drop of reflectance was visible on the 30th of September, that was consistent with *in situ* chlorophyll-a measurements which show that during the summer months, chlorophyll-a concentration is higher and varies little and decreases in autumn. In Lake Širvyys, however, the reflectance pattern is different compared to Lake Jieznas as the lake is characterized by a more diverse dominant phytoplankton species (Savadova et al., 2016). The boA reflectance is the lowest on the 14th of August, consistently with the *in situ* data where a drop of chlorophyll-a concentration in the first half of August is observed.

Satellite images can be used to analyse spatial patterns of phytoplankton distribution in lake, detect areas with the most intense phytoplankton blooms (Bresciani et al., 2014). The central areas of lakes were selected visually to exclude the possible impact on reflectance of the lake bottom and macrophytes. Furthermore, the studied lakes do not have many inflows that could bring water rich in nutrients and suspended particles and affect the reflectance. The reflectance in the band B5 [705 nm] that is often used for chlorophyll-a concentration retrieval (Mishra et al., 2017) was relatively evenly distributed within each analysed lake (Fig. 4(a-d)) with exception



**Fig. 4.** Band 5 boA reflectance distribution in lakes on the 4th of August 2015 (a–d). False Colour Red-Green-Blue (RGB) composite of the lakes studied on the 4th of August 2015, from left to right – eutrophic lakes: Jieznas, Širvyys; oligo-mesotrophic lakes: Guostus, Šventas. Red colour – band B8A [865 nm], green – B4 [665 nm], and blue – B3 [560 nm] (e–h)



of Lake Šventas. In the north-west sector of the lake, a local increase in the boA reflectance (up to 0.074, while the highest reflectance value in eutrophic lakes was 0.066) in B5 was observed. This local increase in reflectance is seen across all five available satellite images for this lake (not shown here). Additionally, a bathymetric map for this lake (Kilkus, 2013) showed that in this part of the lake the depth decreases to  $\sim 1.5$  m. Thus, it is likely that increase in reflectance there is caused by the lake bottom as water in this lake is characterized by high water transparency (the Secchi depth is 7.1 m). In addition, the false colour composite (Fig. 4(e–h)) shows lighter colour in that area. This provides information that this area is not covered by floating/emergent vegetation as then it would be seen in red colour in the False Colour Composite (FCC) (Fig. 4(e–h)).

We suggest using the FCC as a tool to distinguish eutrophic blooming lakes (Fig. 4(e, f)) from lakes with low phytoplankton biomass (Fig. 4(g, h)). Blooming lakes appear blue in the FCC due to higher reflectance in band B3, whereas oligo-mesotrophic lakes look almost black as the reflectance in all the bands is very low.

### Retrieval of chlorophyll-a concentration

The concentrations of chlorophyll-a were calculated using various empirical equations (for equa-

tions' expressions see Table 4) and were compared to *in situ* measurements. The performance of 12 empirical equations, relating the chlorophyll-a concentration calculated from an algorithm and chlorophyll-a concentration, were evaluated using the coefficient of determination ( $R^2$ ) and root mean squared error (RMSE).

All the algorithms performed well ( $R^2 \geq 0.5$ ) when used with the toA data (Table 5). The best algorithms found CHL3\_L1C and CHL4a\_L1C had the  $R^2 = 0.8$ , however, the RMSE were high – 21–53  $\mu\text{g L}^{-1}$ . Both algorithms were calculated using the band ratio between the band B5 and B4. Both algorithms, CHL3 and CHL4a performed worse when used with the boA data ( $R^2 = 0.45$ –0.5). The best algorithm CHL8\_L2A ( $R^2 = 0.76$ ) obtained using the boA data had the lowest RMSE = 9  $\mu\text{g L}^{-1}$ . It was calculated using three bands, B4, B5 and B8A and the empirical equation developed for cyanobacteria rich waters (Matthews et al., 2012).

We derived empirical equations for chlorophyll-a concentration calculation in eutrophic lakes using the band ratio expressions (Table 6). The best algorithm that used the toA data was the band ratio  $P_2$  of bands B5 and B4 ( $R^2 = 0.79$ ) and it gave a lower RMSE = 7  $\mu\text{g L}^{-1}$  than any tested empirical equation from other studies. The derived empirical equations showed similar results when used with the boA data. The derived

Table 5. Descriptive statistics of the relationships between the *in situ* chlorophyll-a concentrations and chlorophyll-a concentrations calculated using the empirical equations (see Table 4 for empirical equations' expressions). The  $p$ -value lower than 0.05 shows the statistically significant relationships with 95% confidence interval

Algorithm code	$R^2$	RMSE, $\mu\text{g L}^{-1}$	Number of samples	$p$ -value (confidence interval 95%)	Algorithm code	$R^2$	RMSE, $\mu\text{g L}^{-1}$	Number of samples	$p$ -value, (confidence interval 95%)
CH1_L1C	0.50	33	8	0.05	CH1_L2A	0.65	17	8	0.02
CH2a_L1C	0.79	26	8	0.00	CH2a_L2A	0.58	22	8	0.03
CH2b_L1C	0.79	32	8	0.00	CH2b_L2A	0.58	27	8	0.03
CHL3_L1C	0.80	21	8	0.00	CHL3_L2A	0.50	42	8	0.05
CHL4a_L1C	0.80	53	8	0.00	CHL4a_L2A	0.45	53	8	0.07
CHL4b_L1C	0.60	53	8	0.02	CHL4b_L2A	0.61	53	8	0.02
CHL5_L1C	0.50	53	8	0.05	CHL5_L2A	0.65	53	8	0.02
CHL6_L1C	0.77	24	8	0.00	CHL6_L2A	0.59	29	8	0.03
CHL7_L1C	0.79	19	8	0.00	CHL7_L2A	0.58	18	8	0.03
CHL8_L1C	0.72	18	8	0.01	CHL8_L2A	0.76	9	8	0.00
CHL9_L1C	0.79	18	8	0.00	CHL9_L2A	0.58	31	8	0.03

Table 6. Descriptive statistics of the linear regression models for chlorophyll-a concentration calculation in eutrophic lakes derived in this study using the band ratio expressions ( $P_1$ – $P_8$ , as described in Table 4) and *in situ* chlorophyll-a data. The  $p$ -value lower than 0.05 shows the statistically significant relationships with 95% confidence interval

Level	$P$	Band ratio	$R^2$	RMSE, $\mu\text{g L}^{-1}$	Number of samples	$p$ -value, (confidence interval 95%)	Equation
L1C	P1_L1C	$P_1 = R_{B5} - \frac{R_{B4} + R_{B6}}{2}$	0.50	11	8	0.05	$Chla = 1933.48 \times P_1 + 26.44$
L1C	P2_L1C	$P_2 = -\frac{R_{B5}}{R_{B4}}$	0.79	7	8	0.00	$Chla = 127.63 \times P_2 - 99.2$
L1C	P4B_L1C	$P_{45} = R_{55}$	0.57	10	8	0.03	$Chla = 1402.91 \times P_{4B} - 30.49$
L1C	P6_L1C	$P_6 = (R_{B4}^{-1} - R_{B5}^{-1}) \times R_{B6}$	0.77	7	8	0.00	$Chla = 180.15 \times P_6 + 28.959$
L1C	P8_L1C	$P_8 = R_{B5} - R_{B4} - ((R_{B8A} - R_{B4}) \times (-664) / (865 - 664))$	0.70	8	8	0.01	$Chla = 2223.18 \times P_8 + 24.03$
L2A	P1_L2A	$P_1 = R_{B5} - \frac{R_{B4} + R_{B6}}{2}$	0.65	9	8	0.02	$Chla = 1830.31 \times P_1 + 8.32$
L2A	P2_L2A	$P_2 = -\frac{R_{55}}{R_{B4}}$	0.58	10	8	0.03	$Chla = 30.43 \times P_2 + 2.4$
L2A	P4B_L2A	$P_{4B} = R_{B5}$	0.43	11	8	0.08	$Chla = 919.71 \times P_{4B} + 12.46$
L2A	P6_L2A	$P_6 = (R_{B4}^{-1} - R_{B5}^{-1}) \times R_{B6}$	0.59	10	8	0.03	$Chla = 112.96 \times P_6 + 38.13$
L2A	P8_L2A	$P_8 = R_{B5} - R_{B4} - ((R_{B8A} - R_{B4}) \times (-664) / (865 - 664))$	0.77	7	8	0.00	$Chla = 1726.50 \times P_8 + 18.29$

equations based on the  $P_2$ ,  $P_{4B}$ ,  $P_6$ , and  $P_8$  band ratios showed better results (higher  $R^2$ ) when used with the toA data than boA data. Only the band B5 peak height algorithm (Toming et al., 2016) showed a better ( $R^2 = 0.65$ ) result when used with the boA data than with the toA data ( $R^2 = 0.5$ ).

The best algorithm that used the boA data was three band algorithm P8\_L2A ( $R^2 = 0.77$ ) that gave the RMSE = 7  $\mu\text{g L}^{-1}$ . The derived equations showed close relationships between the *in situ* chlorophyll-a concentrations and the calculated by algorithm chlorophyll-a concentrations (Fig. 5).

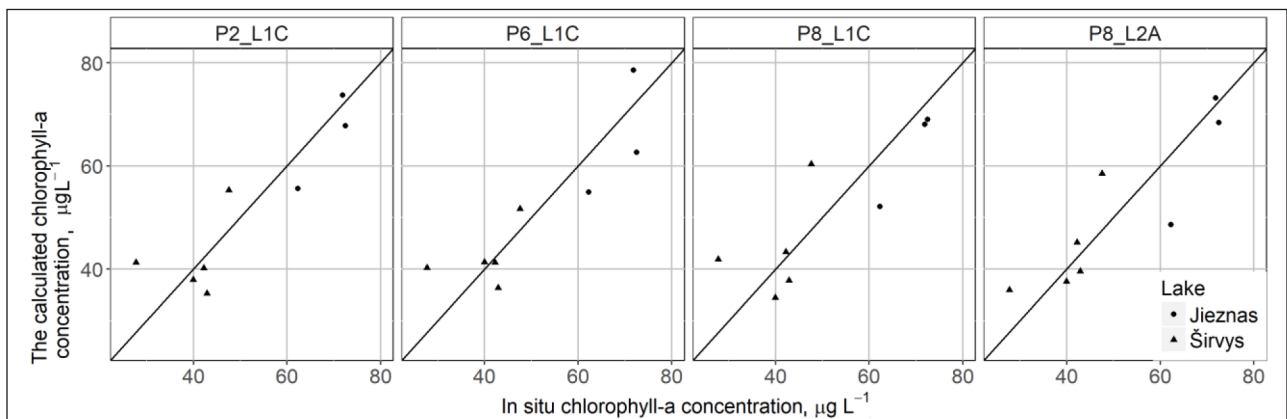


Fig. 5. The graphs of the relationship between the in situ data and chlorophyll-a concentration calculated using the derived empirical equations. The line indicates the 1:1 line

The temporal chlorophyll-a concentration variation in Lake Jieznas was best depicted by the P8\_L1C algorithm as concentration rose on the 14th of August as well as the *in situ* concentration (Fig. 6). In Lake Širvys the same algorithm depicted the chlorophyll-a concentration variation well; however, the drop of chlorophyll-a concentration on the 4th of August was not observed in the chlorophyll-a concentration calculated by the algorithm. The decrease on the 4th of August was depicted better by the algorithm P8\_L2A (Fig. 6).

A larger dataset of *in situ* data is needed to validate algorithms and derive more accurate empirical equations, thus, providing the potential to use satellite data to estimate approximate chlorophyll-a concentrations for a large number of eutrophic lakes.

## DISCUSSION AND CONCLUSIONS

The top of atmosphere (toA) reflectance of lakes varies between 0 in the shortwave infrared wavelength band, B10, and 0.155 in the blue bands, B1-B2, while bottom of atmosphere reflectance varies from 0 to 0.074. Reflectances are higher in the eutrophic lakes than in the oligo-mesotrophic lakes due to the presence of optically active constituent – chlorophyll-a pigment. Reflectances from the studied eutrophic lakes are very low and similar to the reflectances of lakes in Estonia and Sweden studied by Kutser et al. (2016) that are characterized by high chlorophyll-a and organic

matter (CDOM) concentrations.

Reflectance from a lake characterized by a relatively large Secchi depth (up to 7.1 m) can be influenced by the lake bottom that causes reflectance to increase to the values seen in the eutrophic lakes (as seen in the case of Lake Šventas). The Sentinel-2 imagery was shown to be useful in retrieving water depths up to 2.5 m in oligotrophic lake characterized by 4.2 m Secchi depth (Dörnhöfer et al., 2016).

We tested empirical equations derived in other studies to check their potential to be used without adapting them to local conditions as often the available *in situ* database is small. The empirical equations were able to explain a large part of the chlorophyll-a concentration variation in eutrophic lakes ( $R^2 = 0.45-0.8$ ); however, mostly the RMSEs were high – 17–53  $\mu\text{g L}^{-1}$ . The empirical equation CHL8\_L2A ( $R^2 = 0.76$ ) gave a lower RMSE = 9  $\mu\text{g L}^{-1}$  as it was originally derived for the cyanobacteria-rich waters (Matthews et al., 2012).

The temporal changes in reflectance data are consistent with *in situ* chlorophyll-a concentration data. The application of atmospheric correction largely reduces reflectance at the shorter wavelengths, especially in the blue band, and expose the peak of the red-edge (B5, 705 nm) band visible in the eutrophic lake spectra. However, the algorithms tested in this study often showed better results when used with the toA data rather than with boA data. The study by Toming et al. (2016) also found that the band ratio algorithms

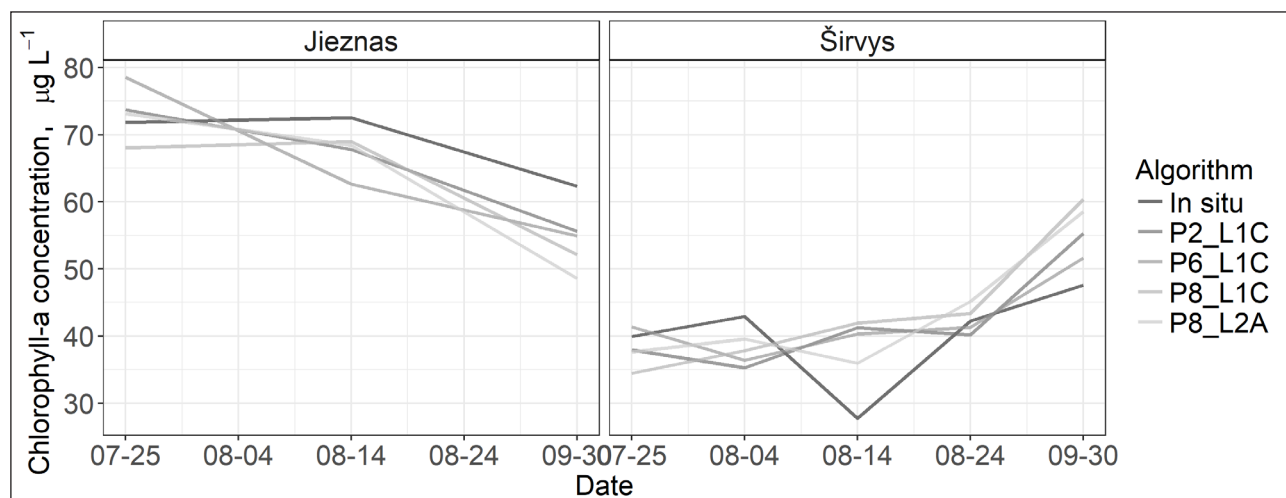


Fig. 6. The *in situ* data and four best linear regression equations ( $R^2 \geq 0.7$ ) derived in this study for chlorophyll-a concentration calculation

perform better when used with the toA data rather than with boA data. We used the Sen2Cor tool that is available within the ESA SNAP tool, however, it was not specifically designed for water bodies (Toming et al., 2016), and therefore, it likely has an error that impedes more accurate retrieval of chlorophyll-a concentration in lakes. Other atmospheric correction algorithms, such as, ACOLITE (Vanhellemont, Ruddick, 2016) need to be tested in order to obtain better results.

We derived empirical equations for eutrophic lakes based on the data from Lake Jieznas and Lake Širvys that are characterized by the recurrent cyanobacteria blooms. The derived equations: based on the band B5 and B4 ratio and based on the three bands B4, B5, and B8A, performed the best in retrieving the chlorophyll-a concentration ( $R^2 = 0.77-0.79$ ,  $RMSE = 7 \mu\text{g L}^{-1}$ ). A larger *in situ* measurement database could improve the results and help to evaluate the algorithm performance in other eutrophic lakes in Lithuania.

The first attempts to calculate chlorophyll-a concentration in eutrophic Lithuanian lakes using the data received from the Sentinel 2 Multispectral Imager sensor show good results as the changes in reflectance, caused by the changes in chlorophyll-a concentration, can be estimated from satellite images. The use of a better atmospheric correction algorithm over the lakes and a larger *in situ* chlorophyll-a concentration database could help to derive more accurate empirical equations for chlorophyll-a concentration retrieval in eutrophic lakes in Lithuania. The algorithms could contribute to the optimization of a monitoring system and improved assessment of the water quality.

Received 30 April 2018

Accepted 24 June 2018

## REFERENCES

- Balevičienė J., Balevičius A., Ciūnys E., Stanevičius V., Vaitkus G., Valiuškevičius G., Kalytė D., Ūselytė R., Šalčiūnienė K. 2009. *Restauruotinių Lietuvos ežerų nustatymas ir preliminarus restauravimo priemonių patinkimas šioms ežerams, siekiant pagerinti jų būklę*.
- Bbe moldanke. 2018. *AlgaeLabAnalyser*. <https://www.bbe-moldaenke.de/en/products/chlorophyll/details/algaelabanalyser.html> Accessed on 2018-06-11.
- Bresciani M., Giardino C., Stroppiana D., Pilkaitytė R., Zilius M., Bartoli M., Razinkovas A. 2012. Retrospective analysis of spatial and temporal variability of chlorophyll-a in the Curonian Lagoon. *Journal of Coastal Conservation*. 16(4): 511–519.
- Bresciani M., Adamo M., De Carolis G., Matta E., Pasquariello G., Vaičiūtė D., Giardino C. 2014. Monitoring blooms and surface accumulation of cyanobacteria in the Curonian Lagoon by combining MERIS and ASAR data. *Remote Sensing of Environment*. 146: 124–135.
- Drusch M., Del Bello U., Carlier S., Colin O., Fernandez V., Gascon F., Hoersch B., Isola C., Laberinti P., Martimort P. et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*. 120: 25–36.
- Dörnhöfer K., Göritz A., Gege P., Pflug B., Oppelt N. 2016. Water constituents and water depth retrieval from Sentinel-2A – a first evaluation in an oligotrophic lake. *Remote Sensing* 8(11): 941.
- Duan H., Zhang Y., Zhang B., Song K., Wang Z. 2007. Assessment of chlorophyll-a concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. *Environmental Monitoring and Assessment*. 129(1–3): 295–308.
- ESA. 2014. *Copernicus Operations Secured until 2021*. [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Copernicus\\_operations\\_secured\\_until\\_2021](http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Copernicus_operations_secured_until_2021) Accessed on 17.05.2018.
- ESA. 2017. *New release of Sen2Cor 2.4.0 is available*. <http://step.esa.int/main/new-release-ofsen2cor-2-4-0-is-available/> Accessed on 17.05.2018.
- ESA. 2018. *SNAP – ESA Sentinel Application Platform v6.0*. <http://step.esa.int/main/> Accessed on 17.05.2018.
- Giardino C., Bresciani M., Pilkaitytė R., Bartoli M., Razinkovas A. 2010. In situ measurements and satellite remote sensing of case 2 waters: first results from the Curonian Lagoon. *Oceanologia*. 52(2): 197–210.
- Giardino C., Bresciani M., Cazzaniga I., Schenk K., Rieger P., Braga F., Matta E., Brando V. E. 2014. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of Lake Garda. *Sensors*. 14(12): 24116–24131.
- Hovis W., Clark D., Anderson F., Austin R., Wilson W., Baker E., Ball D., Gordon H., Mueller J., El-Sayed S. et al. (1980). Nimbus-7 Coastal Zone Colour Scanner: system description and initial imagery. *Science*. 210(4465): 60–63.
- Hunter P., Tyler A., Willby N., Gilvear D. 2008. The spatial dynamics of vertical migration by

- Microcystis aeruginosa in a eutrophic shallow lake: A case study using high spatial resolution time-series airborne remote sensing. *Limnology and Oceanography*. 53(6): 2391–2406.
15. Jiao H., Zha Y., Gao J., Li Y., Wei Y., Huang J. 2006. Estimation of chlorophyll-a concentration in Lake Tai, China using in situ hyperspectral data. *International Journal of Remote Sensing*. 27(19): 4267–4276.
  16. Koreivienė J., Kasperovičienė J., Karosienė J., Vitonytė I. 2012. Ežerų „žydėjimas“. Ką turime žinoti apie „žydinčius“ ežerus ir juose slypinčius pavojus. *Mokslas ir gyvenimas*. 2012. Nr. 10.
  17. Kilkus K., Stonevičius E. 2012. *Lietuvos vandenių geografija*. Vilniaus universitetas.
  18. Kilkus K. 2013. *Ežerai*. 119 Lietuvos ežerų. Alma littera.
  19. Kutser T., Paavel B., Verpoorter C., Ligi M., Soomets T., Toming K., Casal G. 2016. Remote sensing of black lakes and using 810 nm reflectance peak for retrieving water quality parameters of optically complex waters. *Remote Sensing*. 8(6): 497.
  20. LAND 69-2005 „Vandens kokybė. Biocheminių parametrų matavimas. Spektrometrinis chlorofilo-a koncentracijos nustatymas“, patvirtintas 2005-12-28 Lietuvos Respublikos aplinkos ministro įsakymu Nr. D1-648.
  21. Matthews M. W., Bernard S., Robertson L. 2012. An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment*. 124: 637–652.
  22. McClain C. R. 2009. A decade of satellite ocean color observations. *Annual Review of Marine Science*. 1: 19–42.
  23. Mishra D. R., Ogashawara I., Gitelson A. A. 2017. *Bio-optical Modeling and Remote Sensing of Inland Waters*. Elsevier.
  24. Moses W. J., Gitelson A. A., Berdnikov S., Povazhnyy V. 2009. Satellite estimation of chlorophyll-a concentration using the red and NIR bands of MERIS – The Azov sea case study. *IEEE Geoscience and Remote Sensing Letters*. 6(4): 845–849.
  25. Müller-Wilm U. 2016. *Sentinel-2 MSI—Level-2A Prototype Processor Installation and User Manual*. Telespazio VEGA Deutschland GmbH: Darmstadt, Germany.
  26. Müller-Wilm U. 2017. *Sen2Cor Configuration and User Manual*. Telespazio VEGA Deutschland GmbH: Darmstadt, Germany.
  27. Philipson P., Kratzer S., Ben Mustapha S., Strömbeck N., Stelzer K. 2016. Satellite-based water quality monitoring in Lake Vänern, Sweden. *International Journal of Remote Sensing*. 37(16): 3938–3960.
  28. Savadova K., Koreivienė J., Sivonen K., Kasperovičienė J., Suurnakki S., Karosienė J., Wahlsten M., Vitonytė I. 2015. Variation of bloom forming cyanobacteria and microcystins in shallow hypertrophic lake. *European Journal of Phycology*. 50: 204–204.
  29. Savadova K., Koreivienė J., Karosienė J., Kasperovičienė J., Vitonytė I. 2016. Recovery of shallow lake from cultural eutrophication in the light of climate change. Algae in anthropogenically transformed ecosystems. *35th International Conference of the Polish Phycological Society*.
  30. Toming K., Kutser T., Laas A., Sepp M., Paavel B., Nõges T. 2016. First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sensing*. 8(8): 640.
  31. Tranvik L. J., Downing J. A., Cotner J. B., Loisel S. A., Striegl R. G., Ballatore T. J., Dillon P., Finlay K., Fortino K., Knoll L. B. et al. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*. 54(6 part 2): 2298–2314.
  32. Wetzel R. G. 2001. *Limnology: lake and river ecosystems*. Gulf Professional Publishing.
  33. Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <http://ggplot2.org> Accessed on 17.05.2018.
  34. Zambrano-Bigiarini M. 2017. *hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series*. <https://cran.rproject.org/web/packages/hydroGOF/index.html> Accessed on 17.05.2018.

Dalia Grendaitė, Edvinas Stonevičius, Jūratė Karosienė,  
Kenija Savadova, Jūratė Kasperovičienė

### CHLOROFILO-A KONCENTRACIJOS NUSTATYMAS LIETUVOS EUTROFINIUOSE EŽERUOSE IŠ SENTINEL-2 DUOMENŲ

#### *Santrauka*

Vandens telkiniai, naudojami turizmui, žvejybai, industrinėms reikmėms, yra svarbi buveinė augalijai ir gyvūnijai. Vandens telkinių monitoringo sistemos sukūrimas yra pirmas būtinas žingsnis siekiant pagerinti vandens kokybę. Palydovinių duomenų panaudojimas leidžia stebėti daug ežerų didelėje teritorijoje. Šio darbo tikslas – įvertinti Sentinel-2 multispektrinio instrumento duomenų panaudojimo galimybes fitoplanktono žydėjimų pokyčiams ežeruose stebėti ir chlorofilo-a koncentracijai apskaičiuoti. Chlorofilo-a koncentracija yra fitoplanktono biomasės indikatorius, leidžiantis įvertinti ežero trofinę būseną. Tyrimas buvo atliktas remiantis keturių Lietuvos ežerų (dviejų eutrofinių ir dviejų oligomezotrofinių) duomenimis. Dėl fitoplanktono at-

spindėjimas yra didesnis nuo eutrofinių ežerų nei nuo oligomezotrofinių, kuriuose fitoplanktono beveik nėra. Šiame tyrime buvo testuojamos empirinės lygtys chlorofilo-a koncentracijai nustatyti. Dauguma lygčių parodė gerus rezultatus ( $R^2 = 0,5-0,8$ ), tačiau pasižymėjo didelėmis klaidomis (RMSE = 17–53  $\mu\text{g L}^{-1}$ ). CHL8\_L2A lygtis ( $R^2 = 0,76$ ) pasižymėjo mažesnėmis klaidomis, RMSE = 9  $\mu\text{g L}^{-1}$ . Taip pat šiame tyrime buvo išvestos empirinės lygtys eutrofiniams Lietuvos ežerams. Geriausius rezultatus pateikė lygtis, paremta Sentinel-2 kanalų B5/B4 santykiu, ir lygtis, paremta trijų kanalų B4, B5 ir B8A duomenimis ( $R^2 = 0,77-0,79$ ), jos pasižymėjo mažesnėmis klaidomis (RMSE = 7  $\mu\text{g L}^{-1}$ ) nei empirinės lygtys iš kitų tyrimų. Didesnė *in situ* chlorofilo-a koncentracijos duomenų bazė padėtų pagerinti gautus algoritmus. Pirmasis palydovinių duomenų panaudojimo tyrimas ežerų vandens kokybei tirti Lietuvoje parodė gerus rezultatus, kadangi atspindėjimo nuo ežerų pokyčiai yra nulemti chlorofilo-a pokyčių ežeruose.

**Raktažodžiai:** ežerai, vandens kokybė, chlorofilas-a, distanciniai metodai, Sentinel-2