Spatial variability and detection levels for Chlorophyll-a estimates in high latitude lakes using Landsat imagery

Lisboa, Filipe[1]; Brotas, Vanda[1,2]; Santos, Filipe Duarte[3]; Kuikka, Sakari[4]; Kaikkonen, Laura[5]; Maeda, Eduardo[4]

(1) MARE – Marine and Environmental Sciences Centre, Universidade de Lisboa, Portugal

(2) PML – Plymouth Marine Laboratory, Plymouth, United Kingdom

(3) CCIAM-CE3C, Faculty of Sciences, University of Lisbon.

(4) Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, 00014 University of Helsinki, Finland

(5) Department of Geosciences and Geography, P.O. Box 68, FI-00014 University of Helsinki, Finland

ABSTRACT

Monitoring lakes in high-latitude areas can provide a better understanding of freshwater systems sensitivity and accrete knowledge on climate change impacts. Phytoplankton are sensitive to various conditions: warmer temperatures, earlier ice-melt and changing nutrient sources. Satellite imagery can monitor algae biomass over large areas. The detection of chlorophyll a (chl-a) concentrations in small lakes is hindered by the low spatial resolution of conventional ocean colour satellites. The short time-series of the newest generation of space-borne sensors (e.g. Sentinel-2) is a bottleneck for assessing long-term trends. Although previous studies have evaluated the use of high-resolution sensors for assessing lakes' chl-a, it is still unclear how the spatial and temporal variability of chl-a concentration affect the performance of satellite estimates. We discuss the suitability of Landsat (LT) 30-m resolution imagery to assess lakes' chl-a concentrations under varying trophic conditions, across extensive high-latitude areas in Finland. We use in situ data obtained from field campaigns in 19 lakes and generate remote sensing estimates of chl-a, taking advantage of the long-time span of the LT 5 and 7 archives, from 1984 to 2017. Our results show that linear models based on LT data can explain approximately 50 % of the chl-a interannual variability. However, we demonstrate that the accuracy of the estimates is dependent on the lake's trophic state, with models performing in average twice as better in lakes with higher chl-a concentration (> 20 µg/l) in comparison with less eutrophic lakes. Finally, we demonstrate that linear models based on LT data can achieve high accuracy ($R2 = 0.9$; p-value < 0.05) in determining lakes' annual mean chl-a concentration, allowing the mapping of the trophic state of lakes across large regions. Given the long time-series and high spatial resolution, LT-based estimates of chl-a provide a tool for assessing the impacts of environmental change.

1) INTRODUCTION

High latitude lakes are considered sentinels for a climate in change, due to their coupling with ice phenology, response to changes in humidity and precipitation patterns, and the high sensitivity of photosynthetic organisms to changing temperatures (Kraemer et al., 2017). Given Finland's geographical location, the country can experience a faster warmer with respect to the global average. Warming in Finland is estimated to be roughly 50 % higher than the global average, which corresponds to a 0.15-0.20 °C increase per decade (Mikkonen et al., 2015). The effects of eutrophication are still common in Finland with algal blooms being exacerbated by climate change (Malve et al., 2015).

Studies have shown the importance of studying lakes as sentinels for climate change, including variables such as water level, ice phenology, chemical variables, dissolved organic carbon and oxygen (Adrian et al., 2009). Additionally, biota's growth rates, abundance, and species composition can be considered an indicator of climate change (Adrian, 2009; Rühland et al., 2008). Nevertheless, monitoring water bodies has been difficult mainly because *in situ* methods are very localised and cannot always be performed routinely. *In situ* datasets might also contain data gaps and data gathering is not coherent across different regions and field sampling is marked more expensive than using satellite estimates. For these reasons, the evolving methods of satellite monitoring have raised popularity in a broad spectrum from bio-geophysics to marine biology.

Indeed, a powerful way to have synoptic views of these changes is to monitor them from space. One important indicator of change is phytoplankton abundance which can be proxied from the chlorophyll *a* concentration at the water surface (chl *a*). Satellites have now been collecting decades of remotely sensed optical imagery over large areas. Due to the characteristic reflectance of chl *a* pigments, we can estimate concentrations from such images. Earth Observation satellites have been designed and launched with the specific purpose of studying phytoplankton from space with sensors specifically suited to the assessment of aquatic ecosystems. For example, NASA's Sea-viewing Wide Field-of-View Sensor (SeaWiFS) launched in August 1997 onboard the SeaStar satellite, collected data until 2010 at a resolution of 1.1 km. The Terra and Aqua Satellites both collect data through the 36-band MODIS sensor at wavelengths between 0.41 and 14.24 μm applicable to extensive ocean colour algorithms (Gregg et al., 2017). Recently, the European Space Agency (ESA) has launched the Sentinel 3 constellation as part of the Copernicus Programme. Onboard the two sentinels, the Ocean and Land Colour Instruments (OCLI) are collecting data at wavelengths from 0.4 μm to 1.02 μm at 21 spectral bands, allowing for algal pigment discrimination and the further development of phytoplankton functional types characterisation from space (IOCCG, 2014). The resolution of the new OCLI sensor is 300 m, which is not suitable for studying lakes smaller than 600 m.

The use of the above-mentioned satellites hampers the study of Finland's small lakes due to their coarse resolution in relation to the size of the lakes — excluding Lake Saimaa, the average size of lakes is 111 km^2. Additionally, consistent data from the Ocean Colour missions' dates to 1997, which provides a relatively short time series. Given these limitations, the high-resolution Landsat satellites (LT) can be advantageous to use on

small water bodies. LT satellites have collected one of the most comprehensive datasets, with almost 40 years of observations. LT 5 was launched in March 1984 providing data for the following 28 years. LT 5 carried the 7-band Thematic Mapper (TM) multispectral sensor. LT 7, launched in April 1999, is still in operation and carries the 8-band Enhanced Thematic Mapper Plus (ETM+). Exploring this archive for assessing chl *a* concentrations offers an excellent opportunity to evaluate how changes in phenology have occurred in previous decades, due to both climate change and eutrophication. Furthermore, the high spatial resolution of LT data (30 m) allows the assessment of smaller lakes or lakes with complex shapes, minimizing the influence of surrounding land vegetation.

Previous studies have assessed the relationship between satellite imagery and chl *a* on lakes (Giardino et al., 2001; Thiemann et al., 2000). Such studies focused on studying single lakes, under homogeneous environmental conditions. Östlund et al. (2001) analysed LT imagery and samples from a transect on Lake Erken, Sweden achieving R^2 = 0.9. Isenstein, Trescott, and Park (2014) used the band ratio of B2/B1 (green/blue) for chl *a* assessment on Lake Champlain, USA, reporting a R of 0.82. Having such space-derived estimates has been proved useful to study the changing patterns (Giardino et al., 2010). Nevertheless, the challenges increase with smaller lakes. For example, a set of 131 small lakes (averaging 100 ha) in Maine revelled an R^2 of 0.25 (Boucher et al., 2018).

The models derived in these previous studies were limited by temporal and spatial sampling and therefore did not account for temporal or between lakes variations. To take full advantage of the long LT time series, models should robustly incorporate data from both LT5-TM and LT7-ETM+. Vincent et al. (2004) argued that a model developed for LT5 cannot be applied to LT7 images. Nonetheless, their methods involved the use of dark-object-subtracted radiance instead of surface reflectance. Hence, further studies are still needed to fully explore the combined use of LT5-TM and LT7-ETM+, using surface reflectance derived from state-of-the-art atmospheric correction techniques (Masek et al., 2006).

The importance of our study is to demonstrate whether chl *a* assessments using surface reflectance retrieved from LT sensors can be generalised for these different circumstances. In other worlds, we aim to understand how different chl *a* might affect LT reflectance, as a preliminary approach to develop advanced bio-optical models for inland and coastal waters. Such models, in turn, are an essential tool not only for global water quality monitoring, but to also study temporal dynamics of phytoplankton across large areas, allowing the identification of shifts in the temporal characteristics of phytoplankton blooms related to climate variability and land-use changes (IOCCG, 2018). We hypothesize that lakes may be a valuable source of information to monitor the trends over time, as well as local gradients created by climate change.

The water-leaving signal is especially complex for the case of inland waters due to the high concentrations of colored dissolved organic materials (CDOM), tripton, floating and submerged aquatic vegetation, suspended matter (e.g clay) and finally phytoplankton. The tripton component, composed of detritus and minerals from dead phytoplankton and decaying organic matter, varies significantly in optical signatures and concentration

ranges. High concentrations of humic and fulvic acids from surrounding vegetation are leached into lakes and reduce the light availability in the blue spectral region in boreal and brackish waters (IOCCG, 2018 – chapter 5), both the case of most of Finland's lakes. All these factors prove that the lakes considered in this paper are one of the most difficult cases for bio-optical models of phytoplankton proxying from space.

Our objective is to better understand spatiotemporal factors affecting the relationship between LT reflectance data and surface chl *a* concentration over high latitude lakes. We drive our research by asking: 1) how does the performance of models based on LT data vary with different lakes' trophic state? 2) how does the performance of models based on LT data vary at different temporal scales? 3) Can the LT data archive be used to assess the spatial and temporal variability of lakes average chl *a* levels across large areas?

2. Material and methods

2.1) Study area and in situ data

Figure 1 depicts the 20 sampling locations and the 19 sampled lakes used as ground-truth for our model. Lake Saimaa, along the Salpausselkä system, is the only Lake with two sampling spots. The lakes included in this study belong to the intensively monitored lakes of the national routine monitoring network of Finland. Water samples for turbidity were taken from 1 m depth, and chl *a* was determined from a composite sample of 0 – 2 m depth. The concentration of chl *a* was measured with a spectrophotometer after extraction with hot ethanol (ISO 10260, GF/C filter). Turbidity, measured in Formazin Nephelometric Units, was determined by the nephelometric method (EN 27027), based on measurement of light (860 nm) scattered within a 90° angle from a beam directed at the water sample, with formazine used as a standard matching solution.

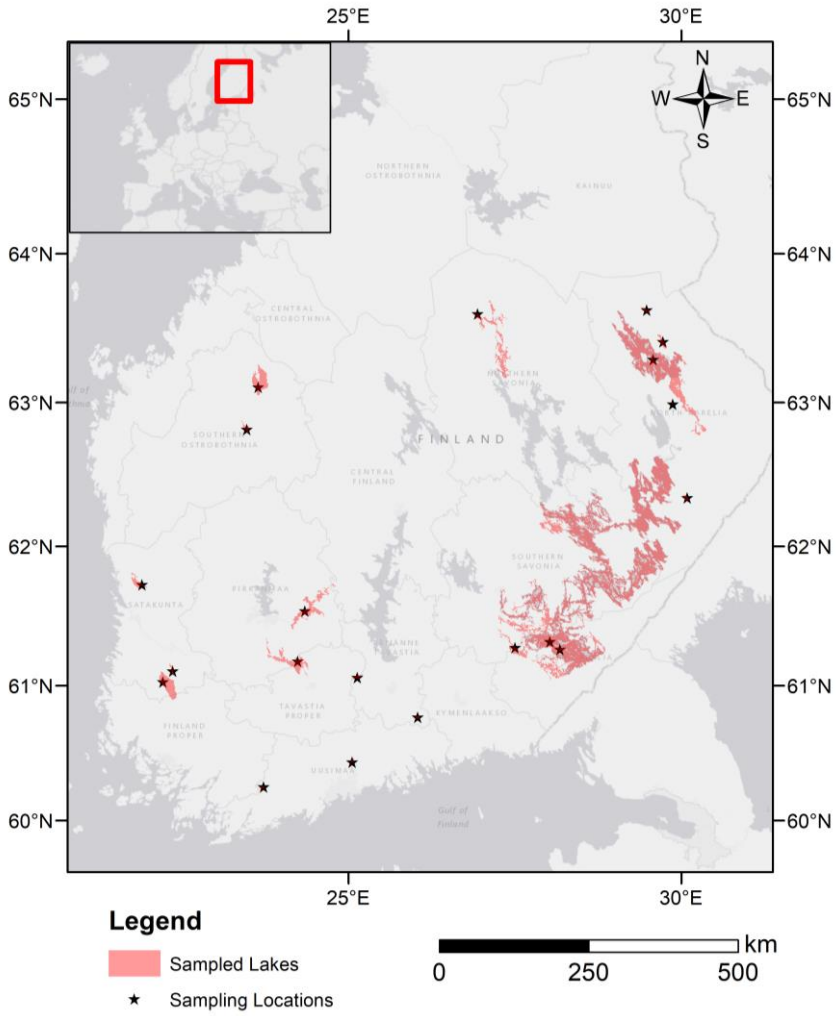

Figure 1 Distribution of the 19 sampled lakes and the 20 sampling locations. Two sampling locations were defined for lake Saimaa, the biggest lake in Finland.

2.2. Landsat Imagery

We used the Surface Reflectance Collection — Tier 1 from LT 5 and 7 provided by the United States Geological Survey (USGS) to the Google Earth Engine archives. Both satellites' lifespans include all our *in situ* sampling dates. These products are processed through the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction — a radiative transfer model that includes water vapor, ozone, geopotential height, aerosol optical thickness and digital elevation (Masek, 2006). Only images with the highest quality factor were used summing a total of 3865. Images are predominantly from summer months; no images were selected from December and January due to high cloud coverage and low – or null – sun elevation. The dates of the images span from February 1999 to April 2017 (错误!未找到引用源。). We have used bands 1-5 from the LT collection corresponding to all visible bands, Near Infrared (NIR-band 4) and Shortwave Infrared (SWIR -band 5).

2.3 Data methods

We evaluated multivariate regression models to identify which LT bands, or band combinations, are more suitable for estimating chl *a*. We also evaluated how time differences between the date of chl *a* field samples and the date of LT images acquisition affect the performance of the models.

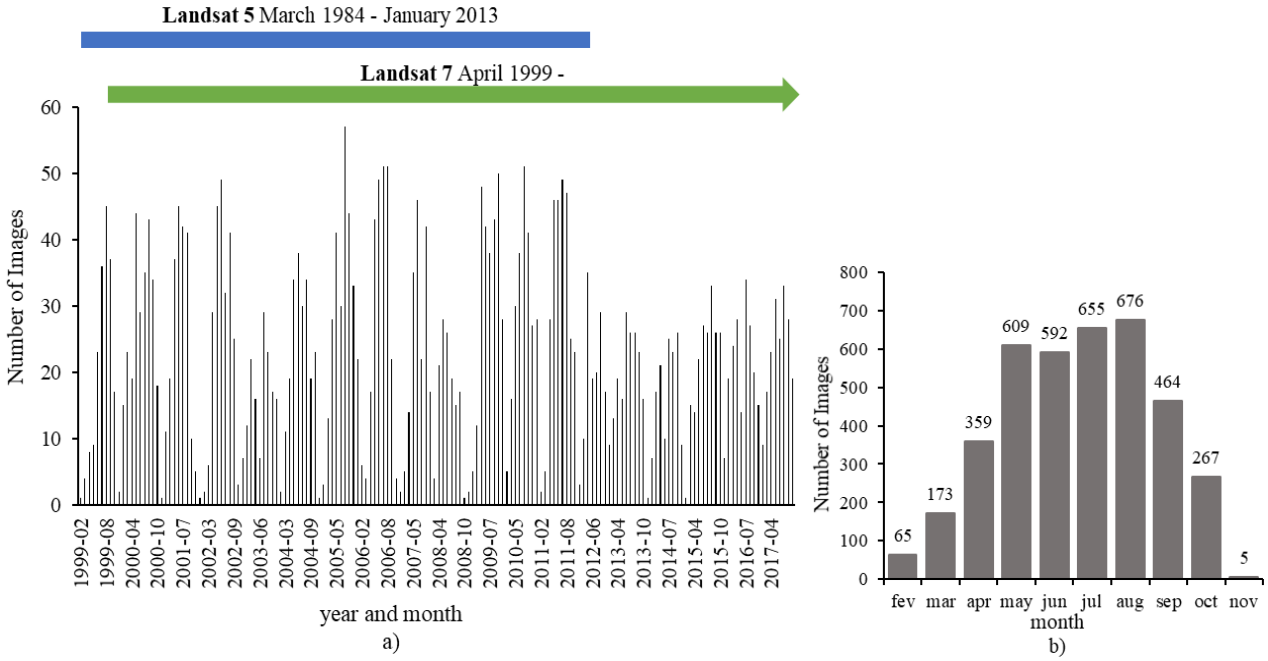

Figure 2 Number of images used for this study throughout the whole imagery collection a) and the distribution per month b).

Our code was developed in Google Earth Engine (Gorelick et al., 2017), which provides the entire dataset from LT imagery and the possibility to define the regions of interest according to our specific needs. The collection of LT data into band reflectance time series over the lakes was carried through two main stages: quality assurance of the images and matching of spatially averaged water reflectances with the *in situ* samples.

Next, we gathered and filtered the image collection. The filter removed cloud cover, cloud shadowing and ice pixels, all masked out based on the C Function of Mask (*cfmask)* algorithm described by Foga et al., 2017. Nevertheless, the *cfmask* cannot account for all imagery artefacts. To solve the issue, we added a normalised difference chlorophyll index (NDCI) to the resulting LT collection. We calculated the 25th and 75th percentiles from the resulting index distribution, thus eliminating pixels with NDCI outside the interquartile range (IQR) and bullet-proofing the *cfmask* filtering.

Next, we spatially averaged the resulting valid pixels per image, as laid down in the area buffer defined in section 0. We gathered data according to the temporal averaging of images collections within ± n days from the date of *in situ* collection. Previous validation studies indicated an ideal time gap between satellite overpass and field measurement of ± 1-2 days (Stadelmann et al., 2001). We have chosen a time gap of maximum ± 1-2 days, from the hour of sampling, to estimate the chl *a* concentration of all the 19 inland water bodies.

All procedures were carried out in GEE and the data exported for further statistical analysis in R (R Development Core Team, 2016) as detailed in section 2.4. For statistical purposes, *in situ* chl *a* data was primarily filtered to remove outliers. To detect these outliers, we used Tukey's method which targets values above and below 1.5 of the interquartile range (IQR). This study also includes an assessment of the relationship between turbidity and chl *a*. We have calculated the overall correlation between both variables and narrowed it down to summer data from the day of the year (doy) 200 to 280.

Defining a permanent water mask.

It is paramount that the spatial average of water-leaving reflectance values is performed over permanent water surfaces. Coarse lake boundaries were defined according to the HydroLAKES dataset from Messager et al., 2016. However, this layer can include small lake islands or areas of intermittent water presence. To further ensure that the imagery corresponds to permanent water surfaces, we intersected the HydroLAKES dataset with the Global Surface Water Data (Pekel et al., 2016). By incorporating these datasets, we ensure that the selected pixels correspond to permanent water surfaces in the period between 1984 and 2017. Our methodology results in the definition of detailed and high-resolution surfaces on which reflectance spatial averaging can be performed. Within these areas, we overlayed a circle, centred around the sampling location, with varying radius (Figure 3). Additionally, to avoid interference from land vegetation and other types of aquatic plants, we have applied a buffer of 50 m away from land when the circle of extraction is above 500 m (Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

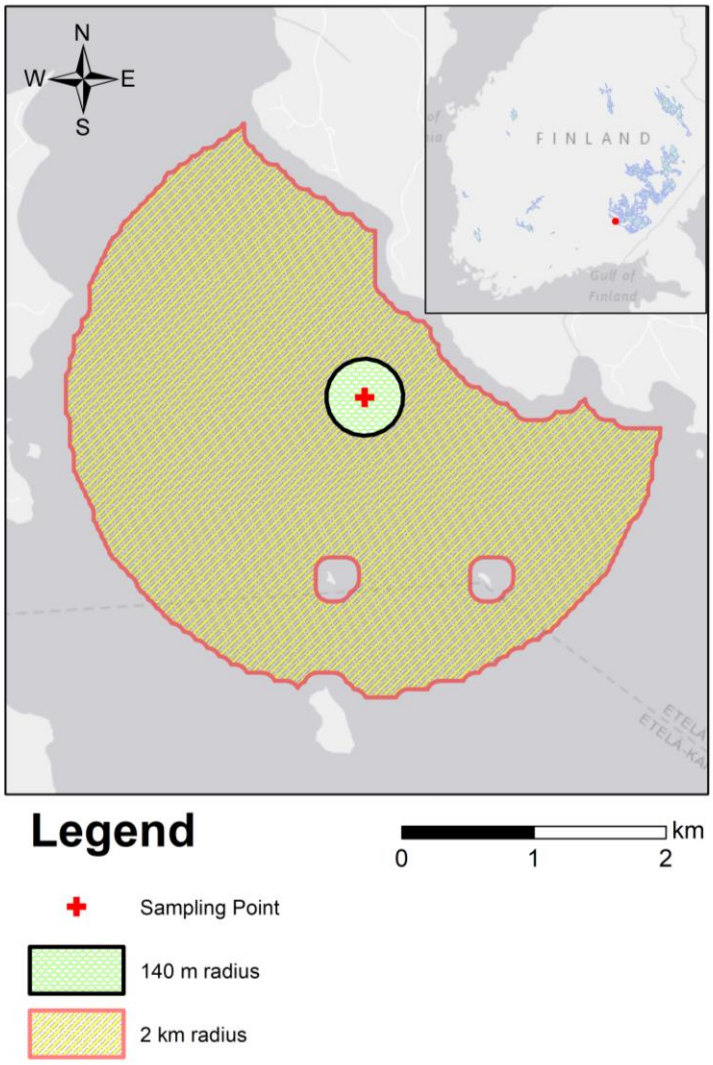

Figure 3 One of the sampled locations at the centre of the area for satellite reflectance band extraction (shaded). Such areas are a variable radius circle around the sampling site and avoiding land.

The shaded areas in Figure 3 depict examples of the surfaces over which the mean reflectance bands were collected. After creating a time-series collection of surface reflectance for each lake, we have generated a database for in situ chl *a* match-up and derived the models.

2.3. Model Selection through Relative Importance Metrics

We assessed the performance of a linear model by merging data from both LT 5 and 7 and all lakes. Multiple linear fits are presented in this analysis to cope with the myriad of band combinations. To have an overview of the variables to incorporate into the model, we used the "Leaps: regression subset selection" exhaustive search, the routine for the best subsets of predictors which uses a linear regression through an efficient branch-and-bound algorithm. The resulting combinations of regressors can be analyzed by either the R^2 of the model or its Schwartz's information criterion or Bayesian information criterion (BIC). The BIC is a criterion for

model selection among a finite set of models (Neath et al., 2012). While fitting models, it is possible to increase the likelihood of a good fitting by adding parameters but doing so may result in overfitting. The BIC resolves this problem by introducing a penalty term for the number of parameters in the model. Lower values of BIC are then preferred for the best model assessment.

Multivariate models for chl *a* were designed using each reflectance band as a regressor. Secondly, we evaluated model performance through relative importance metrics. Relative importance metrics is the quantification of the individual regressor's contributions to a multivariable regression model. For uncorrelated predictors, the multivariate coefficient of determination (R^2) is simply the cumulative result of each R^2 of the single-variate linear model. However, bands of satellite optical imagery are deeply correlated, raising the necessity for a relative importance classification. We have used the *relaimpo* package from R that provides six different methods for assessing relative importance in linear regression (Grömping, 2006). Among these methods, LMG, from Lindeman et al., 1980, is one of the most computationally intensive but highly recommended.

2.4. Spectral indices

Previous studies have demonstrated the usefulness of spectral indices, rather than individual bands, on the detection of vegetation both in land and water. Besides the individual bands, we used two spectral indices to derive our models; the LT adapted version of the normalized difference chlorophyll index, NDCI, and another index, BRG, based on the relative difference of blue and red bands relative to the green band.

The NDCI is given by:

$$NDCI = \frac{\rho_{NIR\ (B4)} - \rho_{red\ (B3)}}{\rho_{NIR\ (B4)} + \rho_{red\ (B3)}} \qquad [eq1]$$

where ρ is the reflectance of the specific band.

The BRG model was adapted from Brivio et al., 2001 who achieved an $R^2 = 0.818$ in Lake Garda, Italy. In such a model, chl *a* is directly proportional to the reflectance index given by:

$$BRG = \frac{\rho_{blue\ (B1)} - \rho_{red\ (B3)}}{\rho_{green\ (B2)}} \qquad [eq2]$$

The same BRG index led to satisfactory results for other small water bodies as in the Malilangwe Reservoir, Zimbabwe with $R^2 = 0.81$ (Dalu et al., 2015).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

2) RESULTS

3.1 Daily Aggregated Data

Model and Variable Selection

Of all models derived through intra-annual data, the one that performed best was achieved with a buffer of 700 m and a time-window of ± 1 day from the satellite acquisition. For this configuration, the results of the LEAPS

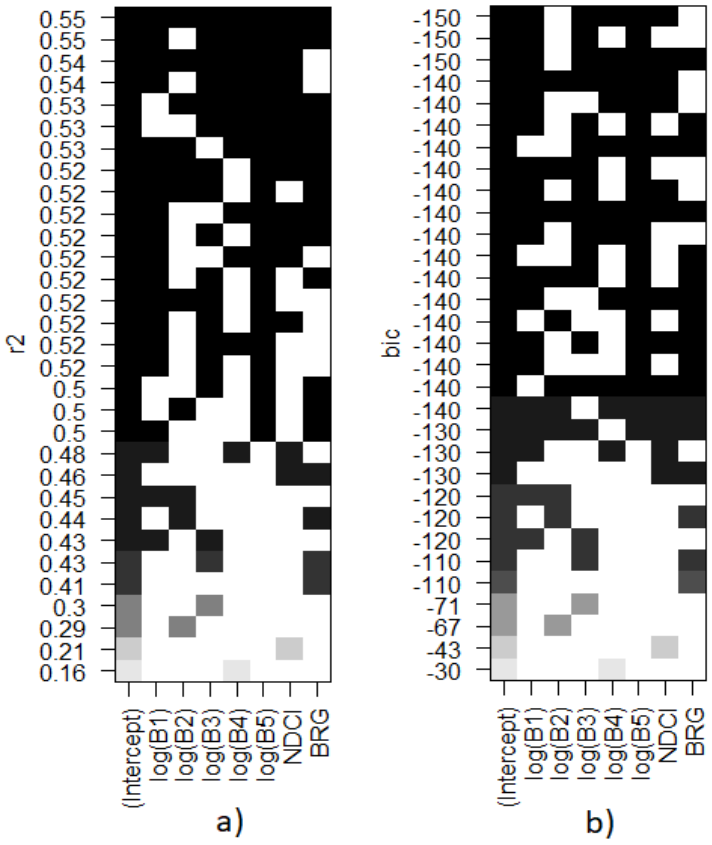

Figure 4 Best linear model assessment through LEAPS variable search. In a) the R2 for the models for the different regressors combinations. In b) the BIC for the respective models.

exhaustive variable search are presented in Figure 4. The R^2 varied from 0.16, using only one explanatory variable (band 4 - NIR), to 0.55, using all possible bands and indices. The highest R^2 value could also be achieved when excluding band 2 (green) and including all other variables. The BIC values varied from -30, using only band 4, to -150, using bands 1, 3, 4 and 5.

错误!未找到引用源。 depicts the relative importance of each band for the overall R^2 of the achieved model. All bands 1-4 are significant (p-value < 0.05) for the construction of the model, although the different models used to evaluate their relative importance did not always agree on the most important variables. Across all methods, band 3 and BRG index has the strongest explanatory power. Band 2 was selected as the second-best explanatory variable by two methods (LMG and First).

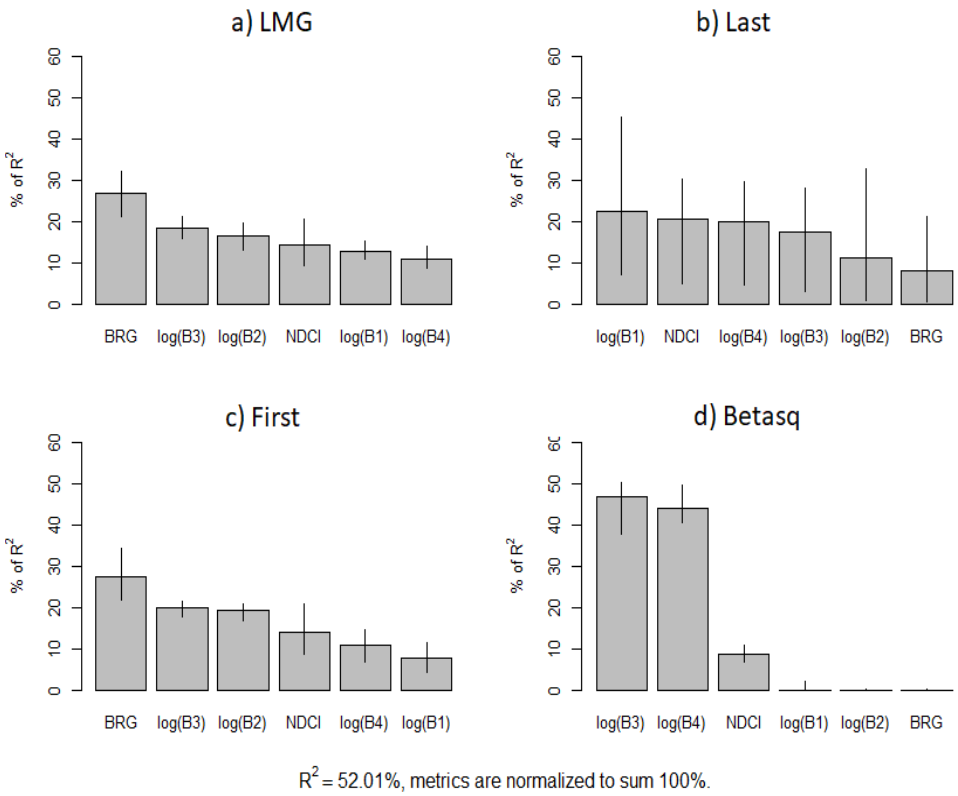

Figure 5 Metric of relative importance using the 700-m radius and a time window of ± 1 day. All bands and the spectral indices included. Relative importance of each LT band with 95% bootstrap confidence intervals.

Figure 6 provides the correlations matrix of all bands and the NDCI and BRG indices. As evidenced, relationships between individual explanatory variables and chl *a* are not linear. Pearson's coefficient of correlation (R) is shown for each band, and indices, with relation to the chl *a* concentration.

Because of the small contribution of the Short-wave Infrared (SWIR) band 5, we have considered a final model for the daily agregated data comprising all these bands and indeces.

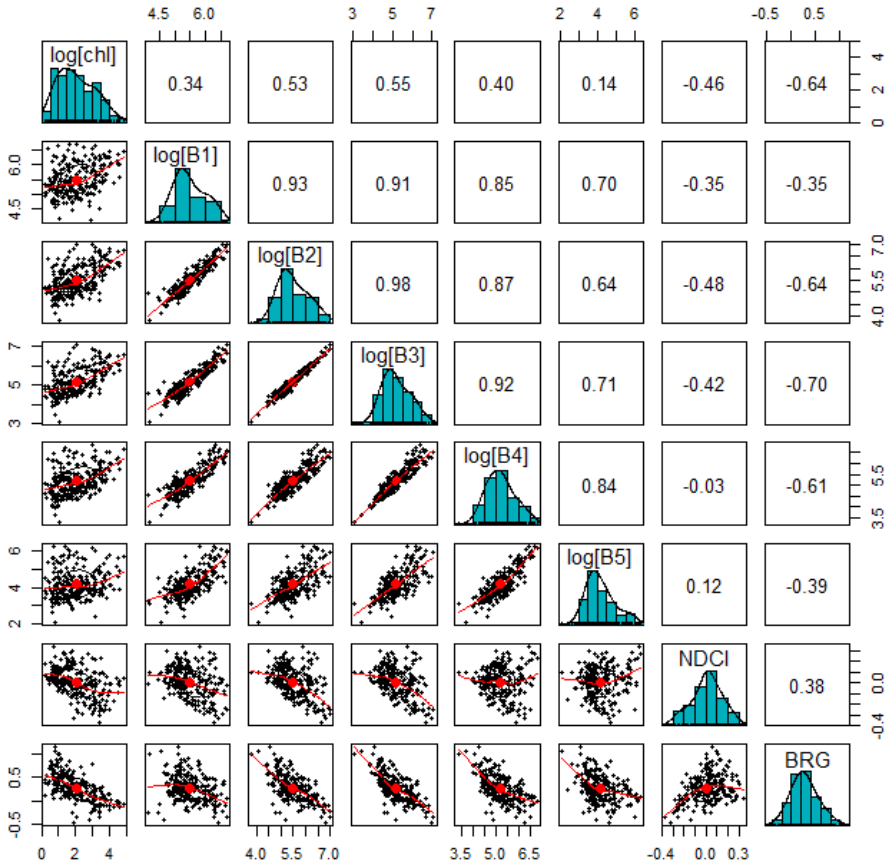

Figure 6 Scatterplots matrix with field measured chl *a* and LT bands the numbers on the diagonal part are the coefficients of correlation, R. The red lines depict polynomial fits to the individual band scatterplots.

All individual bands 1-5 correlate positively with *in situ* chl *a* but bands 2 (R = 0.53) and band 3 (R = 55) are more strongly correlated. Both indices NDCI and BRG are negatively correlated with *in situ* chl *a*.

Model Performance

Short-wave Infrared (SWIR) band 5 of LT 5 and 7 increased the overall performance of the model by 0.04 in the R^2 coefficient. Extracting that band to assess the performance of only visible and NIR bands, also produced important results summarized in Table 1. All models derived from single bands 1 – 4 were significant but not all models that included reflectance indices, NDCI and BRG, are significant. Henceforth, for the best models, we checked if all coefficients are significant, i.e., we checked if all variables significantly contribute for the resulting model. The result is that for the best performing model a R^2 = 52.01 was achieved, including bands 1-4 and both indices. Reducing the time-window between satellite collection and *in situ* sampling has a relevant impact on the overall performance of the model, as it increases the R^2, whilst maintaining a p-value < 0.05.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

Table 1 Buffer size and R^2 for time-windows of ±1 and ±2 days. All tests were performed for the model with bands 1-4, we excluded band 5 as in some cases it leads to non-significant models (p-value 0.05). All bands were significant and all models with p-value < 0.05. For the models derived with the indices, non-significant regressors are shown in brackets.

Buffer Size / t-window	*60 m*	*90 m*	*100 m*	*140 m*	*180 m*	*300 m*	*500 m*	*600 m*	*700 m*	*800 m*	*900 m*
					Model including bands 1 – 4:						
± 1 day	47.81%	46.82%	47.15%	47.68%	43.70%	42.87%	48.27%	47.80%	**48.75%**	48.7%	48.47%
± 2 days	45.15%	37.71%	36.95%	37.51%	37.34%	38.63%	41.4%	41.05%	41.13%	41.77%	41.95%
					Model including bands and indices:						
± 1 day	49.96% [Du]	50.89%	50.77%	50.19%	47.16%	45.15%	51.26%	50.98%	**52.01%**	51.66%	51.92%
± 2 days	45.87% [B2, B3, B4, NDCI, BRG]	39.38% [B1, B2, BRG]	38.83% [B1, BRG]	39.89% [BRG]	39.88%	40.28%	42.53% [B3]	44.39%	44.66%	44.93%	45.21%

From Table 1 we can observe that the choice of the size of the buffer around the sampling location has a small impact when compared to the choice of the time-windows. It is advisable to maintain a buffer size of either 60 m or above 500 meters. Between 90 m to 300 m we observed that the number of additional pixels has created more noise than valuable signal from chl *a* content, thus creating a lower performing model.

Figure 7 shows the best performing model found with a buffer of 700 m and ± 1 day time window; this model includes all visible bands plus Near Infrared (NIR, band 4) and Shortwave Infrared (SWIR, band 5). The inclusion of SWIR slightly increases the overall correlation but also the p-value of the model.

Table 2 Individual band/index importance for the model in figure 7.

	log(B1)	**log(B2)**	**log(B3)**	**log(B4)**	**log(B5)**	**NDCI**	**BRG**
LMG	0.110009	0.149998	0.18466	0.106984	0.075028	0.128362	0.244957

Using satellite images from ± 1 day of *in situ* sampling comes with the disadvantage of a small number of data pairs (Figure 7).

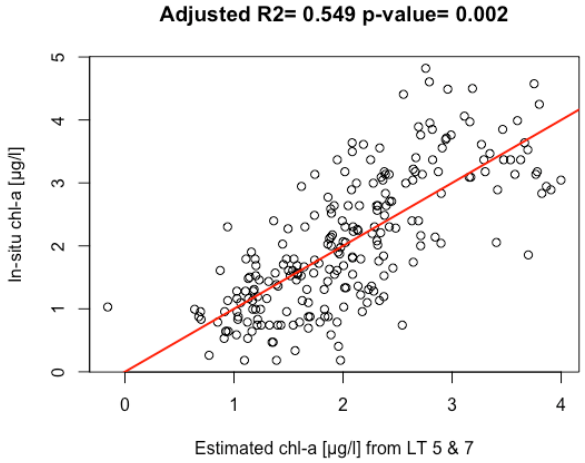

Figure 7 Scatterplot with modelled vs observed chl *a* using the best model (with R^2 and p-value). This model was achieved in the conditions of Table 1: all four bands plus SWIR (band 5) for a time window of ± 1 day and 700 m buffer.

Indeed, the tests in Table 1 and the model in Figure 7 incorporate data from 6-11 out of the 19 lakes – for the remaining lakes, there are no satellite passes free of cloud cover, enough quality control standards, sun elevation or snow-free for the match-up. To assess chl *a* detection levels from different lakes, we have widened the time-window to ± 4 days and lowered the radius to 500 m and included band 5 (SWIR). We have compared results obtained for individual lakes by addressing the physical circumstances of the sampling activities and lakes' characteristics like depth, size and proximity of the sampling to the lakes' margins.

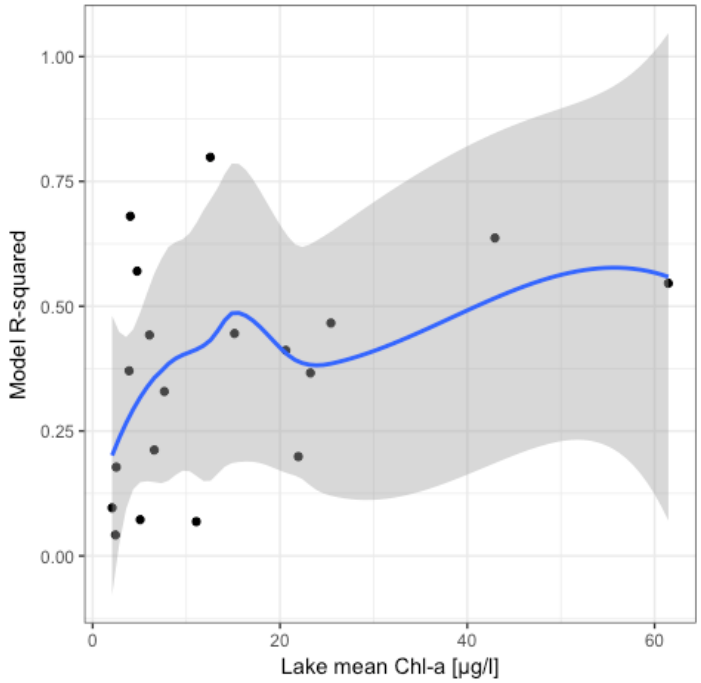

Figure 8 Scatterplot of individual lakes R2 vs average chl *a* concentration in each lake. The graph also shows the root-mean-square error (RMSE) as shadowed area and the blue line calculated as the locally weighted least squares regression. Results were achieved using individual multivariate models for each lake and a time-window of ±4 days.

In Figure 8 we show that the coefficient of determination, R^2, varies depending on the mean chl *a* concentration of each lake. For each trophic state of a lake, the accuracy of the estimates from satellite imagery can vary greatly but there are significantly better performances for eutrophic lakes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

Model performance and trophic state

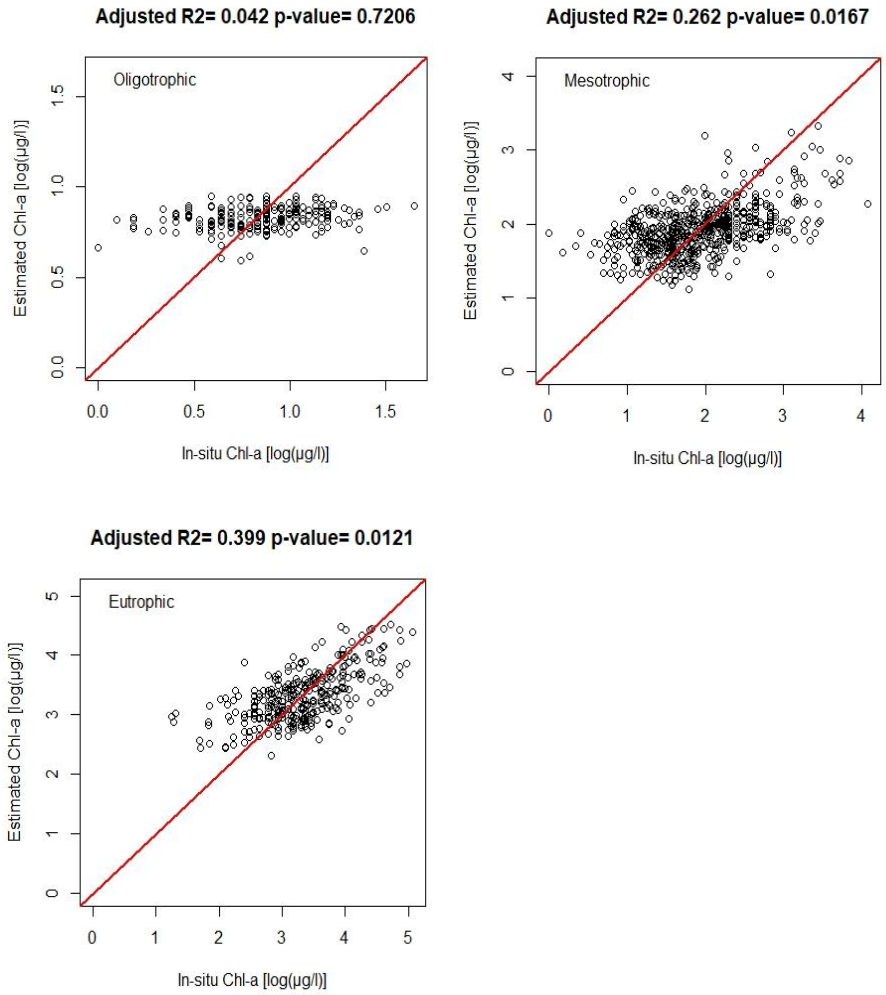

Figure 9 Model performance according to the trophic class of the lakes.

The trophic state of lakes greatly influences the performance of the model built on the individual samples match-up. According to our results in Figure 9, eutrophic lakes have better-resolved chl *a* estimations. Our results show that for eutrophic lakes the model performs twice as better as for mesotrophic lakes. All data points in Figure 10 were filtered to the spring and summer season, i.e., DOY = [100: 280].

Model performance against *in situ* measurements of chl *a* and turbidity

Our results show that turbidity and chl *a* are strongly correlated during summer months, i.e. DOY = [200:280]. Turbidity is therefore highly coupled with chl *a* during ~2.5 months of the year and before that the relationship is not present. The fact that during spring and until July the chl *a* is decoupled from turbidity provides a ground for phenology studies on phytoplankton blooms in lakes. Even in summer period, a model of LT reflectance and turbidity does not perform nearly as well as for chl *a*.

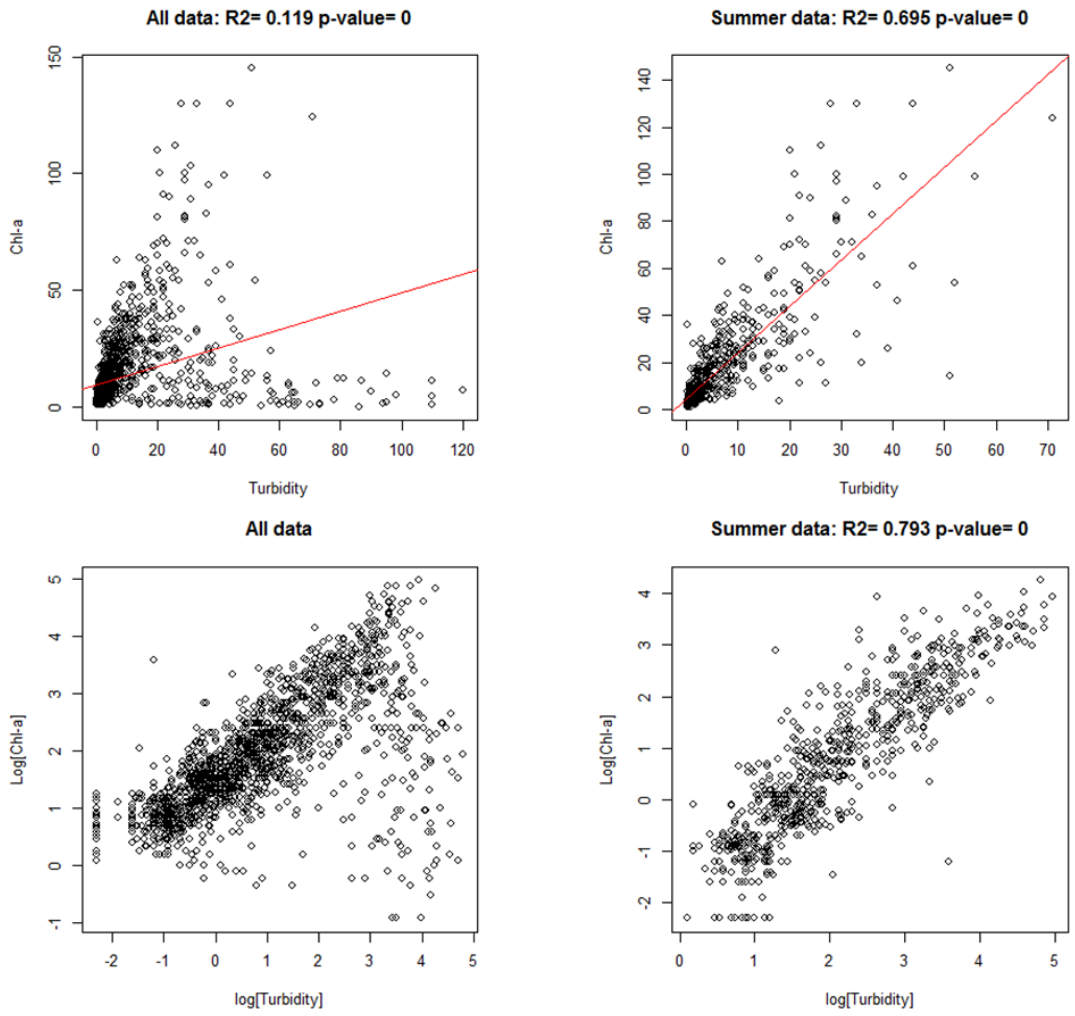

Figure 10 The relationship between chl a and turbidity for all data (left) and summer months (right).

Performance of the model vis-a-vis sampling location and geophysical characteristics

According to the Figure 11, sampling collection set-up greatly influenced how well chl *a* concetrations at particular locations can be estimated using remote sensing. Therefore, the model performance is likely to vary within lakes that are close to each other. There is no linear trend relating depth and R^2, but the smaller the lake the better this model performs. Indeed, small lakes of less than 40 km^2, are better suited to be studied under this model. The relationship between the small lakes signal and the surrounding vegetation will be discussed further. Sampling locations that are very close to the shore (less than 200 m) have poor relation with the model estimates. On the other hand, the performance of the model is better for sampling locations between 200 and 700 m to the lakes' margins (错误!未找到引用源。-c). As it will be discussed further, proximity to the lake's margins can imply higher chl *a* if surrounding vegetation does not interfere with the optical characteristics of the lakes water.

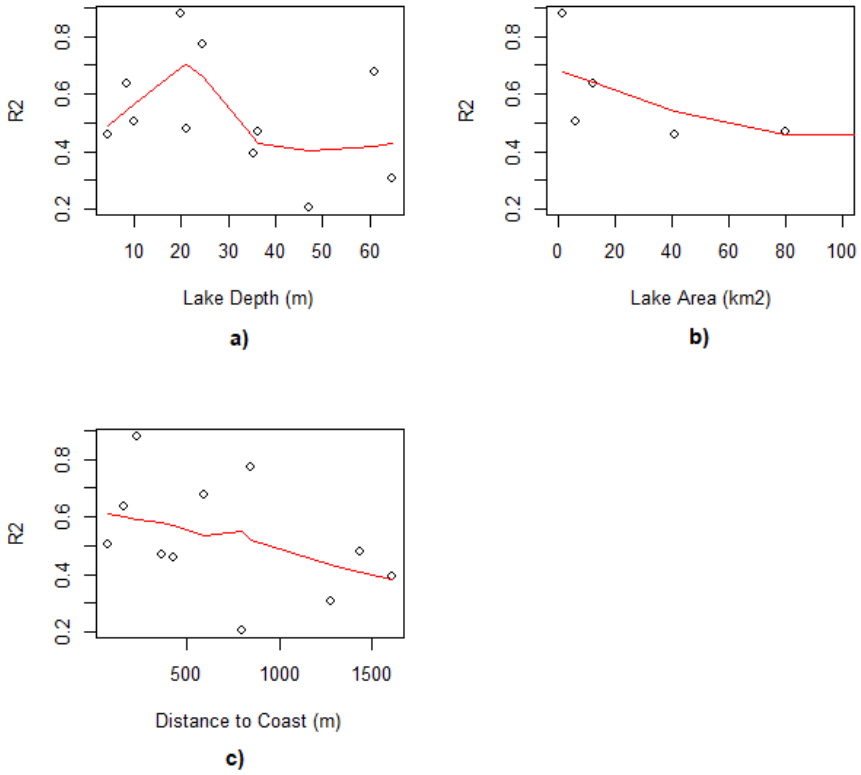

Figure 11 Scatterplot matrix R2 vs lake depth a), lake size b) and distance from the sampling collection site to the coast c).

3.2) Seasonally Aggregated Data

We showed that the signal-to-noise ratio on the remotely sensed chl *a* is higher on the summer months for eutrophic lakes. Hence, we aggregated images collected only during the summer period, i.e., between DOY 100 and 280. These images, when matched-up with the *in situ* samples, produce a well-performing model.

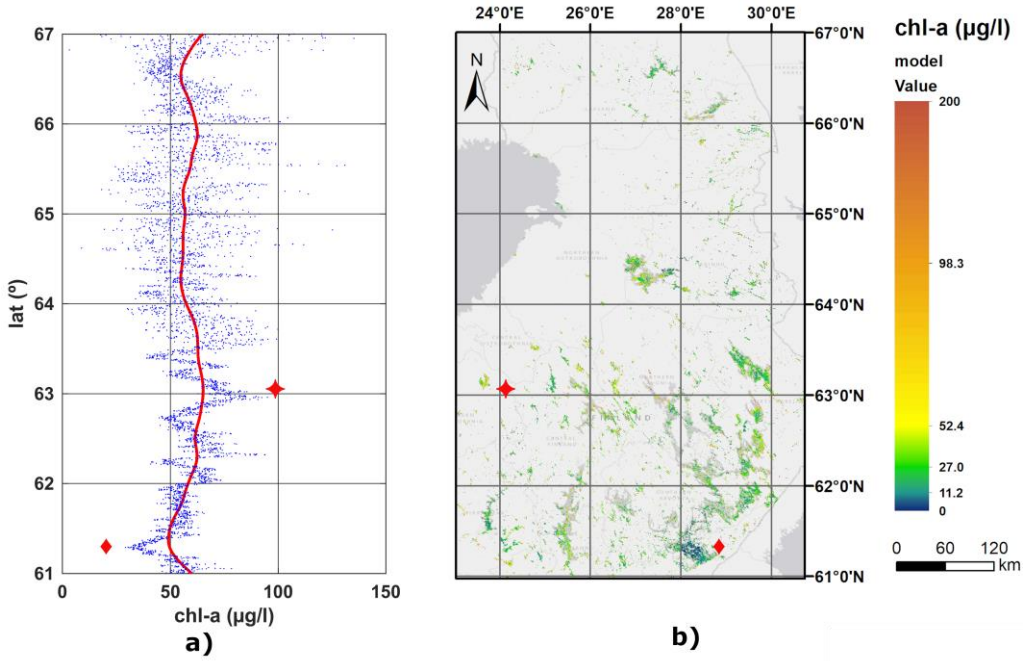

Figure 12 a) blue dots represent the average chl *a* per latitude pixel (0.0013º) and the red line represents the smoothed average per degree in latitude. b) Map with predicted values of mean chl *a* for the Finnish lakes. Lake Saimaa is indicated by the symbol ♦ and lake Lappajärvi by ✦.

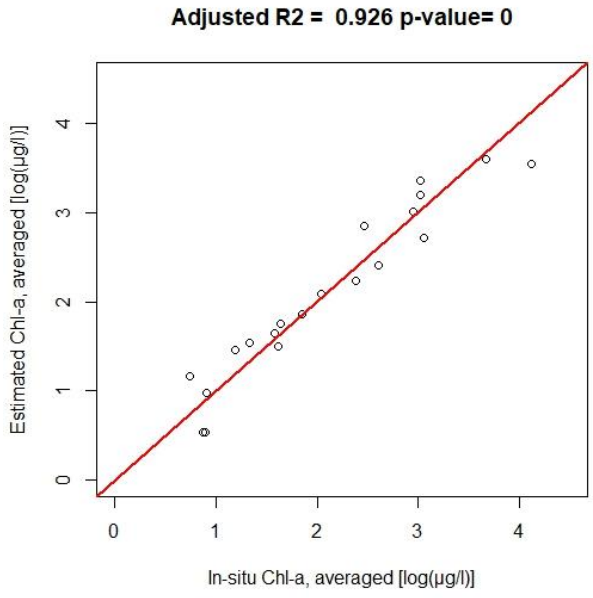

Figure 13 Predicted values of lakes mean surface chl *a*, using multivariate regression.

The seasonally aggregated model (Figure 13) is the result of using a 500-m buffer with a ±2 day with averaged summer results for each lake. This result shows that the model performs well for large-scale estimation of chl *a* between DOY 100 and 280. The results of the chl *a* estimation for the entire Finland, using the corresponding model, is shown in Figure 12. As we expect that a gradient of rising latitudes has

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

large influence on the primary production of the lakes, we averaged the mean chl *a* in our model for the same latitude. For latitudes between 61.5º N to 63.5º N there seems to be a clear positive trend on the increase of seasonal chl *a*. Lake Saimaa (red diamond in Figure 12) is one of the less eutrophic ones, which contributes to the lowest average chl *a* conccentration at the 61.2º N parallel. On the other hand, Lappajärvi (red star in Figure 12 and detailed in Figure 14) is Finland's largest crater lake, it has been given by our model as a very euthrophic lake which corresponds to the same concerns in the litterature.

The clear trend contrasts with the northernmost areas where lakes are less abundant, smaller or sparser. Above 63.5 º N, extreme variance of chl *a* can be seen. Under 63.5ºN latitude, the model correctly identifies lakes that are typically eutrophic, due to the presence of agriculture or arable land.

3) DISCUSSION

In this paper, we aim to better understand how statistical models based on LT perform considering different lakes across an extended time period, from 1984 to mid-2017 and an extended area. There are several differences between the TM and ETM+ sensors, onboard LT 5 and 7, respectively. TM is a multispectral scanning radiometer and ETM+ a whiskbroom scanning radiometer with an additional panchromatic band of 15 m resolution and two 8-bit "gain" ranges. ETM+ also features a 60-m resolution thermal band, replacing the one of 120-m resolution of TM. Nevertheless, these differences do not affect the bands considered in this paper and we assume we can merge data from both sensors.

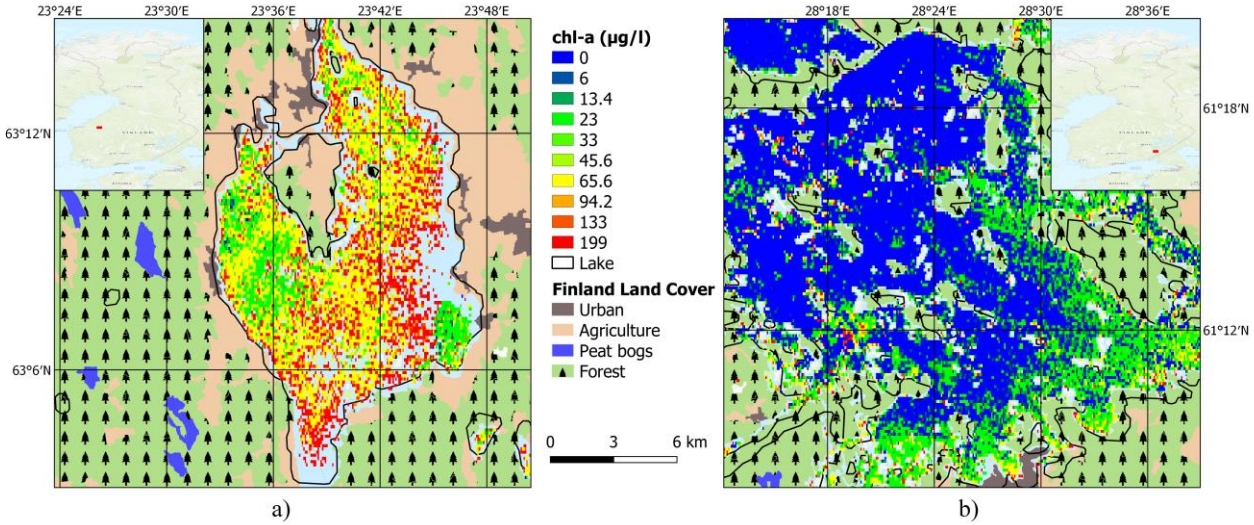

Figure 14 a) Lake Lappajärvi is a eutrophic lake surrounded by arable land and some agriculture land. b) Lake Saimaa is an oligotrophic lake surrounded by forest and some urban areas to the south. Both maps are provided in the same scale.

The choice of the spatial buffer around the sampling location has a small effect on the relationship between chl *a* and satellite data, but a buffer of 500-700 m was found to be the best choice for the models tested in this study. Additionally, a time window of ± 1 day performs better than a ± 2 day, meaning that satellite observations closer to the date when the field samples were collected are preferred. This is in line with previous studies that have assessed chl *a* over lakes, even in different geographic conditions (e.g. Minnesota, USA) (Stadelmann, 2001). We also assessed the impact of extending this time-window up to ±10 days. By doing so, the goodness of predictive chl *a* estimations has decreased significantly. The same is described by other authors, pointing out the decrease in certainty of the model with a longer time-window (Boucher, 2018). Despite this, our models for non-averaged data performed better than other match-up activities for lakes at lower latitudes. As an example, the best model for a selection of water bodies in Maine, USA provided a correlation coefficient of at most 0.25 (Boucher, 2018).

The use of the SWIR 1 band proved not to be useful for the performance of the model derived from intra-annual data. In fact, SWIR bands can be useful for ocean colour on the enhancement of the atmospheric correction algorithms (Vanhellemont et al., 2015). Although useful for calibration of ocean colour models for inland and coastal waters (Wang et al., 2016), the direct use of this band improved the correlation only slightly. Further studies can apply the use of an innovative model comprising red, NIR and SWIR bands as carried on before for MODIS data (Hu, 2009). Nonetheless, the spectral width of the SWIR bands for MODIS (SWIR1: 1628 to 1652 nm and SWIR2: 2105 to 2155 nm) are considerably thinner than the ones on Landsat (SWIR 1: 1570 to 1650 nm and SWIR2: 2110 2290 nm). This difference in radiometric resolution is the reason why we could not use shortwave infrared bands to yield a better model.

On the other hand, as presented in **错误!未找到引用源。** we can determine, with good confidence, the lake specific mean chl *a* concentration in Finnish lakes. This general model for all lakes must be used carefully. From Figure 8 we see that the model performance varies amongst the sampled lakes on the daily aggregation approach. However, it is the level of chl *a* concentration that has the most significant impact on model accuracy. For timely estimations of chl *a*, the LT bands are not ideal as turbidity has a significant influence on the estimated chl *a*. Specifically, during the spring months, it was not possible to detect how turbidity is affecting our results. Despite this, we have seen that seasonally aggregating data from satellite and *in situ* campaigns provides a reliable model. But even in this model, some concerns must be addressed. For example, the land adjacency effect is particularly relevant in small lakes. As seen in **错误!未找到引用源。**-c, preliminary results on the performance of the model with respect to how close the sampling locations are to the margins provide insights that need to be further studied. The adjacency effect is known to reduce apparent surface contrast (Richter et al., 2006) and is particularly severe in the case of dark water bodies surrounded by dense vegetation. As LT imagery is developed mostly for land applications, there is a mask for the lakes' adjacency effect to land but not the converse. In areas near the lakes' margins, the adjacency effect can overestimate chl *a* concentration. The adjacency effect is impactful by lowering the water leaving radiance at shorter wavelengths as showed in Figure 15. It is paramount to have in situ studies that sample a lake at different distances from

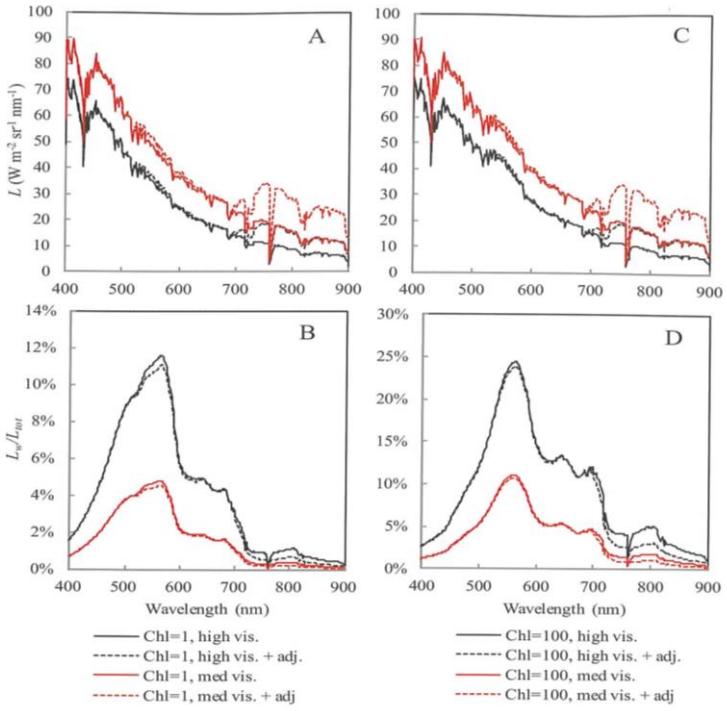

Figure 15 The adjacency effect is most impactful on the shorter wavelength part of the spectrum as evidenced in this adapted figure from Matthews et al 2015.

the coast – such results would greatly improve our capability to quantify the adjacency effect in these types of lakes where sun elevation is low and most of their waters are brown.

From our results, one factor that seems to ameliorate detection performance – including the minimization of the adjacency effect – is the trophic level of the lake.

As seen by the examples of Lappajärvi (⟨chl a⟩ = 86.95 µg/l) and Saimaa lakes (⟨chl a⟩ = 10.1 µg/l), the model correctly identified them as eutrophic and oligotrophic lakes, respectively. Lappajärvi, being a brown water lake, contains humic material and has a high phosphorus content. The phosphorus load into Lappajärvi comes from agriculture and cattle farming and measures have been considered to decrease the impact of eutrophication and mitigate algae bloom occurrence (Teppo et al., 2005). Our results also corroborate the trends in eutrophication such as Lake Keurusselkä (⟨chl a⟩ = 31.83 µg/l) reported as having an increasing trophic status which corresponded to 10 µg/l in 2010 (Räike et al., 2003). In Lake Saimaa the average chl a during the summer period is 10.1 µg/l which also corresponds to the literature, as this is a typical oligotrophic lake (Malve, 2015).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 July 2020 doi:10.20944/preprints202007.0238.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 2898; doi:10.3390/rs12182898

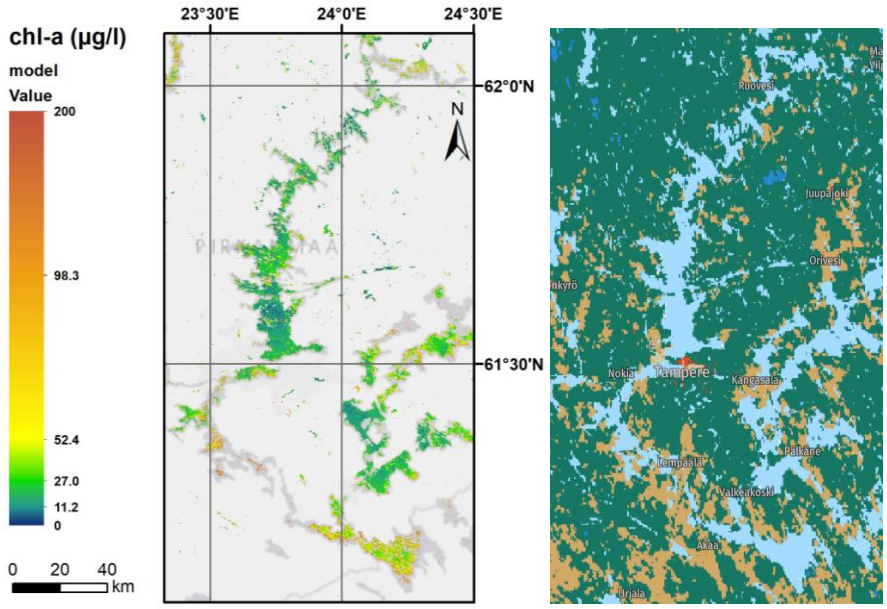

Figure 16 Map with predicted values of mean chl *a (left)* for Pirkanmaa lake and corresponding area for land use: light brown indicating arable land or agriculture, dark green represents forest.

Boreal waters are known for their high concentrations of humic and fluvic acids (CDOM) leached from the

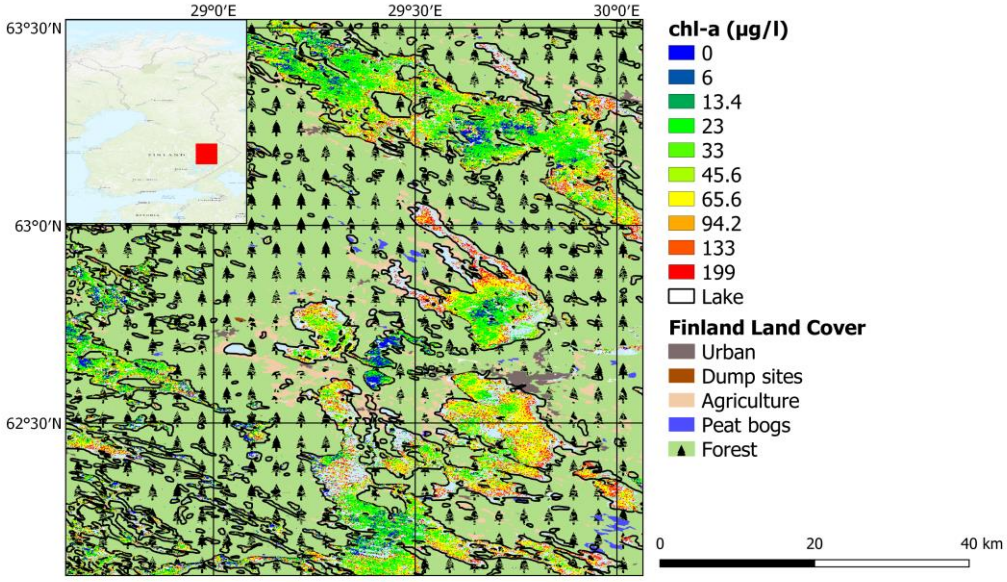

Figure 17 Eastern area of Finland with plenty of small lakes surrounded by dense vegetation and thus where the adjacency effect is most prominent. Land Cover is simplified: "urban" is discontinuous urban fabric, industrial or commercial units; "agriculture" is taken as agriculture and natural vegetation but also non-irrigated arable land; and "forest" comprises transitional woodland-shrub, broad-leaved forest, coniferous forest, mixed forest, and natural grasslands.

surrounding vegetation and soils. These, in turn, reduce the light availability in the blue spectral region

(IOCCG 17). Indeed, from our results, sampling stations placed nearest to the lakes margins present lower correlation with the model.

4) CONCLUSIONS

Despite its relatively high spatial resolution, LT imagery is often overlooked for remote sensing of chl *a*, due to the low spectral resolution of the sensors. Nevertheless, our results demonstrate that estimations over aggregated time windows can be done with high accuracy.

Reducing imagery over monthly means can produce useful results of the model allowing for higher correlations and typical characterizations of the different water bodies within specified timeframes. Assessing chl *a* is possible due to its strong relationship with turbidity although multi-linear models of LT data to estimate turbidity do not produce as good results as the chl *a* models. The relationship between chl *a* and turbidity is most robust during summer, and since the model performs better during this period, the derived model can be used for further studies of phenology shifts during the blooming seasons.

The adjacency effect, when quantified, could explain the high variations of our model when applied to the calculation of chl *a* concentration in small lakes. During this study we came across the difficulty of identifying the impacts of the adjacency effect both through our analysis and the timely literature. Therefore, we believe that the work herewith can open the discussion for retrieving more sophisticated models for chl *a* estimation, giving a better suited measurement of adjacency effects from land on small water bodies remotely sensed through Landsat imagery. Similar studies have already been carried for the adjacency effect on MODIS, SeaWiFS, MERIS, OLCI, OLI and MSI for the case of mid-latitude coastal environments (Bulgarelli et al., 2018). We argue, that satellites can provide valuable information in cases where required sampling density is high, or field surveys are expensive or even impossible to carry out. Especially, satellite images may provide informative prior information for field sampling estimates, in a Bayesian setting. As lakes are sensitive to many climate factors, they may also be useful in climate change monitoring.

ACKNOWLEDGEMENTS

The PhD research grant PD/BD/113932/2015, given by the Portuguese Fundação para a Ciência e a Tecnologia (FCT), funded this work and the authors are thankful for such fundamental support. This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº 810139. Dr. Vanda Brotas received a sabbatical grant from FCT SFRH/BSAB/142981/201. Dr. Eduardo Maeda is funded by the Academy of Finland (decision numbers 318252 and 319905. The authors would like to thank the Suomen Ympäristökeskus (SYKE) for the valuable *in situ* chl *a* data as well as the constructive criticism provided. We also thank the

valuable advice and support from Kari Kallio and Sampsa Koponen. This work was also supported by funding from the European Union's Horizon 2020 Research and Innovation Programme grant agreement N 810139: Project Portugal Twinning for Innovation and Excellence in Marine Science and Earth Observation – PORTWIMS. Finally, the authors would like to thank the Google Earth Engine forum for the help given in producing the tools that are now available for chl *a* remote sensing validation studies from multiple space platforms.

REFERENCES

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., … Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, *54*(6part2), 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

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*(4), 1044–1054. https://doi.org/10.1002/eap.1708

Brivio, P. A., Giardino, C., & Zilioli, E. (2001). Determination of chlorophyll concentration changes in Lake Garda using an image-based radiative transfer code for Landsat TM images. *International Journal of Remote Sensing*, *22*(2–3), 487–502. https://doi.org/10.1080/014311601450059

Bulgarelli, B., & Zibordi, G. (2018). On the detectability of adjacency effects in ocean color remote sensing of mid-latitude coastal environments by SeaWiFS, MODIS-A, MERIS, OLCI, OLI and MSI. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2017.12.021

Dalu, T., Dube, T., Froneman, P. W., Mwazvita, T. B., Clegg, B. W., Nhiwatiwa, T., … Nhiwatiwa, T. (2015). An assessment of chlorophyll-a concentration spatio-temporal variation using Landsat satellite data , in a small tropical reservoir. *Geocarto International*, *6049*(August 2017), 1–14. https://doi.org/10.1080/10106049.2015.1027292

Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley, R. D., Beckmann, T., … Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment*, *194*, 379–390. https://doi.org/10.1016/j.rse.2017.03.026

Giardino, C., Bresciani, M., Villa, P., & Martinelli, A. (2010). Application of Remote Sensing in Water Resource Management: The Case Study of Lake Trasimeno, Italy. *Water Resources*

Management, *24*(14), 3885–3899. https://doi.org/10.1007/s11269-010-9639-3

Giardino, C., U, C. G., Pepe, M., Brivio, P. A., Ghezzi, P., & Zilioli, E. (2001). Detecting chlorophyll , Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery Detecting chlorophyll , Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery, (September 2015). https://doi.org/10.1016/S0048-9697(00)00692-6

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Remote Sensing of Environment Google Earth Engine : Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, (2016). https://doi.org/10.1016/j.rse.2017.06.031

Gregg, W. W., Rousseaux, C. S., & Franz, B. A. (2017). Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. *Remote Sensing Letters*, *8*(12), 1102–1111. https://doi.org/10.1080/2150704X.2017.1354263

Grömping, U. (2006). R package relaimpo: relative importance for linear regression. *Journal Of Statistical Software*, *17*(1), 139–147. https://doi.org/10.1016/j.foreco.2006.08.245

Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, *113*(10), 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012

IOCCG. (2014). Phytoplankton functional types from Space. *Reports and Monographs of the International OceanColour Coordinating Group*, (15), 156.

IOCCG. (2018). *Earth Observations in Support of Global Water Quality Monitoring*. (S. Greb, A. Dekker, & C. Binding, Eds.) (Vol. No. 17). Dartmouth, Canada: IOCCG. https://doi.org/10.25607/OBP-113

Isenstein, E. M., Trescott, A., & Park, M.-H. (2014). Multispectral Remote Sensing of Harmful Algal Blooms in Lake Champlain, USA. *Water Environment Research*, *86*(12), 2271–2278. https://doi.org/10.2175/106143014X13975035526149

Kraemer, B. M., Chandra, S., Dell, A. I., Dix, M., Kuusisto, E., Livingstone, D. M., … McIntyre, P. B. (2017). Global patterns in lake ecosystem responses to warming based on the temperature dependence of metabolism. *Global Change Biology*, *23*(5), 1881–1890. https://doi.org/10.1111/gcb.13459

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Scott, Foresman. Retrieved from https://books.google.fi/books?id=-hfvAAAAMAAJ

Malve, O., Hjerppe, T., Tattari, S., Väisänen, S., Huttunen, I., Kotamäki, N., … Kauppila, P.

(2015). Participatory operations model for cost-efficient monitoring and modeling of river basins - A systematic approach. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2015.06.105

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., … Lim, T. K. (2006). A landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 68–72. https://doi.org/10.1109/LGRS.2005.857030

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*, 13603. https://doi.org/10.1038/ncomms13603

Mikkonen, S., Laine, M., Mäkelä, H. M., Gregow, H., Tuomenvirta, H., Lahtinen, M., & Laaksonen, A. (2015). Trends in the average temperature in Finland, 1847–2013. *Stochastic Environmental Research and Risk Assessment*, *29*(6), 1521–1529. https://doi.org/10.1007/s00477-014-0992-2

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion : background , derivation , and applications, *4*(April), 2–6. https://doi.org/10.1002/wics.199

Östlund, C., Flink, P., Strömbeck, N., Pierson, D., & Lindell, T. (2001). Mapping of the water quality of Lake Erken, Sweden, from Imaging Spectrometry and Landsat Thematic Mapper. *Science of the Total Environment*, *268*(1–3), 139–154. https://doi.org/10.1016/S0048-9697(00)00683-5

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

R Development Core Team. (2016). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria*, *0*, {ISBN} 3-900051-07-0. https://doi.org/10.1038/sj.hdy.6800737

Räike, A., Pietiläinen, O. P., Rekolainen, S., Kauppila, P., Pitkänen, H., Niemi, J., … Vuorenmaa, J. (2003). Trends of phosphorus, nitrogen and chlorophyll a concentrations in Finnish rivers and lakes in 1975-2000. *Science of the Total Environment*, *310*(1–3), 47–59. https://doi.org/10.1016/S0048-9697(02)00622-8

Richter, R., Bachmann, M., Dorigo, W., & Müller, A. (2006). Influence of the adjacency effect on

ground reflectance measurements. *IEEE Geoscience and Remote Sensing Letters*, *3*(4), 565–569. https://doi.org/10.1109/LGRS.2006.882146

Rühland, K., Paterson, A. M., & Smol, J. P. (2008). Hemispheric-scale patterns of climate-related shifts in planktonic diatoms from North American and European lakes. *Global Change Biology*, *84*(2), ???-??? https://doi.org/10.1111/j.1365-2486.2008.01670.x

Stadelmann, T. H., Brezonik, P. L., & Kloiber, S. (2001). Seasonal patterns of chlorophyll a and secchi disk transparency in lakes of east-Central Minnesota: Implications for design of ground- and satellite-based monitoring programs. *Lake and Reservoir Management*, *17*(4), 299–314. https://doi.org/10.1080/07438140109354137

Teppo, A., Tuhkanen, J., & Sivil, M. (2005). Biomanipulation of large moderately eutrophicated Lake Lappajärvi. *SIL Proceedings, 1922-2010*, *29*(2), 841–844. https://doi.org/10.1080/03680770.2005.11902798

Thiemann, S., & Kaufmann, H. (2000). Determination of chlorophyll content and trophic state of lakes using field spectrometer and IRS-1C satellite data in the Mecklenburg Lake District, Germany. *Remote Sensing of Environment*, *73*(2), 227–235. https://doi.org/10.1016/S0034-4257(00)00097-3

Vanhellemont, Q., & Ruddick, K. (2015). Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8. *Remote Sensing of Environment*, *161*, 89–106. https://doi.org/10.1016/J.RSE.2015.02.007

Vincent, R. K., Qin, X., McKay, R. M. L., Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, *89*(3), 381–392. https://doi.org/10.1016/j.rse.2003.10.014

Wang, M., Shi, W., Jiang, L., & Voss, K. (2016). NIR- and SWIR-based on-orbit vicarious calibrations for satellite ocean color sensors. *Optics Express*, *24*(18), 20437. https://doi.org/10.1364/OE.24.020437

ISO 10260. Water quality – Measurement of biochemical parameters – Spectrometric determination of the chlorophyll a concentration. International Organization for Standardization, Geneva, 1992.

EN 27027. Water quality. Determination of turbidity. European Committee for Standardization, Strasbourg, 1994.