

<http://dx.doi.org/10.1590/2318-0331.231820170106>

Can chlorophyll-*a* in meso-oligotrophic shallow waters be estimated using statistical approaches and empirical models from MODIS imagery?

*É possível estimar a Clorofila-*a* em águas rasas meso-oligotróficas usando abordagens estatísticas e modelos empíricos a partir de imagens MODIS?*

Andres Mauricio Munar¹, José Rafael Cavalcanti¹, Juan Martin Bravo¹, David Manuel Lelinho Da Motta Marques¹
and Carlos Ruberto Fragoso Júnior²

¹Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

²Centro de Tecnologia Maceio, Universidade Federal de Alagoas, AL, Brasil

*E-mails: mauriciomunar4@gmail.com (AMM); rafaelcavalcanti@gmail.com (JRC); jumarbra@iph.ufrgs.br (JMB);
dmm@iph.ufrgs.br (DMLDMM); carlosruberto@gmail.com (CRFJ)

Received: August 30, 2017 - Revised: December 29, 2017 - Accepted: February 27, 2018

ABSTRACT

Accurate estimation of chlorophyll-*a* (Chl-*a*) concentration in inland waters through remote-sensing techniques is complicated by local differences in the optical properties of water. In this study, we applied multiple linear regression (MLR), artificial neural network (ANN), nonparametric multiplicative regression (NPMR) and four models (Appel, Kahru, FAI and O14a) to estimate the Chl-*a* concentration from combinations of spectral bands from the MODIS sensor. The MLR, NPMR and ANN models were calibrated and validated using *in-situ* Chl-*a* measurements. The results showed that a simple and efficient model, developed and validated through multiple linear regression analysis, offered advantages (*i.e.*, better performance and fewer input variables) in comparison with ANN, NPMR and four models (Appel, Kahru, FAI and O14a). In addition, we observed that in a large shallow subtropical lake, where the wind and hydrodynamics are essential factors in the spatial heterogeneity (Chl-*a* distribution), the MLR model adjusted using the specific point dataset, performed better than using the total dataset, which suggest that would not be appropriate to generalize a single model to estimate Chl-*a* in these large shallow lakes from total datasets. Our approach is a useful tool to estimate Chl-*a* concentration in meso-oligotrophic shallow waters and corroborates the spatial heterogeneity in these ecosystems.

Keywords: Chl-*a*; MLR; ANN; NPRM; Remote sensing; Lake Mangueira.

RESUMO

A estimativa da concentração de Clorofila-*a* (Chl-*a*) em ambientes aquáticos através de técnicas de sensoriamento remoto é complexa devido as diferenças entre as propriedades óticas da água. O objetivo deste trabalho foi estimar concentrações de Chl-*a* a partir das combinações de bandas espectrais do sensor MODIS, aplicando análise de regressão linear múltipla (MLR), redes neurais artificiais (ANN), regressão multiplicativa não paramétrica e quatro modelos (Appel, Kahru, FAI e O14a). Os modelos de regressão linear múltipla (MLR), NPMR e ANN foram calibrados e validados utilizando medições *in-situ* de Clorofila-*a*. Os resultados demonstraram que um simples e eficiente modelo desenvolvido e validado através de regressão linear múltipla (MLR) produz vantagens (*i.e.*, melhor *performance* e poucas variáveis de entrada) em comparação com os modelos ANN, NPMR e quatro modelos (Appel, Kahru, FAI e O14a). Além disso, foi observado que em um grande lago raso subtropical, no qual o vento (intensidade e direção) e a hidrodinâmica são fatores essenciais na heterogeneidade espacial (distribuição de Chl-*a*), o modelo MLR desenvolvido utilizando um ponto específico do lago teve melhor *performance* que utilizando o total de dados, sugerindo que não seria apropriado generalizar um único modelo para estimar Clorofila-*a* nesses grandes lagos rasos utilizando o total de dados. Esta abordagem é uma ferramenta útil para estimar concentrações de Chl-*a* em ambientes rasos meso-oligotróficos e corroborar a heterogeneidade espacial desses ecossistemas.

Palavras-chave: Chl-*a*; MLR; ANN; NPRM; Sensoriamento remoto; Lagoa Mangueira.



INTRODUCTION

Chl-*a* is an important indicator of the trophic state in lakes and reservoirs, showing patterns associated with internal processes and natural stressors (HONEYWILL; PATERSON; HAGERTHEY, 2002; DUKA; CULLAJ, 2009; SCHALLES et al., 1998). Detection and quantification of Chl-*a* are essential for assessing water quality in these ecosystems, as the concentration of this compound provides critical knowledge of the phytoplankton community. However, most traditional methods used to retrieve Chl-*a* concentrations are based on *in-situ* measurements, and are time-consuming and difficult to apply to large areas (GHOLIZADEH; MELESSE; REDDI, 2016; KASPRZAK et al., 2008), especially with frequent measurements. In recent years, Chl-*a* concentrations have often been estimated by remote sensing, which is an effective means of rapid, high-frequency data acquisition (ALLAN et al., 2011; OLMANSON; BREZONIK; BAUER, 2011; SHI et al., 2015). In addition, this technique allows one to obtain information on remote sites and larger areas.

Currently, bio-optical algorithms based on radiometric measurements (*e.g.*, solar irradiance, sky radiance) have been used to obtain quantitative information of substances present on water bodies (OGASHAWARA; MISHRA; GITELSON, 2017). These algorithms can be classified in empirical, semiempirical, semi-analytical, quasi-analytical or analytical algorithms (OGASHAWARA, 2015), being the empirical and empirical algorithms commonly used to retrieve Chl-*a* concentrations (DUAN et al., 2010; LE et al., 2013; RITCHIE; ZIMBA; EVERITT, 2003; ROSA NETO et al., 2015). Limitations of Chl-*a* detection by remote sensing include atmospheric correction methods and sensor limitations, as well as the influences of detritus, the presence of colored dissolved organic matter (CDOM), and scattering by Total Suspended Matter (TSM), which are difficult to detect because they affect the optical properties of water (DARECKI; STRAMSKI, 2004; HU, 2009; WU et al., 2009). Another limitation is the light reflected off the bottom, which may affect the accuracy of the empirical algorithm because the signal received by the sensor varies as a function of the wavelength and with the clarity of the water (CARDOSO et al., 2012; SHUBHA, 2000; LEE et al., 2001).

Recently, data from MODIS (*Moderate Resolution Imaging Spectroradiometer*) have been used successfully for assessment and continuous monitoring of water quality, including Chl-*a* concentrations, in lakes and reservoirs (DENG et al., 2017; OLMANSON; BREZONIK; BAUER, 2011; SONG et al., 2013; WU et al., 2014; ZHANG et al., 2011). Empirical algorithms based on MODIS data, such as OC3M (O'REILLY et al., 2000), have been used in different environments (GILSON et al., 2010; GURLIN; GITELSON; MOSES, 2011; LESHT; BARBIERO; WARREN, 2013; MATTHEWS; BERNARD; ROBERTSON, 2012; OGASHAWARA et al., 2014). Nevertheless, this algorithm has a tendency to overestimate the Chl-*a* concentrations in oligotrophic environments, and to underestimate it in highly productive environments (KOMICK; COSTA; GOWER, 2009; MOORE; CAMPBELL; DOWELL, 2009).

Recent studies have applied empirical algorithms for retrieving Chl-*a* concentrations in lakes and reservoirs combining MODIS spectral bands (DUAN et al., 2017; CHANG et al., 2012; HUANG et al., 2014; SHI et al., 2017; WANG; SHI; TANG, 2011;

WANG; LIU; TANG, 2017; WU et al., 2009; XIANG et al., 2015; ZHANG et al., 2016). However, these algorithms were developed for eutrophics waters (*e.g.*, Lakes Taihu and Chaohu, China and Lake Okeechobee, USA) and for particular seasons, limiting their uses elsewhere and throughout the year (MATTHEWS, 2011; PALMER et al., 2015a). These approaches can be improved using the appropriate spectral remote-sensing reflectance values and combinations of spectral bands with either regression analysis (CHANG et al., 2012; OGASHAWARA et al., 2014) or an ANN (WU et al., 2009; XIANG et al., 2015). However, only the approach proposed by Ogashawara et al. (2014) was developed in a meso-oligotrophic environment (Itumbiara Reservoir, Brazil).

Monitoring Chl-*a* concentrations from remote sensing in freshwater environments such as shallow lakes has limitations, which are associated with low signal noise ratio and the optical complexity of these waters being highly variable between and even within water bodies (GHOLIZADEH; MELESSE; REDDI, 2016; PALMER; KUTSER; HUNTER, 2015b). In addition, in large shallow lakes, which are rarely studied, especially in meso-oligotrophic environments submitted to intense wind regimes, limitations such as scale factors (*e.g.*, representative monitoring at a single point), spatial gradients and hydrodynamics might significantly affect the interpretation of results from remote-sensing data and the performance of the empirical algorithms (CHAVULA et al., 2009; OGASHAWARA et al., 2014; RUIZ-VERDÚ et al., 2016).

In this study, we (1) proposed and compared multivariable empirical models to retrieve Chl-*a* concentrations, using MODIS data, (2) determined the accuracy of these models in a large shallow subtropical, and (3) verified if generalized models can represent the spatial heterogeneity of the lake. We used multiple linear regression, artificial neural network, nonparametric multiplicative regression, and four models (Appel, Kahru, FAI and O14a) to retrieve the Chl-*a* concentration in the Lake Mangureira, a large shallow subtropical lake located in Southern Brazil. This approach is the first attempt to assess the application of remote sensing techniques in such a system, which the oligotrophic conditions are common most of the time (CROSSETTI et al., 2013; LIMA et al., 2016), and the spatial heterogeneity has large differences between the littoral and pelagic zones (CARDOSO et al., 2012; CROSSETTI et al., 2013; THEY; MARQUES; SOUZA, 2012).

Differently from previous studies, our approach is a first step toward to obtain comprehensive and reliable empirical models for meso-oligotrophic shallow waters, which may be further tested with remote-sensed data in order to retrieve the Chl-*a* concentration and help to understanding the shallow lakes heterogeneity.

METHODS

Study area

Lake Mangureira is a large shallow subtropical waterbody located in Rio Grande do Sul, Brazil. The lake is 820 km² in area and has a mean depth of 2.6 m, a maximum depth of 6.9 m, and is 90 km long and 3–10 km wide (Figure 1). Its trophic state ranges from oligotrophic to mesotrophic (CROSSETTI et al., 2013; FRAGOSO JUNIOR et al., 2011). The regional climate is

subtropical with a mean annual temperature of 16 °C and annual rainfall between 1,800 and 2,200 mm (KOTTEK et al., 2006).

Field data

The field data were collected over a period of 9 years (2001 to 2009) at three sampling points in the North, Central and South parts of Lake Mangueira, TAMAN, TAMAC and TAMAS respectively, measured two to three times a year. The dataset contained concentrations of Chl-*a*, Secchi disk depth and other limnological variables (*i.e.*, Total Suspended Solids (TSS), Dissolved Oxygen (DO), pH and Depth), collected from the three sampling points, and indicated a wide range of Chl-*a* concentrations at the Central point (Table 1). We used the data from the 2001 – 2005

period for calibration and from the 2006 – 2009 period for model validation.

For Chl-*a* determination, surface water samples were filtered through Whatman GF/F with 90% ethanol and measured by spectrophotometry (JESPERSEN; CHRISTOFFERSEN, 1987). Dissolved Oxygen, pH and Depth were measured with a multiparameter probe (YSI 6600). Total Suspended Solids were assessed gravimetrically by water evaporation in porcelain dishes (EATON; CLESCERI, 1999) and the Water Transparency (WT) was measured using a Secchi disk.

Satellite data and image processing

MODIS-Terra (MOD09GA) and MODIS-Aqua (MYD09GA) Level-2 Surface Reflectance products with a daily frequency and a 500-m spatial resolution were downloaded via the web interface *The Land Processes Distributed Interface Active Archive Center* (LP DAAC), from *The US Geological Survey EROS Data Center* (LPDAAC, 2015). These products have been successfully used in different inland waters, including oligo- to mesotrophic reservoirs (surface area: 778 km²) (CURTARELLI et al., 2015; OGASHAWARA et al., 2014), Amazon floodplain lakes (surface area: 2000 km²) (NOVO et al., 2006), Minnesota lakes (lake surface area > 160 ha) (KNIGHT; VOTH, 2012), Poyang Lake (surface area: 3000 km²) (QI et al., 2016), and coastal lagoons (900 km²) (SRICHANDAN et al., 2015). The cloud cover fraction was obtained using a cloud mask product from MODIS product level 2 (named M*D35L2) (ACKERMAN et al., 1998). In addition, MODIS quality product was also used to assess image quality (ROY et al., 2002). The sinusoidal projection images (ISIN) from the LP DAAC platform were transformed to geographic projection (Lat/Long) using the MODIS Reprojection Tool. The MOD09GA, MYD09GA and M*D35L2 products were processed and analysed using a MatLab[®] routine. To retrieve Chl-*a* concentrations, we used the surface reflectance from the first six spectral bands of the MOD09GA and MYD09GA products, with the centre wavelength location respective (band 1: 648 nm, band 2: 858 nm; band 3: 470 nm; band 4: 555 nm; band 5: 1240 nm; and band 6: 1640 nm).

We downloaded a total of 50 MODIS images (32 MODIS-Terra and 18 MODIS-Aqua) matching the *in-situ* measurements of Chl-*a* monitoring days (Table 2). From this total, we selected 16 images for model development and calibration, and 10 images for model verification; the remaining 24 images were discarded due to the presence of clouds, shadows on the measurement

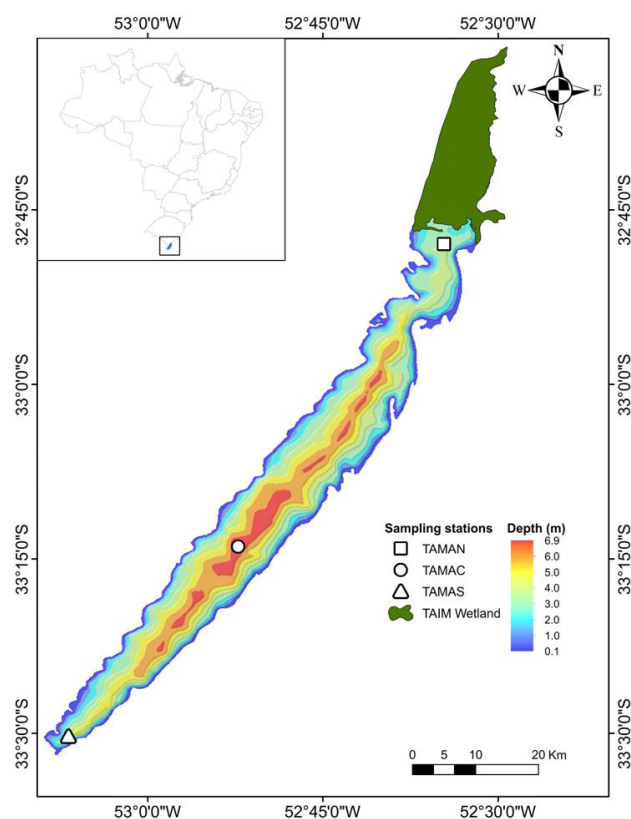


Figure 1. Bathymetry and locations of sampling stations (North - TAMAN, Central - TAMAC and South - TAMAS) in Lake Mangueira, Rio Grande do Sul, Brazil.

Table 1. Chl-*a* concentrations, Secchi disk depth, total suspended solids (TSS), dissolved oxygen (DO), pH and Depth measured in 2001-2009 (N=43) at three points (North, Central, South) in Lake Mangueira.

| Parameter | North | Central | South |
|---|------------------------------|------------------------------|------------------------------|
| | Mean \pm SD (Min – Max) | Mean \pm SD (Min – Max) | Mean \pm SD (Min – Max) |
| Chl- <i>a</i> ($\mu\text{g}\cdot\text{L}^{-1}$) | 10.88 \pm 8.54 (1.04–36.9) | 13.64 \pm 13.52 (0.5–63.3) | 5.54 \pm 4.73 (0.2–19.30) |
| Secchi disk depth (m ⁻¹) | 0.72 \pm 0.33 (0.16–1.55) | 0.79 \pm 0.21 (0.55–1.36) | 1.23 \pm 0.40 (0.72–1.90) |
| TSS ($\mu\text{g}\cdot\text{L}^{-1}$) | 33.04 \pm 21.03 (3.0–80.0) | 26.82 \pm 14.74 (6.0–68.0) | 19.85 \pm 12.97 (3.0–60.0) |
| DO ($\mu\text{g}\cdot\text{L}^{-1}$) | 9.44 \pm 1.50 (6.89–13.45) | 9.62 \pm 1.26 (6.45–12.70) | 9.94 \pm 1.15 (8.12–12.70) |
| pH | 8.06 \pm 0.42 (7.07–8.60) | 8.22 \pm 0.47 (6.76–8.95) | 8.25 \pm 0.53 (6.84–8.92) |
| Depth (m) | 2.88 \pm 0.83 (1.60–4.30) | 2.58 \pm 0.74 (1.04–3.80) | 2.43 \pm 0.73 (1.26–3.80) |

Table 2. Number of available MODIS images for calibration/validation periods for the North (N), Central (C), and South (S) points in Lake Mangueira.

| Id | Field Campaign Date | Station | | | Id | Field Campaign Date | Station | | |
|----|---------------------|---------|---|---|----|---------------------|---------|---|---|
| | | N | C | S | | | N | C | S |
| 1 | 21/03/2001 | | | | 26 | 01/02/2004 | | | |
| 2 | 20/04/2001 | | | | 27 | 19/05/2004 | | | |
| 3 | 23/05/2001 | | | | 28 | 18/08/2004 | | | |
| 4 | 19/06/2001 | | | | 29 | 18/09/2004 | | | |
| 5 | 17/07/2001 | | | | 30 | 10/11/2004 | | | |
| 6 | 21/08/2001 | | | | 31 | 22/03/2005 | | | |
| 7 | 27/09/2001 | | | | 32 | 03/05/2005 | | | |
| 8 | 24/10/2001 | | | | 33 | 01/08/2005 | | | |
| 9 | 20/11/2001 | | | | 34 | 01/11/2005 | | | |
| 10 | 19/12/2001 | | | | 35 | 01/03/2006 | | | |
| 11 | 29/01/2002 | | | | 36 | 17/05/2006 | | | |
| 12 | 21/02/2002 | | | | 37 | 01/08/2006 | | | |
| 13 | 20/03/2002 | | | | 38 | 08/11/2006 | | | |
| 14 | 23/04/2002 | | | | 39 | 09/11/2006 | | | |
| 15 | 27/05/2002 | | | | 40 | 30/05/2007 | | | |
| 16 | 19/06/2002 | | | | 41 | 15/08/2007 | | | |
| 17 | 23/07/2002 | | | | 42 | 07/11/2007 | | | |
| 18 | 21/08/2002 | | | | 43 | 05/03/2008 | | | |
| 19 | 30/10/2002 | | | | 44 | 13/08/2008 | | | |
| 20 | 19/11/2002 | | | | 45 | 12/11/2008 | | | |
| 21 | 17/12/2002 | | | | 46 | 11/03/2009 | | | |
| 22 | 04/05/2003 | | | | 47 | 20/05/2009 | | | |
| 23 | 04/08/2003 | | | | 48 | 03/09/2009 | | | |
| 24 | 01/10/2003 | | | | 49 | 10/11/2009 | | | |
| 25 | 01/12/2003 | | | | 50 | 11/11/2009 | | | |

| | | | | | |
|--|--|--|--|--|------------------------------------|
| <table border="1" style="border-collapse: collapse; width: 100px; height: 30px;"> <tr><td style="background-color: white; width: 50%;"></td></tr> <tr><td style="background-color: #cccccc; width: 50%;"></td></tr> <tr><td style="background-color: #333333; width: 50%;"></td></tr> </table> | | | | Discarded Calibration Validation | N: North C: Central S: South |
| | | | | | |
| | | | | | |
| | | | | | |

sites or to the bad quality of the images from MODIS quality product control.

Multivariable models to retrieve Chl-*a* concentration

Approaches based on MLR (using exclusion-inclusion sequential models), ANN and NPMR were used to retrieve Chl-*a* concentrations, using the MOD09GA and MYD09GA reflectance data products.

An empirical model based on MLR was developed, using two search criteria. The first used a step-wise forward regression from the reflectance for each of the spectral bands from MODIS. This approach was carried out taking into account that the *F* value was low and the *p*-value was high. This method generates a MLR of the bands with the highest correlation. The second development criterion was based on a sequential automatic search of exclusion-inclusion using the Akaike Information Criterion – AIC (BOZDOGAN, 1987). Using this criterion, an R code was developed to fit a sub-model for each of the three datasets (North, Central and South points) and for the total dataset in Lake Mangueira.

An ANN was applied using a hidden layer (Figure 2). The reflectances for each of the spectral bands for three points (North, Central and South) and the total dataset were the input values and the Chl-*a* concentration was the output node. The training algorithm used was the back-propagation method with heuristic techniques acceleration (VOGL et al., 1988) performed in MatLab®. The Purelin function (DEMUTH; BEALE; HAGAN, 2008) was used as the activation function in the hidden and output layers.

NPMR using the software *HyperNiche 1.0* was also used to select the best models describing the relationships between Chl-*a* and the reflectance values from the MODIS spectral bands (McCUNE, 2006; 2004). Two matrices were used as input in NPMR: (1) a matrix of *in-situ* Chl-*a* measurements, and (2) a matrix of reflectance values from the MODIS spectral bands. Here, the local mean NPMR, uniform weights (Species Occurrence model – SpOcc) was used (PETERSON, 2000; McCUNE et al., 2003). This model uses a simple kernel function to estimate the response variable (Chl-*a* in our case) and gives equal weight to all sampling points within the window, while all observations outside the window are given zero weight. In addition, we assessed the performance of four models (three empirical models (Kahru, Appel and O14a) and one semi-empirical model (FAI), Table 3), that are commonly used to estimate Chl-*a* concentrations in aquatic

environments (CURTARELLI et al., 2015; HUANG et al., 2014; HU et al., 2010; TARRANT; NEUER, 2009).

Model accuracy assessment

To evaluate the accuracy of the models, the coefficient of determination (R^2), Bias, Root-Mean-Square-Error (RMSE), and relative RMSE (RMSE%) were calculated between *in-situ* observed and model-estimated Chl-*a* concentrations.

RESULTS

Model development and calibration

The performance of the multiple linear regression, ANN, and NPMR models for the three points (North, Central and South) and the total dataset for Lake Mangureira showed good agreement for the Central and South datasets. In addition, the MLR model performed best for each dataset (Figure 3).

The MLR model assessment indicated that it performed best for the South point ($R^2 = 0.83$, Bias = 0.00, RMSE% = 10.6%), compared to the Central point ($R^2 = 0.77$, Bias = -0.31, RMSE% = 14.6%) and the North point ($R^2 = 0.61$, Bias = 0.00, RMSE% = 19.6%). Also, the MLR model indicated that the most important MODIS channels for the correlation with *in-situ*

Chl-*a* values at North point were the spectral bands 2, 3 and 6, while that for the Central and South points, were the spectral bands 1, 4, 5, 6 and 1, 2, 5, 6 respectively. For the total dataset, spectral bands 1, 3, 5 and 6 were the most important MODIS channels.

The MLR models adjusted to the four datasets used are listed as follows:

$$MLR_{North} = 1.13 + (-584.84 * R(\lambda_{b2})) + (721.14 * R(\lambda_{b3})) + (33.33 * R(\lambda_{b6})) \quad (1)$$

$$MLR_{Central} = 30.26 + (950.36 * R(\lambda_{b1})) + (-1032.97 * R(\lambda_{b4})) + (510.97 * R(\lambda_{b5})) + (-406.31 * R(\lambda_{b6})) \quad (2)$$

$$MLR_{South} = -3.12 + (616.28 * R(\lambda_{b1})) + (-625.52 * R(\lambda_{b2})) + (127.10 * R(\lambda_{b5})) + (-33.21 * R(\lambda_{b6})) \quad (3)$$

$$MLR_{totaldataset} = 9.23 + (36.95 * R(\lambda_{b1})) + (-52.92 * R(\lambda_{b3})) + (14.47 * R(\lambda_{b5})) + (-35.73 * R(\lambda_{b6})) \quad (4)$$

where $R(\lambda_{b1})$, $R(\lambda_{b2})$, $R(\lambda_{b3})$, $R(\lambda_{b4})$, $R(\lambda_{b5})$, $R(\lambda_{b6})$ are the reflectance values of MODIS spectral bands 1, 2, 3, 4, 5 and 6 and MLR is the estimated Chl-*a* in $\mu\text{g}\cdot\text{L}^{-1}$.

The ANN model performed best for the Central point ($R^2 = 0.62$, Bias = 0.01, RMSE% = 16.0%), compared to the South point ($R^2 = 0.41$, Bias = -5.19, RMSE% = 18.7%), the North point ($R^2 = 0.01$, Bias = 2.72, RMSE% = 19.6%), and the total dataset ($R^2 = 0.21$, Bias = 2.42, RMSE% = 18.4%). The ANN model architecture was composed of 3 layers (input, hidden and output layers): the input layer had the same MODIS channel inputs as the MLR models, the hidden layer had several neurons based on each dataset (North: 12, Central: 4; South: 10), and the output had one neuron (Chl-*a*). Thus, the ANN architecture was: North: 3–12–1; Central: 4–4–1 and South: 4–10–1.

The results from the NPMR models indicated the best performance for the South point ($R^2 = 0.54$, Bias = 0.43, RMSE% = 15.4%), compared to the Central point ($R^2 = 0.43$, Bias = 1.37, RMSE% = 11.3%), the North point ($R^2 = 0.06$, Bias = 1.57, RMSE% = 28.4%) and the total dataset ($R^2 = 0.17$, Bias = 0.70, RMSE% = 13.33%). For the three points in the lake, we used as the predictor matrix (input data), the reflectance values from all MODIS spectral bands (1, 2, 3, 4, 5 and 6). The results indicated that, for the Central and South points, the best performance was obtained with the combination of spectral bands 1, 4 and 5, 6 respectively. For the North and the total

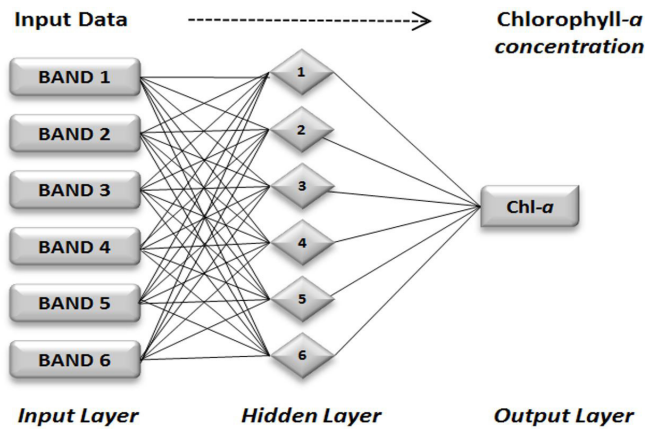


Figure 2. Schematic diagram of an artificial neural-network model.

Table 3. *Kahru*, *APPEL*, *FAI*, and *O14a* models to estimate Chl-*a* concentrations in Lake Mangureira.

| Model | Author |
|---|--------------------------|
| $Kahru = R(\lambda_{b2}) - R(\lambda_{b1})$ | Kahru et al. (2004) |
| $APPEL = R(\lambda_{b2}) - [(R(\lambda_{b3}) - R(\lambda_{b2})) * R(\lambda_{b2}) + R(\lambda_{b2}) - R(\lambda_{b1})]$ | El-Alem et al. (2012) |
| $FAI = R(\lambda_{b2}) - [R(\lambda_{b1}) + (R(\lambda_{b5}) - R(\lambda_{b1})) * \frac{(\lambda_{b2} - \lambda_{b1})}{(\lambda_{b5} - \lambda_{b1})}]$ | Hu and He (2008) |
| $O14a = 5.6897 + (203.5419 * R(\lambda_{b1})) + (-151.0271 * R(\lambda_{b4})) + (83.5912 * R(\lambda_{b5})) + (-170.4078 * R(\lambda_{b6}))$ | Ogashawara et al. (2014) |

where: $R(\lambda_{b1})$, $R(\lambda_{b2})$, $R(\lambda_{b3})$, $R(\lambda_{b4})$, $R(\lambda_{b5})$, $R(\lambda_{b6})$ are the reflectance of bands 1 to 6 respectively with $\lambda_{b1} = 645 \text{ nm}$, $\lambda_{b2} = 859 \text{ nm}$, $\lambda_{b3} = 470 \text{ nm}$, $\lambda_{b4} = 555 \text{ nm}$, $\lambda_{b5} = 1250 \text{ nm}$ and $\lambda_{b6} = 1640 \text{ nm}$.

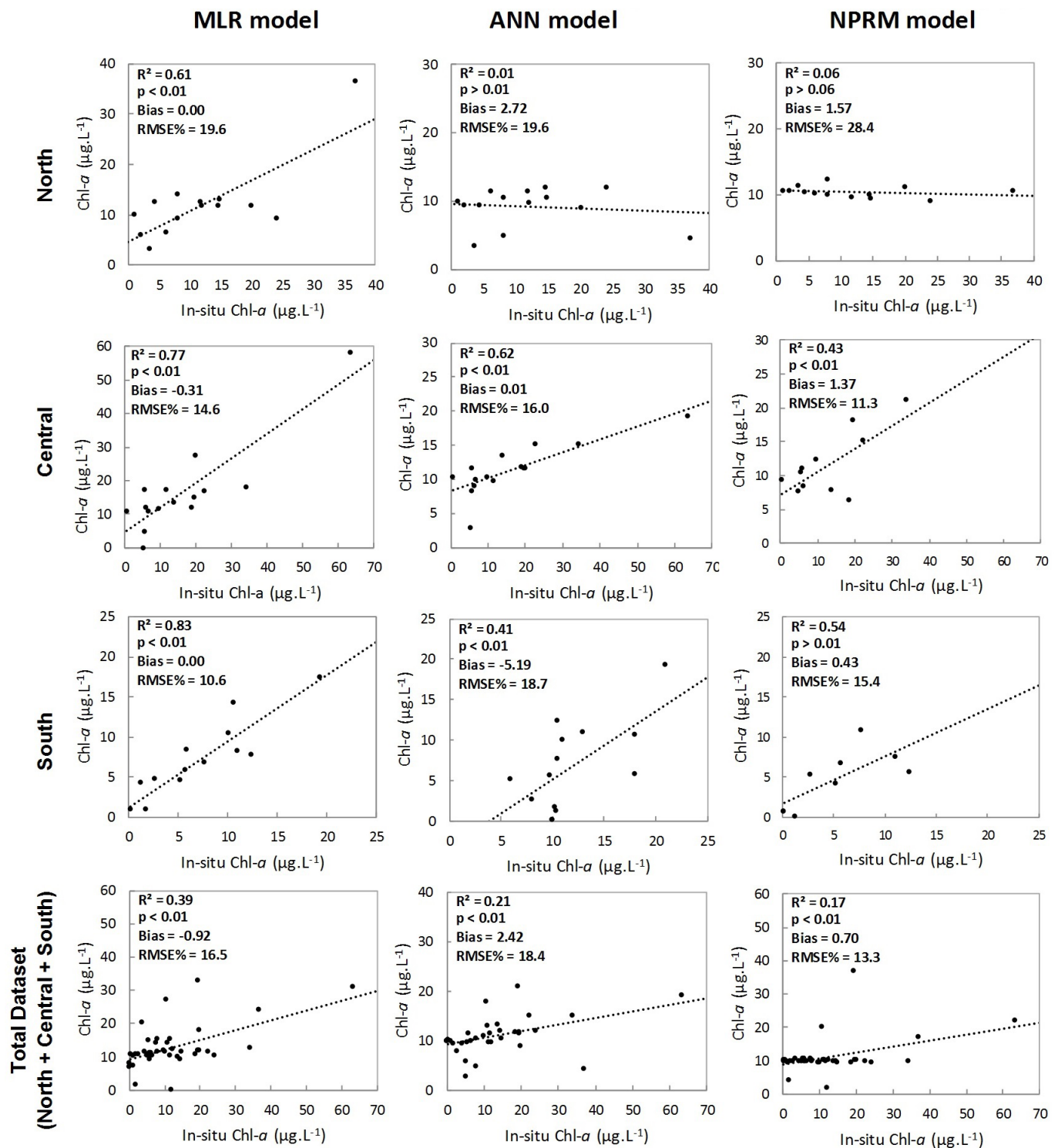


Figure 3. Chl-*a* concentrations predicted by MLR, ANN and NPRM models versus *in-situ* chlorophyll-*a* measurements at three points (North, Central and South) and the total dataset in Lake Mangueira during calibration. The dotted lines represent the best fit.

dataset, the best performance was obtained with the combination of spectral bands 1, 6 and 1, 2, 5, 6 respectively.

Amongst the four models evaluated, the Appel model performed better than the Kahru, FAI and O14a models (Figure 4). The Appel model assessment indicated that it performed best for the Central point ($R^2 = 0.66$, Bias = 0.58, RMSE% = 35.2%), compared

to the North point ($R^2 = 0.38$, Bias = 6.53, RMSE% = 43.9%), the South point ($R^2 = 0.57$, Bias = 5.64, RMSE% = 35.2%) and the total dataset ($R^2 = 0.29$, Bias = 11.6, RMSE% = 26%).

Despite the O14a model showing a good agreement for the North point ($R^2 = 0.52$, Bias = 6.93, RMSE% = 31.4%) and the South point ($R^2 = 0.52$, Bias = 4.04, RMSE% = 29.5%), this

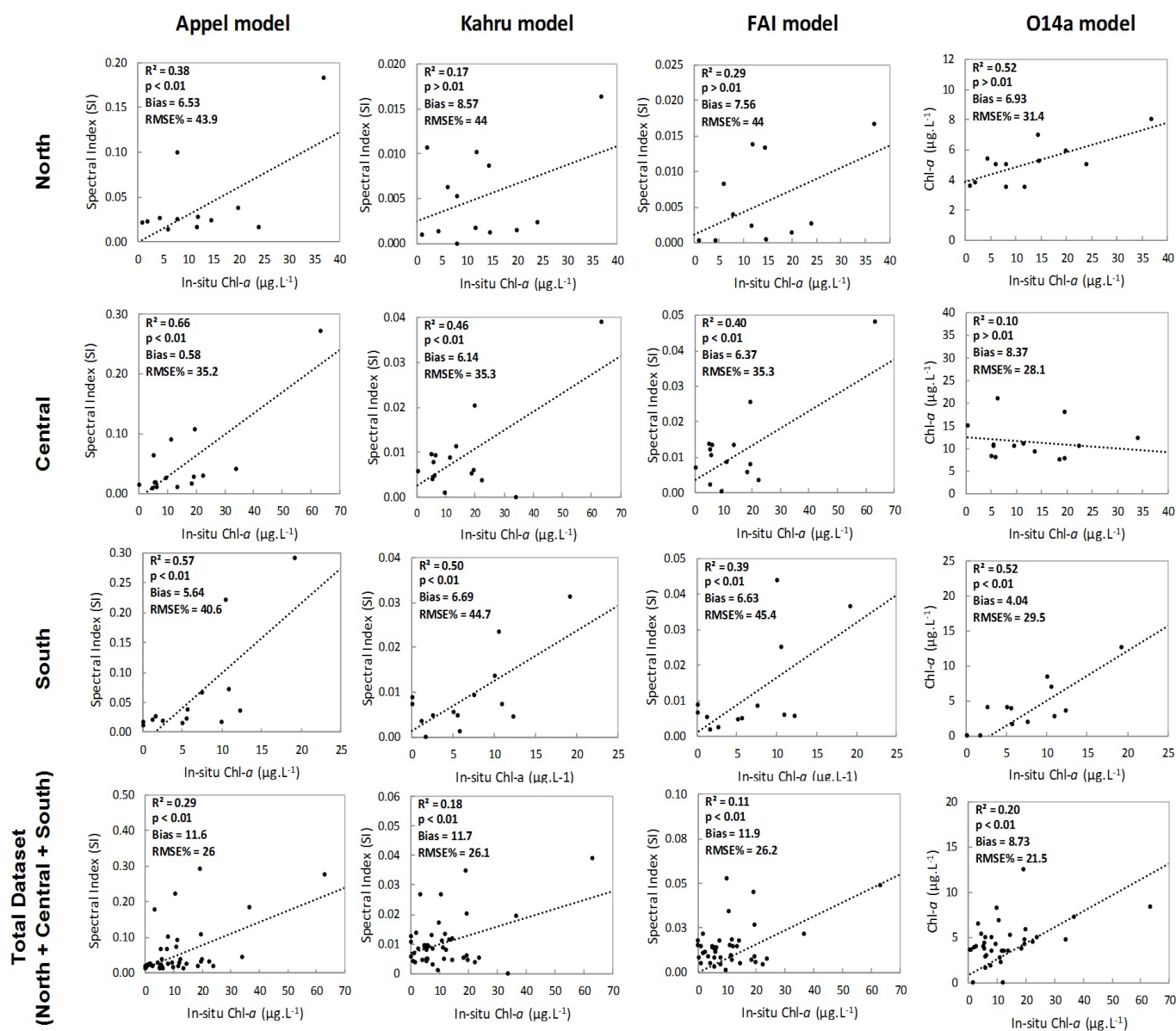


Figure 4. Plots of the Appel, Kahru, FAI and O14a models values versus *in-situ* chlorophyll-*a* measurements at three points (North, Central and South) and the total dataset in Lake Mangueira during calibration. The dotted lines represent the best fit.

model showed the worse results for the Central point ($R^2 \leq 0.10$, Bias ≤ 8.37 , RMSE% $\leq 28.1\%$).

Results from the Kahru and FAI models showed similar performance, being the adjustment for the Kahru model slightly better ($R^2 \leq 0.50$, Bias ≤ 6.69 , RMSE% $\leq 44.7\%$), that the FAI model ($R^2 \leq 0.40$, Bias ≤ 6.37 , RMSE% $\leq 35.3\%$).

Model validation and comparative assessment

The performance of the different models during the validation phase showed, in general, lower values of the accuracy metrics (Figure 5). However, the MLR and ANN models showed a smaller loss of performance ($R^2 = 0.50$, Bias = -1.36 , RMSE% = 29.9% and $R^2 = 0.46$, Bias = -1.78 , RMSE% = 24.6%) and maintained

satisfactory accuracy compared to the NPMR model (*e.g.*, Central point).

In all cases, the performance of our multivariable models using the total dataset was lower than with the specific point dataset. The same occurred in comparisons against four commonly used models, Appel, Kahru, FAI, and O14a (Figure 6).

We found that the MLR was reliable in terms of its estimation of Chl-*a* concentrations for the three points in the lake, although it was less accurate for the South point ($R^2 = 0.37$, Bias = -3.54 , RMSE% = 34.40%). The ANN model performed best for the Central point ($R^2 = 0.46$, Bias = -1.78 , RMSE% = 24.60%), but did not provide satisfactory results for the North and South points. The NPMR model had limitations in retrieving the Chl-*a* concentration for the North (N), Central (C) and South (S) points ($R^2 \leq 0.19$, Bias ≤ 5.37 , RMSE% $\leq 42.50\%$). The assessments with the four

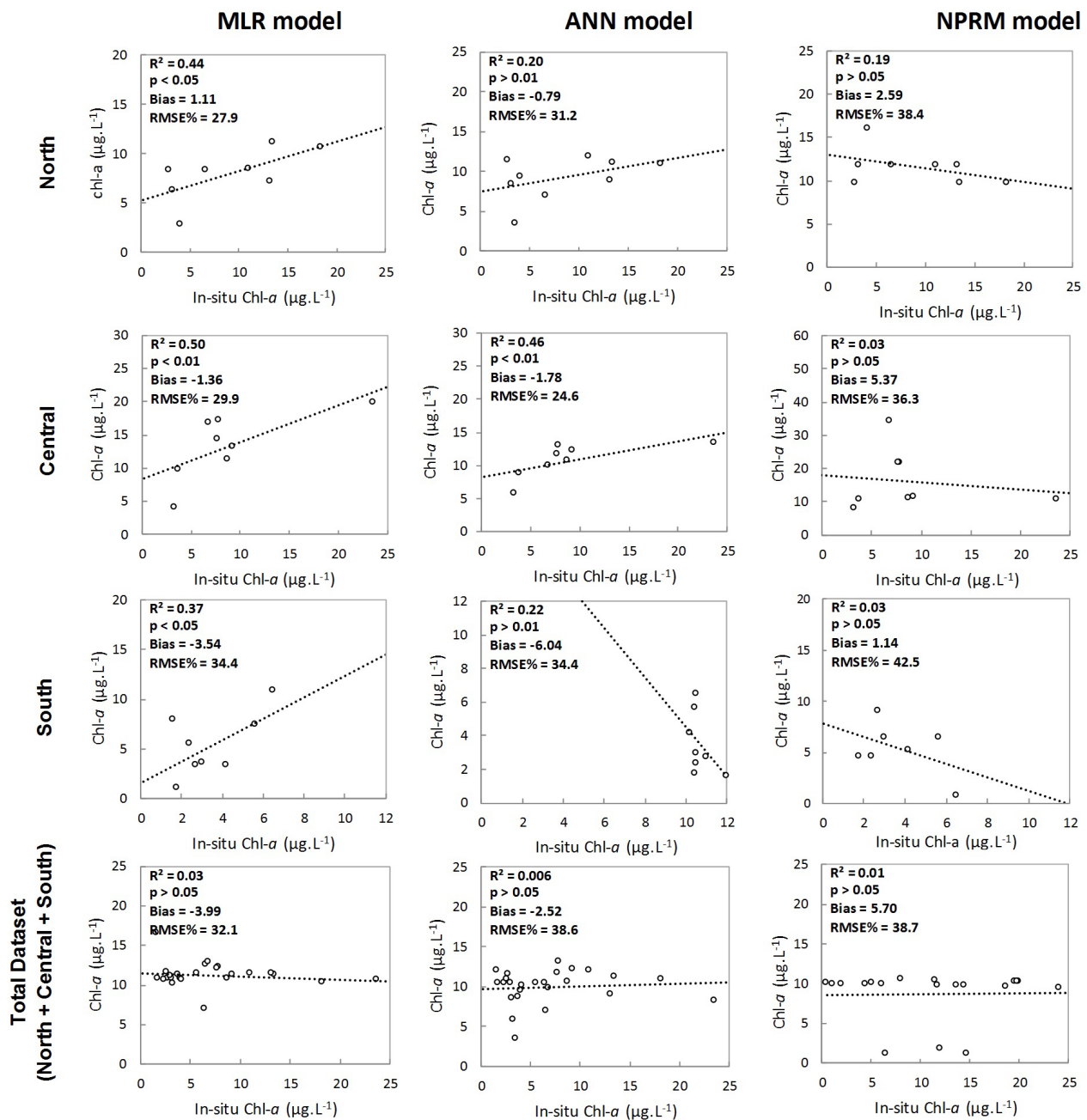


Figure 5. Chl-*a* concentrations predicted by MLR, ANN and NPRM models versus *in-situ* chlorophyll-*a* measurements at three points (North, Central and South) and the total dataset in Lake Mangueira during validation. The solid lines represent the 1:1 (one-to-one) relationship, and the dotted lines represent the best fit.

models (Appel, Kahru, FAI and O14a) showed lower accuracy than the MLR model (Figure 6).

The Kahru model gave the best Chl-*a* retrieval ($R^2 \leq 0.37$, Bias ≤ 9.09 , RMSE% $\leq 36.4\%$), followed by the O14a model ($R^2 \leq 0.35$, Bias ≤ 5.53 , RMSE% $\leq 39.97\%$); the Appel model gave the poorest results ($R^2 \leq 0.26$, Bias ≤ 9.07 , RMSE% $\leq 35.04\%$). Taking into account the total dataset, the accuracy yielded by the MLR, ANN and NPRM models and the four models (Appel, Kahru, FAI and O14a) also showed low performance compared with the *in-situ* Chl-*a* values.

DISCUSSION

Performance of Chl-*a* algorithms

For the models proposed here, the MLR and ANN models performed better than the other models (R^2 and RMSE%, Figure 6), in agreement with the performance found previously for oligo- to mesotrophic waters using MLR models (CURTARELLI et al., 2015; OGASHAWARA et al., 2014) and for eutrophic shallow waters using Back Propagation neural networks (BPs) and

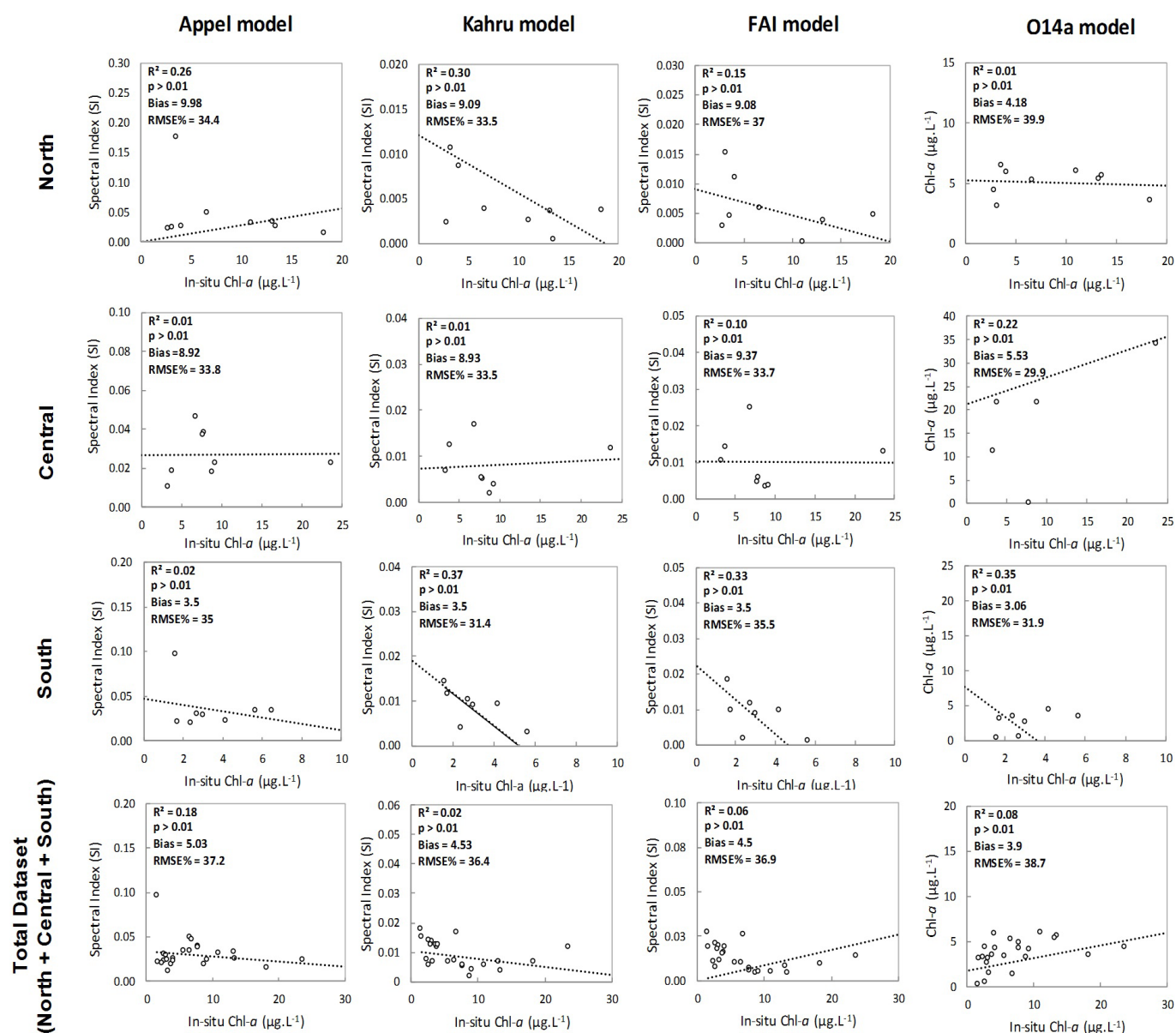


Figure 6. Plots of the Appel, Kahru, FAI and O14a models values versus *in-situ* chlorophyll-*a* measurements at three points (North, Central and South) and the total dataset in Lake Mangueira during validation. The dotted lines represent the best fit.

Radial Basis Function neural networks (RBFs) (WU et al., 2009). Nevertheless, the performance of models proposed and tested during the calibration period was better than for the validation period. This loss of performance during the validation models is a behavior expected, which can be explained by the size of samples and the difference in the datasets used, as has been observed in others studies (GIANCRISTOFARO; SALMASO, 2007; ORTH et al., 2015).

Although the NPMR model is generally stronger than the MLR model (McCUNE et al., 2003; YOST, 2008), our results indicated a significant difference in the performance of the two models. Also, the NPMR model showed lower performance, which was apparent during both the calibration (Figure 3) and the validation (Figure 5) periods. This result indicates that the application of NPMR models to retrieve the Chl-*a* concentration from MODIS

imagery in Lake Mangueira would not be recommended based on the dataset available. For the four models assessed (Appel, Kahru, FAI and O14a), our findings obtained from the total datasets were consistent with those of El-Alem et al. (2012), who obtained satisfactory results for Chl-*a* > 50 $\mu\text{g.L}^{-1}$ using the Kahru and FAI models

In Lake Mangueira, the Chl-*a* concentration estimated using the total dataset indicated that the correlation was the lowest for every model proposed as shown in the calibration (Figure 3), validation (Figure 5) and the four models assessed (Figure 6). Therefore, it would not be appropriate to generalise a single model to retrieve the Chl-*a* concentration in Lake Mangueira from a total dataset, but rather would be necessary to subdivide the datasets into regions. This can be explained by meso-oligotrophic conditions and the spatial heterogeneity, where the Chl-*a* concentration decreases from

north to south, as has been observed or taken into account in other studies (CROSSETTI et al., 2013; FRAGOSO JUNIOR et al., 2011; RODRIGUES, 2009). External forcing factors, such as the wind (both speed and direction) also can affect the temporal and spatial distribution of phytoplankton in large shallow lakes (CARDOSO et al., 2012; CARRICK; ALDRIDGE; SCHELSKE, 1993; KONG; FAO, 2005; WEBSTER; HUTCHINSON, 1994). In these environments, wind conditions impose stress on the system, in which the biological responses (*e.g.*, Chl-*a*) can be interconnected, in contrast to environments such as temperate lakes and oceans. In Lake Mangueira, a previous study (FRAGOSO JUNIOR et al., 2008) showed a trend of phytoplankton aggregation in the southwest and northeast areas, varying with the wind velocity and direction, and coinciding with the orientation of the lake. In addition, another key influencing factor is the larger biomass of submerged macrophytes in the South area, which can inhibit phytoplankton production in this area (THEY et al., 2014) and increases the water transparency (FERREIRA, 2009).

Regarding the proposed models, our results corroborated the spatial heterogeneity in the Chl-*a* distribution in the Lake Mangueira, as observed previously in other studies (CROSSETTI et al., 2013; FRAGOSO JUNIOR et al., 2011; 2008).

The use of models to retrieve Chl-*a* based on MLR from MODIS data is not new. However, our approach showed an improvement over previously published models (*e.g.*, CHANG et al., 2012; OGASHAWARA et al., 2014; WU et al., 2009) due to the implementation of search criteria based on a sequential automatic search of exclusion-inclusion algorithms for selecting parameters. In addition, the estimation of Chl-*a* concentrations based on MLR presented advantages (*e.g.*, few input variables and explicit calculation of the coefficients of selected independent variables) in comparison with the ANN and NPMR models, which is difficult to interpret the relationship between input and output variables (“*blackbox*” models) (WU et al., 2014). Besides the limitations from the MLR model related to the applicability and transferability between different sensors and water bodies, and the biological interpretation of principal components of the regression (GRAHAM, 2003), the MLR model is parsimonious, requiring only a small number of parameters compared to the other models.

Algorithm limitations for estimating Chl-*a* from MODIS imagery

Several studies have shown that algorithms for retrieving Chl-*a* from MODIS imagery may indeed be precluded due to the low signal-to-noise ratio and the noise-equivalent Chl-*a* concentration of MODIS (*e.g.*, CARDER; CANNIZZARO; LEE, 2005; GOWER; BORSTAD, 2004). This is a limitation in meso-oligotrophic environments such as Lake Mangueira, which can clip the performance of algorithms based on MODIS imagery, due to low Chl-*a* contents, and the adjacency effects that can be caused by sun glints, as founded in similar environments (CHAVULA et al., 2009; MATTHEWS; ODERMATT, 2015; WATANABE et al., 2015). Furthermore, at the South point of Lake Mangueira, located in the shallower area of the lake, water transparency is high (Table 1) and the light reflected off the bottom also affects the

accuracy of the empirical algorithms assessed. Another limitation associated with the use of satellite remote-sensing techniques is the detection of Chl-*a* in surface waters only. However, this limitation is not applicable to polyimtic shallow-water systems such as Lake Mangueira, where the Chl-*a* concentrations are relatively uniform throughout the water column except for short time periods (hours, see FRAGOSO JUNIOR et al., 2011).

CONCLUSION

This study evaluated the use of several models to retrieve Chl-*a* concentrations from MODIS Terra and Aqua reflectance data in a large shallow subtropical lake. Our findings suggested the MLR model as the better alternative compared to ANN and NPMR. In addition, the simple MLR model proposed in this study showed two main advantages: better performance and fewer input variables compared with others empirical and semi-empirical models that are commonly used (Appel, Kahru, FAI and O14a). Finally, although the purpose of this study was not to replace standard monitoring methods, the MLR model using MODIS reflectance data can provide a dynamic, broader-scale view of Chl-*a* concentrations. Further investigation could be focused on MLR model application to analyze in long-term scale, the temporal and spatial distribution of Chl-*a* concentration in Lake Mangueira, in order to understand the seasonal dynamic and the inter-annual variations. The overall methodology proposed here can be used as a tool to help reduce the effort and costs associated with conventional limnological monitoring of shallow waters, and as a data source to support distributed ecological mathematical modelling of these ecosystems.

ACKNOWLEDGEMENTS

We are grateful to the Brazilians funding agency *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES, www.capes.gov.br) and *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPQ, www.cnpq.br) for doctoral scholarships awarded to A.M. Munar and J.R. Cavalcanti. The study is part of the research project *Impacto das mudanças climáticas em ambientes fluviais e lacustres do Rio Grande do Sul*, funded by CNPq. We are also grateful to the Global Lake Ecological Observatory Network (GLEON, www.gleon.org) for providing a venue and resources for lake-science discussions, and to Dr. Janet W. Reid (Trumansburg, NY, U.S.A.) for revising the English text.

REFERENCES

- ACKERMAN, S. A.; STRABALA, K. I.; MENZEL, W. P.; FREY, R. A.; MOELLER, C. C.; GUMLEY, L. E. Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research, D, Atmospheres*, v. 103, n. D24, p. 32141-32157, 1998. <http://dx.doi.org/10.1029/1998JD200032>.
- ALLAN, M. G.; HAMILTON, D. P.; HICKS, B. J.; BRABYN, L. Landsat remote sensing of chlorophyll *a* concentrations in central North Island lakes of New Zealand. *International Journal*

- of *Remote Sensing*, v. 32, n. 7, p. 2037-2055, 2011. <http://dx.doi.org/10.1080/01431161003645840>.
- BOZDOGAN, H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika*, v. 52, n. 3, p. 345-370, 1987. <http://dx.doi.org/10.1007/BF02294361>.
- CARDER, K. L.; CANNIZZARO, J. P.; LEE, Z. Ocean color algorithms in optically shallow waters: Limitations and improvements. In: FROUIN, R. J.; BABIN, M.; SATHYENDRANATH, S. REMOTE SENSING OF THE COASTAL OCEANIC ENVIRONMENT, August 22, 2005, San Diego. *Proceedings...* San Diego: SPIE, 2005.
- CARDOSO, L. S.; FRAGOSO JUNIOR, C. R.; SIQUEIRA, R.; MARQUES, D. M. Hydrodynamic control of plankton spatial and temporal heterogeneity in subtropical shallow lakes. In: SCHULZ, H. E.; SIMÕES, A. L. A.; LOBOSCO, R. J. *Hydrodynamics: natural water bodies* InTech, 2012. chap. 2, p. 27-48. <http://dx.doi.org/10.5772/30669>.
- CARRICK, H.; ALDRIDGE, F.; SCHELSKE, C. L. Wind influences phytoplankton biomass and composition in a shallow, productive lake. *Limnology and Oceanography*, v. 38, n. 6, p. 1179-1192, 1993. <http://dx.doi.org/10.4319/lo.1993.38.6.1179>.
- CHANG, N.-B.; YANG, Y. J.; DARANPOB, A.; JIN, K.-R.; JAMES, T. Spatiotemporal pattern validation of chlorophyll-*a* concentrations in Lake Okeechobee, Florida, using a comparative MODIS image mining approach. *International Journal of Remote Sensing*, v. 33, n. 7, p. 2233-2260, 2012. <http://dx.doi.org/10.1080/01431161.2011.608089>.
- CHAVULA, G.; BREZONIK, P.; THENKABAIL, P.; JOHNSON, T.; BAUER, M. Estimating chlorophyll concentration in Lake Malawi from MODIS satellite imagery. *Physics and Chemistry of the Earth Parts A/B/C*, v. 34, n. 13, p. 755-760, 2009. <http://dx.doi.org/10.1016/j.pce.2009.07.015>.
- CROSSETTI, L. O.; BECKER, V.; CARDOSO, L. S.; RODRIGUES, L. R.; COSTA, L. S.; MOTTA-MARQUES, D. Is phytoplankton functional classification a suitable tool to investigate spatial heterogeneity in a subtropical shallow lake. *Limnological Ecology and Management of Inland Waters*, v. 43, n. 3, p. 157-163, 2013. <http://dx.doi.org/10.1016/j.limno.2012.08.010>.
- CURTARELLI, M.; OGASHAWARA, I.; ALCÂNTARA, E.; STECH, J. Coupling remote sensing bio-optical and three-dimensional hydrodynamic modeling to study the phytoplankton dynamics in a tropical hydroelectric reservoir. *Remote Sensing of Environment*, v. 157, p. 185-198, 2015. <http://dx.doi.org/10.1016/j.rse.2014.06.013>.
- DARECKI, M.; STRAMSKI, D. An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea. *Remote Sensing of Environment*, v. 89, n. 3, p. 326-350, 2004. <http://dx.doi.org/10.1016/j.rse.2003.10.012>.
- DEMUTH, H.; BEALE, M.; HAGAN, M. 2008. *Neural Network Toolbox™ 6: user's guide*. Available from: <<https://filer.case.edu/pjt9/b378s10/nnet.pdf>>. Access on: 30 aug 2017.
- DENG, Y.; ZHANG, Y.; LI, D.; SHI, K.; ZHANG, Y. Temporal and spatial dynamics of phytoplankton primary production in lake taihu derived from MODIS Data. *Remote Sensing*, v. 9, n. 3, p. 195, 2017. <http://dx.doi.org/10.3390/rs9030195>.
- DUAN, H.; MA, R.; XU, J.; ZHANG, Y.; ZHANG, B. Comparison of different semi-empirical algorithms to estimate chlorophyll-*a* concentration in inland lake water. *Environmental Monitoring and Assessment*, v. 170, n. 1-4, p. 231-244, 2010. <http://dx.doi.org/10.1007/s10661-009-1228-7>. PMID:19908152.
- DUAN, H.; TAO, M.; LOISELLE, S. A.; ZHAO, W.; CAO, Z.; MA, R.; TANG, X. MODIS observations of cyanobacterial risks in a eutrophic lake: Implications for long-term safety evaluation in drinking-water source. *Water Research*, v. 122, p. 455-470, 2017. <http://dx.doi.org/10.1016/j.watres.2017.06.022>. PMID:28624729.
- DUKA, S.; CULLAJ, A. Evaluation of chlorophyll as the primary index for trophic state classification. *Journal of Environmental Protection and Ecology*, v. 10, n. 2, p. 401-410, 2009.
- EATON, A. D.; CLESCERI, L. S. G. *Standard Methods for the examination of water and wastewater*. 20th ed. Washington: American Public Health Association Inc, 1999.
- EL-ALEM, A.; CHOKMANI, K.; LAURION, I.; EL-ADLOUNI, S. E. Comparative analysis of four models to estimate chlorophyll-*a* concentration in case-2 waters using moderate Resolution Imaging Spectroradiometer (MODIS) imagery. *Remote Sensing*, v. 4, n. 8, p. 2373-2400, 2012. <http://dx.doi.org/10.3390/rs4082373>.
- FERREIRA, T. F. *O papel das macrófitas submersas sobre a qualidade da água, restauração e conservação de lagos rasos subtropicais: estudo de caso, a Lagoa Mangueira, RS*. 2009. 164 f. Tese (Doutorado) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
- FRAGOSO JUNIOR, C. R.; MARQUES, D. M. L. M.; COLLISCHONN, W.; TUCCI, C. E. M.; VAN NES, E. H. Modelling spatial heterogeneity of phytoplankton in Lake Mangueira, a large shallow subtropical lake in South Brazil. *Ecological Modelling*, v. 219, n. 1, p. 125-137, 2008. <http://dx.doi.org/10.1016/j.ecolmodel.2008.08.004>.
- FRAGOSO JUNIOR, C. R.; MOTTA MARQUES, D. M. L.; FERREIRA, T. F.; JANSE, J. H.; VAN NES, E. H. Potential effects of climate change and eutrophication on a large subtropical shallow lake. *Environmental Modelling & Software*, v. 26, n. 11, p. 1337-1348, 2011. <http://dx.doi.org/10.1016/j.envsoft.2011.05.004>.
- GHOLIZADEH, M. H.; MELESSE, A. M.; REDDI, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors (Basel)*, v. 16, n. 8, p. 1298, 2016. <http://dx.doi.org/10.3390/s16081298>. PMID:27537896.

- GIANCRISTOFARO, R. A.; SALMASO, L. Model performance analysis and model validation in logistic regression. *Statistica*, v. 63, n. 2, p. 375-396, 2007.
- GILERSON, A. A.; GITELSON, A. A.; ZHOU, J.; GURLIN, D.; MOSES, W.; IOANNOU, I.; AHMED, S. A. Algorithms for remote estimation of chlorophyll-*a* in coastal and inland waters using red and near infrared bands. *Optics Express*, v. 18, n. 23, p. 24109-24125, 2010. <http://dx.doi.org/10.1364/OE.18.024109>. PMID:21164758.
- GOWER, J.; BORSTAD, G. On the potential of MODIS and MERIS for imaging chlorophyll-*a* fluorescence from space. *International Journal of Remote Sensing*, v. 25, n. 7-8, p. 1459-1464, 2004. <http://dx.doi.org/10.1080/01431160310001592445>.
- GRAHAM, M. H. Confronting multicollinearity in ecological multiple regression. *Ecology*, v. 84, n. 11, p. 2809-2815, 2003. <http://dx.doi.org/10.1890/02-3114>.
- GURLIN, D.; GITELSON, A. A.; MOSES, W. J. Remote estimation of chl-*a* concentration in turbid productive waters — Return to a simple two-band NIR-red model? *Remote Sensing of Environment*, v. 115, n. 12, p. 3479-3490, 2011. <http://dx.doi.org/10.1016/j.rse.2011.08.011>.
- HONEYWILL, C.; PATERSON, D.; HAGERTHEY, S. Determination of microphytobenthic biomass using pulse-amplitude modulated minimum fluorescence. *European Journal of Phycology*, v. 37, n. 4, p. 485-492, 2002. <http://dx.doi.org/10.1017/S0967026202003888>.
- HU, C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, v. 113, n. 10, p. 2118-2129, 2009. <http://dx.doi.org/10.1016/j.rse.2009.05.012>.
- HU, C.; HE, M.-X. Origin and offshore extent of floating algae in Olympic sailing area. *Eos (Washington, D.C.)*, v. 89, n. 33, p. 302-303, 2008. <http://dx.doi.org/10.1029/2008EO330002>.
- HU, C.; LEE, Z.; MA, R.; YU, K.; LI, D.; SHANG, S. Moderate resolution imaging spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China. *Journal of Geophysical Research. Oceans*, v. 115, n. C4, p. CO4002, 2010. <http://dx.doi.org/10.1029/2009JC005511>.
- HUANG, C.; LI, Y.; YANG, H.; SUN, D.; YU, Z.; ZHANG, Z.; CHEN, X.; XU, L. Detection of algal bloom and factors influencing its formation in Taihu Lake from 2000 to 2011 by MODIS. *Environmental Earth Sciences*, v. 71, n. 8, p. 3705-3714, 2014. <http://dx.doi.org/10.1007/s12665-013-2764-6>.
- JESPERSEN, A. M.; CHRISTOFFERSEN, K. Measurements of chlorophyll-*a* from phytoplankton using ethanol as extraction solvent. *Archiv für Hydrobiologie*, v. 109, p. 445-454, 1987.
- KAHRU, M.; MICHELL, B. G.; DIAZ, A.; MIURA, M. MODIS detects a devastating algal bloom in Paracas Bay, Peru. *Eos (Washington, D.C.)*, v. 85, n. 45, p. 465-472, 2004. <http://dx.doi.org/10.1029/2004EO450002>.
- KASPRZAK, P.; PADISÁK, J.; KOSCHEL, R.; KRIENTZ, L.; GERVAIS, F. Chlorophyll-*a* concentration across a trophic gradient of lakes: an estimator of phytoplankton biomass. *Limnologia Ecology and Management of Inland Waters*, v. 38, n. 3, p. 327-338, 2008. <http://dx.doi.org/10.1016/j.limno.2008.07.002>.
- KNIGHT, J. F.; VOTH, M. L. Application of MODIS imagery for intra-annual water clarity assessment of Minnesota lakes. *Remote Sensing*, v. 4, n. 7, p. 2181-2198, 2012. <http://dx.doi.org/10.3390/rs4072181>.
- KOMICK, N.; COSTA, M.; GOWER, J. Bio-optical algorithm evaluation for MODIS for western Canada coastal waters: An exploratory approach using *in situ* reflectance. *Remote Sensing of Environment*, v. 113, n. 4, p. 794-804, 2009. <http://dx.doi.org/10.1016/j.rse.2008.12.005>.
- KONG, F.; FAO, G. Hypothesis on cyanobacteria bloom-forming mechanism in large shallow eutrophic lakes. *Acta Ecologica Sinica Shengtai Xuebao*, v. 25, n. 3, p. 589-595, 2005.
- KOTTEK, M.; GRIESER, J.; BECK, C.; RUDOLF, B.; RUBEL, F. World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, v. 15, n. 3, p. 259-263, 2006. <http://dx.doi.org/10.1127/0941-2948/2006/0130>.
- LE, C.; HU, C.; CANNIZZARO, J.; DUAN, H. Long-term distribution patterns of remotely sensed water quality parameters in Chesapeake Bay. *Estuarine, Coastal and Shelf Science*, v. 128, p. 93-103, 2013. <http://dx.doi.org/10.1016/j.ecss.2013.05.004>.
- LEE, Z.; CARDER, K. L.; CHEN, R. F.; PEACOCK, T. G. Properties of the water column and bottom derived from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data. *Journal of Geophysical Research. Oceans*, v. 106, n. C6, p. 11639-11651, 2001. <http://dx.doi.org/10.1029/2000JC000554>.
- LESHT, B. M.; BARBIERO, R. P.; WARREN, G. J. A band-ratio algorithm for retrieving open-lake chlorophyll values from satellite observations of the Great Lakes. *Journal of Great Lakes Research*, v. 39, n. 1, p. 138-152, 2013. <http://dx.doi.org/10.1016/j.jglr.2012.12.007>.
- LIMA, M. S.; MARQUES, D. M.; THEY, N. H.; MCMAHON, K. D.; RODRIGUES, L. R.; CARDOSO, L. S.; CROSSETTI, L. O. Contrasting factors drive within-lake bacterial community composition and functional traits in a large shallow subtropical lake. *Hydrobiologia*, v. 778, n. 1, p. 105-120, 2016. <http://dx.doi.org/10.1007/s10750-015-2610-9>.
- LPDAAC – The Land Processes Distributed Interface Active Archive Center. *US Geological Survey EROS Data Center*. LPDAAC, 2015. Available from: <<http://lpdaac.usgs.gov>>. Access on: 10 jan 2015.

- MATTHEWS, M. W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing*, v. 32, n. 21, p. 6855-6899, 2011. <http://dx.doi.org/10.1080/01431161.2010.512947>.
- MATTHEWS, M. W.; BERNARD, S.; ROBERTSON, L. An algorithm for detecting trophic status (chlorophyll-*a*), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment*, v. 124, p. 637-652, 2012. <http://dx.doi.org/10.1016/j.rse.2012.05.032>.
- MATTHEWS, M. W.; ODERMATT, D. Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters. *Remote Sensing of Environment*, v. 156, p. 374-382, 2015. <http://dx.doi.org/10.1016/j.rse.2014.10.010>.
- MCCUNE, B. 2004. *Nonparametric multiplicative regression for habitat modeling*. Available from: <<http://www.pcord.com/NPMRintro.pdf>>. Access on: 30 aug 2017.
- MCCUNE, B. Non-parametric habitat models with automatic interactions. *Journal of Vegetation Science*, v. 17, n. 6, p. 819-830, 2006. <http://dx.doi.org/10.1111/j.1654-1103.2006.tb02505.x>.
- MCCUNE, B.; BERRYMAN, S. D.; CISSEL, J. H.; GITELMAN, A. I. Use of a smoother to forecast occurrence of epiphytic lichens under alternative forest management plans. *Ecological Applications*, v. 13, n. 4, p. 1110-1123, 2003. [http://dx.doi.org/10.1890/1051-0761\(2003\)13\[1110:UOASTF\]2.0.CO;2](http://dx.doi.org/10.1890/1051-0761(2003)13[1110:UOASTF]2.0.CO;2).
- MOORE, T. S.; CAMPBELL, J. W.; DOWELL, M. D. A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sensing of Environment*, v. 113, n. 11, p. 2424-2430, 2009. <http://dx.doi.org/10.1016/j.rse.2009.07.016>.
- NOVO, E. M. L. M.; BARBOSA, C. C. F.; FREITAS, R. M.; SHIMABUKURO, Y. E.; MELACK, J. M.; PEREIRA FILHO, W. Seasonal changes in chlorophyll distributions in Amazon floodplain lakes derived from MODIS images. *Limnology*, v. 7, n. 3, p. 153-161, 2006. <http://dx.doi.org/10.1007/s10201-006-0179-8>.
- O'REILLY, J. E.; MARITORENA, S.; SIEGEL, D. A.; O'BRIEN, M. C.; TOOLE, D.; MITCHELL, B. G.; KAHRU, M.; CHAVEZ, F. P.; STRUTTON, P.; COTA, G. F. Ocean color chlorophyll *a* algorithms for SeaWiFS, OC2, and OC4: Version 4. *SeaWiFS Postlaunch Calibration and Validation Analyses*, v. 3, p. 9-23, 2000.
- OGASHAWARA, I. Terminology and classification of bio-optical algorithms. *Remote Sensing Letters*, v. 6, n. 8, p. 613-617, 2015. <http://dx.doi.org/10.1080/2150704X.2015.1066523>.
- OGASHAWARA, I.; ALCÂNTARA, E. H.; CURTARELLI, M. P.; ADAMI, M.; NASCIMENTO, R. F.; SOUZA, A. F.; STECH, J. L.; KAMPEL, M. Performance analysis of MODIS 500-m spatial resolution products for estimating chlorophyll-*a* concentrations in oligo- to meso-trophic waters case study: Itumbiara Reservoir, Brazil. *Remote Sensing*, v. 6, n. 2, p. 1634-1653, 2014. <http://dx.doi.org/10.3390/rs6021634>.
- OGASHAWARA, I.; MISHRA, D. R.; GITELSON, A. A. Remote sensing of inland waters: background and current state-of-the-art. In: MISHRA, D.; OGASHAWARA, I. *Bio-optical modeling and remote sensing of inland waters*. Elsevier, 2017. chap. 1, p. 1-24.
- OLMANSON, L. G.; BREZONIK, P. L.; BAUER, M. E. Evaluation of medium to low resolution satellite imagery for regional lake water quality assessments. *Water Resources Research*, v. 47, n. 9, p. W09515, 2011. <http://dx.doi.org/10.1029/2011WR011005>.
- ORTH, R.; STAUDINGER, M.; SENEVIRATNE, S. I.; SEIBERT, J.; ZAPPA, M. Does model performance improve with complexity? A case study with three hydrological models. *Journal of Hydrology (Amsterdam)*, v. 523, p. 147-159, 2015. <http://dx.doi.org/10.1016/j.jhydrol.2015.01.044>.
- PALMER, S. C. J.; HUNTER, P. D.; LANKESTER, T.; HUBBARD, S.; SPYRAKOS, E.; TYLER, A.; PRÉSING, M.; HORVÁTH, H.; LAMB, A.; BALZTER, H.; TÓTH, V. R. Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. *Remote Sensing of Environment*, v. 157, p. 158-169, 2015a. <http://dx.doi.org/10.1016/j.rse.2014.07.024>.
- PALMER, S. C. J.; KUTSER, T.; HUNTER, P. D. Remote sensing of inland waters: challenges, progress and future directions. *Remote Sensing of Environment*, v. 157, p. 1-8, 2015b. <http://dx.doi.org/10.1016/j.rse.2014.09.021>.
- PETERSON, E. B. *Analysis and prediction of patterns in lichen communities over the western Oregon landscape*. 2000. Thesis (PhD) - Oregon State University, Corvallis, Oregon, 2000.
- QI, H.; LU, J.; CHEN, X.; SAUVAGE, S.; SANCHEZ-PÉREZ, J. M. Water age prediction and its potential impacts on water quality using a hydrodynamic model for Poyang Lake, China. *Environmental Science and Pollution Research International*, v. 23, n. 13, p. 13327-13341, 2016. <http://dx.doi.org/10.1007/s11356-016-6516-5>. PMID:27023820.
- RITCHIE, J. C.; ZIMBA, P. V.; EVERITT, J. H. Remote sensing techniques to assess water quality. *Photogrammetric Engineering and Remote Sensing*, v. 69, n. 6, p. 695-704, 2003. <http://dx.doi.org/10.14358/PERS.69.6.695>.
- RODRIGUES, L. H. R. *Reguladores da dinâmica das comunidades planctônicas e íctica em ecossistemas límnicos subtropicais*. 2009. 126 f. Tese (Doutorado) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
- ROSA NETO, J. L.; FRAGOSO JUNIOR, C. R.; MALHADO, A. C. M.; LADLE, R. J. Spatio-temporal variability of chlorophyll-*a* in the coastal zone of northeastern Brazil. *Estuaries and Coasts*, v. 38, n. 1, p. 72-83, 2015. <http://dx.doi.org/10.1007/s12237-014-9809-2>.

- ROY, D. P.; BORAK, J. S.; DEVADIGA, S.; WOLFE, R. E.; ZHENG, M.; DESCLOITRES, J. The MODIS land product quality assessment approach. *Remote Sensing of Environment*, v. 83, n. 1, p. 62-76, 2002. [http://dx.doi.org/10.1016/S0034-4257\(02\)00087-1](http://dx.doi.org/10.1016/S0034-4257(02)00087-1).
- RUIZ-VERDÚ, A.; JIMENEZ, J. C.; LAZZARO, X.; TENJO, C.; DELEGIDO, J.; PEREIRA, M.; SOBRINO, J. A.; MORENO, J. 2016. Comparison of MODIS and Landsat-8 retrievals of chlorophyll-*a* and water temperature over Lake Titicaca. In GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 10-15 July 2016, Beijing. Nova Jersey: IEEE. *Proceedings...* . p. 7643-7646. <http://dx.doi.org/10.1109/IGARSS.2016.7730993>.
- SCHALLES, J. F.; GITELSON, A. A.; YACOBI, Y. Z.; KROENKE, A. E. Estimation of chlorophyll *a* from time series measurements of high spectral resolution reflectance in an eutrophic lake. *Journal of Phycology*, v. 34, n. 2, p. 383-390, 1998. <http://dx.doi.org/10.1046/j.1529-8817.1998.340383.x>.
- SHI, K.; ZHANG, Y.; XU, H.; ZHU, G.; QIN, B.; HUANG, C.; LIU, X.; ZHOU, Y.; LV, H. Long-Term Satellite Observations of Microcystin Concentrations in Lake Taihu during Cyanobacterial Bloom Periods. *Environmental Science & Technology*, v. 49, n. 11, p. 6448-6456, 2015. <http://dx.doi.org/10.1021/es505901a>. PMID:25936388.
- SHI, K.; ZHANG, Y.; ZHOU, Y.; LIU, X.; ZHU, G.; QIN, B.; GAO, G. Long-term MODIS observations of cyanobacterial dynamics in Lake Taihu: Responses to nutrient enrichment and meteorological factors. *Scientific Reports*, v. 7, p. 40326, 2017. <http://dx.doi.org/10.1038/srep40326>. PMID:28074871.
- SHUBHA, S. *Remote sensing of ocean colour in coastal, and other optically-complex, waters*. Hanover, Nova Hampshire: IOCCG, Dartmouth, 2000. 145 p. (Reports of the International Ocean-Colour Coordinating Group, 3). Available from: <<http://www.ioccg.org/reports/report3.pdf>>. Access on: 30 aug 2017.
- SONG, K.; LI, L.; TEDESCO, L.; LI, S.; DUAN, H.; LIU, D.; HALL, B.; DU, J.; LI, Z.; SHI, K.; ZHAO, Y. Remote estimation of chlorophyll-*a* in turbid inland waters: Three-band model versus GA-PLS model. *Remote Sensing of Environment*, v. 136, p. 342-357, 2013. <http://dx.doi.org/10.1016/j.rse.2013.05.017>.
- SRICHANDAN, S.; KIM, J. Y.; KUMAR, A.; MISHRA, D.; BHADURY, P.; MUDULI, P. R.; PATNAIK, A. K.; RASTOGI, G. Interannual and cyclone-driven variability in phytoplankton communities of a tropical coastal lagoon. *Marine Pollution Bulletin*, v. 101, n. 1, p. 39-52, 2015. <http://dx.doi.org/10.1016/j.marpolbul.2015.11.030>. PMID:26611863.
- TARRANT, P.; NEUER, S. Monitoring algal blooms in a southwestern US reservoir system. *Eos (Washington, D.C.)*, v. 90, n. 5, p. 38-39, 2009. <http://dx.doi.org/10.1029/2009EO050002>.
- THEY, N. H.; MARQUES, D. M. L. M.; SOUZA, R. S. Lower respiration in the littoral zone of a subtropical shallow lake. *Frontiers in Microbiology*, v. 3, p. 434, 2012. <http://dx.doi.org/10.3389/fmicb.2012.00434>. PMID:23293635.
- THEY, N. H.; MARQUES, D. M. L. M.; CROSSETTI, L.; BECKER, V. L. O.; CANTERLE, E.; RODRIGUES, L. R.; DE SOUZA CARDOSO, L.; FRAGOSO JÚNIOR, C. R. Phytoplankton ecological interactions in freshwater ecosystems – integrating relationships in subtropical shallow lakes. In: SEBASTIÁ, M. T. *Phytoplankton: biology, classification and environmental impacts*. New York: Nova Science Publishers, 2014. chap. 4, p. 73–129.
- VOGL, T. P.; MANGIS, J.; RIGLER, A.; ZINK, W.; ALKON, D. Accelerating the convergence of the back-propagation method. *Biological Cybernetics*, v. 59, n. 4–5, p. 257-263, 1988. <http://dx.doi.org/10.1007/BF00332914>.
- WANG, M.; SHI, W.; TANG, J. Water property monitoring and assessment for China's inland Lake Taihu from MODIS-Aqua measurements. *Remote Sensing of Environment*, v. 115, n. 3, p. 841-854, 2011. <http://dx.doi.org/10.1016/j.rse.2010.11.012>.
- WANG, Y.; LIU, D.; TANG, D. Application of a generalized additive model (GAM) for estimating chlorophyll-*a* concentration from MODIS data in the Bohai and Yellow Seas, China. *International Journal of Remote Sensing*, v. 38, n. 3, p. 639-661, 2017. <http://dx.doi.org/10.1080/01431161.2016.1268733>.
- WATANABE, S.; VINCENT, W. F.; REUTER, J.; HOOK, S. J.; SCHLADOW, S. G. A quantitative blueness index for oligotrophic waters: Application to Lake Tahoe, California–Nevada. *Limnology and Oceanography, Methods*, v. 14, n. 2, p. 100-109, 2015. <https://doi.org/10.1002/lom3.10074>.
- WEBSTER, I. T.; HUTCHINSON, P. A. Effect of wind on the distribution of phytoplankton cells in lakes revisited. *Limnology and Oceanography*, v. 39, n. 2, p. 365-373, 1994. <http://dx.doi.org/10.4319/lo.1994.39.2.0365>.
- WU, G.; LIU, L.; CHEN, F.; FEI, T. Developing MODIS-based retrieval models of suspended particulate matter concentration in Dongting Lake, China. *International Journal of Applied Earth Observation and Geoinformation*, v. 32, p. 46-53, 2014. <http://dx.doi.org/10.1016/j.jag.2014.03.025>.
- WU, M.; ZHANG, W.; WANG, X.; LUO, D. Application of MODIS satellite data in monitoring water quality parameters of Chaohu Lake in China. *Environmental Monitoring and Assessment*, v. 148, n. 1–4, p. 255-264, 2009. <http://dx.doi.org/10.1007/s10661-008-0156-2>. PMID:18231871.
- XIANG, B.; SONG, J. W.; WANG, X. Y.; ZHEN, J. Improving the accuracy of estimation of eutrophication state index using a remote sensing data-driven method: A case study of Chaohu Lake, China. *Water S.A.*, v. 41, n. 5, p. 753-761, 2015. <http://dx.doi.org/10.4314/wsa.v41i5.18>.
- YOST, A. C. Probabilistic modeling and mapping of plant indicator species in a Northeast Oregon industrial forest, USA. *Ecological*

Indicators, v. 8, n. 1, p. 46-56, 2008. <http://dx.doi.org/10.1016/j.ecolind.2006.12.003>.

ZHANG, Y.; LIN, S.; QIAN, X.; WANG, Q.; QIAN, Y.; LIU, J.; GE, Y. Temporal and spatial variability of chlorophyll *a* concentration in Lake Taihu using MODIS time-series data. *Hydrobiologia*, v. 661, n. 1, p. 235-250, 2011. <http://dx.doi.org/10.1007/s10750-010-0528-9>.

ZHANG, Y.; MA, R.; DUAN, H.; LOISELLE, S.; ZHANG, M.; XU, J. A novel MODIS algorithm to estimate chlorophyll *a* concentration in eutrophic turbid lakes. *Ecological Indicators*, v. 69, p. 138-151, 2016. <http://dx.doi.org/10.1016/j.ecolind.2016.04.020>.

Authors contributions

Andrés Mauricio Munar, David da Motta Marques and Carlos Ruberto Fragoso Jr developed the idea for the study, as well as its design and paper review.

J. Rafael Cavalcanti and Juan Martin Bravo helped in the methodology for the data analysis, data analysis, construction and validation of the models, support in statistical evaluation and paper review.

All authors were responsible for reviewing the written text and approved the manuscript submission.