

Estimation of chlorophyll-a concentration using Landsat 8 in the Cassaffousth reservoir

María Micaela Ledesma, Matías Bonansea, Claudia Rosa Ledesma, Claudia Rodríguez, Joel Carreño and Lucio Pinotti

ABSTRACT

The physico-chemical and biological composition of a reservoir's effluents directly influences water quality. The values of variables such as high values of concentrations of chlorophyll-a (Chl-a) are indicators of pollution. The objective of this work was to monitor the trophic status and water quality of the Cassaffousth reservoir (Córdoba, Argentina) through the development of statistical models based on field data and satellite information. During 2016 and 2017, samples were taken bimonthly. Seven sampling sites were selected and physico-chemical and biological parameters were assessed. By using regression techniques, Landsat 8 information was related with field data to construct and validate a statistical model to determine the distribution of Chl-a in the reservoir ($R^2 = 0.87$). The generated algorithm was used to generate maps which contained information about the dynamics of Chl-a in the entire reservoir. Remote sensing techniques can be used to expand the knowledge of the dynamics of the Cassaffousth reservoir. Moreover, these techniques can be used as baselines for the development of an early warning system for this and other reservoirs in the region.

Key words | chlorophyll-a, Landsat 8, predictive models, reservoir, water quality

María Micaela Ledesma (corresponding author)
Matías Bonansea
Claudia Rosa Ledesma
Claudia Rodríguez
Joel Carreño
Departamento de Estudios Básico y Agropecuarios, Facultad de Agronomía y Veterinaria (FAYV), Universidad Nacional de Río Cuarto (UNRC), Córdoba, Argentina
E-mail: mledesma@ayv.unrc.edu.ar

Matías Bonansea
Lucio Pinotti
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Río Cuarto, Córdoba, Argentina

INTRODUCTION

Water reservoirs are ecosystems of scientific interest. These ecosystems exist because of their uses, goods and services that benefit humans (Padedda *et al.* 2017). However, due to the growth of economic development and environmental deterioration, lake eutrophication is becoming an increasingly serious phenomenon, which is causing water crisis (Kiefer *et al.* 2015). Eutrophication, which is characterized by algal blooms and results in enrichment of nutrients, has become a worldwide environmental problem (Watanabe *et al.* 2017).

The remote sensing technique is a very useful technique for monitoring large-scale water quality with minimal costs and at more frequent revisit times (Chu *et al.* 2018). This technique has great potential to detect phytoplankton pigments, such as chlorophyll-a concentration (Chl-a) (Watanabe *et al.* 2017). The Chl-a in water is a key indicator of phytoplankton biomass. The increase of Chl-a is the first

manifestation of nutrient pollution, which, then, changes phytoplankton community composition and increases the frequency of nuisance and toxic blooms (Kiefer *et al.* 2015).

Although there has been a lot of research done on the quality of water using different satellites (Li *et al.* 2017), there have been few research studies that use the new Landsat 8 satellite. The Landsat program, developed by NASA (National Aeronautics and Space Administration) in association with the United States Geological Survey (USGS), has provided the longest and the most constant satellite record of the Earth's land surface since 1972. Landsat 8, which is the newest satellite in the Landsat program, was put in orbit to extend the ability to detect and quantitatively characterize changes on the land surface at a scale where natural and human-induced causes of change can be detected and differentiated (Loveland & Irons 2016).

The aim of this work was to develop statistical models based on field data and satellite information to determine and predict Chl-a in the Cassaffouth reservoir.

METHODOLOGY

Study area

Cassaffouth reservoir (32° 10' S, 64° 23' W) (Figure 1), located in the western region of Córdoba province, has an area of 86 ha, a volume of 10.5 hm³ and maximum and mean depths of 28.7 and 11.9 m, respectively. This reservoir is the fourth of a system of six chained reservoirs, which were built between 1930 and 1950 for multiple purposes such as water supply, power generation, flood control, irrigation, tourism and recreational activities (Bonansea et al. 2015a). Cassaffouth reservoir receives water from Río Tercero reservoir, which is the largest artificial lake in the province of Córdoba, and has a single effluent called Tercero river.

Field measurements

The water quality of the reservoir was determined bimonthly during 2016 and 2017 making *in situ* measurements and collecting water samples from seven sampling sites to carry out

the subsequent analysis in the laboratory (Figure 1). Water samples were taken at 20 cm depth. *In situ*, water surface temperature (WST), pH, dissolved oxygen (DO) and Secchi disk transparency (SDT) were evaluated. In the laboratory, we evaluated chlorophyll-a concentration (Chl-a) following standard analytical methods and protocols (APHA 2000).

The samples were carefully collected and handled so that the values of the variables were representative of the reservoir at the time of sampling. Also, samples should provide a real description of temporal and spatial variables of water quality during the sampling program (Bonansea et al. 2015b). The sampling, storage, conservation, transport and analytical techniques were carried out according to standardized methods (APHA 2000).

The bimonthly samples taken at the Cassaffouth reservoir were collected on the same date as the Landsat 8 satellite was passing through the study area. In 2016, samples were collected on June 11, August 30 and November 18. In 2017, samples were collected on February 22, May 29, August 17 and November 5.

Satellite data

Landsat 8, launched in 2013, carries two Earth-observing sensors, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIR). Improved capabilities of this

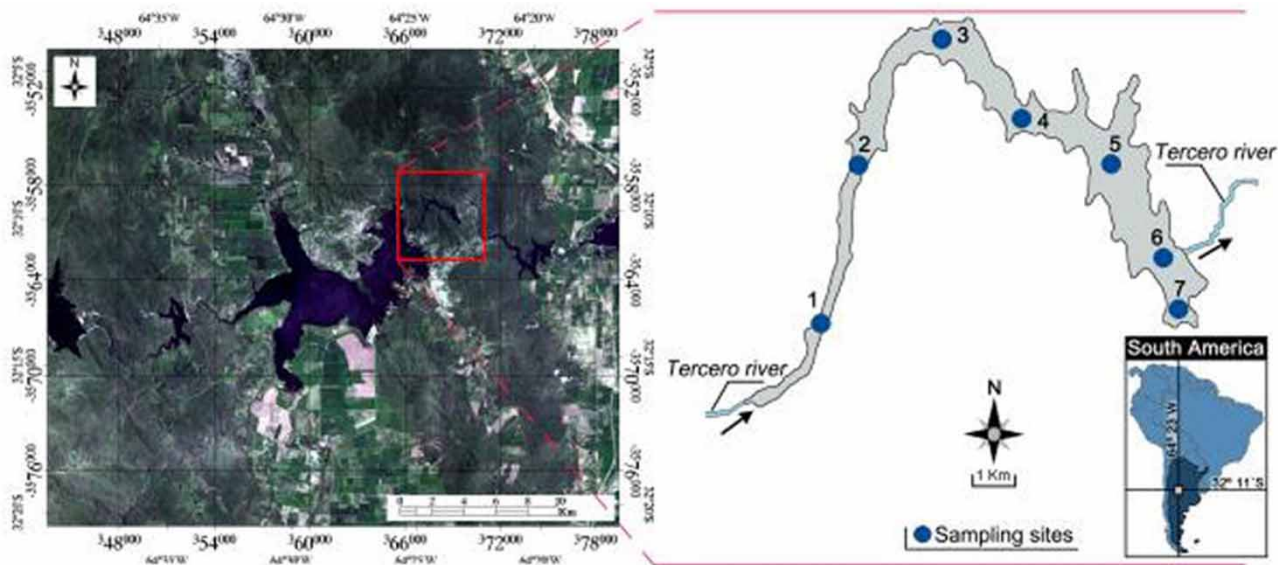


Figure 1 | Study area and location of sampling sites.

satellite include an additional shorter wavelength blue band (aerosol), a narrower near-infrared band, and a 12-bit radiometric resolution in contrast with the 8 bits of previous Landsat satellites (Olmanson *et al.* 2016) (Table 1).

Data collection in the field and satellite overpass must coincide in time before a natural variation of the process is generated (Olmanson *et al.* 2016). In the present study, we downloaded seven cloud-free Landsat 8 images (Path: 229; Row: 82) from the US Geological Survey (<https://glovis.usgs.gov/>). These images, whose date of acquisition correlates with the date of the sampling campaigns, have an L2 processing level, which is characterized by providing surface reflectance data (SR). SR products provide an estimate of the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption (USGS 2016). In addition, these images are terrain corrected, so it was not necessary to apply geometric corrections (USGS 2016).

Model development

Data for Chl-a obtained from each sampling site was loaded in a database, together with its GPS coordinates. Afterwards, a vector file of points was generated with the associated database. The points were transformed to ROIs (Region of Interest) for the extraction of the SR value of each Landsat band (Li *et al.* 2017). Thus, a function of expected values was designed by using multiple linear regression models (Chu *et al.* 2018).

Table 1 | Band specifications for Landsat 8

Electromagnetic region	Spatial (m)	Wavelength (μm)
Aerosol (b_{aerosol})	30	0.43–0.45
Blue (b_{blue})	30	0.45–0.51
Green (b_{green})	30	0.53–0.59
Red (b_{red})	30	0.64–0.67
NIR (b_{NIR})	30	0.85–0.88
SWIR-1 ($b_{\text{SWIR-1}}$)	30	1.57–1.65
TIR-1 ($b_{\text{TIR-1}}$)	100	10.60–11.19
TIR-2 ($b_{\text{TIR-2}}$)	100	11.50–12.51
SWIR-2 ($b_{\text{SWIR-2}}$)	30	2.11–2.29
Pan (b_{pan})	15	0.50–0.68
Cirrus (b_{cirrus})	30	1.36–1.38

Regression analysis is a technique used to investigate the relationship between a response variable (Y) and one or more regressor variables (X), which allows predicting or estimating the value of the variable Y for a given value of the variable X .

This analysis was carried out from the SR of the satellite image bands and from the Chl-a data collected in sampling sites, establishing the estimated response for the population model of linear regression. The validation of the model was carried out by testing the following assumptions: (1) normal distribution of the Chl-a, (2) existence of a linear regression between the variables X and Y , and (3) errors normally distributed with zero mean and variance σ^2 . In the multiple linear regression analysis, the radiance values of the bands or band ratios were considered as regressor variables.

The proposed model has a fixed and random component (Equation (1)):

$$y_i = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \dots + \beta_k * x_{ki} + \epsilon_i \quad (1)$$

where y_i is the value of the Chl-a concentration variable, $x_{1i}, x_{2i}, \dots, x_{ki}$ are the values of SR of the regressor variables, β_0, \dots, β_k are the constants and unknown parameters of the model and ϵ_i is the random error.

The dataset was randomly divided in two parts: 50% of the database was used to generate the model and 50% was used to validate the generated model. Radiance and Chl-a values were combined by linear regression (Equation (2)):

$$\hat{y}_i = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \dots + \beta_k * x_{ki} \quad (2)$$

where \hat{y}_i is the value of Chl-a estimated, $x_{1i}, x_{2i}, \dots, x_{ki}$ are the values of SR of the regressor variables and β_0, \dots, β_k are the estimators of the parameters of the multiple linear population regression model.

To validate the generated model, different statistical metrics were used: the Root Mean Square Error (RMSE), the Normalized Root Mean Error (NRMSE), and the determination coefficient (R^2).

Chl-a mapping

The values of Chl-a predicted by the estimated response were visualized in the Chl-a distribution maps. The color

palette of the Chl-a distribution map was designed so that the intensity of the color correlated with the highest concentration of the variable (Watanabe et al. 2017).

RESULTS AND DISCUSSION

Physical, chemical and biological analysis

The reservoir water was classified as fresh water with low calcium bicarbonate content according to the physico-chemical analysis. The water is suitable for irrigation because it contains sodium salt and residual sodium carbonate in low concentrations. In addition, it is not suitable for animal consumption because it has low salt content (Li et al. 2017). With respect to human consumption, the elements and properties analyzed meet the limits recommended by the World Health Organization (WHO 2006), which must be completed with microbiological analysis.

Table 2 summarizes the mean values, standard deviation, variation coefficients and minimum and maximum values for each of the variables studied in Cassaffouth reservoir in the different sampling campaigns of the year.

Regression model

To predict Chl-a in sites not sampled or far from the sampling points, a multiple linear regression analysis was carried out (Bonansea et al. 2015a). The linear regression model obtained, with the first 50% of the samples, was the following (Equation (3)):

$$\log Chl-a = 1.04 + 1.64 * \frac{b_{blue}}{b_{red}} - 0.53 * \frac{b_{SWIR-1}}{b_{pan}} - 7.26 * b_{pan} - 0.23 * \frac{b_{SWIR-2}}{b_{SWIR-1}} \quad (3)$$

where $\log Chl-a$ is the logarithm of chlorophyll-a concentration ($\mu\text{g/l}$); $\frac{b_{blue}}{b_{red}}$ is the relationship between the SR

Table 2 | Range, mean and standard deviation (std.) of measured parameters at different date of Cassaffouth reservoir

Date (Climate season ^a)		Measured parameters				
		WST (°C)	pH	OD (mg/L)	SDT (m)	Chl-a ($\mu\text{g.l}^{-1}$)
June, 16 (Autumn)	Range	12.10–12.90	8.22–8.90	6.80–7.90	3.00–5.50	1.76–6.23
	Mean	12.38	8.46	7.48	4.17	4.12
	Std.	0.29	0.26	0.38	1.03	1.92
August, 16 (Winter)	Range	11.70–13.20	6.88–7.22	9.90–12.80	4.00–7.00	18.86–40.44
	Mean	12.14	7.06	10.73	5.21	28.74
	Std.	0.52	0.12	0.98	1.07	7.76
November, 16 (Spring)	Range	18.00–20.60	6.90–7.46	4.90–7.40	2.05–3.50	15.25–39.07
	Mean	19.61	7.23	6.53	2.97	25.02
	Std.	0.82	0.19	0.85	0.47	7.39
February, 17 (Summer)	Range	18.70–29.40	6.99–7.32	4.80–8.90	1.00–1.50	2.25–10.76
	Mean	26.16	7.16	7.53	1.25	5.35
	Std.	4.90	0.12	1.43	0.20	2.98
May, 17 (Autumn)	Range	14.60–15.20	7.05–7.25	6.90–8.90	3.25–4.50	1.63–10.78
	Mean	14.89	7.18	8.14	3.96	5.74
	Std.	0.20	0.07	0.68	0.47	3.16
August, 17 (Winter)	Range	11.90–14.50	7.46–8.09	8.30–10.50	3.00–3.50	13.47–50.52
	Mean	12.89	7.88	9.41	3.39	36.86
	Std.	0.85	0.22	0.83	0.20	11.84
November, 17 (Spring)	Range	18.20–20.20	6.97–7.36	4.20–7.30	2.50–3.50	0.66–17.85
	Mean	19.37	7.19	6.03	2.86	12.05
	Std.	0.62	0.14	0.98	0.38	5.79

^aClimate season for the southern hemisphere.

values of the blue and red bands; $\frac{b_{\text{SWIR-1}}}{b_{\text{pan}}}$ is the relationship between the SR values of the SWIR-1 and pan bands; b_{pan} is the radiance value of the panchromatic band and $\frac{b_{\text{SWIR-2}}}{b_{\text{SWIR-1}}}$ is the relationship between the radiance values of the SWIR-2 and SWIR-1 bands.

Figure 2(a) shows the estimated values vs the measured values obtained with the generated model.

Validation procedure

The other 50% of the samples was used to validate the model obtained. In Figure 2(b) it is observed that the estimated values vs the measured values are distributed on both sides of the line 1:1, showing that the validation accuracy of the peak height model in other datasets is good.

In addition, the value of R^2 was acceptable, which shows a strong relationship between the observed and estimated values (Li et al. 2017). In summary, the methodology used for the present study clearly demonstrated the validity of the proposed model. The good performance of the model facilitates the study of temporal and spatial patterns in the studied reservoir (Zheng et al. 2016). However, the generated algorithm is specific to the data set used in the study. Consequently, it would be necessary to obtain a greater calibration set with images Landsat acquired on different dates to generate a multi-temporal study of Chl-a, which could improve the understanding of reservoir processes.

Finally, spatial distribution maps were produced (Figure 3). Based on the response estimated by the selected algorithm, it was possible to predict the concentration of Chl-a in unsampled sites and in sites far from the sampling sites, and thus, to evaluate the spatial variability of this parameter.

Satellite imagery showed that the highest values of Chl-a were recorded on August 30, 2016, and August 17, 2017, coinciding with field data.

Matthews (2011) suggests that there are a large number of investigations that use Landsat sensors to retrieve water quality characteristics. In the present study, a model based on the combination of different bands and band ratios of OLI sensors was applied to estimate Chl-a in Cassaffouth reservoir. Similar bands and band ratios were used in the algorithms developed by Tebbs et al. (2013) and Bonansea et al. (2015a) to make evident the potential use of this sensor for the determination of Chl-a concentration in different reservoirs.

We focused on a specific reservoir, but this powerful technique could be extrapolated to other aquatic systems. In addition, we highlight the importance of the use of remote sensors for the prediction of water quality.

CONCLUSIONS

Water quality is one of the main problems on our planet. In this paper, a water quality characteristic was examined. The

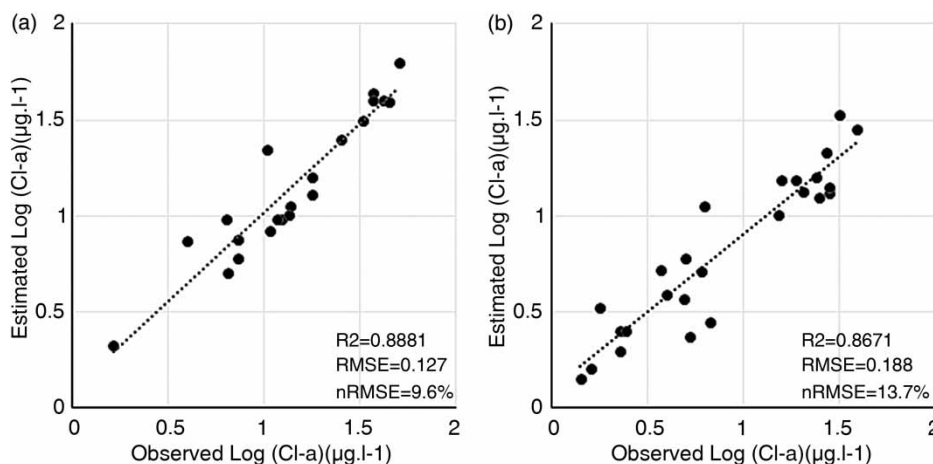


Figure 2 | (a) Log (Chl-a) estimated by the model vs log (Chl-a) observed in the field and (b) validation of log (Chl-a) estimated by the model vs log (Chl-a) observed in the field.

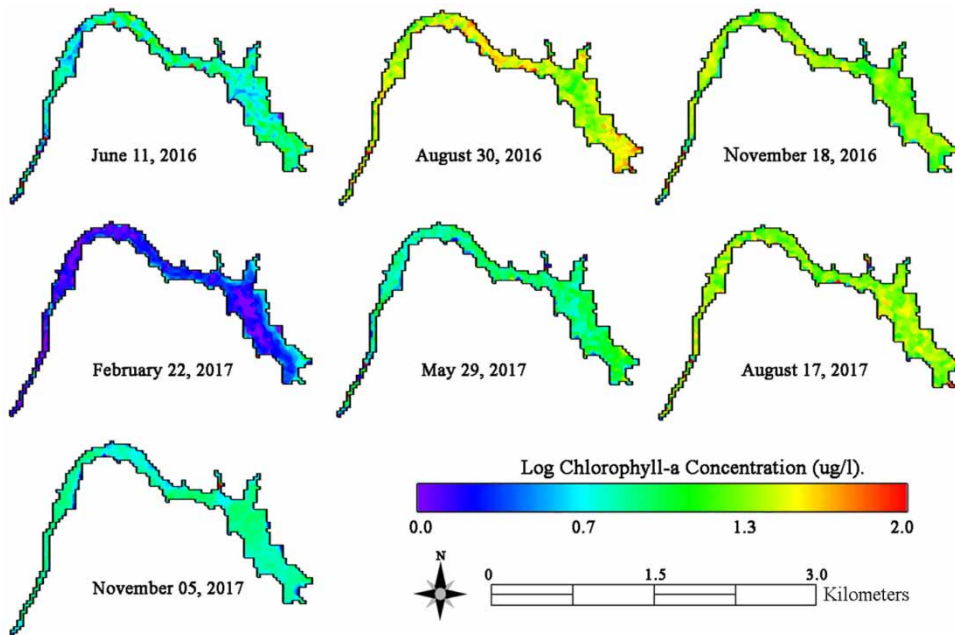


Figure 3 | Distribution of log (Chl-a) in the Cassaffouth reservoir during the study period.

resolutions of the Landsat 8 OLI satellite allowed us to successfully determine and estimate the spatial-temporal distribution of Chl-a in Cassaffouth reservoir. Multiple regression analysis gives significantly consistent information with good agreement of the used statistical metrics ($R^2 = 0.8671$, RMSE = 0.188, NRMSE = 13.7).

Also, the generated spatial distribution maps enabled the interpretation of water quality conditions in the whole study reservoir, demonstrating the potential use of integrating experimental data and remote sensing information for the study and monitoring of aquatic ecosystems.

ACKNOWLEDGEMENTS

This work was supported by the CONICET and the FONCyT (Fondo para la Investigación Científica y Tecnológica) [grant number PICT 1408/15]; and SECyT-UNRC (Secretaría de Ciencia y Técnica) [grant number 18A369]. The authors would like to thank Nautical Security Direction of Córdoba province for their collaboration in sampling campaigns. Furthermore, the anonymous reviewers and the editor are thanked for their insightful comments which helped to improve this paper.

REFERENCES

- American Public Health Association (APHA) 2000 *Standard Methods for the Examination of Water and Wastewater*. APHA/AWWA/WEF, Washington, DC, USA.
- Bonanseña, M., Rodríguez, M. C., Pinotti, L. & Ferrero, S. 2015a Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina). *Remote Sensing of Environment* **158**, 28–41.
- Bonanseña, M., Ledesma, C., Rodríguez, C. & Pinotti, L. 2015b Water quality assessment using multivariate statistical techniques in Río Tercero Reservoir, Argentina. *Hydrology Research* **46** (3), 377–388.
- Chu, H., Kong, S. & Chang, C. 2018 Spatio-temporal water quality mapping from satellite images using geographically and temporally weighted regression. *International Journal of Applied Earth Observation and Geoinformation* **65**, 1–11.
- Kiefer, I., Odermatt, D., Anneville, O., Wüest, A. & Bouffard, D. 2015 Application of remote sensing for the optimization of in-situ sampling for monitoring of phytoplankton abundance in a large lake. *Science of the Total Environment* **527–528**, 493–506.
- Li, Y., Zhang, Y., Shi, K., Zhu, G., Zhou, Y., Zhang, Y. & Guo, Y. 2017 Monitoring spatiotemporal variations in nutrients in a large drinking water reservoir and their relationships with hydrological and meteorological conditions based on Landsat 8 imagery. *Science of the Total Environment* **599–600**, 1705–1717.

- Loveland, T. R. & Irons, J. R. 2016 *Landsat 8: the plans, the reality, and the legacy*. *Remote Sensing of Environment* **185**, 1–6.
- Matthews, M. W. 2011 *A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters*. *International Journal of Remote Sensing* **32** (21), 6855–6899.
- Olmanson, L. G., Brezonik, P. L., Finlay, J. C. & Bauer, M. E. 2016 *Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes*. *Remote Sensing of Environment* **185**, 119–128.
- Padedda, B. M., Sechi, N., Lai, G. G., Mariani, M. A., Pulina, S., Sarria, M., Satta, C. T., Viridis, T., Buscarinu, P. & Lugliè, A. 2017 *Consequences of eutrophication in the management of water resources in Mediterranean reservoirs: a case study of Lake Cedrino (Sardinia, Italy)*. *Global Ecology and Conservation* **12**, 21–35.
- Tebbs, E. J., Remedios, J. J. & Harper, D. M. 2013 *Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline-alkaline, flamingo lake, using Landsat ETM +*. *Remote Sensing of Environment* **135**, 92–106.
- USGS 2016 *Landsat 8 (L8) Data Users Handbook*. Version 2.0 Department of the Interior, US Geological Survey, EROS, Sioux Falls, SD, USA. <https://www.usgs.gov/land-resources/nli/landsat> (accessed 29 April 2018).
- Watanabe, F., Alcântara, E., Rodriguez, T., Rotta, L., Bernardo, N. & Imai, N. 2017 *Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil)*. *Anais da Academia Brasileira de Ciências* **90** (2), 1987–2000.
- WHO (World Health Organization) 2006 *Guidelines for Drinking Water Quality. 1st Addendum to 3rd Edition Vol. 1. Recommendations*. World Health Organization, Geneva, Switzerland.
- Zheng, Z., Ren, J., Li, Y., Huang, C., Liu, G., Du, C. & Lyu, H. 2016 *Remote sensing of diffuse attenuation coefficient patterns from Landsat 8 OLI imagery of turbid inland waters: a case study of Dongting Lake*. *Science of the Total Environment* **573**, 39–54.

First received 1 November 2018; accepted in revised form 14 May 2019. Available online 3 June 2019