

ASSESSMENT OF SATELLITE ALGORITHMS FOR DERIVING CHLOROPHYLL-A FROM TURBID WATERS OF AMAZON FLOODPLAIN LAKES

Rogério Flores Júnior¹, Daniel Andrade Maciel¹, Carolline Tressmann Cairo¹, Felipe Menino Carlos¹, Felipe de Lucia Lobo¹, Lino Augusto Sander de Carvalho², Evlyn Márcia Leão de Moraes Novo¹, Claudio Clemente Faria Barbosa¹

¹Instituto Nacional de Pesquisas Espaciais; Av. dos Astronautas, 1.758 - Jardim da Granja, São José dos Campos - SP, 12227-010; rogerio.floresjr@gmail.com; damaciel_maciel@hotmail.com; carolline.cairo@inpe.br; felipe.carlos@fatec.sp.gov.br; felipellobo@gmail.com; evlyn.novo@inpe.br; claudio.barbosa@inpe.br

²Universidade Federal do Rio de Janeiro, Departamento de Meteorologia, lino.sander@gmail.com

ABSTRACT

*The Amazon floodplain represents one of the most important terrestrial ecosystems being a highly complex and dynamic environment, with a key role in the global carbon cycle. Therefore, the monitoring and management of their aquatic systems is vital to increase the knowledge on the biogeochemistry involving water components. Optically Active Components (OAC's) as chlorophyll-a (chl-a) can be a proxy to environmental parameters such as water trophic status and primary productivity. Standard methods to determine chl-a are based on *in situ* measurements being expensive and time consuming, alternatively, remote sensing can be a viable option through the calibration of chl-a algorithms. Therefore, this work aims the assessment of empirical algorithms for chl-a retrieval in Amazon lakes with turbid waters using Remote Sensing reflectance (R_{rs}) from *in situ* data gathered in four campaigns between 2015 and 2017. *In situ* R_{rs} was then used to simulate Landsat 8/OLI and Sentinel 2/MSI images which were calibrated and validated by Monte Carlo simulation. The best algorithms were validated using images acquired almost concurrently to *in situ* data acquisition for both sensors. Preliminary results pointed out the ability to estimate chl-a with errors smaller than 30% for MAPE for simulated data.*

Key words — Water quality, Chlorophyll-a, Empirical algorithms, Remote sensing, Landsat, Sentinel.

1. INTRODUCTION

The Amazon basin has an area of near 6.5 million km² and it is considered the biggest watershed in the world [1]. About 17% of the basin is composed by a wide floodplain with rich biodiversity and providing habitat and ecosystem services for flora and fauna, being essential for the biogeochemical cycles of carbon and nutrients [2].

However, these resources have been threatened by a number of anthropic factors with large areas of primary forest being degraded by habitat fragmentation, edge effects, selective cutting, fires, illegal gold mining, and other activities [3]. The progressive increase of these

anthropogenic factors can affect the quality of water resources, leading to an imbalance in aquatic system contributing for the occurrence of algae blooms what potentially may have impacts on public health [4]. Therefore, the monitoring of this resource is necessary to understand the effects of these changes, helping in decision-making to present and future.

The monitoring of the aquatic ecosystems is traditionally based on *in situ* data acquisition and laboratory analysis which are time consuming and costly. Moreover, it is constrained by the access to the study site and to sampling designs which do not cope with the spatial and temporal variability inherent to aquatic environments limiting large scale studies.

Remote sensing can be a tool for estimating chl-a and as an alternative to traditional methods, reducing costs with its high spatial and temporal resolution. Among water quality parameters apt to be assessed by remote sensing, chl-a concentration (chl-a) is important which can be used for phytoplankton biomass determination and primary productivity studies. In addition, it can be used as a proxy of the trophic state of the water body, being the main pigment found in all species of phytoplankton. The derivation of chl-a from remote sensing can be performed through empirical and semi-analytical algorithms. Empirical algorithms have the benefit of being simple and easy to implement, based on band ratios without needing knowledge of AOC's inherent properties. The optically active components interaction with the electromagnetic radiation (REM), can be used as a proxy for the determination of chl-a.

There are several algorithms available in the literature for chl-a estimation [5]. However, most algorithms were developed for oceanic waters mainly influenced by the chl-a concentration. Floodplain lakes of the Amazonian are highly complex being influenced by "flood pulses" which modifies water composition in space and time, altering the proportion of dissolved and suspended sediments [6]. This dynamic has great influence on the accuracy of the algorithms for estimating chl-a since they have been developed for waters with low turbidity rates. Therefore, this work aims to assess several empirical algorithms [7–13] for turbid waters of the Amazon floodplain.

2. MATERIAL AND METHODS

2.1. Study area

The study area is located in the lower Amazon region, northwest of the city of Santarém and at about 900 km from Amazon River mouth. The Curuai Lake (CL) was selected from more than 30 lakes present in the floodplain area because of its size and data availability. CL is a highly complex environment with a diverse land use and land cover on its banks, and an increasing occupation for agriculture and livestock on the former flooded forest areas. With a dynamic flooding pulse, the surface of the open water ranges from around 600 km² in the dry season up to 3500 km² at the flood peak [14]. The hydrological regime can be defined in four stages, the flooding in January and February during which increasing Amazon water input to the floodplain up to the overbank flow peaking between April and June as a function of the climatic controls on the precipitation over the Amazon basin. The water starts to recede from August to October up to the river reaches its minimal stages from October to December when the floodplain lakes become very shallow with depths below to one meter [1].



Figure 1: Location of study lakes and sampling points.

2.2. Limnological data

Limnological data were acquired in four field missions, June 2015 during the flood season, March, 2016, during the rising season, July 2016 during the flood season, and August 2017 during the receding season. The chlorophyll-a concentration measurements were determined in duplicate according to the methodology described by [15] using Whatman GF/C (1,2 μm) glass fiber filter. The Total Suspended Solids (TSS) were determined based on [16] also in duplicate, and was separated in its organic (TSO) and inorganic fractions (TSI). Both measurements (chl-a and TSS) were calculated using simple mean of the duplicates.

2.3. Radiometric data

The radiometric data were collected using the radiometer TriOS Ramses with three sensors to measure Water- leaving radiance (L_t), sky diffuse radiance (L_{sky}) and irradiance towards water surface (E_d) in the spectral interval from 320 to 950 nm and resolution of 3,33 nm. The remote sensing reflectance (R_{rs}) was calculated (Equation 1) using the glint correction proposed by [17] with inputs of wind, latitude, longitude, time and date of field sampling points.

$$R_{rs} = \frac{L_t - \rho * L_{sky}}{E_d} \quad (1)$$

A representative spectrum was selected among over 150 measurements for each sampling station, first with a visual outlier removal and then based on the minimum sum value of the difference between median R_{rs} values at each wavelength per true R_{rs} value at each wavelength. Then, the spectral bands from OLI and MSI were simulated using the spectral response function for each sensor [18, 19].

2.4. Satellite Data

Orbital data were acquired for OLI (Landsat 8) and MSI (Sentinel 2A e 2B) sensors. Available images were selected considering cloud cover and temporal interval between field (August 08 to 12) and satellite acquisitions. The selected images for both sensors are from August/2017, being august 10 to OLI and 08 to MSI. All data were acquired from USGS (United States Geological Survey) and Copernicus (ESA European Space Agency) databases. The images were downloaded as level-1 product already with radiometric calibration and orthorectified but lacking atmospheric correction. Thus 6S (Second Simulation of The Satellite Signal in The Solar Spectrum [20]) radiative transfer code was applied using input parameters of Water Vapor, Ozone, AOT from MODIS data. An additional glint correction was applied by subtracting SWIR band R_{rs} values from visible and NIR (VNIR) bands considering that the signal should be close to zero at SWIR spectral region [21].

2.5. Empirical Algorithms

Empirical algorithms from literature for chlorophyll-a retrieval were assessed for turbid waters. The models were applied to simulated OLI, MSI bands using *in situ* spectra. A total of 11 algorithm with variations of 6 pre-selected algorithms (Table 1).

Table 1: Selected semi-empirical Algorithms.

Equations	Authors
$Chl - a = 418,88 \frac{R_{rs}(720)}{(R_{rs}(684) - R_{rs}(700))} + 19,275$	[7,8]
$Slope = \frac{R_{rs}(\lambda_2) - R_{rs}(\lambda_1)}{(\lambda_2) - (\lambda_1)}$	[9]
$MCI = RL_w(\lambda_2) - \left[(RL_w(\lambda_1) + \left(RL_w(\lambda_3) - RL_w(\lambda_1) * \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_1} \right)) \right]$	[10]
$CHL = \frac{R_{rs}(708) - (R_{rs}(665) + R_{rs}(740))}{2}$	[11]
$3B = R_{rs}(665)^{-1} - (R_{rs}(708)^{-1} * R_{rs}(753))$	[12]
$NDCI = \frac{R_{rs}(red) - R_{rs}(NIR)}{R_{rs}(red) + R_{rs}(NIR)}$	[13]

2.6. Statistical Analysis

To calibrate and validate the algorithms, a Monte Carlo simulation with 10000 interactions was implement for each algorithm and sensor. Two datasets were created, one to calibrate the algorithm with 70% of the R_{rs} samples and another with 30% for validation. Linear and Polynomial fit were tested for each algorithm. Three types of datasets were used in the simulation: i) a set composed for the data acquired in all field campaigns, ii) one consisting of data acquired in a

specific campaign iii) one set consisting of samples acquired in similar hydrograph phase. The statistical methods criteria for selecting the best models were Mean absolute percentage error (MAPE), root-mean-square error (RMSE) and R².

3. RESULTS

3.1. Optical Active Components

The limnological data for the main OAC's are presented in Figure 2 with minimum, mean and maximum for all field campaigns.

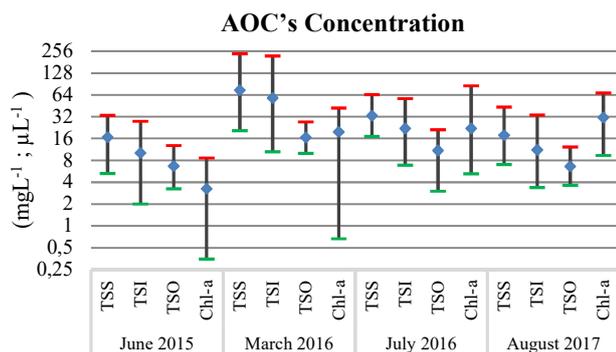


Figure 2: AOC's minimum, maximum and mean for each campaign.

3.2. Field data calibration and validation

From the algorithms tested for all data sets, satisfactory results were observed only for August/2017 (A17) campaign and for data acquired July/2016 and August/2017 (J16A17) campaigns (MAPE <50%) (Table 3). The algorithms chl-a [7], MCI [10] and 3B [12] did not reliable calibration results for both campaigns.

The results of the A17 campaign show MAPE values ranging from 25.57% up to 40.09%, with a minimum RMSE of 2.05 µgL⁻¹ and maximum of 16.50 µgL⁻¹. The R² scored between 0.70 to 0.90 for most models. However, for J16A17 campaign, the MAPE values varied from 39.82% to 64.85%. R² values ranged between 0.30 to 0.80 with some algorithms for MSI sensor with R² > 0.75 (Table 2).

Table 2: MC simulation results for field calibration.

Lake Curuai (August 17)				Lake Curuai (July 16 and August 17)			
Algorithms	MAPE (%)	R ²	RMSE (µgL ⁻¹)	Algorithms	MAPE (%)	R ²	RMSE (µgL ⁻¹)
NDCI G/R OLI Lin	25.57	0.79	10.83	NDCI G/R OLI Lin	39.82	0.29	12.49
Slope N/R S2A Lin	26.46	0.92	6.72	NDCI G/R OLI Pol	41.61	0.33	18.04
NDCI N/R S2A Lin	27.27	0.89	2.05	Slope R/G OLI Lin	42.02	0.41	12.05
NDCI N/R S2A Pol	27.47	0.87	8.64	Slope R/G S2A Pol	42.13	0.44	18.26
Slope N/R S2A Pol	27.70	0.92	7.76	Slope R/G S2A Lin	42.16	0.44	18.26
Slope R/G S2A Lin	29.38	0.76	11.45	Slope R/G OLI Pol	42.21	0.41	18.49
Slope R/G S2A Pol	29.38	0.76	2.27	NDCI N/R S2A Lin	46.69	0.81	10.69
Slope R/G OLI Pol	29.71	0.74	12.14	Slope N2/R S2A Lin	48.33	0.39	17.76
Slope R/G OLI Lin	30.03	0.76	12.76	Slope N2/R S2A Pol	48.95	0.39	17.19
NDCI G/R OLI Pol	32.48	0.71	10.40	NDCI N/R S2A Pol	50.04	0.76	17.49
CHL MBR2012 S2A Pol	33.20	0.81	9.81	Slope N/R S2A Lin	51.84	0.87	10.69
Slope N2/R S2A Lin	36.42	0.56	14.93				
Slope N2/R S2A Pol	38.08	0.44	16.50				
CHL MBR2012 S2A Lin	40.09	0.73	11.69				
A3B moses2009 S2A Lin	45.85	0.04	18.00				

3.2.1. Landsat 8 / OLI

For OLI, the best algorithm for both campaigns is NDCI [13] using green and red bands with linear fit (MAPE = 25.57% and 39.82% for A17 and J16A17, respectively). Moreover, for both fits (linear and Polynomial) satisfactory results were found (MAPE of 29% for A17 and 42% for J16A17) to the Slope [9] algorithm with green and red bands. Slope R² values for A17 campaign ranged from 0.74 up to 0.76. However, for J16A17 it had a low performance (<0.45). RMSE values also showed differences between the datasets with a wider range for J16A17 (12.05 to 18.49 µgL⁻¹) than A17 (10.40 to 12 µgL⁻¹).

3.2.2. Sentinel 2 / MSI

For the MSI, the algorithms calibrated with the field data also showed better performance in the A17 campaign. The NDCI Algorithms using the NIR (B5) and Red (B4) bands obtained a good fit for A17 (MAPE < 28%) e R² > 0.87 with low RMSE values (2.05-8.64 µgL⁻¹). The Slope algorithm also obtained good performance for NIR and red bands with MAPE < 28% and R²= 0.92 for this campaign. Using the red and green bands the algorithm got good scores with MAPE = 30% and R² = 0.76. But for NIR2 (740nm) and red bands the adjustments rendered poorer results with MAPE above 33% and R² ~ 0.50 and also with high values of RMSE (14.93-16.50 µgL⁻¹) for A17 campaign. The 3-band [11] algorithm had a MAPE = 33.20% with a correlation of 0.81 and RMSE = 9.81 µgL⁻¹ for polynomial fit at A17 campaign. The linear fit had the poorest result with MAPE=40.09%, R² = 0.73 and RMSE = 11.69 µgL⁻¹.

In the J16A17 campaign only the NDCI and Slope algorithms presented satisfactory results for MSI, but with MAPE values ranging from in 42% up to 52%. However, both fits of NDCI with NIR and red show good R² (0.76 and 0.81).

3.3. Image Validation

Algorithms with MAPE's lower than 45% from MC simulation with field data were applied to the selected images from Sentinel 2/MSI (14 variants) and Landsat8/OLI (7 variants). The best algorithms were determined through validation over field data from CL in August 2017 (n=18).

Table 3: Best algorithms from image (MAPE<100%) selected for J16A17 and A17 campaigns (n=18).

MSI Algorithms	MAPE	R ²	RMSE
	(%)		(µgL ⁻¹)
NDCI NR MSI A17 Lin	42.70	0.26	14.22
NDCI NR MSI A17 Pol	45.71	0.27	14.41
NDCI NR MSI J16A17 Lin	48.49	0.26	15.44
slope N2R MSI A17 Lin	77.71	0.39	17.50
OLI Algorithms	MAPE	R ²	RMSE
	(%)		(µgL ⁻¹)
NDCI GR OLI J16A17 Lin	37.33	0.28	0.80
NDCI GR OLI A17 Pol	40.82	0.18	4.25
NDCI GR OLI J16A17 Pol	42.14	0.31	1.83

Table 3 presents only the algorithms with MAPE values lower than 100% when applied to the images. The NDCI with NIR and red bands for MSI (MAPE = 42.70% and 45.71% for linear and polynomial, respectively), but with $R^2 = 0.27$ and $RMSE = 14.30 \mu\text{gL}^{-1}$. For OLI only the NDCI algorithm using the green and red bands had good results with MAPE values of 40.82% for A17 and 37.33% and 42.14% for J16A17. However, for both sensors the values of R^2 obtained low performances ranging from 0.18 to 0.39. For the RMSE values, the OLI models presented values between 0.80 e $1.83 \mu\text{gL}^{-1}$ while for the MSI the values varied between 14.22 e $17.50 \mu\text{gL}^{-1}$.

4. DISCUSSION

It is possible observe a better performance (low MAPE, highest R^2 and lowest RMSE) for the algorithms using red-edge NIR (B5) band in A17 campaign when validated with field data, due to scattering by phytoplankton's cell structure around 705nm [10]. These results corroborate with the results obtained by [13] in three bays in the United States ($R^2 = 0.90$) with concentrations up to $30 \mu\text{gL}^{-1}$ and also for simulated data ($R^2 = 0.95$) with $60 \mu\text{gL}^{-1}$. Even without the red-edge bands the OLI sensor presented positive results for chl-a estimation algorithms with the field data, using the green and red band ratios.

Even though A17 and J16 campaigns have similar AOC's concentration, they are in different hydrological conditions. Where J16 campaign (High water) had low water levels due to a hydrological year of drought while for the A17 campaign (receding) the water level was for a typical year.

The best algorithm applied to image was NDCI [13] in green and red bands on OLI sensor and NIR/red bands on MSI sensor. But, this algorithm presented a low relation with the data collected in the field. The temporal variability between the field collections, highly interaction between OAC's and REM, the stratification of chl-a in the water column and even differences in sensor characteristics (TriOS and satellites) may lead to modeling inaccuracies.

5. CONCLUSIONS

The results pointed out the difficulty to decorrelate OAC's through empirical algorithms in turbid environments. However, the algorithms had good agreement with chl-a estimative when applied to field data thus confirming its feasibility for use in turbid waters. The temporal variability between field and satellite data may have been incompatible for the detection of the chl-a dynamics in the lake, thus justifying the discrepancies of the field and satellite calibrations.

5.1. Acknowledgments

The authors thank CNPq and CAPES for their scholarship. Also, the project "Environmental Monitoring by Satellite in the Amazon Biome" from the Brazilian development bank (MSA-BNDES 1022114003005).

6. REFERENCES

- [1] Barbosa, C. C. F. *Sensoriamento remoto da dinâmica de circulação da água do sistema planície de Curuaí/Rio Amazonas*. Tese (Doutorado) — INPE, 2005.
- [2] Melack, J. M. et al. Regionalization of methane emissions in the amazon basin with microwave remote sensing. *Global Change Biology*, Wiley Online Library, v. 10, n. 5, p. 530–544, 2004.
- [3] Laurance, W.F.; peres, C. A. *Emerging threats to tropical forests*. [S.l.]: University of Chicago Press, 2006.
- [4] Junk, W. J.; Mello, J. Impactos ecológicos das represas hidrelétricas na bacia amazônica brasileira. *Estudos avançados*, SciELO Brasil, v. 4, n. 8, p. 126–143, 1990.
- [5] Matthews, M. W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing*, Taylor & Francis, v. 32, n. 21, p.6855–6899, 2011.
- [6] Barbosa, C.C.F.; Novo, E. M. L. M.; Melack, J. M.; Gastil-Buhl, M.; and Pereira, Filho, W. "Geospatial analysis of spatiotemporal patterns of pH, total suspended sediment and chlorophyll-a on the Amazon floodplain," *Limnology*, vol. 11, no. 2, pp. 155–166, 2010.
- [7] Gitelson, A. A. et al. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sensing of Environment*, Elsevier, v. 112, n. 9, p. 3582–3593, 2008.
- [8] Gitelson, A. A. et al. Estimation of chlorophyll-a concentration in productive turbid waters using a hyperspectral imager for the coastal ocean—the azov sea case study. *Environmental Research Letters*, IOP Publishing, v. 6, n. 2, p. 024023, 2011.
- [9] Mishra, D. R.; Mishra, S. Plume and bloom: effect of the mississippi river diversion on the water quality of lake pontchartrain. *Geocarto International*, Taylor & Francis, v. 25, n. 7, p. 555–568, 2010.
- [10] Gower, J. et al. Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing*, Taylor & Francis, v. 26, n. 9, p. 2005–2012, 2005.
- [11] 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*, Elsevier, v. 124, p. 637–652, 2012.
- [12] Moses, W.J. et al. Satellite estimation of chlorophyll-a concentration using the red and nir bands of MERIS—the azov sea case study. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 6, n. 4, p.845–49,2009.
- [13] Mishra, S.; Mishra, D. R. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment*, Elsevier, v. 117, p. 394–406, 2012.
- [14] Bourgoin, L. M. et al. Temporal dynamics of water and sediment exchanges between the curuaí floodplain and the amazon river, brazil. *Journal of Hydrology*, Elsevier, v. 335, n. 1-2, p. 140–156, 2007.
- [15] Nush, E. Comparison of different methods for chlorophyll and phaeopigment determination. *Arch. Hydrobiol. Beih.*, v. 14, p. 14–36, 1980.
- [16] Wetzel, R. G.; Likens, G. E. *Limnological analysis*. [S.l.]: Springer Science & Business Media, 2013.
- [17] Mobley, C. D. Polarized reflectance and transmittance properties of windblown sea surfaces. *Applied optics*, Optical Society of America, v. 54, n. 15, p. 4828–4849, 2015.
- [18] Barsi, J. A. et al. The spectral response of the landsat-8 operational land imager. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 6, n. 10, p. 10232–10251, 2014.
- [19] Sentinel-2 Spectral Response Functions(S2-SRF). Available at: <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-si/document-library/-/asset_publisher/Wk0TKajISaR/content/sentinel-2a-spectral-responses>.
- [20] Vermote, E. F. et al. Second simulation of the satellite signal in the solar spectrum, 6s: An overview. *IEEE transactions on geoscience and remote sensing*, IEEE, v. 35, n. 3, p. 675–686, 1997.
- [21] Wang, M.; Shi, W. The nir-swir combined atmospheric correction approach for MODIS ocean color data processing. *Optics Express*, Optical Society of America, v. 15, n. 24, p. 15722–15733, 2007.