ATMOSPHERIC CORRECTION METHODS FOR THE RETRIEVAL OF THE ORANGE *CONTRA*-BAND FROM MULTISPECTRAL DATA

Rafael Grinberg Chasles[1], *Matheus de Souza Pinto Oliveira*[2], *Felipe Nincao Begliomini* [3], *Daniel Andrade Maciel* [4], *Claudio Clemente Faria Barbosa* [5], *Cláudia Maria De Almeida* [6], *Thainara Munhoz Alexandre de Lima* [7]

[1] [2]Instituto Nacional de Pesquisas Espaciais – INPE, Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil,
[1]rafaelchasles@gmail.com, [2]matheus.oliveira@inpe.br, [3]felipe.begliomini@inpe.br, [4]daniel.maciel@inpe.br,
[5]claudio.barbosa@inpe.br, [6]claudia.almeida@inpe.br, [7]thainara.lima@inpe.br

ABSTRACT

The purpose of this work is to evaluate two methods of atmospheric correction (AC) applied to multispectral images from the Operational Land Imager (OLI) in order to retrieve the orange contra-*band (590–635 nm). It is useful to obtain spectral information related to the presence of phycocyanin (PC) in inland waters. This pigment concentration indicates the presence of cyanobacteria in the water and can be detected by remote sensors with bands covering the absorption peak of the PC (620 nm). The AC algorithms tested the Acolite and 6S processors and were applied to multispectral data from the OLI multispectral sensor, onboard Landsat 8. The orange* contra-*bands calculated with the images corrected by those two methods were compared with the one estimated with reflectance data obtained in-situ for validation. It showed that, for this dataset, Acolite has performed better (MAE = 1.38; BIAS = 0.87).*

Key words – *Acolite, 6S,* contra-*band, atmospheric correction, Operational Land Imager.*

1. INTRODUCTION

The Landsat series of satellites are very important for multiple areas, such as observation of biomes, land use and land cover studies, and many others. It has global coverage and a 50-year-long time series. The imaging sensor Operational Land Imager (OLI), which is onboard Landsat 8, features 30 m spatial resolution for the visible light range, near infra-red (NIR) and short-wave infrared (SWIR). These characteristics are welcome in inland water quality monitoring, where many water bodies are too small to be seen in images acquired with coarser spatial resolution.

However, the broad multispectral (MS) bands may limit application of these data for some studies, such as, for instance, monitoring phycocyanin (PC) in inland water bodies. This is because the absorption peak of some optically active pigments is not covered by the range of MS bands. OLI sensor features 11 bands, of which 5 covers the visible light range: Coastal-Blue, 435–451 nm; Blue, 452–512 nm; Green, 533–590 nm; Red, 636–673 nm; and Panchromatic (Pan, 503–676 nm). The PC absorption peak (620 nm) is only covered by Pan, which is a very broad band. Therefore, this band set makes the retrieval of spectral information directly related to PC difficult.

This pigment indicates the presence of cyanobacteria in inland waters, which can produce toxic compounds and contaminate drinking water. Hence, the development of methods for observing cyanobacterial blooms is a major concern in water quality studies.

In order to take advantage of the qualities that belong to the OLI sensor for monitoring PC in eutrophic inland waters, while circumventing its limitations, a method was developed for retrieving the orange *contra*-band from this sensor data [1]. It consists in offsetting the signal from the Pan band against the green and red MS bands, leaving only the range covering the PC absorption peak.

To accomplish this goal, it is necessary to correct the atmospheric effects in data from the OLI L1 bundle. It means to convert the top-of-atmosphere reflectance (TOA) values to ground reflectance. Atmospheric correction (AC) methods aim to separate the signal coming from the atmosphere scattering from the signal coming from the ground reflectance. On this subject, this article will compare the performances of two AC algorithms.

The first is described in [1], and is available through the Acolite software. The second is Second Simulation of the Satellite Signal in the Solar Spectrum (6S), a radiative transfer model (RTM) designed to simulate the solar radiation path through the atmosphere [2].

After this correction, it is possible to extract two orange *contra*-bands and then match-up them with the reflectance dataset acquired in-situ.

2. MATERIAL E METHODS

2.1. Data

The in situ data were made available by the Laboratory of Instrumentation of Aquatic Systems (LABISA). The field campaigns were conducted on four dates when Landsat 8 was passing over the study area: Ibitinga reservoir, São Paulo state (Path 221, Row 075), making the match-up possible. The dates are 02/06/14, 03/26/14, 06/16/14 and 08/12/18. On each one of these days, 4 samples were collected, summing a total of 16 samples. As mentioned, Landsat OLI sensor dataa were acquired from the United States Geological Survey (USGS) Earth Explorer website.

2.2. Study area

Ibitinga reservoir, shown in Fig.1, is located in the city with the same name, in the central area of São Paulo state, Brazil. It is an impoundment of Tietê river, one of the most important in the state, built for the construction of a hydroelectric power plant between 1960 and 1970 decades. It lies at an altitude of 404 meters (SRTM) and its area covers about 93.6 km². The use of the lands surrounding the lake are mainly characterized

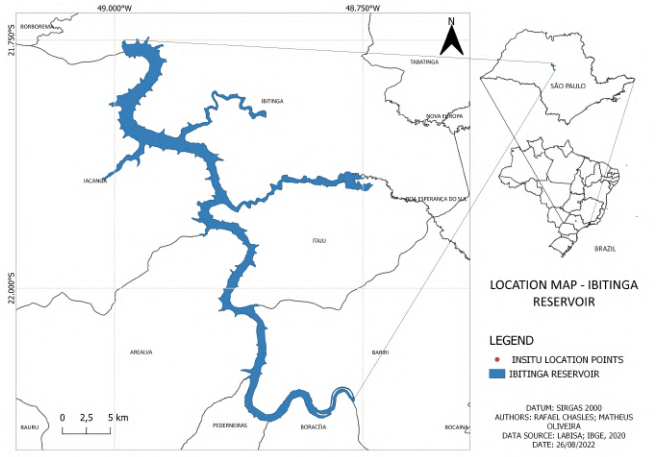

Figure 1: Ibitinga Location Map.

by agriculture and pasture. An eventual dumping of fertilizers in the lake could lead to high concentration of PC in inland water bodies.

2.3. Orange Band retrieval - Castagna's method

The method of retrieving the *orange contra-band* is based on the divisibility of the area formed by the normalized spectral response function (SRF) of the Panchromatic band to extract the spectral information between 590 and 635 nm, which is the wavelength range of interest. According to [1], the area within the SRF, $f(\lambda)$ of a band x can be divided in an sum of the integrals of i spectral sub-regions that constitute this band:

$$L_x = \int_{\lambda \in W_x} L(\lambda) f_x(\lambda) d\lambda = \sum_{i=1}^{n} \int_{\lambda \in W_{x_i}} L(\lambda) f_x(\lambda) d(\lambda) = 1 \tag{1a}$$

$$\int_{\lambda \in W_x} f(\lambda) f_x(\lambda) d(\lambda), \tag{1b}$$

where W_x is the wavelength range of waveband x, and W_{x_i} are the wavelength subsets of each spectral sub-region. According to [1], Equation (1a) is valid only if i sub-regions do not overlap and their collection covers the entire extension of wavelengths in the SRF.

Based on this idea, it is possible to divide the OLI Panchromatic band spectrum in four sub-regions defined by the Full Width at Half Maximum of the Green and Red MS bands (Fig.2).

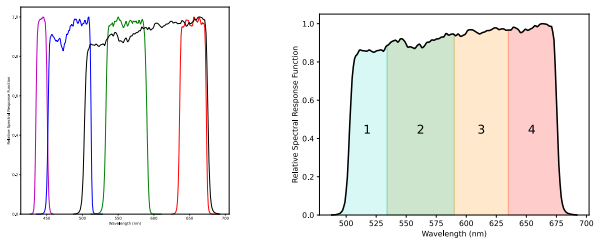

Figure 2: (A) Relative SRFs for Coastal-Blue (magenta), Blue, Green and Red MS OLI bands, plus Pan band (black); and (B) Spectral subregions defined in the Pan band.

We can name the regions as: Turquoise (Region 1; 503–533 nm); Green (Region 2; 534–589 nm); Orange (Region 3; 590–635 nm); and Red (Region 4; 636–673 nm). Considering that the bands Green and Red are the limits of the Regions, their radiances are similar in scale to those of Regions 2 and 4. These are the sub-regions formed between the wavelength ranges of the bands Green and Red themselves. As a consequence, when subtracting the radiance of the Green band (L_{green}) and the Red band (L_{red}) from the Pan band (L_{pan}), considering the respective scale coefficients (S), the result is the radiance from the two other regions ($L_{composite}$), namely Regions 1 and 3. The Eq. (2a) shows this concept.

$$L^{composite} = \frac{(L^{pan} - S_{green} L^{green} - S_{red} L^{red})}{S_{composite}} \tag{2a}$$

$$S_x = \int_{\lambda \in W_x} f_{pan}(\lambda) d\lambda \tag{2b}$$

Nevertheless, Orange is the only of the four sub-regions which includes the PC absorption peak, in the wavelength 620 nm. In this way, it is necessary to isolate the spectral information in this Region. The mathematical relation expressed in Equation 2 makes it possible, but is imperative to replace the scale coefficients (S) by empirically calculated ones in order to calculate the L_{orange} instead of the $L_{composite}$. These empirical S must be calculated by multilinear regression with remote sensing reflectance R_{rs} values obtained from the available bands. This procedure is indicated in Equation (3):

$$R_{rs}^{orange} = \beta_{pan} R_{rs}^{pan} + \beta_{green} R_{rs}^{green} + \beta_{red} R_{rs}^{red} \tag{3}$$

Where β corresponds to the empirical coefficients containing the scale information for the MS bands and Regions. The R_{rs} values are a measure of the water-leaving radiance in each band normalized by the at-surface downwelling solar irradiance and present steradian (sr^{-1}) as their units.

The coefficients used in this work are shown in Equation (4). Those were calculated in [3] for a dataset of radiometric samples from optically distinct water bodies distributed throughout the Brazilian territory and based on the method presented in [1].

$$R_{rs}^{orange} = 2.0020 R_{rs}^{pan} + 0.7188 R_{rs}^{green} + 0.2031 R_{rs}^{red} \tag{4}$$

Fig.3 presents the methodological flowchart of the previously described process.

2.4. Atmospheric correction methods

2.4.1. Acolite

Acolite is a generic image processor developed at the Royal Belgian Institute of Natural Sciences (RBINS) for application in water studies with different sensors. In this work, Landsat 8/OLI images were employed to perform the atmospheric

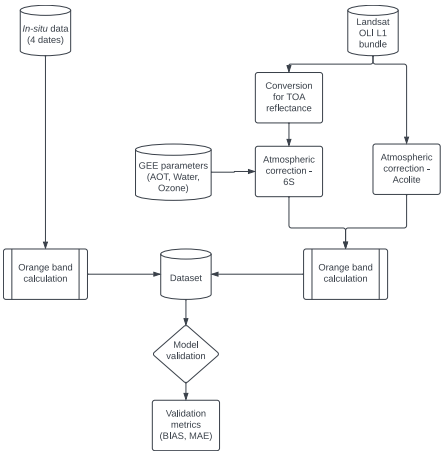

Figure 3: Methodological flowchart.

correction using the default method — the dark spectrum fitting (DSF).

According to [4], simpler methods use the signal of SWIR to find the correction parameters. Given that the water-leaving reflectance in the short-wave infrared wavelengths is zero, these algorithms assume that the signal in SWIR bands is produced solely by aerosol and Rayleigh scattering. Then, after a Rayleigh correction, the signal can be extrapolated to the visible and NIR bands by an exponential function (EXP method). However, the adjacency effects and sun glint will be mistakenly considered as aerosols, causing the parameters to be incorrect. Therefore, unlike this method, the DSF algorithm selects the best-fitting combination of wavebands and aerosol model to retrieve the parameters required for the atmospheric correction.

This approach, described in [4], consists of the construction of a dark spectrum (P_{dark}) from the object with the lowest top-of-atmosphere (TOA) reflectance observed in each band. Also, an spectrum of the atmospheric path reflectance (P_{path}) is computed for the scene-specific sun and viewing geometry, as well as the Continental and Maritime aerosol models and the aerosol optical thickness at 550 nm. Then, the Root Mean Squared Difference (RMSD) between P_{dark} and P_{path} is computed for each band pair containing the fitted band. The combination of aerosol model and band with the lowest RMSD is selected as the best fitting for the atmospheric correction.

2.4.2. 6S

The Second Simulation of the Satellite Signal in the Solar Spectrum (6S) is a Radiative Transfer Model (RTM) which has established itself as one of the standard RTMs used for both remote sensing research and the creation of operational products [5]. 6S can simulate the atmospheric radiative transfer of polarized and non-polarized visible and infrared radiation under different atmospheric conditions. The parameters include the atmospheric conditions, the sensor's and target's altitude, wavelength, and ground reflectance. In addition, an atmospheric correction mode allows the calculation of a ground reflectance, given an at-sensor radiance or reflectance value and a set of atmospheric parameters [2]. In this work, it was used Py6S, which is a Python interface for 6S.

The primary simulation outputs are at-sensor reflectance and radiance, subdivided into their individual components, as well as a number of other calculated atmospheric parameters, such as CO transmittance, scattering angle, water transmittance, atmospheric intrinsic reflectance, Rayleigh transmittance and scattering, CH_4 transmittance, optical depth, spherical albedo, pixel reflectance and others.

The atmospheric inputs to do the processing were acquired in Google Earth Engine (GEE). Aerosol optical thickness at 550nm (AOT) and water vapor column data were extracted from MODIS sensor, calculated in g/cm², and Ozone (cm-atm) data that was obtained from Aura sensor.

The altitude of the target was acquired from Shuttle Radar Topography Mission (SRTM), with 30 m spatial resolution. The lighting and sensor geometry (Solar Azimuth Angle, Solar Zenith Angle, View Azimuth Angle, View Zenith Angle) are described by the Landsat 8 Solar Azimuth Angle (SAA), Solar Zenith Angle (SZA), Sensor Azimuth Angle (VAA), and Sensor Zenith Angle (VZA) bands, respectively. These are the geometry input parameters for 6S processing.

For radiance inputs, it was calculated the spectrum with the smaller value of the difference between the median of all the spectra at each wavelength and the remaining spectra measurements. Eq. (5) was used to represent each station in order to calculate the spectrum with the smallest value for the difference between the median of all the spectra at each wavelenght and the remaining spectra measurements.

$$Dif(R_{rs,i}) = \sum_{\lambda=400}^{900} |Rrs_{i,\lambda} - (Rrs_{Median,\lambda})| \quad (5)$$

$Dif_{Rrs,i}$ is the sum of the difference between median R_{rs} and each in situ-derived R_{rs} at each wavelength for the i sample. $Dif_{i,\lambda}$ is the R_{rs} at sample i and wavelength $\lambda, and Rrs_{Median,\lambda}$ is the median R_{rs} for each station and wavelength (i.e., the median for all spectra measured at each station). The spectrum having the smallest $Dif_{Rrs,i}$ was then selected [1].

3. RESULTS AND DISCUSSION

Performance assessment of ocean color satellite data has generally relied on statistical metrics chosen for their common use and the rationale for selecting certain metrics is infrequently explained. Commonly reported statistics based on mean squared errors, such as the coefficient of determination (r^2), root mean square error, and regression slopes, are most appropriate for Gaussian distributions without outliers and, therefore, are often not ideal for ocean color and inland water algorithm performance assessment, which is often limited by sample availability. In contrast, metrics based on simple deviations, such as bias and mean absolute error (MAE), as well as pair-wise comparisons, often provide more robust and straightforward quantities for evaluating ocean color algorithms with non-Gaussian distributions and outliers [6].

Bias quantifies the average difference between this estimator and expected value, and estimates systematic error. Often based on mean, median error can also be used if a more robust metric is needed. MAE is a measurement of accuracy,

which means it does not amplify outliers and accurately reflects error magnitude. Compared to mean, median absolute estimates are less sensitive to outliers. Similar metrics include mean/median absolute percentage error. [7].

$$BIAS = 10^{\sum_{i=1}|(\log_{10}(Mt) - log_{10}(Ot)/n|} \quad (6)$$

$$MAE = 10^{\sum_{i=1}|(\log_{10}(Mi) - log_{10}(Oi)/n|}, \quad (7)$$

where M represents the modeled value, O is the observation, and n represents sample size, respectively.

The 6S estimator was wrong by an average of 82%, while the Acolite was wrong by 38%. The bias values indicate a systematic deviation of underestimation in both cases of 42% and 13% for 6S and Acolite, respectively. In these work, the r^2 and slope were not calculated because the number of the *in-situ* samples is very low ($n = 16$). Fig.4 shows a scatter plot comparing the R_{rs}^{orange} computed by both processors to those acquired *in-situ* for each of the samples. In this diagram, the blue line highlights where points should be if in-situ data were equal to those computed from remote sensing data.

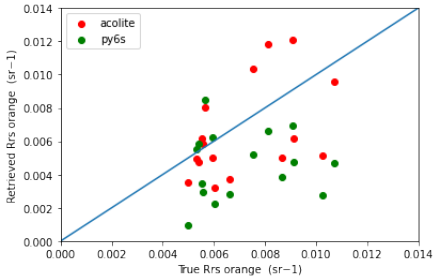

Figure 4: Scatter plot relating R_{rs}^{orange} calculated by Acolite and 6S.

Fig.5 shows an example of the orange band for the Ibitinga reservoir. Darker tones of orange express higher values of R_{rs}^{orange}, which could indicate less concentration of PC in this area. Therefore, higher values in orange band might be influenced by Colored Dissolved Organic Matter (CDOM) presence. This optically active component is commonly found in eutrophic inland water bodies, and has stronger light absorption in the blue region, thus influencing the signal in Pan band.

4. CONCLUSIONS

- Validation metrics were calculated so as to compare the orange *contra*-band estimated with in-situ acquired data with those retrieved with satellite images from OLI L1 bundle processed with Acolite and 6S.
- The results indicated that Acolite performed better, achieving a MAE of 38% and a BIAS of 13%. The data processed with 6S attained a MAE of 82% and a BIAS of 42%.
- Although 6S is a robust atmospheric processor, it is very sensitive to input parameters. Because the MODIS sensor data do not have parameters such as AOT and water vapor for the study area on two dates (7/16/2014 and 8/12/2018), it was necessary to get data from one previous day. This may have contributed to the poorer performance of 6S in this study.
- For a better analysis, it would be interesting to have more in-situ field samples. Also, it would be recommendable to perform the sun glint and adjacency effects correction in the data before processing the atmospheric corrections, since this might result in more accurate outputs.
- In order to determine if lower values in orange band are caused by PC absorption, it would require in-situ collected data particularly focused on PC and CDOM concentrations in these waters.

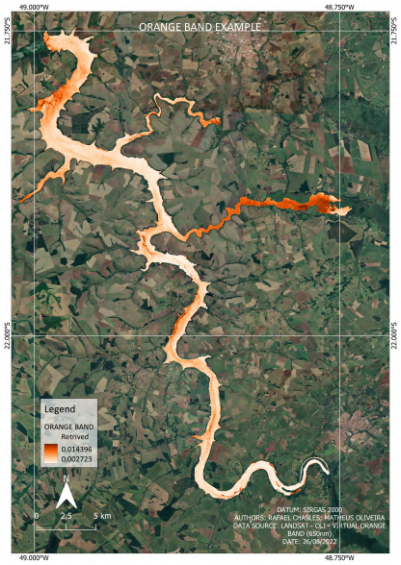

Figure 5: Orange Band used to assess the level of PC in the Ibitinga Reservoir sumperimposed on a Landsat 8-OLI true color composite image. Darker tones of orange could indicate less concentration of PC.

5. REFERENCES

[1] Alexandre Castagna, Stefan Simis, Heidi Dierssen, Quinten Vanhellemont, Koen Sabbe, and Wim Vyverman. Extending landsat 8: Retrieval of an orange contra-band for inland water quality applications. *Remote Sensing*, 12(4):637, 2020.

[2] Robin Timothy Wilson. Py6s: A python interface to the 6s radiative transfer model. *Comput. Geosci.*, 51(2):166–171, 2013.

[3] Felipe Nincao Begliomini. Validating the landsat-8/oli virtual orange band for the brazilian inland waters. 2022.

[4] Quinten Vanhellemont. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the landsat and sentinel-2 archives. *Remote Sensing of Environment*, 225:175–192, 2019.

[5] Eric F Vermote, Didier Tanré, Jean L Deuze, Maurice Herman, and J-J Morcette. Second simulation of the satellite signal in the solar spectrum, 6s: An overview. *IEEE transactions on geoscience and remote sensing*, 35(3):675–686, 1997.

[6] Blake A. Schaeffer Keith A. Loftin Bridget N. Seegers, Richard P. Stumpf and P. Jeremy Werdell. Performance metrics for the assessment of satellite data products: an ocean color case study.

[7] Marie-Hélène Forget, Venetia Stuart, Trevor Platt, et al. Remote sensing in fisheries and aquaculture. 2009.