AlgaeMAp: Algae Bloom Monitoring Application for Inland Waters in Latin America

Felipe de Lucia Lobo*
Department of Technological Development (CDTec), Federal University of Pelotas
and
Gustavo Willy Nagel
Remote Sensing Program, National Institute for Space Research (INPE)

July 10, 2023

Abstract

Due to increasing algae bloom occurrence and water degradation on a global scale, there is a demand for water quality monitoring systems based on remote sensing imagery. This paper describes the scientific, theoretical, and methodological background for creating a cloud-computing interface on Google Earth Engine (GEE) which allows end-users to access algae bloom related products with high spatial (30 m) and temporal (~5 day) resolution. The proposed methodology uses Sentinel-2 images corrected for atmospheric and sun-glint effects to generate an image collection of the Normalized Difference Chlorophyll-a Index (NDCI) for the entire time-series. NDCI is used to estimate both Chl-a concentration, based on a non-linear fitting model, and Trophic State Index (TSI), based on a tree-decision model classification into five classes. Once the Chl-a and TSI algorithms had been calibrated and validated they were implemented in GEE as an Earth Engine App, entitled Algae Bloom Monitoring Application (AlgaeMAp). AlgaeMAp is the first online platform built within the GEE platform that offers high spatial resolution of water quality parameters. The App benefits from the huge processing capability of GEE that allows any user with internet access to easily extract detailed spatial (30 m) and long temporal Chl-a and TSI information.

Keywords: Google Earth Engine, Sentinel-2, water quality, chlorophyll-a, Trophic State Index, Earth Engine App

*CNPq, Capes, INPE

1 Introduction

Human activities on a global scale have significantly contributed to the degradation of the water quality of inland aquatic systems by increasing their nutrient levels. Particularly, reservoirs are under high pressure due to the increasing water demand for urban areas, including irrigation, industrial use, and energy production, while still needing to maintain their ecological function [1,2,3]. The quasi-lentic nature of reservoirs leads to a higher phosphorus accumulation, which may trigger phytoplankton production, abundance, and frequency of algae blooms [4]. Algae bloom (AB) is a rapid increase or accumulation in the population of algae, characterized by the blue-green water coloration caused by algae's pigments, that can cause serious consequences to human health and aquatic biogeochemistry due to the production of toxins [5,6]. The increase in aquatic ecosystems' productivity has been monitored by a series of different Trophic State Indexes (TSIs) based on information regarding the nutrient input (generally, phosphorus), water transparency (usually Secchi depth), and chlorophyll concentration (a proxy for phytoplankton biomass) [7]. The advantage of using Chl-a for AB monitoring is that it can be estimated with satellite data and, therefore, used to systematically monitor phytoplankton abundance at a relatively low cost. The interaction between light and photoactive pigments in phytoplankton cells, such as Chlorophyll-a (Chl-a), enables the detection of algae bloom from space-borne imagery [8]. For instance, the Normalized Difference Chlorophyll Index (NDCI) has been used to detect phytoplankton presence by using a ratio of near-infrared (705 nm—maximum phytoplankton reflectance sensitivity) and red bands (665 nm—high Chl-a absorption peak) [9]. In phytoplankton-rich waters around the globe, NDCI has shown to detect a wide Chl-a concentration range when using Sentinel-2 imagery [8,10,11,12]. Watanabe et al. [8] observed that NDCI performed better than other algorithms (2-band, 3-band, and slope) for Chl-a retrieval in the Tietê cascade system, which is the study area of this research. The development of algorithms using remote sensing for Chl-a estimation relies largely on in situ Chl-a measurements. Unfortunately, the lack of in situ data, in Latin American countries for example, limits water quality management due to the scarce information on the current status of surface waters. This is because most of the monitoring programs provide water quality data at a low frequency (monthly or bi-monthly). Alternatively, the spatial coverage and increased frequency of image acquisition make satellite remote sensing a key tool for AB monitoring programs. Recently, new processing capabilities have used remote sensing historical information, such as Sentinel-2, to track seasonal patterns and degradation tendencies [4,13,14]. The Google Earth Engine (GEE), for instance, offers high performance cloud computing resources to process large amounts of geospatial datasets [15,16]. According to Hirt et al. [17], monitoring water resources using GEE is promising since it offers access to a large image database of several satellites (Sentinel and Landsat programs) thus permitting temporal analysis over large areas. Recent studies have used GEE to map Chl-a concentration and its temporal dynamics [18,19,20,21], showing its ability to monitor phytoplankton abundance and TSI using satellite images. To meet the need to incorporate satellite remote sensing into the TSI monitoring system in Latin American (LA) inland waters, several researchers from several countries are developing Algae Bloom (AB) mapping tools under the support of GEE and Earth Observation Data Science (EO) [22]. This research aims to develop a user-friendly application (GEE App) to retrieve algae bloom-related products towards a monitoring system for inland waters in Latin America. More specifically, the objective is to calibrate and validate predictive algorithms for Chlorophyll-a and Trophic State Index classes using both in situ data and Sentinel-2/MSI data available in Google Earth Engine. Once parametrized, these predictive algorithms were implemented in the GEE App to compute the spatial and temporal variation of Chl-a concentration and TSI classes. This paper gives the theoretical and methodological background supporting the tool's development and reports the first results of their application to the monitoring of Tietê River Basin, located in the State of São Paulo, Brazil (Figure 1).

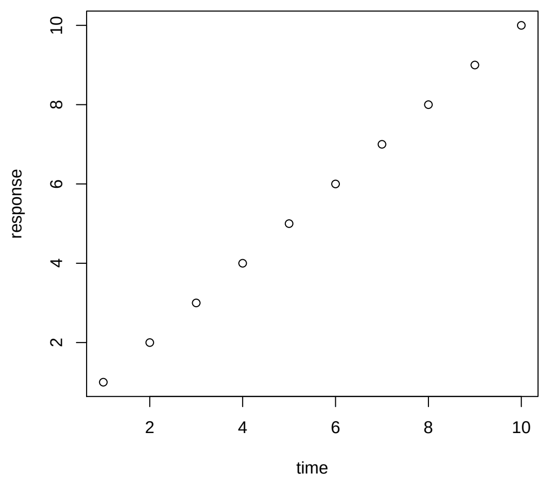

Figure 1: Consistency comparison in fitting surrogate model in the tidal power example.

Table 1: D-optimality values for design X under five different scenarios.

one	two	three	four	five
1.23	3.45	5.00	1.21	3.41
1.23	3.45	5.00	1.21	3.42
1.23	3.45	5.00	1.21	3.43

- Note that figures and tables (such as Figure 1 and Table 1) should appear in the paper, not at the end or in separate files.
- In document front matter, you may set the key `blinded` under a `journal` key to hide

the authors and acknowledgements, producing the required anonymized version.

- Remember that in the anonymized version, you should not identify authors indirectly in the text. That is, don't say "In Smith et. al. (2009) we showed that …". Instead, say "Smith et. al. (2009) showed that …".
- These points are only intended to remind you of some requirements. Please refer to the instructions for authors at http://amstat.tandfonline.com/action/authorSubmission?journalCode=uasa20&page=instructions#.VFkk7fnF_0c
- For more about ASA style, please see https://files.taylorandfrancis.com/asa-style-guide.pdf.
- If you have supplementary material (e.g., software, data, technical proofs), identify them in the section below. In early stages of the submission process, you may be unsure what to include as supplementary material. Don't worry—this is something that can be worked out at later stages.

2 Methods

Don't take any of these section titles seriously. They're just for illustration.

3 Verifications

This section will be just long enough to illustrate what a full page of text looks like, for margins and spacing.

Gelman & Vehtari (2021) offer some guidance about key ideas about statistical ideas. On an unrelated note, spreadsheets are important to use correctly (Broman & Woo 2018). Log-linear models are an attractive way to model categorical data (Bishop et al. 1975).

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. **With this spacing we have 25 lines per page.** The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog. The quick brown fox jumped over the lazy dog.

4 Conclusion

5 Disclosure statement

The authors have the following conflicts of interest to declare (or replace with a statement that no conflicts of interest exist).

6 Data Availability Statement

Deidentified data have been made available at the following URL: XX.

SUPPLEMENTARY MATERIAL

Title: Brief description. (file type)

R-package for MYNEW routine: R-package MYNEW containing code to perform the diagnostic methods described in the article. The package also contains all datasets used as examples in the article. (GNU zipped tar file)

HIV data set: Data set used in the illustration of MYNEW method in Section 3 (.txt file).

7 BibTeX

We encourage you to use BibTeX. If you have, please feel free to use the package natbib with any bibliography style you're comfortable with. The .bst file agsm has been included here for your convenience.

References

Bishop, Y. M. M., Fienberg, S. E. & Holland, P. W. (1975), *Discrete Multivariate Analyses: Theory and Practice*, Boston, MA: MIT Press.

Broman, K. W. & Woo, K. H. (2018), 'Data Organization in Spreadsheets', *The American Statistician* **72**(1), 2–10.
URL: *https://doi.org/10.1080/00031305.2017.1375989*

Gelman, A. & Vehtari, A. (2021), 'What are the Most Important Statistical Ideas of the Past 50 Years?', *Journal of the American Statistical Association* **116**(536), 2087–2097.
URL: *https://doi.org/10.1080/01621459.2021.1938081*