Martech 2024.
Marine Technology Workshop
06, 07th June. Palma, Balearic Islands, Spain

Chlorophyll-a Level Data Update from SeaWiFS Images Using Landsat-7 Data

Gabriel E. Suárez-Fernández[1*], Pedro Arias[1], Andrea Martínez-Movilla[1], Joaquín Martínez-Sánchez[1]

[1] *Applied Geotechnologies Research Group, University of Vigo, ETSE Minas, 36310 Vigo (Spain)*
**gabrieleduardo.suarez@uvigo.gal*

***Abstract* – The current lack of continuous and temporal monitoring of Chlorophyll-a in certain regions is evident. There are some platforms that collected this data in the past, without current update. For this reason, a model based on a Random Forest classification was trained using the limnological classification based on chlorophyll-a concentration from SeaWiFS data and extrapolated to Landsat-7. The model achieved an accuracy of 62.29% with a moderate kappa coefficient. This approach allowed the development of a well-performing model capable of efficiently monitoring and updating chlorophyll-a data in a region lacking in-situ sensors for this variable.**

***Keywords* –** *Remote Sensing, Random Forest, Chlorophyll-a, Landsat, SeaWiFS.*

I. INTRODUCTION

Chlorophylls play a crucial role in photosynthesis. However, when the growth of certain organisms formed by chlorophylls, such as phytoplankton, is excessive, the "bloom" can cause several problems. For this reason, the concentration of Chlorophyll-a (Chl-a) is often used for assessing the level of eutrophication [1].

Nowadays, there are many traditional methods of manual sampling, which can obtain an accurately measure of Chl-a concentration. Nevertheless, these methods are expensive, time-consuming, and not very useful for large spatial extents. Hence, Remote Sensing (RS) technologies are an effective method for monitoring water bodies [2]. Despite this, the RS platforms that have been developed, such as SeaWiFS [3], have not been updated, which adds to the lack of in-situ data.

In this context, we propose a study to update the available RS data and create a model. This model will allow to follow the limnological classification of Chl-a in areas where in-situ data are missing to adjust the information. It will start from SeaWiFS data and using machine learning techniques will be extrapolated to multispectral satellite, such as Landsat, which ensure temporal continuity.

II. METHODOLOGY

The adopted methodology is applied to the study area located as shown in Figure 1. It involves the extrapolation of data using two satellite platforms of different spatial resolution: SeaWiFS and Landsat-7. The data is downloaded through Google Earth Engine platform, with a maximum cloudiness limit set at 25% for multispectral images. Then, a single image is generated for each satellite platform by calculating the average of all the values in each pixel along all the coincident bands for each year and month available. Lastly, these images are downloaded.

After a random sample of 2500 points is used to extract the values of specific bands corresponding to each month and year. NA values are removed from the sample. Then, four spectral indices are calculated according to the following equations:

$$NDWI = (NIR - SWIR1)/(NIR + SWIR1)$$

$$NDWI2 = (GREEN - NIR)\ /(GREEN + NIR)$$

$$MNDWI = (GREEN - SWIR1)\ /(GREEN + SWIR1)$$

$$NDVI = (NIR - RED)/(NIR + RED)$$

Regarding the water bodies categorization, they are classified according to limnological levels based on Chl-a concentration. These levels are categorized as: ultra-oligotrophic (<1 mg/m3), oligotrophic (between 1 and 2.6 mg/m3), mesotrophic (between 2.6 and 7.2 mg/m3), eutrophic and hypertrophic (>7.2 mg/m3) [4].

Lastly, a Random Forest (RF) classification model is trained. For this purpose, the dataset is divided into training and testing in a ratio of 80/20. The training data are balanced by oversampling and undersampling to adjust each class up to 10000 points. The model assessment is performed by analyzing the error rate during training and the accuracy and kappa coefficient values of the confusion matrix during testing, as well as the ROC curve.

III. RESULTS

Concerning the overall results, the RF classifier showed a training error of 14.89%. In testing, it achieved an accuracy of 62.29% with moderate kappa agreement (kappa = 0.4122). The ROC curve results were 0.7698, 0.5936, 0.8091 and 0.6625, respectively for each class. Some current predictions are shown in Figure 1.

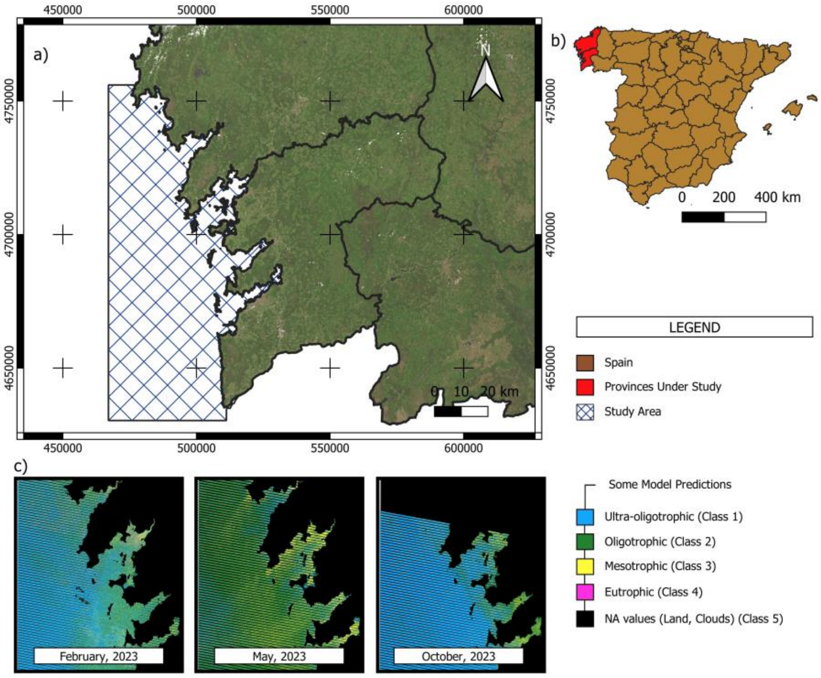

Fig 1. Study area and some predictions with the model. a) Location of study area (EPSG:32629). b) Provinces under study. c) Model prediction for Landsat – 7.

IV. CONCLUSIONS

The developed model showed a good performance in rapidly assessing and updating eutrophication levels using RS based on Chl-a concentration. This methodology enabled comprehensive insights into marine water quality, facilitating extensive temporal monitoring and updates, a crucial requirement in regions where such data is notably scarce. However, the lack of in-situ sensing data limited the capacity to thoroughly assess the success of resolution adjustments among satellite platforms. Therefore, future studies should prioritize applying these techniques in areas with comprehensive coverage of both satellite and in-situ sensing data.

V. ACKNOWLEDGEMENTS

This work has been funded by “Ayudas para la Formación de Profesorado Universitario (FPU)”, Spanish Ministry of Education [FPU21/03038] and by NextGenerationEU through the grant TED2021-132000B-I00 of investment line I3.

REFERENCES

[1] W. Zou, G. Zhu, H. Xu, M. Zhu, Y. Zhang, and B. Qin, “Temporal dependence of chlorophyll a–nutrient relationships in Lake Taihu: Drivers and management implications,” *J Environ Manage*, vol. 306, Mar. 2022, doi: 10.1016/j.jenvman.2022.114476.

[2] G. J. Lakshmi, V. Bodapati, S. S. Baji, and T. S. Panuganti, “Water Quality Monitoring Using Remote-Sensing,” in *Proceedings of International Conference on Computational Intelligence and Sustainable Engineering Solution, CISES 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 624–628. doi: 10.1109/CISES58720.2023.10183395.

[3] C. R. McClain, G. C. Feldman, and S. B. Hooker, “An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series,” *Deep Sea Res 2 Top Stud Oceanogr*, vol. 51, no. 1–3, pp. 5–42, 2004, doi: 10.1016/j.dsr2.2003.11.001.

[4] R. E. Carlson, “Estimating Trophic State,” *Lakeline*, pp. 25–28, 2007.