

Satellite remote sensing of phytoplankton biomarker pigments: a statistical learning approach

Andy Stock & Ajit Subramaniam, Lamont-Doherty Earth Observatory, Columbia University, New York, USA

Background. Phytoplankton are the base of the marine food web and have a major climate-regulating function. However, these ecosystem services depend on phytoplankton community composition, which is sensitive to climate change. One of the most important challenges in satellite monitoring of the world's oceans is thus the development of algorithms that can distinguish different phytoplankton functional types (PFTs) from space. Researchers have proposed various algorithms for this purpose (Mouw et al., 2017). Types of existing algorithms include abundance-based (using only chlorophyll-a concentrations as predictor), spectral (using remote sensing reflectances) and ecological (using environmental predictors like sea surface temperature in addition to spectral data). However, these algorithms have not been fully validated, and past comparisons between different algorithm types have been inconclusive.

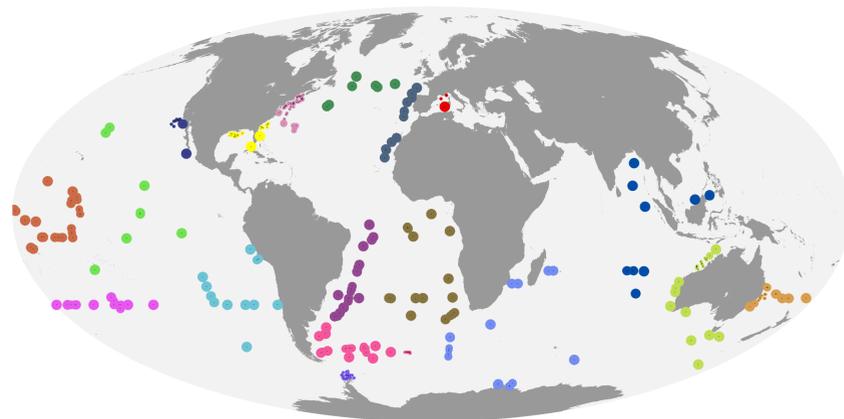


Fig. 1. In-situ HPLC samples. Colors indicate spatial blocks, areas of bubbles are proportional to spatial declustering weights.

Methods. We obtained in-situ HPLC measurements of the concentrations of eight pigments serving as biomarkers for PFTs (Vidussi et al., 2001) from NASA SeaBASS. We matched these observations with MODIS-Aqua spectral and other satellite data, yielding a data set of 442 observations (Figs. 1, 2). We then compared the performance of different abundance-based algorithms (implemented as smoothing splines), spectral algorithms and ecological algorithms (implemented as random forests). Given sparse and spatially clustered in-situ observations, we tuned and tested the different algorithms by means of hierarchical spatial block cross-validation (Roberts et al., 2017), an approach for estimating the extrapolation error of statistical models when observations are not independent. We used the best models to generate global maps of relative pigment concentrations.

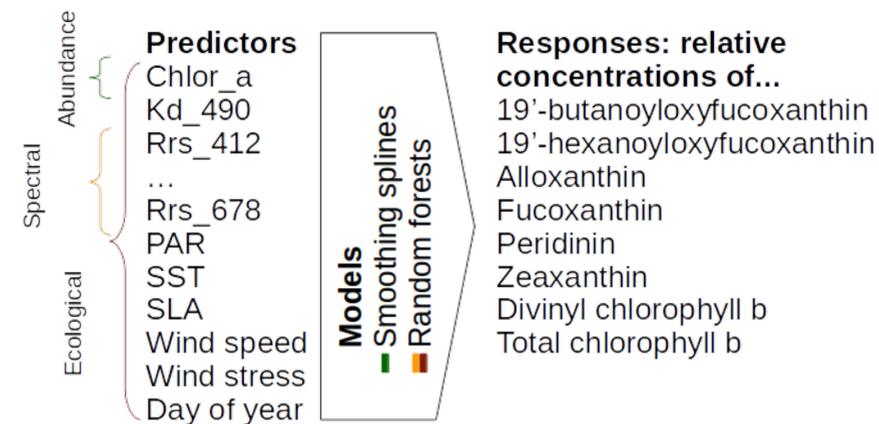


Fig. 2. Overview of predictors, responses and models.

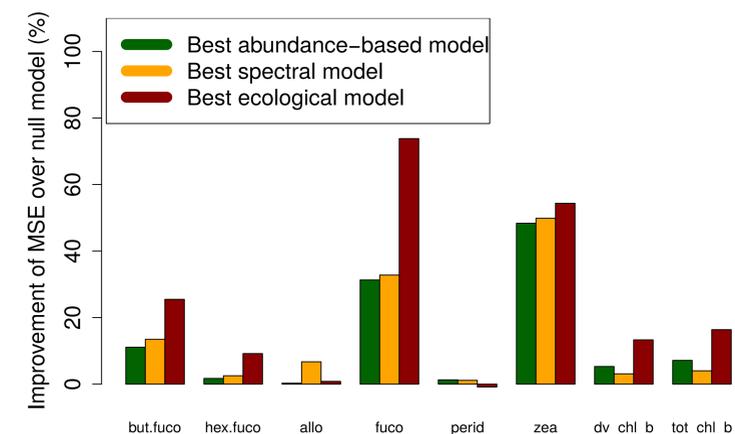
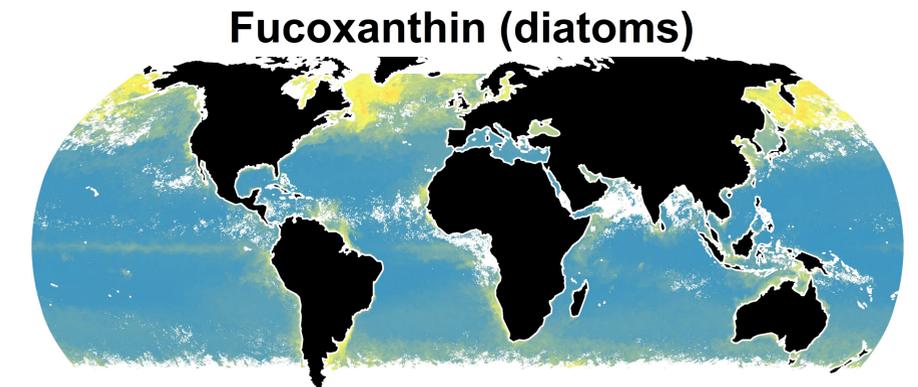


Fig. 3. Reduction of cross-validated error (MSE) for different algorithm types and pigments in comparison to a null model.

Results. Compared to a null model always predicting the mean of the training blocks, the best models identified by our approach reduced the cross-validated mean squared error (MSE) by 74% for fucoxanthin, 59% for zeaxanthin, and 26% for 19'-butanoyloxyfucoxanthin. For these 3 pigments, ecological models worked best, and abundance-based models worst, but the differences were small for zeaxanthin. For all other pigments, improvements over the null model were small.

References

- Mouw CB et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Frontiers in Marine Science*, 4, 41.
- Roberts DR et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913-929.
- Vidussi F et al. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *Journal of Geophysical Research: Oceans*, 106(C9), 19939-19956.



Zeaxanthin (cyanobacteria, prochlorophytes)

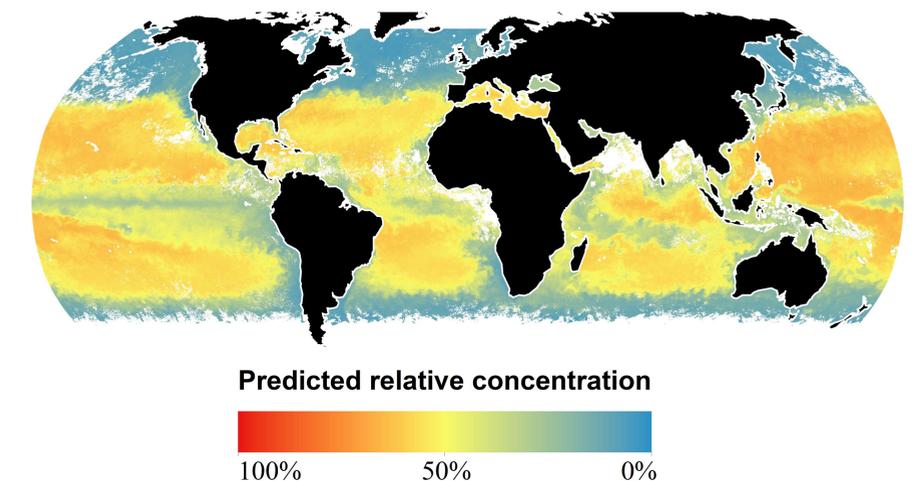


Fig. 4. Mean predicted relative concentrations (proportion of total pigment) of fucoxanthin and zeaxanthin for June 2017.

Summary and conclusions.

1. We identified good models predicting relative concentrations of fucoxanthin and zeaxanthin.
2. For the other biomarker pigments, all tested models performed only little better than the null model (if at all), suggesting that multi-spectral and environmental data like SST are insufficient predictors of the associated phytoplankton communities.

Acknowledgments

This work was supported in part by the Gulf of Mexico Research Initiative's "Ecosystem Impacts of Oil and Gas Inputs to the Gulf" (ECOGIG) program. This is ECOGIG contribution #529. This work was also supported by NASA OBB grant NNX16AAJ08G.