

Supporting Information for Constraining the ocean's biological pump with in situ optical observations and supervised learning. Part 1: particle size distributions

D.J. Clements¹, S. Yang¹, T. Weber², A.M.P. McDonnell³, R.Kiko⁴,

L.Stemmann⁴, D.Bianchi¹

¹Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, Los Angeles, CA, USA.

²Department of Earth and Environmental Sciences, University of Rochester, Rochester, New York, USA

³College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, Fairbanks, Alaska 99775-7220, USA.

⁴Sorbonne Université, CNRS, UMR 7093, Institut de la Mer de Villefranche sur mer, Laboratoire d'Océanographie de Villefranche,

Villefranche-sur-Mer, France

Contents of this file

1. Figures S1 to S4

Corresponding author: D.J. Clements (dclements@atmos.ucla.edu)

Corresponding author: D.Bianchi (dbianchi@atmos.ucla.edu)

Correlation Coefficient (r^2) of the powelaw fits through the observed particle size distribution. Each dot represents an individual profile used. Fits with $R^2 < 0.90$ are removed.

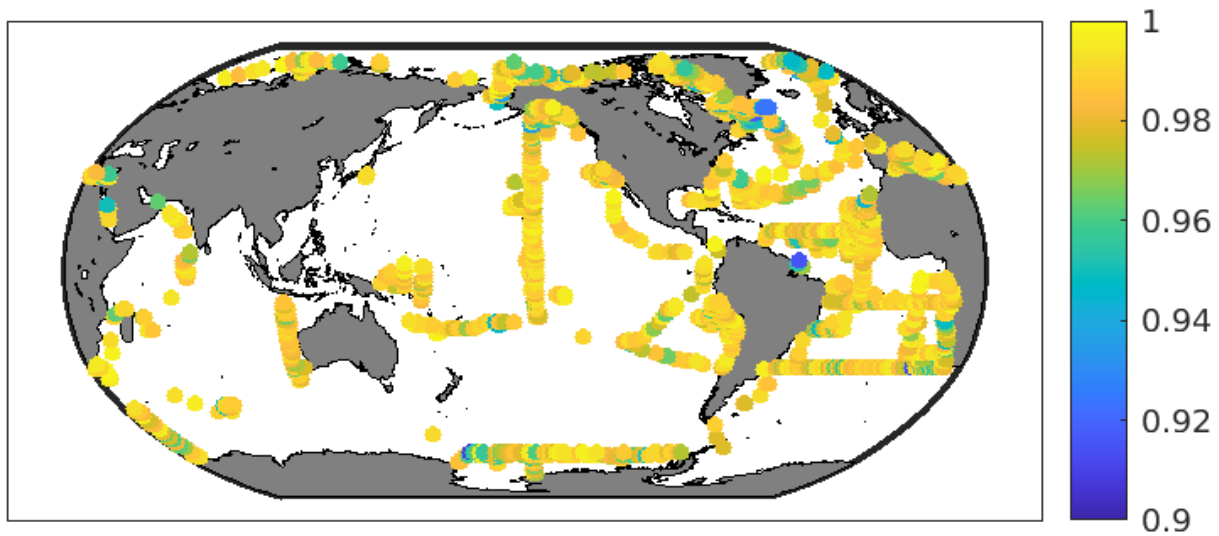


Figure S1. The correlation coefficient (R^2) of the linear least squares fits between the log of the particle counts and the log of their size. The slope and intercept of these fits are the powerlaw parameters to describe the PSD. r -squared less than 0.9 are removed from the dataset, before extrapolating globally.

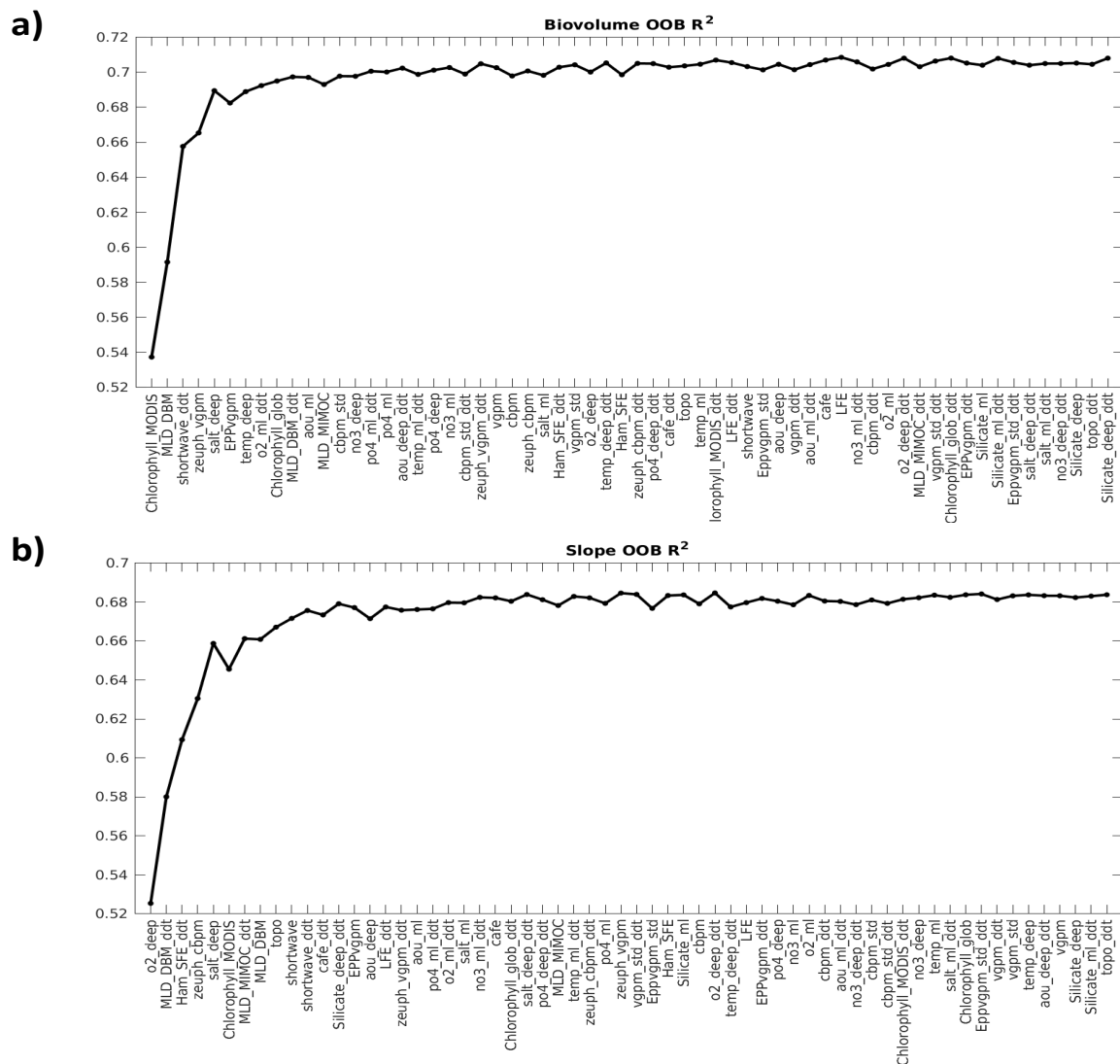


Figure S2. The relative importance of each predictor for each predicted variable from the Euphotic zone. The y-axis shows the r -squared of a random forest, using all predictors at and to the point along the x-axis. It shows the strength of adding additional predictors. A) shows the ranking for the Biovolume, and B) for the Slope.

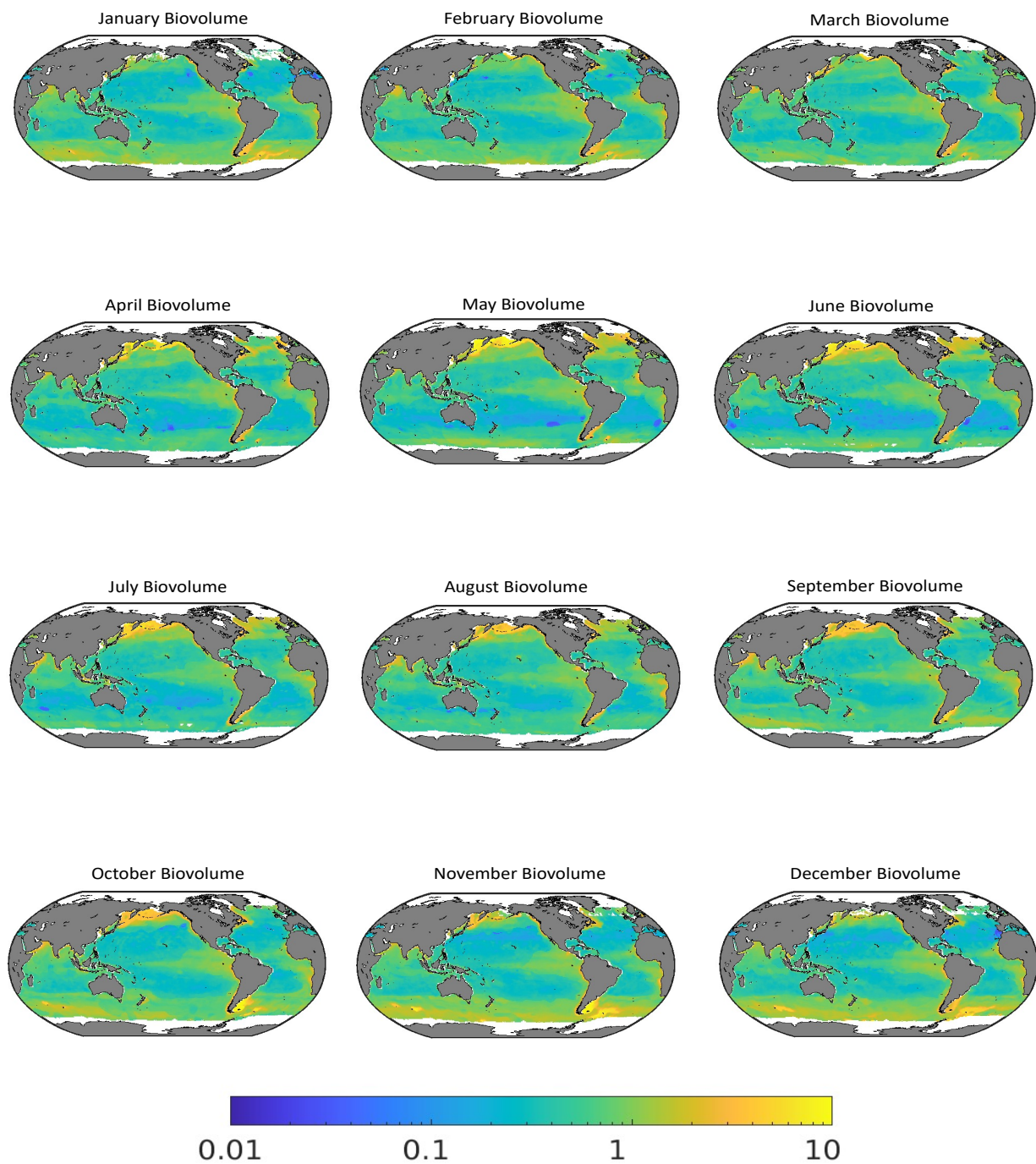


Figure S3. Particulate carbon biovolume at the base of the euphotic zone reconstructed from the random forest calculations, showing monthly climatologies.

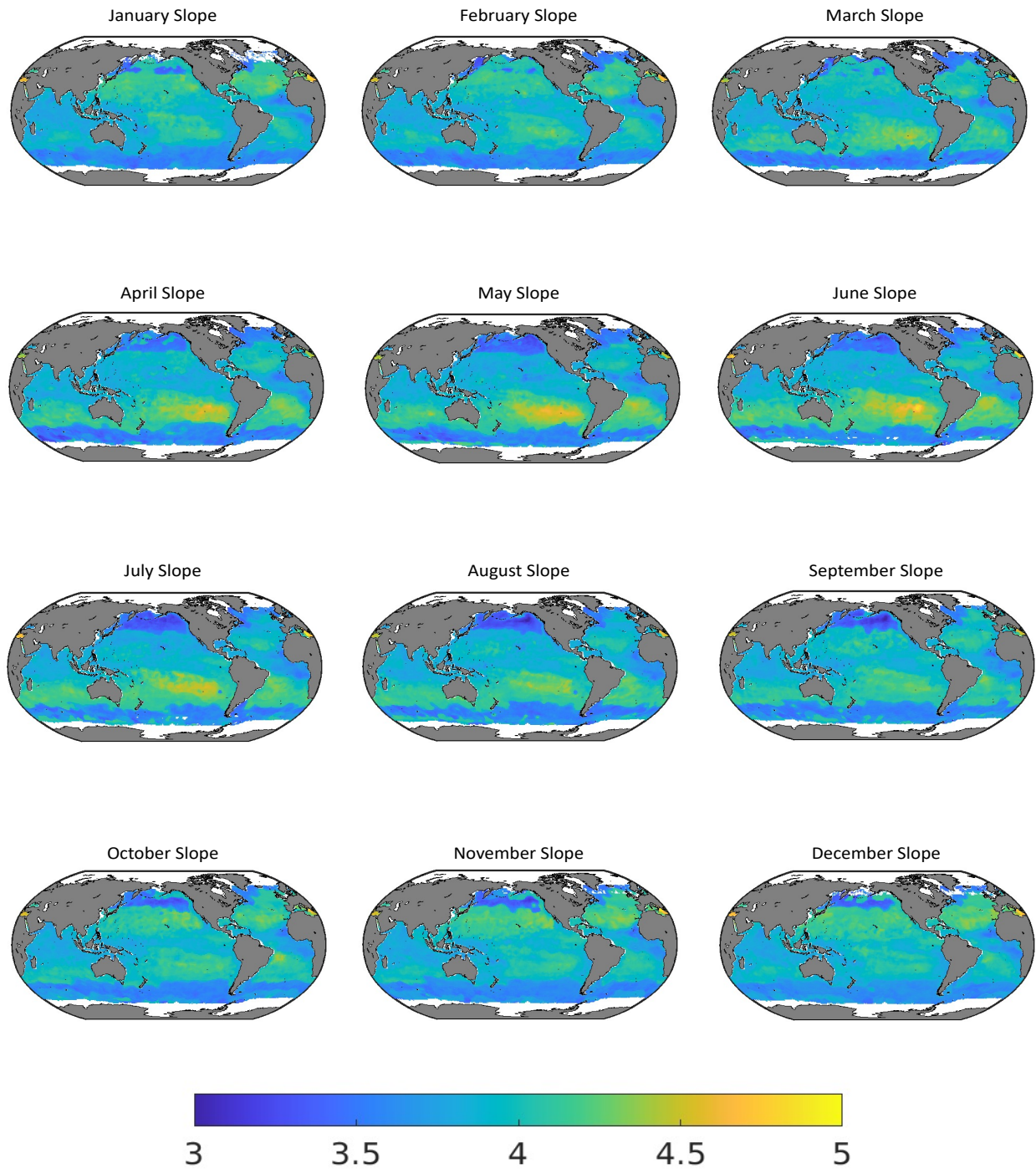


Figure S4. Particulate carbon PSD slope at the base of the euphotic zone reconstructed from the random forest calculations, showing monthly climatologies.