Exploratory Analysis of Phytoplankton's Ecological Niche in Hawaii Ocean Time-series (HOT) Data

Weili Luo

University of North Carolina at Chapel Hill

August 15, 2025

Keywords: Phytoplankton, ecological niche, stratification, Hawaii Ocean Time-series (HOT), biogeochemical cycles, chlorophyll, primary production

Abstract

Phytoplankton are the most basic components of marine food webs and are very important for biogeochemical cycles around the world. This research does an exploratory data analysis of the public Hawaii Ocean Time-series (HOT) dataset to determine the multi-dimensional biological niche of phytoplankton. Several ways of visualization, including a correlation heatmap, vertical depth profiles, and an interactive 3D scatter plot, are used to show how temperature, depth, and macronutrient (nitrate, phosphate, and silicate) are related.

Introduction

There are billions of tiny living things in the ocean, called phytoplankton. They are at the base of the sea food web, with billions of other tiny ocean animals eating them and then bigger animals eating them. While they are alive, they work just like plants on land: they take sunlight, turn carbon dioxide into energy for growth, and release oxygen into the Earth's atmosphere, except that, according to Field et al. (1998), they are responsible for approximately half of the global primary productivity. When the phytoplankton dies, they slowly drift down into the deep ocean, taking the carbon and nutrients away from the surface and participating in the material cycle in the ocean. Additionally, through their growth, death, and decomposition, phytoplankton influence the movement of minor elements (nitrogen, phosphorus, silicon, and minor metals) and therefore their abundance and distribution in the water column.

The growth of phytoplankton is regulated by the combination of light and nutrient availability. The deeper sunlight gets into the water, the darker it becomes because water absorbs and scatters light. There is a layer of water that exists called the euphotic zone. In this layer, the light is intense enough for plants like algae to photosynthesize and make more nutrients than they consume.

This project is based on the publicly available Hawaiian Ocean Times Series data (HOT) to further investigate how phytoplankton in the North Pacific survive and adapt to an oligotrophic environment. Specific questions that will be explored include: First, examining the seawater temperature, salinity, depth, and key nutrients to understand how these environmental conditions correlate with phytoplankton biomass. Second, what the phytoplankton "niche" in this system is - what environmental conditions they need in order to survive. Third, how ratios of different nutrients and surface vs. stratification of the water column affect the composition and population of phytoplankton.

Data and Method

The data investigated ranges from a depth of 0 to 500m in the date range from October 2000 to December 2023 and all comes from public bottle data of several stations from Hawaii Ocean Time-series Data Organization & Graphical System (HOT-DOGS), including ALOHA, HALE-ALOHA, Kaena, Kahe Point, and WHOTS, and saved as NetCDF (.nc) files.

All data processing was done in Python and Jupyter Notebook. The library xarray was imported to process the .nc file to extract core variables: pressure, temperature, salinity, chlorophyll a, nitrate, phosphate, and silicate. These individual data sets were then merged into one pandas DataFrame for better demonstration. The invalid value here is -9.0, which appears in the data column of nutrients. So they were all converted to NaN (Not a Number) and got excluded before doing the following calculations and visualization. For exploratory analysis, the Python libraries below were used: matplotlib, seaborn, numpy, and plotly. And three of the visualization methods are described below: (1) correlation heatmaps to show linear relationships between variables, (2) multi-panel depth profiles to demonstrate vertical water column structure, and (3) interactive 3D scatter plots to capture multidimensional ecological niches of phytoplankton.

Results

Overview of Variable Correlations

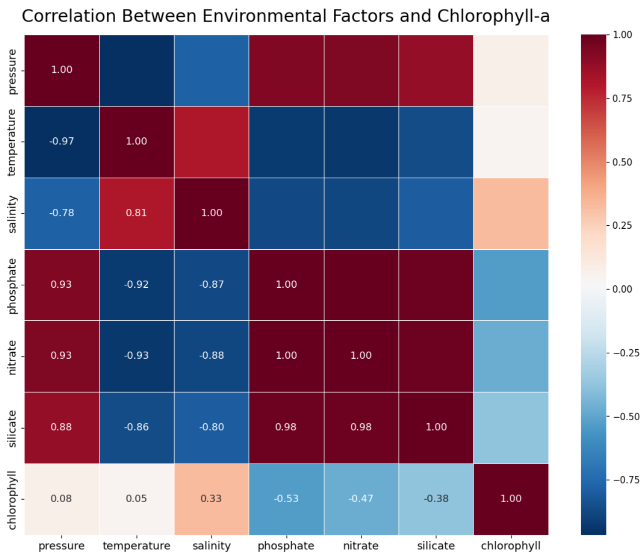

Fig. 1. Correlation heatmap of environmental factors and chlorophyll a.

The primary goal of this correlation heatmap is to illustrate and assess the linear relationships between chlorophyll a and all the other environmental factors in the dataset. The Pearson correlation coefficient between the two variables is calculated in each cell of the chart, ranging from -1 to +1. One noticeable correlation is that chlorophyll and macronutrients are relatively more negatively correlated with each other. This is because phytoplankton grows by taking a lot of nutrients from the water column, which then lowers the concentration of nutrients. On the other hand, chlorophyll is low with high nutrient levels (normally at deeper sea where the pressure is high and the temperature is low) because light and other resources are limited. The correlation between nitrate, phosphate, and silicate also shows deep red colors with coefficients all above +0.95, meaning they are consumed by phytoplankton at about the same time, which aligns with the notion of "Redfield Ratio". Another observation is that the pressure has a substantial negative connection with temperature (-0.87), whereas it has a positive association with all nutrients (around +0.9), meaning the deeper, the colder the water, and the more nutrients that could be found. Finally, the relationship between chlorophyll and temperature (0.05) or pressure (0.08) is far weaker than that of nutrients. That information indicates that we cannot simply conclude "warmer is better" or "shallower is better". And we might need more complicated models and mechanisms to examine the correlation.

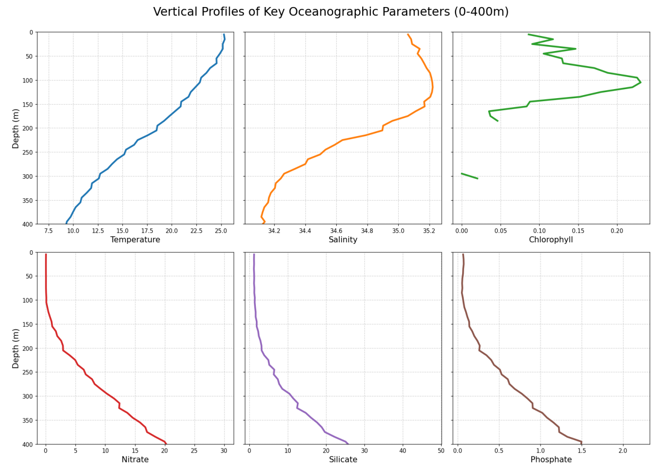

Fig. 2. Vertical profiles of key oceanographic parameters (0-400m).

Vertical Structure of the Water Column

The vertical structure of the North Pacific Subtropical Gyre is shown on the subplots above. The blue curve on the left-top demonstrates that the water temperature drops quickly between around 50 and 150 meters, which is a classic example of a thermocline. The thermocline is a layer in the ocean where the temperature drops quickly and in a non-linear way as you go deeper, and it makes a stable line that separates the warm upper ocean from the cooler, deeper waters (Gordon, 1986). Another clear trait is that these three primary nutrients - nitrate (red), silicate (purple), and phosphate (brown) - all share relatively similar distribution patterns and mark the top of a nutrient salt layer (nutricline). In the top 100 meters of water, these nutrients are almost gone, with levels close to zero, while they suddenly soar with depth, which is in line with the pattern of nutrient delivery from the surface to the surface that Cermeño et al. (2008) outlined. The chlorophyll graph (the green curve) is the most interesting one. The peak concentration of chlorophyll is where it's called Deep Chlorophyll Maximum (DCM), emerging at the depths of about 100 to 125 meters. The depth of this DCM exactly matches the top border of the nutricline and the bottom of the thermocline, which fully supports the old hypothesis of how it formed: it is an "optimal niche" for phytoplankton to balance between getting both the light coming from above and nutrients spreading from below (Latasa et al., 2017).

The Phytoplankton Ecological Niche

By capturing one frame from this interactive 3D scatter plot, we can have a better view of chlorophyll a concentration based on the color and size of each dot. The most noticeable thing is the biggest and brightest yellow dots that are arranged in the optimal niche when all the environmental factors are combined and balanced at the perfect amount. From the depth axis, the largest concentration of chlorophyll a revolves around 50 to 125

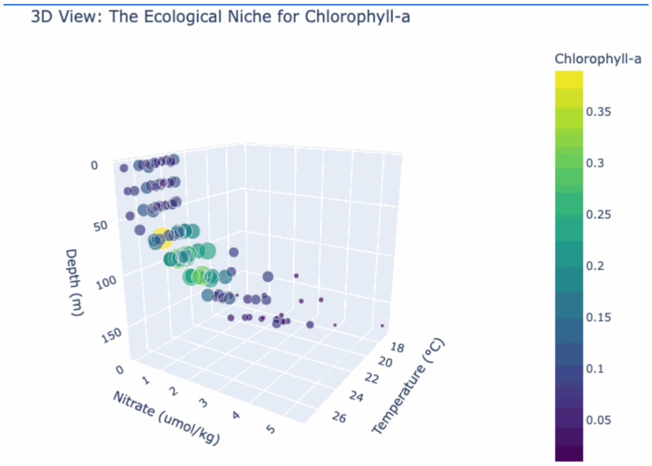

Fig. 3. A 3D scatter plot showing the phytoplankton ecological niche.

meters deep, which is exactly where DCM occurs. Data indicate that at this depth, although sunlight is less prevalent than at the sea surface, it is adequate to facilitate photosynthesis and it's deep enough to reach the top of the nutricline and get "food" from the deep ocean (Latasa et al., 2017). From the nitrate axis, the major concentration of chlorophyll exists around the nitrate level of 0-2µmol/kg, where nitrate concentration starts to rise from zero, which, again, provides additional confirmation of the mechanism of DCM. Lastly, at the temperature axis, the best biological niche is limited to a temperature range of 20 to 25 °C. That indicates the significant role temperature plays in this ecological niche, even when light and nutrition levels are good.

Discussion

As we start with a correlation heatmap, we can see that chlorophyll a is negatively related to the three main nutrients: nitrates, phosphate, and silicate. In this case, it seems that phytoplankton do well when all the nutrients are used up, and that they don't do well when there are a lot of nutrient salts around. We might also notice how pressure and temperature affect each other. This gives phytoplankton more things to think about when they are trying to decide whether to give up some light energy for photosynthesizing in order to get more food supplies. Vertical depth maps provide us with a better view of how these factors change the different layers of the ocean where phytoplankton lives. A clear thermocline can be seen in the temperature curve. It separates the warm, well-lit, yet oligotrophic surface layer from the cold, dark, yet nutrient-rich deep ocean. According to the graph, DCM is between 100 and 125 meters deep, not on the top, where there is a lot of light. This shows that phytoplankton need to live in deeper places to get enough light and food. The last part is the interactive 3D scatter plot, which shows all of this data in more than one dimension. Taking into account temperature, pressure, and nutrients, we can say that the best place for phytoplankton to live is somewhere with a lot of light (about 100 meters deep), a lot of nutrients but almost fully used up, and temperatures between 20 and 25 °C.

This analysis used multidimensional visualization to define a phytoplankton-rich niche based on macronutrients and physical conditions in the HOT dataset, but it lacked micronutrient data, particularly iron. Iron availability controls primary production in the North Pacific Subtropical Gyre (NPSG), a typical marginal zone of high nitrate, low chlorophyll (HNLC) seas (Moore et al., 2002). Therefore, further study lays the groundwork for two distinct future directions. A simple improvement would be to integrate measured or modeled iron concentration data into the analysis and create a four-dimensional niche model with temperature, depth, nitrate, and iron to better understand phytoplankton biomass restrictions. Second, using molecular indicators to directly analyze phytoplankton iron status in niches would also be a forward-thinking research strategy. As this will allow research to move from the macroscopic perspective of "what's in the environment" to the microscopic physiological mechanism of "what the cells are doing."

Conclusion

This data-driven study of the HOT dataset successfully found and showed a unique, multidimensional biological niche for phytoplankton. These conditions - a depth of 50 to 125 meters, temperatures between 20 and 25 °C, and the start of nitrate availability - always led to the greatest amounts of chlorophyll a. Besides, this analysis needs to include more about how micronutrients, like iron, affect the community structure and physiological state of phytoplankton in this marine system.

Acknowledgement

Data were provided by the Hawaii Ocean Time-series (HOT) program, which is supported by the U.S. National Science Foundation under grant OCE-1756517. We thank the captain and crew of the research vessel and the HOT program team for their efforts in collecting and processing these data.

References

P. Cermeño, S. Dutkiewicz, K. G. Harrison, M. J. Follows, and P. G. Falkowski. The role of the nutricline depth in regulating the ocean carbon cycle. *Proceedings of the National Academy of Sciences*, 105(51): 20344–20349, 2008.

C. B. Field, M. J. Behrenfeld, J. T. Randerson, and P. Falkowski. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281(5374):237–240, 1998.

A. L. Gordon. Interocean exchange of thermocline water. *Journal of Geophysical Research: Oceans*, 91(C4): 5037–5046, 1986.

M. Latasa, A. M. Cabello, X. A. G. Morán, and R. T. Barber. Distribution of phytoplankton groups within the deep chlorophyll maximum in the equatorial pacific. *Limnology and Oceanography*, 62(1):177–190, 2017.

J. K. Moore, S. C. Doney, D. M. Glover, and I. Y. Fung. Iron cycling and nutrient-limitation patterns in surface waters of the world ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49 (1-3):463–507, 2002.