Machine-Learning Assessment of Chlorophyll-a Responses to Atmospheric Dust and Environmental Factors Using Remote Sensing Data in the Marmara Sea

This presentation participates in OSPP

Outstanding Student & PhD candidate Presentation contest

Sharing not permitted

Basak Demir[1*], Yusuf Aydin[1], Nazli Olgun[1]

[1] Department of Climate and Marine Sciences, Eurasia Institute of Earth Sciences, Istanbul Technical University, 34469 Maslak, Istanbul, Türkiye.

[*]Corresponding author:demirb19@itu.edu.tr

Abstract

The Marmara Sea, covering approximately 11,350 km² in northwestern Turkey, links the Black Sea and the Aegean Sea via the Bosporus and Dardanelles straits. The Marmara Sea is facing eutrophication and mucilage outbreaks, necessitating the monitoring of key indicators, including chlorophyll-a, which serves as an indicator of phytoplankton abundance.

In this study, spatio-temporal dynamics of chlorophyll-a (Chl-a), Aerosol Optical Depth (AOD), Sea Surface Temperature (SST), Particulate Organic Carbon (POC), Photosynthetically Active Radiation (PAR), and precipitation were investigated on a monthly scale using MODIS-derived products from 2005 to 2020. Time series analysis and machine learning models such as HGB (Histogram Gradient Boosting), Random Forest, and Multiple Linear Regression were performed for exploring temporal patterns, relationships, and modeling Chl-a, respectively. Chl-a showed a moderate negative correlation with SST (r = –0.52) and a strong positive correlation with POC (r = 0.80), while its relationship with AOD was negligible. The observed Chl-a values ranged between 0.6 and 19.50 mg/m³ over the study period, with the highest values observed in April and the lowest values occurring between June and November. Modeling Chl-a based on satellite-derived environmental variables showed that the Random Forest algorithm achieved the highest performance, yielding R = 0.90, R^2 = 0.80, RMSE = 0.13, MAE = 0.11, and MBE = 0.04.

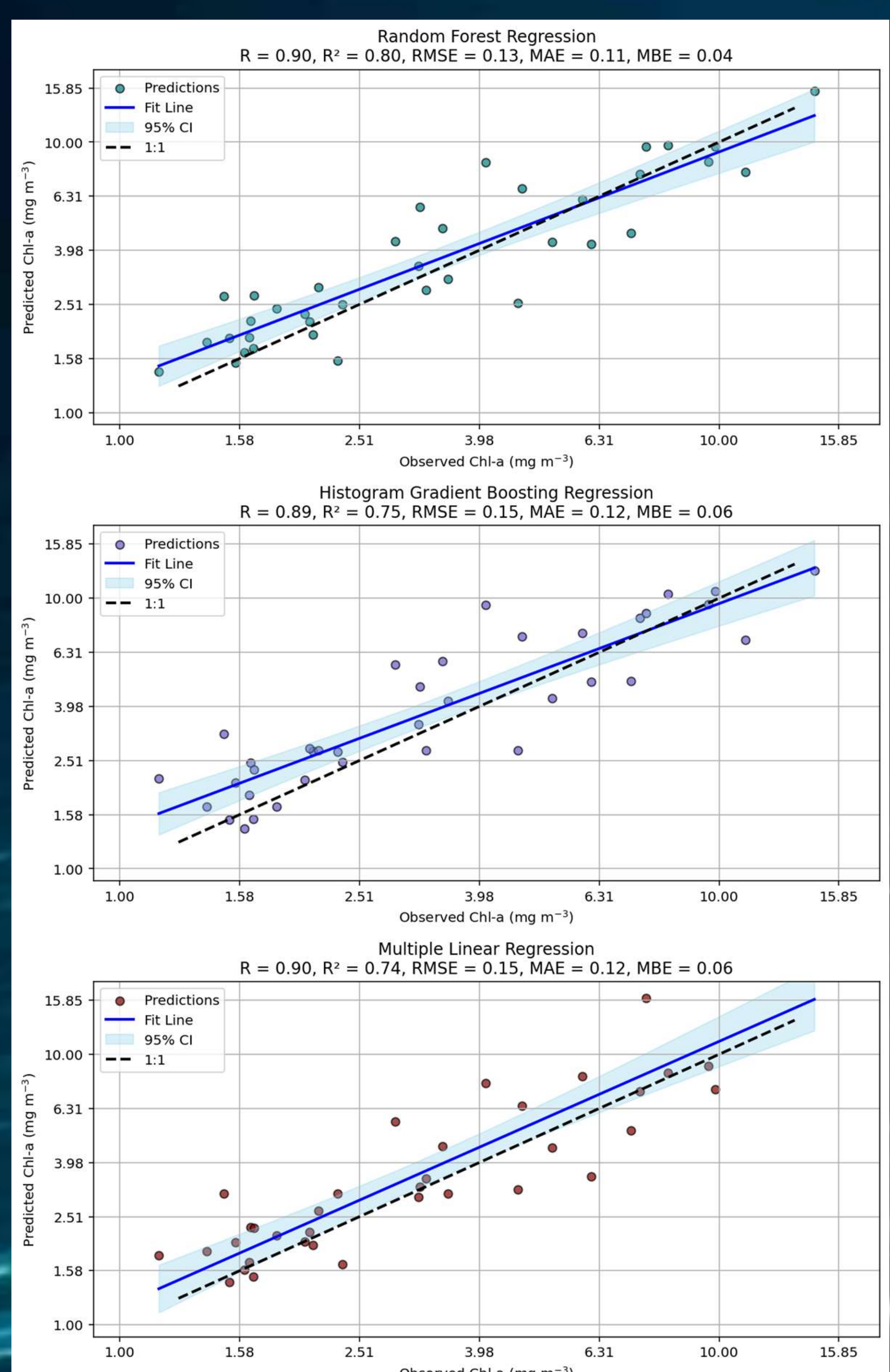

Figure 3: 95% confidence interval plots of observed and predicted chlorophyll-a concentrations generated using machine learning algorithms.

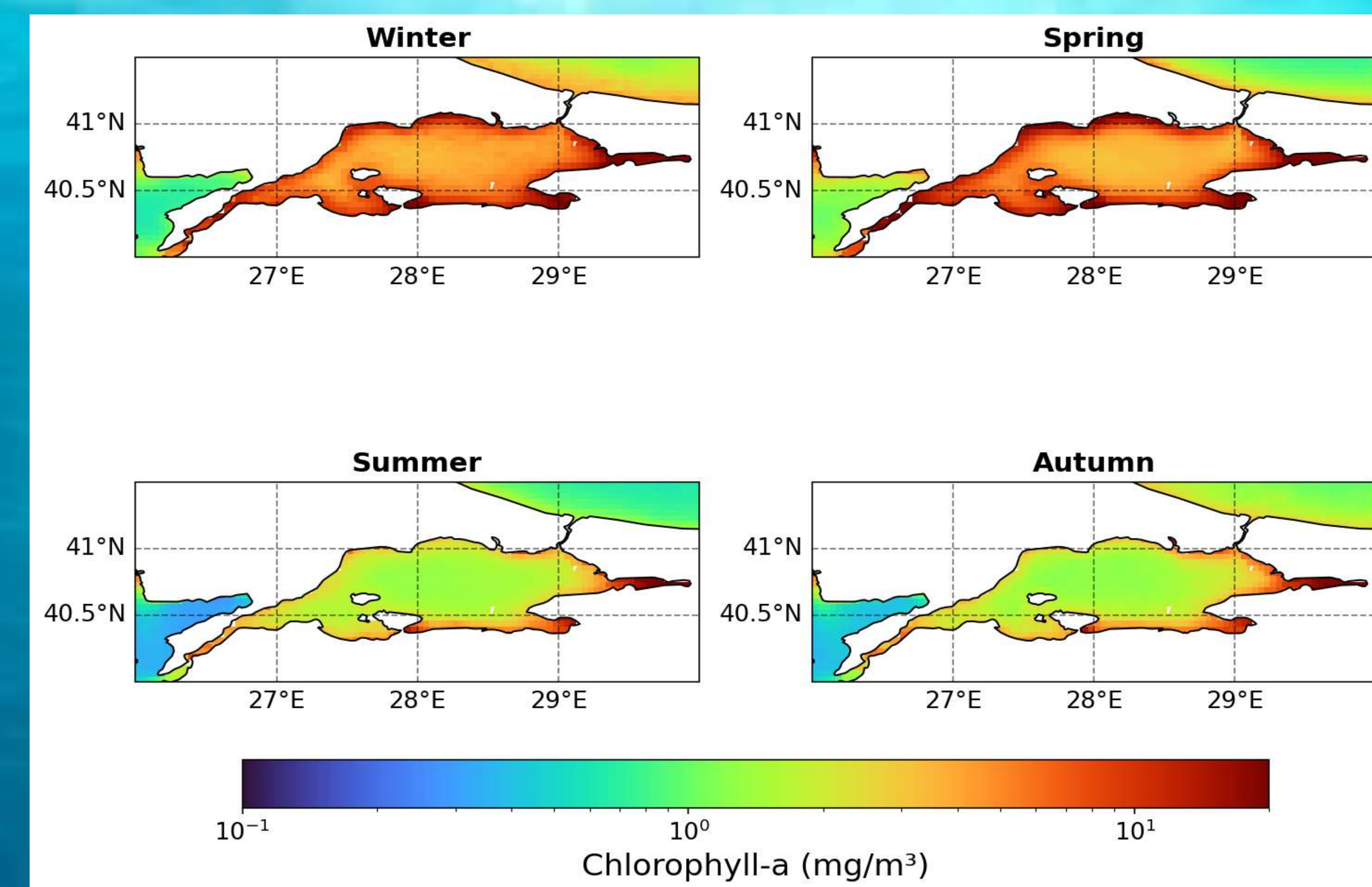

Figure 1 : Seasonal distribution of chlorophyll-a concentration in the Sea of Marmara between 2005 and 2020.

Introduction

Chlorophyll-a is a key indicator of eutrophication and water quality, reflecting phytoplankton response to nutrient enrichment. In the Marmara Sea, concentrations generally increase from the Dardanelles Strait to the semi-enclosed Izmit Gulf due to terrestrial inputs. Its variability is influenced by atmospheric and oceanographic factors such as aerosol optical depth (AOD), particulate organic carbon (POC), sea surface temperature (SST), photosynthetically active radiation (PAR), and precipitation. Aerosols, including mineral dust, can enhance phytoplankton growth by supplying nutrients like iron and phosphorus, but may also reduce light penetration and limit photosynthesis. Ocean systems act as major carbon sinks, while POC, SST, and PAR play important roles in regulating marine productivity.

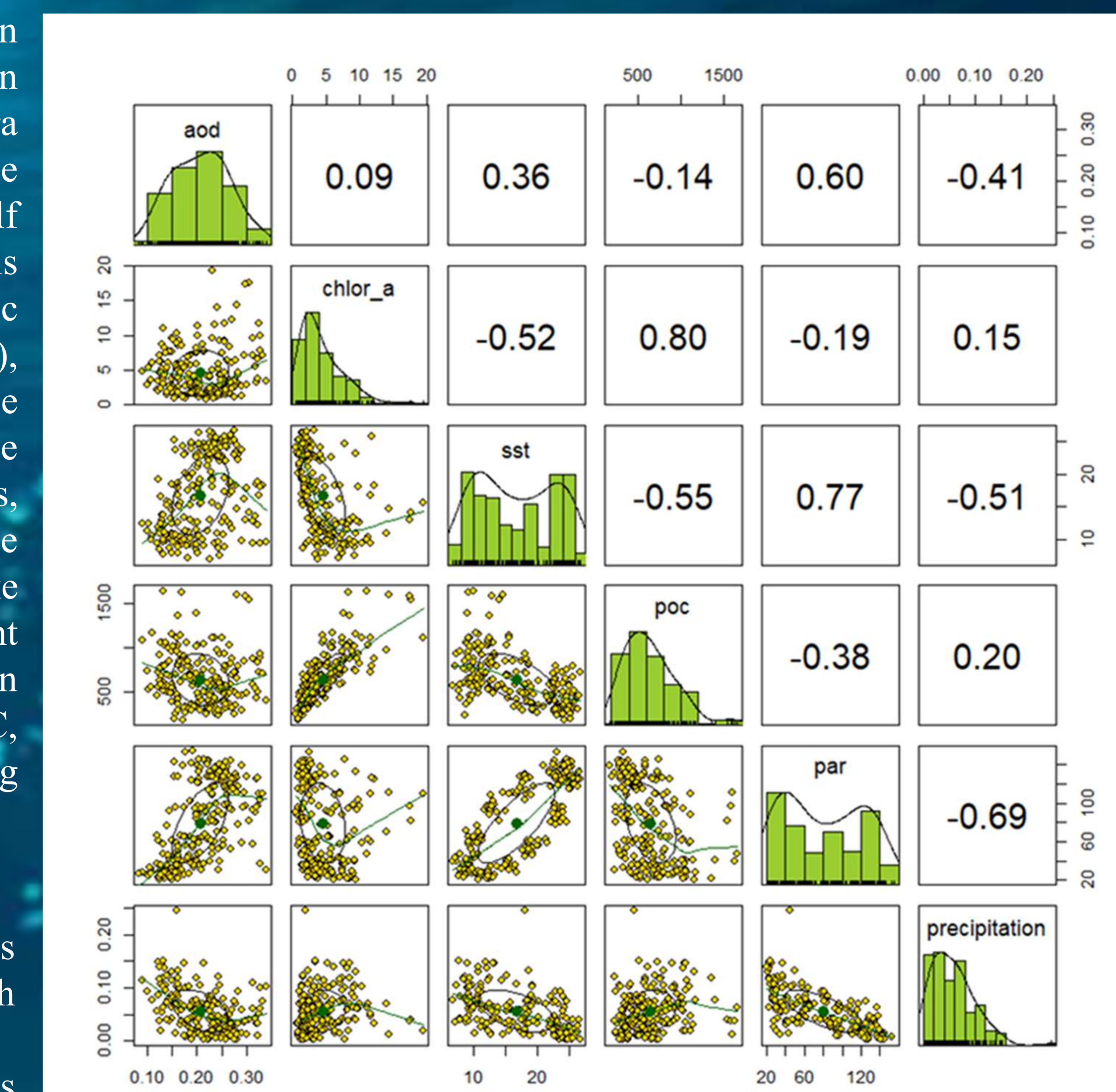

Figure 2: Scatter plot matrix generated from monthly data for the Marmara Sea parameters between 2005 and 2022.

Methodology

In this study, monthly data of Chl-a, POC, PAR, and SST at a spatial resolution of 4 km were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) SMI Level-3 ocean products derived from the MODIS Aqua and Terra satellites, while AOD data were obtained from the MODIS Level-3 atmospheric products. In addition precipitation data were obtained from the ERA5 reanalysis product provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). All datasets used in this study cover the period from 2005 to 2020.

Multiple linear regression, as well as the machine learning algorithms Random Forest regression and Histogram-based Gradient Boosting regression, were employed.

Results

MODIS data were analyzed at daily, monthly, and seasonal scales to map Chl-a, SST, AOD, POC, and PAR. The 2005–2020 time series reveal strong seasonal variability across all variables (Fig 4). It has been observed that, in general, chlorophyll-a (Chl-a) concentrations reach their highest levels during the winter and spring seasons (Fig 1).

AOD shows clear seasonal fluctuations with a slight decreasing trend, while Chl-a and POC exhibit high variability with occasional peaks but overall declining tendencies. SST presents a pronounced seasonal cycle and a long-term increase consistent with warming trends. PAR follows a stable seasonal pattern with a slight increase, whereas precipitation shows irregular variability with a weak upward trend (Fig 5).

According to the scatter plot matrix, the correlation between chlorophyll-a and POC reaches its highest value Fig (2).

Among the models, Random Forest Regression performs best (R = 0.90, R^2 = 0.80) with the lowest errors. Histogram Gradient Boosting captures general trends but shows greater dispersion, while Multiple Linear Regression performs worst due to its inability to model nonlinear relationships (Fig 3).

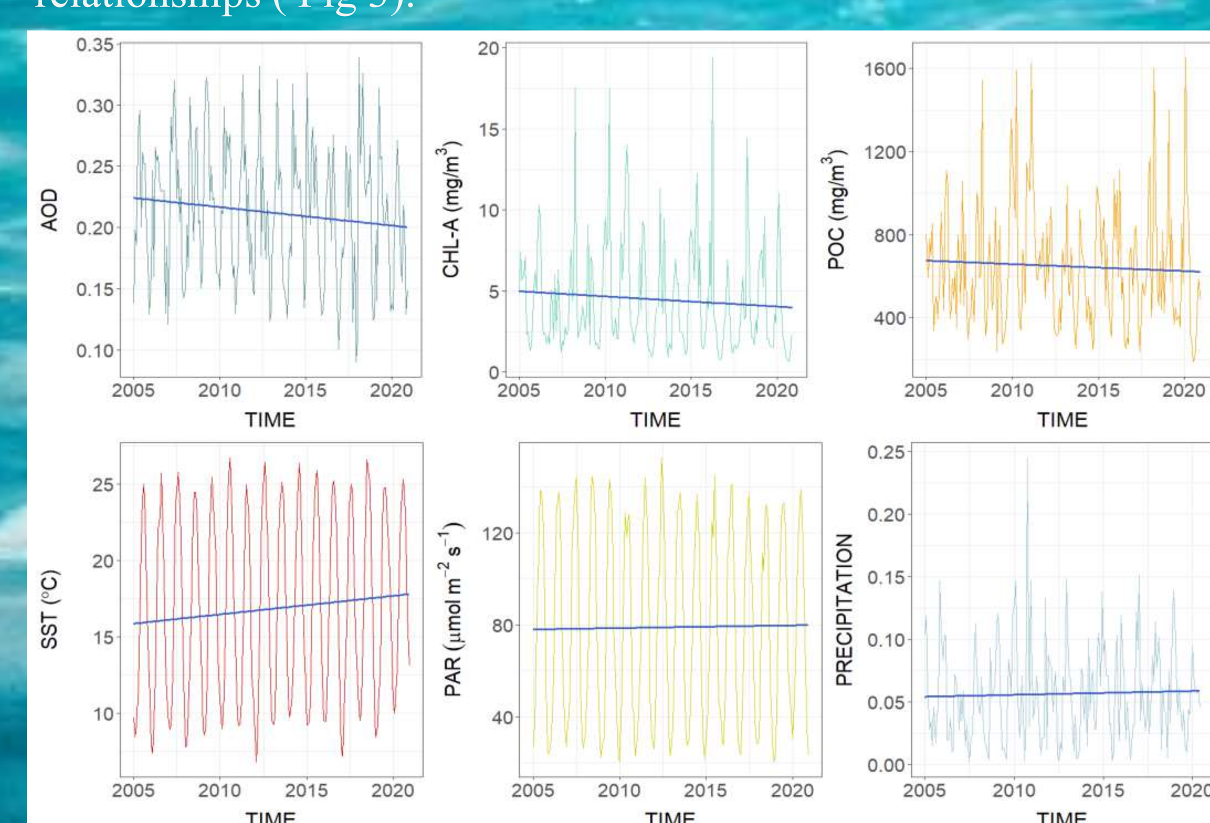

Figure 4: Time series analysis of environmental variables in the Sea of Marmara between 2005 and 2020.

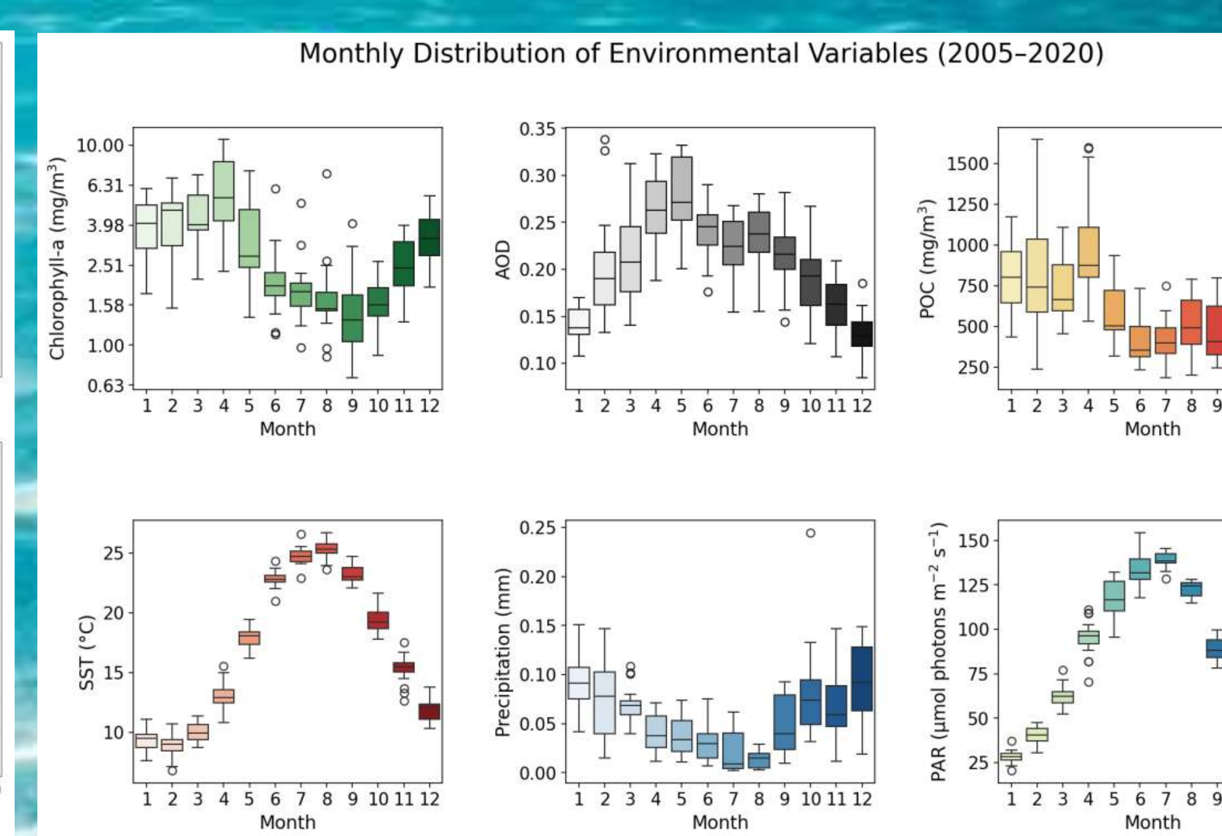

Figure 5: Box plot analysis of environmental variables in the Sea of Marmara between 2005 and 2020.

Conclusion

Analysis of MODIS data from 2005 to 2020 reveals that while Sea Surface Temperature (SST) exhibits a rising trend consistent with global warming, Chlorophyll-a (Chl-a) and Particulate Organic Carbon (POC) levels show a declining trajectory, suggesting a reduction in marine biological productivity. In Figure 5, chlorophyll-a (Chl-a) concentrations are observed to be low during the summer months and increase in the autumn, whereas aerosol optical depth (AOD) values are higher in summer compared to autumn.

Given these complex, nonlinear interactions, the Random Forest model's superior performance (R^2 = 0.80) underscores the necessity of machine learning for accurately predicting ecosystem dynamics, highlighting the critical need to monitor the long-term impact of atmospheric dust shifts on marine primary production.