Summary

This report analyzes satellite-derived chlorophyll-a concentration data from the Global Ocean Colour (Copernicus-GlobColour) product, covering the period from 1997-09-04 to 2024-08-14 in the Antofagasta Region, Chile. The goal is to understand temporal and spatial trends in ocean chlorophyll concentrations. Through time series and spatial analysis, we uncover seasonal patterns, identify chlorophyll concentration peaks, and explore their correlation with coastal regions. These insights are valuable for understanding marine ecosystem health and can inform further research into environmental factors influencing chlorophyll distribution.

Key Findings

- **Seasonal Trends:** The data shows a clear yearly trend with higher chlorophyll concentrations in summer and lower concentrations in winter.
- **Chlorophyll Peaks:** Chlorophyll concentration peaks are observed closer to the coast, though further investigation is needed to determine the causes, such as temperature, shallower waters, or proximity to nutrient sources like rivers.
- **Peak Detection & Optical Flow:** Initial peak detection methods show promise, but require further validation. Optical flow is also a potentially useful tool for analyzing chlorophyll movement, but additional testing is necessary.

Data Summary

The dataset used is the *Global Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (monthly and interpolated)* product derived from satellite observations. It provides satellite images of **chlorophyll-a** concentrations in ocean waters. The data contains daily measurements with a spatial resolution of approximately 4 km, including latitude, longitude, UTC timestamp, and chlorophyll-a concentration in mg/m^3.

The sampled region covers latitudes -27.98 to -21.02 and longitudes -72.98 to -69.02, corresponding to the Antofagasta Region in Chile. It spans from 1997-09-04 to 2024-08-14, with no missing days, totaling 9,842 daily records.

The dataset includes measurements over both land and sea. While sea measurements represent valid chlorophyll observations, values over land show invalid values. Completeness analysis revealed consistent coverage over the sea and a single outlier point on land, likely corresponding to a large inland water body **(Fig. 1)**. By using invalid land cells as terrain markers, the data was spatially segmented into three regions: sea, land, and coast.

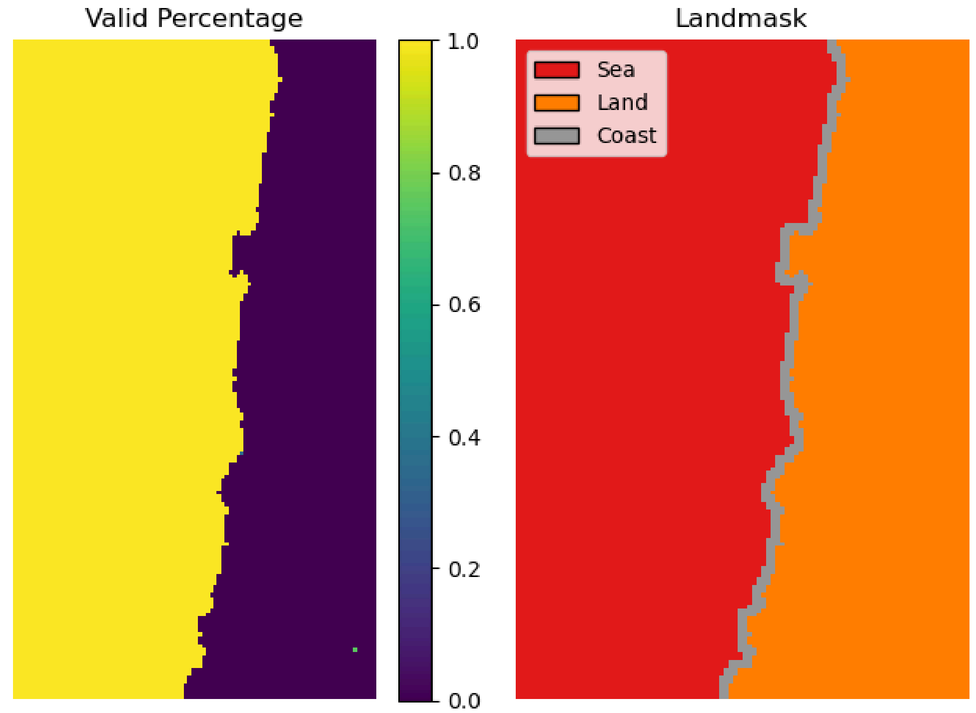

Figure 1: Land classification based on data completeness. An outlier over land likely corresponds to a large inland water body. Cells were classified as land if more than 95% of their observations were NaN, improving robustness over the initial approach of using only fully NaN cells.

Methodology

The methodology involves both temporal and spatial analysis. The temporal analysis focuses on identifying periodic increases in chlorophyll concentrations that follow seasonal trends. The spatial analysis aims to identify peaks in chlorophyll concentrations and provides an initial approach to assess the daily displacement of these concentrations.

Temporal Analysis

To assess temporal trends, a 5-year daily regional mean was compared with the overall dataset's daily mean **(Fig. 2)**. This revealed a seasonal trend, with higher chlorophyll concentrations during summer and the lowest in winter.

To corroborate, the region's mean chlorophyll concentration was correlated with itself using *Pearson correlation*, with a 30-day lag, to assess monthly autocorrelation **(Fig. 3)**. The results show strong correlations within a few months, as expected, with a slight yearly increase likely due to seasonal factors, aligning with previous findings.

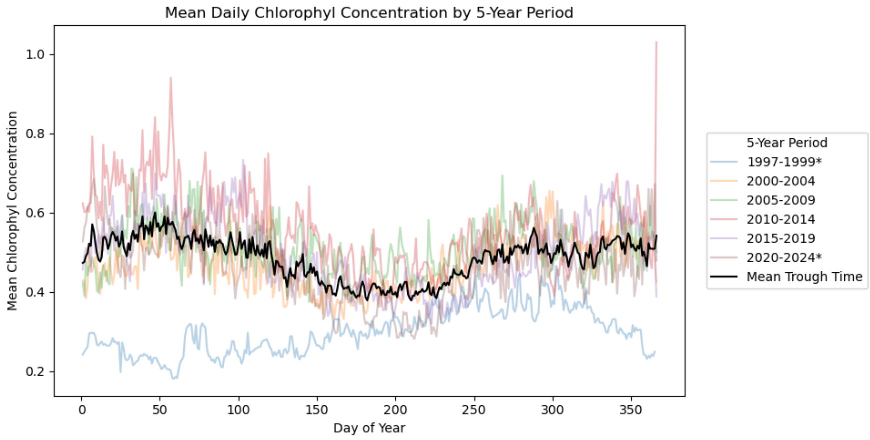

Figure 2: Daily fluctuation of Chlorophyll. Higher chlorophyll concentrations are observed around day 50, corresponding to February (Southern Hemisphere summer), while the lowest concentrations occur around day 200, in July (Southern Hemisphere winter).

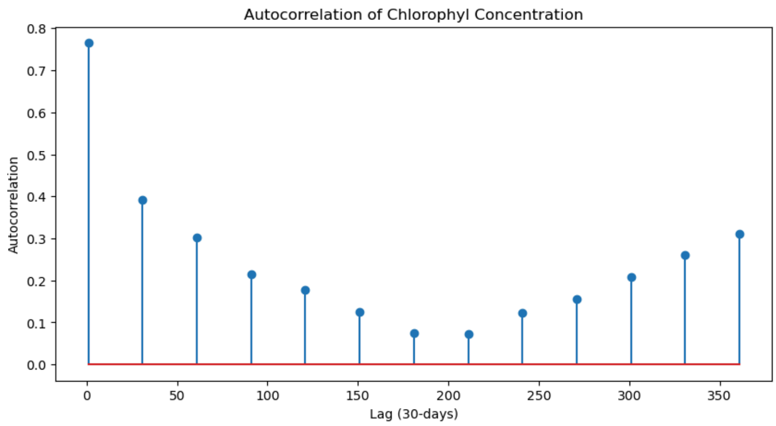

Figure 3: Autocorrelation analysis. It shows strong correlations within a few months, as expected, due to the relationship between chlorophyll concentrations at nearby times. A slight yearly increase in correlation is observed, likely driven by seasonal factors.

Spatial Analysis

Our initial analysis focuses on identifying where maximum chlorophyll concentrations occur. To do this, we calculate the per-cell maximum across time **(Fig. 4)**, revealing that the highest concentrations are found along the coast.

Next, we examine peaks on a single day, selecting 1998-09-05 as a random date **(Fig. 5)**. Two approaches were tested: applying a maximum filter before detecting local peaks and using a Hanning window before detecting local peaks. Several parameter values were manually tested for the latter.

Finally, using optical flow on the consecutive day after the selected date, we explore the movement of chlorophyll concentrations in the sea **(Fig. 6)**.

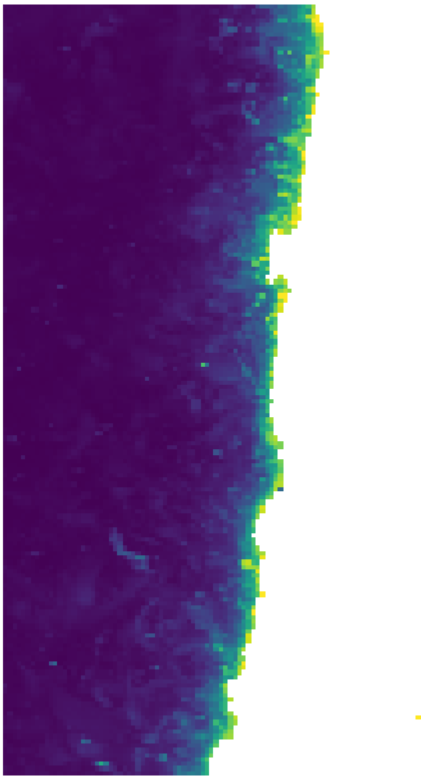

Figure 4: Per-cell global maximum. The naïve approach suggests that chlorophyll concentration peaks occur closer to the coast.

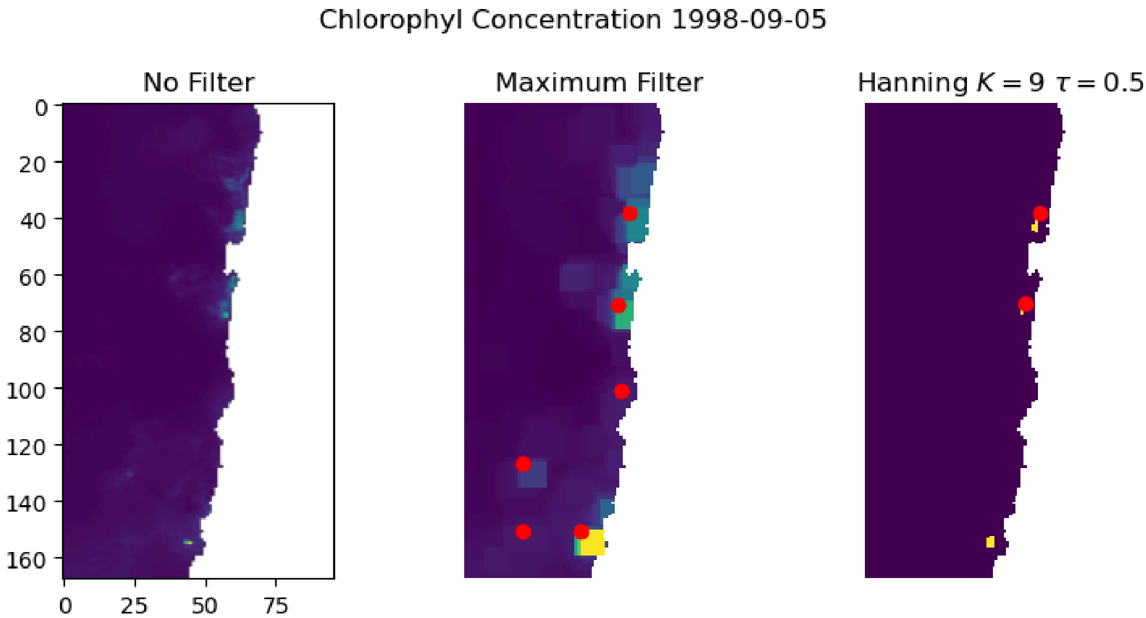

Figure 5: Peak detection. Two filtering approaches were tested to enhance peaks: a maximum filter and a Hanning window with a kernel size of 9 and a threshold of 0.5.

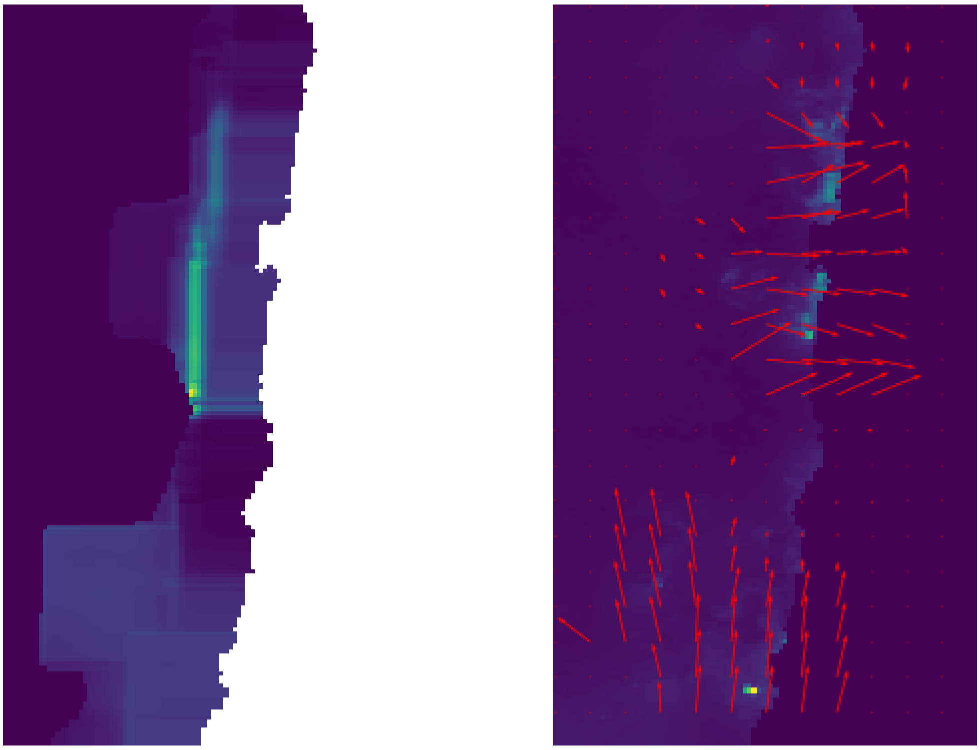

Figure 6: Optical flow. An initial approach to analyze the movement of chlorophyll concentrations in the ocean.

Conclusions

We confirmed the presence of yearly seasonal trends, with higher chlorophyll concentrations in summer and lower concentrations in winter. Chlorophyll peaks are observed closer to the coast, though the exact cause—whether due to temperature, shallower waters, or proximity to nutrient sources like rivers—requires further investigation.

Peak detection shows potential, though additional testing is needed. While optical flow appears promising, it also requires further validation.

Reproducibility

The base code, datasets, and requirements are provided alongside this document in a Jupyter notebook to ensure reproducibility.