Exploring Plankton Health and Its Implications within the Arctic

Module Code: MTHM507

Module Name: Communicating Data Science

Student Number: 710026691

Word Count: 4395

Individual Report

AI-supported/AI-integrated use is permitted in this assessment. I acknowledge the following uses of GenAI tools in this assessment:

- I have used GenAI tools to proofread and correct grammar or spelling errors

- I have used GenAI to help cross-check coding/plotting

I declare that I have referenced use of GenAI outputs within my assessment in line with the University referencing guidelines.

Table of Contents

Introduction 3
Scientific Context and Uncertainty Surrounding Phytoplankton 4
Methodology 5
  Data Sources 5
  Data Science Methods 6
Results and Analysis 6
  Cross Correlation Analysis 6
  Vector Autoregression Model 8
  SARIMA Models 9
  Granger Causality Test 10
  Spectral Analysis 11
Limitations and Future Directions 12
Policy Recommendations 13
Conclusion 14
References 15
Appendices 17
  Appendix A – Data Sources 17

Introduction

Plankton are a diverse and fundamental group of organisms inhabiting aquatic environments, categorised phytoplankton (plant-like) and zooplankton (animal-like). They play a vital role in marine ecosystems, acting as the foundation of the food chain, producing oxygen and absorbing carbon dioxide. Zooplankton are typically smaller organisms that feed on phytoplankton and other small organic particles. Phytoplankton, specifically, are microscopic organisms that convert inorganic carbon into organic matter through photosynthesis. Phytoplankton significantly influence global ecological and climatic processes due to their role in the carbon cycle and climate regulation. Phytoplankton have had considerable impact with their role in carbon sequestration – acting as a carbon sink by transferring carbon dioxide from the atmosphere into oceanic depths through biological and physical processes (Falkowski, 2012). Assessing the health and productivity of phytoplankton is essential in understanding and predicting broader environmental and climatic outcomes in the future.

This report aims to examine the plankton health status within the Arctic region, which is experiencing climate change at an accelerated rate, warming up to four times faster than the rest of the world. This phenomenon, known as Arctic amplification results in rapid environmental change and provides a unique opportunity to study the impacts of rising temperatures and associated environmental stressors on phytoplankton populations (Rantanen et al., 2022).

There are five common indicators that are studied to assess the health status of plankton including Biomass, Abundance, Diversity, Phenology and Morphology. This report will be primarily focusing on Biomass which refers to the total amount of phytoplankton present within a given area, serving as a robust indicator of overall phytoplankton productivity and ecosystem health (Falkowski, 2012). Biomass is a particularly advantageous metric because it provides a stable and consistent representation of phytoplankton populations over time, allowing for precise and reliable analysis using various data science techniques. Through employing advanced analytical techniques such as time series analysis, cross-correlation, vector autoregression (VAR), spectral analysis and Granger causality tests, we can quantify the relationships between phytoplankton biomass productivity and key environmental factors in the Arctic. These insights will provide a comprehensive assessment of climate change impacts in this region, supporting future policy recommendations aimed at preserving the Arctic ecosystem.

To assess the health status of Arctic phytoplankton, this report will first present a background review of current scientific understanding of plankton dynamics and their responses to environment stressors - particularly sea surface temperature (SST) and salinity. This review will also explore the key uncertainties surrounding plankton that remain in literature such as long-term observational data and regional variability in responses. The methodology section will outline the data collection process and then discuss the data science techniques chosen for the analysis. These methods will then be applied to evaluate the impact of key environmental variables on phytoplankton biomass and bloom dynamics within the Arctic. The goal is to understand the underlying dynamics and determine whether observed changes are due to direct effects or mediated by seasonal or cyclical processes. Finally, the analytical insights will be contextualised with a policy framework, offering informed strategic recommendations aimed at preserving the Arctic marine ecosystems under accelerating climate pressures.

Scientific Context and Uncertainty Surrounding Phytoplankton

The term ‘Phytoplankton’ encompasses diverse taxonomic groups, including single-celled algae (e.g. diatoms, green algae, dinoflagellates) and photosynthetic bacteria (cyanobacteria). Despite their microscopic size, phytoplankton are incredibly diverse in form and function. Some possess mineral type shells such as silica frustules in diatoms or calcium carbonate plates in coccolithophores, while others are naked cells (Lindsey & Scott, 2010). This diversity is what allows phytoplankton to thrive in almost all sunlit aquatic habitats. Even though phytoplankton make up a less than 1% of the total living plant material (photosynthetic biomass) on Earth, they are incredibly productive. They produce around 50% of the planet’s oxygen and are responsible for nearly half of global primary production, rivalling terrestrial plants in their contribution to carbon fixation (Field et al., 1998). Certain species of phytoplankton, however, can form harmful algal blooms (HAB’s), which produce toxins which can lead to hypoxic “dead zones” in nutrient-rich coastal regions (Amaneesh et al., 2023).

Phytoplankton are central to regulating the Earth’s climate through the biological carbon pump which is a mechanism whereby carbon fixed during the photosynthesis process is transported to deep ocean layers, removing it from the atmosphere (Falkowski & Raven, 2007). As phytoplankton die or are consumed, part of the organic matter sinks into the deep ocean which contributes to long-term carbon sequestration. As a result, it is estimated that phytoplankton drive approximately 10 gigatonnes of carbon annually into the deep ocean, further signifying their involvement in moderating atmospheric carbon dioxide levels and climate regulation (Falkowski et al., 1998). However, the efficiency of this carbon pump depends heavily on the composition of phytoplankton communities which is difficult to measure. Larger, heavier diatoms tend to sink faster and therefore transport carbon more efficiently than smaller ‘picoplankton’ which contribute less to deep ocean carbon flux (Tremblay et al., 2015).

Many long-term studies have reported declines in global phytoplankton abundance (Boyce et al., 2010), while others suggest these findings are uncertain due to limited temporal and spatial observational overage. Satellite observations and data collected since the late 1990’s show regional variability in trends and highlight the sensitivity of phytoplankton to environmental drivers such as temperature, nutrient supply, and light availability (Lewis et al., 2020). Despite growing interest and developments in observational tools, there is significant scientific uncertainty that remains surrounding the current and future state of phytoplankton, particularly in remote and rapidly changing regions such as the Arctic. Satellite measures are obstructed during winter months by sea ice both spatially and temporally (Johnston et al., 2020). Additionally, there is uncertainty surrounding the accuracy of the satellites that detect surface-level chlorophyll concentrations as a proxy for biomass, which may not be able to capture deeper blooms or shifts in species compositions.

Recent studies suggest there have been shifts in phytoplankton community composition as a direct result of climate change, especially in polar regions such as the Arctic. Smaller species such as picoplankton and nanoflagellates have been observed more frequently in parts of the Arctic Ocean, likely due to surface freshening and intensified stratification (Li et al., 2009). Stratification refers to the vertical layering of plankton within a water column based on density differences. The differences in water density occur as a result from differences in temperature and salinity changes (Queensland Government, n.d.). These smaller species of plankton that thrive in the Arctic can live under poor nutrient condition and are less efficiently grazed by larger zooplankton, potentially weakening the trophic transfer efficiency of the Arctic food web. Additionally, changing bloom timing due to earlier sea ice retreat – driven by ongoing climate change - can disrupt ecological synchrony, leading to mismatches between the peak availability of phytoplankton and the life cycles of dependent organisms such as zooplankton and fish larvae (Ardyna & Arrigo, 2020). The changes in bloom timings reveal how Arctic ecosystems can respond in different and unexpected ways, highlighting the need to use long-term data partnered with advanced modelling tools to better understand and predict how these systems will respond to climate change.

There is a complex interplay between multiple environmental divers such as sea surface temperature (SST), salinity, nutrient availability, and pollution. For example, while warmer waters may enhance phytoplankton growth in some high-latitude regions due to increased light and reduced ice cover, it may also increase stratification which can limit nutrient mixing and ultimately suppress productivity elsewhere (Li et al., 2009). This adds further challenges in isolating causality and forecasting for long-term ecological responses.

Given the intricate interplay between these environmental factors, the application of robust quantitative methods is essential for understanding how they influence phytoplankton biomass. This is particularly true in the Arctic, where rapid climate change is occurring and traditional ecological assumptions surrounding phytoplankton may no longer hold. Data science methods therefore provide a powerful framework for identifying temporal patterns, detect lagged relationships and understand the interplay between key environmental drivers toward phytoplankton productivity within the Arctic.

Methodology

Data Sources

This study draws upon multiple publicly available datasets to examine the correlation between phytoplankton biomass and key environmental variables within the Arctic, focusing on the region north of Iceland. Our primary objective as previously mentioned is to evaluate the extent to which sea surface temperature (SST), salinity, and associated time-series indicators relate to phytoplankton productivity over time, using statistical and data science techniques that are suited to temporal ecological analysis.

The core datasets used originate from the Ocean Colour Climate Change Initiative (OCCI) which contains monthly phytoplankton data in the Arctic. Monthly sea surface temperature data was scarped from HadiSST (Hadley Centre Sea Ice and Sea Surface Temperature dataset). These time-series span from 1998 to 2022, covering both Arctic and Antarctic regions. Noticeably, the winter months are absent in some years due to polar night and the unavailability of satellite-derived colour data. This is a limitation that is addressed during data preprocessing.

Supplementary data was used to improve the robustness of models used in the analysis including the Hadley EN4 Salinity Dataset, providing monthly salinity time-series data which is crucial for examining the stratification and freshwater influence. Further data on SST was scraped from COPEPOD and is used to validate patterns across Arctic marine regions. All datasets were standardised and merged, ensuring temporal and spatial alignment. Missing data point, especially in winter months were addressed using seasonal decomposition and interpolation where appropriate, though missing data caused by polar nights has been unaltered. All data sources used can be found in Appendix A.

Data Science Methods

To discover temporal dependencies and potential relationships between phytoplankton biomass (measured as chlorophyll-a concentration) and environmental indicators, we conducted a range of time-series analysis techniques using R studio. The `tidyverse`, `forecast`, `lubridate`, `vars` and `TSA` packages were used to handle data wrangling, visualisation, model fitting and diagnostic testing.

The first method used is a cross-correlation analysis which aims to identify the lag and strength of associations between SST, salinity and chlorophyll-a concentrations across different time intervals. This method allows us to understand the changes in environmental variables that precede phytoplankton responses.

Additionally, we will use Vector Autoregression (VAR), a multivariate technique which is used to capture interdependencies and dynamic feedback between SST and phytoplankton biomass. This method is useful for simultaneous modelling of multiple time-series data, allowing for a better understanding of how phytoplankton dynamics both influence and are influenced by environmental factors.

We will then apply a Seasonal Autoregressive Integrated Moving Average (SARIMA) model that tracks seasonality. By removing seasonality, we hope to isolate underlying non-seasonal relationships. Through capturing seasonal cycles within the Arctic ecosystems (bloom peaks during summer), SARIMA ensures that identified correlations are not only a result of repeating annual patterns.

Using a Granger Causality Test in conjunction with VAR models, we can evaluate whether SST or salinity time series can statistically predict changes in chlorophyll-a concentrations. This approach tests directional causality by indicating whether one variable contains predictive information about another.

Finally, a spectral analysis was conducted to detect recurring seasonal cycles or periodicities in the phytoplankton data, which should provide insights into bloom dynamics and external forcing patterns, such as ice melt timing or light cycles. This method should reveal how regular phytoplankton blooms are over the observed 25-year period.

Results and Analysis

Cross Correlation Analysis

To examine and identify the relationship between sea surface temperature (SST) and phytoplankton productivity, we performed a cross-correlation analysis between chlorophyll-a and SST anomalies in the Arctic Ocean. The multi-axis time-series plot below in Figure-1 displays the interplay between the different environmental indicators being observed over the period of 1998 to 2022. Notably, Chlorophyll-a (CHLA) and Net Primary Production (NPP) both show increasing trends after 2010, while salinity declines, SST anomalies peak in 2006 before trending downward. These patterns may suggest shifts in phytoplankton dynamics caused by freshening conditions, despite the recent cooling in SST, potentially linked to Arctic environmental changes and stratification effects.

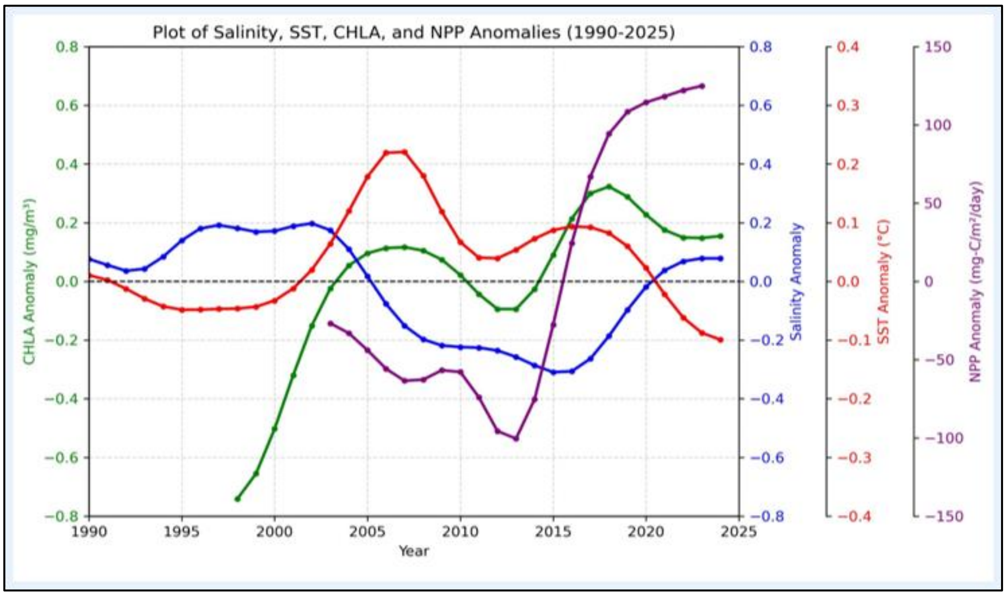

Figure-1 – Long Term Trends of Different Environmental Indicators

The cross-correlation function (CCF) plot in Figure-2 demonstrates a cyclical correlation pattern between SST and CHLA in the Arctic, with significant correlations shown at both positive and negative lags. Specifically, the correlation varies over a 10-lag period which suggests a recurring temporal structure in the relationship between the variables. The positive peaks in correlation are shown at around +7 to +10 months, whereas the negative correlations occur at lag 0 or shortly after around lag 5. These patterns indicate that SST anomalies not only have a significant immediate effect on phytoplankton activity but also could affect CHLA levels months later. This could be explained through delayed physical or biological processes such as stratification, nutrient mixing, or bloom timings (Ardyna & Arrigo, 2020; Johnston et al., 2020).

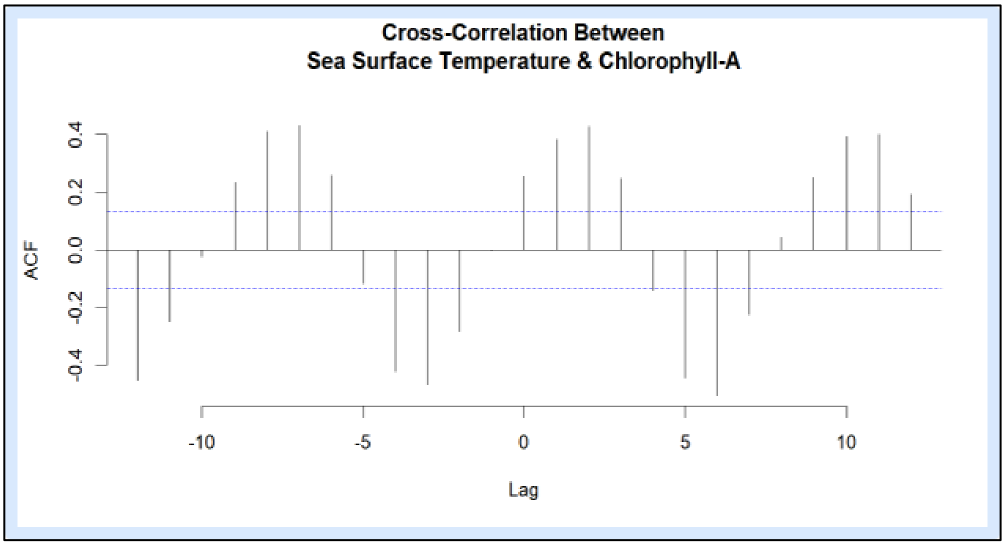

Figure-2 – Cross-Correlation Function Plot Between SST and CHLA

Pearson Correlation Analysis	
CHLA vs Salinity	**r = 0.328**
CHLA vs SST	**r = -0.399**
CHLA vs NPP	**r = 0.568**
SST vs Salinity	**r = -0.628**

Table-1 – Pearson Correlation Analysis Results

The moderate negative correlation observed at lag 0 suggests that the higher SST anomalies are associated with suppressed phytoplankton biomass in the same month. This is supported by the Pearson correlation coefficient of -0.399, derived from the bivariate analysis of monthly SST and chlorophyll-a (CHLA) anomalies. Therefore, it is shown that elevated SST may improve ocean stratification, reducing nutrient availability in surface waters and limiting phytoplankton growth. Ocean stratification limits phytoplankton growth by restricting the mixing of nutrient-rich deep waters with the sunlit surface, reducing the availability of essential nutrients such as nitrogen and phosphorus, crucial for phytoplankton growth (Li et al., 2009). In contrast, the positive correlations shown at future lags may reflect compensatory responses such as where biological mechanisms or delayed mixing processes have restored biomass levels after the initial suppression.

Vector Autoregression Model

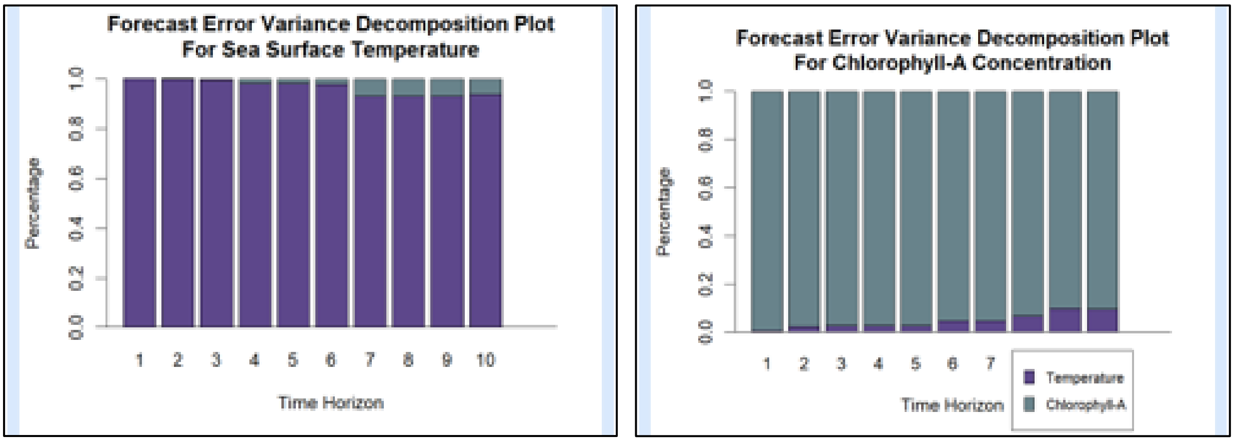

Figure-3 – VAR Model Forecasted Error Variance Decomposition Plots for SST and CHLA

The VAR model was used to evaluate the bidirectional temporal relationships between SST and phytoplankton biomass. Results from the model indicate a strong autoregressive component for SST with an R squared value of 0.93. This suggests a significant relationship with most of its own lags and with some of chlorophyll's lags, indicating that SST is highly predictable from its own past values. On the other hand, chlorophyll-a, showed a moderate R squared value of 0.63 displaying a significant relationship with a few of SST's lags but to a weaker extent. This suggests that phytoplankton biomass is less self-driven and more susceptible to external influences.

While both SST and CHLA had statistically significant relationships with their own lags, only a small portion of CHLA variance was explained by past SST values, as shown in the forecast error variance decomposition plot in Figure-3. Across a 10-month forecast horizon as shown, CHLA was mostly driven by its own past values, with SST contributing only a minor share to its variability. This asymmetry is mirrored with the SST decomposition plot, where CHLA plays a similarly limited role in predicting SST variability.

These results suggest that SST and CHLA do exhibit lagged correlations, but not necessarily predictive power in a causative way. This is further proven through ecological studies within the Arctic that highlight that phytoplankton responses to temperature are often mediated by intermediate factors such as nutrient mixing, light availability, and salinity-driven stratification (Li et al., 2009). Therefore, though SST has been proven to influence phytoplankton productivity within the Arctic, it is unlikely to be the prominent driver of bloom dynamics alone.

The VAR results corroborate the findings from the cross-correlation analysis, reinforcing that the relationship between SST and phytoplankton biomass is structured temporally but is influenced by multiple ecological processes. Rather than acting in isolation, sea surface temperature interacts with a range of environmental factors – such as salinity, light availability, and nutrient dynamics that collectively influence phytoplankton productivity levels within the Arctic.

SARIMA Models

To isolate seasonal patterns and assess their contribution to variance in sea surface temperature and chlorophyll-a (CHLA) concentration, SARIMA models were fitted to both variables. Prior to model fitting, stationarity was confirmed in both datasets using the Augmented Dickey-Fuller (ADF) test, which ensures that the data is appropriate for time-series modelling. The Mean Absolute Percentage Error (MAPE) metric is used to quantify the forecasting accuracy of the models by measuring the average absolute error between predicted and actual values, expressed as a percentage of the actual values.

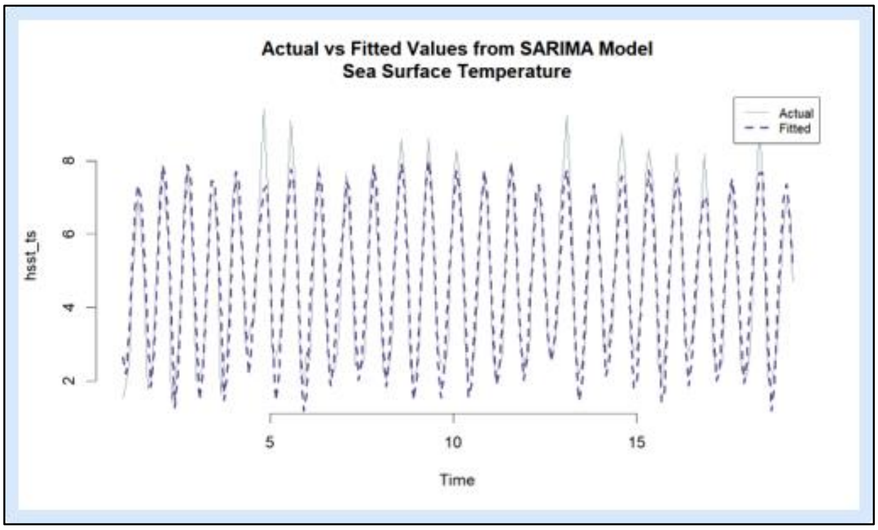

Figure-4 – SARIMA Model's Actual vs Fitted Values for SST

The SARIMA model for SST in Figure-4 above, effectively captures the seasonal cycles. The SST model achieved a Mean Absolute Percentage Error (MAPE) of 16.83%, a moderately strong score that indicates the model explains a substantial proportion of SST's monthly variability through recurring seasonal patterns. This strong seasonal pattern is consistent with previous Arctic studies that identify temperature as a strongly cyclical variable, driven by solar radiation and ice cover fluctuations throughout each year (Lewis et al., 2020).

In contrast, the CHLA SARIMA model in Figure-5 below, revealed a MAPE of 39.17%, which suggests a moderately weak predictive accuracy. Despite the model capturing some seasonal bloom dynamics, such as peaks during summer months, there was considerably more irregularity and unexplained variance in phytoplankton biomass. This result is significant, suggesting that although phytoplankton in polar regions do exhibit seasonal blooms, their dynamics are heavily influenced by multiple interacting factors as previously mentioned including nutrient supply, ice melt timing, and water column stratification (Li et al., 2009). These factors do not strictly follow seasonal patterns and therefore are more difficult to capture through the SARIMA approach alone.

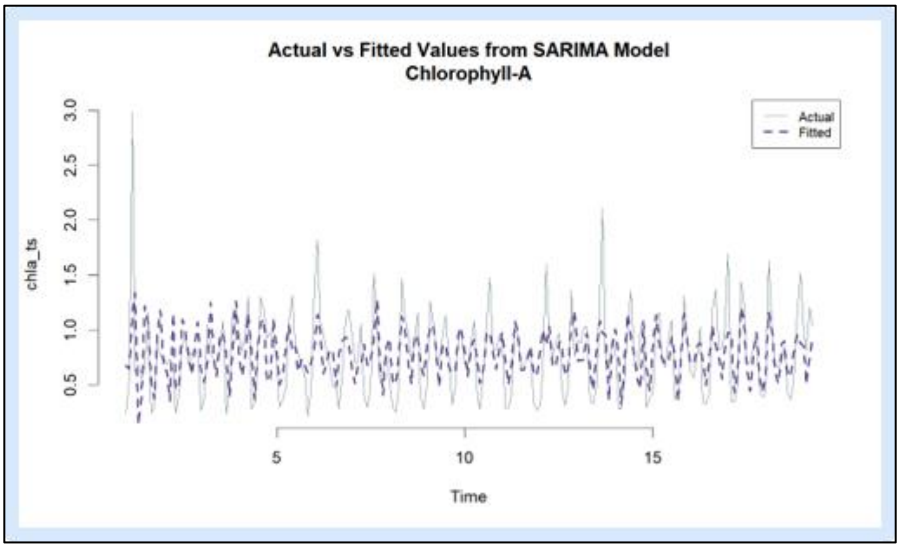

Figure-5 – SARIMA Model's Actual vs Fitted Values for Chlorophyll-A

After fitting the SARIMA models, the residuals from both SST and CHLA were analysed to assess the strength of their non-seasonal relationship. The cross-correlation of residuals revealed that some weak seasonal structure remained, though it was largely reduced. A subsequent VAR model applied to these residuals produced R squared values of 0.30 for SST and 0.29 for CHLA respectively, suggesting a weak-to-moderate relationship that may still reflect the lingering seasonal effects.

These findings imply that seasonality explains a large portion of SST variability, but plays a less prominent role in phytoplankton dynamics, which are subject to more complex and potentially non-linear ecological drivers.

Granger Causality Test

The Granger causality test is applied to assess whether SST can predict chlorophyll-a concentration or vice versa, across a range of 10 lag periods. The results shown below in Table-2 show significant p-values (much smaller than 0.05), implying past values of one series are useful for predicting the other in both cases.

Granger Causality Test Results

Test Direction	Model 1	Model 2	F Statistic	P Value
SeaTemp to Plant	Plankton ~ Lags(Plankton, 1:10) + Lags(SeaTemp, 1:10)	Plankton ~ Lags(Plankton, 1:10)	-10	5.0407 × 10^-6
Plant to SeaTemp	SeaTemp ~ Lags(SeaTemp, 1:10) + Lags(Plankton, 1:10)	SeaTemp ~ Lags(SeaTemp, 1:10)	-10	4.5243 × 10^-6

Table-2 – Granger Causality Test Results

From the test direction SeaTemp to Plankton – the F-statistic was 5.0407 x 10^{-6}, and for Plankton to SeaTemp, the F-statistic was 4.5243 x 10^{-6}. These low p-values indicate that both SST and CHLA exhibit Granger causality in either direction, meaning that changes in past Sea Surface Temperature values are statistically useful for predicting future cholorophyll-a concentrations and vice versa.

These findings suggest that temperature changes in the Arctic do play an important part in predicting phytoplankton dyanmics, and that phytoplankton biomass also has some degree of predictive power on sea surface temperature changes, likely due to feedback mechanisms or indirect effects including light availability or nutrient cycling.

Spectral Analysis

Spectral density analysis is performed on both Arctic Sea Surface Temperature (SST) and phytoplankton data to identify the frequency components of a time series. This analysis is useful for detecting periodic signals in the data, such as seasonal cycles or oscillations that repeat and regular intervals.

The spectral density plot for SST in Figure-6 displays significant peaks at frequencies 1 and 2, which indicate the presence of annual and semi-annual cycles. The first peak at frequency 1 is particularly prominent, corresponding to the annual seasonal cycle driven by solar radiation, sea ice dynamics, and the Earth's axial tilt, which strongly affect Arctic temperatures (Lewis et al., 2020). The secondary peak at frequency 2 could be linked to atmospheric oscillations such as the Arctic Oscillation, which impacts sea surface temperature biannually (Ardyna & Arrigo, 2020). As the frequencies increase, the lower amplitude suggests that short-term fluctuations in temperature are less significant compared to the long-term seasonal cycles.

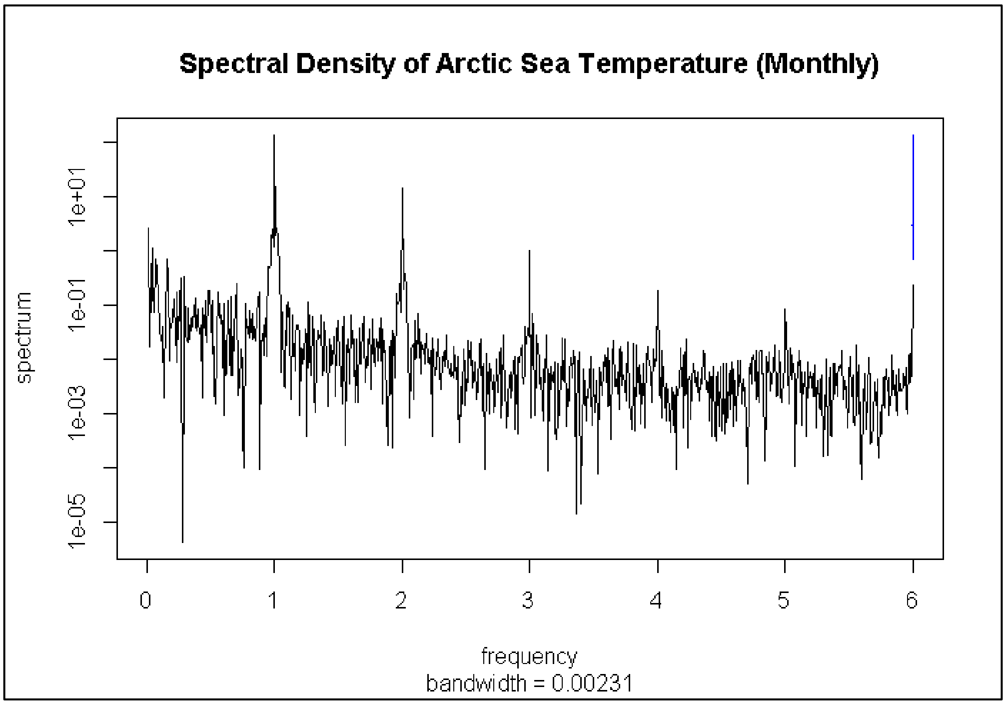

Figure-6 – Spectral Density Plot of Arctic Sea Surface Temperature

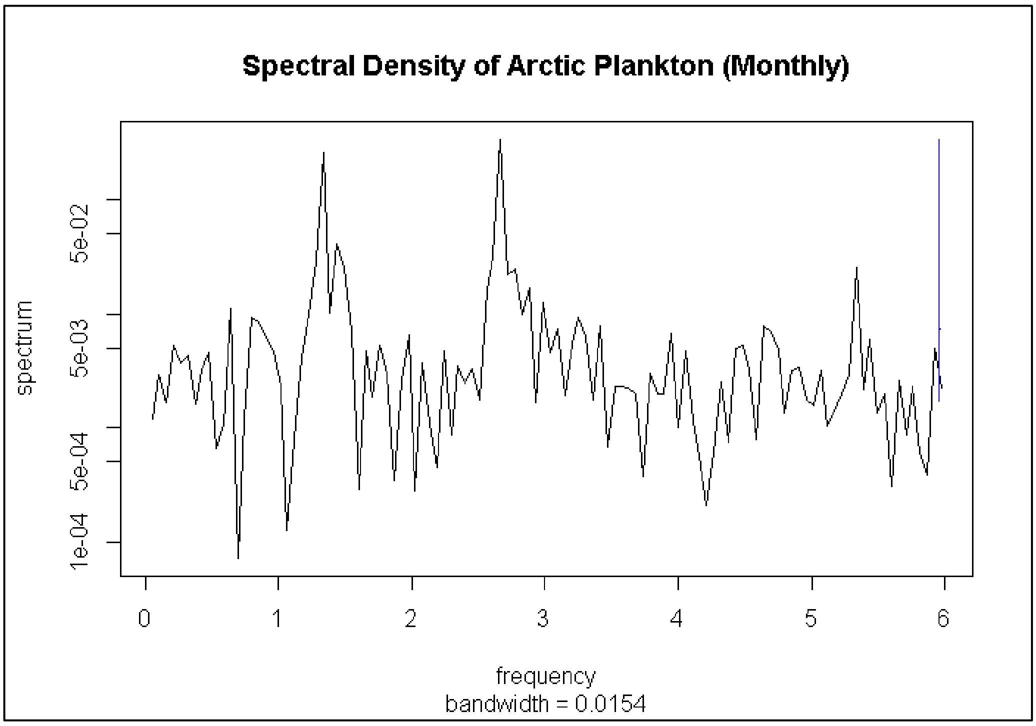

Figure-7 – Spectral Density Plot of Arctic Phytoplankton CHLA

The spectral density plot in Figure-7 for Arctic phytoplankton displays a complex profile, with multiple peaks across a range of frequencies. There are dominant peaks at frequencies between 1 and 3, reflecting annual seasonal bloom dynamics of phytoplankton, which are strongly related to the annual cycle of sunlight and temperature changes. The broader and more variable spectrum in phytoplankton data suggest that other factors, such as nutrient availability, sea ice dynamics, and water column mixing, play a role in modulating bloom timing and intensity. For phytoplankton, the bandwidth is 0.0154, which shows a broader frequency range that captures the complex variability in phytoplankton productivity. This wider variability in the phytoplankton spectrum compared to SST underpins the point that phytoplankton in the Arctic are subject to more complex, interacting drivers beyond just temperature changes, including biological processes and ecological feedback (Li et al., 2009).

The spectral density analysis confirms that seasonality plays a central role in both Arctic Sea Temperature changes and phytoplankton dynamics. However, the more complex spectral pattern for phytoplankton dynamics highlights that their productivity is influenced by a variety of factors beyond temperature. This complexity suggests that predictive models of phytoplankton biomass should account for multiple environmental drivers in conjunction with temperature.

Limitations and Future Directions

While this report offers valuable insights into the relationships between Arctic Phytoplankton biomass and relevant indicators such as Arctic Sea Surface Temperature, several limitations must be addressed.

There are multiple data gaps during polar winter months, particularly for chlorophyll-a (CHLA) concentrations, due to a lack of sunlight. Despite these gaps being addressed using seasonal decomposition and interpolation techniques, missing winter data can affect the accuracy of the analysis, especially for phytoplankton productivity in the winter period, when temperature and light availability play the most significant roles (Johnston et al., 2020). Satellite-derived data for high-latitude regions is also limited by sea ice and cloud cover which restricts the availability of accurate measurements during periods (Lewis et al., 2020).

The analysis relies on linear statistical models such as VAR, SARIMA, and Granger causality tests which assume that relationships between variables (SST and CHLA) are linear and stationary. However, non-linear interactions may exist, particularly in complex Arctic ecosystems where multiple drivers influence phytoplankton dynamics. For example, this is shown in ecological studies where the non-linear nature of environmental interactions between temperature and nutrient availability is not fully captured by linear methods (Henson et al., 2015). Future research could benefit from incorporating non-linear modelling and machine learning techniques such as using random forests or neural networks to better capture the relationship between SST and CHLA dynamics. These methods have been useful for ecological research of large, complex datasets and uncovered non-linear dependencies that traditional linear models may miss (Kearney et al., 2018).

The data used in this assignment is limited to the region north of Iceland. While this study provides insights into this specific Arctic region, SST and phytoplankton dynamics can vary significantly throughout other part of the Arctic due to differences in ocean currents, regional temperature anomalies, and ice dynamics (Ardyna & Arrigo, 2020). Through expanding spatial coverage to include different Arctic sub-regions, the results can be averaged, and more robust conclusions can be formed around the Arctic region (Johnston et al., 2020). The data used in this study was available at monthly resolution, which is useful for capturing long-term trends. However, to capture short term fluctuations in SST and phytoplankton biomass, a higher resolution would allow for a more detailed exploration of how seasonal transitions and abrupt changes (e.g. light availability or ice melt timing) influence phytoplankton growth.

To improve the accuracy and applicability of this research for future studies, the implementation of field-based observations with satellite data is imperative, particularly during period when satellite-derived data is missing. Therefore, using in situ sensors and remote sensing technologies for long-term monitoring of phytoplankton blooms would provide a better understanding of local variations in phytoplankton productivity and seasonal patterns across different Arctic regions.

Furthermore, given the accelerating pace of climate change in the Arctic, further research should focus on modelling how shifting temperature baselines and ice retreats impacts phytoplankton biomass. Using climate models and ocean circulation models, we can predict how changing environmental conditions can affect Arctic food webs, marine eco system health and carbon cycling (Falkowski & Raven, 2013).

Policy Recommendations

The implementation of Ecosystem-Based Management is crucial given the importance of phytoplankton as the foundation of Arctic marine food webs. This management would formulate strategies to address the impacts of climate change on marine ecosystems in a comprehensive manner, balancing the need for environmental conservation with the challenges posed by rapid Arctic warming.

Policy making should aim to protect these organisms to preserve their carbon sequestration capacity. This could include measures preventing further degradation of phytoplankton habitats due to pollution, overfishing, or other anthropogenic pressures. Further policies could include funding for blue carbon initiatives, where Arctic marine ecosystem can be preserved or restored to enhance their capacity to sequester carbon.

Through increasing international collaboration between Arctic nations, research organisations, and indigenous communities, this can improve the monitoring, data collection, and analysis of Arctic ecosystems. This could include partnerships for shared satellite data platforms, data sharing on climate models or field-based research on Arctic ecosystems.

Conclusion

This report provides valuable insights into the relationship between phytoplankton biomass and key environmental factors in the Arctic, particularly sea surface temperature (SST). Through the application of advanced data science techniques, such as cross-correlation analysis, VAR models, and SARIMA models, we have demonstrated that phytoplankton productivity is influenced by complex, non-linear ecological interactions. The findings underscore that phytoplankton is a complex variable that cannot be accurately predicted through linear models alone. While temperature is a significant factor, it is also heavily influenced by nutrient mixing, stratification, and other biological processes, making it difficult to predict without considering the interplay of these additional environmental variables.

The results highlight the crucial role of the Arctic in global carbon cycling, as phytoplankton acts as a carbon sink, but also reveal that predicting changes in phytoplankton biomass requires a broader understanding of the factors driving its dynamics. Nutrient availability, stratification, and ice dynamics all play pivotal roles in shaping phytoplankton blooms, which are not governed by seasonal patterns alone. As a result, further research is needed to better understand these non-linear relationships and improve predictions of how climate change will affect Arctic marine ecosystems.

Given the complexity of phytoplankton dynamics in the Arctic, there is a clear need for adaptive, multi-faceted policies to safeguard Arctic ecosystems. By addressing the limitations in data gaps, improving collaborative research efforts, and implementing holistic ecosystem management practices, future policies can better support the resilience of Arctic marine resources. This integrated approach will be essential in ensuring that the Arctic continues to function as a vital part of the global climate system.

References

Amaneesh, C., Balan, S. A., Silpa, P. S., Kim, J. W., Greeshma, K., Mohan, A. A., ... & Ramanan, R. (2023). *Gross negligence: Impacts of microplastics and plastic leachates on phytoplankton community and ecosystem dynamics*. Environmental Science & Technology, 57(1), 5–24. https://doi.org/10.1021/acs.est.2c05817

Ardyna, M., & Arrigo, K. R. (2020). Phytoplankton dynamics in a changing Arctic Ocean. *Nature Climate Change, 10*(10), 892–903. https://doi.org/10.1038/s41558-020-0905-y

Boyce, D. G., Lewis, M. R., & Worm, B. (2010). *Global phytoplankton decline over the past century.* Nature, 466(7306), 591–596. https://doi.org/10.1038/nature09268

Falkowski, P. (2012). *Ocean Science: The power of plankton. Nature, 483*(7387), S17–S20. https://doi.org/10.1038/483s17a

Falkowski, P. G., & Raven, J. A. (2013). Aquatic Photosynthesis: Second Edition. In *Google Books*. Princeton University Press. https://books.google.co.uk/books?hl=en&lr=&id=kUCrAQAAQBAJ&oi=fnd&pg=PP1&dq=Falkowski

Falkowski, P. G., Barber, R. T., & Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science, 281*(5374), 200–206. https://doi.org/10.1126/science.281.5374.200

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). *Primary production of the biosphere: Integrating terrestrial and oceanic components*. Science, 281(5374), 237–240. https://doi.org/10.1126/science.281.5374.237

Henson, S. A., et al. (2015). A global overview of marine phytoplankton dynamics and the implications for ocean food webs. *Nature Climate Change*, 5(3), 199–208. https://doi.org/10.1038/nclimate2578

Johnston, N. T., Hamel, J., Tremblay, J. É., & Gosselin, M. (2020). *Under-ice phytoplankton blooms and long-term observations reveal recent changes in Arctic primary productivity*. Global Change Biology, 26(11), 6369–6385. https://doi.org/10.1111/gcb.15285

Kearney, M. R., et al. (2018). The predictive ecology of species distributions: Model selection, model averaging, and predicting the effects of global change. *Ecology Letters*, 21(2), 122-137. https://doi.org/10.1111/ele.12904

Lindsey, R., & Scott, M. (2010). *What Are Phytoplankton?* Nasa.gov; NASA Earth Observatory. https://earthobservatory.nasa.gov/features/Phytoplankton

Li, W. K. W., McLaughlin, F. A., Lovejoy, C., & Carmack, E. C. (2009). *Smallest algae thrive as the Arctic Ocean freshens*. Science, 326(5952), 539. https://doi.org/10.1126/science.1179798

Lewis, K. M., van Dijken, G. L., & Arrigo, K. R. (2020). *Changes in phytoplankton concentration now drive increased Arctic Ocean primary production*. Science, 369(6500), 198–202. https://doi.org/10.1126/science.aay8380

Queensland Government. (n.d.). *Stratification*. WetlandInfo. Department of Environment, Science and Innovation. https://wetlandinfo.des.qld.gov.au/wetlands/ecology/processes-systems/stratification/

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., ... & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications earth & environment*, *3*(1), 168.

Tremblay, J. É., Bélanger, S., Barber, D. G., Asplin, M., Martin, J., Darnis, G., ... & Fortier, L. (2015). Climate forcing multiplies biological productivity in the coastal Arctic Ocean. *Geophysical Research Letters, 42*(18), 7749–7756. https://doi.org/10.1002/2015GL065049

Appendices

Appendix A – Data Sources

- CSV files containing OCCI monthly phytoplankton data and SST data from HadiSST: https://iop.eventsair.com/hackathon2023/data-sets
- NOAA OI v2.1 Sea Surface Temperature (SST) time series: https://www.st.nmfs.noaa.gov/copepod/data/ru-03101/index.html
- PSMSL Sea Level Data: https://psmsl.org/data/