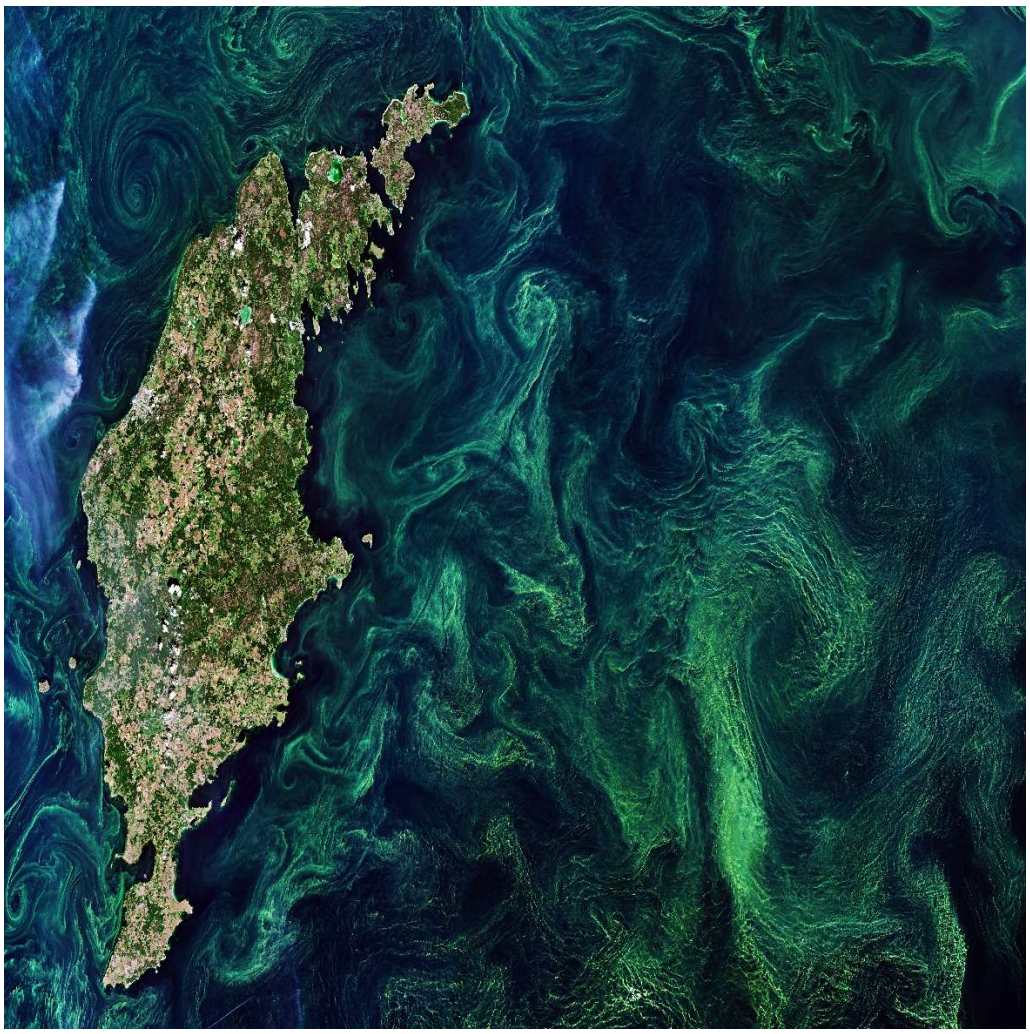
Hypoxia increase and the correlation to Cyanobacteria Blooms in the Western Gotland Basin from 1998 to 2024



Baltic Blooms by European Space Agency, ESA, used under CC BY-SA 3.0 IGO.

Alva Johansson

2026

Department of
Earth and Environmental Sciences
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Alva Johansson (2026).
Hypoxia increase and the correlation to Cyanobacteria Blooms in the Western Gotland Basin from 1998 to 2024.
Ökningen av syrefattiga bottnar och korrelationen till blomning av cyanobakterie-alger i Västra Gotlandbassängen från 1998 till 2024.
Bachelor degree thesis, 15 credits in *Physical Geography and Ecosystem Analysis*
Department of Earth and Environmental Sciences, Lund University

Level: Bachelor of Science (BSc)

Course duration: *March* 2026 until *June* 2026

Disclaimer

Hypoxia increase and the correlation to Cyanobacteria Blooms in the Western Gotland Basin from 1998 to 2024

Alva Johansson

Bachelor thesis, 15 credits, in Physical Geography and Ecosystem Analysis

Supervisor
José Beltrán
PhD Department of Earth and Environmental Sciences, Lund University

Course coordinator
Jonas Ardö
Professor Department of Earth and Environmental Sciences, Lund University

Examinator:
Helena Filipsson
Professor Department of Earth and Environmental Sciences, Lund University

Acknowledgement

I want to especially thank my dear friend and biggest academic inspiration, Savvina Leontakianakou. She once told me that if I don't know what to do next, choose the path where you can have the greatest positive impact on others. This has stuck with me while studying my bachelor's degree and writing this thesis. You inspire me to do better and be better, in science and in life. Thank you for believing in me Sav.

I wish to express gratitude to all of my family, friends and course mates for their kindness and support through this. I am so lucky to be surrounded by such competent, kind and patient people. Especially when I was deep in the writing you were there for me. Thank you.

I also wish to express my gratitude to my academic institution, the Department for Physical Geography and Ecosystem Science, now called Department of Earth and Environmental Sciences at Lund University, for the opportunity of a global learning experience through these three years, it taught me more than a book could.

Finally, I wish to thank my sister, Ida. You always inspire me, always has and always will. Thank you for being my reminder in life of what's important.

Abstract

Cyanobacterial blooms and bottom-water hypoxia have long been studied and connected to the eutrophication in the Baltic Sea. This thesis investigates the dynamic of the spatiotemporal correlation between cyanobacterial summer bloom and dissolved oxygen at the sea floor in the Western Gotland Basin, as per defined by HELCOM 2022. The thesis is conducted using weekly and monthly data from 1998 to 2024, provided by Copernicus Marine Service, on known geophysical and geochemical drivers of Harmful Algal Blooms, HAB's. Basin wide Spearman correlations, Mann-Kendall trend analysis, cross-correlation and partial correlation methods were applied to aggregated monthly data of June, July and August. The basin was divided by the approximate halocline depth of 80 m, comparing correlations for areas with different vertical mixing prerequisites. No statistically significant correlation for pixel-wise and basin mean correlations between sea floor oxygen minimum and the cyanobacteria bloom frequency was found. However a lagged cross-correlation between the same variables of -12 and + 12 months, not statistically significant, still indicated a negative relationship, with different Spearman ρ peaks for waters deeper than 80 m and above it. Dissolved oxygen at the sea floor showed a statistically significant decrease over time, with a p value of p = 8.15e-06 and Sen slope of −21.7 mmol m^{-3}/decade while SST, salinity and chlorophyll a exhibited consistently increasing values over the 27 years.

Table of Contents

1 Introduction

1.1 Study Area

This study was conducted on the brackish inland sea the Baltic Sea, located in northern Europe. It is landlocked and intracontinental, surrounded by Germany, Denmark, Sweden, Finland, Russia, Estonia, Latvia, Lithuania and Poland. The study focuses on the Western Gotland Basin, which has a maximum depth of about 460m (Feistel et al., 2008).

1.1.1 Scope of the study

This study was conducted on data covering the Western Gotland Basin in the Baltic Sea from 1998 to 2024. It was done using Copernicus geospatial data, with National Centers for Environmental Information, NCEI, bathymetry data and HELCOM basin border data to observe the change in the spatial and temporal interaction between hypoxia and cyanobacteria blooms on a weekly and monthly temporal scale. Variables considered were O_2 at the sea floor and occurrence of cyanobacteria bloom together with the biggest drivers of Harmful Algal Blooms, HAB's. The drivers considered in this study are sea surface temperature (SST), salinity, nitrate (NO_3), ammonium (NH_4), phosphate (PO_4), chlorophyll a, wind direction and speed at sea surface combined with the mixed layer depth and temperature at the sea floor. The study area of the Western Gotland Basin was chosen because of the frequent occurrence of cyanobacterial blooms and hypoxia in the area (Hansson & Viktorsson, 2025).

1.1.1.1 Data extent

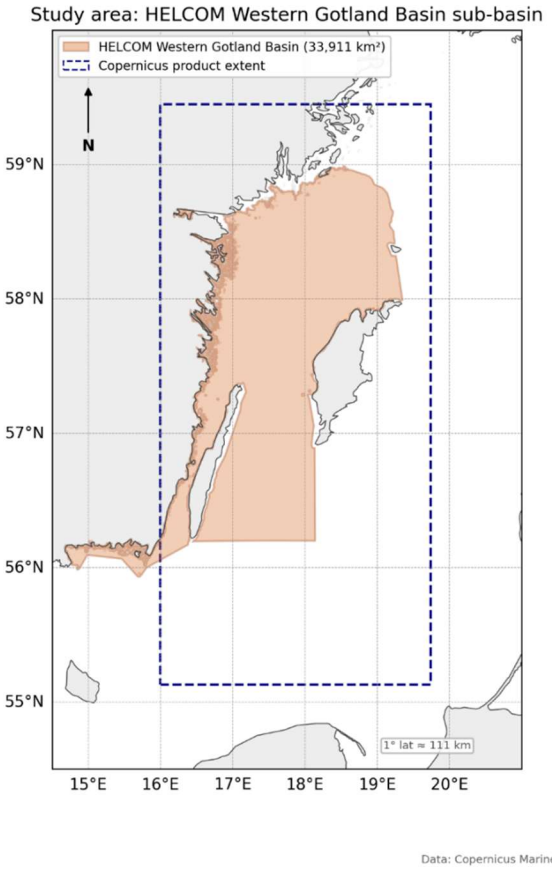


Figure 1. Study area, showing the spatial extent of the Western Gotland Basin as per defined by HELCOM 2022. The blue perforated line indicates the extent of the raster data extracted from Copernicus. This study will be done in the intersecting area of the Western Gotland Basin and the Copernicus data extent.

1.2 Background

1.2.1 Oceanography of the Baltic Sea

The nutrient pools and distribution in the deepwater parts of the Baltic basins depend on inflow events, stagnation and larger variations in oxygen concentration (Nausch et al., 2008). The Baltic Sea has one inflow of saline water from the North Sea through the Danish straits and Kattegat.

The currents transport saline water in the upper layers out of the Baltic Sea into the North Sea, while a counteracting current moves saline and oxygenated water from the North Sea into the Baltic in the deeper layers of the water table. The Baltic Sea is permanently stratified due to salinity differences between the deep and shallow waters, caused by the flow of saline water from the North Sea through the Danish straits. This reinforces the strong halocline and thermocline hindering vertical circulation of the water masses in the water column. Which in turn causes the Baltic Sea to have low ventilation of its bottom layers of water (Lass & Matthäus, 2008). Salinity variations across the basin are also caused by the inflow of freshwater from river runoff from the catchment area for the Baltic Sea (Börgel et al., 2026).

1.2.2 Eutrophication and Hypoxia

Due to anthropogenic activities, mainly agriculture but also industry and forestry etc, the input of available nutrients has increased to the Baltic Sea. Coasts and water bodies that are semi-enclosed like the Baltic Sea, are more susceptible to the possible changes in the ecosystem functioning when nutrient availability increases (Nausch et al., 2008). Over the years from 1850 until 1980, changes in excess phosphorous, P and nitrogen, N loads to the Baltic Sea from fertilizers used in the agricultural factor and anthropogenic input from urban development increased 4.5 times in the Baltic Sea (Funkey et al., 2014). The increase in hypoxia in the Baltic has been linked to eutrophication (Stal et al., 2003 & Vahtera et al., 2007). A common definition of hypoxia used in scientific papers is dissolved oxygen below 2 mg per litre (Funkey et al., 2014). That is the definition used in this work, where areas with oxygen < 2mg/L will be considered hypoxic.

1.2.2.1 Baltic Sea historical hypoxia data

In the current geological time, the Holocene, nitrogen fixing cyanobacterial blooms have been observed more often and as more prevalent during periods of sea floor hypoxia. Using geological sediment samples, the increase in nitrogen fixing cyanobacteria coupled with sea floor hypoxia has been observed to occur during the Littorina transgression, the Medieval Climate Transgression and in current time (1950 - to now circa) (Funkey, et al., 2014).

1.2.3 Algae blooms in the Baltic Sea

Algae blooms in the Baltic Sea is a natural and vital part of the ecosystem. The climatic and environmental circumstances favour a spring bloom and summer bloom of different species. During the spring the Baltic Sea experiences a spring bloom of diatoms and some dinoflagellates (Wasmund & Siegel 2008). The summer blooms tend to be made up of cyanobacteria. The two most common species of cyanobacteria that bloom in the Baltic Sea during the summer bloom are Aphanizomenon Flos-aquae and Nodularia Spumigena. Both species are known N_2 (nitrogen) fixing cyanobacteria, and Nodularia Spumigena is known to create a neurotoxin called nodularin (Stal et al., 2003).

1.2.3.1 Spring Blooms

According to Vahtera et al., (2007) the algae bloom in early spring is triggered by light availability and stratification depth in the surface layer. When the spring bloom has concluded, plenty of the nutrients in the Baltic Sea have been exhausted, specifically the supply of

dissolved inorganic nitrogen, DIN, and dissolved inorganic phosphorous, DIP. Eutrophication increases the availability of DIN and DIP, which is considered to have a direct impact on the spring blooms increasing in productivity (Vahtera et al., 2007).

1.2.4 Nutrient loops connected to hypoxia

1.2.4.1 Phosphorous feedback loop

In 2002 Conley et al., showed that DIP water pools and its annual changes in the Baltic Sea had a *positive correlation* to the basin area which had hypoxic conditions in its bottom water. However, the changes in the DIP pool in the Baltic was not found to have a statistically significant correlation to the total phosphorus load (Conley et al., 2002).

1.2.4.2 Nitrogen feedback loop

The occurrence of nitrogen and hypoxia in the Baltic Sea water seem to exhibit a statistically significant negative relationship. This suggests that DIN losses due to the nitrogen fixing cyanobacteria increase in circumstances of hypoxic or anoxic seafloor environment. (Vahtera et al., 2007)

These findings indicated that the nitrogen pool in the Baltic Sea is controlled to a higher degree by its internal processes, than external loading such as nutrient input from agricultural practices. They found a statistically significant negative relationship between the pool of DIN and hypoxic water. Each processes' influence when it comes to controlling the nitrogen pool of the Baltic Sea and its size is not quantified, but the biological and physiological transformations of nutrients P and N caused by microbes is well known to be indirectly dependent on oxygen availability. In turn oxygen availability is controlled by hydrography and sedimentation of organic matter. Short term changes of the external load of nutrients does not have much effect on the P and N budgets in the pelagic of the Baltic Sea. Concluding that, one can say that cyanobacteria blooms are also controlled by the internal processes, which have been changed by long term external input of nutrients (Vahtera et al., 2007).

1.2.4.3 Redfield ratio

The average Redfield ratio for phytoplankton is C:106:N:16:P:1 (Redfield et al., 1963). A Redfield ratio of N to P of 16:1 is not a competitive advantage for the nitrogen fixing cyanobacteria that blooms in summer (Stal et al., 2003). The mean N:P ratio in the surface layer of the Baltic Sea was found to be around 7-9:1 in winter in the late 90's and its thought to be caused by denitrification (Nausch et al., 2008). In their studies, Stal et al. (2003) mentions having measured ratios as low as 1:1, N:P. From more current studies, the Nitrate is observed to be all used up in the water layer above the halocline after the spring bloom period has passed (Börgel et al., 2026)

1.3 Focus of this study

1.3.1 HAB's knowledge gaps

The current understanding of cyanobacteria blooms and their interaction with physical, chemical and biological factors is for the most part based on empirical field correlations and not on an understanding of the blooming physiological response to the environmental factors around it. This is what is used as input for current models used to model future and current bloom dynamics. These models are used for making decisions regarding implementation tactics used to create a healthier Baltic Sea ecosystem considering the anthropogenic stressors. When the environment changes, these correlations do not necessarily look the same (Munkes et al., 2021).

The relationship between the growth of cyanobacteria and excess P, & bloom intensities relation to nitrogen fixing is a reason for current models projecting cyanobacterial bloom differing from one another. The dynamics of cyanobacteria blooms in the Baltic Sea and the controls and interaction between these is still not understood well enough, which causes models used in the field to model future predictions of cyanobacteria blooms have a certain level of uncertainty connected to them (Munkes et al., 2021).

1.3.2 Objective of this thesis

The water surrounding the island of Gotland is known to have large areas of hypoxia and anoxia, a total lack of oxygen at the sea floor. Both the subbasins Eastern and Western Gotland Basins are considered to get progressively worse by the Swedish Meterological and Hydrological Institute, SMHI's Oxygen Survey (Hansson & Viktorsson, 2025).

The current eutrophication of the Baltic Sea spurs on the spread of hypoxia and permanently anoxic areas along the sea floor. The main instigator of this is the sedimentation of the spring bloom, which when broken down creates more internal loading of P, and whose use of nitrogen promotes N_2 fixing cyanobacteria to grow in the summer. The N_2 fixing cyanobacteria, which makes the reduction of external P load inadequate as a slowdown measurement on shorter time scales (Vahtera et al., 2007).

Funkey et al., showed that climate variability, anthropogenic activity and stratification have influenced hypoxia occurrence during the Holocene, and their results imply coupling between cyanobacteria blooms intensity and hypoxia existence during our current geological time. Understanding the lead lag and succession of hypoxia and cyanobacterial blooms is difficult (Funkey et al, 2014).

This study looks at correlations between cyanobacteria blooms and oxygen depleted sea floors. As well as cyanobacteria and its most well-known drivers and these drivers' correlation to and oxygen at sea floor, at a weekly and monthly temporal scale for 27 years, in hopes of understanding the interactions between important drivers of the cyanobacteria blooms and spread of hypoxia in the Baltic Sea over time.

1.4 Aim & Hypothesis

1.4.1 Research question

How will the correlation between hypoxia and cyanobacterial blooms look and change in the years of 1998 to 2024 in the Western Gotland Basin subbasin, in the Baltic Sea? How will the main drivers of cyanobacteria bloom, stated in section 1.1.1 and their correlation to hypoxia and cyanobacteria bloom change over the years 1998 to 2024 on a spatial and temporal scale? Variables considered were O_2 at the sea floor and occurrence of cyanobacteria bloom together with the biggest drivers of HAB's, Harmful Algal Blooms.

1.4.1.1 Research Questions

Research Questions 1: Is there coupling between cyanobacteria bloom and sea floor hypoxia? If so, at what timescale does the bloom × bottom-hypoxia coupling appear?

Research Questions 2: Which of the studied environmental drivers co-vary most strongly with bloom frequency and with bottom-water oxygen?

Research Questions 3: Is the coupling of cyanobacteria bloom and hypoxia spatially homogeneous across the Western Gotland Basin?

Research Questions 4: How has bloom frequency, bottom-water oxygen, and their drivers changed in the study area from 1998 to 2024?

Research Questions 5: Is the apparent lagged coupling a real biological signal, or an artefact of overlapping trends or wind-driven detection bias?

1.4.2 Aim

This thesis looks at the spatiotemporal relationships between detected cyanobacteria bloom and reanalysed amounts of dissolved oxygen at sea bottom level, over 27 years. The aim is to see if there is any correlation between them, indirect or direct, as well as looking at if there is a lagged correlation between them. Since the phenomenon of hypoxia and cyanobacteria blooms in the Baltic Sea are known to interact indirectly through the biological and chemical nutrient cycling, this thesis aims to map and identify the relationships between the drivers of the blooms and hypoxia, as well as the drivers and the blooming.

1.4.3 Hypothesis

Based on previous literature and the current scientific understanding of the problem at hand, a correlation between oxygen concentration at sea floor and cyanobacteria bloom occurrence is not expected. A positive correlation between SST and cyanobacteria bloom is expected, as well as between chlorophyll a and cyanobacteria. While a negative correlation between mixing depth and cyanobacteria is expected. An overall trend showing decrease in dissolved oxygen at the sea floor from 1998 to 2024 is expected to be found.

2 Materials & Methods

2.1 Materials

2.1.1 Datasets

The oceanographic and biogeochemical data was obtained from Copernicus Marine Data Store (CMEMS), while basin borders were acquired from HELCOM Map and Data service and bathymetry data was taken from National Centers for Environmental Information, NCEI a part of NOAA. For the CMEMS data, cyanobacteria were obtained from the dataset *Baltic Sea Multiyear Ocean Colour Plankton, Reflectances and Transparency L3 daily observations,* within 0 represent no occurrence, 1, subsurface occurrence, 2 surface occurrence and 3 equates to both subsurface and surface occurrence of cyanobacteria has been detected. Wind data was obtained from *Global Ocean Hourly Reprocessed Sea Surface Wind and Stress from Scatterometer and Mode*. Physical variables, mixed layer depth, surface salinity, sea surface temperature and sea floor temperature were obtained from *Baltic Sea Physics Reanalysis.* The data on oxygen, nutrients and chlorophyll a was obtained from *Baltic Sea Biogeochemistry Reanalysis* dataset. The subbasin, Western Gotland Basin, was extracted from the *HELCOM subbasin with coastal WDF waterbodies or watertypes 2022 level 4a.* Bathymetry for the area was collected from *NCEI's ETOPO 2022 30-arc-second global bedrock elevation product.* Datasets resolution and datatype can be seen in *table 1*, and technical specifics can be seen in *table 2*. DOI for datasets can be found in the reference section.

2.1.1.1 Tables

Table 1. Table displaying data from the same datasets which variable they were used to conjure for the analysis, their spatial resolution and how the data was originally constructed by their authors.

Variable Group	**Resolution (rasters)**	**Datatype**	**Temporal Extent**
Cyanobacteria	1 × 1 km	Reprocessed	4 September 1997 – 13 May 2026
Wind	0.25° × 0.25° & 0.125° × 0.125°	Observation corrected model	1 June 1994 - 21 January 2026
Mixed layer depth, surface salinity, surface T, bottom T	2 × 2 km	Reanalysis	1 January 1993 - 31 December 2024
Surface O_2, bottom O_2, Chl A, NO_3, NH_4, PO_4,	2 × 2 km	Reanalysis	1 January 1993 – 31 December 2024
Basin extent	Vector layer, Polygons.	Digital Map	2022
Bathymetry	30-arc-seconds	Composite DEM	2022

Table 2. Table displaying each variable separately with their specific name and unit from the Copernicus marine data store.

Variable	**Copernicus label name**	**Unit**	**Temporal resolution**
Cyanobacteria	“CYANOBLOOM”	[0, 1, 2, 3]	Daily
Northward wind	"northward_wind"	[m/s]	Hourly
Eastward wind	"eastward_wind"	[m/s]	Hourly
Mixed layer depth	“mlotst”	[m]	Daily
Surface salinity	”so”	[0.001] [g/kg]	Daily
Surface T	“thetao”	[°C]	Daily
Bottom T	“bottomT”	[°C]	Daily
Surface O_2	"o2"	[mmol/m^3]	Daily
Bottom O_2	"o2b"	[mmol/m^3]	Daily
Nitrate, NO_3	"no3"	[mmol/m^3]	Daily
Ammonium, NH_4	"nh4"	[mmol/m^3]	Daily
Phosphate, PO_4	"po4"	[mmol/m^3]	Daily
Chlorophyll A	"chl"	[mg/m^3]	Daily

2.1.2 Software

Compiling, downloading, aggregating and analysing datasets were done in Python 3.12.6 (Python Software Foundation 2024). The programming scripts were developed with the assistance of the AI language model Claude, version claude-sonnet-4-6 (Anthropic, 2025). The data processing performed in Python 3.12.6 was done using packages xarray 2026.4.0 (for working with NetCDF and regridding data), numpy 2.4.6 (for numerical computations), pandas 3.0.3 (framing resulting data and tables displaying time series), geopandas 1.1.3 (for computing the non NetCDF data), copernicusmarine 2.4.1 (used to download larger data files from the Copernicus marine data store), dask 2026.3.0 (to do parallel and chunked computing for large datasets), netcdf4 1.7.4 (to read and write NetCDF files), shapely 2.1.2 (for clipping polygons), pyproj 3.7.2 (to make transformations in coordinate reference systems), rioxarray 0.22.0, Cartopy 0.25.0, rasterio 1.5.0, scipy 1.17.0, pymannkendall 1.4.3, statsmodels 0.14.6, matplotlib 3.10.9, pyarrow 24.0.0. All the data processing was done via Python using uv 0.11.8 in the notebook environment Marimo 0.23.6 (Agrawal & Scolnick, 2023). The infrastructure and helper packages can be found in *Appendix 1*, which contains all main packages used for the processing of data in python.

2.2 Methods

2.2.1 Data collection, aggregation and harmonising

2.2.1.1 Literature Search

Relevant sources of literature were obtained through the search motors Finn by Lund University Web of Science and Google Scholar. Physical books used for literature research were found at the Geolibrary in Lund University, Sweden. Since the Baltic Sea is a well-documented body of water, only literature on the Baltic Sea was used. For physical books the latest available version was used.

2.2.1.2 Data collection

Many databases contain data regarding the state of the Baltic Sea. For this analysis the main source of data used was the Copernicus Marine Data Store. The data available there had a long temporal span, as well as it was already interpolated into raster data. That meant an already tested, scrutinized and appropriate interpolation of the data had been done. The time frame, January 1st, 1998, to December 1st, 2024, was chosen due to data availability for all the datasets.

Variables chosen depended on the main known drivers of cyanobacteria from literature. Cyanobacteria blooms are believed to be connected to eutrophication and oxygen depletion at the sea floor, but they are not directly correlated. The drivers that are known to have a big impact on the blooming is sea surface temperature, wind, mixing depth and nutrients. Most datasets were chosen based on availability, resolution and temporal coverage. Sea surface temperature was derived from the *Baltic Sea Physics Reanalysis* for spatial homogeneity data on the physical variables instead of the SST observational product from Copernicus marine data store.

2.2.1.3 Data harmonisation

Before starting the analysis, the data needed to be temporally and spatially harmonised. All datasets were downloaded in the finest temporal resolution available. Some data downloaded had a temporal scale of hours, while some had daily. To do statistical testing over such a long time period, this data needed to be aggregated into weekly and monthly datasets. The wind data had to be combined, since there where two differing datasets within the same product ID with different spatial resolution for different years. When aggregated, the wind data was merged to have the lower resolution of the two datasets, 0.25° × 0.25°. All datasets had to be cut to the basin and regridded to the extent and resolution of the cyanobacteria data cells, of 1 km x 1 km.

Downloading and aggregating the hourly and daily data was done with the toolbox for python, Copernicus marine (*Appendix 1),* using a python script provided by my supervisor, José Beltrán, and adjusted using Claude AI (Anthropic, 2025). The weekly and monthly data created were on the mean, minimum, maximum and standard deviation of each variables value for each raster cell per week and month. These were saved separately as geotiffs and all four values per variable combined as NetCDF files. For weekly data, each week was stored as 1 band in a raster, for all the weeks in the raster, and the same for months. These are referred to as "variable X" + "mean/min/max/std" in this report.

Due to the structure of the original dataset, cyanobacteria was aggregated from daily into different categories than all its driver variables. The cyanobacteria raster layer had values 0-3, with 0 meaning no observed occurrence of cyanobacteria, 1 observed subsurface occurrence, 2 observed surface occurrence and 3 both subsurface and surface occurrence of cyanobacteria had been observed. The weekly and monthly aggregated values from this dataset was instead minimum, maximum, mode, frequency & valid count. Minimum displayed lowest value observed in one cell of that raster during one of the time periods, max the highest. Valid count is the number of valid observances, days with no atmospheric interference. Frequency adds all

values above 0 for a week or month and divides it by the valid count of days. This creates a decimal number for each raster cell that represents the fraction of valid days with a bloom occurring.

When all data was aggregated into weekly and monthly dataset with min, max, mean, std & min, max, mode, freq & valid count, they were regridded to have the same grid size as the cyanobacteria bloom raster, 1x 1km, and cut to the spatial extent of the cyanobacteria data and the Western Gotland Basin polygon.

2.2.2 Analytical Methods

2.2.2.1 Basin wide

When doing basin wide correlations, each variables mean value across all cells in the raster gets averaged to a mean per band. For the cyanobacteria bloom the variable that got averaged over the whole basin was bloom frequence, and for bottom oxygen the variable used is minimum, chosen as an indicator of hypoxia. All other variables' basin wide average was calculated on their mean value.

2.2.2.2 Pixel by Pixel

Pixel wise Spearman correlation was calculated on the same values of variables as for the basin wide. The pixel-by-pixel correlations were only done on monthly data and during the months of June, July and August (JJA), when cyanobacteria is known to bloom. Any pixels that had less than ≈ 55 % valid monthly observations of their JJA data were completely excluded.

The datasets on nitrogen, NO_3 + NH_4 were combined to create the data for dissolved inorganic nitrogen, DIN. DIN was used to create a variable called N:P ratio, also used for pixel by pixel cross correlation and Mann-Kendall Trend analysis.

For pixel wise Spearman correlation, p values were taken from the t-statistic approximation. Since a lot of data is processed the probability of false discoveries is increased, specifically type 1 errors. To minimise and control the number of false discoveries the Benjamini-Hochberg procedure, also called False Discovery Rate procedure was applied. It was used to adapt p-values to decrease type 1 errors, at $\alpha = 0.05$ for each driver variable on the valid pixels.

A sensitivity analysis was done on the pixel wise correlations by leaving data from one year out, computed once for every year to confirm no specific years are drivers for the result, and cause for anomalies, visible in *Appendix 8*.

2.2.2.3 Depth Split

The study area has both shallow and deep water. Knowing the thermocline and halocline play a big part in the ability for vertical mixing in the stratified layers, results were split into deep basin areas, where the depth of the sea floor were below 80 m, the halocline depth threshold used in this study, and sea bottom depth of less than 80 m. The areas are called deep basin and shallow shelf in this report. This split was done using bathymetry data from NCEI. The Bathymetry data was interpolated into the grid of the cyanobacteria layer. The deep basin covers ≈ 14 500 km² of the Western Gotland Basin while the shallow shelf covered ≈ 18 700 km².

2.2.3 Statistical Analysis

All *interannual* analyses are done purely on the months of June, July and August each year. This is because JJA is the bloom season for cyanobacteria, and to get as representable correlations as possible of reality, months known to have seasonal environment that promotes cyanobacteria growth were used. This decision also reduces changes in variables and their correlations that are caused by seasonal variability, since that is not the focus of this study. A total of 81 months per variable (assuming all data is valid) are computed, covering the 27 years, 1998 – 2024.

2.2.3.1 Spearman correlation

Spearman correlation is used to study the correlation between two variables, especially when the relationship is expected to be nonlinear. It is used to understand both direction and strength of the correlation between the two variables, which is why it was used for this analysis. The Spearman rank correlation coefficient looks at the uniformity relationship between two variables. It computes a Pearson correlation on the two variables' data ranking, and not the raw values. This makes the Spearman correlation coefficient more resistant to outliers (Feng et al., 2026). Since the data used for this report is reanalysed, reprocessed or created from an observation corrected model and harmonized, that makes the Spearman correlation appropriate to use. In this study, Spearman correlations were used to create Pairwise Spearman correlations, Partial Spearman correlations and Pixel wise Spearman correlations.

Pairwise Spearman correlation was used to calculate a pairwise rank correlation matrices, one with all months in a year visible in *Appendix 2*, and one with June, July & August, to demonstrate the difference in results depending on choice of months and justify the use of JJA in this thesis. The main result used is the JJA matrix. It was also used to create heat maps, seen in *Appendix 6* and *Appendix 7*.

Partial Spearman was used to look at the correlation between cyanobacteria bloom frequency and volume of dissolved oxygen at the sea floor, but without the influence of the surface conditions dependences shared between them. This is done by using ordinary least squares to reduce the influence of variables that when ranked, cyanobacteria bloom and bottom oxygen have most dependency to. After doing ordinary least squares, the output is then correlated via a Pearson correlation. For comparison, a pairwise correlation (no reduced dependencies) were created, and one where the dependency on SST and mixed layer depth was reduced, as well as one where dependency on SST, mixed layer depth and wind speed was reduced. To assure Spearman was appropriate in this case, a scatterplot was created for visual check of monotonic relationships were made, visible in *Appendix 5*.

2.2.3.2 Mann-Kendall & Sen's slope

The Mann-Kendall test is used to look at data over time, specifically to analyse increase and decreasing trends. For this analysis as before, only the months of June, July and August were used. Instead of having 81 values (for 81 months) the data was collapsed into one value per year representing the mean of that variable over JJA each year. As previously, cyanobacteria bloom is collapsed to the mean of bloom frequency, bottom oxygen is collapsed to the mean of its minimum values and all other variables area means of their means. Mann-Kendall reports a trend direction and a p value, but to estimate the weight of the trend direction and p value the Sen slope was calculated. It was changed from per year to per decade values, because the combined yearly data makes monthly p values trivial, since each year has been turned from JJA into one observation/year.

2.2.3.3 Cross Correlation

To assess possible temporal lag between cyanobacteria blooms and hypoxia at the sea floor, a Spearman cross correlation was done with time lags of -12 to +12 months, on the continuous monthly time series. The lagged correlations used all months in the year, not only June, July and August, as for most other analysises. This was done on both raw data and detrended data to compare results. Due to the nature of the data, and previous analyses done, both bloom frequency and bottom oxygen have long term trends. To prevent trend overlap the data was detrended, but the raw data is kept for visual comparison.

3 Results

3.1 Workflow introduction

The different analyses explained in the method section all relate to different parts of the research questions. Therefore, the results will start with a short description of the progression pinpointing what research questions the analytical tests relate to.

3.1.1 Analytical workflow and research questions

Spearman cross correlations are done to answer research question one and determine whether sea floor hypoxia and cyanobacteria bloom frequency exhibit any statistical coupling. This is displayed in the cross-correlation matrix in *figure 4*, lead lag cross correlation, *table 8* and *table 10*. The variable cross-correlation in *figure 4* also works to answer the second research question, with important findings highlighted in *figure 4* and *table 3*.

To investigate homogeneity coupling of cyanobacteria bloom frequency and sea floor hypoxia in the studied basin, the split at the halocline was done. This is of importance for the third research question. The same Spearman cross correlation as previous was then performed, but on the two different areas of the studied basin to answer the third research question. The result is displayed in *table 9* and *table 10*.

Whether there are any trends in the change of the main variables over the time span of 27 years was investigated to answer the fourth research question using the Mann-Kendall trend and Sen's slope. This result is displayed in *figure 5* with the statistically significant values displayed in *table 6*.

The fifth and last research question aimed to see if the lagged coupling between cyanobacteria bloom frequency and sea floor hypoxia could be due to wind driven detection data, overlapping trends creating an artifact or possibly a real biological signal. To understand this, Partial Spearman Pairwise rank correlations were conducted, to compare resulting correlations and see if the relationships remained or changed dependant on removal of SST and mixing layer depth or wind. This can be seen in *table 4* and *table 5*.

3.2 Weekly variability 1998 to 2024

3.2.1 The average seasonal change in mean and variability over 27 years

3.2.1.1 Figures

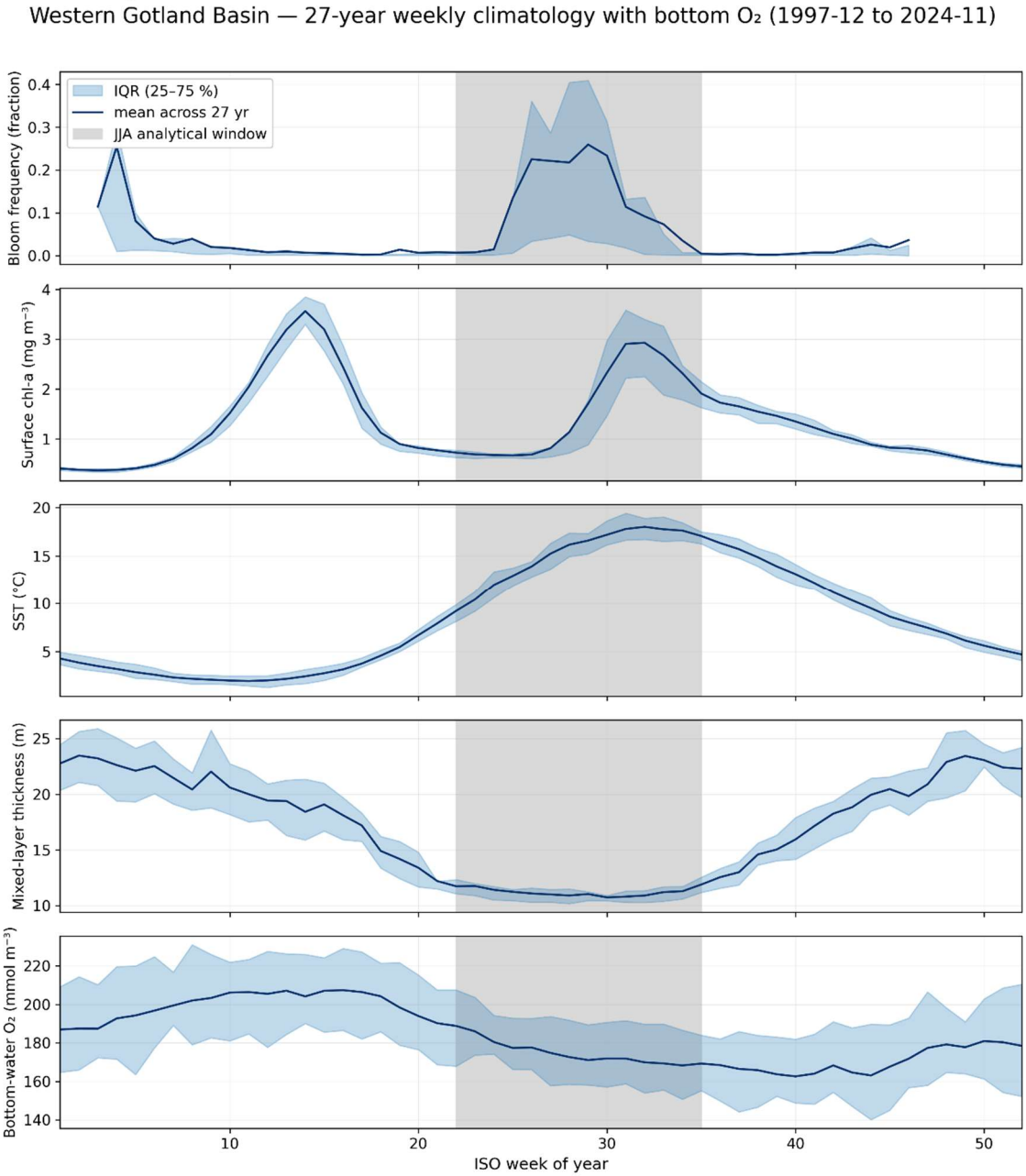


Figure 2. Average seasonal variability over 27 years, 1998 to 2024 for cyanobacteria bloom, chlorophyll a, SST, mixed layer depth & dissolved oxygen at the sea floor in the Western Gotland Basin in the Baltic Sea. The average Interquartile range, IQR, shows values between the 25th and 75th quartile, displaying variability from the mean during different weeks/seasons in a year. Grey highlights the months of June, July and August. This is based on basin wide average of the weekly data for all the years of 1998 – 2024.

In *figure 2* above, the mean is signified by the dark blue line. Variation in the mean is observed throughout the weeks and seasons for all variables. A larger light blue Interquartile range, IQR, indicates larger variability for the throughout the years.

The first graph in *figure 2* is on weekly bloom frequency data averaged over the subbasin of Gotland Western Basin. For all 27 years, the basin-wide average has been averaged to display the average bloom frequency for e.g. week 10 over the years 1998 to 2024. This graph has two notable peaks. One in early spring/late winter and one during summer. These are also the periods with the highest variability in the graph and so the biggest range in IQR.

The second graph, displayed in *figure 2* is on chlorophyll a at the sea surface. It displays two peaks, in spring and summer. It peaks 1 – 5 weeks after the graph for cyanobacteria bloom displays its peak in the summer season. The spring blooms are not recorded in the data used for the graph above this. The spring peak starts around week 7 and drops before week 20, while maintain a larger plateau base value during the summer than winter. The come down from the peak of chlorophyll a in summer is much less abrupt and linear than the come down from the spring bloom. This graph has a relatively stable IQR interval, but where variability seems most common is during both peaks, with a larger interval over the second peak.

The third graph is the seasonal change in sea surface temperature. The variability is small and stable but slightly larger during summer and autumn. The mean goes from approximately 2-3 °C in and around week 5 – 15, to 18°C between week 30 – 34 during late summer.

The fourth graph is mixed layer depth. It peaks during late autumn, winter and early spring. During said time periods the mixed layer depth displayed both a higher variability and mean. The weeks during June, July and August display the lowest mean for mixed layer depth and variability, at slightly above 10 meters depth.

The last graph in *figure 2* is the mean of minimum oxygen values in the basin. Out of all graphs, this variable shows the most consistently strong variability. The mean curves to display a slight peak during spring, decrease over summer and show a slight increase again very late in the year, in early winter.

3.2.2 Cyanobacteria bloom and Sea floor oxygen occurrence over 27 years

3.2.2.1 Figures

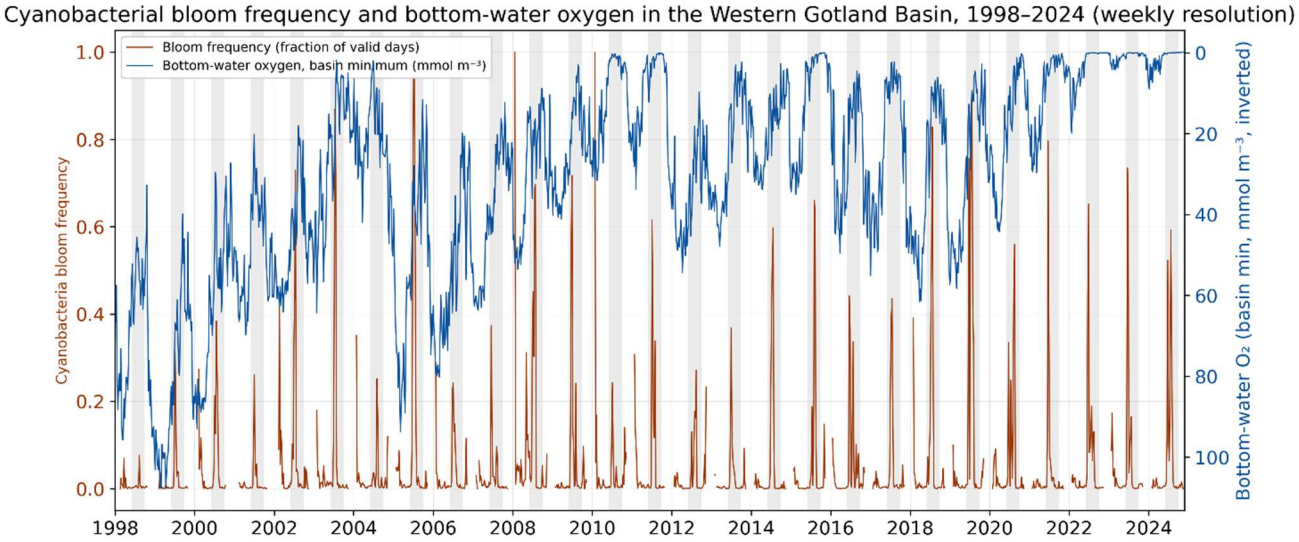


Figure 3. Weekly basin wide averages values of the variable bloom frequency and minimum dissolved oxygen at the seafloor (mmol m^{-3}) in the Western Gotland Basin, plotted over the temporal scale of 1998 to 2024. Grey bands highlight the summer months of June, July and August for each year.

In *figure 3* blooms are seen clearly as the orange line peaking. It is mainly visible inside the grey band displaying the summer months. Some summers with larger peaks are 2005, 2007 and 2019. Whenever the orange line breaks it means there is no data. Interestingly large peaks in data can be observed outside of the summer months highlighted in grey. Such as in the beginning of the year in 2008 and 2010. Cyanobacteria bloom frequency displays the mean of the bloom frequency across the full basin for each week. The frequency is the days above 0 divided with valid days (where data is not covered by clouds). A value of 1 indicates that there is bloom detected 100% of the valid observation days a week (no clouds). Since these values

are averaged over the full basin a value closer to 1 for the bloom frequency in *figure 3,* does not equate harmful algal bloom but indicates cyanobacteria presence in majority of the basin.

The basin mean value of weekly minimum oxygen is displayed as a blue line. It is inverted on the y- axis to bloom frequency. The higher the blue line reaches, the closer the mean of the minimum value is to 0 mmol m^{-3} in the basin. There are some different trends to be observed in this graph. Observing the seasonal variation, there is a dip in oxygen during almost all grey highlights, June July and August. As well as an increase in oxygen volume over the winter months. There are some trends on the larger temporal scale. From 1998 to 2004 there is a general decrease in the amounts of oxygen, followed by a large peak in 2005. From 2005 there is once again a general decrease until 2012, where it once again displays a peak in value, but smaller than the peak in 2005. After this, both decrease and increase in oxygen show a more incremental change. With a decrease from 2012 to 2016, and an increase from 2016 to 2018 instead of a large spike as in 2005 and 2012 it increases over two years. From 2018 it continually decreases until 2024, with the lowest mean in 2023.

3.3 Spearman correlation on monthly data for JJA over 27 years

3.3.1 Co-varying variables

3.3.1.1 *Correlation matrix*

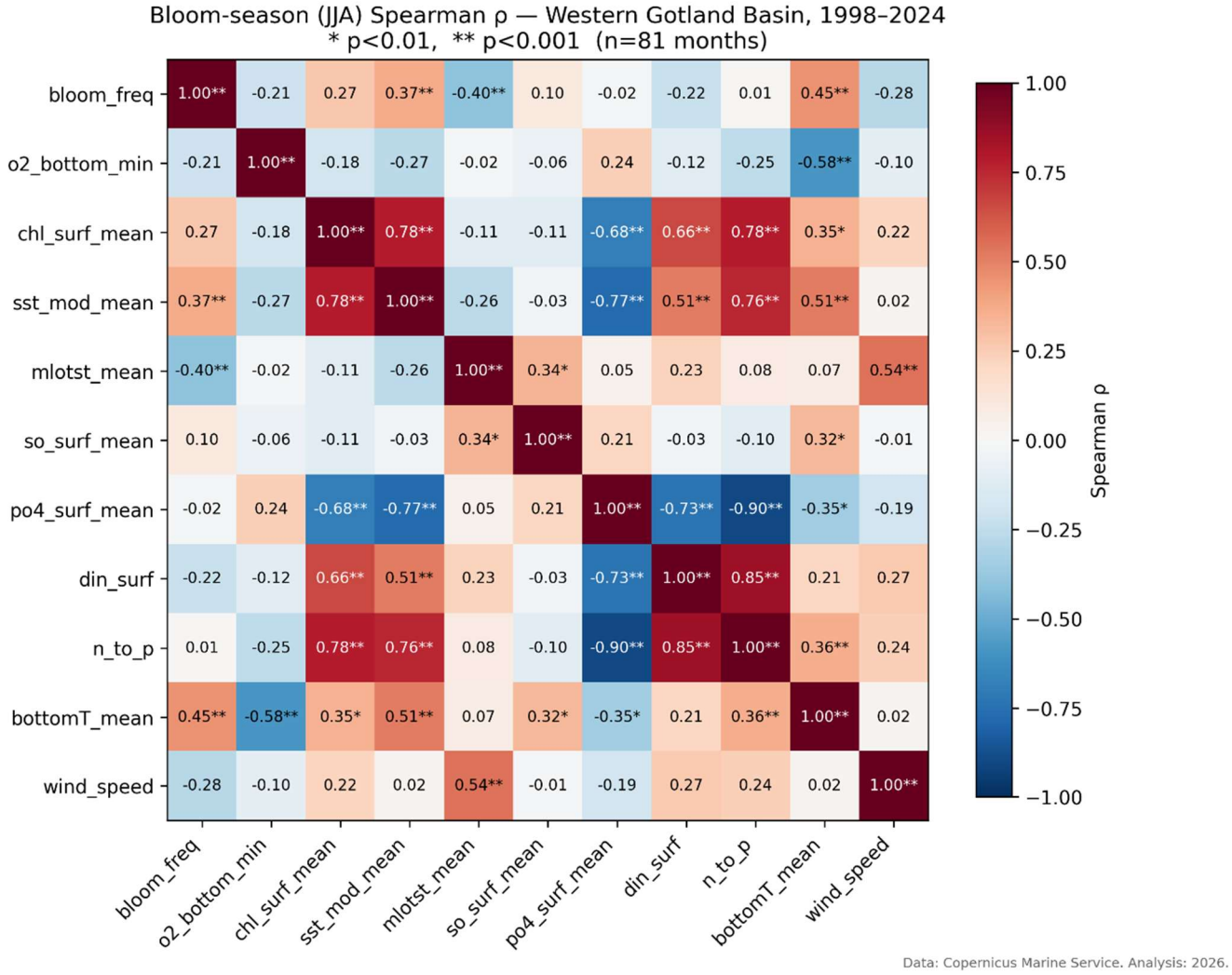


*Figure 4. Spearman correlation ρ values calculated on the basin wide average of the months of June, July & August from 1998 to 2024, presented in a correlation matrix. The matrix is created using the monthly basin wide means for all variables except cyanobacteria which uses the basin wide mean of bloom frequency and bottom oxygen which uses a basin wide mean calculated on the minimum values for each cell. Red indicates a positive correlation and blue a negative. Stars, * indicate statistically significant values, one star * means a p value < 0.01 and two stars, ** indicates a p value <0.001.*

This correlation matrix presents basin wide monthly correlations for the summer months, June, July and August from 1998 to 2024. In the appendix one other correlation matrix can be seen, displaying the monthly basin wide correlation for the same variables, but using all months in the year and not only the summer months, see *Appendix 2*.

Bloom frequency shows statistically significant correlation to SST, mixed layer depth and sea floor temperature. While oxygen levels at the sea floor only had a statistically significant correlation to sea floor temperature. All nutrients show statistically significant correlation to each other. The Spearman correlation between all variables provides an overview, but correlation does not mean causation. All the strongest co-varying variables can be seen in *table 3* below. Contrary to what was stated in the hypothesis, chlorophyll a and bloom frequency does not have a statistically significant correlation.

Table 3. The strongest co-varying variable pairs from Spearman correlation matrix. All correlations have a p value <0.001 and are statistically significant. Based on the data in figure 4.

Co-varying variabels	**Spearman ρ**	**p-value**	**p value <0.001**
PO_4 mean ↔ N:P ratio	−0.902	1.49e-30	**
DIN ↔ N:P ratio	+0.848	1.55e-23	**
Chl-a mean ↔ N:P ratio	+0.782	6.77e-18	**
Chl-a mean ↔ SST mean	+0.782	6.91e-18	**
SST mean ↔ PO_4 mean	−0.770	4.73e-17	**
SST mean ↔ N:P ratio	+0.758	2.73e-16	**
PO_4 mean ↔ DIN	−0.730	1.13e-14	**
Chl-a mean ↔ PO_4 mean	−0.680	2.96e-12	**

The bloom frequency and minimum oxygen at sea floor did not have a statistically significant correlation. In *table 3,* displaying the strongest co-varying correlations found, the strongest were a negative correlation of−0.902 between PO_4 and the N to P ratio and the strongest positive where a correlation of +0.848 between DIN and the N to P ratio. The strongest co-varying correlation displayed in *table 3* that is not between two nutrients variables is the correlations between chlorophyll a, and the sea surface temperature.

3.3.2 Pairwise Partial Spearman rank correlation

3.3.2.1 Tables

Table 4. Partial Spearman rank correlation of cyanobacteria bloom frequency and minimum oxygen at sea floor in the Western Gotland Basin. The correlation is done over the summer bloom season months of June, July and August from 1998 to 2024, n = 81 months. One pairwise correlation between the two variables, bloom frequency and minimum oxygen was computed, and two different correlations where controls for the main shared driver variables influence was removed. Variables whose influence was removed was SST, mixed layer depth (MLD) and wind speed.

Model	**Spearman ρ**	**p-value**
Pairwise	−0.207	0.064
Controlling for SST + MLD	−0.162	0.148
Controlling for SST + MLD + wind speed	−0.172	0.124

For the Pairwise Partial Spearman correlation between cyanobacteria bloom frequency and minimum sea floor oxygen, the regular test was not statistically significant as seen in *table 4*. When removing possible influence from SST and mixed layer depth it remained statistically insignificant. When removing wind along with SST and mixed layer it remained statistically insignificant.

Table 5. Spearman rank correlation for wind detectability check between bloom frequency and wind speed in Western Gotland Basin. Done on the months of June, July and August from 1998 to 2024, n = 81 months.

Model	Spearman ρ	p-value
Pairwise	−0.279	0.012
Controlling for SST + MLD	−0.153	0.173

The results from the wind detectability test are seen in *table 5*. Cyanobacteria bloom frequency was pairwise Spearman rank correlated to wind speed and showed no statistical significance. When controlling for influence from SST and mixed layer depth the values remained statistically insignificant. This indicate that the bloom signal is not observational bias.

3.4 Mann-Kendall trends and Sen's slope

3.4.1 Mann-Kendall trends and Sen's slope

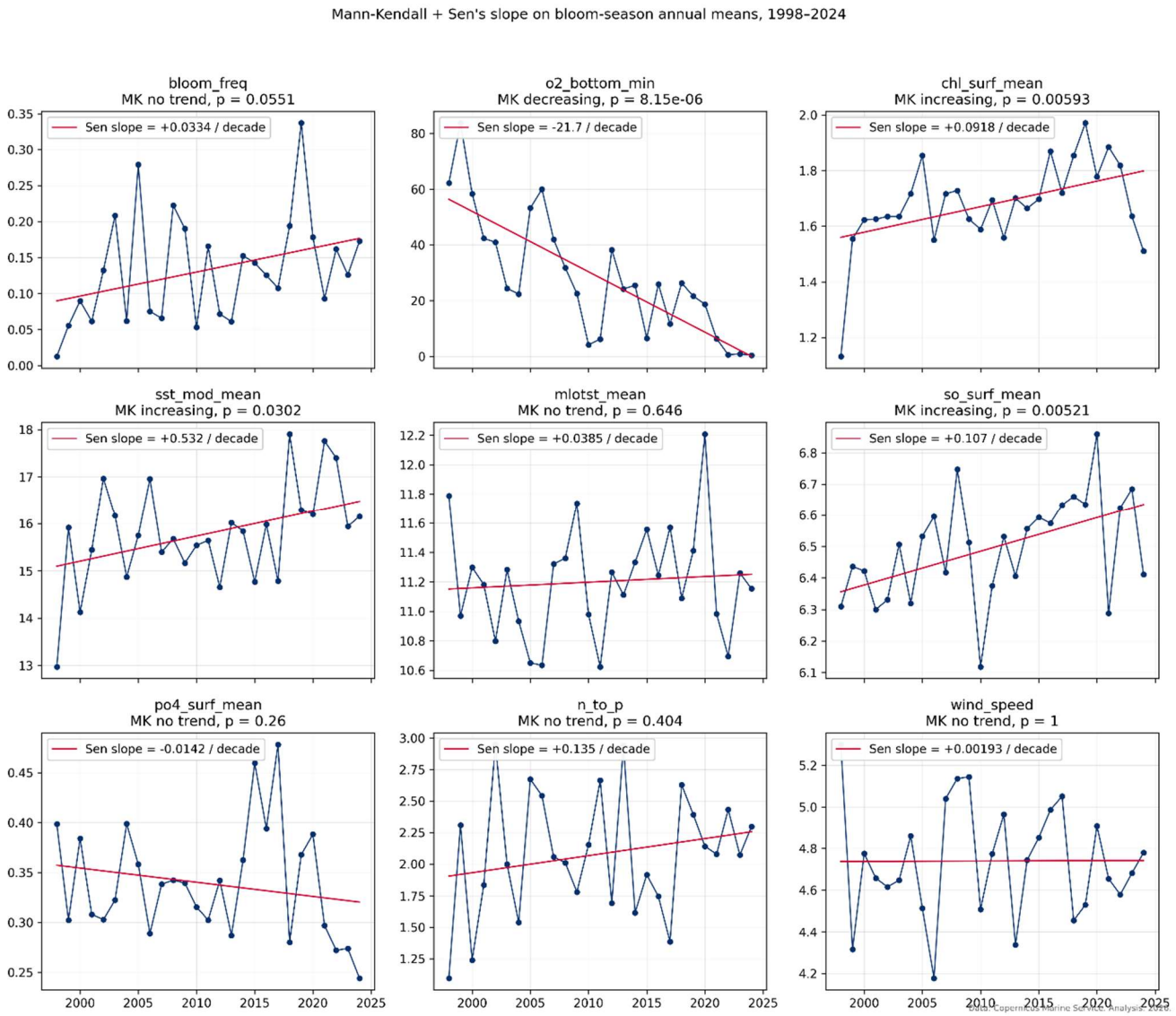


Figure 5. This figure displays the Mann-Kendall trend for bloom frequency, minimum bottom oxygen, mean chlorophyll a, SST, mixing depth, salinity, phosphorous, nitrogen to phosphorous ratio and windspeed. The Sen slope trend line is visible in red in each individual graph.

The Mann-Kendall trend with Sen's slope in *figure 5* above displays 9 variables. Wind speed, N to P ratio, phosphorous, mixed layer depth and bloom frequency display no statistically significant trend, although the bloom frequency borders on being statistically significant. Chlorophyll a, SST, and salinity display a positive trend. Sea floor oxygen is the only negative trend displayed, and it is the strongest trend visible. In *table 6* below all statistically significant variables, their p-value and decadal decrease in value is displayed.

Table 6. Statistically significant Mann-Kendall trends for the annual mean of the bloom season, June, July and August in Western Gotland Basin from 1998 to 2024. n = 27 years. Variables are Oxygen at sea floor, chlorophyll a, SST and salinity. Sen's slope is expressed as per decade, and only the variables who were found significant at $p < 0.05$ are shown.*

Variable	Trend	p-value	Sen's slope per decade
O_2 min	Decreasing	8.15e-06	−21.7 mmol m^{-3}
Chl-a mean	Increasing	0.006	+0.092 mg m^{-3}
SST mean	Increasing	0.030	+0.532 °C
Salinity mean	Increasing	0.005	+0.107

In *table 6* the four statistically significant Mann-Kendall trends are displayed. The most significant is the decrease in oxygen at sea floor over the years. The Sen slope value for the decrease in oxygen at the sea floor is -21.7 mmol/m^3. The three other statistically significant values are *chlorophyll a, SST* and *salinity* at the sea surface. They are all showing an increase per decade. Sea surface temperature shows an increase of +0.532 °C per decade, while chlorophyll showed +0.092 mg m^{-3} and salinity +0.107 practical salinity unit, psu, per decade.

3.4.1.1 Depth split

Table 7. The Western Gotland Basin, split at 80 m depth, the average halocline depth, and Mann-Kendall trend is performed on the sea floor oxygen minimum for the annual mean of June, July and August from 1998 to 2024, n = 27 years.

Region	Trend	p-value	Sen's slope per decade
Deep >80 m	Decreasing	2.11e-05	−21.4 mmol m^{-3}
Shelf , ≤80 m	Decreasing	1.07e-06	−29.2 mmol m^{-3}

For comparison of different areas within the subbasin Western Gotland Basin, the area was divided into areas with a depth below the halocline, >80 m, the deeper water, and above, <80 m, the shelf or shallow water. The Mann-Kendall trend was performed on annually averaged data on the sea floor oxygen for the months of June, July and August for each area, resulting in the values presented in *table 7*. Both the shelf and the deep areas had a statistically significant trend value, both were decreasing but their Sen slopes differed. The shallower area displays a larger decrease per decade of 29.2 mmol m^{-3}, while the deep had a decrease of 21.4 mmol m^{-3}. In *table 6*, which displays the Sen slope for the averaged minimum oxygen at the sea floor for the whole basin, the average for the whole basin is −21.7 mmol m^{-3}, which is closer to the values of the deep basin areas than the shallow.

3.5 Lagged correlations

3.5.1 Cross correlation

Table 8. Bloom frequency and minimum oxygen at sea floor peak cross correlation, between 1998 to 2024 during bloom season, June, July and August for Western Gotland Basin. The table displays both the raw data and linearly detrended data.

Series	Peak lag (months)	Peak Spearman ρ
Raw	+4	−0.386
Linearly detrended	+7	−0.256

Table 8 above shows the raw data when cross correlated showing a peak lag at +4 months and the linearly detrended data shows a peak lag at +7 months. *Appendix 3* and *Appendix 4* displays graphs with the same data. There is also a shift in Spearman ρ, with a lower value for the detrended dataset.

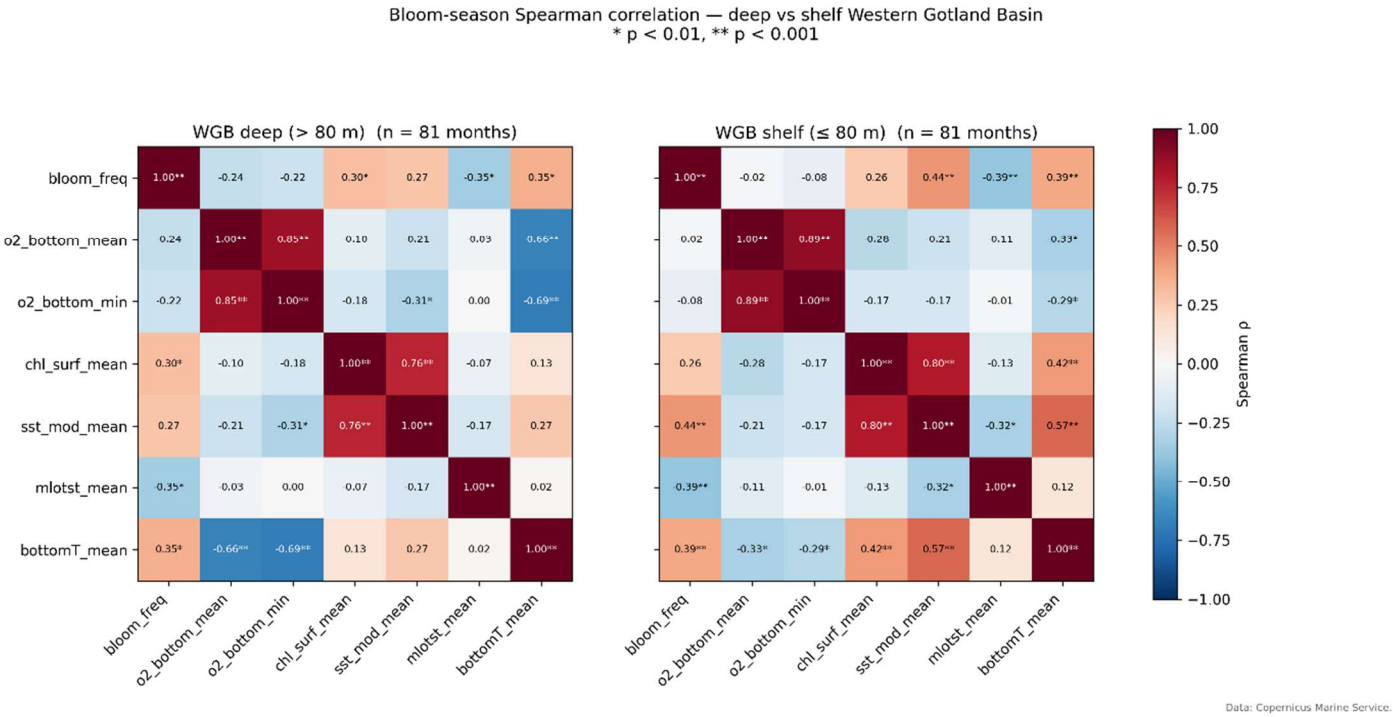


Figure 6. Spearman correlation done on basin wide average of monthly data, for deep and shallow areas. One matrix only shows the correlations for areas where the depth of the Baltic Sea is larger than 80m, and the other one shows the averages for when the depth is less than 80m. Data is on the summer months of June, July and August.

In *figure 6*, the Spearman correlation on all variables is done on different bottom depths. The Western Gotland Basin was split into two areas, the shallow shelf, and the deep basin. The split was done at 80, by the halocline. The values for each variable in cells in the deep basin were correlated to the other variables in the deep basin, and same for the shelf areas. The Spearman correlation was done on the monthly mean data of bloom season, June, July and August from 1998 to 2024. The basin wide averages were made on the mean values for each variable except for two, sea floor oxygen was averaged on its mean and minimum value and cyanobacteria bloom was averaged on its frequency.

The correlations appear different between the deep basin and the shallow shelf. In the deep basin the sea floor temperature and oxygen display a stronger correlation than it does in the shelf. As for correlation between bloom frequency and sea floor oxygen, neither the averaged mean or minimum vales were statistically significant in either the deep basin or the shelf.

However, there is a large difference between the Spearman ρ values between the deep basin and the shelf. Neither is statistically significant, but in the deep basin the Spearman ρ values are a lot bigger than for the shelf. Another interesting difference is that for the deep basin, the correlation between bloom frequency and SST was not statistically significant, but for the shelf it was, with a $p < 0.001$. In both areas the mixed layer depth had a statistically significant correlation to the bloom frequency.

3.5.1.1 Depth split

Table 9. Spearman rank correlation on the Western Gotland Basin, split at 80 m depth, the average depth for the halocline, for the sea floor oxygen min and bloom frequency, during bloom season, June, July and August from 1998 to 2024, n = 81 months.

Sub-region	**Spearman ρ**	**p-value**
Deep >80 m	−0.215	0.054
Shelf ≤80 m	−0.083	0.462
Δρ (Deep − Shelf)	+0.132	—

In *table 9* the Spearman rank correlation for the depth split is presented. Deep areas are ones below 80 m depth, and shelf or shallow areas have their sea bottom at or above 80 m depth. Both have negative Spearman values.

Table 10. Cross correlation done on deep and shallow areas in the Western Gotland Basin on bloom frequency and sea floor oxygen min during June, July and August 1998 to 2024. The table displays the peak lag values for both areas.

Sub-region	**Peak lag (months)**	**Peak Spearman ρ**
Deep >80 m	+7	−0.410
Shelf ≤80 m	+6	−0.318

In *table 10* above, the peak lag is presented. The areas of the Western Gotland Basin which had a sea bottom deeper than 80 m had a peak lag at + 7 months, and the shallower areas had a peak lag at +6 months.

4 Discussion

4.1 Lagged bloom – oxygen signal

4.1.1 Cross correlation of hypoxia and blooms

In *figure 4*, the correlation between all variables can be observed. Notably, looking at the correlation between the minimum oxygen and cyanobacteria bloom frequency, there is no statistical correlation. The hypothesis stated that no correlation between the two variables were expected, as on smaller timescales previous literature links oxygen availability in the water column above and below the halocline to atmospheric forcing (Börgel et al., 2026). This also aligns with how the relationship is thought to indirectly interact as parts of the internal feedback loop from eutrophication to oxygen depletion (Vahtera et al, 2007).

In *table 8*, the cross correlation peak lag and ρ values between sea floor oxygen and bloom frequency with time lags of + 12 to – 12 months is shown. For the raw data, the peak ρ showed at 4 months. When the data was detrended, peak ρ persisted which can be interpreted as indicating a genuine lagged correlation between the datasets and not trend overlap. The Spearman ρ indicates strength and direction of the relationship. In *table 8* it is visible that the raw data exhibits a stronger relationship between variables at −0.386 vs detrended data at −0.256. The relationship remained negative, but the peak lag was at 7 months instead of 4 for the detrended data.

In *table 4*, and *table 5,* the values from the Pairwise Partial Spearman rank correlation is stated. In *table 4*, neither of the correlations exhibited statistical significance. In other words, when removing shared driver variables influence, the correlation between cyanobacteria bloom frequency and sea floor hypoxia remained negative and not statistically significant. This in turn can be interpreted as the shared drivers of SST, mixed layer depth and wind speed are not the cause of this relationship, removing that uncertainty factor. The Pairwise Spearman correlation in *table 5* was executed to check for wind detectability. The bloom x wind pairwise correlation is not statistically significant after controlling for SST and mixed layer depth influence. This also supports the lagged correlation results, by removing one uncertainty factor and displaying that wind is not the cause.

It is likely this mechanism is visible as a multi-month lag and not as an immediate, synchronous correlation because of the cycle of cyanobacteria dying, sinking and getting consumed which is not an immediate process (Vahtera et al, 2007). For future studies, it would be interesting to look at the cross correlation of the variables at longer time lags than 1 year prior and after.

There are limitations to be aware of when interpreting the result. The data used for sea floor hypoxia is reanalyzed, actual long-term measurements of the sea floor oxygen levels might show different results. The data used is calculated based on in situ measurements, but there is no adequate long-term dataset containing enough datapoints to use only in situ measurements. As for the cyanobacteria, the dataset itself is based on reprocessed satellite imagery. Days with atmospheric disturbances are not included in the cyanobacteria frequency data. The data is then averaged over the full basin and over months. This all creates a broad estimation of the circumstances to begin with and is a limitation, throughout most calculations done on basin wide monthly averages.

4.1.2 Variables co-varying with hypoxia and blooms

Figure 4 displays a cross-correlation matrix, in which all the variables co-variance to each other, hypoxia and bloom can be seen. It is interesting to note some known drivers that did not show

a correlation to neither bloom frequency nor sea floor oxygen. None of the nutrient variables showed a statistically significant correlation to either and neither did wind speed. In the case of cyanobacteria bloom, lacking a correlation to wind makes logical sense, since strong winds create an unfavourable environment for blooms to appear.

4.1.2.1 Bloom Frequency, temperature and mixed layer depth

Going back to *figure 4*, it can also be used to discern which variables that the bloom frequency had its strongest correlation to and where statistical significance could be noted. The cyanobacteria bloom frequency had statistically significant correlations to three variables. Sea floor mean temperature, with the strongest relationship, ρ 0.45, the mixed layer depth, the only negative relationship out of the three, ρ -0.4, and finally sea surface temperature, to which it had a positive relationship with Spearman ρ at 0.37. All three correlations had $p < 0.001$. Cyanobacteria are known to thrive in shallow mixed water and with higher temperatures which supports these findings. The sea floor temperature can be due to a general increase in temperature of the water body during the summer, which is when cyanobacteria bloom.

When interpreting these correlations, it is important to keep in mind that Spearman ρ does not measure linear correlation and therefor it will not show a unimodal relationship. Spearman correlation measures monotonic relationships, meaning that it recognises a consistent direction for the relationship. If variable y decreases and variable x increases it will recognise the relationship but the increase/decrease in the variables does not need to be linear, only directionally consistent. Since natural phenomenon like algae blooms are impacted by so many variables and the environment, they exist within always differ, it is a useful statistic. But it does leave some to be desired, since it cannot detect preferable ranges of variables for the bloom, specifically when looking at sea surface temperature.

A limitation that is important to mention when interpreting the results displayed in *figure 4*, is the fact that the p values likely overestimate the significance of the correlation between variables. The correlation values in *figure 4* and *table 3* are calculated on data aggregated to a monthly average for each pixel, which in turn has been averaged over the whole Western Gotland Basin to represent 1 month. The data is on three consecutive months and are therefore not independent from one another. The climate in June does affect the environmental factors in July and so forth. For that reason, $p < 0.01$ are considered suggestive, with $p < 0.001$ being the conservative significance, for all values in *figure 4* and *table 3*.

4.1.2.2 Bottom Oxygen and sea floor temperature

As for the sea floor oxygen, only one variable was found to have a statistically significant correlation, the sea floor temperature. This was statistically significant with $p < 0.001$, meaning it does pass our conservative threshold. The relationship is negative, with a Spearman ρ of -0.58. The relationship of this correlation can also be observed in *figure 6*, that divides the basin into depths above or below the halocline. In *figure 6*, both the shallow shelf and deep basin display correlations between oxygen minimum/oxygen mean and the temperature at the sea floor. The deeper basin does have a stronger relationship and higher statistical significance than the shallow shelf, whose p value < 0.01.

There could be many reasons as to why the sea floor oxygen and temperature correlates. The most likely explanation is due to the seasonal variability. Oxygen solvability in water is decreased with temperature increase and vice versa (Andersson et al, 2015).

4.2 Spatial homogeneity of bloom and hypoxia

4.2.1 Depth split differences

4.2.1.1 Bottom Oxygen and Bloom frequency

The results so far have shown no synchronous correlation between hypoxia and bloom frequence. The Spearman ρ values consistent direction after being detrended visible in *table 8* indicate a lagged correlative relationship between the bloom and hypoxia. Looking at the lagged correlative relationship between hypoxia and bloom, in *table 10* the difference between the deep basin, whose peak lag is at +7 months with a ρ of – 0.410, and the shallower shelf which displayed a peak lag at +6 months and slightly smaller ρ of - 0.318. The depth split can be seen as an image in *Appendix 9* and the difference between the deep basin and shallow shelf in *Appendix 10*.

4.2.1.2 Co-varying in shallow and deep water

From *figure 6* the difference in co-varying variables in basin with depth reaching above or below the halocline (80m) makes it deductible that neither the coupling of bloom and hypoxia, nor bloom and its drivers or hypoxia and the variables sustain a homogenous relationship over the Western Gotland Basin. It becomes quite evident that sea floor oxygen and bloom frequency interact differently with each other and other variables depending on the basin depth. This can be observed in *table 9* as well.

The overarching impression that is given by observing the combined results is that the coupling between hypoxia and cyanobacteria bloom is stronger in the deeper areas of the basin than on the shallow shelf. The results themselves are descriptive, where the coupling is strongest can be interpreted, but there is no casual mechanism indicating why those are results. It is known that the Baltic Sea has a low ventilation of its bottom layers of water, specifically below the halocline due to the patterns of inflow of saline waters from the North Sea through the Danish straits reinforcing the thermocline and the halocline (Lass & Matthäus, 2008). The Western Gotland Basin is located relatively close to these inflows, despite of that in recent years when larger inflows of oxygenated water have been recorded, it has not been able to reach any of the deeper parts in the basins surrounding Gotland (Hansson & Viktorsson, 2025).

4.3 1998 to 2024 – What has changed?

4.3.1 Mann-Kendall trends and Sen's slope's

4.3.1.1 Changes in Sea floor oxygen from 1998 to 2024

To examine if any long-term trends of change were present over the studied years, the statistical Mann-Kendall trend test was done on each variable. Resulting graphs can be seen in *figure 5*. There were 4 variables that presented a statistically significant trend. In *table 6* it is visible that the strongest and most significant is the only decreasing trend found, the decrease in oxygen at sea floor. The Sen slope gives a decrease of −21.7 mmol/m³/decade with a highly significant p value of p = 8.15e-06. A decrease in oxygen availability at the sea floor from 1998 to 2024 in the Western Gotland Basin was hypothesised since the decrease of oxygen availability, increase in anoxia and hypoxia is a well-documented phenomenon of the sea floor in the Baltic Sea in previous studies (Börgel et al, 2026; Funkey et al, 2014; Munkes et al, 2021).

The decrease of sea floor oxygen can also be observed in *figure 3*, as a visual tool for how both bloom frequency and sea floor oxygen have varied over 27 years. In *table 7* which displays resulting values of Mann-Kendall trend conducted on the deep basin and shallow shelf, separately. Both the shallow and deep areas of the basin display statistically significant values indicating a trend of oxygen decrease from 1998 to 2024 on the sea bottom. The deep basin and the shelf had p-values of 2.11e-05, 1.07e-06 and a Sen slope indicating decadal decrease of 29.2 mmol m^{-3}, and 21.4 mmol m^{-3} respectively.

5 Conclusion

This thesis examined spatiotemporal correlation between cyanobacteria blooms and hypoxia in the Western Gotland Basin of the Baltic Sea between 1998 to 2024. It is concluded that there was no synchronous correlation between the cyanobacteria summer blooms occurrence and dissolved oxygen availability at the sea floor. Sea surface temperature, mixed layer depth and bottom temperature were found to be the strongest co-varying variables to the bloom frequency, while the sea floor oxygen had one statistically significant co-varying variable, sea floor temperature. Oxygen at the sea floor showed a highly significant decrease from 1998, to 2024, with the fastest decrease where depth was less than 80m, and a basin average decrease of −21.7 mmol m^{-3} per decade. The lagged cross correlation between hypoxia and cyanobacteria bloom sustains coupling after being detrended, controlled for shared environmental drivers and checked for wind bias, but remains non statistically significant. The peak detrended lag at the shallow shelf area, above 80 m depth was 6 months, and 7 months at the deep basin below 80 m depth. However, the lagged cross correlation is descriptive and so it does not establish causation. Future research should investigate possible lead lag patterns over a larger temporal scale, as well as establishing causations.

6 References

Agrawal, A., & Scolnick, M. (2023). marimo - an open-source reactive notebook for Python [Computer software]. https://github.com/marimo-team/marimo

Andersson, A., Meier, H. E. M., Ripszam, M., Rowe, O., Wikner, J., Haglund, P., Eilola, K., Legrand, C., Figueroa, D., Paczkowska, J., Lindehoff, E., Tysklind, M., & Elmgren, R. (2015). Projected future climate change and Baltic Sea ecosystem management. AMBIO, 44(3), 345–356. https://doi.org/10.1007/s13280-015-0654-8

Anthropic. (2025). *Claude* (claude-sonnet-4-6) [Large language model]. https://claude.ai

Börgel, F., Ruvalcaba Baroni, I., Barghorn, L., Borchert, L., Cahill, B., Dutheil, C., Esters, L., Falarz, M., Filipsson, H. L., Gröger, M., Hänninen, J., Hieronymus, M., Jakobson, E., Karami, M. P., Kuliński, K., Liblik, T., Meier, H. E. M., Messori, G., Naumov, L., Neumann, T., Post, O., Rehder, G., Rutgersson, A., & Voelker, G. S. (2026). Large-scale atmospheric circulation and its impact on the Baltic Sea region: Controls, predictability and consequences. *Earth System Dynamics*, *17*(3), 415–450. https://doi.org/10.5194/esd-17-415-2026

Conley, D. J., Humborg, C., Rahm, L., Savchuk, O. P., & Wulff, F. (2002). Hypoxia in the Baltic Sea and Basin-Scale Changes in Phosphorus Biogeochemistry. Environmental Science & Technology, 36(24), 5315–5320. https://doi.org/10.1021/es025763w

E.U. Copernicus Marine Service Information (CMEMS). (2024). *Baltic Sea multiyear ocean colour plankton, reflectances and transparency L3 daily observations* [Data set]. Marine Data Store. Retrieved April 17 2026 from https://doi.org/10.48670/moi-00296

E.U. Copernicus Marine Service Information (CMEMS). (2024). *Baltic Sea physics reanalysis* [Data set]. Marine Data Store. Retrieved May 3, 2026 from https://doi.org/10.48670/moi-00013

E.U. Copernicus Marine Service Information (CMEMS).. (2024). *Global ocean hourly reprocessed sea surface wind and stress from scatterometer and model* [Data set]. Marine Data Store. Retrieved May 16 2026 from https://doi.org/10.48670/moi-00185

E.U. Copernicus Marine Service Information (CMEMS). (2023). *Baltic Sea biogeochemistry reanalysis* [Data set]. Marine Data Store. Retrieved May 3 2026 from, https://doi.org/10.48670/moi-00012

Feistel, R., Nausch, G., & Wasmund, N. (2008). *State and evolution of the Baltic Sea, 1952-2005 A detailed 50- year survey of Meteorology and Climate, Physics, Chemistry, Biology and Marine Environment.* (pp. 1 -3). John Wiley &Sons, Inc. ISBN 978-0-471-97968-5

Feng, C., Wang, H., Liu, H., & He, Z. (2026). Generalized Spearman rank correlation coefficient and its application. *Communications in Statistics: Case Studies, Data*

Analysis and Applications, *0*(0), 1–9. https://doi.org/10.1080/23737484.2026.2657333

Funkey, C. P., Conley, D. J., Reuss, N. S., Humborg, C., Jilbert, T., & Slomp, C. P. (2014). Hypoxia Sustains Cyanobacteria Blooms in the Baltic Sea. Environmental Science & Technology, 48(5), 2598–2602. https://doi.org/10.1021/es404395a

Hansson, M., & Viktorsson, L. (2025). *The oxygen situation in the Baltic Sea 2024* (SMHI Oceanography Report No. 80). Swedish Meterological and Hydrological Institute (SMHI). https://www.smhi.se/en/publications-from-smhi/publications/2025-04-14-the-oxygen-situation-in-the-baltic-sea-2024

Lass, H., Matthäus, W. (2008). General oceanography of the Baltic Sea. In Feistel, R., Nausch, G., & Wasmund, N. (Eds.). *State and evolution of the Baltic Sea, 1952-2005* (pp. 5 - 7). John Wiley &Sons, Inc. ISBN 978-0-471-97968-5

Munkes, B., Löptien, U., & Dietze, H. (2021). Cyanobacteria blooms in the Baltic Sea: A review of models and facts. Biogeosciences, 18(7), 2347–2378. https://doi.org/10.5194/bg-18-2347-2021

Nausch, G., Nehring, D., & Nagel, K. (2008). Nutrient concentrations, trends and their relation to eutrophication. In Feistel, R., Nausch, G., & Wasmund, N. (Eds.). *State and evolution of the Baltic Sea, 1952-2005* (pp. 337- 347). John Wiley & Sons, Inc. ISBN 978-0-471-97968-5

NOAA National Centers for Environmental Information. 2022: ETOPO 2022 15 Arc-Second Global Relief Model [Data set]. NOAA National Centers for Environmental Information. Retrieved May 21 2026. From https://data.noaa.gov/metaview/page?xml=NOAA/NESDIS/NGDC/MGG/DEM//iso/xml/etopo_2022.xml&view=getDataView&header=none

Python Software Foundation. (2024). *Python* (Version 3.12.6) [Computer software]. https://www.python.org

Redfield, A. C., Ketchum, B. H., Richards, F.A. (1963) The influence of organisms on the composition of sea-water (pp.27). ISBN9780674017283. https://documentatiecentrum.watlab.be/imis.php?module=ref&refid=28944

Stal, L. J., Albertano, P., Bergman, B., Bröckel, K. von, Gallon, J. R., Hayes, P. K., Sivonen, K., & Walsby, A. E. (2003). BASIC: Baltic Sea cyanobacteria. An investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea—responses to a changing environment. *Continental Shelf Research, European Land-Ocean Interaction*, *23*(17), 1695–1714. https://doi.org/10.1016/j.csr.2003.06.001

The Baltic Marine Environment Protection Commission (HELCOM). (2022). *HELCOM subbasin with coastal WDF waterbodies or watertypes 2022level 4a.* [Data set]. ICES Metadata Catalogue. Retrieved May 20, 2026 from https://maps.helcom.fi/arcgis/rest/services/MADS/Monitoring/MapServer/109

Vahtera, E., Conley, D. J., Gustafsson, B. G., Kuosa, H., Pitkänen, H., Savchuk, O. P., Tamminen, T., Viitasalo, M., Voss, M., Wasmund, N., & Wulff, F. (2007). Internal Ecosystem Feedbacks Enhance Nitrogen-fixing Cyanobacteria Blooms and Complicate Management in the Baltic Sea. AMBIO: A Journal of the

Human Environment, 36(2), 186–194. https://www-jstor-org.ludwig.lub.lu.se/stable/4315813?seq=1

Wasmund, N., & Siegel, H. (2008). Phytoplankton. In Feistel, R., Nausch, G., & Wasmund, N. (Eds.). *State and evolution of the Baltic Sea, 1952-2005* (pp. 445-446). John Wiley & Sons, Inc. ISBN 978-0-471-97968-5

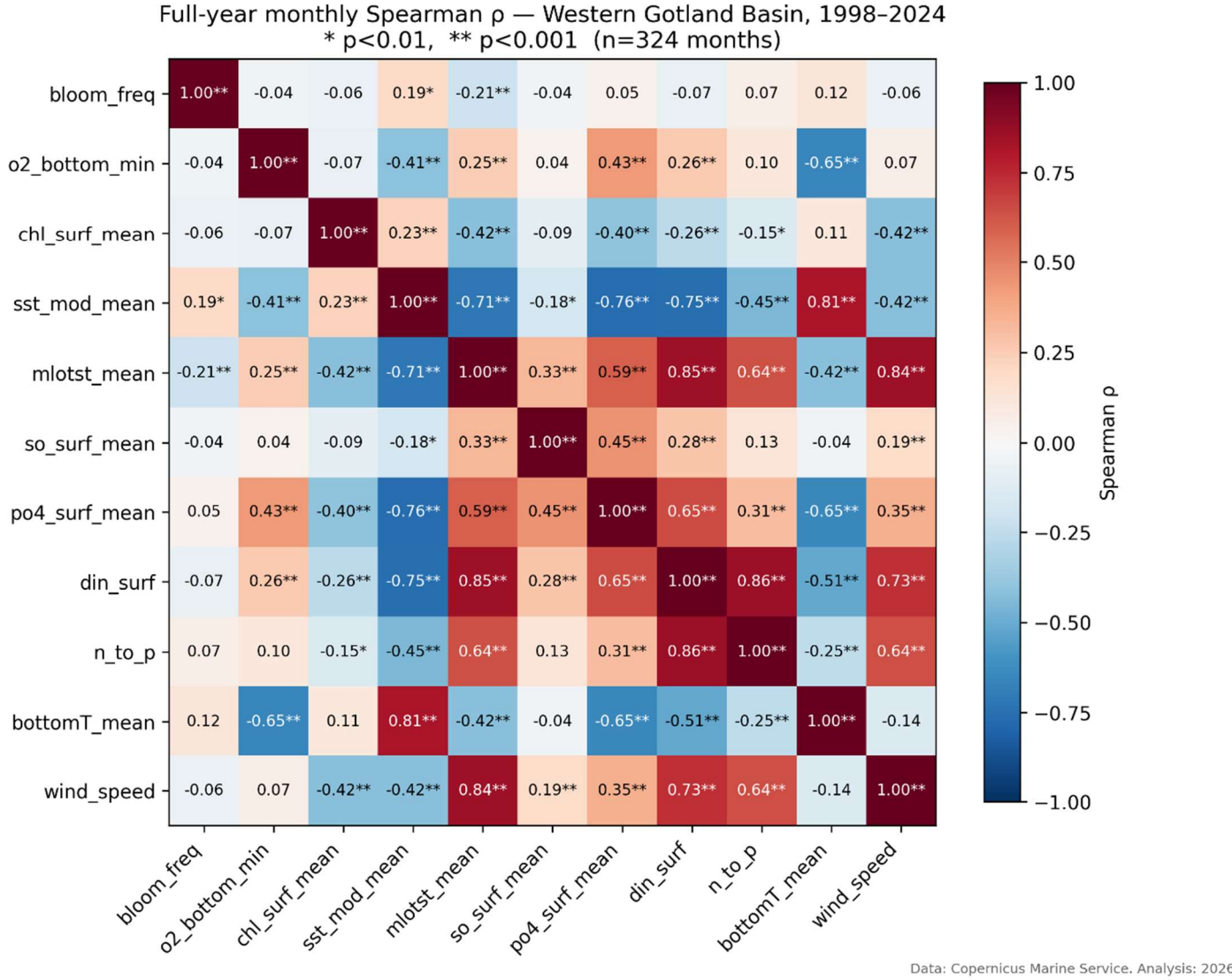
7 Appendix

Appendix 1. Python Software Environment (Packages). Extended list including packages not mentioned in methods description. Repository /DOI provides links to their GitHub webpage or DOI provided by package creators per their webpages stated requests.

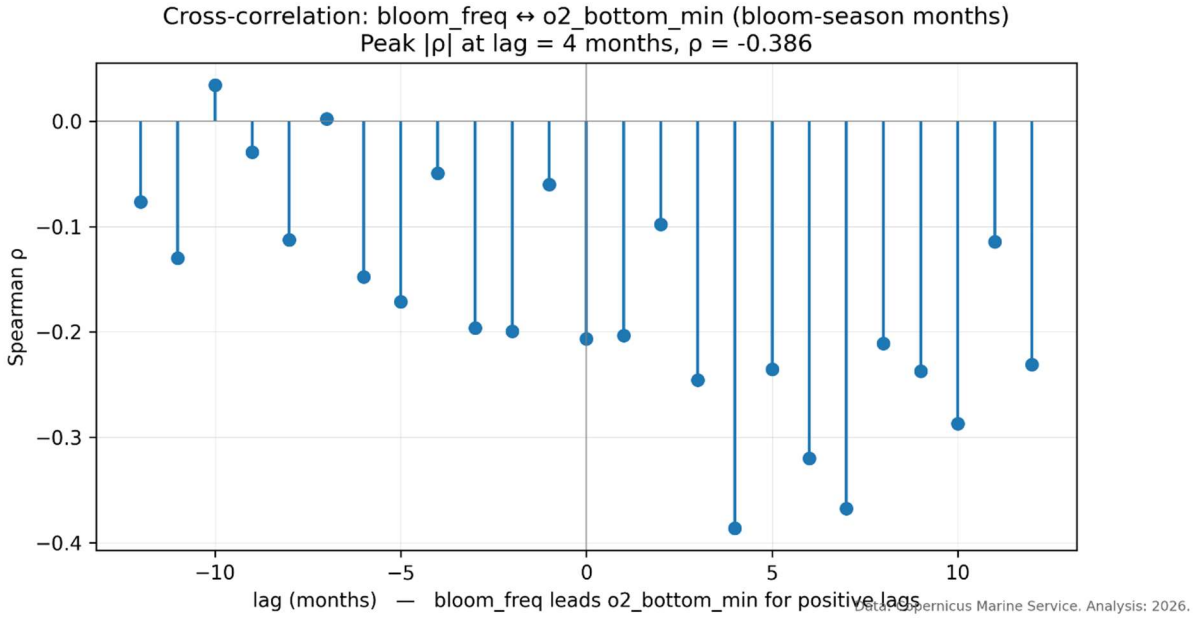
Python package	**Version**	**Repository/DOI**
Python	3.12.6	https://www.python.org
uv	0.11.8	https://github.com/astral-sh/uv
marimo	0.23.6	https://github.com/marimo-team/marimo
xarray	2026.4.0	https://github.com/pydata/xarray/tree/v2026.04.0 doi.org/10.5281/zenodo.598201
numpy	2.4.6	https://github.com/numpy/numpy
pandas	3.0.3	https://github.com/pandas-dev/pandas doi.org/10.5281/zenodo.3509134
netCDF4	1.7.4	https://github.com/Unidata/netcdf4-python
dask	2026.3.0	https://github.com/dask/dask
geopandas	1.1.3	https://github.com/geopandas/geopandas
shapely	2.1.2	https://github.com/shapely/shapely
Cartopy	0.25.0	https://github.com/SciTools/cartopy
rioxarray	0.22.0	https://github.com/corteva/rioxarray
pyproj	3.7.2	https://github.com/pyproj4/pyproj
scipy	1.17.0	https://github.com/scipy/scipy

pymannkendall	1.4.3	https://github.com/mmhs013/pyMannKendall doi.org/10.21105/joss.01556
statsmodels	0.14.6	https://github.com/statsmodels/statsmodels
matplotlib	3.10.9	https://github.com/matplotlib/matplotlib doi.org/10.1109/MCSE.2007.55
copernicusmarine	2.4.1	https://github.com/mercator-ocean/copernicus-marine-toolbox
pyarrow	24.0.0	https://github.com/apache/arrow
rasterio	1.5.0	https://github.com/rasterio/rasterio.git

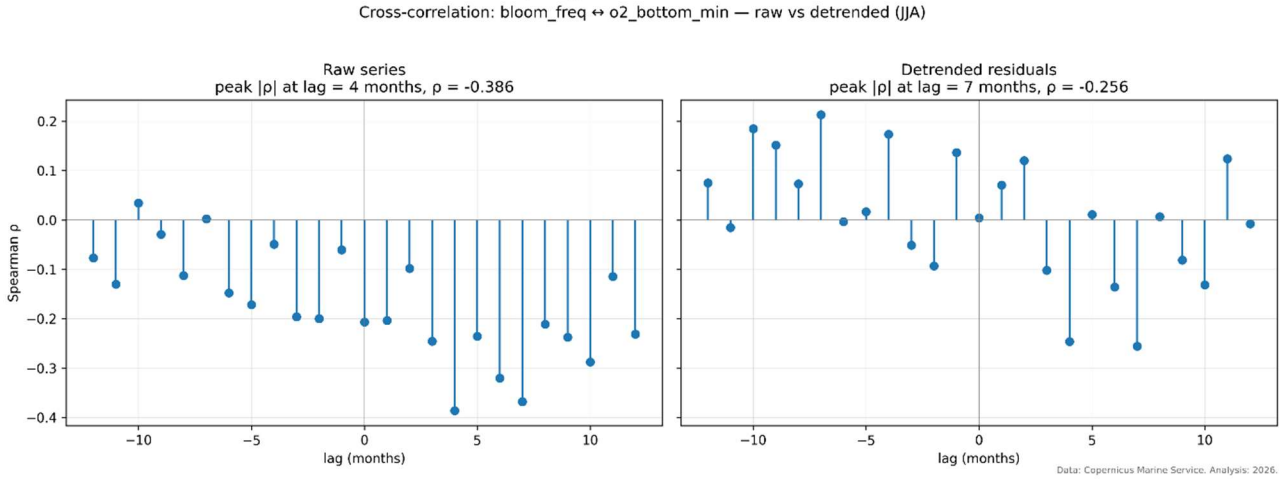
Appendix 2. Full year monthly Spearman cross correlation on all variables. Includes all months for the years 1998 to 2024. Done for visual comparison on cross correlation values for all variables when using only June, July & August or the full year. The resulting values of *Appendix 2* were not used for the final results.



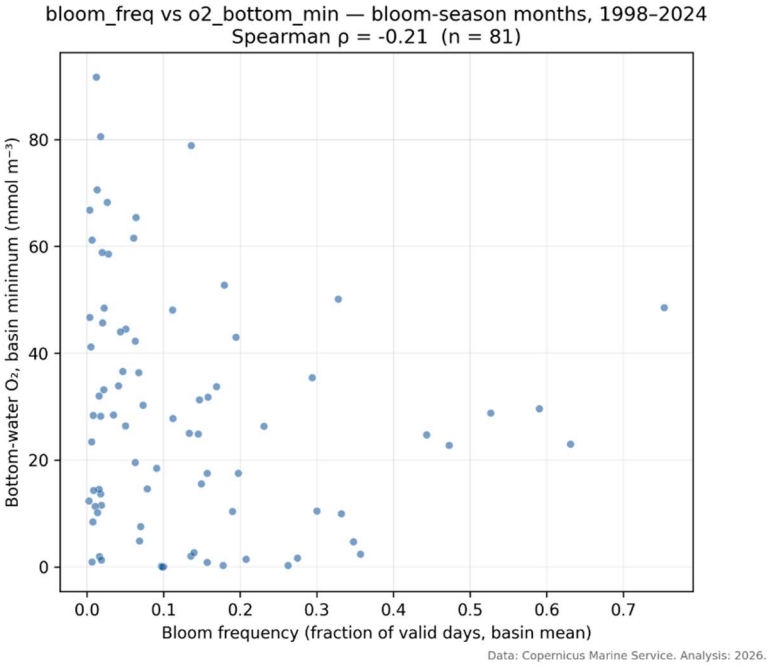
Appendix 3. Bloom Frequency and sea floor minimum oxygen, cross correlated raw data on possible lead lag. The cross correlated data is averaged over the Western Gotland Basin. The cross correlation exhibited a negative ρ, with peak lag at 4 months. The lag was computed with + & - 12 months.



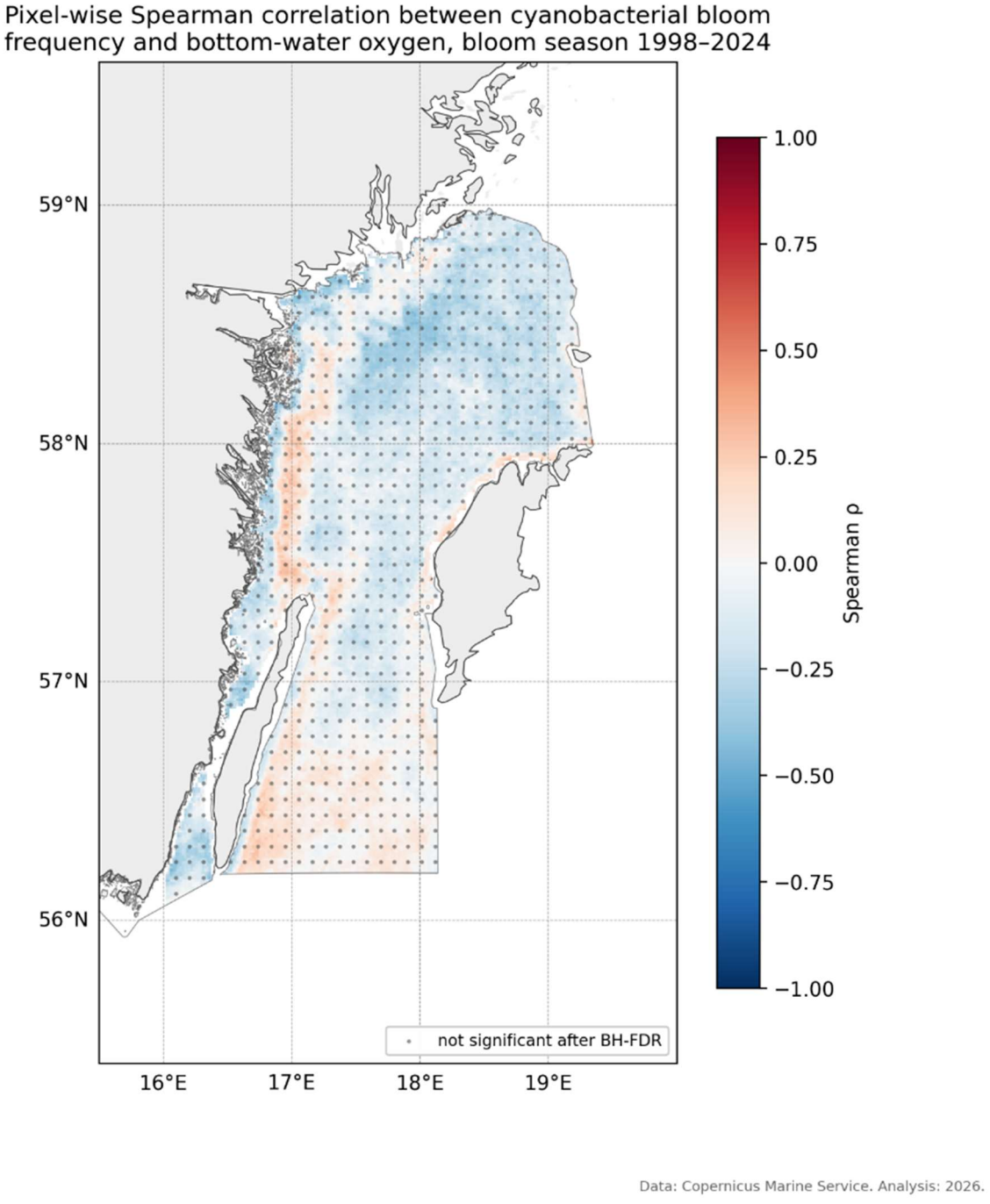
Appendix 4. Comparing the graphs of cross correlated and raw and detrended data on oxygen and bloom lag, basin wide on the Western Gotland Basin from 1998 to 2024. The left graph is the same as exhibited above in *Appendix 3*, while the right image is a cross correlation with the same data but detrended. The right image exhibits more variation between positive and negative ρ, and a peak negative cross correlation at 7 months. Neither is statistically significant.



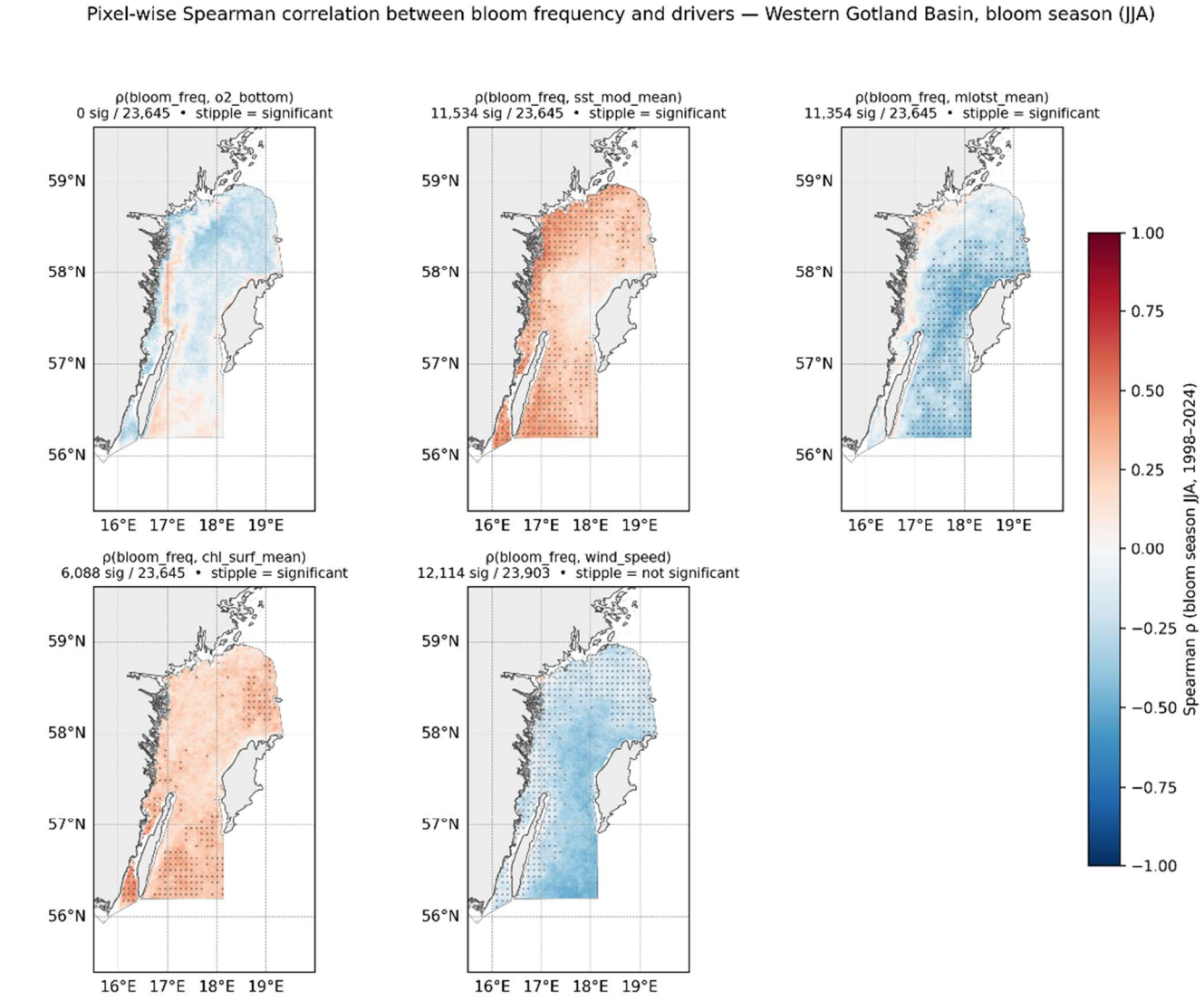
Appendix 5. Visualisation of sea floor oxygen and bloom frequency. Computed as a visual aid for establishing a non-linear relationship to determine suitability for Spearman correlation as a methodological tool.



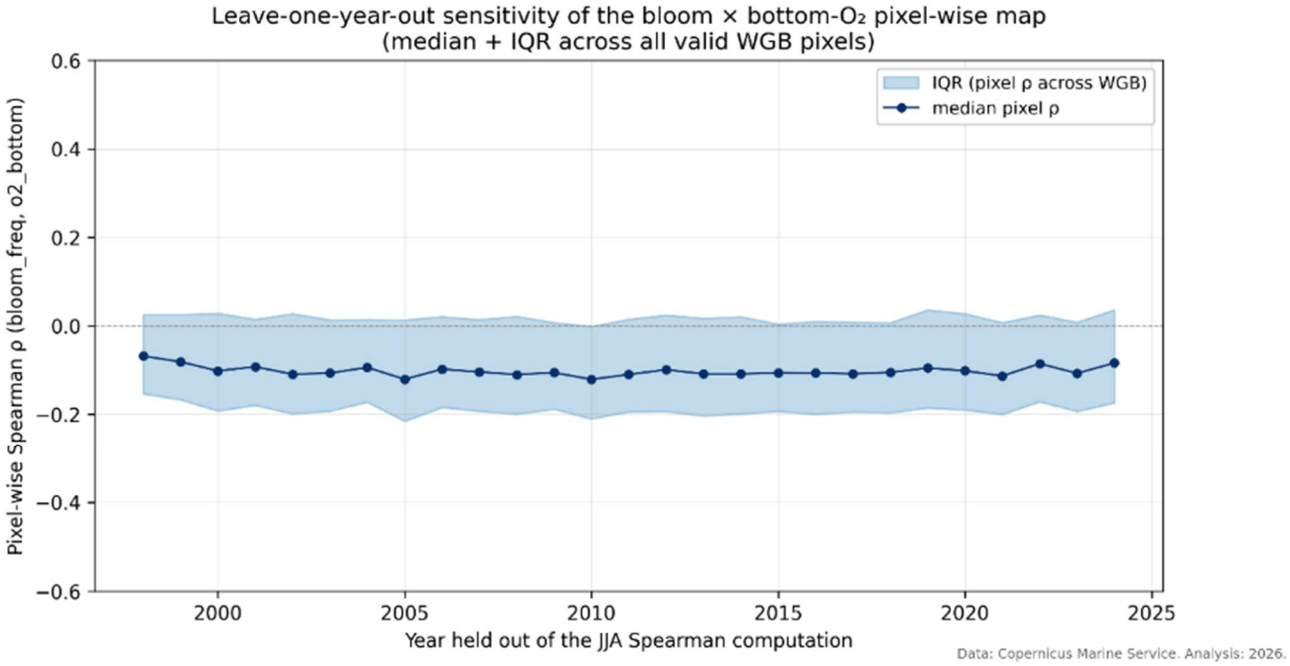
Appendix 6. Pixel wise Spearman correlation of bloom and oxygen. Doing cross correlation per pixel between sea floor minimum oxygen and sea surface cyanobacteria bloom frequency on all monthly data for June, July and August from 1998 to 2024. The results are visible in the *Appendix 6*. Because there is no statistically significant area, as signified by the dots in the image, it exhibited the same results as the basin wide averages for cross correlating the same two variables. Therefore, the image was not displayed in the results section since it can be considered redundant but rather included here in the appendix.



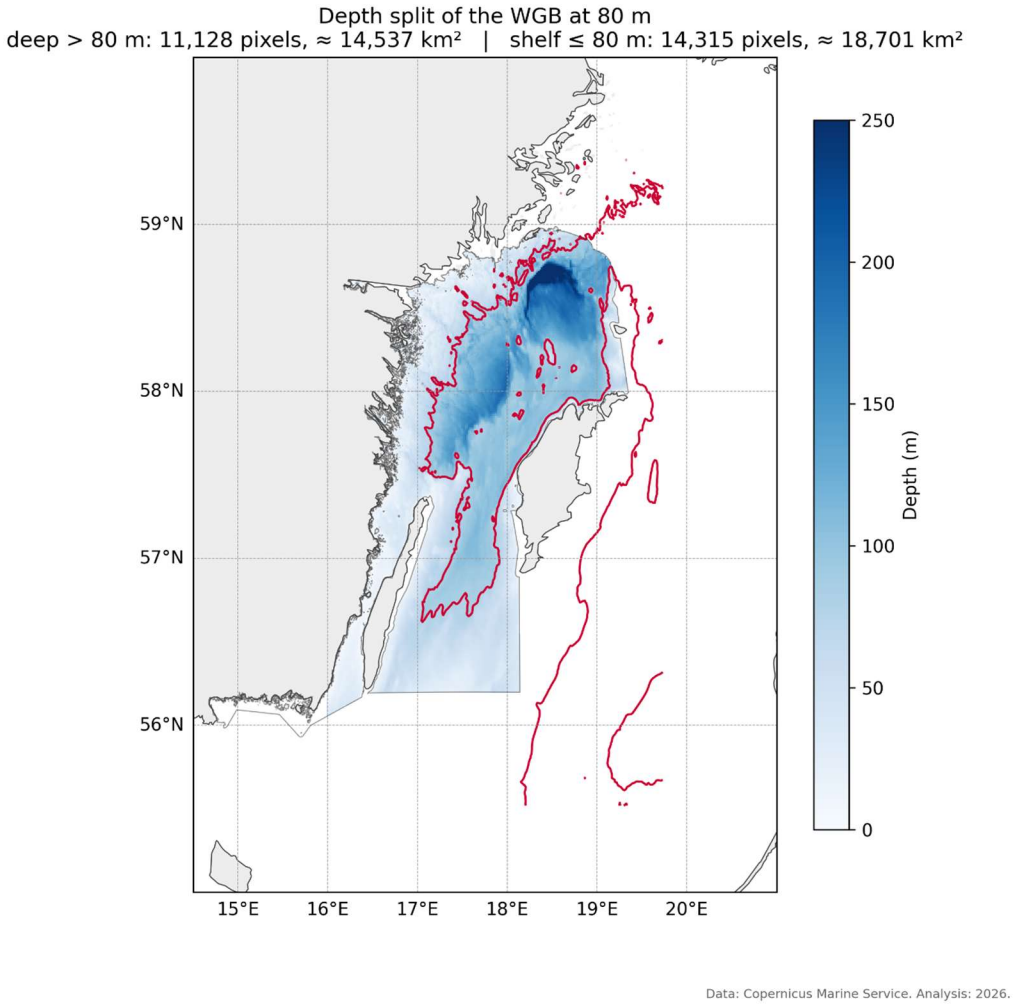
Appendix 7. Pixel wise Spearman correlation of multiple variables. The same pixel wise cross correlation as visible in *Appendix 6* but done on the bloom frequency and its biggest drivers, min O_2 at sea floor, SST mean, mixing depth mean, chlorophyll a mean, and wind speed mean. The cross correlation is conducted on the Western Gotland Basin, monthly averaged data for June, July and August between 1998 to 2024. The colour of the image indicates the Spearman ρ for each pixel, while dots show statistical significance in image one, two, three and four, while in the fifth image, wind speed correlated to bloom frequency a dot indicates that the cell is not statistically significant. Over each image the number of statistically significant pixels are displayed as "0 sig/ 23,645" for the first image of O2 bottom and cyanobacteria bloom. These images were not included in the results sections since the relevant findings could be retrieved from the cross-correlation matrix as seen in *figure 4* and therefore considered redundant.



Appendix 8. This graph displays a sensitivity analysis test on sea floor oxygen and cyanobacteria bloom frequency. This test was conducted to assess whether there was any summer season over the studied years which caused a large difference to the resulting values. This is also called a Benjamini-Hochberg procedure, here it is done as a Spearman cross correlation with one year taken out at the time. In the graph, the median ρ and interquartile range, IQR, indicating variability, values remain stable, which in turn indicate that no years observed exerted an especially high impact on the results.



Appendix 9. The map displays the subbasin of Western Gotland Basin, with the red lines outlining the depth split at 80 m, since this is where the average halocline depth is. Lighter blue areas, specifically near the coast are referred to as shallow shelf areas in text, while deeper, dark blue areas are referred to as deep basin in the running text. The image displays the basins bathymetry, sea floor depth, as provided by the by the Bathometry data from NCEI, ETOPO 2022 15 Arc-Second Global Relief Model (NOAA, 2022).



Appendix 10. Lagged cross correlation of cyanobacteria bloom frequency and sea floor oxygen below and above approximated halocline at 80 m depth. The lagged cross correlation was done on the detrended data, in the Wester Gotland Basin for the months of June, July and August between 1998 to 2024. The time interval investigated was +12 to – 12 months. The different depths showed differing peak Spearman ρ at 7 months for the deep basin and 6 months for the shallow shelf. The correlation remained statistically insignificant as per previous lead lag correlations.