LONG -TERM DYNAMICS OF ZOOPLANKTON ABUNDANCE IN RELATION TO ENVIRONMENTAL CONDITIONS AND PHYTOPLANKTON BIOMASS

Master's Thesis

Systems Biology Master's program

Vilnius University

STUDENT NAME: Živilė Grublytė
STUDENT NUMBER: 1916238

SUPERVISOR: Johannes Werner

SUPERVISOR'S DECISION: ....

FINAL GRADE   ....

DATE OF SUBMISSION: 12 05 2021

CONTENTS

LIST OF ABBREVIATIONS

Aut – autumn

CTD - conductivity, temperature, and depth

Dbar - decibar (equal to 1 meter sea water)

F – females

HELCOM - Helsinki Commission

ICES - International Council for the Exploration of the Sea

J1 – juveniles, 1'st stage

J2 – juveniles, 2'nd stage

Lat - latitude

Long – longitude

M - males

MDS - multidimensional scaling

nMDS – non-metric multidimensional scaling (nMDS) analysis

NTRA - nitrate molar concentration (mmol/m^3)

OXY - oxygen concentration (ml/L)

PC – principal component

PCA – principal component analysis

PHOS - dissolved inorganic phosphate molar concentration (mmol/m^3)

PRES - pressure at which samples were collected (dbar)

PSU - Practical Salinity Unit, equivalent to per thousand or ($^0/_{00}$) or to g/kg

SAL – salinity (PSU)

SLCA - silicate molar concentration, (mmol/m^3)

Spp - multiple species

Spr – spring

SST – sea surface temperature

Sum – summer

TEMP - Temperature (^{o}C)

V. – version

Win – winter

INTRODUCTION

There are many ways in which climate change can affect the ocean. For instance, water temperature is rising and consequentially the level of dissolved oxygen is decreasing. Furthermore, the concentration of carbon dioxide in the atmosphere is increasing, which causes acidification of water. The regional changes are also important: on the sites, where ice is melting, regional freshening occurs while regional salinification is detected in the areas of vast evaporation (Käse and Geuer, 2018). Moreover, using nitrogen fertilizers can cause increasing levels of nitrogen compounds in water (Rhein et al., 2013). All these hydrographic changes can affect marine ecosystems. Plankton being small, sensitive to their environment water organisms, having fast reproduction rate and short life cycle quickly respond to these environmental changes. These features make plankton to be good bioindicators, suggesting us how environmental variations affect marine ecosystems. Therefore, the desire to better understand climate change influence on living water stocks has led to increased interest into long term changes in plankton dynamics (Beaugrand and Reid, 2003, 2012). There are already a number of studies analysing changes in plankton abundance or diversity in various seas and lakes, which demonstrated significant changes in plankton communities (Kozak and Gołdyn, 2004; Rousseaux and Gregg, 2015). However, those changes are usually dependent on the regional hydrographic conditions, thus the results cannot be applied globally.

In the Baltic Sea long term changes in plankton dynamics were also analysed. There were a number of studies investigated changes in phytoplankton diversity (Wasmund, Nausch and Feistel, 2013; Hjerne et al., 2019) and its responses to environment changes (Griffiths et al., 2020). Furthermore, studies on zooplankton trends in the Baltic Sea included the analysis of changes in zooplankton size and stock in the Northern Baltic sea (HELCOM, 2018) and response to hydrographic environment variability in the Southern (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019) and Southeastern Baltic Sea (Aleksandrov, Zhigalova and Zezera, 2009). However, even though the Western Baltic Sea is unique in its complex hydrographic conditions due to the in- and out-flux of water from the North Sea (Lehmann, Krauss and Hinrichsen, 2002), plankton dynamics in this area are only poorly studied and understood. Phytoplankton trends studies from the Mecklenburg Bight covered only the period from 1979 until 2005 (Wasmund and Uhlig, 2003; Wasmund et al., 2011) and there are no comprehensive studies of zooplankton trends in the Western Baltic Sea. Therefore, this study focuses on time series data from Mecklenburg Bight (Western Baltic Sea). It may reveal whether climate change (temperature) and related hydrographic shifts (salinity) or eutrophication (nutrients) as well as changes in the phytoplankton biomass have influenced the zooplankton abundance in the Western Baltic Sea

Aim and tasks:

The aim of the study is to detect long-term (1998-2016) shifts in the Western Baltic Sea (Mecklenburg Bight) ecosystem and to indicate how links between plankton abundance and environmental conditions as well as between different plankton groups can be established.

In order to achieve this aim, the following tasks will be performed:

1. Analysing whether intraannual environmental variability has an influence on the abundance of plankton communities.
2. Detecting monotonic increasing and decreasing long term (1998-2016) trends in the time series of the zooplankton abundance, phytoplankton biomass and hydrography variables.
3. For the plankton taxa, demonstrating shifting trends, relationships between plankton abundance and environmental conditions have to be analysed and possible biological implications discussed.
4. Investigating the relationship between zooplankton abundance and phytoplankton biomass.

LITERATURE REVIEW

Plankton – definition, classification and role in the ecosystem

Plankton are marine and freshwater organisms, which are unable to swim well enough to move against forces of tides and currents, therefore they exist in a drifting state. Organisms of some species are classified as plankton for their entire life, while others are qualified as plankton only in the beginning of their life and eventually become large enough to move against currents (National Oceanic and Atmospheric Administration, 2020). Most plankton species are microscopic organisms however they also include larger species like some crustaceans and jellyfish. Plankton organisms can be classified in several ways, including classification by size (macro, meso, micro, nano, and pico), metabolism (autotrophic vs. heterotrophic), by life history (meroplankton - temporary members of the plankton, mostly larval forms of some animals - and holoplankton - permanent members of the plankton) or by taxonomy (i.e., Crustaceous vs. Gelatinous zooplankton) (Lumini and Nanni, 2019). However, the most basic categories divide plankton simply into two groups: phytoplankton (primary producers) and zooplankton (secondary producers) (Figure 1), although a clear separation between these two groups is not always possible.
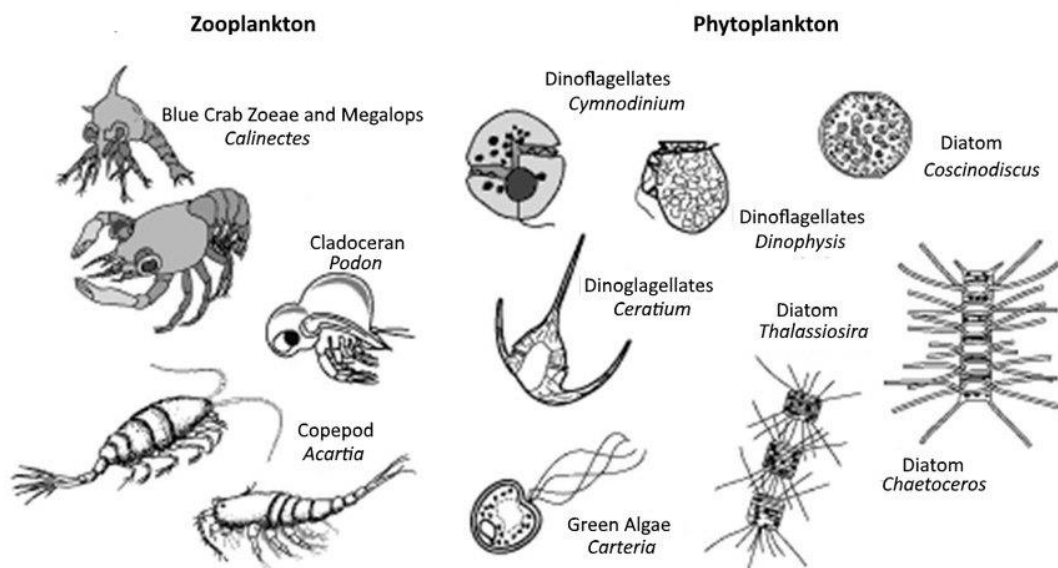
Figure 1 A simple plankton taxonomy according to Lumini and Nanni, 2019.

Phytoplankton are defined as a group of microscopic, unicellular organisms that play an essential role as the base of food webs in aquatic ecosystems (D'Costa and Naik, 2019). They perform photosynthesis to produce energy, therefore phytoplankton are found close to the water surface (NOAA, 2017). Chlorophyll-containing photosynthetic protists may also occasionally have a heterotrophic nutrition and are called mixotrophs. One prominent example of mixotrophic organisms

are dinoflagellates, which include both photosynthetic and heterotrophic flagellates (Flynn et al., 2013). Since the microscopic phytoplankton analysis also allows the quantification of heterotrophic organisms of the typical phytoplankton size (nano- and microplankton), sometimes unicellular heterotrophic organisms are also involved in phytoplankton analysis. Therefore mixotrophic and heterotrophic organisms were also included in the diverse phytoplankton classes, such as Cyanobateria and Dionphyceae, in this study.

Zooplankton are small (usually less than 2 mm long) aquatic invertebrates which usually drift in deeper water throughout the daytime hours and venture up to the surface at night to feed on phytoplankton (NOAA, 2017). The best-known examples of zooplankton groups are copepods, cladocerans (water fleas) and rotifers (Hendrey, 2001). Zooplankton also sometimes includes protozoans, coelenterates, larval flatworms, mites, insect larvae and fish larval stages (Stanley Gregory, 2003). Cladocerans and copepods are the main crustacean zooplankton (Sommer and Sommer, 2006). These groups of zooplankton differ in their reproduction: cladocerans reproduction occurs by parthenogenesis of diploid females whereas males are often absent for many generations, while copepods reproduction occurs by sexual fertilization with males always being present. Therefore, when data about zooplankton is collected, adult copepods are usually distinguished by their sex while cladocerans are not.

Phytoplankton and zooplankton are closely linked to each other and are dependent on several environmental and biological factors, which affect the growth of each community and the interaction between them (Aziz, Gharib and Dorgham, 2006). Phytoplankton rapidly reacts to environmental changes such as temperature, salinity and nutrient concentration of the water. Zooplankton usually consumes phytoplankton as energy source, thus changes of phytoplankton abundance can quickly affect zooplankton populations. In addition to their interaction, phytoplankton and zooplankton also play an important role in marine ecosystem in general. For instance, plankton is a beneficial food source for animals, especially in their larval stages, hence plankton studies aid in understanding the natural aquatic ecosystem and in predicting the number of fish which can be received from the ocean (Rodríguez *et al.*, 1989). On top of that, studies of plankton may allow to grasp the impact of overfishing and the climate change. These facts altogether make plankton analysis a relevant and interesting topic in environmental research.

Collection and storage of the plankton data

Monitoring plankton

There are several techniques to sample plankton from the water pond. Integrated per depth samples can be collected using a hose (Lindahl, 1986) or by pooling equal amounts of water collected from

various depths (Majaneva *et al.*, 2009). In order to obtain plankton samples with high concentrations, nets are used. Net mesh-size depends on the plankton which should be collected. For example, to collect phytoplankton a net with a 10 µm mesh-size is recommended (HELCOM, 2020), while for mesozooplankton sampling net with 100 µm mesh-size is usually used (HELCOM, 2017a). Large mesh size can cause losing small-sized species as well as zooplankton nauplii from the sample, while choosing small mesh-size might reduce the efficiency of filtration and make large organisms to avoid the net (Garcia *et al.*, 2020). Not only devices used to collect data might have an influence on the sampling results, there are other factors such as speed, direction and depth at which the net is hauled, subsampling techniques among other things. The differences in plankton collection methods can make it difficult to interpret the data and compare the results of sampling in time or in different locations. To avoid this, common guidelines for the plankton monitoring have to be developed and adopted by the research groups.

Data quality control and storage

Baltic Sea zooplankton time series data have been collected since the 1950s (Viitasalo, 1992; Ojaveer, Lumberg and Ojaveer, 1998). Sampling techniques and analysis methods have to be harmonized in order to obtain comparable data. Therefore, the Helsinki Commission (HELCOM) started Baltic Sea basin-wide coordination of zooplankton monitoring in 1979 (Helcom, 2017). HELCOM provides guidelines for monitoring plankton data, including techniques on describing how to sample plankton, the sample collection frequency, as well as quality control procedures which is required to generate relevant and reliable data. Currently, the HELCOM members (Denmark, Estonia, Finland, Germany, Latvia, Lithuania, Poland, Russia, Sweden and the European Union) are obliged to conduct monitoring of coastal waters and to report the data to the Commission. All data collected in the Baltic Sea are available for each of the HELCOM sub-basins (Figure 2) from the International Council for the Exploration of the Sea (ICES) data portal. The main objective of ICES is to improve the scientific knowledge of marine environments and its living resources as well as to provide advice to competent authorities (*ICES – HELCOM*, no date). Consequently, various datasets related to marine environments are stored in this portal and can be freely accessed by researchers.

Figure 2 Sub-basins of the Baltic Sea (Helcom, 2001).

The main challenges with plankton time series statistical analysis

There are several common issues that scientists face when working with plankton time series data, especially unevenly spaced data, seasonality in time series and complicated interpretation of the results. Researchers developed different approaches to overcome these problems and to present informative statistical analysis results which are discussed below.

Unevenly spaced data

Time series can be defined as a sequence of observations occurring in time. When measurements are performed in regular time intervals, then the time series are evenly spaced. Most methods dealing with time series can only handle regular time series (Stanley Gregory, 2003). However, unevenly sampled time series are common in many real-life situations. Plankton data are mostly collected in irregular time intervals. The main aim of phytoplankton sampling is to cover the growth season, therefore winter phytoplankton measurements are only conducted rarely. In zooplankton

time series data gaps also occurs for a variety of logistical reasons common to work at sea (Valdés and Moral, 1998).

There are different approaches to deal with unevenly spaced data. One of the ways is to aggregate data calculating arithmetic mean over all samples for a station within a season (Wasmund and Uhlig, 2003). As the timing of blooms differs among individual areas, the definition of season is needed. HELCOM recommended the definition of seasons by months in the Baltic Sea (Table 1). Also, plankton data can be aggregated for two-month periods (Beaugrand and Reid, 2003).

Table 1 Definition of seasons (HELCOM, 1996)

Season	Belt Sea, Sound, Kattegat	Baltic proper
Spring	February-April	March-May
Summer	May-August	June-September
Autumn	September-November	October-December
Winter	December-January	January-February

The gaps in plankton series can also be filled by imputing missing values (Valdés and Moral, 1998). There are various interpolation methods, but in general, good interpolation technique should meet these criteria: it should not require a lot of data to fill in missing values; estimation of model parameters and missing values should be permitted at the same time; large-series computation must be efficient and fast, procedure should be applicable to stationary and non-stationary time series (Beveridge, 1992) and selected interpolation methods should be accurate and robust (Lepot, Aubin and Clemens, 2017).

Finally, gaps in time series can be overcome by local smoothing (Fryer and Nicholson, 1999; Kane, 2007; Bode *et al.*, 2015). Local smoothing operates on the assumption that between a criterion variable and a set of predictor variables, there is an unknown but continuous function which estimates the value of each point as a weighted average of points defined as "close" by the predictor variables (Huber, 1977) (Figure 3).
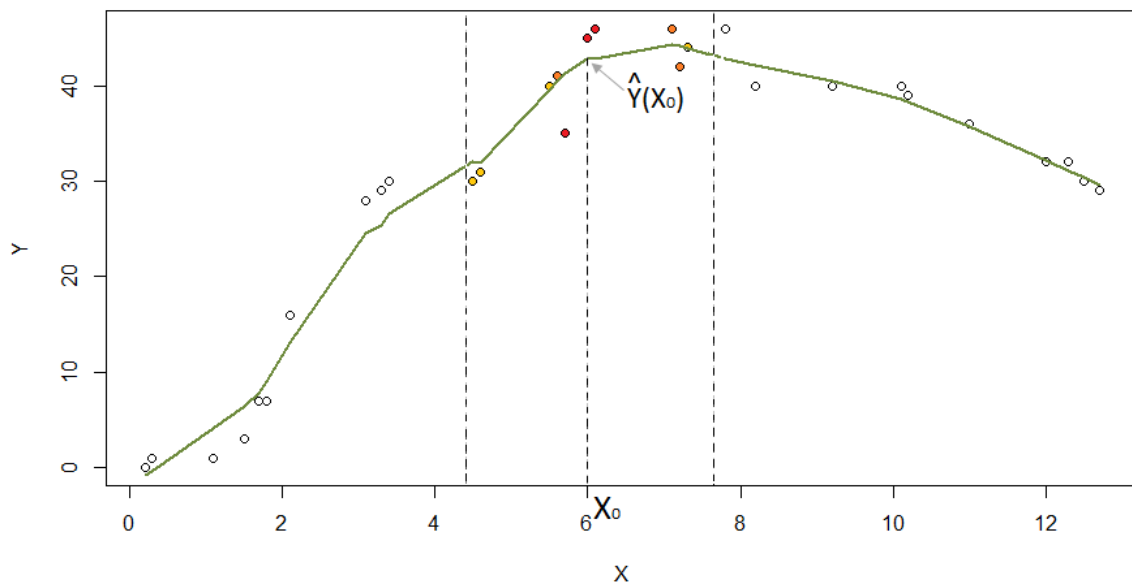
Figure 3 Example for local smoothing. Only a subset of data points (1/3 here) are used to evaluate the function for a specific predictor. The closest points have the largest weight for the evaluation.

In summary, the issue of irregularities in plankton sampling can be solved by different methods. The main groups of techniques, used to convert irregular plankton time series into evenly spaced time series are data aggregation, interpolation of missing values and imputing missing values by local smoothing.

Seasonality in time series

Plankton is very sensitive to environmental changes and responds quickly to variations in water temperature or salinity (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019). As a consequence, it also reacts to environment seasonality. It was shown, that both phytoplankton and zooplankton exhibit consistent seasonal patterns in taxonomic composition and size structure (Benedetti *et al.*, 2019). For example, in spring the water temperature is increased by sunlight which creates a thermocline trapping nutrients at the ocean surface. Phytoplankton quickly responds to these conditions, letting them absorb energy and intake nutrients, which they need for photosynthesis and growth, which eventually causes spring blooms and consequently strong increase in phytoplankton abundance. Similarly, as autumn begins, vertical water mixing may bring nutrients up, resulting in smaller fall blooms (Garrison, 2012). If data is aggregated as described in the previous section, every season can be analysed separately (Wasmund and Uhlig, 2003). It is important to recognize and remove seasonal components from the time series so they would not be confused with long-term trends.

13

Time series can be decomposed into three components: a trend component T_t, a seasonal component S_t and a remainder (error) component E_t. A trend is a long-term change (increase or decrease) in time series. Seasonality in time series data can be described as a regularly repeating pattern that happens on fixed intervals (such as the calendar year) (Gelder, 2018). The remainder component comprises the remaining variation in a time series once any trend and seasonal variation have been removed (Gary Napier, 2020). The decomposition can be either additive or multiplicative (Equation 1).

Equation 1. Time series decomposition. 1) Additive decomposition 2) Multiplicative decomposition. Y_t – time series value (actual data) at period t, T_t – trend component, S_t – seasonal component, E_t – remainder component.

$$1)\ Y_t = T_t + S_t + E_t \qquad 2)\ Y_t = T_t * S_t * E_t$$

The additive model is used when the trend and seasonal variation act independently, while a multiplicative model should be chosen when seasonal effect depends on the size of the trend (Gary Napier, 2020) (Figure 4).
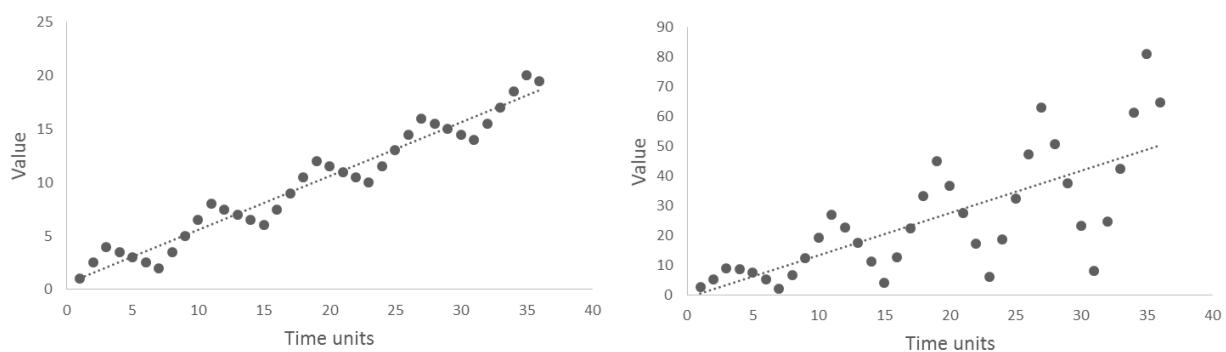
Figure 4 Examples of data for additive (left) and multiplicative (right) models.

Seasonal movements can be large enough that they mask other characteristics of the data that are object of interest for researchers. For example, seasonality in time series can hamper the analysis of long term trends. The process of estimating and removing seasonal effects from the time series in order to better reveal certain non-seasonal features is known as seasonal adjustment. Removing the seasonal component directs focus on characteristics, which seasonal movements tend to mask, such as long-term changes of the series (*MCD - Seasonal Adjustment Frequently Asked Questions*, 2017). In plankton time series, seasonally adjusted data make the comparison of long-term trends of different plankton species and environmental variables possible.

Challenges interpreting the results of the analysis

Even when statistical analysis is already performed, the interpretation of the results is usually not straightforward. Although currently clear guidelines for plankton sampling have been developed to obtain harmonized data, some challenges related to plankton data analysis remain. First, technologies used to monitor plankton may advance over time which may alter the sampling results. Moreover, although the zooplankton are collected by towing fine-meshed nets through the water, not all species are representatively captured by this method. Some species are too small and cannot be captured using nets, others are too fragile and do not survive collection by nets and subsequent processing in chemical preservatives (Lalli and Parsons, 1997). Also, changes in zooplankton abundance might occur because younger forms of organisms escape the net and are therefore not detected. Taken together, it becomes clear that the detected plankton abundance cannot always represent the actual plankton abundance. Therefore, other factors that might have an influence on the detection of plankton should be kept in mind interpreting the results of the statistical analysis.

Research on plankton long-term trends

Plankton and environmental changes

Both phytoplankton and zooplankton respond strongly on their surrounding environment. Phytoplankton depends on many physical features of their environment such as water salinity, temperature, availability of nutrients or light (Käse and Geuer, 2018). Various physiological processes of zooplankton are also sensitive to water temperature (Mauchline, 1998). Moreover, their reproduction rate is fast, and they have short life cycle, so they quite quickly respond to the changes in the environment and therefore are good bioindicators (Mitsch and Gosselink, 2011). For these reasons, long-term trends of phytoplankton and zooplankton are analysed to provide a better understanding about the influence of changes in environment on marine ecosystems.

Previous plankton time series studies have shown that dynamics in marine ecosystems are constantly changing. For example, assessing the trends of phytoplankton composition at a global scale have shown, that there is significant decline in diatoms percentage of total phytoplankton (Rousseaux and Gregg, 2015). It was also shown that primary production in the Arctic Ocean has increased by 57% between 1998 and 2018 due to the ice loss (Lewis, Van Dijken and Arrigo, 2020). Zooplankton communities are also changing: the increasing water temperature leads to shift towards zooplankton communities dominated by small copepods (Rice, Dam and Stewart, 2015; Kelly *et al.*, 2016). Demonstrated changes in marine ecosystems prove that long-term trends of the plankton data can provide valuable information about how climate change affects marine ecosystems and might help to predict their alteration in the future.

Zooplankton trends in relation to phytoplankton

Relationships between zooplankton and phytoplankton have already been studied for a long time. Phytoplankton, serving as a food for zooplankton organisms, suggest that peaks in zooplankton abundance should coincide with peaks of phytoplankton abundance or immediately follow it. However, the dynamics of these two groups of plankton are usually more complicated. Two main hypotheses were suggested explaining the inverse correlation between phytoplankton and zooplankton. The grazing hypothesis (Harvey et al., 1935) suggests that high concentrations of zooplankton lead to grazing down phytoplankton while low zooplankton concentration let phytoplankton to thrive. The animal exclusion hypothesis (Hardy, 1936) suppose that when concentration of phytoplankton is high, zooplankton just refrain from coming up to the water surface or come up for a shorter time.

Most recent studies usually describe relationships between zooplankton and phytoplankton either by bottom-up or top-down control. Bottom-up control refers to the nutrients, which phytoplankton consumes as food: higher concentrations of nutrients let phytoplankton to thrive and thus increase zooplankton abundance as well. Oppositely, top-down cascades refer to the grazing of phytoplankton by higher organisms in the food chain.

Most studies examining top-down cascades involve freshwater ecosystems. The reason could be historical – lake studies usually were performed to learn about the regulation of algal blooms induced by cultural eutrophication (Hessen and Kaartvedt, 2014). For example, it was shown, that in Malta Lake (artificial lake in Poland) the abundance of nanophytoplankton was limited by zooplankton grazing, whereas the development of large phytoplanktonic organisms was stimulated (Kozak and Gołdyn, 2004). However, this tendency does not always apply: in a highly eutrophic lake only nanophytoplankton species from certain taxonomic groups are suppressed by zooplankton (Gołdyn and Kowalczewska-Madura, 2008).

The studies in marine ecosystems are more oriented to fish yield and consequently, to bottom-up processes (Hessen and Kaartvedt, 2014). It was shown that changes in phytoplankton communities can affect the fish yield (Daniel M. Ware and Richard E. Thomson, 2005; Frederiksen *et al.*, 2006). Originally it was assumed that in contrast to freshwater plankton, top-down cascades are weak in marine ecosystems (Shurin *et al.*, 2002). However, later it was demonstrated, that these relationships exist, and predators have an influence on copepods abundance, while copepods itself reduce large phytoplankton species but simultaneously promote small algae (Stibor *et al.*, 2004). The difference of marine ecosystems to lake trophic cascades can be explained by the fact, that dominant zooplankton in lakes are cladocerans while in marine ecosystems copepods are the most abundant organisms. Cladocerans suppress small phytoplankton, while copepods suppress large

ones (Sommer and Sommer, 2006). Even more, despite phytoplankton showing very few links to the physical environment, zooplankton have positive relationship with phytoplankton throughout the North Atlantic, which suggests that studying marine ecosystems by the bottom-up approach alone might be an over-simplification (Allen, 2019).

In summary, phytoplankton and zooplankton dynamics are closely connected. Both positive and negative correlation are observed between abundances of these two plankton groups. Although the relationship of phytoplankton and zooplankton is subject of interest for a long time, there are only few studies that analyse their direct relationship.

Plankton in the Baltic Sea

The Baltic Sea is the largest inland brackish sea in the world. It is connected to the North Sea and the Atlantic Ocean only via Skagerrak and Kattegat straits. The salinity of the water in the Baltic Sea depends on the introduction of saline marine water from the North Sea (usually during winter storms) and freshwater supply from the rivers (Graham and Bergström, 2001). Therefore, a strong marine salinity gradient is characteristic to it, changing from highly saline water near the traits to near freshwater in the innermost parts (Snoeijs-Leijonmalm, Schubert and Radziejewska, 2017). The climate change has caused significant changes in the hydrographic properties of the Baltic Sea: in certain areas water salinity has decreased due to increased influx of freshwater (HELCOM, 2017b), surface water has warmed in all seasons since 1985 (HELCOM, 2013) and the number of anoxic areas has significantly increased (Hansson, Andersson and Axe, 2011). These characteristics encouraged scientists to explore the dynamics of phytoplankton and zooplankton in the Baltic Sea.

Dominant phytoplankton in the Baltic Sea are diatoms, dinoflagellates and cyanobacteria. The study of the long-term data from the Baltic Sea has revealed a relationship between water temperature and phytoplankton composition - mild winters were associated with a decrease in diatom growth afterwards and allowed motile dinoflagellates to bloom instead (Wasmund, Nausch and Feistel, 2013). Similarly, it was demonstrated that changes in water temperature are associated with earlier spring blooms of diatoms and dinoflagellates and decreased diatom bloom magnitude (Hjerne *et al.*, 2019). Recent research showed that the evidence of summer phytoplankton biomass relation to regional climate is relatively weak, and emphasized the importance of local environmental conditions (Griffiths *et al.*, 2020).

A recent experimental study has shown that most common Baltic Sea benthic organisms prefer diatoms over cyanobacteria (Hedberg *et al.*, 2020), however lately less diatoms are reaching benthic fauna while cyanobacteria abundance trends are increasing. This could cause changes in zooplankton mean size and total stock. An evaluation of these properties in the Northern Baltic Sea

has shown that in some areas zooplankton mean size and/or total biomass have declined during the last decades (HELCOM, 2018). Hydrographic environment is also important for zooplankton: previous studies focusing on the Southern (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019) and Southeastern (Aleksandrov, Zhigalova and Zezera, 2009) Baltic Sea have demonstrated that zooplankton dynamics are strongly associated with water temperature and salinity.

In summary, the Baltic Sea has a long history of investigations of its response to climate change. Being the largest inland sea and having low salinity compared to many other seas makes it unique and attracts the attention of researchers, who investigate the changes in hydrographical variables or dynamics of phytoplankton or zooplankton over time. However, not much is known about the zooplankton dynamics in the Western Baltic Sea. The fact that plankton communities are highly dependent on the local environment suggests that the time series analysis of the Western Baltic Sea are also worth deeper investigation.

METHODS

Data

The data was collected at station TF-0012 (Mecklenburg Bight) which is in the Belt Sea (Figure 2) and provided by The Leibniz Institute for Baltic Sea Research Warnemünde, Germany. Samples were collected and analysed according to the HELCOM COMBINE Baltic Sea monitoring guidelines (Helcom, 2017). All datasets were provided as Excel documents.

Hydrography data

Hydrography data covered the period between 1998 and 2018. A list of measured variables including their explanation is provided in the Table 2.

Table 2. Hydrography variables.

Variable	Unit	Designation
NTRA	mmol/m^3	Dissolved nitrate molar concentration
OXY	ml/L	Oxygen concentration
PHOS	mmol/m^3	Dissolved inorganic phosphate molar concentration
PRES	dbar	Pressure at which samples were collected
SAL	PSU	Salinity
SLCA	mmol/m^3	Dissolved silicate molar concentration
TEMP	^{o}C	Temperature

The surface and the bottom water differs in its hydrographic properties (e.g. sea surface water temperature (SST) can be higher than the bottom water temperature). Also, the surface water has a higher seasonal variability than the bottom water. For this reason, vertical structuring has to be taken into account, when long-term variation in the plankton abundance is studied. However, the zooplankton data is vertically integrated, thus any high vertical resolution in environmental factors cannot be linked to zooplankton responses. Therefore, it was decided to create two separate data frames: one representing surface water (PRES ϵ [0, 4]) and one representing deep water (PRES ϵ [15, 20]). The average values of variables in the specified depth on the specific day were used.

Phytoplankton biomass time series

Phytoplankton biomass time series covering the period between 1980 and 2016 contained 515 phytoplankton taxa assigned to the classes Bacillariophyceae, Chlorophyceae, Choanoflagellatea, Chrysophyceae, Ciliophora, Cryptophyceae, Cyanophyceae, Dictyochophyceae, Dinophyceae, Ebriophyceae, Euglenophyceae, Incertae sedis, Prasinophyceae, Prymnesiophyceae and Trebouxiophyceae. For diverse classes such as Cyanophyceae and Dinophyceae, taxa assigned to

orders were provided, because orders represent deeper taxonomic groups which use different strategies to generate energy (autotrophic, mixotrophic and heterotrophic taxa). Also, a group of "others" was introduced, where unidentified phytoplankton as well as phytoplankton not belonging to any of the previously mentioned classes were assigned. Wet weight (µg/L) was provided as phytoplankton biomass parameter. Before transferring the information into Python data frames, the area of the sampling was filtered according to the coordinates (lat. [54.3063, 54.3198], long. [11.55, 11.5567]) (TF-0012 station, Mecklenburg Bight). Two separate data frames (one for phytoplankton classes and one for Cyanobacteria and Dinophyceae orders) were created, where each column represented class or order respectively and rows represented samples, collected at a specific point in time.

Zooplankton abundance time series

Zooplankton abundance (counts/m³) data from TF-0012 station (Mecklenburg Bight) covered the period between 1998 and 2018. The data was structured into three main groups: copepods, cladocerans and "others" (Table 3).

Table 3 Zooplankton taxa and groups kept for further analysis. The groups labelled as removed were not included into downstream analysis due to infrequent detection. Abbreviations: L-larvae, N-nauplii, T-trochophore.

Copepods
Kept for the analysis: *Acartia* spp. (*Acartia bifilosa*, *Acartia longiremis*, *Acartia tonsa*); Centropages spp. (*Centropages hamatus*, Centropages nauplii, juvenilles); *Eurytemora* spp. (*Eurytemora affinis*); Harpacticoidea; *Microsetella*; *Oithona similis*; *Pseudo/Paracalanus* spp. (*Paracalanus parvus*, *Pseudocalanus* spp.); *Temora longicornis*; others. Removed: *Calanus* spp.; *Centropages typicus* (adults), *Euterpina*; *Limnocalanus macurus*; *Longipedia*.
Cladocerans
Kept for the analysis: *Bosmina* spp*;* Cladocera*; Evadne nordmanni; Podon intermedius; Podon leuckartii; Pleopsis polyphemoides.* Removed: *Cercopagis pengoi; Evadne anonyx; Evadne spinifera;* others.
Others
Kept for the analysis: *Alaurina;* Anthoathecatae*; Asterias; Balanus* N+L*; Bivalvia* L*; Carcimus; Crangon; Ctenophora; Euphysa; Fritellaria; Gastropoda* L*; Gymnoleata; Oikopleura; Pectenariidae;* Polychaeta (*Harmothoe*, Polychaeta L, Polychaeta T, *Spionidae* L)*; Sagittidae; Synchaeta; Teleostei; Tintinnida;* others. Removed: *Aurelia; Diastylis; Facetotecta* L*;* Gammaridea*; Keratella; Mysidae; Nematoda; Ostracoda;; Phoronis; Rathkea; Sagittidae; Sarsia; Trematoda; Trochophora; Turbellaria.*

Not all zooplankton data was analysed further. Some of the species were only sparsely detected and therefore are not suitable for analysing long term trends in their abundance. Hence, only the taxa, containing at least five measurements in the same month per time series were included in the further analysis. Since most often samples were not collected more than once per month, this cut-off also excluded taxa which were not detected at least in five distinct years. After applying this threshold,

Cercopagis pengoi, *Evadne anonyx*, *Evadne spinifera* and "others" were removed from the cladocerans group. In the plankton group of species, not belonging to cladocerans or copepods ("others") nearly half of the given taxa did not meet this criterion, therefore fifteen taxa (Table 3) were not analysed further. *Polychaeta* class larvae (*Harmothoe*, Polychaeta L, Polychaeta T, *Spionidae* L) were summed up and further analysed as continuous Polychaeta class. In the copepods group *Calanus* spp.; *Centropages typicus* (adults), *Euterpina*; *Limnocalanus macurus* and *Longipedia* were not detected often enough to be kept for further analysis.

For the copepods the stage of life (N - nauplii, J1 (juveniles 1'st stage), J2 (juveniles 2'nd stage) and sex for the adults (females, males) were provided. Usually for nauplii and juveniles, distinct species cannot be distinguished from each other. Therefore nauplii and juveniles were assigned to a specific genus, but not specific species.

Software

The data was provided as Excel documents (.xlsx format). For further analysis Jupyter notebooks (Version of the notebook server - v. 6.0.0) were used running a Python (v. 3.7.3) kernel. All data was imported into Pandas v. 0.24.2 (The pandas development team, 2020) data frames. For the calculations and plots other Python packages and libraries, such as Matplotlib v. 3.2.1 (Hunter, 2007), Numpy v. 1.18.4 (Harris *et al.*, 2020), Seaborn v. 0.9.0 (Waskom *et al.*, 2017), Scipy v. 1.4.1 (Virtanen *et al.*, 2020) also were used.

Data preprocessing

Python modules (Appendix 3.1, Appendix 3.2) were written to import the data from the Excel files so that all data frames would be represented in a similar structure. Several issues, common to time series analysis, such as unevenly spaced data, missing values and seasonality in time series were dealt with before applying statistical methods to it. Main data preprocessing steps and created new data frames are represented in Figure 5, details of which will be presented later in this section.
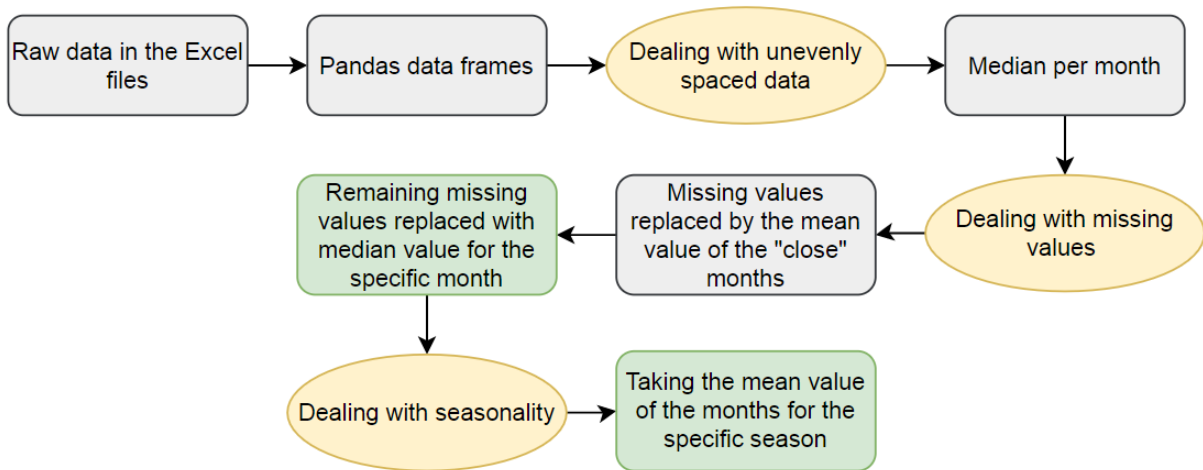
Figure 5 Data preprocessing steps and created data frames. Yellow ovals indicate the challenges during data processing, intermediate data frames are shown in grey, green rectangles represent data frames used for the statistical analysis.

Outliers in the data

For each variable (plankton abundance/biomass, hydrography variables) box plots were created (Appendix 3.3), representing the range of variable values for each month (e.g. Supplementary Figure 1). There are some outliers in the raw data, however, it was decided to keep them, since it was assumed, that they did not appear due to sensor errors or human mistakes (data was validated) and are important to describe the variability of the system.

Dealing with unevenly spaced data

All time series (hydrography variables, phytoplankton and zooplankton abundance) were unevenly spaced. From the bar plot (Supplementary Figure 2) it can be seen, that hydrography samples were not collected in June, September and December and only very few zooplankton samples were collected during these months.

To obtain evenly spaced data frames were reshaped so that each row would represent the median value for the particular month (Appendix 3.4). The example of the formed monthly data frame is represented in Table 4.

Table 4 Hydrography variables data frame (1998-2018) where each row represents one month. For clarity, only a small subset of the data is shown.

Year	Month	TEMP	OXY	SAL	NTRA	PHOS	SLCA
	1	4.101	7.877	14.902	4.670	0.670	16.400
	2	2.580	7.358	16.897	6.590	0.670	16.000
	3	3.659	9.008	10.425	3.180	0.300	13.100
1998	...	...	...	...	...	...	...
	10	10.252	6.790	17.135	1.360	0.450	17.300
	11	NA	NA	NA	NA	NA	NA
	12	NA	NA	NA	NA	NA	NA
	1	2.752	8.070	13.933	6.800	0.580	15.200
1999	2	2.465	8.348	14.789	7.860	0.600	16.900
	3	3.874	9.767	8.529	4.050	0.390	13.300
	...	...	...	...	...	...	...

Dealing with missing values

Missing values in the time series were replaced by the mean value of the two closest months (e.g. mean of May and July for June) using a moving average function (window size – 3, centre - true). Remaining missing values were replaced by the median value of specific month (Appendix 3.5).

Detecting and removing seasonality

For the Mann-Kendall test (see p. 25) data cannot be seasonal. To check for seasonality of the data, autocorrelation plots were used ("autocorrelation_plot" function from Pandas (Appendix 3.3)).To avoid seasonality in time series the mean value for the months of the specific season (Table 1 (Belt Sea)) was calculated for each variable using monthly data (Dealing with unevenly spaced data p. 23) with filled missing values (Dealing with missing values p. 23). Newly created data frames (e.g. Table 5) were used to perform Mann-Kendall test on each season time series separately.

Table 5 Data frame representing hydrography variables for spring season (1998-2016). The mean value of the hydrography variables in February, March and April is calculated per year.

Year	TEMP	OXY	SAL	NTRA	PHOS	SLCA
1998	4.277	8.290	12.388	3.797	0.378	12.617
1999	4.287	8.988	10.688	4.673	0.400	13.475
2000	4.152	9.093	12.523	2.680	0.322	10.283
...	...	...	...	...	...	...
2014	4.878	8.351	13.672	1.737	0.313	7.675
2015	5.009	8.401	12.079	1.878	0.275	7.925
2016	5.680	8.418	12.269	0.947	0.250	4.950

Statistical methods

Kruskal-Wallis test and Post hoc Dunn's test

The Kruskal-Wallis test was used to find out if median of phytoplankton biomass or zooplankton abundance is statistically significantly different on any of the seasons. To perform this test SciPy "stats.kruskal" function was used. Dunn's multiple comparison post-hoc test (scikit-posthocs v. 0.6.7 (Terpilowski, 2019) Python package) was used to determine which seasons differ from each other season. Both tests were performed with a chosen significance level of $p = 0.05$.

The results of these tests were used to find seasons when species are most abundant (Appendix 3.6). For example, species are considered as "spring" species not only if the median value of their abundance is the highest in spring, but also if the median value in spring is not statistically significantly different from the season with highest median abundance value).

Non-metric multidimensional scaling analysis

The non-metric multidimensional scaling (nMDS) was used to identify the most similar years and those which show the largest difference comparing hydrography variables, phytoplankton biomass data and zooplankton abundance data (Appendix 3.7). Firstly, distances ($dist_1, dist_2, \ldots dist_n$) between each variable for every pair of years were calculated using the dynamic time warping (DTW) technique. For this purpose, the DTAIDistance v. 2.0.0 library was used. The general distance between each pair of years was calculated using the Euclidian distance formula (Equation 2). These distances were used to fill a symmetric distance matrix representing distances between years (Figure 6). Finally, using Scikit-Learn's "Manifold.MDS", nMDS plots were created. Stress values are used to represent the difference between distances in the two-dimensional space compared to the complete multidimensional space. Stress values calculated using the "Manifold.MDS" were not normalized, therefore Kruskal's Stress-1 formula (Equation 3) was used to calculate normalized stress values.

Equation 2. Euclidian distance formula.

$$General\ distance = \sqrt[2]{dist_1^2 + dist_2^2 + \cdots + dist_n^2}$$
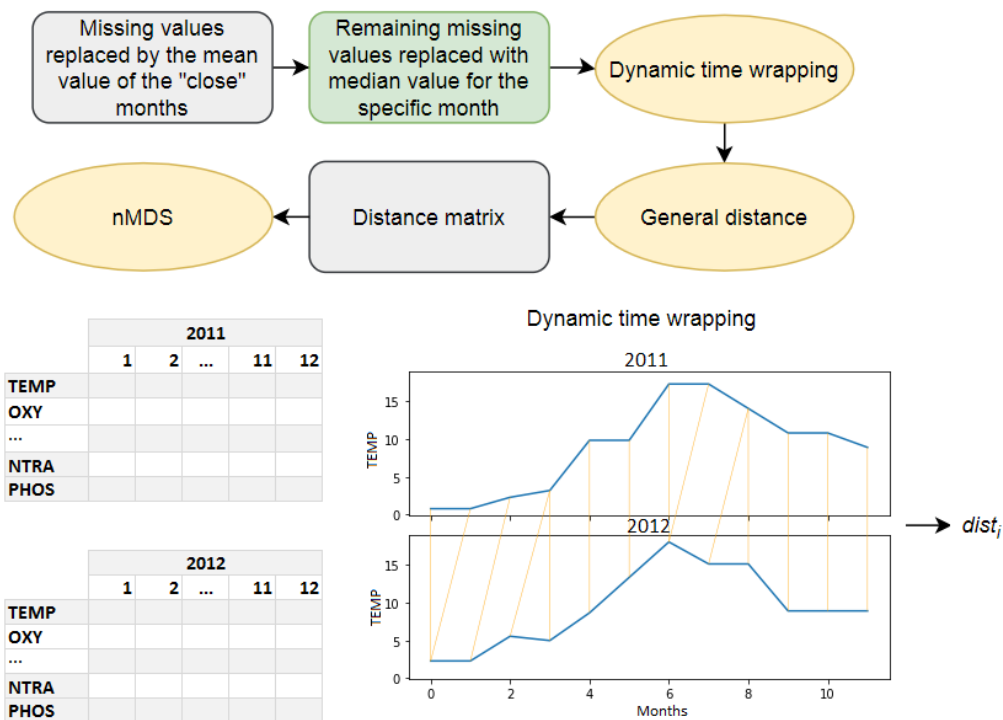
Figure 6 Creation of the distance matrix for the nMDS analysis (p. 24).

Equation 3. Kruskal's Stress-1 definition. δ_{ij} refers to the Euclidean distance, across all dimensions, between points *i* and *j*, d_{ij} represents the disparities that are the result from the transformation of the dissimilarities.

$$Stress = \sqrt{\frac{\sum(d_{ij} - \delta_{ij})^2}{\sum d_{ij}^2}}$$

Hierarchical clustering

Hierarchical cluster analysis was used to illustrate the most distant points in two-dimensional graphs received by nMDS. The single linkage clustering was applied using SciPy's "cluster.hierarchy" (Appendix 3.7).

Mann-Kendall test

pyMannKendall v. 1.4.1 (Hussain and Mahmud, 2019) package was used to perform a Mann-Kendall test (chosen significance level p = 0.05) (Appendix 3.8). This test was used to identify monotonic upward or downward trends in the time series. The Mann-Kendall test can be applied to the data,

which is not seasonal (see Detecting and removing seasonality, p. 23) and has only one data point per time period (see Dealing with unevenly spaced data p. 22) (Glen, 2016).

Calculating correlation

The "spearmanr" function from SciPy was used to calculate the Spearman correlation coefficient with a chosen significance level of $p = 0.05$ (Appendix 3.8).

RESULTS

Influence of short-term changes in environmental conditions on plankton communities

Plankton abundance patterns were scanned for differences to elucidate the effect of short term changes in environmental conditions on plankton communities. For that purpose, nMDS (p. 24) was performed. Even though the stress values > 0.2 indicate that the representation of the distances between years in 2-dimensional graphs are not perfect, it still can be noticed, that 2013 shows a large difference from the other years in phytoplankton classes and orders as well as in zooplankton cladocerans and copepods groups abundance data (Figure 7).

No significant differences between 2013 and the other years were noticed with respect to the median values of the hydrography variables (Supplementary Table 1). However, comparing median values for each month with multiannual median values, significant differences in five out of six studied hydrography variables became apparent for March 2013 (Table 6). Moreover, it was shown, that for twelve out of sixteen studied phytoplankton classes, biomass in March 2013 was lower than usually in March during the studied period (Supplementary Table 2). The general phytoplankton biomass also was lower in March of 2013 (386.026 µg/L) than usual (median value 987.335 (503.866-1842.980) µg/L) for March 1998-2016). Analysis of the copepods species, for which the stage of development was also provided, have shown, that for most species nauplii abundance was lower and adults' abundance was higher in March of 2013 than other years from the period between 1998 and 2018 (Table 7).

Table 6 Hydrography variables in March 2013 compared to March in 1998-2018. Only oxygen concentration (OXY) does not differ significantly from the concentrations detected in other years.

Variable	March 2013	Median (March 1998-2018)	IQR(March 1998-2018)
NTRA	4.155	0.280	0.070 - 1.572
OXY	9.288	9.008	8.690 - 9.563
PHOS	0.620	0.180	0.120 - 0.360
SAL	9.418	11.55	10.13 - 12.17
SLCA	17.50	7.500	3.000 - 12.75
TEMP	0.774	3.874	2.822 - 4.379

Table 7 Copepods abundance in March 2013 compared to March during 1998-2018. Median values with IQRs provided in Supplementary Table 3.

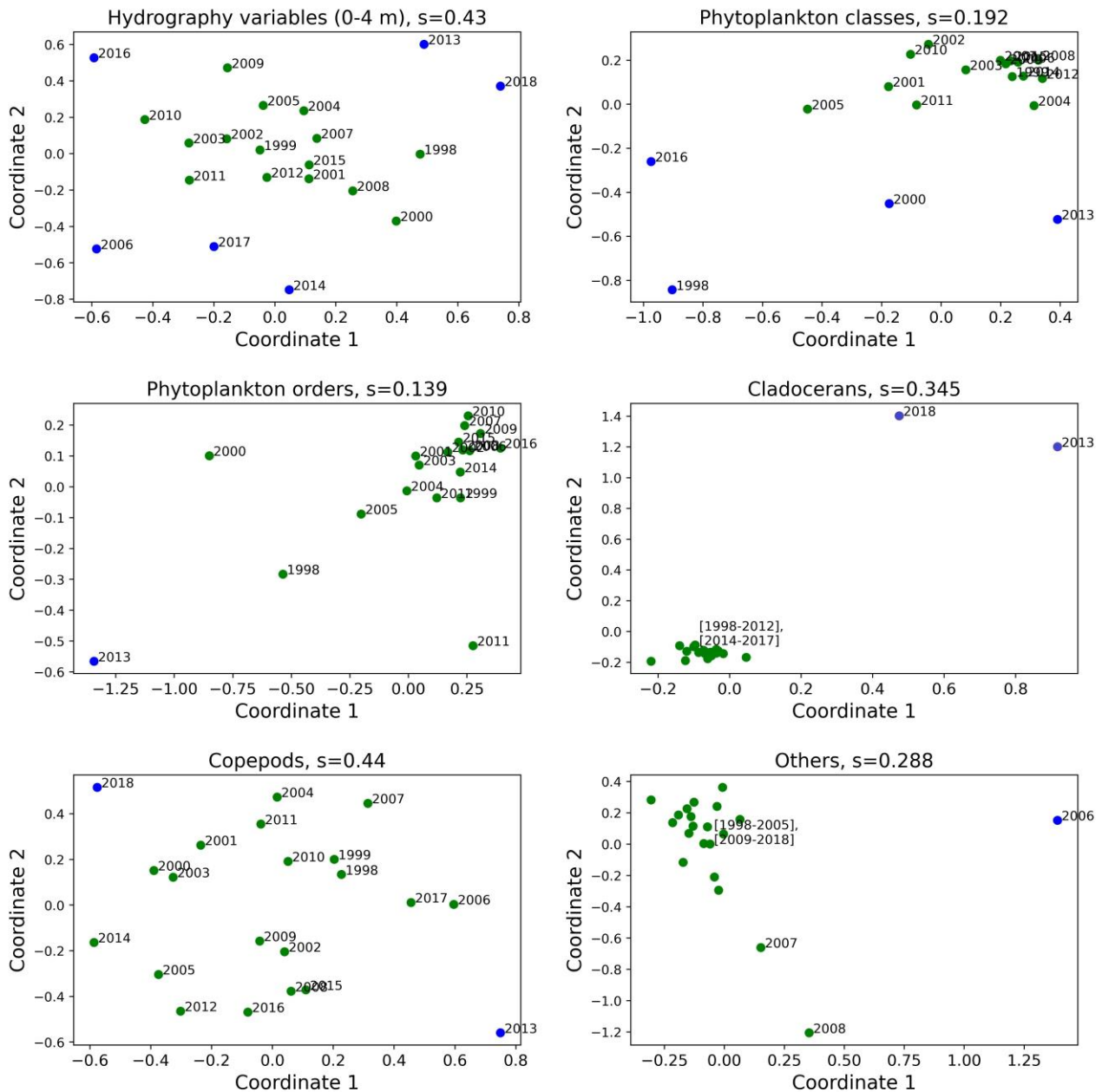| Copepods species	Nauplii	Adults
Acartia spp.	less	more
Calanus spp.	more	-
Centropages spp.	less	more
Oithona similis	less	-
Pseudo/Paracalanus spp.	more	more
Temora longicornis	less	more

Figure 7 Results of the nMDS analysis. Biomass patterns of phytoplankton classes and orders as well as abundance patterns of copepods and cladocerans groups were different in 2013 compared to other years. Plot colours represent hierarchical clustering results (p. 25). s – stress values.

Long term shifts in plankton communities and changes in environmental conditions

Exploratory data analysis has shown, that data is seasonal. For example, autocorrelation plots demonstrate a yearly seasonality in hydrography data (Supplementary Figure 3). Some of the zooplankton species are more abundant in spring, while others appear more frequent in autumn (Supplementary Figure 4). As expected, the highest phytoplankton biomass values are present during spring and autumn blooms. Moreover, it could be shown that "spring" species (see Kruskal-Wallis test and Post hoc Dunn's test, p. 24) show higher correlation to hydrography variables in spring than to the complete time series (). Altogether, including the results from nMDS analysis suggest, that trends of plankton abundance and changes in hydrographic conditions should be analysed separately for each season.

The results of the Mann-Kendall test have shown, that in sea surface (0-4 m) water nitrates concentration is decreasing in spring, water temperature is increasing in summer and water salinity along with nitrates and inorganic phosphates concentration is decreasing in winter. Deep water (15-20 m) hydrography variables did not show any significant monotonic trends in the studied period.

Significant monotonic shifting trends of phytoplankton classes and orders are represented in Table 8. For seasons with detected trends, the Spearman correlation coefficient to hydrography variables was calculated. In some cases, correlation values between hydrography data and phytoplankton data showed coincidence with trends of hydrography variables. For example, in summer water temperature is increasing and negative correlation exists between Chrysophyceae class biomass and surface water temperature. Therefore, it is possible, that due to increasing water temperature, biomass values of Chrysophyceae class are decreasing.

Table 8 Phytoplankton classes and orders with significant monotonic increasing or decreasing trends which were detected at least in one of the seasons. Spearman correlation coefficient shown for hydrography variables in seasons with detectable trends. Spr – spring, Sum – summer, Aut – autumn, Win – winter.

	Decreasing	Increasing	Correlation
Phytoplankton classes			
Bacillariophyceae	Sum	-	Sum: SLCA -0.6316
Chrysophyceae	Sum	-	Sum: **TEMP -0.4596**
Cryptophyceae	Spr, Sum, Aut, Win	-	Sum: SLCA -0.714, **TEMP -0.5772**
Dictyochophyceae	-	Win	Win: **PHOS -0.5178**
Euglenophyceae	Sum	-	Sum: SLCA -0.6614
Prymnesiophyceae	-	Sum	-
Trebouxiophyceae	-	Sum	-
Cyanophyceae class orders			
Chroococcales	-	Spr, Aut, Win	Spr: **NTRA 0.5316**
Nostocales	-	Spr, Aut, Win	Aut: SAL -0.4827
Oscillatoriales	-	Spr, Sum, Aut	-
Dinophyceae class orders			
Dinophysales, AU	-	Spr, Sum, Win	-
Gonyaulacales, AU	Aut	Spr	-
Gymnodiniales, AU	-	Aut, Win	-
Gymnodiniales, HT	-	Spr, Aut, Win	-
Peridiniales, AU	-	Spr, Sum, Aut	-
Peridiniales, HT	Sum	-	-
Prorocentrales, AU	Aut	Spr	-

Similarly, zooplankton trends in relation to environmental conditions are described in Table 9 (cladocerans, others) and Table 10 (Copepods). Most of the zooplankton taxa are showing increasing trends in their abundance. For the copepods, the stage of development and sex for the adults is also provided. Copepods nauplii, juveniles and adults could not be analysed together as one continuous group, because they may show different preferences for environmental conditions at different stages of development. Moreover, copepods males and females sometimes show different trends (Table 10), and analysing all adults together these differences can be overlooked.

Since copepods are the dominating zooplankton group in the Baltic Sea (Dzierzbicka-Glowacka *et al.*, 2018), the relationship between copepods abundance and phytoplankton biomass was investigated as well. In some cases, it could be shown that even when no correlation to hydrographic variables was found (e.g. *Eurytemora* spp.), correlation to at least to one of phytoplankton classes' biomass was detected (Supplementary Table 5).

Table 9 Zooplankton taxa (belonging to cladocerans and "others" groups), with significant monotonic increasing or decreasing trends, detected at least in one of the seasons. Spearman correlation coefficient shown for hydrography variables in seasons with detectable trends.

	Decreasing	Increasing	Correlation
Cladocerans			
Bosmina spp	-	Aut	
Cladocera	Spr, Aut	-	Spr: **NTRA 0.4657** Aut: NTRA -0.6028
Pleopsis polyphemoides	-	Sum, Aut, Win	Aut: SLCA -0.4764 Win: **PHOS -0.5444**
Podon intermedius	-	Sum, Aut, Win	Sum: **TEMP 0.4722**
Podon leuckartii	-	Spr, Sum, Win	Spr: **NTRA -0.4678**
Others			
Anthoathecatae	-	Sum, Aut	
Asterias	-	Spr, Sum, Aut, Win	Spr: TEMP 0.506
Carcimus	-	Sum, Aut	
Crangon	-	Sum	
Ctenophora	-	Spr, Win	
Euphysa	-	Sum, Aut	
Pectenariidae	-	Sum, Aut, Win	Win: **NTRA -0.4969**
Polychaeta	-	Spr	
Sagittidae	-	Spr, Sum, Aut, Win	
Synchaeta	Sum	-	Sum: **TEMP -0.4895**
Teleostei	-	Spr, Sum, Aut, Win	
Tintinnida	Spr	-	
others	-	Spr, Sum, Aut, Win	

Table 10 Copepods species with significant monotonic increasing or decreasing trends, detected at least in one of the seasons. Spearman correlation coefficient shown for hydrography variables in seasons with detectable trends. F – females, M - males

	Decreasing	Increasing	Correlation
Acartia spp.			
Acartia indeterminate	-	Spr, Sum, Aut, Win	Sum: **TEMP 0.4653** Aut: TEMP 0.517
A. bifilosa F	Aut	-	
A. tonsa F	-	Aut, Win	Aut: TEMP 0.6346
A. tonsa M	-	Win	Win: **PHOS -0.4916**
Calanus spp.			
Nauplii	-	Spr, Sum, Aut, Win	Spr: SAL 0.4905
Centropages spp.			
J1	-	Spr	
J2	-	Spr	Spr: SAL -0.5982
C. hamatus F	Win	-	
Eurytemora spp.			
Nauplii	-	Sum, Aut	
J2	-	Aut	
E. affinis M	-	Sum, Aut	
Others			
Harpacticoidea	Spr	-	
Microsetella	-	Spr, Sum, Aut, Win	
Pseudo(Para)calanus spp.			
Nauplii	Win	-	Win: **PHOS 0.5947, SAL 0.4842**
J2	Spr, Win	-	Win: OXY -0.5702, **PHOS 0.6105**, SLCA 0.5088
Par. parvus F	-	Spr, Sum, Aut, Win	
Par. parvus M	-	Sum, Aut	
Temora longicornis			
J1	-	Spr	Spr: OXY -0.5965, TEMP 0.6807
J2	-	Win	Win: **NTRA -0.6982, SAL -0.4561**
T. longicornis F	Aut	-	
T. longicornis M	-	Spr	

DISCUSSION

In this study, data collected in 1998-2016 was analysed to detect long-term shifts in the Western Baltic Sea (Mecklenburg Bight) ecosystem. Additionally, the influence of short-term anomalies in hydrographic conditions on plankton communities was analysed. It was shown that the plankton abundance patterns in 2013 were significantly different from other years in the studied period (Figure 7, Table 7, Supplementary Table 2). Moreover, long-term (1998-2016) time series analysis revealed that both increasing and decreasing trends in phytoplankton biomass or zooplankton abundance could be detected depending on the analysed taxonomic group. Calculated Spearman correlation coefficients suggest an existing relationship between organisms in plankton groups itself (Supplementary Table 5) as well as associate plankton abundance to hydrographic conditions (Table 8, Table 9). However, observational analysis and used statistical tests results are not sufficient to explain why detected trends are present. Thus, I discuss here how this study results are similar or different to the previous studies and provide possible ecologic explanation why particular trends were detected and what the reason behind the exceptionality of the year 2013 is.

Influence of environmental conditions in 2013 on plankton communities

nMDS analysis showed that values of hydrography variables in 2013 do not differ a lot from the other years, however there is a striking difference in the plankton communities' abundance and biomass patterns in 2013 compared to all other years in the studied period. Since values of hydrography variables are quite continuous over the year, short-term anomalies do not lead to high distance values between years. Additionally, the median values of the hydrography variables for the 2013 does not differ significantly from overall median values, because the higher values in some months compensate the lower values in other months (e.g., higher median temperature in August and September and lower temperature than usual in March compensate each other and lead to not significantly different overall median temperature value in 2013 (Supplementary Table 1, Supplementary Table 6)). In contrast, plankton blooms occur only in specific months and if there are changes in plankton bloom season, they can be easily detected.

Checking the median values for the specific months, it was found that March in 2013 SST was much lower than usual (Table 6). This corresponds to multiple reports from different areas in Europe, indicating that March 2013 was exceptionally cold, in some areas being the second coldest March over the hundred years period (Andrews, 2013; BBC, 2013; World Meteorological Organization, 2014). This detected short-term change in 2013 had significant influence on plankton communities. Lower phytoplankton biomass values in March (Supplementary Table 2) suggest that the spring bloom probably started later. Interestingly, less than usual copepods nauplii were detected in March

2013, but adults' abundance was higher (Table 7). This can be explained by the fact, that if the conditions for zooplankton development are not favourable, they development to nauplius stage can be delayed and copepods can exist in the resting egg stage (M. Engel, 2013). Moreover, the animal exclusion hypothesis (Hardy, 1936) suggest, that the lower density of phytoplankton in the surface water might lead to zooplankton coming to feed on the surface of water more frequently, or they might stay there longer, to get the same amount of energy. For this reason, more zooplankton might get caught using a mesh and this would lead to report of higher zooplankton abundance.

To summarize, these findings show that plankton communities react to temporal changes in their environment. The largest difference in plankton abundance and biomass patterns was noticed in 2013, when March was exceptionally cold, because plankton is most sensitive to changes in the environment occurring in bloom season. Also, the fact that plankton react to short-term changes in their environment, supports the hypothesis that slight changes in hydrographic conditions in a long-term might also have a distinct effect.

Changes in hydrographic conditions in the Baltic Sea

As just discussed, plankton species show clear seasonality in their abundance and are sensitive to changes in hydrographic conditions at the period of their appearance. Therefore, the analysis of shifts in hydrographic conditions and the ecosystem was performed for each season separately. Indicated trends in hydrographic conditions in the Western Baltic Sea (Mecklenburg Bight) showed that SST in summer is increasing while no significant trend was detected in other seasons (p. 29). This corresponds to the findings of other researchers who have shown, that SST trends are strongest in summer months, especially in the Southern Baltic Sea, while in the Northern Baltic Sea warming trends were most obvious in winter and autumn (Siegel, Gerth and Tschersich, 2006; Lehmann, Getzlaff and Harlaß, 2011; Reckermann *et al.*, 2014; Stramska and Białogrodzka, 2015). Several possible reasons for water temperature increase in summer were summarised by Kniebusch *et al.*, (2019), including increasing warm summer inflow events bringing warm surface water from the North Sea to the Baltic Sea (Leppäranta and Myrberg, 2009; BACC and Team, 2015), higher absorption due to increased turbidity (cloudiness) at water surface as well as increased length of the sea surface water warming period.

The Mann-Kendall test has shown, that there is significant decreasing trend of sea surface water salinity in winter (p. 29). Several previous studies expected the Baltic Sea water to get less saline, because mild winters might lead to less snow and ice-cover and thus to the greater fresh water run-off to the Baltic Sea (Störmer, 2011; Gräwe, Friedland and Burchard, 2013). Moreover, it was shown, that during the second half of the 20th century, the precipitation was increasing in the Baltic Sea area (Reckermann *et al.*, 2014) which could also have led to freshening of the Baltic Sea water.

Finally, decreasing concentrations of inorganic phosphates in winter and nitrates in winter and spring were detected in the studied area. Previous studies of nutrients in the Baltic Sea have shown that both, inorganic phosphorus and nitrogen compounds concentrations in the Southwestern Baltic Sea are decreasing (based on long time series (> 10 years)), while in other areas of the Baltic Sea increasing trends were noticed (European Environment Agency, 2019). Assumably, the concentrations of nutrients are reduced due to reduced loads from the influx of rivers (European Environment Agency, 2015).

To sum up, SST trends in the Mecklenburg Bight are similar as in southern parts of the Baltic Sea, however the seasonality differs from the trends in the Northern Baltic Sea. Similarly, findings about nutrient concentrations correspond to the previous studies with focus on nutrients near the coasts of the Southwestern Baltic Sea but not other parts of the Baltic Sea. This suggests, that trends of phytoplankton biomass and zooplankton abundance in the Western Baltic Sea might differ from the other areas in the Baltic Sea.

Phytoplankton trends in 1998-2016 Mecklenburg Bight

The most abundant and therefore most analysed phytoplankton classes in the Baltic Sea in previous studies were Bacillariophycea (diatoms), Dinophyceae (belongs to Dinoflagellata superclass) and Cyanophyceae (Cyanobacteria). In this study, trends of less abundant phytoplankton classes' biomass were also investigated. Below it is discussed how detected increasing or decreasing trends for distinct phytoplankton classes correspond to the results from studies in other areas of the Baltic Sea as well as older studies from the Western Baltic Sea (analysing trends in (1979–1999) and (1979–2005) (Wasmund and Uhlig, 2003; Wasmund *et al.*, 2011)) along with ecologic results interpretation.

Previous studies of the diatoms (Bacillariophycea) biomass in the Western Baltic Sea have shown that their concentration from 1979 to 2005 decreased in spring (Wasmund and Uhlig, 2003; Wasmund *et al.*, 2011) most likely due to competing for food sources with increasing concentrations of more motile dinoflagellates after mild winters (Wasmund, Nausch and Feistel, 2013). It was also shown that diatoms growth can be estimated by silicates consumption (Wasmund, Nausch and Feistel, 2013). Similarly, negative correlation between silicates concentration and Bacillariophycea biomass was detected in this study (Table 8), however the decreasing trend was detected not in spring but in summer. In a recently published paper (Griffiths *et al.*, 2020) it was suggested that summer phytoplankton biomass observations and in turn the ability to detect trends are affected by grazing pressure. Most likely, it would have the greatest effect on diatoms which are less common in summer, but are preferred food source for zooplankton (Hedberg *et al.*, 2020). This suggest that

in more recent years phytoplankton biomass is more affected by large shifts in the whole ecosystem than in competition for resources or changes in hydrographic conditions.

Similarly, decreasing biomass trends were also detected for Chrysophyceae and Cryptophyceae classes (Table 8). Interestingly, previous analysis of the Mecklenburg Bight have shown, that Cryptophyceae biomass trends in summer were increasing from 1991 until 2005 (Wasmund and Uhlig, 2003; Wasmund *et al.*, 2011). In a previous experimental study (Wirth, Limberger and Weisse, 2019) it was shown, that grow rates of organisms belonging to Chrysophyceae and Cryptophyceae classes are increasing until the optimum temperature, but starts to decrease if the temperature is too high. In combination with the detection of increasing SST in summer as well as negative correlation between SST in summer and biomass of these two phytoplankton classes, this suggest that in Mecklenburg Bight SST in summer is becoming higher than optimal for these species.

In contrast to previously discussed classes, most orders of Cyanophyceae and Dinophyceae classes show increasing trends (Table 8). In previous studies, analysing the period from 1979 to 2005, no or even decreasing trends of Cyanophyceae and Dinophyceae biomass were detected in Mecklenburg Bight (Wasmund and Uhlig, 2003; Wasmund *et al.*, 2011). The increasing abundance of these classes in other areas of the Baltic Sea is usually associated with mild winters and increasing SST (Hedberg *et al.*, 2020; Wasmund, Nausch and Feistel, 2013). Additionally, increasing trends of Dictyochophyceae biomass were detected (Table 8). The organisms belonging to this class occur exclusively in the Western Baltic Sea and only after mild winters (Wasmund, Nausch and Feistel, 2013).

In summary, a comparison of this study results to the phytoplankton trends detected in 1979-2005 have shown that major shifts in the Mecklenburg Bight, causing increase or decrease of the detected trends, appeared in the last two decades. Ecological interpretation of the results suggest that biomass of phytoplankton classes, which are sensitive to higher water temperatures or the ones which are highly consumed by zooplankton in summer are decreasing. On the other hand, biomass of phytoplankton classes, which prefers mild winters, are increasing in the Western Baltic Sea.

Zooplankton trends in 1998-2016 Mecklenburg Bight

To date, the latest zooplankton trends in the Baltic Sea are not comprehensively analysed. There were some studies analysing long-term zooplankton dynamics in the Southeastern Baltic Sea during 1998–2007 (Aleksandrov, Zhigalova and Zezera, 2009) as well as in the Southern Baltic Sea in the periods from 1969 to 2001 (Feike *et al.*, 2007) and from 2006 to 2012 (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019). In this study, zooplankton trends (1998-2016) in the Western-Baltic Sea (Mecklenburg Bight) were analysed. Below, a possible ecological explanation

for the detected trends together with comparisons to previous studies from other areas of the Baltic Sea is provided.

When zooplankton trends are analysed in relation to hydrographic conditions, usually the most attention is directed to water temperature and salinity, because they are considered as the factors with the strongest effect on zooplankton abundance. Previous studies in the Southern and Southeastern Baltic Sea have shown that increasing water temperature can be associated to increasing trends of thermophilous and eurythermous (able to tolerate a wide range of environmental temperatures), organisms of both cladocerans (*Evadne nordmanni, Bosmina* spp.) and copepods (*T. longicornis*, *C. hamatus*, and *Acartia* species) groups (Möllmann *et al.*, 2002; Aleksandrov, Zhigalova and Zezera, 2009; Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019). Also it was shown that warmer years in the Baltic Sea are associated with higher biomass of copepods and lower biomass of rotifers (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019). The biomass of some of the rotifers, such as *Synchaeta* spp., most likely decreased not because of increasing temperature itself, but due to the presence of more cyclopoid copepods, which consume rotifers as food (Zhang *et al.*, 2015). In addition to that, time series (1960-1997) analysis has shown that abundance of *Podon* spp. and *Evadne nordmanni* were positively associated to salinity in summer and autumn in the Gdansk Deep (Central Baltic Sea) (Möllmann *et al.*, 2002). The decrease of *Pseudocalanus* spp. abundance in the Southeastern Baltic Sea (1990-2000) was associated to the decrease of water salinity (Aleksandrov, Zhigalova and Zezera, 2009).

Similar trends with some discrepancies were detected in the Mecklenburg Bight (Table 9, Table 10). As in other areas of the Baltic Sea, the abundance of thermophilic zooplankton species (*Bosmina* spp., *T. longicornis*, *C. hamatus*, and *Acartia spp.*) is increasing at least at one of the developing stages. Moreover, decreasing abundance trends and negative correlation to the increasing SST for rotifers *Synchaeta* spp. were detected. However, in contrast to previous studies, where increasing trends (1998–2007) of *Bosmina* spp. abundance were detected in summer in the Southeastern Baltic Sea (Aleksandrov, Zhigalova and Zezera, 2009), in this study increasing trends of *Bosmina* spp. abundance were exclusively detected in autumn. Normally, *Bosmina* spp. are most abundant in August and their abundance decrease in autumn, when the water cools down (Aleksandrov, Zhigalova and Zezera, 2009). Increasing trends of *Bosmina* spp. abundance in autumn suggest that the water temperature stays optimal for *Bosmina* spp. for a longer time after summer.

Contrary to the previous studies from more eastern regions of the Baltic Sea, significant trends of *Evadne nordmanni* and *Pseudocalanus* could not be identified. These species were positively associated with water salinity (Möllmann *et al.*, 2002; Aleksandrov, Zhigalova and Zezera, 2009; Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019). In addition, even though positive trends of *Podon* spp. were detected in the Mecklenburg Bight, unlike in the Central Baltic

Sea (Möllmann *et al.*, 2002), it did not show any correlation to water salinity. Keeping in mind, that usually only salinity values beyond the optimal range have an influence on plankton development, it is very likely that salinity is not the limiting factor for the abundance of these species as the western parts of the Baltic Sea are more saline.

In this study, similarly as in the research from the Southern Baltic Sea (Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann, 2019) diverse trends for different plankton stages for both males and females were detected. Musialik-Koszarowska, Dzierzbicka-Głowacka and Weydmann associated trends of copepods stages to different environment preferences of copepods at different developing stages, accelerated copepods' metabolism and life cycle, but did not provide an explanation why copepods males and females trends differ. In this study, the trends *A. bifilosa*, *C. hamatus, T. longicornis* females were decreasing, while for males either no or increasing trends were detected. Also increasing trends of *E. affinis* males, but not females were detected in the Central Baltic Sea (Möllmann *et al.*, 2002). Gusmão and McKinnon (2009) suggested, that since skewed sex ratios and intersexuality are observed in many copepod species, environmental sex determination may be widespread in this group. Their observations suggested that food limitation is an important determinant for sex change. However, skewed sex ratios can also be explained by higher predation on males than females (Hirst *et al.*, 2010). A similar reasoning might explain why copepods females and males of the same species show different trends, but this would require more investigation. Since usually male planktonic copepods are smaller in size than females (Gusmão and McKinnon, 2009), decreasing trends of the abundance of females could lead to decrease of zooplankton mean size or total biomass. This corresponds to the findings of other areas of the Baltic Sea, showing that zooplankton communities become dominated by smaller size copepods (Rice, Dam and Stewart, 2015; Kelly *et al.*, 2016; HELCOM, 2018).

To summarize, zooplankton trends which were detected in the Southern and Southeastern Baltic Sea and associated to increasing water temperature in previous studies, are usually detected in the Western Baltic Sea as well. However, zooplankton abundance trends associated to changes in water salinity do not share similar patterns in the Western and Eastern parts of the Baltic Sea. Furthermore, the analysis of the copepods abundance suggest that copepods at different stages of development and of different adults sex do not prefer the same environmental conditions. However, this study is only observational, therefore further studies will be necessary to explain why these trends are different.

Zooplankton in relation to phytoplankton

As previously discussed, abiotic conditions mostly associated with zooplankton trends are water temperature and salinity. However, in this study correlations between zooplankton abundance and

phosphate and nitrate concentrations were also detected. Assumably, these correlations do not exist due to direct relationship between water nutrients and zooplankton, but because of the linkage through phytoplankton. This idea can be supported by the fact that increasing or decreasing trends for zooplankton species were present, even if no correlation with hydrographic variables could be found. Most often in those cases, correlations with at least one of the phytoplankton classes' biomass were detected (Supplementary Table 5).

However, only few examples can be found, where specific seasonal trends can be explained by correlation between hydrography data and phytoplankton trends and correlation between zooplankton abundance and phytoplankton biomass (Figure 8). The lack of such examples suggest that analysing plankton communities based solely on a bottom-up approach is an oversimplification and experimental studies are required to explain the causality of the detected relationships.
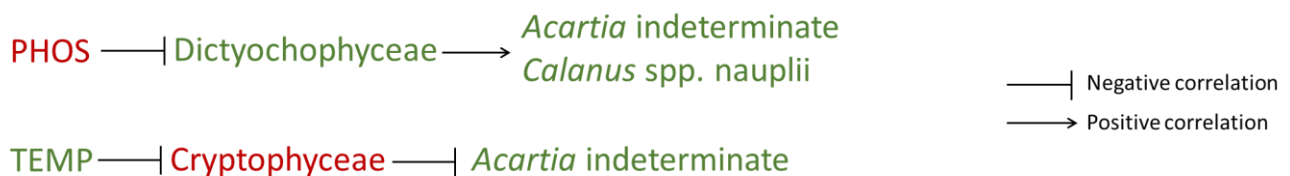

Figure 8 Examples where specific seasonal trends can be explained by relationships between environmental conditions, phytoplankton and zooplankton. Concentration of phosphates (PHOS) in winter is decreasing (red). It shows negative correlation to Dictyochophyceae biomass. Dictyochophyceae biomass is increasing (green) and so is *Acartia* indeterminate and *Calanus* spp. nauplii organisms. Similar interpretation for temperature (TEMP), Cryptophyceae and *Acartia* indeterminate relationship in summer (red represents decreasing trends, green - increasing).

Limitations

The ability to detect trends in hydrographic conditions and plankton abundance as well as to find the relationships between them is greatly affected by the data itself. In this study, samples collected in one station of the Western Baltic Sea (Mecklenburg Bight) were analysed. Since plankton are drifting organisms, the detected abundance of plankton at the specific station can be influenced by sea currents, wind direction and the sample collection technique. Therefore, this study could be extended by adding more sampling points from several more stations from the Western Baltic Sea, to check if common trends at close stations could be detected. Moreover, uneven sample collection always introduces some uncertainty about the validity of the results, because missing values have to be imputed, and it is not clear how well they represent the reality. After the collection of water samples, the next step before providing data for the statistical analysis is examination of collected samples using a microscope, counting how many organisms of specific species can be detected and extrapolating value to number of organisms per m³ or counting the approximate phytoplankton biomass (µg/L). Even though currently more modern techniques, related automated plankton image

analysis are being developed (Luo *et al.*, 2018), to this day the quality of data is relying on the experience of scientists, analysing plankton samples, and consequentially normal human errors are unavoidable, especially if multiple institutions/research groups are involved. Finally, interpretation of the results is not straightforward. Detected correlation values just suggest that a relationship between variables might exist, however the causality of the events cannot be explained only by observational analysis. To fully understand, why detected trends might be present, and explain observed correlations further experimental studies are needed.

CONCLUSIONS

1. Extraordinary cold March in 2013 and associated changes in plankton abundance revealed that short-term anomalies occurring in bloom season have the highest influence on plankton abundance in this particular year.
2. In the last two decades, the biomass of phytoplankton taxa, negatively associated to higher water temperature in summer is decreasing, while biomass of phytoplankton taxa preferring mild winters is increasing.
3. The abundance of zooplankton taxa, playing a key role in ecosystem shifts in the Southern and the Southeastern Baltic Sea, share similar trends in the Western Baltic Sea if the increase or decrease is associated to the water temperature, but not to water salinity.
4. Relation between zooplankton abundance and phytoplankton biomass was detected, however the underlying reasons could only be explained by experimental studies.

EBERHARD KARLS UNIVERSITY TÜBINGEN (GERMANY)

Systems Biology Master's program at Vilnius University

Živilė Grublytė

Master's Thesis

Long -Term Dynamics of Zooplankton Abundance in Relation to Environmental Conditions and Phytoplankton Biomass

SUMMARY

Plankton organisms are useful bioindicators which can provide information whether changing environmental conditions have an influence on marine ecosystems. Previous phytoplankton trends studies from the Mecklenburg Bight covered only the period from 1979 until 2005. Moreover, there are no comprehensive studies of zooplankton trends in the Western Baltic Sea yet available. Therefore, this study focuses on plankton abundance trends during 1998-2016 in the Mecklenburg Bight (Western Baltic Sea). Decreasing and increasing trends were detected using Mann-Kendall test and associated with changing environmental conditions based on Spearman correlation coefficient values. In addition, non-metric multidimensional scaling was used to detect unusual years in order to identify anomalies of plankton abundance that are associated with environmental variability. The results of this study indicate that shifts in plankton communities in 2013 can be associated with an exceptionally cold March this year. It is likely that environmental conditions during the bloom season might have the highest influence on plankton abundance. Comparison of phytoplankton biomass trends in 1998-2016 with those detected until 2005 in the Mecklenburg Bight revealed some changes which occurred due to increasing surface water temperature. Due to the salinity gradient in the Baltic Sea, decreasing zooplankton trends of species susceptible to water freshening in the Southern and Southeastern Baltic Sea (e.g. *Evadne nordmanni* and *Pseudocalanus*) were not identified in the Western Baltic Sea. However, long-term trend analysis revealed that similar zooplankton abundance trends as in other Baltic Sea areas are detected when they are associated with increasing water temperature (e.g. *T. longicornis*, *C. hamatus*, and *Acartia* species). Overall observed trends during the last two decades indicate significant shifts in the Mecklenburg Bight ecosystem which are related to changes in environmental conditions there.

EBERHARD KARLS UNIVERSITY TÜBINGEN (VOKIETIJA)

Sistemų biologijos magistro programa, Vilniaus Universitas
Živilė Grublytė

Magistro darbas

Ilgalaikiai zooplanktono kiekio pokyčiai ir jų sąryšis su aplinkos sąlygomis bei fitoplanktono biomasės pokyčiais

SANTRAUKA

Planktono organizmai yra vertingi bioindikatoriai, galintys suteikti informacijos apie tai ar besikeičiančios aplinkos sąlygos turi įtakos jūrinėms ekosistemoms. Ankstesni Meklenburgo įlankos fitoplanktono biomasės pokyčių tyrimai apėmė laikotarpį tik nuo 1979 iki 2005 m. Be to, iki šiol nebuvo išsamių tyrimų nagrinėjančių zooplanktono kiekio pokyčius Vakarų Baltijos jūroje. Tai paskatino šiame darbe analizuoti planktono kiekio pokyčius Meklenburgo įlankoje (Vakarų Baltijos jūra) nuo 1998 iki 2016. Planktono kiekio didėjimas ar mažėjimas buvo nustatytos naudojant Mann-Kendall testą ir siejamas su besikeičiančiomis aplinkos sąlygomis, remiantis Spearmano koreliacijos koeficiento reikšmėmis. Taip pat, siekiant nustatyti, ar planktono gausos anomalijos yra susijusios su aplinkos kintamumu, išsiskiriantys planktono kiekiu metai buvo aptikti naudojantis nemetrinių daugiamačių skalių (nMDS) analize. Šio tyrimo rezultatai parodė, kad planktono kiekio pokyčiai 2013 m. gali būti siejami su išskirtinai šaltu šių metų kovo mėnesiu. Tai rodo, kad žydėjimo sezono metu esančios aplinkos sąlygos daro didžiausią įtaką planktono gausai. Palyginus fitoplanktono biomasės tendencijas Meklenburgo įlankoje 1998–2016 m. su tendencijomis, nustatytomis iki 2005 m., pastebėti pasikeitimai, įvykę dėl kylančios paviršiaus vandens temperatūros. Dėl nevienodo druskingumo skirtinguose Baltijos jūros regionuose, kai kurių zooplanktono organizmų kiekis sumažėjo Pietų ir Pietryčių Baltijos jūroje (pvz. *Evadne nordmanni* ir *Pseudocalanus*), tačiau sumažėjimas nebuvo nustatytas Vakarų Baltijos jūroje. Vis dėlto, ilgalaikė planktono kiekio pokyčių analizė parodė, kad panašios zooplanktono gausos tendencijos, kaip ir kituose Baltijos jūros regionuose, nustatomos, jei jas galima susieti su kylančia vandens temperatūra (pvz., *T. longicornis*, *C. hamatus*, and *Acartia* rūšys). Per pastaruosius du dešimtmečius nustatyti planktono kiekio pokyčiai rodo reikšmingus Meklenburgo įlankos ekosistemos pasikeitimus, susijusius su besikeičiančiomis aplinkos sąlygomis.

ACKNOWLEDGEMENTS

REFERENCES

1. Alan B. Gelder (2018) *Lecture 3 Time Series I Notations.* Available at: https://agelder.weebly.com/uploads/2/3/2/5/23257230/time_series1.pdf (Accessed: 4 May 2021).

2. Aleksandrov, S., Zhigalova, N. and Zezera, A. (2009) 'Long-term dynamics of zooplankton in the southeastern Baltic Sea', *Russian Journal of Marine Biology*, 35, pp. 296–304. doi: 10.1134/S106307400904004X.

3. Allen, S. (2019) 'Long-term Trends and Variability of North Atlantic Phytoplankton Populations'. University of Southampton.

4. Andrews, J. (2013) 'How Cold Europe Was in March 2013.', *http://www.accuweather.com/en/weather-blogs/ andrews/how-cold-europe-was-in-march-2013/9313909 (Accessed 15 January 2015).* Available at: https://www.accuweather.com/en/weather-news/how-cold-europe-was-in-march-2013/195300 (Accessed: 17 April 2021).

5. Aziz, N. E. A., Gharib, S. M. and Dorgham, M. M. (2006) 'The interaction between phytoplankton and zooplankton in a Lake-Sea connection, Alexandria, Egypt', *International Journal of Oceans and Oceanography ISSN*, 1(1), pp. 973–2667. Available at: http://www.ripublication.com/ijoo.htm.

6. BACC, I. I. and Team, A. (2015) *Second assessment of climate change for the Baltic Sea Basin*. SpringerOpen.

7. BBC (2013) *March weather was second coldest on record - Met Office.* Available at: https://www.bbc.com/news/uk-22010852 (Accessed: 17 April 2021).

8. Beaugrand, G. and Reid, P. C. (2003) 'Long-term changes in phytoplankton, zooplankton and salmon related to climate', *Global Change Biology*, 9(6), pp. 801–817. doi: 10.1046/j.1365-2486.2003.00632.x.

9. Benedetti, F. *et al.* (2019) 'The Seasonal and Inter-Annual Fluctuations of Plankton Abundance and Community Structure in a North Atlantic Marine Protected Area ', *Frontiers in Marine Science* , p. 214. Available at: https://www.frontiersin.org/article/10.3389/fmars.2019.00214.

10. Beveridge, S. (1992) 'Least squates estimation of missing values in time series', *Communications in Statistics - Theory and Methods*. Taylor & Francis, 21(12), pp. 3479–3496. doi: 10.1080/03610929208830990.

11. Bode, A. *et al.* (2015) 'Annual trend patterns of phytoplankton species abundance belie homogeneous taxonomical group responses to climate in the NE Atlantic upwelling', *Marine Environmental Research*. Elsevier Ltd, 110(September), pp. 81–91. doi: 10.1016/j.marenvres.2015.07.017.

12. D'Costa, P. M. and Naik, R. K. (2019) 'Chapter 31 - Advances in sampling strategies and analysis of phytoplankton', in Meena, S. N. and Naik, M. M. B. T.-A. in B. S. R. (eds). Academic Press, pp. 501–521. doi: https://doi.org/10.1016/B978-0-12-817497-5.00031-8.

13. Daniel M. Ware and Richard E. Thomson (2005) 'Production in the Northeast Pacific', *Science*, 308(May), pp. 1280–1284. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15845876.

14. Dzierzbicka-Glowacka, L. *et al.* (2018) 'Seasonal changes in the abundance and biomass of copepods in the south-eastern Baltic Sea in 2010 and 2011', *PeerJ.* PeerJ Inc., 6, pp. e5562–e5562. doi: 10.7717/peerj.5562.

15. European Environment Agency (2015) *Nutrients in transitional, coastal and marine waters — European Environment Agency.* Available at: https://www.eea.europa.eu/data-and-maps/indicators/nutrients-in-transitional-coastal-and-4/assessment (Accessed: 26 April 2021).

16. European Environment Agency (2019) *Nutrients in transitional, coastal and marine waters.* Available at: https://www.eea.europa.eu/data-and-maps/indicators/nutrients-in-transitional-coastal-and/nutrients-in-transitional-coastal-and (Accessed: 23 April 2021).

17. Feike, M. *et al.* (2007) 'Studies on the zooplankton community of a shallow lagoon of the Southern Baltic Sea : long-term trends , seasonal changes , and relations with physical and chemical parameters', pp. 95–106. doi: 10.1007/s10750-006-0420-9.

18. Flynn, K. J. *et al.* (2013) 'Misuse of the phytoplankton-zooplankton dichotomy: The need to assign organisms as mixotrophs within plankton functional types', *Journal of Plankton Research*, 35(1), pp. 3–11. doi: 10.1093/plankt/fbs062.

19. Frederiksen, M. *et al.* (2006) 'From plankton to top predators: Bottom-up control of a marine food web across four trophic levels', *Journal of Animal Ecology*, 75(6), pp. 1259–1268. doi: 10.1111/j.1365-2656.2006.01148.x.

20. Fryer, R. J. and Nicholson, M. D. (1999) 'Using smoothers for comprehensive assessments of contaminant time series in marine biota', *ICES Journal of Marine Science*, 56(5), pp. 779–790. doi: 10.1006/jmsc.1999.0499.

21. Garcia, T. M. *et al.* (2020) 'Plankton net mesh size influences the resultant diversity and abundance estimates of copepods in tropical oligotrophic ecosystems', *Estuarine, Coastal and Shelf Science.* Elsevier Ltd, 249. doi: 10.1016/j.ecss.2020.107083.

22. Garrison, T. S. (2012) *Oceanography: an invitation to marine science.* Cengage Learning.

23. Gary Napier (2020) 'Chapter 1 Time Series Fundamentals | Time Series'. Available at: https://bookdown.org/gary_a_napier/time_series_lecture_notes/ChapterOne.html#time-series-modelling (Accessed: 4 February 2021).

24. Glen, S. (2016) *Mann Kendall Trend Test: Definition, Running the Test - Statistics How To.* Available at: https://www.statisticshowto.com/mann-kendall-trend-test/ (Accessed: 27 March

2021).

25. Gołdyn, R. and Kowalczewska-Madura, K. (2008) 'Interactions between phytoplankton and zooplankton in the hypertrophic Swarzędzkie Lake in western Poland', *Journal of Plankton Research*, 30(1), pp. 33–42. doi: 10.1093/plankt/fbm086.

26. Graham, L. and Bergström, S. (2001) 'Water Balance Modelling in the Baltic Sea Drainage Basin – Analysis of Meteorological and Hydrological Approaches', *Meteorology and Atmospheric Physics*, 77, pp. 45–60. doi: 10.1007/s007030170016.

27. Gräwe, U., Friedland, R. and Burchard, H. (2013) 'The future of the western Baltic Sea: two possible scenarios', *Ocean Dynamics*. Springer, 63(8), pp. 901–921.

28. Griffiths, J. R. *et al.* (2020) 'Limited evidence for common interannual trends in Baltic Sea summer phytoplankton biomass', *PloS one*. Public Library of Science San Francisco, CA USA, 15(4), p. e0231690.

29. Gusmão, L. F. M. and McKinnon, A. D. (2009) 'Sex ratios, intersexuality and sex change in copepods', *Journal of Plankton Research*, 31(9), pp. 1101–1117. doi: 10.1093/plankt/fbp059.

30. Hansson, M., Andersson, L. and Axe, P. (2011) 'Areal Extent and Volume of Anoxia and Hypnoxia in the Baltic Sea, 1960-2011'. SMHI.

31. Hardy, A. C. (1936) 'Plankton ecology and the hypothesis of animal exclusion.', in *Proceedings of the Linnean Society of London*. Wiley Online Library, pp. 64–70.

32. Harris, C. R. *et al.* (2020) 'Array programming with {NumPy}', *Nature*. Springer Science and Business Media {LLC}, 585(7825), pp. 357–362. doi: 10.1038/s41586-020-2649-2.

33. Harvey, H. W. *et al.* (1935) 'Plankton production and its control', *journal of the Marine Biological Association of the United Kingdom*, 20(2), pp. 407–441.

34. Hedberg, P. *et al.* (2020) 'Effects of changing phytoplankton species composition on carbon and nitrogen uptake in benthic invertebrates', *Limnology and Oceanography*. John Wiley & Sons, Ltd, n/a(n/a). doi: https://doi.org/10.1002/lno.11617.

35. Helcom (2001) 'Environment of the Baltic Sea area 1994-1998', *Baltic Sea Environment Proceedings*, 82a(82), pp. 1–26.

36. Helcom (2017) 'Manual for Marine Monitoring in the COMBINE', (July), pp. 309–332.

37. HELCOM (1996) *Third Periodic Assessment of the State of the Marine Environment of the Baltic Sea, 1989-1993*. Helsinki Commission, Baltic Marine Environment Protection Commission.

38. HELCOM (2013) 'Climate Change in the Baltic Sea Area, HELCOM thematic assessment in 2013', *Baltic Sea Environment Proceedings*, 137(137), p. 54.

39. HELCOM (2017a) *Guidelines for Monitoring of Mesozooplankton*.

40. HELCOM (2017b) 'State of the Baltic Sea- Second HELCOM holistic assessment, 2011-2016', *Baltic Sea Environment Proceedings 155*, 155, pp. 4–7. Available at: http://stateofthebalticsea.helcom.fi/pressures-and-their-status/hazardous-substances/.

41. HELCOM (2018) *Zooplankton mean size and total stock (MSTS) Key Message*.

42. HELCOM (2020) *Guidelines for monitoring phytoplankton species composition, abundance and biomass.*

43. Hendrey, G. R. (2001) 'Acid Rain and Deposition', in Levin, S. A. B. T.-E. of B. (ed.). New York: Elsevier, pp. 1–15. doi: https://doi.org/10.1016/B0-12-226865-2/00001-8.

44. Hessen, D. O. and Kaartvedt, S. (2014) 'Top–down cascades in lakes and oceans: different perspectives but same story?', *Journal of Plankton Research*, 36(4), pp. 914–924. doi: 10.1093/plankt/fbu040.

45. Hirst, A. G. *et al.* (2010) 'Does predation control adult sex ratios and longevities in marine pelagic copepods?', *Limnology and Oceanography*, 55(5), pp. 2193–2206. doi: 10.4319/lo.2010.55.5.2193.

46. Hjerne, O. *et al.* (2019) 'Climate Driven Changes in Timing, Composition and Magnitude of the Baltic Sea Phytoplankton Spring Bloom', *Frontiers in Marine Science*, 6, p. 482. doi: 10.3389/fmars.2019.00482.

47. Huber, J. (1977) 'Local Smoothing: a Method of Controlling Error and Estimating Relationships in Consumer Research', *ACR North American Advances*, NA-04. Available at: https://www.acrwebsite.org/volumes/9323/volumes/v04/NA-04/full (Accessed: 1 February 2021).

48. Hunter, J. D. (2007) 'Matplotlib: A 2D graphics environment', *Computing in Science & Engineering*. IEEE COMPUTER SOC, 9(3), pp. 90–95. doi: 10.1109/MCSE.2007.55.

49. Hussain, M. and Mahmud, I. (2019) 'pyMannKendall: a python package for non parametric Mann Kendall family of trend tests.', *Journal of Open Source Software*. The Open Journal, 4(39), p. 1556. doi: 10.21105/joss.01556.

50. *ICES – HELCOM* (no date). Available at: https://helcom.fi/about-us/partners/ices/ (Accessed: 28 January 2021).

51. Kane, J. (2007) 'Zooplankton abundance trends on Georges Bank, 1977-2004', *ICES Journal of Marine Science*, 64(5), pp. 909–919. doi: 10.1093/icesjms/fsm066.

52. Käse, L. and Geuer, J. K. (2018) 'Phytoplankton Responses to Marine Climate Change – An Introduction BT - YOUMARES 8 – Oceans Across Boundaries: Learning from each other', in Jungblut, S., Liebich, V., and Bode, M. (eds). Cham: Springer International Publishing, pp. 55–71.

53. Kelly, P. *et al.* (2016) 'Zooplankton responses to increasing sea surface temperatures in the southeastern Australia global marine hotspot', *Estuarine, Coastal and Shelf Science*, 180, pp. 242–257. doi: https://doi.org/10.1016/j.ecss.2016.07.019.

54. Kniebusch, M. *et al.* (2019) 'Temperature Variability of the Baltic Sea Since 1850 and Attribution to Atmospheric Forcing Variables', *Journal of Geophysical Research: Oceans*, 124(6), pp. 4168–4187. doi: 10.1029/2018JC013948.

55. Kozak, A. and Gołdyn, R. (2004) 'Zooplankton versus phyto- and bacterioplankton in the Maltański Reservoir (Poland) during an extensive biomanipulation experiment', *Journal of Plankton Research*, 26(1), pp. 37–48. doi: 10.1093/plankt/fbh006.

56. Lalli, C. and Parsons, T. R. (1997) *Biological oceanography: an introduction*. Elsevier.

57. Lehmann, A., Getzlaff, K. and Harlaß, J. (2011) 'Detailed assessment of climate variability in the Baltic Sea area for the period 1958 to 2009', *Climate Research*, 46(2), pp. 185–196.

58. Lepot, M., Aubin, J. B. and Clemens, F. H. L. R. (2017) 'Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment', *Water (Switzerland)*, 9(10). doi: 10.3390/w9100796.

59. Leppäranta, M. and Myrberg, K. (2009) *Physical oceanography of the Baltic Sea*. Springer Science & Business Media.

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

61. Lindahl, O. (1986) 'A dividable hose for phytoplankton sampling', *Report of the working group on phytoplankton and management of their effects. International Council for the Exploration of the Sea, CM*, 50, p. 26.

62. Lumini, A. and Nanni, L. (2019) *Ocean ecosystems plankton classification*, *Studies in Computational Intelligence*. Springer International Publishing. doi: 10.1007/978-3-030-03000-1_11.

63. Luo, J. Y. *et al.* (2018) 'Automated plankton image analysis using convolutional neural networks', *Limnology and Oceanography: Methods*. John Wiley & Sons, Ltd, 16(12), pp. 814–827. doi: https://doi.org/10.1002/lom3.10285.

64. M. Engel (2013) 'Calanoid copepod resting eggs - a safeguard against adverse environmental conditions in the German Bight and the Kara Sea?', *Journal of Chemical Information and Modeling*, 53(9), pp. 1689–1699.

65. Majaneva, M. *et al.* (2009) 'Phytoplankton monitoring: the effect of sampling methods used during different stratification and bloom conditions in the Baltic Sea'. Boreal Environment Research Publishing Board.

66. Mauchline, J. (1998) *Adv. Mar. Biol. 33: The biology of calanoid copepods*.

67. *MCD - Seasonal Adjustment Frequently Asked Questions* (2017). Available at: https://web.archive.org/web/20170113192428/http://www.census.gov/const/www/faq2.html#two (Accessed: 5 February 2021).

68. Mitsch, W. J. and Gosselink, J. G. (2011) *Wetlands*. John Wiley & Sons.

69. Möllmann, C. *et al.* (2002) 'Long-term trends in abundance of cladocerans in the Central Baltic Sea', *Marine Biology*, 141(2), pp. 343–352. doi: 10.1007/s00227-002-0822-9.

70. Musialik-Koszarowska, M., Dzierzbicka-Głowacka, L. and Weydmann, A. (2019) 'Influence

of environmental factors on the population dynamics of key zooplankton species in the Gulf of Gdańsk (southern Baltic Sea)', *Oceanologia*, 61(1), pp. 17–25. doi: https://doi.org/10.1016/j.oceano.2018.06.001.

71. National Oceanic and Atmospheric Administration (2020) *What are Plankton?*, *U.S. Department of Commerce*. Available at: https://oceanservice.noaa.gov/facts/plankton.html (Accessed: 6 May 2021).

72. NOAA (2017) 'What are phytoplankton?', *What are phytoplankton?*, pp. 1–8. Available at: https://earthobservatory.nasa.gov/features/Phytoplankton (Accessed: 16 January 2021).

73. Ojaveer, E., Lumberg, A. and Ojaveer, H. (1998) 'Highlights of zooplankton dynamics in Estonian waters (Baltic Sea)', *ICES Journal of Marine Science*, 55(4), pp. 748–755. doi: 10.1006/jmsc.1998.0393.

74. Reckermann, M. *et al.* (2014) *Climate Change in the Baltic Sea region - What do we know ?*

75. Rice, E., Dam, H. G. and Stewart, G. (2015) 'Impact of Climate Change on Estuarine Zooplankton: Surface Water Warming in Long Island Sound Is Associated with Changes in Copepod Size and Community Structure', *Estuaries and Coasts*, 38(1), pp. 13–23. doi: 10.1007/s12237-014-9770-0.

76. Rodríguez, E. *et al.* (1989) *Biological Research in Aquatic Science*, *Intech*. Available at: https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics.

77. Rousseaux, C. S. and Gregg, W. W. (2015) 'Recent decadal trends in global phytoplankton composition', *Global Biogeochemical Cycles*. John Wiley & Sons, Ltd, 29(10), pp. 1674–1688. doi: https://doi.org/10.1002/2015GB005139.

78. Shurin, J. B. *et al.* (2002) 'A cross-ecosystem comparison of the strength of trophic cascades', *Ecology letters*. Wiley Online Library, 5(6), pp. 785–791.

79. Siegel, H., Gerth, M. and Tschersich, G. (2006) 'Sea surface temperature development of the Baltic Sea in the period 1990-2004', *Oceanologia*. -, 48(S).

80. Snoeijs-Leijonmalm, P., Schubert, H. and Radziejewska, T. (2017) *Biological oceanography of the Baltic Sea*. Springer Science & Business Media.

81. Sommer, U. and Sommer, F. (2006) 'Cladocerans versus copepods: the cause of contrasting top–down controls on freshwater and marine phytoplankton', *Oecologia*. Springer, 147(2), pp. 183–194.

82. Stanley Gregory (2003) *Zooplankton*. Available at: http://oregonstate.edu/instruction/fw556/lecture/pdf/zooplankton.pdf (Accessed: 31 January 2021).

83. Stibor, H. *et al.* (2004) 'Copepods act as a switch between alternative trophic cascades in marine pelagic food webs', *Ecology Letters*. John Wiley & Sons, Ltd, 7(4), pp. 321–328. doi: https://doi.org/10.1111/j.1461-0248.2004.00580.x.

84. Störmer, O. (2011) 'Climate Change Impacts on Coastal Waters of the Baltic Sea', *Coastal Research Library*, 1, pp. 51–69. doi: 10.1007/978-94-007-0400-8_4.

85. Stramska, M. and Białogrodzka, J. (2015) 'Spatial and temporal variability of sea surface temperature in the Baltic Sea based on 32-years (1982–2013) of satellite data', *Oceanologia*. Elsevier, 57(3), pp. 223–235.

86. Terpilowski, M. (2019) 'scikit-posthocs: Pairwise multiple comparison tests in Python', *The Journal of Open Source Software*, 4(36), p. 1169. doi: 10.21105/joss.01169.

87. The pandas development team (2020) 'pandas-dev/pandas: Pandas'. Zenodo. doi: 10.5281/zenodo.3509134.

88. Valdés, L. and Moral, M. (1998) 'Time-series analysis of copepod diversity and species richness in the southern Bay of Biscay off Santander, Spain, in relation to environmental conditions', *ICES Journal of Marine Science*, 55(4), pp. 783–792. doi: 10.1006/jmsc.1998.0386.

89. Viitasalo, M. (1992) 'Mesozooplankton of the gulf of finland and northern Baltic proper a review of monitoring data', *Ophelia*. Taylor & Francis, 35(2), pp. 147–168. doi: 10.1080/00785326.1992.10429976.

90. Virtanen, P. *et al.* (2020) '{SciPy} 1.0: Fundamental Algorithms for Scientific Computing in Python', *Nature Methods*, 17, pp. 261–272. doi: 10.1038/s41592-019-0686-2.

91. Waskom, M. *et al.* (2017) 'mwaskom/seaborn: v0.8.1 (September 2017)'. Zenodo. doi: 10.5281/zenodo.883859.

92. Wasmund, N. *et al.* (2011) 'Long-term trends in phytoplankton composition in the western and central Baltic Sea', *Journal of Marine Systems*, 87(2), pp. 145–159. doi: 10.1016/j.jmarsys.2011.03.010.

93. Wasmund, N., Nausch, G. and Feistel, R. (2013) 'Silicate consumption: An indicator for long-term trends in spring diatom development in the Baltic Sea', *Journal of Plankton Research*, 35(2), pp. 393–406. doi: 10.1093/plankt/fbs101.

94. Wasmund, N. and Uhlig, S. (2003) 'Phytoplankton trends in the Baltic Sea', *ICES Journal of Marine Science*, 60(2), pp. 177–186. doi: 10.1016/S1054-3139(02)00280-1.

95. Wirth, C., Limberger, R. and Weisse, T. (2019) 'Temperature × light interaction and tolerance of high water temperature in the planktonic freshwater flagellates Cryptomonas (Cryptophyceae) and Dinobryon (Chrysophyceae)', *Journal of Phycology*, 55(2), pp. 404–414. doi: 10.1111/jpy.12826.

96. World Meteorological Organization (2014) *The March 2013 Cold Wave Over Europe | World Meteorological Organization*. Available at: https://public.wmo.int/en/resources/meteoworld/march-2013-cold-wave-over-europe (Accessed: 17 April 2021).

97. Zhang, H. *et al.* (2015) 'Counteracting effects of recruitment and predation shape

establishment of rotifer communities under climate change', *Limnology and Oceanography*. Wiley Blackwell, 60(5), pp. 1577–1587. doi: 10.1002/lno.10122.

APPENDICES

Appendix 1. Supplementary tables

Supplementary Table 1 Median of hydrography data for 2013 and IQR for all years

Variable	Median 2013	IQR 1998-2018
NTRA	0.650	0.040 - 1.815
OXY	7.006	6.771 - 8.347
PHOS	0.393	0.115 - 0.553
SAL	10.03	9.773 - 13.52
SLCA	14.08	6.850 - 14.20
TEMP	11.03	4.379 - 14.09

Supplementary Table 2 Phytoplankton biomass (µg/L) in March 2013 and median values with IQRs in March 1998-2016. All phytoplankton classes except Bacillariophyceae, Choanoflagellatea, Cilioph and Ebriophyceae have significantly lower biomass in March of 2013 than in March during 1998-2016.

Variable	2013 March	Median (1998-2016 March)	IQR (1998-2016 March)
Bacillariophyceae	240.346	387.537	124.799 - 1278.174
Chlorophyceae	0.013	0.277	0.117 - 0.434
Choanoflagellatea	0.768	0.764	0.281 - 1.696
Chrysophyceae	1.116	4.508	1.943 - 11.282
Ciliophora	115.469	59.771	40.19 - 117.87
Cryptophyceae	11.254	30.918	15.391 - 42.238
Cyanophyceae	0.526	1.138	0.550 - 3.762
Dictyochophyceae	1.850	6.137	5.382 - 21.394
Dinophyceae	6.661	106.839	50.55 - 292.89
Ebriophyceae	4.455	1.226	0.64 - 3.069
Euglenophyceae	1.315	4.364	1.866 - 6.238
Incertae sedis	0.853	1.391	1.264 - 1.573
Others	0.456	20.696	11.868 - 39.321
Prasinophyceae	0.302	1.545	0.465 - 3.601
Prymnesiophyceae	0.610	5.338	2.123 - 28.583
Trebouxiophyceae	0.032	0.284	0.148 - 0.285

Supplementary Table 3 Copepods abundance (counts/m³) in March 2013 and median values with IQRs in March 1998-2018.

Copepods species	Stage of development	2013 March	Median (March 1998-2018)	IQR (March 1998-2018)
Acartia spp.	Nauplii	949.93	1792.44	1038.53 - 3326.387
	Adults	1094.32	372.86	192.0 - 659.17
Calanus spp.	Nauplii	760.60	0.00	0.0 - 0.0
Centropages spp.	Nauplii	17.47	408.20	156.17 - 1097.05
	Adults	285.09	63.68	45.72 - 135.22
Oithona similis	Nauplii	82.31	231.13	106.68 - 741.16
	Adults	465.57	342.42	155.70 - 480.06
Pseudo(Para)calanus spp.	Nauplii	8652.78	1849.09	1137.91 - 2410.13
	Adults	585.09	86.81	41.95 - 151.98
Temora longicornis	Nauplii	55.99	792.18	305.79 - 1080.28
	Adults	717.98	134.13	100.33 - 204.57

Supplementary Table 4 Comparison of "spring" species correlation to hydrography variables in spring seasons and the entire time series. Spearman correlation between abundance of "spring" species in spring and values of hydrography variables in spring was calculated. Spearman correlation was also calculated between the same "spring" species abundance and hydrography variables in the complete time series. Median values with IQRs of significant correlation coefficients are represented in the table.

Analyzed group	Median correlation to hydrography variables in spring	Median correlation to hydrography variables during the entire year
Phytoplankton classes	0.217 (0.174 - 0.269)	0.518 (0.466 - 0.571)
Phytoplankton orders	0.219 (0.170 - 0.282)	0.582 (0.582 - 0.582)
Cladocera	0.222 (0.178 - 0.286)	0.574 (0.574 - 0.574)
Copepods	0.247 (0.182 - 0.361)	0.492 (0.482 - 0.533)
Others	0.217 (0.158 - 0.298)	0.510 (0.506 - 0.516)

Supplementary Table 5 Copepods species with significant monotonic increasing or decreasing trends, detected at least at one of the seasons. Spearman correlation coefficient is shown for phytoplankton classes' biomass in seasons with detectable trends

	Decreasing	Increasing	Correlation
Acartia spp.			
Acartia indeterminate	-	Spr, Sum, Aut, Win	Spr: Ebriophyceae 0.5207 Sum: Cryptophyceae -0.4825, Others -0.5083 Aut: Cryptophyceae -0.5133 Win: Ciliophora -0.4997, Cryptophyceae -0.5588, Dictyochophyceae 0.479, Ebriophyceae 0.5268
A. bifilosa F	Aut	-	Aut: Ebriophyceae 0.5328
A. tonsa F	-	Aut, Win	Aut: Prymnesiophyceae 0.4572 Win: Ciliophora -0.5738
A. tonsa M	-	Win	

|	Decreasing	Increasing	Correlation
Calanus spp.			
Nauplii	-	Spr, Sum, Aut, Win	Sum: Incertae sedis 0.607, Prasinophyceae 0.5361, Trebouxiophyceae 0.6041 Aut: Cryptophyceae -0.6263, Prymnesiophyceae 0.4563 Win: Dictyochophyceae 0.6486, Dinophyceae 0.4756
***Centropages* spp.**			
J1	-	Spr	Spr: Chrysophyceae -0.5772, Cryptophyceae -0.7246, Prasinophyceae 0.7228, Prymnesiophyceae 0.5204
J2	-	Spr	Spr: Euglenophyceae -0.4895, Trebouxiophyceae 0.6819
C. hamatus F	Win	-	Win: Choanoflagellatea -0.4596
***Eurytemora* spp.**			
Nauplii	-	Sum, Aut	Sum: Trebouxiophyceae 0.4582, Aut: Cryptophyceae -0.6535, Cyanophyceae 0.5114
J2	-	Aut	Aut: Cryptophyceae -0.7275, Incertae sedis 0.5488
E. affinis M	-	Sum, Aut	Sum: Prasinophyceae 0.6719, Aut: Cryptophyceae -0.615
Others			
Harpacticoidea	Spr	-	Spr: Chrysophyceae 0.4707, Prymnesiophyceae -0.4925
Microsetella	-	Spr, Sum, Aut, Win	Spr: Chlorophyceae -0.4934, Sum: Incertae sedis 0.4579, Prasinophyceae 0.5699, Trebouxiophyceae 0.5027, Aut: Cryptophyceae -0.679
***Pseudo(Para)calanus* spp.**			
Nauplii	Win	-	
J2	Spr, Win	-	Spr: Cryptophyceae 0.4842, Ebriophyceae 0.486
Par. parvus F	-	Spr, Sum, Aut, Win	Spr: Ebriophyceae 0.4698, Sum: Others -0.5061, Prasinophyceae 0.4869
Par. parvus M	-	Sum, Aut	Aut: Cryptophyceae -0.6634
Temora longicornis			
J1	-	Spr	Spr: Incertae sedis 0.4858
J2	-	Win	
T. longicornis F	Aut	-	Aut: Cryptophyceae 0.5439
T. longicornis M	-	Spr	Spr: Chrysophyceae -0.4649, Cryptophyceae -0.6211, Prasinophyceae 0.5211

Supplementary Table 6 Temperature (ºC) median values (ºC) for 2013 and IQR range for temperature values in 1998-2018 for all months.

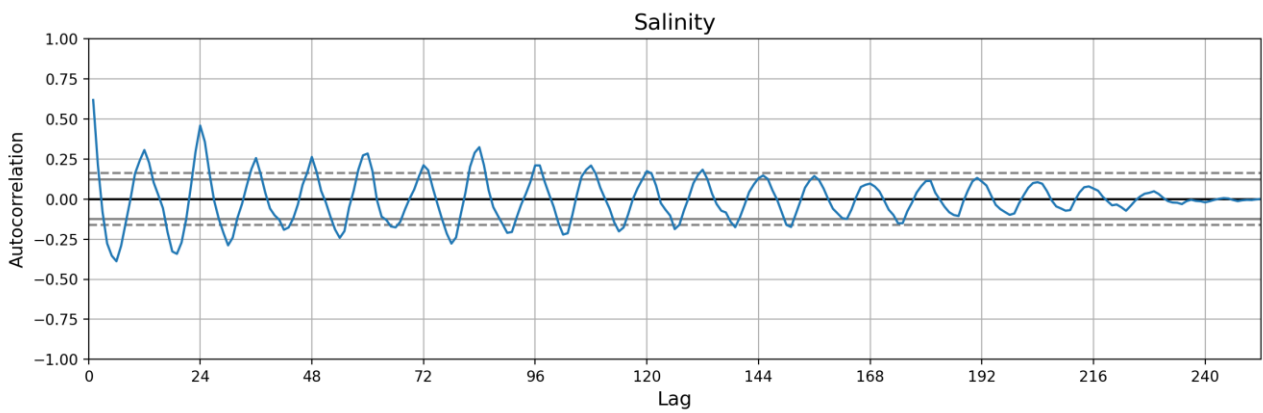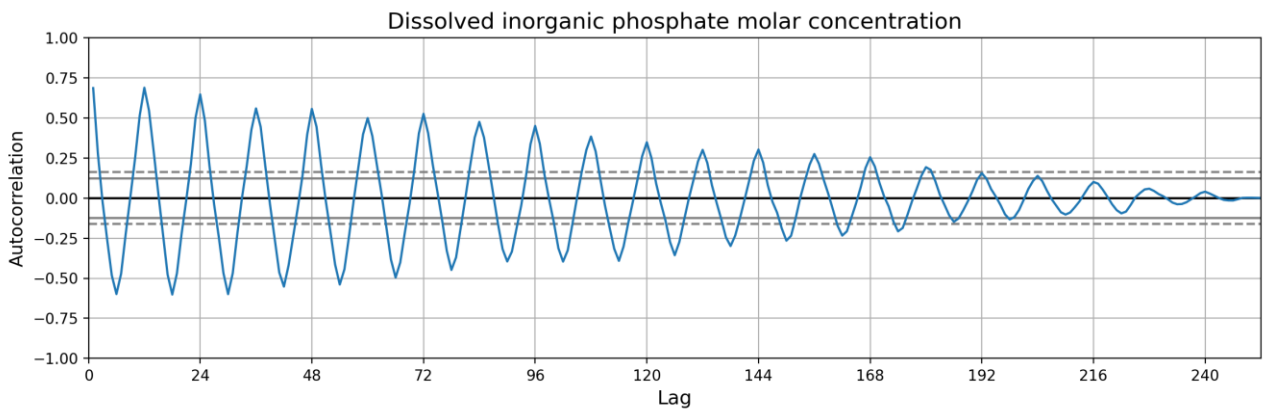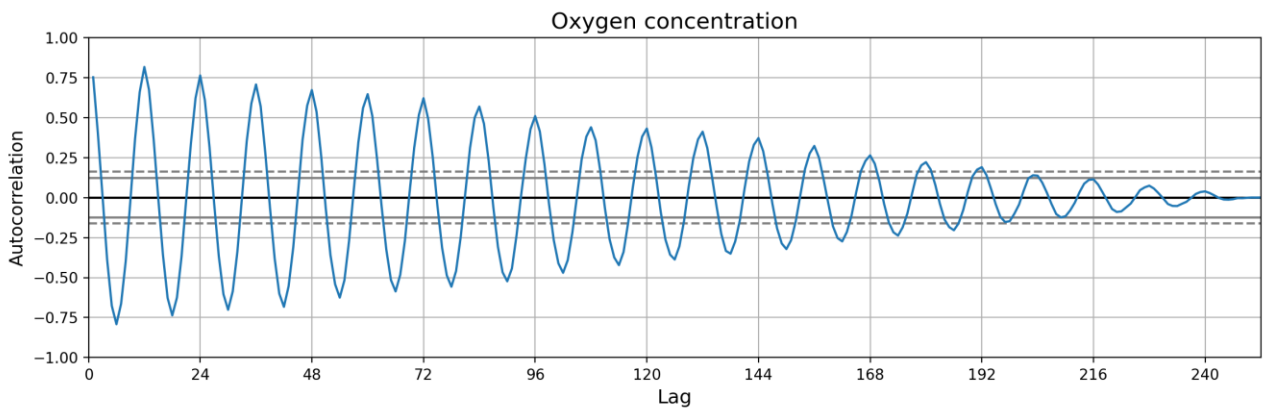| Month	TEMP in 2013	TEMP IQR 1998-2018
Jan	2.129	1.648 - 3.303
Feb	2.129	1.587 - 3.216
Mar	0.774	2.822 - 4.379
Apr	4.893	4.945 - 7.295
May	9.011	9.011 - 10.226
June	14.319	12.889 - 14.298
July	19.626	17.323 - 19.065
Aug	20.528	16.875 - 18.493
Sep	20.528	12.054 - 15.146
Oct	11.027	10.163 - 11.256
Nov	11.027	9.625 - 11.027
Dec	11.027	6.356 - 10.163

Appendix 2. Supplementary plots
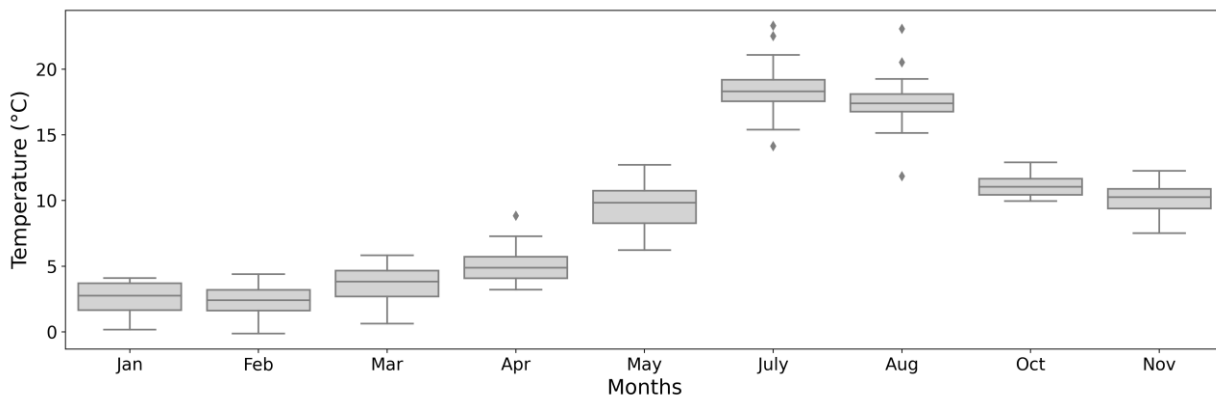
Supplementary Figure 1 Box plots representing water surface (0-4 m) temperature (°C) distribution based on the raw data. No information about the temperature was provided for June, September and December.
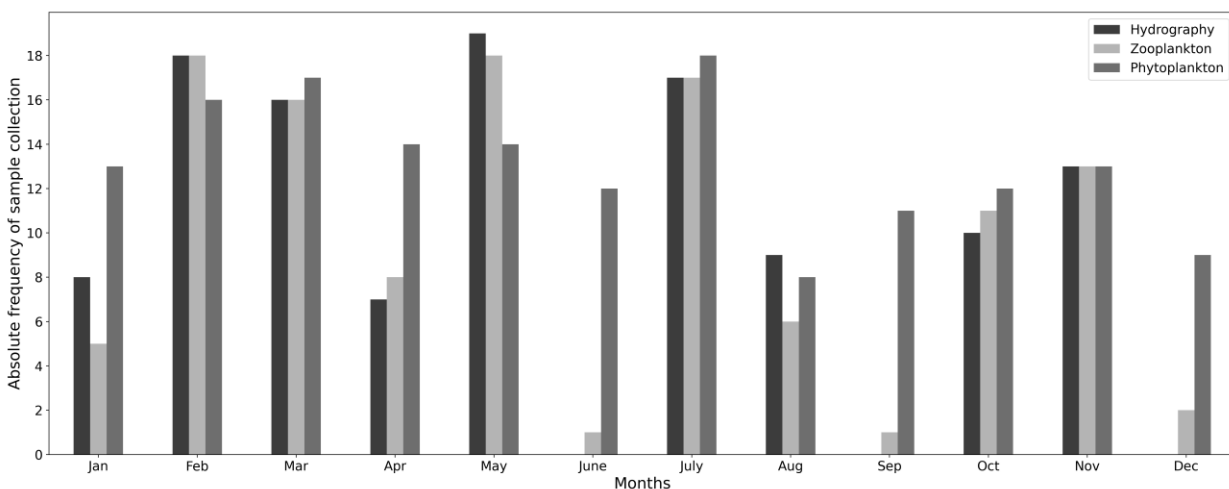
Supplementary Figure 2 Bar plot showing how many samples were collected in each month in the period 1998-2016. Hydrography samples were not collected in June, September and December and only very few zooplankton samples were collected during these months.

Dissolved nitrate molar concentration

Oxygen concentration

Dissolved inorganic phosphate molar concentration

Salinity
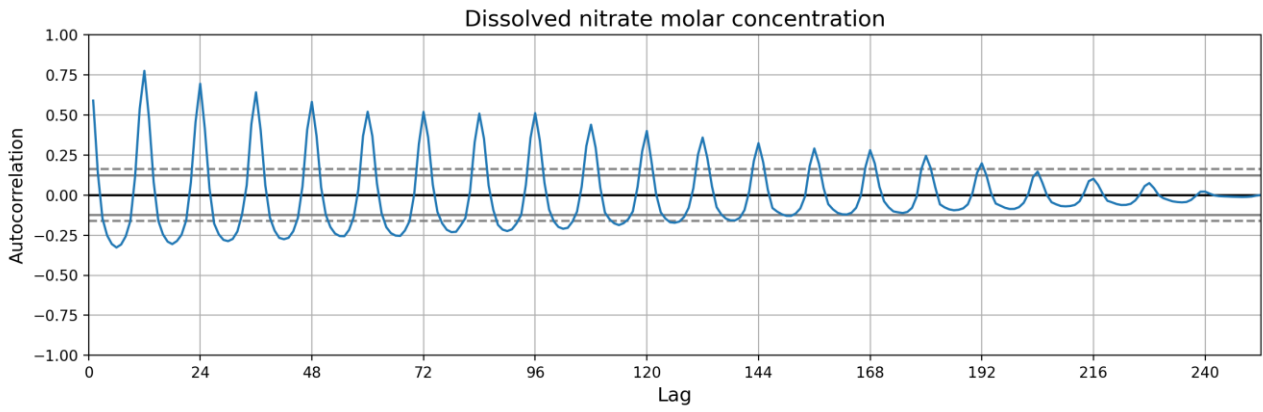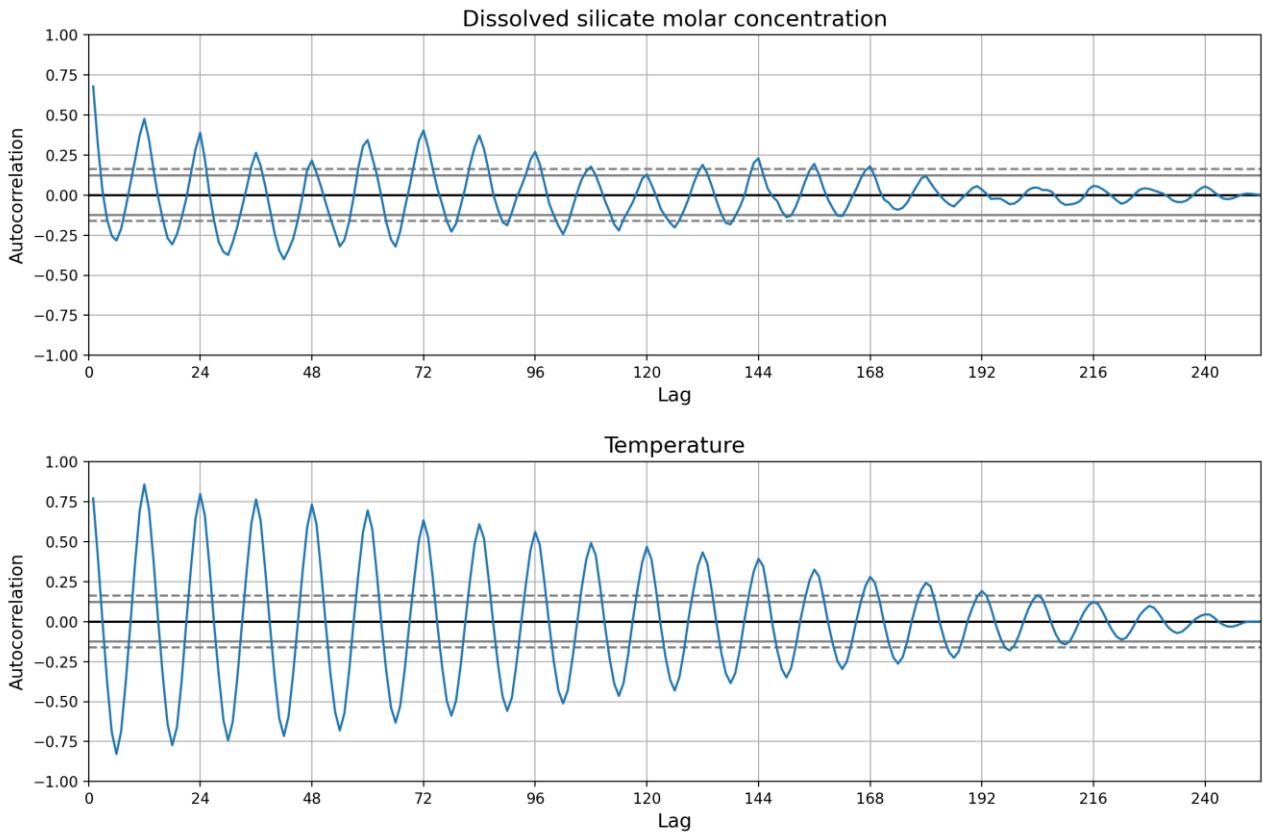
Supplementary Figure 3 Autocorrelation plots demonstrating hydrography data seasonality. The peaks in the autocorrelation plots rep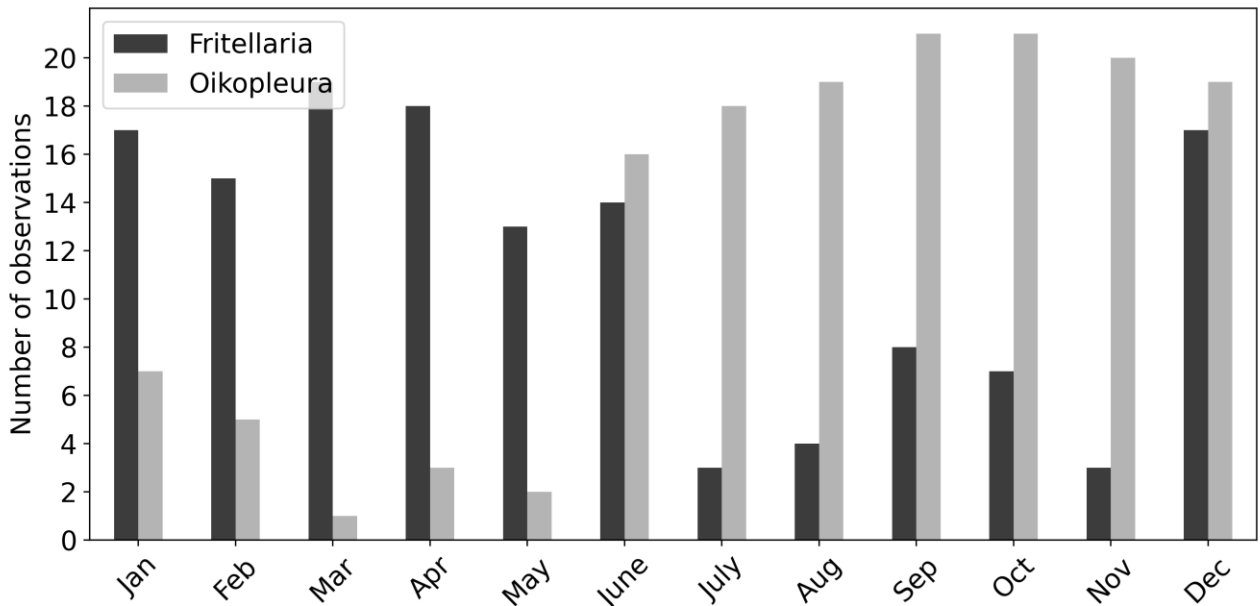eating at lag=12 period indicate yearly seasonality in monthly data with filled missing values. The horizontal lines in the plot correspond to 95% and 99% confidence intervals.

Supplementary Figure 4 Number of observations of *Fritellaria* (spring species) and *Oikopleura* (autumn species) in the particular month. Monthly data with imputed missing values were used for this bar plot.
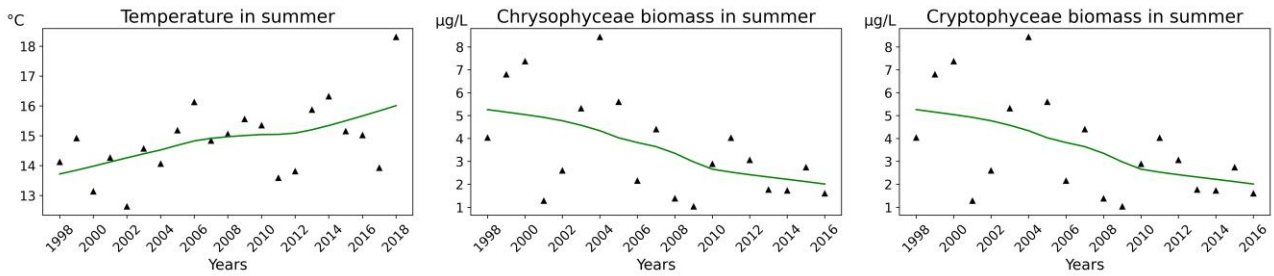
Supplementary Figure 5 Temperature (°C) and biomass (µg/L) of Chrysophyceae and Cryptophyceae phytoplankton classes in summer. Fraction of data used for the locally weighted scatterplot smoothing - 0.75.

Appendix 3. Python modules

Appendix 3.1 Module to import data from the Excel files

```python
# get data
import pandas as pd
#path to data
path_hydrography = 'data/Station_TF-0012-1998-2018_Hydrography.xlsx'
path_zooplankton = 'data/Station_TF-0012_-_1998-2018_Zooplankton.xlsx'
path_phytoplankton_classes = 'data/Phytoplankton/Classes_TF0012.xlsx'
path_phytoplankton_orders = 'data/Phytoplankton/Cyano_Dino_orders_TF0012.xlsx'


def create_dataframe_from_depth(df,minv,maxv):
    '''Selets rows where pressure is between given values'''
    columns=df.columns[1:len(df.columns)-1]
    df=df.loc[df.loc[:,'PRES7PRD']>=minv]
    df=df.loc[df.loc[:,'PRES7PRD']<=maxv]
    df=df.loc[:,columns]
    df=df.reset_index()
    df=df.groupby(df['Time'].dt.date).agg(['mean'])
    df=df.droplevel(axis=1, level=1)
    df.index= pd.to_datetime(pd.Series(list(df.index)))
    df['Month'] = df.index.month
    return df


def hydrography():
    '''Gets hydrography data: df_h_data, df_h_station, df_h_expediton'''
    df_h_data = pd.read_excel (path_hydrography, sheet_name='Data')
    df_h_station = pd.read_excel (path_hydrography, sheet_name='Station Details')
    df_h_expediton = pd.read_excel (path_hydrography, sheet_name='Expedition Details')
    df_h_data=df_h_data.set_index('Time')
    df_h_data=df_h_data.loc[:,'PRES7PRD':'TEMP7STD']
    df_h_data=df_h_data.dropna(axis=0, how='all')
    df_h_data['Month']=df_h_data.index.month
    return df_h_data


def hydrography_depth():
```

```python
    '''Returns hydrography data dataframes where pressure is: 0-4; 15-20'''
    df_h_data=hydrography()
    depth0_4=create_dataframe_from_depth(df_h_data,0,4)
    depth15_20=create_dataframe_from_depth(df_h_data,15,20)
    return depth0_4, depth15_20


def zooplankton():
    '''Gets zooplankton data: df_zooplankton'''
    df_zooplankton_raw = pd.read_excel (path_zooplankton, skiprows=5)
    df_zooplankton = df_zooplankton_raw.dropna(how='all') #dropping empty lines
    df_zooplankton = df_zooplankton.drop(axis=0, index=1)
    df_zooplankton.iloc[:,[0,1]] = df_zooplankton.iloc[:,[0,1]].fillna(method='ffill')
    df_zooplankton = df_zooplankton.rename(columns = {'Unnamed: 0': 'Subclass/order','Unnamed: 1': 'Genus','month': 'Stage','04.08.002':
'04/08/2002'})
    df_zooplankton = df_zooplankton.set_index(['Subclass/order','Genus','Stage'])
    df_zooplankton.columns = pd.to_datetime(pd.Series(list(df_zooplankton.columns)))
    df_zooplankton = df_zooplankton.transpose()
    return df_zooplankton


def phytoplankton_class():
    '''Gets phytoplankton classes data, taxa which belong in classes saved in the dictionary'''
    df_phy_class = pd.read_excel(path_phytoplankton_classes, parse_dates=[5])
    i = 11
    taxa_in_classes={}
    while i < (len(df_phy_class.columns)):
            key = (df_phy_class.iloc[0,i])
            value = df_phy_class.columns[i]
            taxa_in_classes[key] = value
            i+=1
    df_phy_class = pd.DataFrame(df_phy_class[:].values)
    df_phy_class = df_phy_class.rename(columns=df_phy_class.iloc[0])
    df_phy_class = df_phy_class.drop(0, axis=0)
    df_phy_class = df_phy_class.set_index('DATE')
    df_phy_class = df_phy_class.drop('Month', axis=1)
    df_phy_class.index = pd.to_datetime(pd.Series(list(df_phy_class.index)))
    df_phy_class_for_an=pd.DataFrame(df_phy_class.loc[:,'Bacillariophyceae':])
    df_phy_class_for_an['Month'] = df_phy_class_for_an.index.month
    df_phy_class_for_an=df_phy_class_for_an.drop(columns=['Raphidophyceae'])
    return df_phy_class_for_an, taxa_in_classes


def phytoplankton_order():
    '''Gets phytoplankton orders data, to which class order belong is saved in the dictionary'''
    df_phy_orders = pd.read_excel(path_phytoplankton_orders, parse_dates=[5])
    class_of_order={}
    class_of_order['Cyanophyceae']=list(df_phy_orders.iloc[0,7:10])
    class_of_order['Dinophyceae']=list(df_phy_orders.iloc[0,10:])
    df_phy_orders = pd.DataFrame(df_phy_orders[:].values)
    df_phy_orders = df_phy_orders.rename(columns=df_phy_orders.iloc[0])
    df_phy_orders = df_phy_orders.drop(0, axis=0)
    df_phy_orders = df_phy_orders.set_index('DATE')
    df_phy_orders = df_phy_orders.drop('Month', axis=1)
```

```python
    df_phy_orders.index = pd.to_datetime(pd.Series(list(df_phy_orders.index)))
    df_phy_order_for_an=pd.DataFrame(df_phy_orders.loc[:,'Chroococcales':])
    df_phy_order_for_an['Month'] = df_phy_order_for_an.index.month
    return df_phy_order_for_an, class_of_order


# Functions to return hydrography data with combined columns

def hydro_with_combined_columns():
    '''Combines columns measuring the same property'''
    df=hydrography()
    df['TEMP']=df['TEMP7IDD'].fillna(df.TEMP7STD)
    df['OXY']=df['DOXY4IDD'].fillna(df.DOXY6TID)
    df['SAL']=df['PSAL7IDD'].fillna(df.PSAL5IBD)
    df.drop(labels=['TEMP7IDD', 'TEMP7STD', 'DOXY4IDD', 'DOXY6TID', 'PSAL7IDD', 'PSAL5IBD'], axis=1)
    df=df[['PRES7PRD', 'TEMP', 'OXY', 'SAL', 'NTRA7DXD', 'PHOS7DXD', 'SLCA7DXD']]
    df=df.rename({'PRES7PRD': 'PRES', 'NTRA7DXD': 'NTRA', 'PHOS7DXD': 'PHOS', 'SLCA7DXD': 'SLCA'}, axis=1)
    df['Month']=df.index.month
    return(df)


def create_dataframe_from_depth_PRES(df,minv,maxv):
    '''Selets rows where pressure is between given values'''
    columns=df.columns[1:len(df.columns)-1]
    df=df.loc[df.loc[:,'PRES']>=minv]
    df=df.loc[df.loc[:,'PRES']<=maxv]
    df=df.loc[:,columns]
    df=df.reset_index()
    df=df.groupby(df['Time'].dt.date).agg(['mean'])
    df=df.droplevel(axis=1, level=1)
    df.index= pd.to_datetime(pd.Series(list(df.index)))
    df['Month'] = df.index.month
    return df


def hydrography_depth_comb():
    '''Returns hydrography data dataframes where pressure is: 0-4; 15-20'''
    df_h_data=hydro_with_combined_columns()
    depth0_4=create_dataframe_from_depth_PRES(df_h_data,0,4)
    depth15_20=create_dataframe_from_depth_PRES(df_h_data,15,20)
    return depth0_4, depth15_20
```

Appendix 3.2 Module to get data frames, representing zooplankton groups (copepods, cladocerans, others)

```python
# functions to get dataframes of zooplankton groups
import pandas as pd
import get_data


df_zooplankton = get_data.zooplankton()
df_zooplankton['Year'] = df_zooplankton.index.year
df_zooplankton['Month'] = df_zooplankton.index.month
```

```python
df_zooplankton_new=df_zooplankton.drop([('Cladocera', 'Cercopagis pengoi'),('Cladocera', 'Evadne anonyx'),('Cladocera', 'Evadne
spinifera')], axis=1)
df_zooplankton_new=df_zooplankton_new.drop([('others', 'Ostracoda'),('others', 'Gammaridea'),('others', 'Trematoda')], axis=1)
df_zooplankton_new=df_zooplankton_new.drop([('Copepods', 'Limnocalanus macurus')], axis=1)


#Dictionary to have 3 seperate dataframes for Copepods, Cladocera, Others:
di={}
for i in df_zooplankton_new.columns.levels[0]:
    di[i] = df_zooplankton_new[i]


#Defining Copepods dataframe:
Copepods=di.get('Copepods')
columns_to_drop_cop=['Calanus spp.', 'Euterpina', 'Longipedia']
New_Copepods=Copepods
for i in columns_to_drop_cop:
    New_Copepods=New_Copepods.drop(i, axis=1)


#Creating a dictionary to rename columns of Copepods:
abbrev={}
for i in New_Copepods.columns.get_level_values(0):
    if i=='Pseudo/Paracalanus':
        abbrev.update( {i : 'Pseu/Para'} )
    elif i != 'others':
        split_name = i.split()
        if len (split_name)>1:
            if  split_name[1]=='spp.' or split_name[1]=='spp':
                abbrev.update( {i : split_name[0][0:4]} )
            else:
                abbrev.update( {i : split_name[0][0] + split_name[1][0:3]} )
        else:
            abbrev.update( {i : i[0:4]} )


#Renaming columns:
for i in abbrev:
    New_Copepods=New_Copepods.rename(columns={i:abbrev.get(i)})
New_Copepods.head()


def Others_group():
    '''Function gets Others group without irrelevant species for the analysis'''
    Others=di.get('others')
    Others=Others.droplevel('Stage', axis=1)
    columns_to_drop=['Diastylis', 'Mysidae', 'Facetotecta L', 'Keratella ', 'Aurelia', 'Rathkea', 'Sarsia', 'Ophiura', 'Spionidae L', 'Trochophora',
'Nematoda', 'Turbellaria', 'Phoronis']
    #Summing up those columns: polychaetes (sum of polychaeta L, polychaeta T, Spionidae, Harmothoe),
    Others['Polychaetes']=Others.loc[:,['Polychata L', 'Polychata T', 'Harmothoe', 'Spionidae L']].sum(axis=1)
    Others['Month'] = Others.index.month
    columns_to_drop.extend(['Polychata L', 'Polychata T', 'Harmothoe'])
    New_others=Others
    for i in columns_to_drop:
        New_others=New_others.drop(i, axis=1)
    return New_others
```

```python
def Cladocera_group():
    '''Function gets Cladocera group without irrelevant species for the analysis'''
    Cladocera=di.get('Cladocera')
    Cladocera=Cladocera.droplevel('Stage', axis=1)
    Cladocera['Month']=Cladocera.index.month
    New_Cladocera=Cladocera.drop(['others'], axis=1)
    return New_Cladocera


def Copepods_sep():
    '''Function gets Copepods where adults, copepodites and nauplii are searated'''
    Copepods_sep=New_Copepods.copy()
    for i in New_Copepods.columns.get_level_values(0):
        #separating Copepodites
        if (i,'Copepodites') in Copepods_sep:
            name=i+'J'
            Copepods_sep[name, 'J']=Copepods_sep[i]['Copepodites']
            Copepods_sep=Copepods_sep.drop([(i, 'Copepodites')], axis=1)
        if (i,'J1') in Copepods_sep and (i,'J2') in Copepods_sep:
            name=i+'J1'
            Copepods_sep[name, 'J1']=Copepods_sep[i]['J1']
            name=i+'J2'
            Copepods_sep[name, 'J2']=Copepods_sep[i]['J2']
            name=i+'J1/J2'
            Copepods_sep[name, 'J1+J2']=Copepods_sep[i]['J2']+Copepods_sep[i]['J1']
            Copepods_sep=Copepods_sep.drop([(i, 'J1')], axis=1)
            Copepods_sep=Copepods_sep.drop([(i, 'J2')], axis=1)
        #separating Nauplii
        if (i,'Nauplii') in Copepods_sep:
            name=i+' Nauplii'
            Copepods_sep[name, 'Nauplii']=Copepods_sep[i]['Nauplii']
            Copepods_sep=Copepods_sep.drop([(i, 'Nauplii')], axis=1)
        #separating indetermined
        if (i,'indet') in Copepods_sep:
            name=i+' indet'
            Copepods_sep[name, 'indet']=Copepods_sep[i]['indet']
            Copepods_sep=Copepods_sep.drop([(i, 'indet')], axis=1)
        #separating males/females
        if (i,'females') in Copepods_sep and (i,'males') in Copepods_sep:
            name=i+'F'
            Copepods_sep[name, 'females']=Copepods_sep[i]['females']
            name=i+'M'
            Copepods_sep[name, 'males']=Copepods_sep[i]['males']
            name=i+'F/M'
            Copepods_sep[name, 'F/M']=Copepods_sep[i]['females']+Copepods_sep[i]['males']
            Copepods_sep=Copepods_sep.drop([(i, 'females')], axis=1)
            Copepods_sep=Copepods_sep.drop([(i, 'males')], axis=1)
        #summing up females/males/copepodites (for Temora longicornis, Oithona similis, Limnocalanus macurus)
        name=i+'J1/J2'
        name2=i+'F/M'
        name3=i+'F/M/J1/J2'
```

```
        name4=i+'J'
      if (name,'J1+J2') in Copepods_sep and (name2,'F/M') in Copepods_sep:
         Copepods_sep[name3,'F/M/J1/J2']=Copepods_sep[name]['J1+J2']+Copepods_sep[name2]['F/M']
      elif (name4,'J') in Copepods_sep and (name2,'F/M') in Copepods_sep:
         Copepods_sep[name3,'F/M/J']=Copepods_sep[name4]['J']+Copepods_sep[name2]['F/M']
   #Summing up sopepodites with adults where copepodites of different species can't be separated
   #for Acartia summing up all adults + copepodites:

Copepods_sep['AcarF/M/J1/J2','F/M/J1/J2']=Copepods_sep['AcarJ1/J2']['J1+J2']+Copepods_sep['AbifF/M']['F/M']+Copepods_sep['AlonF
/M']['F/M']+Copepods_sep['AtonF/M']['F/M']
   #for Centropages

Copepods_sep['CentF/M/J1/J2','F/M/J1/J2']=Copepods_sep['CentJ1/J2']['J1+J2']+Copepods_sep['ChamF/M']['F/M']+Copepods_sep['Cty
pF/M']['F/M']
   #for Eurytemora
   Copepods_sep['EuryF/M/J1/J2','F/M/J1/J2']=Copepods_sep['EuryJ1/J2']['J1+J2']+Copepods_sep['EaffF/M']['F/M']
   #for Pseudo/Paracalanus

Copepods_sep['Pseu/Para','F/M/J1/J2']=Copepods_sep['Pseu/ParaJ1/J2']['J1+J2']+Copepods_sep['PparF/M']['F/M']+Copepods_sep['Ps
euF/M']['F/M']
   Copepods_sep_sum=Copepods_sep.sum(level=0, axis=1)
   Copepods_sep_sum['Month']=Copepods_sep.index.month
   return Copepods_sep_sum


def Copepods_comb():
   '''Function gets Copepods where adults are treated together with not adult forms'''
   Copepods_comb=New_Copepods.copy()
   Copepods_comb_sum=Copepods_comb.sum(level=0, axis=1)
   Copepods_comb_sum['Acar']= Copepods_comb_sum.loc[:, ['Acar', 'Abif', 'Alon', 'Aton' ]].sum(axis=1)
   Copepods_comb_sum['Cent']= Copepods_comb_sum.loc[:, ['Cent', 'Cham', 'Ctyp']].sum(axis=1)
   Copepods_comb_sum['Eury']= Copepods_comb_sum.loc[:, ['Eury', 'Eaff']].sum(axis=1)
   Copepods_comb_sum['Pseu/Para']= Copepods_comb_sum.loc[:, ['Pseu/Para', 'Ppar', 'Pseu']].sum(axis=1)
   Copepods_comb_sum=Copepods_comb_sum.drop(['Abif', 'Alon', 'Aton' ], axis=1)
   Copepods_comb_sum=Copepods_comb_sum.drop(['Cham', 'Ctyp'], axis=1)
   Copepods_comb_sum=Copepods_comb_sum.drop(['Eaff'], axis=1)
   Copepods_comb_sum=Copepods_comb_sum.drop(['Ppar', 'Pseu'], axis=1)
   Copepods_comb_sum['Month']=New_Copepods.index.month
   Copepods_comb_sum.head()
   return Copepods_comb_sum
```

Appendix 3.3   Module to perform descriptive analysis

```
# functions to analyze data
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
import matplotlib.dates as mdates
import seaborn as sns
import numpy as np
from pandas.plotting import autocorrelation_plot
```

```
from scipy.stats import spearmanr

def plot_boxplots(df):
    ''' Function plots boxplots for the dataframe, where the last column is month'''
    num_of_col=len(list(df.columns))-1
    cols_to_plot=list(df.columns[0:num_of_col])
    fig, axes = plt.subplots(num_of_col, 1, figsize=(20, num_of_col*5), sharex=True)
    for name, ax in zip(cols_to_plot, axes):
        sns.boxplot(data=df, x='Month', y=name, ax=ax)
        ax.set_title(name)
    return axes
def one_column(df, i):
    '''Function to return i'th the column of the dataset without NA values'''
    df = df.dropna()._get_numeric_data() ##droping missing values
    column = pd.DataFrame(df[df.columns[i]])
    return column


def plot_autocorrelation(df):
    ''' Function plots autocorrelation plots'''
    cols=df.columns[:len(df.columns)-1]
    df_plot=df.loc[:,cols]
    df_plot=df_plot.astype(float) #type of values-float
    for i in range(len(df_plot.columns)):
        column_to_plot=one_column(df_plot,i)
        autocorrelation_plot(column_to_plot)
```

Appendix 3.4   Module to get data frames with median value per month

```
# functions to get dataframes with median for the specific period
import pandas as pd

def add_period_column(df, period):
    '''Function adds 'two-month period' and 'Year' columns to the dataframe. The second parameter should be one of the 'Season'/'Two-
months'/'Month'/'''
    df=df.apply(pd.to_numeric)
    if  period == 'Month': #part of the code was deleted, because two-monts and seasonal periods are not used anymore
        df['Month2'] = df.index.month
    else:
        print('Period parameter is not correct for add_period_column function.')
    df['Year'] = df.index.year
    return df


def genus_median_per_period (df, period):
    '''Function returns a dataframe were median for the column is calucalated for the specific period. The second paramater is the name of
the column of period'''
    a=df.groupby(['Year', period])[df.columns[0]].median()
    for i in range(1, len(df.columns)-3):
        b=df.groupby(['Year', period])[df.columns[i]].median()
        a=pd.concat([a, b], axis=1, sort=False)
    a=pd.DataFrame(a)
```

```
    return (a)


def form_multiindex_season(df, period):
    '''Function returns multiindex. The second parameter should be one of the 'Season'/'Two-months'/'Monthly'/'''
    min_y=min(df.index).year
    max_y=max(df.index).year+1
    years=[]
    for i in range(min_y, max_y):
        years.append(i)
    if period == 'Month':
        iterables =[years, [1,2,3,4,5,6,7,8,9,10,11,12]]
    else:
        print('Period is incorrect')
    indexes=pd.MultiIndex.from_product(iterables)
    return indexes


def return_monthly(df):
    ''' Returns a dataframe where indexes are monthly periods for every year'''
    df_mon = add_period_column(df, 'Month')
    df_mon_med = genus_median_per_period(df_mon, 'Month')
    new_index = form_multiindex_season(df_mon, 'Month')
    df_mon_med = df_mon_med.reindex(new_index)
    return df_mon_med
```

Appendix 3.5   Module to fill the missing values in monthly data

```
# filling monthly data
import pandas as pd


def first_rolling(df):
    '''Fills data-frame rolling average, than median for the specific month'''
    df_filled=df.copy()
    for col in df.columns:
        ser=df.loc[:, col]
        ser=ser.fillna(ser.rolling(3,min_periods=1, center=True).mean())
        df_filled.loc[:, col]=ser
        for month in list(df_filled.index.levels[1]):
            ser=df_filled.loc[pd.IndexSlice[:, month], col]
            ser.fillna(ser.median(), inplace=True) ###includes previously filled values into median calculation
            df_filled.loc[pd.IndexSlice[:, month], col]=ser
    return(df_filled)
```

Appendix 3.6   Module to find species of specific season and correlation to hydrography variables

```
# Finding species for specific season

import pandas as pd
```

```python
import numpy as np
import scikit_posthocs as sp
import scipy
from scipy.stats import spearmanr
import statistics


def season_with_highest_abundance(df_filled):
    '''Returns the dataframe with median values for the seasons when species are most abundant'''
    list_of_rows=[]
    seasons=['spring', 'summer', 'autumn', 'winter']
    for col in df_filled.columns:
        row={}
        row['Variable']=col
        spring=df_filled.loc[pd.IndexSlice[:, [3,4,5]], col]
        summer=df_filled.loc[pd.IndexSlice[:, [6,7,8,9]], col]
        autumn=df_filled.loc[pd.IndexSlice[:, [10,11,12]], col]
        winter=df_filled.loc[pd.IndexSlice[:, [1,2]], col]
        #Kruskal test:
        Kruskal_p=scipy.stats.kruskal(spring, summer, autumn, winter, nan_policy='omit')[1]
        if Kruskal_p>0.05:
            print (col, 'No difference between seasons')
        else:
        #post hoc test
            x = [spring, summer, autumn, winter]
            df=sp.posthoc_dunn(x, p_adjust = 'holm')
            df=df[df<0.05]
        #finding when the median is the highest:
            medians=[spring.median(), summer.median(), autumn.median(), winter.median()]
            highest_idx=[medians.index(max(medians))]
        #checking if there are some not significantly different:
            highest_seasons=list(np.where(np.isnan(df.iloc[medians.index(max(medians)),:]))[0])
            for i in highest_seasons:
                if i not in highest_idx:
                    highest_idx.append(i)
            for idx in highest_idx:
                row[seasons[idx]]=medians[idx]
            list_of_rows.append(row)
    df_median_on_season=pd.DataFrame(list_of_rows)
    df_median_on_season=df_median_on_season.set_index('Variable')
    df_median_on_season.loc[:,['spring', 'summer', 'autumn', 'winter']]
    return(df_median_on_season)


def season_species_table(df_filled, df_hydro, season):
    '''Returns the data for the specific season and species, which are the most abundant at that particular season'''
    df_median_on_season=season_with_highest_abundance(df_filled)
    if season == 'spring':
        season_species=df_filled[df_median_on_season[df_median_on_season.spring.notnull()].index]
        season_species=season_species.loc[pd.IndexSlice[:, [3,4,5]], :]
        season_hydro=df_hydro.loc[pd.IndexSlice[:, [3,4,5]], :]
    elif season == 'summer':
        season_species=df_filled[df_median_on_season[df_median_on_season.summer.notnull()].index]
```

```python
            season_species=season_species.loc[pd.IndexSlice[:, [6,7,8,9]], :]
            season_hydro=df_hydro.loc[pd.IndexSlice[:, [6,7,8,9]], :]
        elif season == 'autumn':
            season_species=df_filled[df_median_on_season[df_median_on_season.autumn.notnull()].index]
            season_species=season_species.loc[pd.IndexSlice[:, [10,11,12]], :]
            season_hydro=df_hydro.loc[pd.IndexSlice[:, [10,11,12]], :]
        elif season == 'winter':
            season_species=df_filled[df_median_on_season[df_median_on_season.winter.notnull()].index]
            season_species=season_species.loc[pd.IndexSlice[:, [1,2]], :]
            season_hydro=df_hydro.loc[pd.IndexSlice[:, [1,2]], :]
    season_species=season_species.groupby(level=0).mean() #mean value for the season in particular year
    season_hydro=season_hydro.groupby(level=0).mean()
    season_hydro=season_hydro.loc[:2016,:]
    return season_species, season_hydro


def correlation_table(season_species, season_hydro):
    '''returns the correlation between 2 dataframes variables'''
    list_of_dict=[]
    for taxa in season_species.columns:
        series1 = season_species[taxa].loc[1998:2016]
        row={'Variable':taxa}
        for var in season_hydro.columns:
            series2 = season_hydro[var].loc[1998:2016]
            corr = spearmanr(series1, series2)
            if corr[1]<0.05:
                row[var]=corr[0]
            else:
                row[var]=np.nan
        list_of_dict.append(row)
    cor_to_hydro=pd.DataFrame(list_of_dict)
    cor_to_hydro=cor_to_hydro.set_index('Variable')
    return (cor_to_hydro)


def compare_corr_spring_whole_year(dataframe_to_compare, depth0_4_comb_monthly_full):
    'Compares correlation coefficients calculated for spring or for the whole year, when correlation is calculated for the species most
abundant in spring'
    plankton_spring, hydro_spring=season_species_table(dataframe_to_compare, depth0_4_comb_monthly_full, 'spring')
    corr_spring=correlation_table(plankton_spring, hydro_spring)
    cor_all=correlation_table(dataframe_to_compare, depth0_4_comb_monthly_full)
    cor_all=cor_all.loc[plankton_spring.columns,]#comparison only for species abundant in spring
    abs_val=np.abs(cor_all)
    data=abs_val.values[pd.notna(abs_val.values)]
    abs_val2=np.abs(corr_spring)
    data2=abs_val2.values[pd.notna(abs_val2.values)]
    #Correlation through the year
    median_corr_spring=round(statistics.median(data), 3)
    IQR_corr_spring=str(round(np.quantile(data, 0.25),3)) + ' - ' + str(round(np.quantile(data, 0.75),3))
    #Correlation on spring
    median_corr_year=round(statistics.median(data2), 3)
    IQR_corr_year=str(round(np.quantile(data2, 0.25),3)) + ' - ' + str(round(np.quantile(data2, 0.75),3))
    return(median_corr_spring, IQR_corr_spring, median_corr_year, IQR_corr_year)
```

```python
def return_dict(df_analysed, median_corr_spring, IQR_corr_spring, median_corr_year,IQR_corr_year):
    dict_of_medians={}
    dict_of_medians['Data frame analysed']=df_analysed
    dict_of_medians['Median_corr_spring']=median_corr_spring
    dict_of_medians['IQR_corr_spring']=IQR_corr_spring
    dict_of_medians['Median_corr_year']=median_corr_year
    dict_of_medians['IQR_year']=IQR_corr_year
    return dict_of_medians
```

Appendix 3.7   Module for non-metric multidimensional scaling analysis

```python
# nMDS analysis
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import math
from sklearn import manifold
from sklearn.preprocessing import StandardScaler
from dtaidistance import dtw_visualisation as dtwvis
from dtaidistance import dtw
from sklearn.metrics import euclidean_distances
from scipy.cluster.hierarchy import linkage, dendrogram

# To represent how years differ from each other:

def calculate_dist_betw_years(df, year1, year2):
    '''Calculates the distance between two years'''
    dist_list=[]
    for col in df.columns:
        s1 = np.array(df.loc[year1, col])
        s2 = np.array(df.loc[year2, col])
        distance = dtw.distance(s1, s2)
        dist_list.append(distance)
    dist=0
    for i in dist_list:
        dist=dist+i*i
    gen_dist=math.sqrt(dist)
    return(gen_dist)

def distance_matrix(df_filled):
    '''Calculates distance matrix'''
    D=[]
    years=set(df_filled.index.get_level_values(0))
    for year1 in years:
        dist_ar=[]
        for year2 in years:
            dist_ar.append(calculate_dist_betw_years(df_filled, year1, year2))
        D.append(dist_ar)
```

```python
    D=np.array(D)
    return(D)


def plot_2dim_years_in_clusters(df_filled, title): ## idea from: https://stackoverflow.com/a/61964297
    'Function plots 2-D plots and colours points based on dendrogram colours'
# DATA EXAMPLE
    mds = manifold.MDS(2, dissimilarity='precomputed', metric=False, n_init=20)
    max_value=distance_matrix(df_filled).max()
    D=(distance_matrix(df_filled))/max_value #normalizing distances
    coords = mds.fit_transform(D)


    #stress value:
    # Calculation of stress according to the https://stackoverflow.com/a/64271501
    points=mds.embedding_
    ## Manual calculus of sklearn stress
    DE = euclidean_distances(points)
    DE = DE/DE.max()  #normalizing distances
    stress = 0.5 * np.sum((DE - D)**2)
    ## Kruskal's stress (or stress formula 1)
    stress1 = np.sqrt(stress / (0.5 * np.sum(D**2)))


    # DENDROGRAM
    labels_for_graph=list(set(df_filled.index.get_level_values(0)))
    #plt.figure(figsize=(5, 5))
    plt.subplot(121)
    z = linkage(coords, 'single')
    d = dendrogram(z, labels=labels_for_graph)
    plt.yticks([])


    # COLORED PLOT
    plt.subplot(111)
    points = d['leaves']
    colors = ['none'] * len(points)
    for xs, c in zip(d['icoord'], d['color_list']):
        for xi in xs:
            if xi % 10 == 5:
                colors[(int(xi)-5) // 10] = c


    #labels for scatter plot
    all_years=list(set(df_filled.index.get_level_values(0)))
    new_years=[]
    for i in d['leaves']:
        new_years.append(all_years[i])


    title_to_show=title+', s='+str(round(stress1,3))
    for point, color, year in zip(points, colors, new_years):
        plt.plot(coords[point, 0], coords[point, 1], 'o', color=color)
        plt.text(coords[point, 0], coords[point, 1], f' {year}')
        plt.title(title_to_show, fontsize=15)
        plt.xlabel('Coordinate 1', fontsize=15)
        plt.ylabel('Coordinate 2', fontsize=15)
```

```
    plt.xticks(fontsize=12)
plt.tight_layout()
saving_name='output/mds/'+title+'.png'
plt.savefig(saving_name, dpi=400)
plt.show()
```

Appendix 3.8   Module to perform Mann-Kendall test and detect correlations

```
# Python module to return Mann Kendal test results and correlation
import pandas as pd
import numpy as np
import pymannkendall as mk
import scipy
from scipy.stats import spearmanr


def return_seasons(df_filled):
    '''Function mean for specific season values'''
    spring=df_filled.loc[pd.IndexSlice[:, [2,3,4]], :].groupby(level=0).mean()
    summer=df_filled.loc[pd.IndexSlice[:, [5,6,7,8]], :].groupby(level=0).mean()
    autumn=df_filled.loc[pd.IndexSlice[:, [9,10,11]], :].groupby(level=0).mean()
    winter=df_filled.loc[pd.IndexSlice[:, [12,1,2]], :].groupby(level=0).mean()
    return (spring, summer, autumn, winter)


def Mann_Kendal_1_season(season_df):
    '''Performs Mann Kendall test for seasonal data'''
    columns=[]
    results=[]
    for col in season_df.columns:
        ser=season_df.loc[:,col]
        result = mk.original_test(ser)
        if result[1]==True:
            #print (col, '', result[0])
            columns.append(col)
            results.append(result[0])
    return(columns, results)


def rows_Mann_kendal_for_season(columns, results, season_name, df_correlation_table):
    '''Creates list of dictionaries for each season with plankton with changing trends and correlation in interest'''
    list_of_dict_per_season=[]
    dict_row={}
    for i in range(len(columns)):
        dict_row={}
        dict_row['Season']=season_name
        dict_row['Plankton with changing trends']=columns[i]
        dict_row['Mann Kendal test result']=results[i]
        cor_values=return_correlation_values(df_correlation_table, columns[i])
        dict_row['Correlation']=cor_values
        list_of_dict_per_season.append(dict_row)
    return (list_of_dict_per_season)
```

```python
def correlation_table(season_species, season_hydro):
    '''returns the correlation between 2 dataframes variables'''
    list_of_dict=[]
    for taxa in season_species.columns:
        series1 = season_species[taxa].loc[1998:2016]
        row={'Variable':taxa}
        for var in season_hydro.columns:
            series2 = season_hydro[var].loc[1998:2016]
            corr = spearmanr(series1, series2)
            if corr[1]<0.05:
                row[var]=corr[0]
            else:
                row[var]=np.nan
        list_of_dict.append(row)
    cor_to_hydro=pd.DataFrame(list_of_dict)
    cor_to_hydro=cor_to_hydro.set_index('Variable')
    # cor_to_hydro=cor_to_hydro.dropna(how='all')
    return (cor_to_hydro)


def return_correlation_values(correlation_table, column):
    '''Return correlation variables and values from the correlation table'''
    cor_values=correlation_table.loc[column,:][correlation_table.loc[column,:].notna()]
    string_to_return=[]
    for i in range(len(cor_values.index)):
        var=cor_values.index[i]
        value=round(cor_values.values[i],4)
        string_var_value=var + ' ' + str(value)
        string_to_return.append(string_var_value)
    return(string_to_return)


def return_seasonal_dataframes_list(df_monthly_filled):
    '''returns list of dataframes for specific season'''
    spring, summer, autumn, winter = return_seasons(df_monthly_filled)
    df_list=[spring, summer, autumn, winter]
    return(df_list)


def table_trends_and_correlations(df_plankton_filled_monthly, df_1_filled_monthly):
    '''Takes plankton monthly data and hydrography (or other plankton) data and returns trends and correlations'''
    dfs_season=return_seasonal_dataframes_list(df_plankton_filled_monthly)
    df_hydro_season=return_seasonal_dataframes_list(df_1_filled_monthly)
    df=pd.DataFrame()
    season_names=['Spring', 'Summer', 'Autumn', 'Winter']
    for i in range(len(dfs_season)):
        columns, results=Mann_Kendal_1_season(dfs_season[i])
        df_correlation_table=correlation_table(dfs_season[i], df_hydro_season[i])
        list_of_dict_per_season=rows_Mann_kendal_for_season(columns, results, season_names[i], df_correlation_table)
        if len(list_of_dict_per_season)>0 and i==0:
            df=pd.DataFrame(list_of_dict_per_season)
        elif len(list_of_dict_per_season)>0:
            df1=pd.DataFrame(list_of_dict_per_season)
            df=pd.concat([df, df1], axis=0)
```

```
idx=df.loc[:,['Season', 'Plankton with changing trends']]
idx=pd.MultiIndex.from_frame(idx)
df.index=idx
df=df.drop(['Plankton with changing trends', 'Season'], axis=1)
df=df[['Mann Kendal test result','Correlation']]
return(df)
```