

Predictability of Marine Ecosystems in a Changing Climate

DISSERTATION

zur

Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Rostock

vorgelegt von

Frau Dipl.-Oz. Karin Junker
geb. Hartig am 06.06.1967 in Hamburg
aus Stockholm, Schweden

Datum der Einreichung: 30. April 2013

Tag der mündlichen Prüfung: 28. Oktober 2013

Referent: PD Dr. Joachim Dippner

Koreferent: Prof. Dr. Ilppo Vuorinen

Meinem Mann Roland und unseren Söhnen Moritz und Till.

Acknowledgements

This dissertation would not have been possible without the guidance and the help of many people who contributed in one way or another to this study.

First and foremost, my utmost gratitude belongs to my supervisor and mentor Joachim W. Dippner. Without his broad scientific expertise, ideas and patience this thesis would never have been completed. He trusted my capabilities sometimes more than I did myself and gave me that little extra encouragement whenever needed. Thank you, Joachim.

I am also very grateful to Ilppo Vuorinen. He supported my work with valuable comments whenever we met, wrote a great letter of support and at even kindly agreed to be assessor of this thesis.

Gratitude also to Heide Schulz-Vogt who managed to find financial support for my travel to the ASLO conference in New Orleans, so I could chair my session. Many thanks also for kindly agreeing to assess this dissertation.

Then, I would like to thank the many unnamed people who take the samples and do the measuring and counting in the labs. Their data was kindly provided me by Mario von Weber and Sigrid Sagert, who supplied the monitoring data of the Landesamt für Umwelt, Naturschutz und Geologie of Mecklenburg-Western Pomerania, Georg Kornilovs and Saskia Otto, who provided the zooplankton data, Ingrid Kröncke supplied the data on benthic macrofauna. My gratitude belongs also to Norbert Wasmund, who is in charge of the monitoring of phytoplankton in the Baltic Sea and supplied valuable information on the monitoring procedure. Miguel Rodriguez-Medina kindly provided the runoff data and Deliang Chen supplied the vorticity for the construction of the Baltic Sea Environmental index. The contact to Dr. Chen was made by Thorsten Blenckner, who also provided valuable comments on my experiments.

Further, I would like to thank Dusan Sovilj, who made the uncounted cal-

culations with the neural networks trying to model the macrozoobenthos with often frustrating results.

I should also not forget to mention Christoph Humborg who supported me by providing space to work at the Baltic Nest Institute in Stockholm. Without that, it wouldn't have been possible to spend so much time in Stockholm with my family.

My husband Roland and my sons Moritz and Till belongs my warmest gratitude. They accepted patiently long weeks without their wife and mother and supported me morally even when I was in Warnemünde in hours of "skyping". Thank you guys!

Erklärung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig angefertigt und ohne fremde Hilfe verfasst habe. Dazu habe ich keine außer den von mir angegebenen Hilfsmitteln und Quellen verwendet und die den benutzten Werken inhaltlich und wörtlich entnommenen Stellen habe ich als solche kenntlich gemacht.

Declaration

I hereby declare under oath that I have completed the work submitted here independently and have composed it without outside assistance. Furthermore, I have not used anything other than the resources and sources stated and where I have taken sections from these works in terms of content or text, I have identified this appropriately.

Rostock, 30. April 2013

Am 30. September 2013 wurden gemäß der Gutachten einige sprachliche Fehler korrigiert.

Abstract

Ecosystem predictability is the basis for ecosystem management to reach a good environmental status and to mitigate negative effects from anthropogenic impact. To shed light on the influence of climate on the variability and predictability of marine ecosystem, various climate indices are related to ecosystem descriptors in this thesis. The various aspects influencing ecosystem predictability are studied and discussed. It is shown that predictability of marine ecosystems is altered by large-scale transitions in the atmosphere. A regime shift in 2001/2002 is identified in the atmosphere as well as in the marine ecosystem, which resulted in an increase in non-linearity. The resulting decrease of predictability could be compensated by choosing a multivariate climate descriptor that combines large-scale climate indices with regional-scale descriptors.

Zusammenfassung

Die Vorhersagbarkeit von Ökosystemen ist die Grundlage für erfolgreiches Ökosystemmanagement, um einen guten Zustand des Ökosystems trotz negativer anthropogener Einflüsse zu erreichen. Um die Bedeutung des Klimas auf die Variabilität und Vorhersagbarkeit auf marine Ökosysteme zu untersuchen, werden in dieser Arbeit Klimaindizes mit verschiedenen Deskriptoren mariner Ökosysteme in Beziehung gesetzt. Die verschiedenen Faktoren, welche die Vorhersagbarkeit von Ökosystemen beeinflussen, werden untersucht und diskutiert. Es kann gezeigt werden, dass großskalige Veränderungen in der Atmosphäre die Vorhersagbarkeit von Ökosystemen stark beeinflussen können. Es kann eine Zustandsänderung des Systems in 2001/2002 sowohl in der Atmosphäre als auch im marinen Ökosystem gezeigt werden, die eine stärkere Nichtlinearität zur Folge hatte. Es wird gezeigt, dass die resultierende Herabsetzung der Vorhersagbarkeit durch geeignete Kombination von großskaligen und regionalen Klimadeskriptoren kompensiert werden kann.

Contents

1	Introduction	1
1.1	Regime Shifts and Predictability	2
1.2	Climate change and impact on ecosystems	6
1.3	Objectives and Outline of this Thesis	9
2	Material and Methods	11
2.1	Data Basis	11
2.1.1	North Sea Benthos	11
2.1.2	Baltic Sea Zooplankton	12
2.1.3	Baltic Sea Phytoplankton	14
2.1.4	Nutrients	17
2.1.5	Climate Predictors and Physical Variables	17
2.2	Methods	20
2.2.1	Analysis of Empirical Orthogonal Functions (EOF)	20
2.2.2	Wavelet Analysis	22
2.2.3	Statistical Downscaling	25
2.2.4	Neural Networks	29
3	Experiments	35
3.1	Shifts in Benthos	35
3.1.1	Observed shifts in benthos and climate	35
3.1.2	Benthos Predictability Using Neural Networks	37
3.2	The Baltic Sea Environmental (BSE) Index	44
3.2.1	Construction of the BSE	44
3.2.2	Performance Tests	49
3.3	Baltic Sea Phytoplankton Predictability	53

3.3.1	Results EOF Analysis	53
3.3.2	Downscaling	56
4	Discussion	63
4.1	Occurrence of regime shifts	63
4.2	Enhancing Prediction: the Baltic Sea Index	66
4.3	Predictability of Phytoplankton	68
4.3.1	System Dynamics	70
4.3.2	Data Issues	72
5	Conclusion	77
	Bibliography	80
A	Abbreviations	95
B	Supporting Information	99
B.1	Arkona Sea EOF Analysis	99
B.2	LUNG data EOF Analysis	113
B.2.1	All Stations	113
B.2.2	All Stations except S66 and KHM	121
B.2.3	Only Outer Coastal Zone Stations	129
B.2.4	Only Inner Coastal Zone Stations	137
B.3	Arkona Sea Phytoplankton Downscaling 1979-2005	145
B.4	Arkona Sea Phytoplankton Downscaling 1979-2000	158

Chapter 1

Introduction

Ecosystems are complex adaptive systems exhibiting homeostatic and self-organizing features and inherent non-linear dynamics that imply a certain degree of unpredictability (Levin 1999). Assessing the degree of predictability is of vital importance for future ecosystem management aiming at conserving the ecosystems ability to provide goods and services to society. These management actions are relying on predictions and projections of future ecosystem state to estimate their potential for success. Moreover, not only predictions, but also the estimation of the degree of uncertainty in these predictions are just as important for management actions.

Today, it is well known that human action has far-reaching impact on marine ecosystems. Heavy exploitation, eutrophication and toxic substances are often threatening the ecosystems function and resilience, especially in densely inhabited areas. Nutrient inflow can cause eutrophication (Wulff et al. 2001) and hazardous substances are a major threat for the biodiversity (HELCOM 2009). Intense fishing is removing the reproductive fish and thus alters fish stocks capability to reproduce in necessary amounts (Stenseth et al. 2004). All these influences have extensive impact on ecosystems by altering the food web and internal processes.

On top of these, in many areas there are quite serious threats for ecosystem functioning and climate change is increasingly stressing ecosystem resilience, possibly pushing it over the brink, causing it to shift to a new state. These shifts are not only undesirable because the new state is usually not favorable, but also because they are increasing non-linearity and decreasing predictabil-

ity of ecosystems. Depending on the type of shift, they can even render an ecosystem unpredictable (Fig. 1.1).

The next sections in this chapter give a short introduction on predictability and regime shifts, the possible pathways climatic influence on ecosystems and their predictability in the North and Baltic Sea.

1.1 Regime Shifts and Predictability

Whether a dynamical system can be predicted depends basically only on two things: one is amount of noise from sampling error and unpredictable behaviour of a forcing variable, the so-called predictor, the other is the degree of complexity of the system inherent dynamics. I.e. the degree of non-linearity of the dynamics determines how fast the skill of a prediction decreases. In a system, where a physical forcing variable exhibits chaotic dynamics, the non-linearity of the system determines whether it remains predictable despite of the disturbances. Usually, predictive skill declines with prediction time - unless the system is dominated by noise, where the error remains constant. This decay of prediction skill can be used to estimate the complexity of a system. If the decay is slow, the system can be considered as only weakly non-linear, as opposed to strongly non-linear, where the decay is fast. That constant remaining part of the error can be attributed to noise. A number of methods have been developed for this purpose (e.g. Sugihara and May 1990; Tsonis and Elsner 1992; Wales 1991). However, the methods proposed by the mentioned authors require a fair amount of data points (Wales (1991) called time series with ~ 1000 points sparse) in order to resolve the dynamical structure of the system. Unfortunately, only few time series for the marine environment have this quality.

Recently, the term *regime shift* has come to some popularity. It is a representation of a non-linear response of an ecosystem to a disturbance, where the system moves from one stable state, in terms of chaos theory *basin of attraction*, to another. Depending on the degree of non-linearity in the system and the systems status, this shift might occur in a chaotic way, thus decreasing predictability of an ecosystem. Following the concept of Collie et al. (2004), the possible equilibriums of an ecosystem can be pictured as points on a topological

catastrophe manifold, where type of shift depends on the systems complexity i.e. non-linearity and external forcing (Fig. 1.1). The extent of predictability depends on the degree of non-linearity between the forcing variable, or predictor, and the response. Especially when the system responds in a discontinuous manner to a disturbance will predictability decrease. However, not all regime shifts have to be discontinuous and triggered by single events. Gradual change might gradually deprive an ecosystem of its resilience until it eventually shifts to a new state.

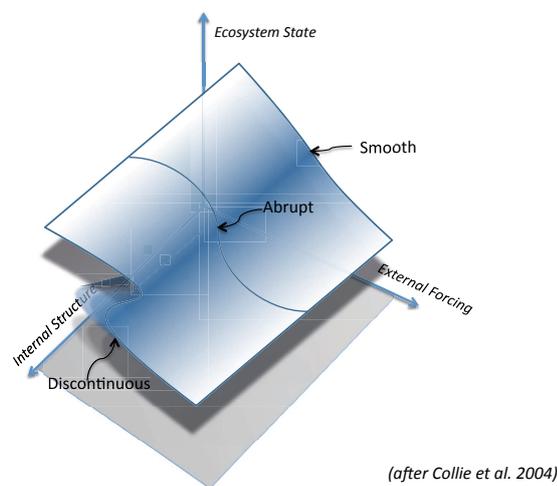


Figure 1.1: Topological catastrophe manifold in analogy to Collie et al. (2004). Each point on the surface represents equilibrium values of a variable representative for an ecosystem e.g. abundance reacting to external forcing or internal dynamics. In case of a discontinuous regime shift, the pathway of the shift is undefined and several possible stable states exist for the same external forcing.

However, the exact definition and theory of regime shifts is still subject to discussion. The next few paragraphs will explain the scope in which the term is used in this study and clarify some of its concepts.

Regime Shift Definition

It is widely accepted that the term *regime shift* refers to rapid changes in the structure of a dynamic system. Defining a *regime* of a quasi-stable dynamic system as a characteristic behaviour of a natural phenomenon (sea level pres-

sure, recruitment, etc.) over time, the regime *shift* can be defined as a fast transition from one dynamic regime to another (Hare and Mantua 2000). Collie et al. (2004) gave an overarching definition of a regime shift for marine ecosystems as low-frequency, high-amplitude changes in oceanic conditions that may propagate through several trophic levels and be especially pronounced in biological variables. This definition follows Bakun (2005) who defined a regime shift as "a persistent radical shift in typical levels of abundance or productivity of multiple important components of the marine biological community structure, occurring at multiple trophic levels and on a geographical scale that is at least regional in extent".

Some differences in definition can be found regarding climate and biological regime shifts. Climate regime shifts refer to changes in trend of a physical atmospheric property, such as temperature, with regimes lasting decades (Swanson and Tsonis 2009), while in biological systems, the term regime shift usually implies changes over several trophic levels. The duration of the new biological regime is only defined relatively, i.e. the time span in which an ecosystem shifts from one regime to another has to be much shorter than the duration of a regime (Reid et al. 2001).

Owing to the differences in definition, the term "biological regime shift" (BRS) will be used according to Reid et al. (2001) and the term "climate regime shift" (CRS) according to Swanson and Tsonis (2009) in the following.

Regime Shift Theory

Forced by external drivers such as climate, nutrient inputs or toxic substances, ecosystems can shift in different ways to a new regime that can be classified into three major types: smooth, abrupt and discontinuous shift (Fig. 1.1) (Collie et al. 2004; Scheffer et al. 2001). These types correspond to different degrees of non-linearity in the system, ranging from quasi-linear or weakly non-linear to highly non-linear.

An ecosystem can respond to a gradual trend of a driver by taking on a new state in a smooth and likewise gradual way, which is indicating quasi-linear relationships between driver and response variables, or it can exhibit a certain degree of resilience until it shifts abruptly to a new state, which is indicative of a non-linear relationship between driver and response variable.

The abrupt change can occur in a predictable way, with only one possible stable state depending on the driving forces, which would apply to a relatively weak non-linear system, but also in an unpredictable, catastrophic way, which corresponds to a highly non-linear complex system. The state of a complex system is inherently unpredictable during such a shift, since several possible stable states exist over a certain range of the forcing variable. This discontinuous shift exhibits also hysteresis, i.e. in order to return to a previous state it is not enough to restore the environmental conditions to the level before the shift. Instead, it is necessary to reduce the forcing further until the system eventually switches back to its previous state. This is a crucial difference to an abrupt shift, which is non-linear, but nevertheless predictable, since at all times only one possible ecosystem state exists. The hysteresis has far-reaching significance for ecosystem management: if we looked at the forcing in terms of nutrients, this means that if a discontinuous shift occurred due to eutrophication, nutrient levels would have to be reduced to a lower level than before the shift to restore the original state (Scheffer et al. 2001). In such case, it is even uncertain whether the original state can ever be reached again. Changes in the internal structure, can be so fundamental, that a system never returns to its previous state (Duarte et al. 2009).

Examples of Regime Shifts

Various BRSs were identified during the last century all over the world (e.g. Beaugrand 2004; Drinkwater 2006; Mantua et al. 1997; Peterson and Schwing 2003a; Reid et al. 1998). These were often associated with high losses in fisheries industries, like e.g. after the collapse of the peruvian sardine stock sardine at the beginning of 1970s due to a climate regime shift (Alheit and Bakun 2010) or the collapse of the Haddock stock at Georges Bank, a discontinuous shift (Collie et al. 2004), most likely triggered by a high fishing mortality (Fogarty and Murawski 1998).

For the Baltic Sea region, a focus area of this study, past shifts of the marine ecosystem had severe impact on fisheries and tourism industries. Especially the well-documented shift at the end of the 1980s, which occurred on all trophic levels (Casini et al. 2009) and was associated with a shift from cod to clupeids and a decrease in herring growth (Casini et al. 2010), had pronounced

consequences for the Baltic Sea fishery and its management (MacKenzie and Schiedek 2007). The transitions in the ecosystem were a consequence of anthropogenic stress, but climate has also been shown to be an important factor (Dippner 1997a; Möllmann et al. 2003; Skogen et al. 2011; Stenseth et al. 2002; Stenseth and Mysterud 2002).

Finding Regime Shifts

The possibility of discontinuous regime shifts has long been known from theory and modelling (Scheffer et al. 2001), however finding these shifts in "natural world" time series can be challenging. Scheffer and Carpenter (2003) listed a set of criteria that identify biological regime shifts: A necessary condition is a significant step in the time series. Multimodality of the frequency distribution is an affirmative indicator. Further, the two anticipated regimes may be characterized by different relationships to the control factor. Several possibilities exist to check for these characteristics (Collie et al. 2004; Mantua 2004; Rudnick and Davis 2003). These methods were applied to identify BRSs in the time series of macrofauna communities (see Chap. 3.1).

Empirical approaches to identify regime shifts are necessary since the functional relationships in an ecosystem are unknown to a greater extent. Therefore, we need to approach the problem by analysing ecological time series and the skill of predictions using state-of-the-art statistical methods to identify forcing variables, their relationship with the ecosystem and possible break points, where the character of these relationships change or other driving factors become dominant.

1.2 Climate change and impact on ecosystems

Climate change is stressing the homeostatic capabilities of an ecosystem by altering the physical environment. Recent changes in large-scale atmospheric circulation, e.g. the eastward shift of the centres of action of inter-annual variability of the North Atlantic Oscillation (NAO) (Jung et al. 2003), influence the Baltic and North Sea climatology: the summer sea surface temperature (SST) has increased by 1.4.°C in the period 1985-2002 in the Baltic Sea, which is about triple the global warming trend (Lehmann et al. 2002, 2011). The SST

has also for the other seasons as well as annual mean, a positive trend. This is expressed in milder winters and associated significant decrease in ice cover in the Baltic Sea during winter (Omstedt et al. 2004). Changes in storminess (Donat et al. 2011; Jaagus et al. 2008) and temperature could also be found, as well as an increase in westerly winds (Dippner 1997b; Heyen and Dippner 1998; BACC Author Team 2008).

Milder winters can inhibit spring convection, if the surface water temperature does not cool down to the temperature of maximum density. This will on one hand reduce the nutrient supply from deeper layers, but it will also cause earlier and stronger stratification of the water column. This enhances light availability for the plankton in the surface layer and the limited nutrient supply from deeper layers will favour those phytoplankton species that are capable of vertical movement in the water column. Nitrogen fixing, floating cyanobacteria also have an advantage, since they do not depend on nitrate supply while receiving the highest light dosages. Immotile non-floating phytoplankton like diatoms get the short end of the stick, when the stability of the water column reduces vertical turbulence. While diatoms are usually the first to profit from increasing light and temperature increase and plenty nutrients in spring, they now sink due to reduced vertical turbulence and get shaded by the other phytoplankton, that are quickly depleting nutrients in the euphotic zone. This change in species composition might result in reorganization of the food web. Since changing trophic interactions in turn strongly modulate the response of the ecosystem to climate variability (Winder and Sommer 2012), non-linearity of the system can be expected to increase, and reduce predictability.

The occurrence of now unusually high temperatures will continue to rise, as well as the number of storms (IPCC 2012; EEA 2012). These extreme events further alter ecosystems by eventually pushing them over the brink of their resilience and causing them to shift to a new status. Ecosystems are usually resilient towards extreme events and will return to their previous status relatively fast. It can be shown however, that the return to an earlier status is slowing down in ecosystems deprived of their resilience (Carpenter et al. 2011), which is in accordance with the catastrophe theory for dynamical systems (Scheffer and Carpenter 2003; Collie et al. 2004). Thus, extreme events can be a source for abrupt and discontinuous regime shifts, leading to

unpredictability of the ecosystem.

The large-scale atmospheric circulation pattern is also determining the occurrence and strength of major Baltic Sea inflows of saltier and oxygen-rich Atlantic Water from the North Sea through the Kattegat and the outflow of Baltic Sea surface water. Alterations thus affect the salinity and oxygen content especially of the deeper layers of the Baltic Sea. Meier et al. (2012) projected a southward movement of the horohalinicum, i.e. the zone with a salinity of 5-8 and with a minimum in biodiversity (Remane and Schlieper 1971; Vuorinen et al. 2012). A decrease of salinity would further limit the cod reproductive volume so that fishing quotes might have to be reduced in order to sustain the cod stock in the Baltic Sea (Lindegren et al. 2010; Möllmann et al. 2011).

Changes in the trend of atmospheric variables are especially of interest, since they are often followed by non-linear responses of ecosystems that might reduce ecosystem predictability. Multiple changes in the trend in the global mean temperature could be identified in the past 100 years, termed climate regime shifts (CRS) according to Swanson and Tsonis (2009): 1910–1920, 1938–1945, 1976–1981, and 2001/2002. Especially the CRS around 1976 (Graham 1994; Hurrell 1995; Trenberth and Hurrell 1994) and the CRS around 2000 were well documented (Bond et al. 2003; Peterson and Schwing 2003b). Several of these CRS triggered shifts in biological communities, e.g. the shift around 1920 in the North Atlantic, where exceptional warming caused massive migration of several species and strong increase of recruits of commercially important species in many areas (Drinkwater 2006), or 1976 in the North Pacific, where shifts were strong alterations in abundance and distribution were linked to a shift of the Pacific Decadal Oscillation from a negative to a positive phase (Hare and Mantua 2000).

Climate change and the associated temperature increase affects also directly the physiology of organisms and biogeochemical processes. Due to the temperature preferences of the individual species (Wasmund 1994), a change in temperature may lead to shifts in species composition. A changing seasonality is leading to decoupling of phenological processes by inducing mismatch between trophic levels and functional groups (Edwards and Richardson 2004; Sommer et al. 2012).

Higher temperatures also affect the metabolic rates of organisms. Metabolic theory of ecosystems predicts a decrease in biomass if the temperature increases while limiting basal sources are being left unchanged (Brown et al. 2004). Higher metabolic rates also increase the demand in nutrients or other limiting sources per unit biomass. According to this theory, the size of organisms will decrease, leading to lower total biomass while abundance might not change. However, higher temperatures also speed up the recycling processes and the microbial loop which might then lead to higher net primary production (Taucher and Oschlies 2011). These processes are little understood yet and Taucher and Oschlies (2011) argue, that not even the trend of phytoplankton abundance might be predictable.

1.3 Objectives and Outline of this Thesis

Purpose of the presented study is to investigate the influence of climate on the predictability and variability of marine ecosystems in the North and Baltic Sea. Climate certainly has strong influence on ecosystem variability, and climate change puts the ecosystem into a non-steady state. It is therefore hypothesized that the large-scale atmospheric transitions from climate change will increase the degree of non-linearity in the ecosystem, leading to reduced predictability or vice versa. It is further hypothesized, that a reduced predictability can be partly compensated by choosing a better representation of the climatic forcing.

The degree of ecosystem predictability is estimated relating various large-scale hemispheric climate indices as well as multivariate regional-scale climate indices to various ecosystem descriptors. Biological and climate regime shifts are identified and studied using well studied time series of macro-zoobenthos of the southern North Sea as a descriptor for the ecosystem (Dippner et al. 2010) (Chap. 3.1.1). The predictability despite the presence of regime shifts is estimated for these time series. The well known regime shift in 1989/1990 is identified as a smooth shift, while another biological regime shift in 2001/2002 is identified as a discontinuous shift, triggered by a shift in climate. This finding is the motivation for tackling the resulting increase in non-linearity using state-of-the-art neural network methods, "Optimally Pruned Extreme Learning Machines" (OPELM) and "Optimally Pruned k-Nearest Neighbours" (OP-

KNN). Depending on the type of regime shift, the skill of prediction should increase using these non-linear and non-parametric methods. However, experiments showed, that a significant increase in prediction skill could only be achieved using a multivariate descriptor of climatic forcing. This study is documented in Chap. 3.1.

As a consequence of the increased non-linearity caused by climate, predictability of Baltic Sea ecosystem time series also decreases. Dippner et al. (2000) demonstrated that a major amount of the inter-annual variability in zooplankton time series can be explained using anomalies of the sea surface temperature of the COADS data set for the period 1960-1992. This relationship fails for the updated time series of the same zooplankton species. Also here, the development of a new index as a better predictor for the Baltic Sea ecosystem is chosen to enhance predictions. The performance of the new index, called the Baltic Sea Environmental (BSE) Index, is tested using various physical response variables and the updated zooplankton time series. This is presented in Chap. 3.2.

Encouraged by the performance of the BSE, phytoplankton predictability is examined for the Arkona Sea and in the German part of the coastal zone of the southern Baltic Sea, using the BSE as predictor. These results are shown in Chap. 3.3.

Chapter 2

Material and Methods

2.1 Data Basis

In the subsequently described experiments, time series of benthos, zooplankton and phytoplankton are used as descriptors for ecosystems. They are included as response variables in the downscaling experiments. Predictors are physical parameters and climate indices as well as nutrients. To test the performance of a new index, also physical time series were included as response variables, i.e. Landsort gauge, ice extent in the Baltic Sea and sea surface temperature in the Gotland Sea (GS). The used time series are described below.

2.1.1 North Sea Benthos

The well studied time series of North Sea benthic macrofauna was used to study the feasibility of neural networks to overcome the reduced predictability associated with a regime shift around 2001/2002 as published in Junker et al. (2012). Macrofaunal samples were collected in the 2nd quarter of the years 1978 to 2005 in the sublittoral zone off the island of Norderney, one of the East Frisian barrier islands, at five different stations located in water depths of between 12 m and 20 m (Fig. 2.1). A 0.2 m² van Veen grab was used for sampling. A single grab was taken at each of the five stations. The samples were sieved over 0.63 mm mesh size and fixed in 4% buffered formaldehyde. After sorting, the organisms were preserved in 70% alcohol. Biomass was determined as Ash-free dry weight (AFDW) per m². Samples were dried for 24 h at 85 °C and burned for 6 h at 485 °C. Species number, abundance, and

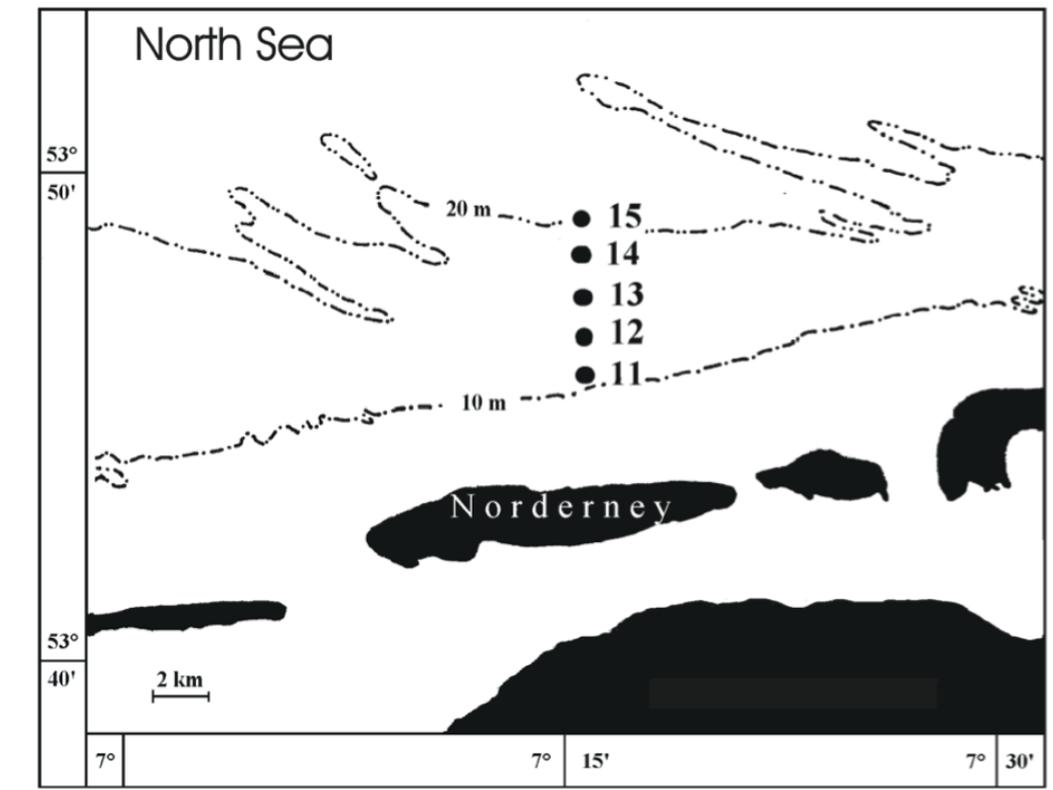


Figure 2.1: Study area off the island of Norderney with stations sampled in spring from 1978 to 2005 (Junker et al. 2012)

biomass from the five stations were pooled and treated as replicates for the area, since the multivariate comparison had shown no significant difference between the macrofauna assemblages (Kröncke et al. 1998).

2.1.2 Baltic Sea Zooplankton

Zooplankton time series from the central Baltic Sea (ICES subdivision (SD)28-2, Fig 2.2) were used to study the feasibility of a new index to enhance prediction in presence of a regime shift as published in Dippner et al. (2012).

The samples were collected seasonally (May, August, October) with Juday net, the opening diameter of 0.36 m and the diameter of the middle section 0.5 m, mesh size 0.16 mm. The hauls were carried out vertically from the depth of 100 m or from the bottom at the stations shallower than 100 m and from different depth layers: mainly 0-25, 25-50 and 50-100 m. Taking into account some sampling deeper than 100 m, it could be considered that mesozooplankton is distributed till the depth of 100 m and only some smaller

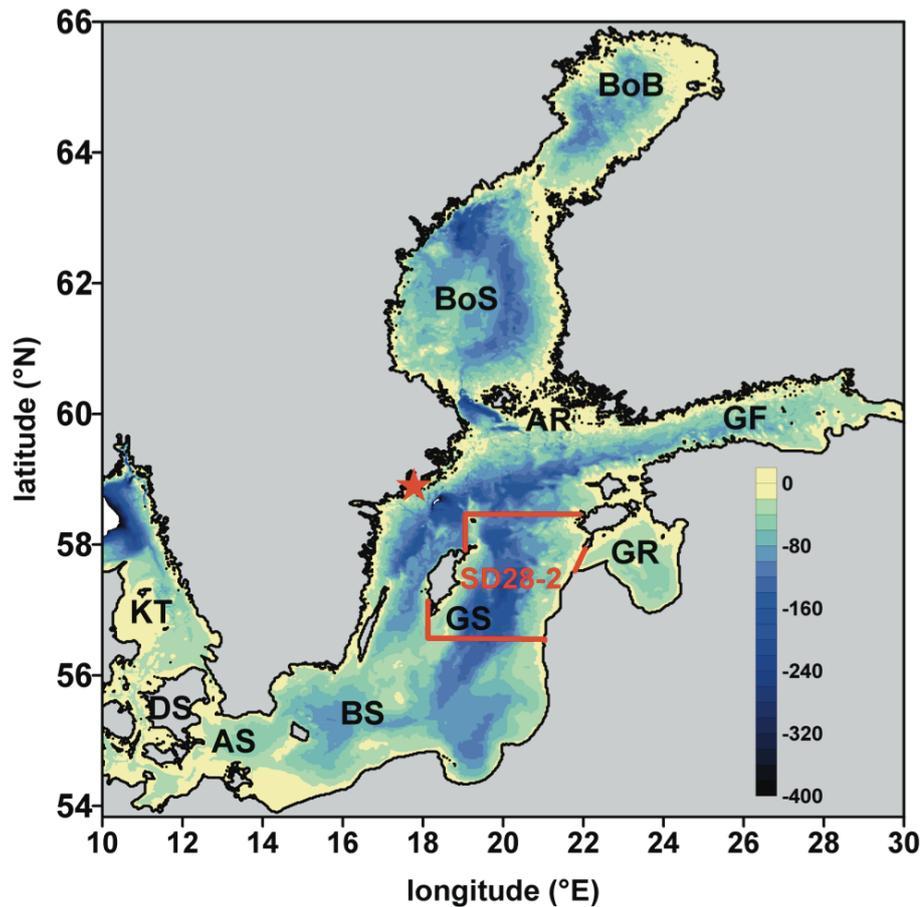


Figure 2.2: Depth distribution of the Baltic Sea and its basins. Bothnian Bay (BoB), Bothnian Sea (BoS), Archipelago Sea (AR), Gulf of Finland (GF), Gulf of Riga (GR), Gotland Sea (GS), Bornholm Sea (BS), Arkona Sea (AS), Danish Sounds (DS), Kattegat (KT). The red asterisk marks the position of Landsort gauge. The red lines mark the area of ICES subdivision 28-2. (Dippner et al. 2012)

part of adult *Pseudocalanus elongatus* can be met deeper. The sampling is conducted throughout the daytime. Zooplankton samples are preserved in formaldehyde and later treated in the laboratory under binocular microscope. All zooplankton specimens are determined as to species or in some cases to genus, and, for Copepoda species, seven stages nauplii, five copepodite stages and adults are distinguished. The biomass per cubic metre is calculated assuming that the filtered volume was 1 m³ per 10 m of the water layer and the filtering coefficient was regarded as 1 (UNESCO 1968). Biomasses are estimated from values of individual wet weights (Hernroth 1985). Averages are calculated for the shallow and deep (100 m) stations separately.

The adult stages of all zooplankton species with sufficient abundance are considered for the experiment. These are *Acartia* spp., *Bosmina longispina*, *Evadne nordmanni*, *Pseudocalanus* spp., *Syncheata* spp. and *Temora longicornis*. These data have been used in a previous article (Dippner et al. 2000) to show the climate-driven variability in a downscaling experiment using Comprehensive Ocean-Atmosphere Data Set (COADS) sea surface temperature (SST) as climate predictor. The time series cover the period 1960-1997 and were chosen to test the performance of the Baltic Sea Environmental index (BSE) in comparison to the results of the previous study by Dippner et al. (2000). The average biomass and abundance of the adult and copepodite stages of the three major species (*Acartia* spp., *T. longicornis* and *Pseudocalanus* spp.) of the depth interval 0-100 m for the deeper stations, and 0 m - bottom of the shallower stations were selected from the 1960-2008 monitoring data.

2.1.3 Baltic Sea Phytoplankton

Phytoplankton data of the monitoring program of the Landesamt für Umwelt, Naturschutz und Geologie of Mecklenburg-Western Pomerania (LUNG) (Landesamt für Umwelt, Naturschutz und Geologie Mecklenburg-Vorpommern 2006) is chosen to represent the coastal ecosystem of the southern Baltic Sea. All sampling sites are along the coast of Mecklenburg-Western Pomerania (Fig. 2.3).

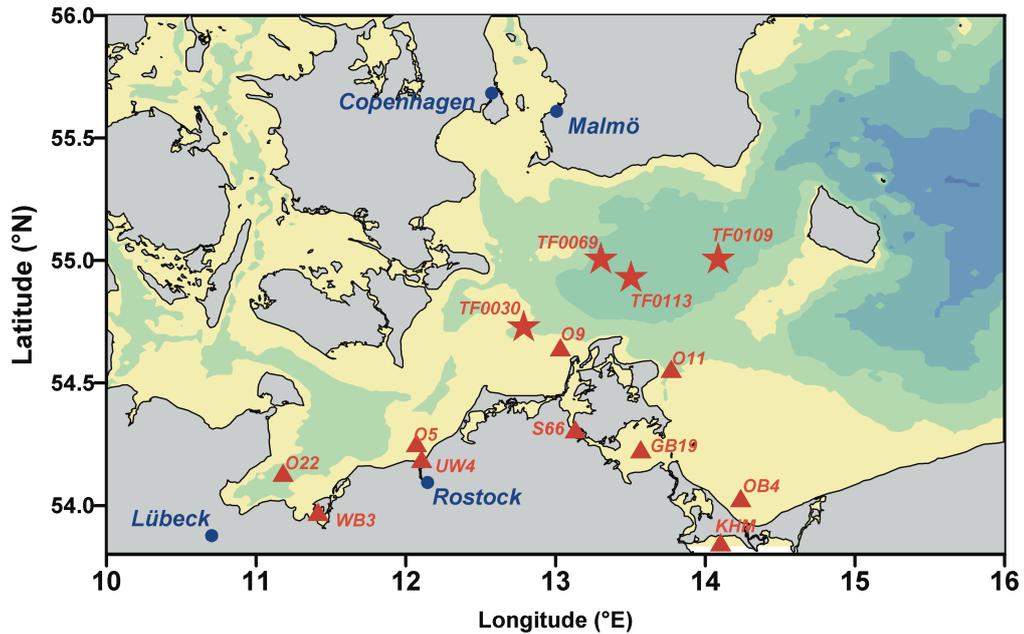


Figure 2.3: Study area with stations: Arkona Sea and coast of Mecklenburg-Western Pomerania. Stations marked with stars are from the IOW data set, stations marked with triangles are from the LUNG data set

The time series for the presently sampled LUNG-stations start in 1986. Older data is available, but longer gaps in the data set and inhomogeneities make it necessary to restrict the study to data after 1986. LUNG monitors 66 coastal stations, however sufficiently long and homogenous time series of phytoplankton abundance and biovolume are available for only 10 stations (Fig. 2.3). The chosen stations are grouped into 4 data sets according to availability and water body types:

- Group A All available stations
(GB19, KHM, O11, O22, O5, O9, OB4, S66, UW4, WB3)
- Group B Only mesohaline stations
(GB19, O11, O22, O5, O9, OB4, UW4, WB3)
- Group C Only stations of the outer coastal zone
(O11, O22, O5, O9)
- Group D Only stations from the inner coastal zone
(GB19, KHM, OB4, S66, UW4, WB3)

Data from LUNG monitoring is referred to as "LUNG data" in the following.

Additionally to the LUNG data, phytoplankton abundance and biomass from the Arkona Sea is chosen to represent the pelagic ecosystem. The data has been extracted from the oceanographic database (IOWDB) of the Leibniz Institute for Baltic Sea Research Warnemünde (IOW), which holds all data collected by the IOW as part of the Helsinki Commission (HELCOM) monitoring program. Only the most regularly sampled non-coastal stations are considered: TF0109, TF0113, TF0030, TF0069. The earliest measurements are from 1979, the latest from 2010. This data set is referred to as "IOW data" in the following.

The available data of both sources covers physical parameters, nutrients and phytoplankton on the species level as abundance, biovolume (for the LUNG data) resp. biomass (for the IOW data). The measurements were performed according to the HELCOM (2008) COMBINE manual. The phytoplankton data was sampled with a Niskin bottle sampler, fixed in formaldehyde and counted according to Utermöhl (1958). Only surface samples (0.5 m depth) are considered. Phytoplankton biovolume was calculated by approximating the cell shape to simple geometrical solids according to Rott (1981). As far as possible, taxonomic resolution was performed on the species level. In order to ensure a consistent database, all samplings with uncertain quantitative values or uncertain taxonomic determinations were eliminated (Sagert et al. 2008).

The autotrophic phytoplankton species are grouped into functional and taxonomic groups. Mixotrophic and heterotrophic species are not considered and excluded from the data set. Seasonal average abundance [ind/ml] and biovolume [ml/m³] resp. biomass [mg/m³] is calculated for each of the following heterotrophic phytoplankton classes: Bacillariophyceae, Chlorophyceae, Chrysophyceae, Craspedophyceae, Cryptophyceae, Dinophyceae, Euglenophyceae, Prasinophyceae, Prymnesiophyceae (Haptophyceae) and Cyanobacteria. Besides these groups are the sums of total phytoplankton and flagellates considered as well as the ratio of diatoms:flagellates to see whether a shift in this ratio is associated with climate.

The seasonal data is normalized by log-transformation and standardized by removing the seasonal signal and calculating the standard scores (z-scores: $z = \frac{X - \bar{X}}{\sigma}$ with X a random variable, \bar{X} the expected value and the standard deviation σ .) Pre-runs of downscaling experiments showed no significant dif-

ference whether means or medians are used to calculate the seasonal data. In the follow up experiments are means chosen.

Data exploration showed that some phytoplankton groups are most of the time of the considered time span absent. As the method used in the down-scaling experiments is not suitable for this type of time series we concentrate on the major groups and consider only those groups that are most of the time (>50%) present. Applying this rule only phytoplankton data from spring (MAM), summer (JJA) and fall (SON) remains. Some of the less common groups also drop out for some seasons.

2.1.4 Nutrients

In addition to the phytoplankton data of the LUNG and HELCOM monitoring program, the following nutrients and derived properties are chosen as predictors.

- Local Phosphate concentrations ($PO_4[mmolP/m^3]$)
- Local total Phosphorous concentrations ($totP[mmolP/m^3]$)
- Local Nitrate concentrations ($NO_3[mmolN/m^3]$)
- Local Nitrite concentrations ($NO_2[mmolN/m^3]$)
- Local Ammonium concentrations ($NH_4[mmolN/m^3]$)
- Local total Nitrogen concentrations ($totN[mmolN/m^3]$)
- Local Silicate concentrations ($SiO_4[mmolSi/m^3]$)
- Local molar ratio of N:P derived from NO_3 and PO_4
- Local molar ratio of N:P derived from total N and total P

2.1.5 Climate Predictors and Physical Variables

Predictors

The following data sets are used in the subsequently described experiments as predictors. The large scale climate indices are used in all experiments, whereas the regional scale Baltic Sea Index (BSI) and Chen index (Chen 2000) are only used in experiments involving time series from the Baltic Sea. All indices are considered as monthly mean values.

The considered climate indices are:

- The Arctic Oscillation (AO) index from 1899 to 2007 (Thompson and Wallace 1998) which describes the leading Empirical Orthogonal Function (EOF) of monthly geopotential height anomalies at the 1000 hPa level on the Northern Hemisphere poleward from 20°N.
- The North Atlantic Oscillation (NAO) index from 1864 to 2009 (Hurrell 1995) defined as the difference between the normalized monthly sea level pressure (SLP) anomalies at Lisbon and Stykkisholmur.
- The Atlantic Multidecadal Oscillation (AMO) index from 1856 to 2009 (Enfield et al. 2001), defined as the monthly SST anomalies in the North Atlantic area weighted from 0° to 70°N.
- The extended (1948-2009) BSI (Lehmann et al. 2002) which is created by combining data from the Swedish Meteorological and Hydrological Institute (SMHI) and data from the National Centre of Environmental Predictions (NCEP) (Kalnay et al. 1996). There are no inhomogeneities between these two data sets.
- Monthly mean values of the updated (1780-2010) Chen index . The Chen index consists of three time series: the components of the geostrophic wind and the relative vorticity of the geostrophic wind over the area 50°-70°N and 0°-30°E.

Time series of physical properties from the Baltic Sea area:

- Monthly mean SST fields derived from the COADS on a $2 \times 2^\circ$ grid for the period 1900-1992 (Woodruff et al. 1987). In addition, an updated version of the COADS for the period 1960-2010 is also used.
(<http://icoads.noaa.gov/data.icoads.html>)
- Monthly mean salinity on oceanographic standard depth has been compiled on $1 \times 1^\circ$ grid for the whole Baltic Sea (Feistel et al. 2008). The data from 1900 to 2005 between 120 and 200 m in the area of the central Gotland Sea 16°-22°E and 55°-59°N (Fig. 3.10) is averaged to create a salinity time series.
- Monthly mean runoff data for the whole Baltic Sea area for the period 1921-1993 has been compiled by Bergström and Carlsson (1994). The data set for river runoff consists of the data from Mikulski (1982) for the period 1921-1949, and the data compiled by the SMHI for the period

1950-1993 and extended later up to 2002. There are no inhomogeneities between these two data sets.

- Monthly mean SST from 1902 to 2005 in the Gotland Sea area averaged between 16°-22°E and 55°-59°N (Feistel et al. 2008).
- Local (station data) temperature anomalies.
- Local (station data) salinity anomalies.

For the experiments with benthos the following time series from the North Sea are considered:

- SLP anomalies poleward of 30°N on the Northern Hemisphere from NCEP/NCAR reanalysis (Kalnay et al. 1996).
- Area averaged monthly meridional wind anomalies (1948-2010) in the southern North Sea (53°-56°N, 2°-9°E) from NCEP/NCAR reanalysis.
- Monthly precipitation rate anomalies (1948-2010) averaged over the area 50°-57°N, 4°W-9°E from NCEP/NCAR reanalysis.
- Area averaged monthly SST anomalies (1948-2009) in the southern North Sea (53°-56°N, 2°W-9°E) from NCEP/NCAR reanalysis (Kalnay et al. 1996).
- Salinity from International Council for the Exploration of the Sea (ICES), Marsden square 96668.
- Temperature from ICES, Marsden square 96668.
- Weekly SST data for the German Bight from 1968 to 2007 south of 55.5°N and east of 6.5°E from ship-of-opportunity programs, commercial vessels, light vessels, fixed stations and buoys, coastal stations, research vessels and monitoring programs provided by the Federal Maritime and Hydrographic Agency German Federal Maritime and Hydrographic Agency (BSH) Hamburg (Becker et al. 1986).

Physical Response Variables

The following data sets are used in the subsequently described experiments as response variables:

- Monthly mean values from Landsort gauge (Fig. 2.2) for the period 1897-2002 from SMHI;

- Annual average sea ice extent for the period 1720-2006 (Feistel et al. 2008).

2.2 Methods

Dippner and Kröncke (2003) have shown, that the partial linearity in the ecosystem response allows the analysis and prediction of biodiversity, structure, function and dynamics of ecosystems using linear methods. Thus, it can be expected that methods like the analysis of empirical orthogonal functions (EOF) (Preisendorfer and Mobley 1988), statistical downscaling (von Storch et al. 1993) and the analysis of principal oscillation patterns (Hasselmann 1988; von Storch et al. 1988; Von Storch and Zwiers 2001) are feasible to address the previously stated hypotheses and to study the variability of the ecosystems and to predict its future state. However, downscaling experiments showed that the considered time series are unsuitable for the analysis of principal oscillation patterns. Due to high variability and gappyness yielded the analysis of principal oscillation patterns of major phytoplankton groups in the Baltic Sea in first runs no meaningful results.

A common tool to study the dominant modes in a time series and their variability in time is also the wavelet analysis (Torrence and Compo 1998). The method and the motivation for using it is discussed in Section 2.2.2. The methods described in the following will be used on climatological and ecological time series with the seasonal cycle removed, i.e. anomalies.

2.2.1 Analysis of Empirical Orthogonal Functions (EOF)

EOF analysis is a multivariate eigen-technique used to reduce the dimensionality of a dataset and thereby reducing the amount of noise. The EOF analysis seeks structures that explain the maximum amount of variance in a dataset. The idea behind it is to find patterns and coefficients (i.e. principal components), that describe the variability of a data set in such a way, that the first pattern describes most variability, the second pattern second most, etc.. Since most of the variability is usually represented by the first few EOFs the amount of data can be vastly reduced using only these few EOFs. The other EOFs

usually only contain the noise or small scale variability (sometimes also called station characteristics) and can thus be cut off.

The EOF analysis is done in such a way that the anomalies $\vec{G}'(r, t)$ of a data set $\vec{G}(r, t)$ with $t = 1..T$ observations and $r = 1..R$ stations are expanded to a series of K EOFs:

$$\vec{G}'(r, t) = \sum_{i=1}^K \vec{\Gamma}_i^{EOF}(r) \gamma_i^{EOF}(t) + \sum_{i=K+1}^R \vec{\Gamma}_i^{EOF}(r) \gamma_i^{EOF}(t) \quad (2.1)$$

where $\vec{\Gamma}_i^{EOF}(r)$ are the EOF patterns and $\gamma_i^{EOF}(t)$ the time coefficients.

To calculate the EOF it is common to do an eigenvalue-analysis on the covariance matrix $\mathbf{C} = \mathbf{G}\mathbf{G}^T$ where \mathbf{G}^T is the transposed matrix G . Any symmetric matrix, like the covariance matrix, can be decomposed using an eigenanalysis

$$\mathbf{C}\mathbf{E} = \mathbf{\Lambda}\mathbf{E} \quad (2.2)$$

where \mathbf{E} is the matrix with the eigenvectors \mathbf{e}_i and $\mathbf{\Lambda}$ is the matrix with the eigenvalues λ_i along its diagonal and zeros elsewhere. The EOFs $\vec{\Gamma}_i^{EOF}(r)$ are now the eigenvectors \mathbf{e}_i and λ_i describe the amount of explained variance of each eigenvector. It is important to note that the eigenvectors are orthogonal to each other and therefore independent.

The evolution of the system in time can now be described by the EOF (or time) coefficients, which can be estimated in the following way: Considering that the original matrix \mathbf{G} can be reproduced by the dot product of the EOFs and some matrix \mathbf{Z} with the time coefficients:

$$\mathbf{G} = \mathbf{E}\mathbf{Z} \quad (2.3)$$

we get the time coefficients from

$$\mathbf{Z} = \mathbf{E}^T\mathbf{G} \quad (2.4)$$

Since most of the variance is explained by the leading K EOFs, it is possible to use EOF analysis to filter a dataset by differentiating between a signal space described by the first K EOFs and a noise space spanned by the trailing $R - K$ EOFs. These trailing patterns describe a minor part of the variability and are

usually station details. So, it is usually enough to consider

$$\vec{G}'(r, t) = \sum_{i=1}^K \vec{\Gamma}_i^{EOF}(r) \gamma_i^{EOF}(t) \quad (2.5)$$

2.2.2 Wavelet Analysis

For the analysis of localized variations of power within a time series spectrum, wavelet analysis is a well established method. It has been applied to time series from all kinds of research fields, among them climatological and ecological time series (e.g. Cazelles et al. 2008; Mi et al. 2005). By decomposing a time series into time-frequency space, it is possible to determine both the dominant modes of variability and how those modes vary in time (Torrence and Compo 1998). In contrast, the Fourier transform only yields the information about which major modes occur in the series, but not when. A similar method, the windowed Fourier transform (WFT), also provides a decomposition into time-frequency space, but since the frequency resolution is the same for all frequencies, it is less efficient than the wavelet approach. The WFT tends to overrepresent the higher frequencies and underrepresent the lower frequencies (Lau and Weng 1995). In a wavelet transform (WT), the signal is decomposed using a specific function $\Psi(t)$ (called wavelet) that is scaled according to the frequency in focus (Daubechies 1992; Lau and Weng 1995) (Fig. 2.4).

$$\Psi_{b,a}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (2.6)$$

b is the position (translation) and $a > 0$ is the scale (dilation) of the wavelet $\Psi_{b,a}(t)$. The factor $1/\sqrt{a}$ is a normalization factor that keeps the energy of the daughter wavelet $\Psi_{b,a}(t)$ the same as the mother wavelet $\Psi(t)$.

A time series $x(t)$ is then decomposed in a WT by calculating the wavelet coefficients $W(b, a)$ as follows:

$$W(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \Psi^*\left(\frac{t-b}{a}\right) x(t) dt \quad (2.7)$$

where $*$ denotes the complex conjugate form.

For a discrete time series $x_n, n = 1, \dots, N$ with uniform time steps δt this

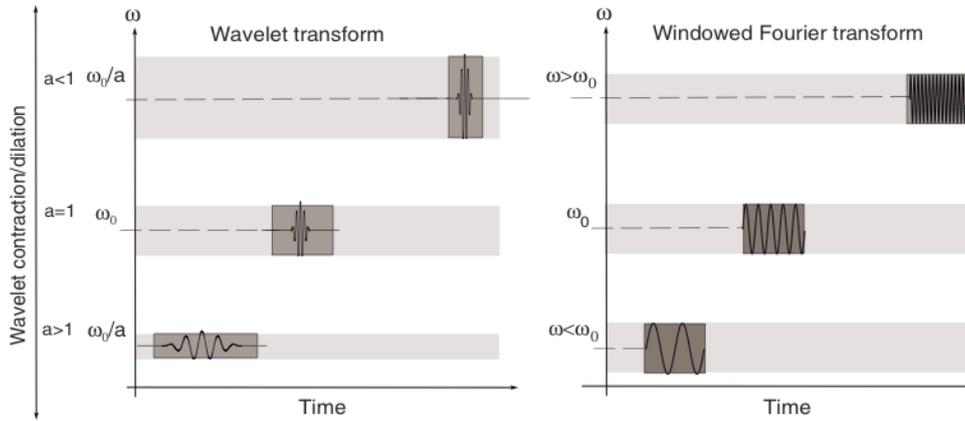


Figure 2.4: WT (left) and WFT (right). For higher frequencies is the wavelet narrower, which gives good time resolution, but the scale is larger and the frequency resolution reduced compared to the lower frequencies. In WFT are the windows always the same size, which yields unsatisfactory time and frequency resolution (Cazelles et al. 2008).

can be rewritten as

$$W_n^X(a) = \frac{\delta t}{a} \sum_{n'=1}^N x_{n'} \Psi_0 \left[(n' - n) \frac{\delta t}{a} \right] \quad (2.8)$$

Due to the localized decomposition of the signal, wavelet transforms do not require stationary time series (Cazelles et al. 2008; Lau and Weng 1995) like other spectral or correlation methods do. This is a major advantage in studying ecological time series, because they often show substantial changes in variance, periodicity and species interactions as well as trends (Stige et al. 2006; Cazelles and Hales 2006). By performing a local time-scale decomposition of the time series wavelet analysis can be used to localize and quantify abrupt changes like regime shifts (Cazelles et al. 2008).

In order to receive qualitatively and quantitatively meaningful results, and to interpret them correctly, it is necessary to consider a number of different aspects, e.g. the type of wavelet transform (WT), wavelet function or base and scale. The choice of these depends on the aim of the study and the type of the signal.

There exist two types of WT: the discrete (DWT) and the Continuous Wavelet Transform (CWT). The Discrete Wavelet Transform (DWT) is ba-

sically a compact representation of the data and useful for denoising of data and data compression. It is often used to characterize spatial patterns and in image processing. However, it may not resolve the physically most important information in a time series, because it does not increase the scales of the wavelet continuously like the CWT but in steps of the power of 2. The CWT on the other hand, is better for feature extraction purposes (Grinsted et al. 2004) because the continuous wavelets provide better timescale localization information. However, the results of a CWT have to be interpreted with this redundancy in mind, because it may cause spurious features. Nevertheless this drawback of CWT, it is in this study considered to give better results because the ecological and climatological time series under investigation are of wavelike structure and the emphasis of this study lies in finding the prominent features in them. Two proven wavelets in climatological and ecological studies are the Morlet and the Mexican Hat (Mi et al. 2005). In this study we will focus on the former, because it provides a good balance in time and frequency localization. Another advantage of the Morlet is that the relationship between frequency and scale can be derived analytically.

The Morlet is defined as

$$\Psi(\eta) = \pi^{-1/4} \left(e^{i\omega_0\eta} - e^{-\frac{\omega_0^2}{2}} \right) e^{-\frac{\eta^2}{2}} \quad (2.9)$$

with ω_0 the nondimensional frequency and η a nondimensional time factor. For a Morlet usually $\omega_0 > 5$, so the correction term $e^{-\frac{\omega_0^2}{2}}$ can be neglected and it is enough to say

$$\Psi(t) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\frac{\eta^2}{2}} \quad (2.10)$$

The Morlet wavlet is a plane wave modified by a Gaussian envelope (Lau and Weng 1995). Compared with other non-orthogonal wavelet functions, (e.g. Mexican Hat), the Morlet has the advantage of providing a good frequency resolution and localization of scale. However, this is at the cost of lower time resolution. The Mexican Hat provides a better detection and localization of single events. Since the Morlet is a complex wavelet, it will also return information about both amplitude and phase and is thus better adapted for capturing oscillatory behavior. A more comprehensive discussion about the properties of the different wavelet functions can be found in (Torrence and Compo 1998).

Wavelets are mostly used on univariate time series. However, the wavelet approach is also feasible for the analysis of bi-variate time series. With the help of wavelet cross-spectra and wavelet coherency can relationships between time series be identified (Grinsted et al. 2004; Jevrejeva et al. 2003). The wavelet cross-spectra can be produced by performing a cross wavelet transform (XWT) of two time series x_n and y_n (Torrence and Compo 1998). It is defined as $W^{XY} = W^X W^{Y*}$, where W^{Y*} denotes the complex conjugate form of W^Y . The cross wavelet power is $|W^{XY}|$. Peaks in the cross wavelet transform indicate high power in both wavelet transforms. However, it lacks information about the phase relationship. This information can be derived from wavelet coherence. It is similar to the Fourier squared coherence, where the frequency bands are identified within which the two time series covary (Torrence and Webster 1999).

Torrence and Webster (1999) defined wavelet coherence as

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{XY}(s))|^2}{S(s^{-1}|W_n^X(s)|^2)S(s^{-1}|W_n^Y(s)|^2)} \quad (2.11)$$

S is a smoothing operator that applies to the cross-spectrum (in the numerator) and the wavelet power spectra (in the denominator) of the two time series. The resemblance of the definition of wavelet coherence with the traditional correlation coefficient is striking. In fact, the wavelet coherence can be thought as a localized correlation coefficient in frequency space (Grinsted et al. 2004).

In this dissertation will only wavelet coherences be shown. All wavelet coherences are calculated using the toolbox of A. Grinsted.

2.2.3 Statistical Downscaling

In statistical downscaling, a regional scale response variable is estimated from one, or a set of, large scale predictors by finding statistical relationships between the two and using these relationships for prediction. This statistical modelling of spatial data can be used to downscale to a finer grid and for prediction under the assumption that the large scale predictor is sufficiently accurate. It has been termed in climatology where it is mostly used to downscale climatological data from global scale data, e.g. the output of a global climate model, to regional scale.

There exist several methods for statistical downscaling. A comparison of methods can be found in Wilby and Wigley (1997), Zorita and von Storch (1999) and Huth (1999). Statistical downscaling can be based on linear methods like EOF and canonical correlation, or non-linear methods like clustering, self organizing maps or neural-networks (Zorita 2009). Being aware that some ecosystem interactions are non-linear the latter methods are seemingly more appropriate. Furthermore, Neural Networks have the advantage of being non-parametric and thus do not require normally distributed data. However, non-linear methods also have major drawbacks: overfitting and misinterpretation is even more a problem than in linear downscaling and the demand on computing power and number of data points is higher. Since the available time series are fairly short and the Baltic Sea and North Sea ecosystem has proven to react linearly (or that it can be linearised by careful data transformation) in larger parts (Kröncke et al. 1998; Dippner et al. 2000), a linear method is thought to be sufficient for the initial experiments.

In this study, the term statistical downscaling refers to the method of von Storch et al. (1993), which has been proven to be also suitable for estimating biological data from climatological data. Applying this method Kröncke et al. (1998) and Dippner and Kröncke (2003) found strong relationships between macrozoobenthos and climate variability. Dippner and Ottersen (2001) found high correlation between cod recruitment and large scale climate forcing through sea surface temperature. Dippner et al. (2001) showed that the winter anomalies in abundance of *Acartia* spp. and *Eurytemora affinis* are driven by climate variability, but not the larger zooplankton species.

In principle, the downscaling method used here consists of applying EOF analysis to extract major patterns and subsequently applying a Canonical Correlation Analysis (CCA). As EOF is used to study the variability, the CCA is used to study the correlation pattern of data. The basic concept of the downscaling method is shown in Figure 2.5.

As mentioned, the noise will be eliminated by selecting only the major EOF of the datasets of the predictors and predictands. For the predictor $\vec{G}(r, t)$, the EOF-subspace of the anomalies is $\vec{G}'(r, t)$ (see section 2.2.1) (Dippner et al.

2001):

$$\vec{G}'_S(r, t) = \sum_{i=1}^K \vec{\Gamma}_i^{EOF}(r) \gamma_i^{EOF}(t) \quad (2.12)$$

and for the anomalies of the response variable $\vec{L}'(s, t)$ with $s = 1, \dots, S; t = 1, \dots, T_2$, S are the stations, the EOF-subspace is accordingly:

$$\vec{L}'_S(s, t) = \sum_{i=1}^N \vec{\Lambda}_i^{EOF}(s) \lambda_i^{EOF}(t) \quad (2.13)$$

where $\vec{\Lambda}_i^{EOF}(s)$ is the i th EOF pattern fixed in space and $\lambda_i^{EOF}(t)$ are the corresponding time coefficients. Note that $N \leq S$ and usually $N \ll S$. After that, a CCA will be performed on the resulting leading eigenmodes of the predictors and predictands. This is done by first combining the time coefficients $\gamma_i^{EOF}(t)$ and $\lambda_i^{EOF}(t)$ into new CCA time coefficient $\gamma_j^{CCA}(t)$ and $\lambda_j^{CCA}(t)$:

$$\gamma_j^{CCA}(t) = \sum_{i=1}^I g_{ij} \gamma_i^{EOF}(t) \quad (2.14)$$

and

$$\lambda_j^{CCA}(t) = \sum_{i=1}^I l_{ij} \lambda_i^{EOF}(t) \quad (2.15)$$

where the weights g_{ij} and l_{ij} are selected such that

$$\langle \gamma_i^{CCA}(t), \lambda_j^{CCA}(t) \rangle = \rho_j * \delta_{ij} \quad (2.16)$$

ρ_j is at maximum for $j = 1$. For $j > 1$, ρ_j is maximum under the constraint that the CCA time coefficients of the global and the local parameters are orthogonal:

$$\langle \lambda_i^{CCA}(t), \lambda_j^{CCA}(t) \rangle = \langle \gamma_i^{CCA}(t), \gamma_j^{CCA}(t) \rangle = \delta_{ij} \quad (2.17)$$

The CCA patterns are then:

$$\vec{G}'_{SS}(r, t) = \sum_{i=1}^I \vec{\Gamma}_i^{CCA}(r) \gamma_i^{CCA}(t) \quad (2.18)$$

and

$$\vec{L}'_{SS}(s, t) = \sum_{i=1}^I \vec{\Lambda}_i^{CCA}(s) \lambda_i^{CCA}(t) \quad (2.19)$$

Note the similarity to the construction of the EOF subspace (eq. 2.12 and eq. 2.13). Here however: $[G'_{SS}(r, t) \subseteq G'_S(r, t) \subseteq G'(r, t)]$ and $[L'_{SS}(s, t) \subseteq L'_S(s, t) \subseteq L'(s, t)]$.

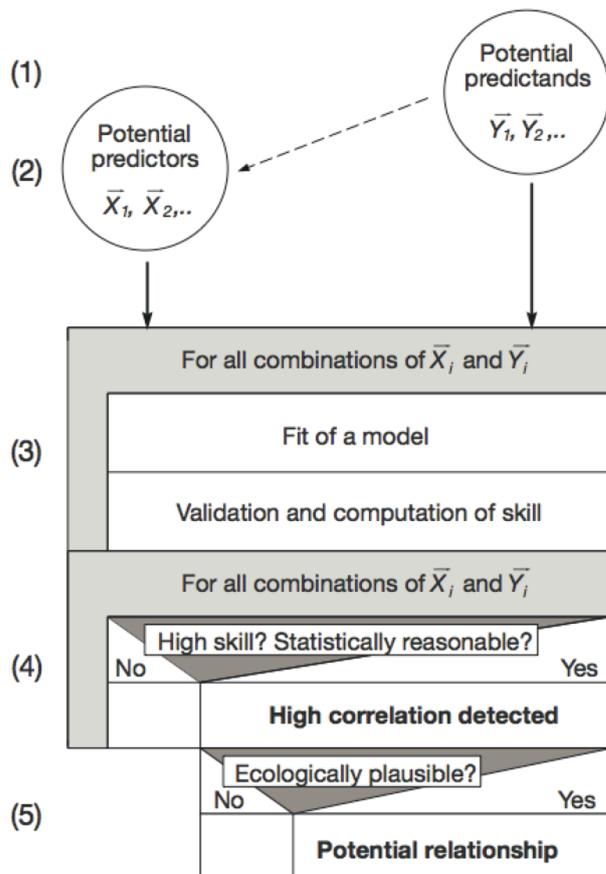


Figure 2.5: Outline of the statistical downscaling method after Kröncke et al. (1998)

The results of the CCA are then validated and the skill calculated. If a high skill has been found, it will have to be decided whether the results are ecologically plausible (Kröncke et al. 1998). CCA has the tendency to over fit peaks (Heyen et al. 1996). Filtering of time series can mitigate this problem. For the new Baltic Sea Environmental (BSE) index (Chap. 3.2) a moving average filter with 1-2-1-filter weights is applied (Dippner et al. 2012).

Validation and Skill

The validation of the found correlations has to be done with independent data. von Storch et al. (1993) split the data set into fit and validation periods. However, this is inconvenient if the time series is short. Thus a cross-validation technique (Michaelsen 1987) is used here. In cross-validation, if n time steps of data are available, n models are fitted by using $n - 1$ different time steps each. For each model, the n th step of the predictand is regressed from the predictor. Finally, the n estimations are compared with the observations of the response variable. The significance level of selected results is calculated with the Monte Carlo technique, with 10000 generated series of random numbers with the same statistical properties (mean, standard deviation and autocorrelation) as the EOF coefficients of the predictands. As skill factors, correlation coefficient r between the regional observations and the cross-validated estimations and Brier-based score β are used. The Brier-based score skill is defined as: $\beta = 1 - \frac{\sigma_E^2}{\sigma_O^2}$ where σ_E^2 and σ_O^2 are the variances of the error (i.e. observation minus model) and observations. $\beta = 1$ means that model and observation are identical, $\beta = 0$ that the error of the model has the same size as the variance of the observations (Livezey 1995). The Brier-based score is used throughout this work to estimate the skill of the fitted models, also the neural nets.

2.2.4 Neural Networks

Linear methods as the above downscaling method are only suitable if linear relationships between predictors and predictands exist. It is often possible to linearise the system by carefully transforming the available data, but sometimes it is impossible due to higher degrees of non-linearity in the system. Therefore, it is advisable to use non-linear methods for prediction in such cases. Neural networks are proven methods for prediction of environmental properties (e.g. brown trout reproduction (Lek et al. 1996) or river runoff (Minns and Hall 1996)). Numerous learning algorithms exist, of which supervised learning algorithms are considered the most appropriate for prediction of environmental time series and thus used most for prediction of ecological properties (Lek and Guégan 1999). Of the many possible learning algorithms two methods, the Optimally Pruned Extreme Learning Machine (OPELM) and

the Optimally Pruned K-Nearest Neighbours (OPKNN) were chosen. The following descriptions of the OPELM and OPKNN are cited from (Junker et al. 2012).

OPELM Method

The OPELM method belongs to the single layer neural network methods (Miche et al. 2008) and is based on the Extreme Learning Machine (ELM) algorithm from which it borrows the original single layer feed-forward network construction (Huang et al. 2006). The OPELM (Fig. 2.6) algorithm is introduced as a more robust method to tackle the problem of irrelevant neurons and to be more adaptive to both linear and non-linear problems. The OPELM algorithm consists of three steps:

1. Construction of the single hidden layer
2. Ranking the neurons in the hidden layer
3. Selecting the appropriate number of top ranked neurons

Ranking in step 2 is performed with least angle regression (LARS) algorithm (Efron et al. 2004), an algorithm used to rank variables in regression problems in a linear setting. LARS provides exact ranking when the problem is linear, which is the case in ELM between hidden layer and output variable. Once the ranking is obtained, the selection is done with Leave-one-out (LOO) estimation based on the outputs of hidden neuron. To demonstrate how the neurons are selected, Fig. 2.7 shows for the sake of clarity the LOO error versus the number of neurons for a run in which the AO index is projected on macrofauna biomass in the southern North Sea (see chapter 3.1). The upper limit of the number of neurons is given by the number of data pairs itself. The optimal number of neurons is estimated by searching the local minimum of the LOO error, which appears at 12 neurons in this case (Fig. 2.7). Finally, this number of neurons is used for the simulation.

The three steps above are the core of the original OPELM algorithm. However, when using data sets with few samples, the LOO estimate can have high variance and can lead to complex models and over-fitting. Instead of LOO, other model selection criteria can be employed, such as information theoretic criteria. In the experiments, we resort to the corrected Akaike information cri-

terion (Hurvich and Tsai 1989; Sugiura 1978). The advantage of ELM based models over other types of neural networks is their very quick training time, and at the same time they have comparable generalization capabilities.

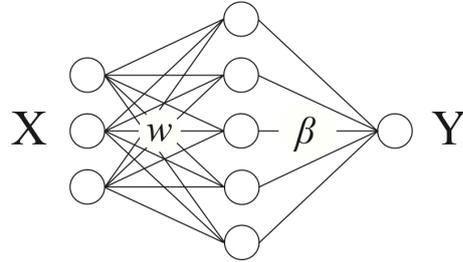


Figure 2.6: General structure of the OPELM model: X are the climate input data, Y the benthic output data, w the input weights and β the output weights.

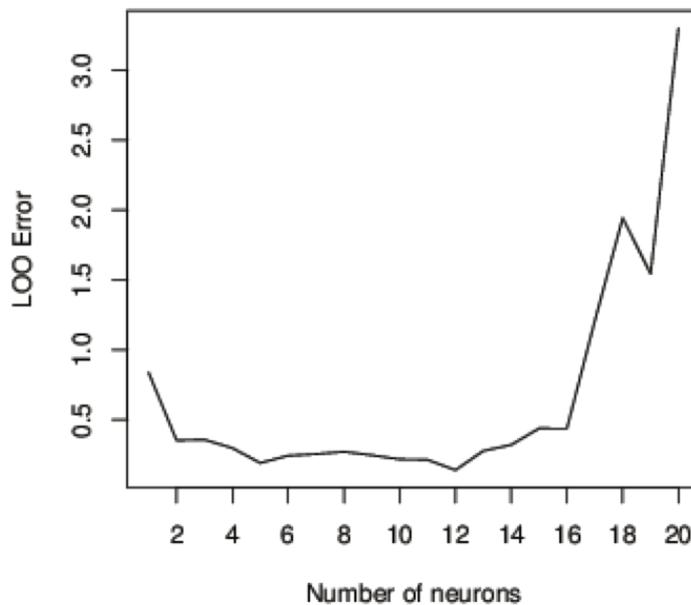


Figure 2.7: LOO error versus the numbers of neurons. The local minimum gives the optimal number of neurons.

ELM construction phase The main concept behind the ELM lies in the random initialization of the hidden layer input weights and biases. According to Huang et al. (2006), the input weights and biases do not need to be adjusted and it is possible to calculate implicitly the hidden layer output matrix and

hence the output weights. The network is obtained with very few steps and very low computational cost. Consider a set of M distinct samples (x_i, y_i) with $x_i \in \mathfrak{R}$ and $y_i \in \mathfrak{R}$, where d is the number of input features. For each sample $x_j, j \in \{1, \dots, M\}$, the model produces the prediction \hat{y}_j based on the sum

$$\sum_i^N \beta_i f(w_i x_j + b_i) = \hat{y}_j \quad (2.20)$$

With f being the activation function, w_i the input weights, b_i the biases and β_i the output weights. Input weights w_i and biases b_i are randomly generated. In our case, x_i is the climate time series and y_i the benthos time series, respectively. The aim is to match the prediction to the actual outputs y_j which can be written in matrix form $\mathbf{H}\beta = \mathbf{y}$, with

$$\mathbf{H} = \begin{pmatrix} f(w_1 x_1 + b_1) & \cdots & f(w_N x_1 + b_N) \\ \vdots & \ddots & \vdots \\ f(w_1 x_M + b_1) & \cdots & f(w_N x_M + b_N) \end{pmatrix} \quad (2.21)$$

$\beta = (\beta_1, \dots, \beta_N)^T$ and $y = (y_1, \dots, y_N)^T$. The output weights β are computed with the Moore-Penrose generalized inverse of the matrix \mathbf{H} and the target values, i.e. $\beta = \mathbf{H}^+ \mathbf{Y}$. The original paper focuses on sigmoid and sine activation functions, but the kernels of neurons are not limited to these two only. In the OPELM, beside sigmoid activation function, the linear and Gaussian kernels are utilized as well. The linear kernel enables OPELM to adapt to the problems that are highly linear by adequate ranking and selection.

LARS ranking phase As mentioned, LARS provides exact ranking when the problem is linear. This is the case in the basic ELM, where the connection is linear between hidden layer and the output. The output is given from the data y_i , while the "variables" are the outputs of the kernels $h_k = (f(w_k x_1 + b_k), \dots, f(w_k x_m + b_k))^T$ (the columns of \mathbf{H}). With this ranking, we can also assess the importance of different kernels in the model.

Selection phase

The main importance behind any model selection criteria is to estimate how good a model can predict future data. The error during training phase is a poor indicator of the generalization properties of the model. A penalty term has to be added to account for model complexity and number of samples available

in the data. As mentioned above, we use the corrected Akaike information criterion. The Akaike information criterion (AIC) (Akaike 1974) is based on the information theoretic concept of Kullback–Leibler information (Kullback and Leibler 1951) and is expressed with the formula:

$$AIC = M \ln(\hat{\sigma}^2) + 2K \quad (2.22)$$

where $\hat{\sigma}^2$ is the mean square error on the training data and K the number of parameters of the model. In OPELM case, this is the number of adjustable output weights of the hidden layer. In the situation when K is large relative to the number of samples, or when number of samples is quite low, the second-order correction term is added to the AIC formula giving the corrected AIC_c criterion (Sugiura 1978).

$$AIC_c = M \ln(\hat{\sigma}^2) + 2K + \frac{2K(K+1)}{M-K-1} \quad (2.23)$$

The second order term accounts for the finite sample size, and when $M \rightarrow \infty$ both AIC and AIC_c are equal. The goal is to use the model, which minimizes the AIC_c value. For OPELM, the number of neurons giving lowest AIC_c value is chosen as the appropriate complexity of the network.

OPKNN Method

The Optimally Pruned K-Nearest Neighbours (OPKNN) shares a similar approach to the OPELM (Yu et al. 2008). Instead of using random initialization of input weights w and various kernels, OPKNN uses simple k -nearest neighbour (KNN) model as its kernel. The key idea behind KNN is that similar training samples should have similar outputs. The similarity is based on some form of distance metric, and the usual approach is to use the Euclidean metric in the input space. In OPKNN, matrix \mathbf{H} is defined as:

$$\mathbf{H} = \begin{pmatrix} y_{P(1,1)} & \cdots & y_{P(1,N)} \\ \vdots & \ddots & \vdots \\ y_{P(M,1)} & \cdots & y_{P(M,N)} \end{pmatrix} \quad (2.24)$$

Where $P(i, j)$ is the index of the j th nearest neighbour of sample x_i and $y_{P(i,j)}$ is the output of that j th nearest neighbour. An important feature of

OPKNN is that the model is deterministic, as it does not have any stochastic elements. A comparison of these two models in time series domain is given in (Sovilj et al. 2010).

Chapter 3

Experiments

3.1 Shifts in Benthos

The contents of the following section are already published in, and taken from Dippner et al. (2010) and Junker et al. (2012).

3.1.1 Observed shifts in benthos and climate

Atmospheric winter circulation over the North Atlantic area has been proven to be an optimal predictor to forecast the structure of the macrofauna communities in the following spring in the southern North Sea (Kröncke et al. 1998). The mediator between climate and benthic macrofauna is SST which is highly correlated to the NAO index (Becker and Pauly 1996). Based on those relationships, (Dippner and Kröncke 2003) developed forecast equations for the prediction of macrofauna community structure in spring from the climate during the winter before. Despite presence of a regime shift in 1989/1990, the time series of macrozoobenthos is well predictable during the studied period. This indicates a linear relationship with the predictor, the NAO. Because of the good results of the older study downscaling experiments were performed with updated time series. Surprisingly, the predictions failed. A possible reason is the presence of an abrupt or discontinuous biological regime shift. Swanson and Tsonis (2009); Wang et al. (2009) found a climate regime shift (CRS) in 2001/2002. The possibility that the reduced predictability originates from this CRS is studied here.

As mentioned in Chap. 1.1, several possibilities exist to identify regime

shifts. Collie et al. (2004) compiled a set of questions to identify the existence of biological regime shifts from the criteria noted by Scheffer and Carpenter (2003) that are applied here.

The time series for benthos biomass reveals significant steps in the 1989/1990 and 2001/2002 (Fig. 3.1a). The timing of the first step is in agreement with the well documented biological regime shift in 1989/1990 (e.g. Alheit et al. (2005)). The timing of the second step corresponds with the climate regime shift in 2001/2002. Steps in a time series do not prove a regime shift, but they are a necessary condition. The distribution of the biomass time series shows also bimodality (Fig. 3.1b), which corroborates the presence of a shift in the time series. A further affirmation are the strong peak in 1990 and two minor peaks in 1995 and 2002 in the 2nd time derivative (Fig. 3.1c) of the biomass time series. However, the 2nd derivative of the macrofauna abundance time series only shows a pronounced peak in 2002 (Fig. 3.1d). These results should not be seen as a proof for a regime shift, but they can be seen as an indicator for rapid change. Figs. 3.1c and 3.1d further suggest that the step in 2001/2002 is obviously of a different type than that in 1989/1990.

Statistical downscaling was performed for the NAO index as driver variable and for six response variables: the median and mean of macrofauna species number, log abundance and log biomass for the periods 1978-1999, 1978-2005, and 2000-2005. The correlation coefficient r between the regional observations and the cross-validated estimations and Brier-based score skill β were used as skill factors and computed for all combinations (Livezey 1995). For 1978–1999 (Table 3.1), highly significant correlations and Brier-based score skills indicated a relatively high potential predictability of benthic macrofauna community structure (Dippner and Kröncke 2003). For the whole period 1978–2005 only few, and for 2000–2005 no significant correlations and no meaningful skill exist. The same holds for the AMO, AO, and for the NECP reanalysis SST of the southern North Sea as driver variables (Table 3.1).

The fact that 2001/2002 BRS was different to the 1989/1990 BRS can be demonstrated by relating the SST to climate time series. The wavelet analysis (Fig. 3.2) shows a significant in phase coherence between NAO index and NCEP SST anomalies for the wavelet period of 14–15 years between the mid 1970s and the end of the 1990s. After ~ 2000 this coherence disappeared. In

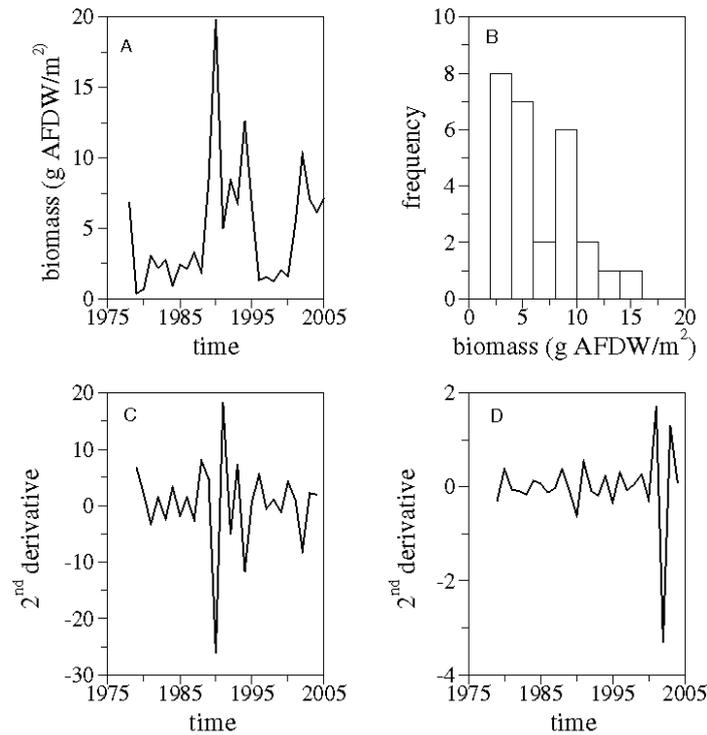


Figure 3.1: (a) Time series of benthic macrofauna biomass in the southern North Sea; (b) frequency distribution of biomass; (c) 2nd time derivative of biomass time series; (d) 2nd time derivative of abundance time series times 10^{-4} (Dippner et al. 2010)

contrast, wavelet coherence between AO index and NCEP SST anomalies indicates a strong in phase coherence for periods of 6-16 years (Fig. 3.2 bottom). A composite analysis of SLP DJFM anomalies (December-March average) north of 30°N shows for the period 1978-1999 a strong zonal orientation, but, a meridional orientation of SLP anomalies over North Atlantic and Eurasia for the period 2000-2009 (Fig. 3.3).

3.1.2 Benthos Predictability Using Neural Networks

In such situations where the non-linearity of a system strongly increases so that linear methods are not anymore feasible for prediction, two possibilities exist to enhance the skill of a forecast: either the development of a new predictor that has a more linear relationship with the response variable, or the use of non-linear statistical methods. Here are two relatively new methods applied

Predictor	Benthic Predictand	1978-1999	1978-2005	2000-2005
NAO	Med. log abund.	0.73 (0.46)	0.49 (0.17)	0.19 (-)
NAO	Med. species number	0.80 (0.61)	0.48 (0.17)	0.38 (-)
NAO	Med. log biomass	0.67 (0.42)	0.46 (0.17)	0.72 (-)
NAO	Mean log abund.	0.59 (0.22)	0.40 (0.07)	0.12 (-)
NAO	Mean species number	0.80 (0.58)	0.50 (0.17)	0.13 (-)
NAO	Mean log biomass	0.68 (0.42)	0.47 (0.18)	0.62 (-)
AMO	Med. log abund.	0.65 (0.37)	0.13 (-)	0.09 (-)
AMO	Med. species number	0.39 (0.07)	0.03 (-)	0.49 (-)
AMO	Med. log biomass	0.49 (0.18)	0.16 (-)	0.35 (-)
AMO	Mean log abund.	0.65 (0.38)	0.19 (-)	0.02 (-)
AMO	Mean species number	0.42 (0.09)	0.01 (-)	0.29 (-)
AMO	Mean log biomass	0.45 (0.13)	0.09 (-)	0.62 (0.22)
AO	Med. log abund.	0.47 (0.16)	0.36 (0.08)	0.29 (-)
AO	Med. species number	0.73 (0.52)	0.52 (0.24)	0.54 (0.01)
AO	Med. log biomass	0.53 (0.25)	0.39 (0.12)	0.40 (-)
AO	Mean log abund.	0.38 (0.10)	0.31 (0.06)	0.32 (-)
AO	Mean species number	0.67 (0.41)	0.51 (0.22)	0.50 (-)
AO	Mean log biomass	0.58 (0.29)	0.45 (0.15)	0.36 (-)
NCEP-SST	Med. log abund.	0.28 (-)	0.26 (-)	0.57 (-)
NCEP-SST	Med. species number	0.56 (0.22)	0.48 (0.17)	0.45 (-)
NCEP-SST	Med. log biomass	0.52 (0.25)	0.48 (0.15)	0.55 (0.38)
NCEP-SST	Mean log abundance	0.13 (-)	0.15 (-)	0.59 (-)
NCEP-SST	Mean species number	0.54 (0.19)	0.48 (0.16)	0.50 (-)
NCEP-SST	Mean log biomass	0.44 (0.10)	0.45 (0.13)	0.09 (-)

Table 3.1: Correlation coefficients and the Brier based score skill in parenthesis between climate predictors and benthic macrofauna predictands for different periods. Bold numbers mark significant correlations with respect to the 99% confidence level.

to study the possibility to enhance the prediction of the benthic macrofauna time series from the southern North Sea: OPELM and OPKNN as described in section 2.2.4.

OPELM and OPKNN are applied in a hindcast mode to the median of the biomass, abundance and species number for the whole period 1978-2005 using all climate indices and different SST time series of the North Sea as single predictors as well as in combination in a multivariate predictor. To test the performance of the hindcast, we apply OPELM and OPKNN to the same data set, but skipped the benthic data for the period 2000-2005, in order to predict this period with the models fitted for the previous years. As an example,

Figure 3.4 shows the prediction of median of the biomass obtained from the AO winter index as predictor.

Trends and amplitudes of inter-annual variability are well reproduced for the fitting period (1978-1999). However, the prediction of the neural networks for 2000-2005 is as poor as the linear statistical downscaling method (Dippner et al. 2010).

In a next step, we combine AO, SST of the southern North Sea, precipitation and meridional wind in a multivariate predictor. This predictor is used to train new models with OPKNN with possible lags of up to 11 years. We skipped the OPELM because the OPKNN method proved to be slightly superior regarding the skills over OPELM. The OPELM inferiority is due to inherent randomness of the method, coupled with small number of samples, which leads to high variability in predictions. The OPELM method seems more susceptible to limited data and the randomness of the model. This combination makes any kind of result for OPELM very variable, and therefore inferior to OPKNN. Moreover, since OPKNN is deterministic, the model output is much easier to interpret. We compare the results with the results of a prediction using only one predictor, the AO (Figs. 3.5, 3.6). Again, trends and amplitudes are well reproduced for both runs but the amplitudes are significantly better reproduced using the multivariate predictor. Here, the LOO error is 10.79 [number/m²] for species number time series, while for biomass it is $1.7 * 10^6$ [mgAFDW/m²]. Both errors are for OPKNN model.

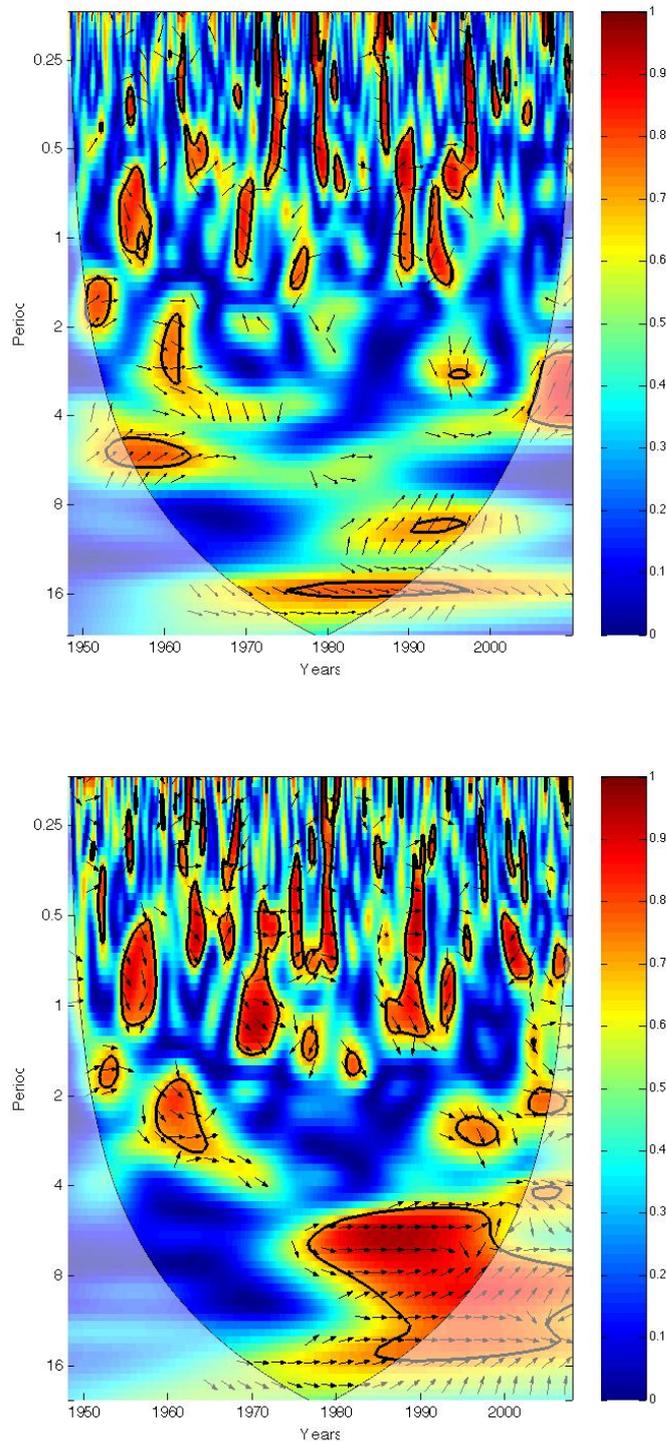


Figure 3.2: Wavelet coherence and phase between NAO index and NCEP SST (top) and AO index and NCEP SST (bottom) in the southern North Sea. Contours are wavelet squared coherencies. The vectors indicate the phase difference between NAO and SST (a horizontal arrow pointing from left to right signifies in phase and an arrow pointing vertically upward means the second series lags the first by 90° , i.e. the phase angle is 270°). The thick black line is the 5% significance level using the red noise model and the thin black line indicates the cone of influence.

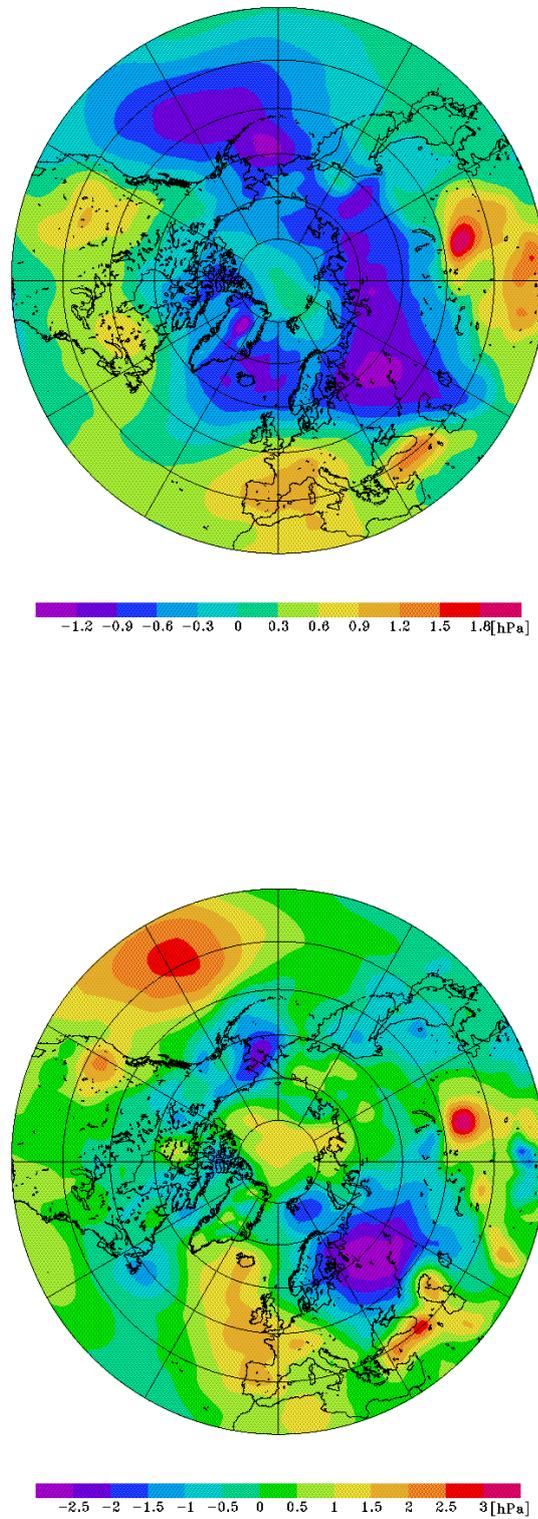


Figure 3.3: SLP Composite (December-March average) for the period 1978-1999 (top) and the period 2000-2009 (bottom).

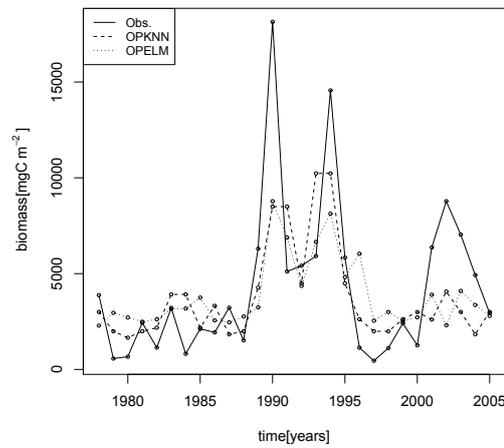


Figure 3.4: Median of biomass (mg AFDW /m²) of benthic macrofauna (full line) and their prediction from AO index using OPELM model (dashed line) and OPKNN model (dotted line).

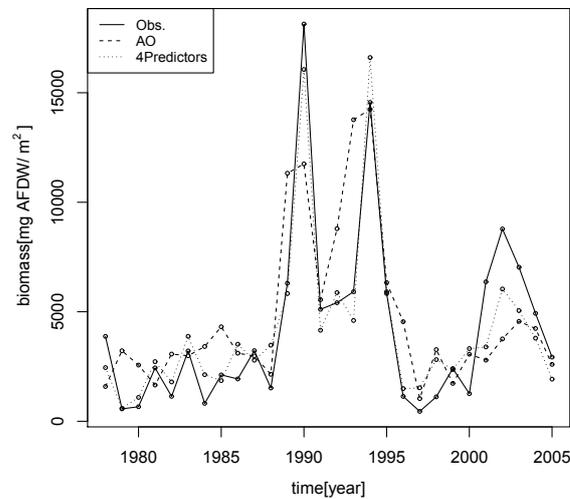


Figure 3.5: Median of biomass (mg AFDW /m²) of benthic macrofauna (full line) and their prediction from a multivariate predictor using OPKNN model and the AO as a single predictor (dashed line) and a multivariate predictor consisting of AO, SST of the southern North Sea, precipitation and meridional wind (dotted line). Fitting period is the complete time series; shown are the results of the LOO validation.

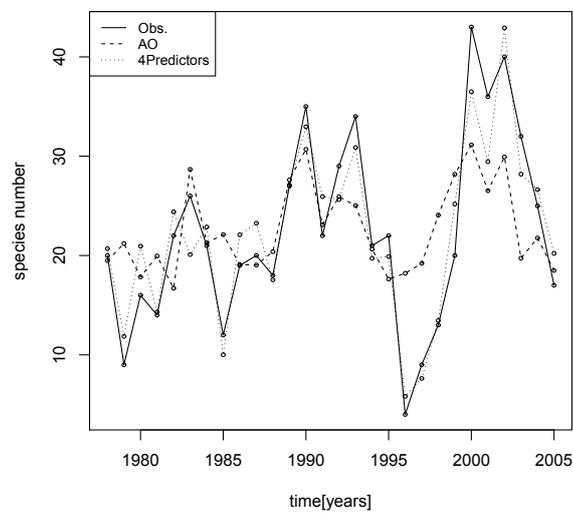


Figure 3.6: Median of species number of benthic macrofauna (full line) and their prediction from a multivariate predictor using OPKNN model and the AO as a single predictor (dashed line) and a multivariate predictor consisting of AO, SST of the southern North Sea, precipitation and meridional wind (dotted line). Fitting period is the complete time series; shown are the results of the LOO validation.

3.2 The Baltic Sea Environmental (BSE) Index

The construction and the results are based on Dippner et al. (2012) and completed with results from the poster Junker and Dippner (2011) and analysis of wavelet coherence.

3.2.1 Construction of the BSE

Analogous to the results from macrozoobenthos in the North Sea, the CRS around 2001/2002 is expected to decrease predictability also of the Baltic Sea ecosystem. Downscaling experiments following the work of Dippner et al. (2000), but with updated time series, confirmed the decreased predictability after 2000. Therefore, having shown that multivariate predictors are yielding better predictions, the development of a new index for the Baltic Sea area seems feasible to increase prediction skill.

Being a semi-enclosed intra-continental brackish water basin, the Baltic Sea has some unique characteristics that need to be taken into account for the selection of potential ecosystem drivers. It has a closed basin circulation (Voss et al. 2005), that causes a long water residence time. It is also characterized by strong horizontal as well as vertical salinity gradients and pronounced heterogeneity in ecosystem variables. Due to its horizontal salinity gradient, the Baltic Sea has marine species in the transition area to the North Sea and freshwater species at the end of the Gulf of Finland and the Bothnian Bay (Fig. 2.2). Further, the inter-annual and inter-decadal variability of the Baltic Sea is characterized by the climate variability on the northern hemisphere and major Baltic inflows of water with relatively high salinity propagating from the North Atlantic through the North Sea into the deeper parts of the Baltic Sea (Matthäus and Franck 1992).

A statistical analysis of the seasonal and inter-annual variations in the regional temperature anomalies of Sweden during 1861–1994 shows a strong relation to the NAO for the period 1985–1994 (Chen and Hellström 1999). However, correlation analysis over different periods shows that the strength of association varies with time and region (Chen and Hellström 1999). To improve statistical downscaling models, (Chen 2000) derived circulation indices for Scandinavia from monthly sea level pressure data based on the classifica-

tion system of Lamb (1950). These indices allow a reproduction of 70 % of the total variance in the January air temperature for Sweden during 1887–1994.

The large-scale atmospheric circulation patterns in the Arctic and North Atlantic described by the Arctic oscillation (AO) and NAO significantly control ice extent and ice thickness in the Baltic Sea (Omstedt and Chen 2001). AO and NAO are highly correlated if the atmospheric dynamics is driven by the North Atlantic (Deser 2000); however, the AO appears to describe more of the dynamics of the Baltic Sea ice conditions than the NAO (Jevrejeva et al. 2003). This is affirmed by wavelet coherence (Fig. 3.7), that shows strong anti-phase coherence of the AO over a wide spread of periods, whereas the also anti-phase coherence of the NAO with the ice extent is much weaker and ceases at the end of the 1980s, which coincides with the BRS in 1989/1990. The coherence of the NAO and ice extent in the >16 year period band is from 1980 outside the cone of influence and thus uncertain.

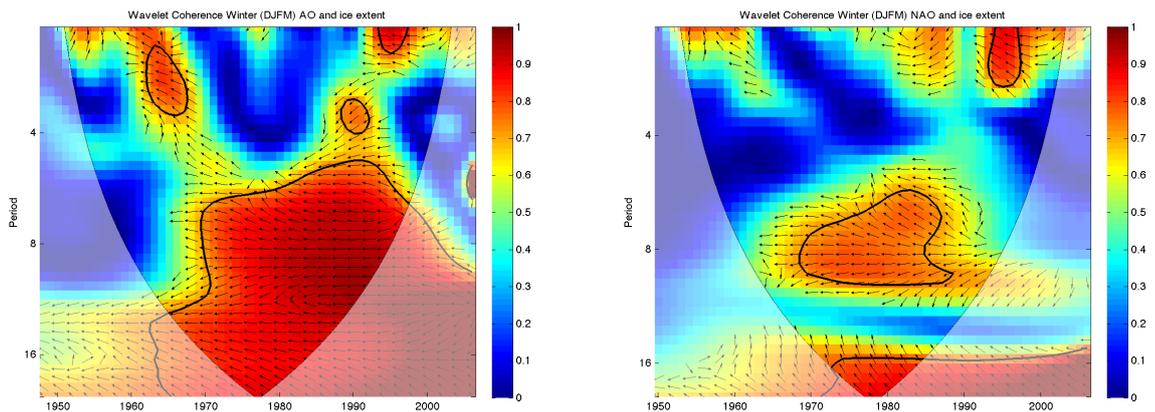


Figure 3.7: Wavelet coherences and phase between AO (left) and ice extent and NAO (right) and ice extent of the Baltic Sea. For further explanation see Fig. 3.2

Similar results can be found for Landsort gauge anomalies (Fig. 3.8) and the SST anomalies of the Gotland Sea (SST-GS) (Fig. 3.9). Also here the AO shows stronger, here in-phase, wavelet coherence than the NAO.

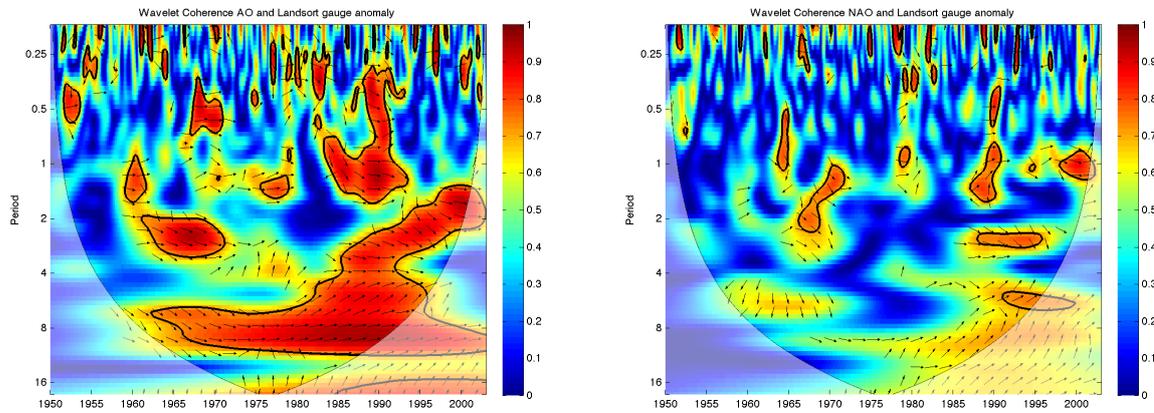


Figure 3.8: Wavelet coherences and phase between AO (left) and Landsort gauge anomalies and NAO (right) and Landsort gauge anomalies of the Baltic Sea. For further explanation see Fig. 3.2

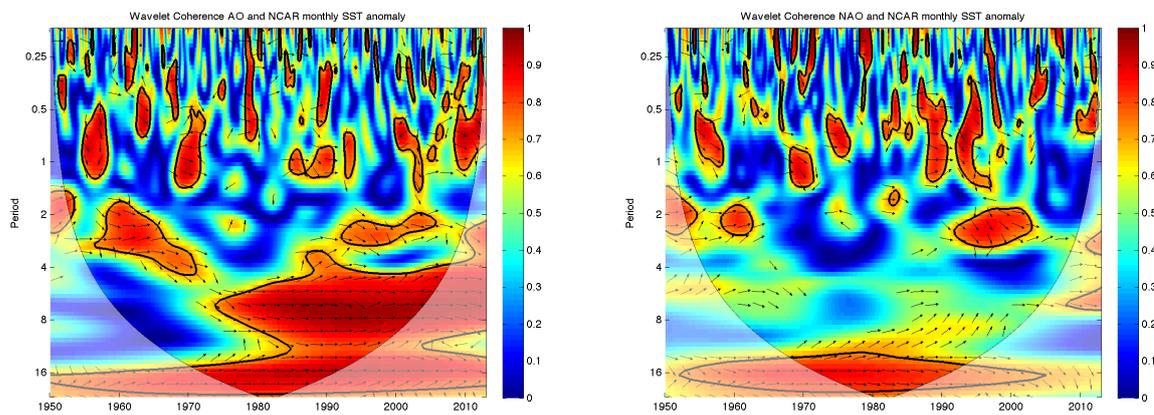


Figure 3.9: Wavelet coherences and phase between AO (left) and SST-GS and NAO (right) and SST-GS. For further explanation see Fig. 3.2

Salinity and oxygen concentration in the Baltic Sea also depend strongly on large-scale atmospheric circulation (Zorita and Laine 2000). A strong meridional sea level pressure gradient over the North Atlantic causes positive rain fall anomalies, increasing river runoff and decreasing salinity. Due to the weakened stratification, deep water oxygen concentrations increase (Gerlach 1994; Zorita and Laine 2000). An increase in precipitation will also result in higher input of nutrients or dissolved organic matter by rivers and enhanced eutrophication in near coastal areas with higher phytoplankton and benthic biomass (Dippner and Ikauniece 2001).

A long-term analysis of 100 years of hydrographic data with focus on the freshwater budget (Winsor et al. 2001, 2003) indicates, that freshwater supply to the Baltic Sea has large variations on time scales up to several decades. A similar result has been obtained by Omstedt et al. (2004). They argued that it is rather problematic to clearly define ‘trends’ or ‘regime shifts’ on shorter time scales because the Baltic Sea has decadal climate modes on the order of 30–60 years. Analysis of a cumulative Baltic winter index shows that during the last 350 years six climate regime shifts have occurred (Hagen and Feistel 2005).

From these findings, it was decided to select four time series for the construction of the BSE (Fig. 3.10):

- The AO index to represent the large-scale atmospheric circulation pattern.
- The salinity between 120 and 200 m in the Gotland Sea, to take account for the major Baltic Sea inflows
- The integrated river runoff of all rivers draining into the Baltic Sea to cover the influence from precipitation
- The relative vorticity of geostrophic wind over the Baltic Sea area from the Chen index to have a better representation of large scale blocking patterns.

All these time series have different length and the salinity time series is infrequent before World War II. To construct a consistent index, we select the period 1948–2002. Since over fitting of peaks is a common problem in the

selected downscaling method (Chap. 2.2.3), all time series are filtered with a 1-2-1-weighted moving average filter. An EOF analysis of the normalized anomalies of these four time series is computed and the 1st Principal Component score of the 1st EOF mode serves as the new multivariate BSE index (Fig. 3.11).

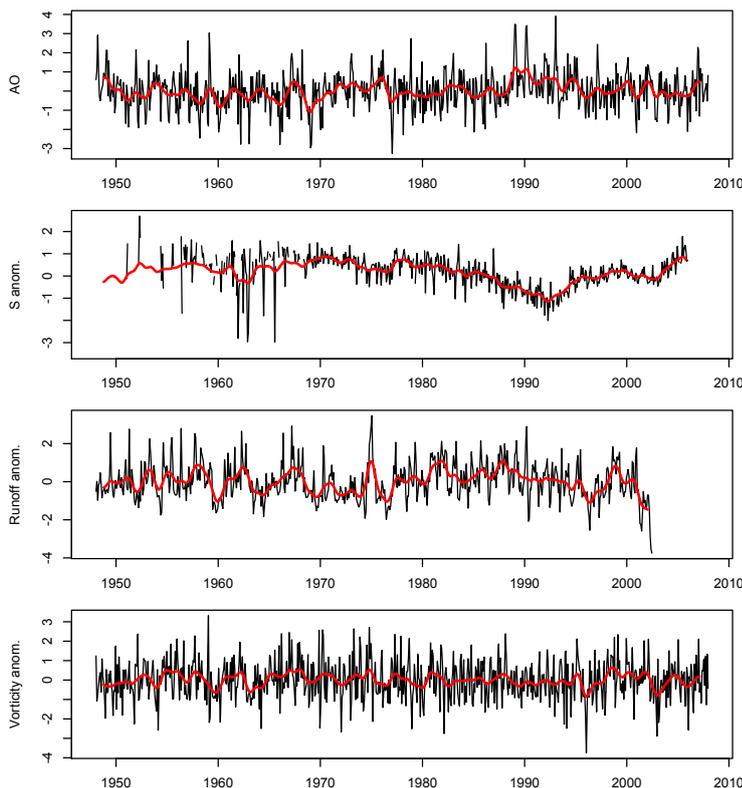


Figure 3.10: The four time series used to construct the BSE: monthly values of a) AO, b) anomaly of the salinity in the Gotland deep c) anomaly of the integrated river runoff, d) anomaly of the vorticity. The red line is the monthly time series filtered with a cut off of 25 months.

The analysis of the EOFs showed that the 1st EOF has an amount of explained variance of 38.5 %, the 2nd EOF of 28.8 %, the 3rd EOF of 17.3 % and the 4th EOF 15.4 %. The EOF patterns show that the first mode is dominated by the river runoff, the second mode by the AO, the third mode by the vorticity and the fourth mode by the deep water salinity. Due to the fact that the 1st EOF is controlled by the river runoff which ends in 2002, the performance tests using statistical downscaling can only be performed for periods until 2002.

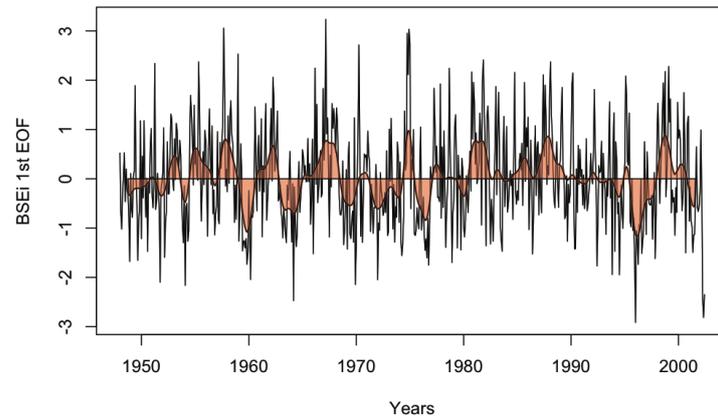


Figure 3.11: The time coefficients of the 1st EOF of the BSE. The red line is the time series filtered with a cut off of 25 months

3.2.2 Performance Tests

To test the performance of the new BSE index, statistical downscaling was used. In the first test, Atlantic multi-decadal oscillation, AO, NAO, Baltic Sea index, Chen and BSE index serve as climate predictor and SST in the Gotland Sea, Landsort gauge and sea ice extent serve as environmental response variable. To test the performance of the BSE index with biological data, in addition to the climate indices, COADS SST is used as climate predictor since it was the most powerful predictor in the older study by Dippner et al. (2000) and the various mesozooplankton time series from Latvian monitoring as regional predictands.

Results for Physical Response Variables

No significant correlation has been found between the Atlantic multi-decadal oscillation and the physical data. The regional Baltic Sea index has a better performance than the global indices AO and NAO. The same holds for the Chen and the BSE index. A comparison of Baltic Sea index, Chen index with three EOFs and BSE index with four EOFs shows similar results for the SST in the Gotland Sea and the ice extent for winter time. Concerning Landsort gauge, all indices show a clear correlation for winter data. The Baltic Sea index (BSI) has a highly significant correlation ($r = 0.66$), but, the Chen index (r

= 0.84) and the BSE index ($r = 0.87$) are significantly better. The Chen and the BSE index, both indicate high correlations and skills throughout the year in the case of Landsort gauge (Fig. 3.12). The only exceptions are April for the Chen index and April, May and June for the BSE index. During these period no correlations can be identified (Fig. 3). Altogether, the performance test with the above-mentioned physical data indicates that the BSE index has a better correlation and a better skill than global indices and the Baltic Sea index and is equivalent to the Chen index with respect to the model skill (Table 3.2).

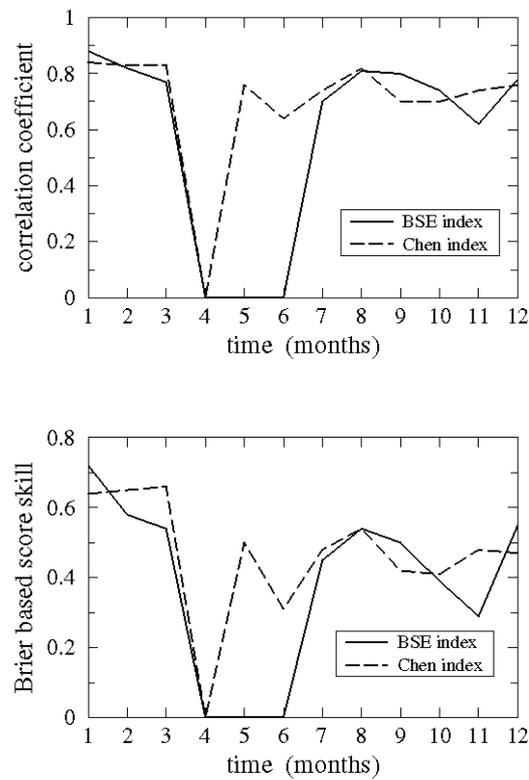


Figure 3.12: Correlation coefficient and Brier-based score skill as function of the month of the year for the downscaling projections of BSE and Chen index on Landsort gauge

	SST-GS (2/-1)	LG (1/0)	IE (1/0)
AMO	ns	ns	ns
AO	0.64 (0.39)	0.62 (0.37)	0.61 (0.34)
NAO	0.62 (0.35)	0.65 (0.41)	0.54 (ns)
BSI	0.71 (0.49)	0.66 (0.42)	0.67 (0.43)
Chen (1)	ns	0.59 (0.31)	0.37 (ns)
Chen (2)	0.61 (ns)	0.84 (0.63)	0.46 (ns)
Chen (3)	0.75 (0.45)	0.84 (0.64)	0.72 (0.42)
BSE (1)	0.69 (0.45)	0.73 (0.51)	0.67 (0.44)
BSE (2)	0.72 (0.48)	0.87 (0.67)	0.68 (0.44)
BSE (3)	0.73 (0.48)	0.87 (0.67)	0.70 (0.44)
BSE (4)	0.73 (0.48)	0.87 (0.68)	0.70 (0.45)

Table 3.2: Correlation coefficients and the Brier-based score skill in parentheses between the climate predictors: AMO, AO, NAO, BSI, Chen index, BSE index and Baltic Sea predictands: the SST in the GS, the mean sea level at LG and the IE. Bold numbers mark significant correlations with respect to the 99.9% confidence level for the correlation coefficient and for the 99% confidence level for the skill. The number in parentheses in the predictor column denotes the number of considered EOFs. The numbers in parentheses in the predictand row display the considered month of the years and the time lag in month. ns: no significance.

Results for Zooplankton

The performance of the BSE index is also tested using Latvian zooplankton time series. In a previous article, Dippner et al. (2000) demonstrated that a major amount of inter-annual variability in zooplankton time series can be explained by climate variability using COADS SST as climate predictor. This downscaling experiment is repeated using COADS SST and BSE index as predictor for the period 1960–1992 (Table 3.3). The results indicate that the prediction is equivalent in case of *Acartia* spp. whereas for *E. nordmanni*, the BSE index has a higher correlation. For the period 1960–1992, a downscaling experiment using COADS SST as predictor shows a better correlation and skill for *Synchaeta* spp. and *T. longicornis*. No meaningful correlation has been found for *B. longispina* and *Pseudocalanus* spp. In the final step, all considered climate indices like Atlantic multi-decadal oscillation, AO, BSI, BSE, Chen, NAO and an updated COADS SST were used as predictor for the Latvian mesozooplankton time series for the period 1960–2002. The BSE

index was used twice, with one and with four considered EOFs, respectively (Table 3). This experiment showed that the BSE index using only the 1st EOF as predictor performs better than all other indices

	BSE (1960–1992)	COADS-SST (1960–1992)
Acar (S1/M2)	0.70 (0.37)	0.70 (0.37)
Bos (S2/M7)	ns	ns
Evad (S1/M2)	0.80 (0.55)	0.78 (0.60)
Syn (S1/M3)	0.75 (0.32)	0.76 (0.50)
Temo (S1/M2)	0.64 (0.27)	0.71(0.42)

Table 3.3: Correlation coefficients and the Brier-based score skill in parentheses between climate predictors and zooplankton biomass as used in the previously performed downscaling experiment (Dippner et al. 2000) The fitting period is 1960–1992. *Acartia* spp. (Acar), *Bosmina longispina* (Bos), *Evadne nordmanni* (Evad), *Syncheata* spp. (Syn) and *Temora longicornis* (Temo) as predictands. The season is denoted with an "S" and the season number, with S1 as spring (May values) and S2 as summer (August values). The month of the predictor is denoted with an "M" and the number. Bold numbers mark significant correlations with respect to the 99.9% confidence level for the correlation coefficient and for the 99% confidence level for the skill. ns: no significance

3.3 Baltic Sea Phytoplankton Predictability

Encouraged by the results of the downscaling of the physical and zooplankton time series, phytoplankton time series from German monitoring programs are chosen to study the potential predictability. One data set comes from the LUNG (Landesamt für Umwelt, Naturschutz und Geologie Mecklenburg-Vorpommern 2006), the other from the IOW, in charge of the federal monitoring program for the Baltic Sea (Chap. 2.1.3)

Due to the differences of coastal and open sea ecosystem, it is expected that the two systems have different dynamics and possibly a different predictability. The coastal zone of Mecklenburg-Western Pommern and the adjacent Arkona Sea 2.1.3 are taken as examples of these two systems and studied separately. The time series of the phytoplankton groups are first run through EOF analysis and then statistical downscaling. The used data sets and the statistical methods are thoroughly described in Chapter 2.

3.3.1 Results EOF Analysis

EOF analysis showed that more than half of the variance is already explained by the 1st EOF for the Arkona data in spring (MAM) and fall (SON) phytoplankton (Tab. B.2). The 1st EOF explained even more variance in nutrients, salinity, temperature and chlorophyll-a (Tab. B.1). The only exceptions are Dinophyceae abundance, Crysophyceae biomass and the composite parameters total flagellates, and phytoplankton spring abundances and the ratio of Bacillariophyceae:Flagellates. Dinophyceae and the ratio of Bacillariophyceae:Flagellates have relatively high amount of variance explained also by the higher EOFs, which indicates a high noise level due to spatial heterogeneity. For the others, the first two EOF together usually accounted for more than 75% of the variance.

For LUNG phytoplankton data, the picture is a bit different. Taking all stations into account (group A), the first EOF accounts for less than 30% of the explained variance of phytoplankton abundance. Exceptions are Bacillariophyceae and the Diatom:Flagellate ratio where the first EOF accounts for almost 40% in the three considered seasons. The first EOF explains only slightly more variance if stations S66 and KHM are removed (group B). For

group A and B, the first 3 EOF are needed to account for more than half of the variance. If only the outer coastal stations are considered, the first EOF explains at least 30% of the variance and up to 67% for fall Dinophyceae. The second EOF explains also around 30% of the variance for all phytoplankton groups. Thus, the first two EOF explain more than 60% of the variance. (Tables B.3 - B.6)

Overall, variance of biomass data is slightly (around 5%) better explained by the first EOF than for abundance. This is valid for all phytoplankton groups.

Examination of the time coefficients of the 1st EOF (Appendix B) and the loads affirms the high spatio-temporal variability in phytoplankton. Table 3.4 gives a quick overview of the linear trends in the time coefficients of the 1st EOF, plots of the time coefficients and loads can be found in Appendix B. The trends are very inhomogenous, especially for the composite parameters. It is difficult to find a common pattern across the different station groups. Only Cryptophyceae abundance shows an increasing trend for all station groups, while biomass resp. biovolume are exhibiting a slightly decreasing trend.

Spring Chlorophyceae biomass is decreasing significantly in the Arkona Sea, while Cryptophyceae and Dinophyceae increase in abundance but not in biomass. The total of Flagellate biomass also decreases as well as the total phytoplankton biomass. In summer, the Bacillariophyceae abundances and the biomasses of Cyanophyceae, Dinophyceae, Euglenophyceae, Prasinophyceae, total flagellates and total phytoplankton decrease significantly. The decrease in cyanobacteria biomass is mainly caused by the very low amounts in 2005 and 2011. The abundances remain relatively stable. Just like in spring Cryptophyceae abundances increase also in fall significantly. The increasing trend in Cryptophyceae abundance is to some extent caused by a few instances with very low cell numbers in the first quarter of the time series, however, the increasing trend remains even after omitting these instances. Cyanobacteria biomass time series shows a significant decreasing trend in fall, while a trend in abundance cannot be identified.

The LUNG data set displayed a very variable picture. Most prominent is the increase in Cryptophyceae abundance in all station groups in spring, just as noticed already in the Arkona Sea. For the group D, with only stations of the inner coastal area, even biovolume increased in spring. Cryptophyceae

	MAM				JJA				SON							
	Ar.	A	B	C	D	Ar.	A	B	C	D	Ar.	A	B	C	D	
Bacillariophyceae	a	↗	→	↗	↗	↘ ⁹⁵	↗	↗	↗	↗	↗	→	↗	↗	↗	
	b	↘	→	↗	↗	↗	↘	↗	↗	↗	↗	→	↗	↗	↗	↗
Chlorophyceae	a	→	↘ ⁹⁵			→	→									
	b	↘ ⁹⁵				→	→									
Chrysophyceae	a	↗	↗			↗	↗									
	b	↗	↗			↗	↗									
Craspedophyceae	a	→	→			→	→									
	b	↗	↗			→	→									
Cryptophyceae	a	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁰	↗ ⁹⁵	↗	↗	↗	↗	↗	↗ ⁹⁵	↗ ⁹⁵	↗	↗	↗ ⁹⁵	↗
	b	→	→	↗ ⁹⁰	↗ ⁹⁰	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
Cyanophyceae	a	↗	↗			↗	↗	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁵	↗	↗	↗	↗	↗
	b	↗	↗			↗	↗	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁵	↗	↗	↗	↗	↗
Dinophyceae	a	↗ ⁹⁰	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
	b	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
Euglenophyceae	a	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
	b	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
Prasinophyceae	a	→	→			→	→									
	b	↗	↗			↗	↗									
Total Flagellates	a	↗ ⁹⁵	↗ ⁹⁵	↗ ⁹⁰	↗	↗	↗	↗	↗	↗	↗ ⁹⁰	↗	↗	↗	↗	↗
	b	↘ ⁹⁰	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
Total Phytoplankton	a	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
	b	↘ ⁹⁰	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
Bacil.:Flagel.	a	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗
	b	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗	↗

Table 3.4: Trends of the time coefficients of the 1st EOF of phytoplankton. ↑ and ↓ mark a strong trend but with significance level <90%, ↗ and ↘ mark significant trends with the level of significance as superscript, ↗ and ↘ a weak trend, → no trend. a is abundance, b is biomass for Arkona data or biovolume for LUNG data. Ar. marks Arkona data, the letters A-D the station groups of the LUNG data.

biovolume trend was not significant for the other stations in spring, but in summer, where a significant decrease could be registered. The increase in Cryptophyceae abundance in fall as in the Arkona Sea could also be found for the LUNG stations, but it was only significant for the station groups A, B and D. Due to the strong increase in Cryptophyceae, total Flagellate abundance also increases in spring and in fall. The most striking feature in the summer time coefficients of the 1st EOF is the increase in Cyanobacteria abundance after 1992. This is not as pronounced in the Arkona data, where an increase in cyanobacteria abundance can be seen in the SON averages after 1992. The increasing trends for data after 1992 are significant for all station groups. Since the LUNG dataset is characterized by a large amount of missing values and zero abundance and biovolume, only a few phytoplankton classes and groups were suitable for EOF and also for the subsequently described downscaling procedure.

Another interesting feature is the occurrence of a turning point in the curve resulting from local polynomial regression fitting, the LOESS smoother, for most phytoplankton classes and composites in the Arkona Sea as well as the LUNG data set at the end of the 90s and an increase in variance. An increase in variance can be interpreted as a sign of a coming regime shift (Carpenter et al. 2011) and a change in trend direction is indicative of a regime shift (Scheffer and Carpenter 2003).

3.3.2 Downscaling

To estimate the influence of climate on the registered phytoplankton variability, statistical modeling is often used to correlate the response time series with climate indices. It has predictive capabilities and is thus feasible to study the predictability. Since the results of statistical downscaling from the previous studies are encouraging, and because of its robustness and ease, it was decided to use linear statistical downscaling (2.2.3) also for phytoplankton development.

Pre-runs of statistical downscaling of phytoplankton abundances and biomass show that the performance of the BSE as predictor is usually superior over the large scale indices and even local nutrient concentrations as predictors, which do not perform at all. Other indices are only for very few phytoplankton classes

and only for selected seasons equally good or marginally better descriptors. An exception are Cryptophyceae where the overall best Brier-score β of a model fit was found for spring abundance with $\beta = .21$ with the AMO as predictor in contrast to $\beta = 0.2$ with the BSE. However, the AMO does not perform for any season of any other phytoplankton class. The BSE is clearly the most versatile index and therefore the BSE is used as climatological predictor for all phytoplankton classes and composites.

Since many phytoplankton classes and groups of the LUNG data set have large gaps or zeros, the number of groups to downscale is strongly reduced. According to the results of EOF analysis, also the remaining groups are highly variable. This already indicates the lower predictability of the coastal ecosystem. In accordance, no meaningful correlations with a significance level of at least 90% were found for any of the remaining groups of the LUNG data set. Due to the weak results of the downscaling experiments, further experiments with this data set were discontinued.

Tables 3.5 and 3.6 list the best correlation coefficients, best Brier-based scores and predictive months of the best model fits for 1979-2005 and 1979-2000 respectively for each season and phytoplankton group of the Arkona Sea. The data density is too low for the winter season, so it is excluded. The best correlation coefficients and Brier-scores are taken from the model results for the station data, i.e. while one station might have a good correlation of fit with observed data, the others might have bad correlations. Only those data sets with more than 70% data points (presence) were attempted to fit. Only the best results for all meaningful lags (here month of the predictor) are shown.

Recalling that the runoff time series ends in 2002 and the potential regime shift in 2000 (Dippner et al. 2010) these two different fitting periods are necessary to see whether the fit for the period (1979-2005) is enhanced by restricting the fitting period. Extending the fitting period for downscaling with the BSE as predictor would not make any sense since the second time series, salinity, ends after 2005. Downscaling with large scale climate indices also for the period 1986-2010 did not yield any better results than for the shorter fitting periods, as expected. A single large scale predictor clearly does not suffice to explain a significant amount of variance in phytoplankton in the Baltic Sea.

As EOF analysis has shown, the first two EOF of phytoplankton time series

account for more than 70% of variance. Therefore, two EOF are considered enough for the phytoplankton as response in downscaling. The predictor, the 1-2-1-filtered BSE, is used with all 4 possible EOF, following the results of the performance tests with zooplankton and the physical responses (Chap. 3.2.2). Pre-runs have shown that the unfiltered BSE might give slightly better results for some few specific phytoplankton groups, seasons and lags. Given that the results are only marginally better and for only so very few predictands the unfiltered BSE is not further included in this study.

Most data of the Arkona Sea can be fitted for the period 1979-2005 so that the variance of the error is somewhat smaller than the variance of the observation for at least one station (the maximum Brier-score in Tables 3.5 and 3.6). For some phytoplankton classes and seasons no correlations can be found for any station: spring Euglenophyceae, spring biomass of Cryptophyceae, summer Craspedophyceae, and fall Chlorophyceae, Chrysophyceae, total flagellates and the fall biomass of Cryptophyceae, Cyanophyceae, Dinophyceae and Euglenophyceae, and the fall abundance of Bacillariophyceae. Here the intermediate results were lower than necessary to continue calculation of the algorithm.

For the period 1979-2000 the picture regarding the found correlations is similar, the differences are mainly that correlations could be found additionally for Euglenophyceae spring abundance and biomass and fall biomass, fall Dinophyceae and the ratio of Cyanophyceae abundance in fall. No correlations could be additionally found for some summer abundances: Chlorophyceae, Cryptophyceae and the ratio of Diatoms:Flagellates.

Correlations, are fairly low and the variance of the error is often close to the variance of the observations even the best values. One has to keep in mind that the correlations and Brier-scores in tables 3.5 and 3.6 are the best values of all stations. Calculating mean correlations and Brier-scores give a different picture. Here, the Brier-scores are often even negative and good scores might be based on only few data points(Fig 3.13).

Thus, comparison of the time coefficients of the 1st EOF of modeled and original data is probably giving a better impression of the model skill. The plots of the time coefficients and the Brier-scores of the model results are shown in the Appendix B. Despite that the maximum Brier-score is usually

positive, the model skill based on the Brier-score of the time coefficients of the 1st EOF is negative, indicating meaningless results. The trends and the turning points however, are well met for some response variables: e.g. the trend and the turning point for Cryptophyceae data and for spring abundance of Bacillariophyceae the turning point around the 1990s.

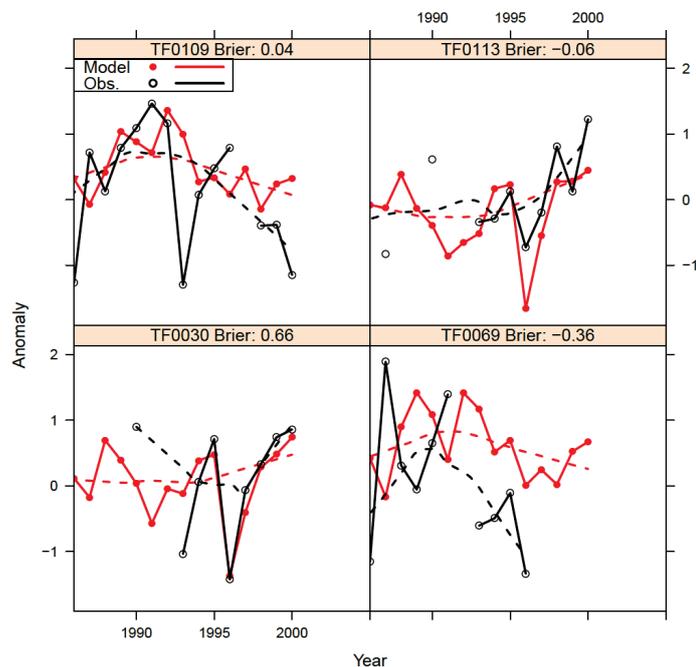


Figure 3.13: Prasinophyceae spring abundance. Observations (black) and model fit (red) for the station data. Dashed curves are the respective LOESS filtered series.

The differences in model skill become clear in the plot of station data for spring Prasinophyceae abundance (Fig. 3.13). Plots of summer Cyanobacteria abundance for both modeling periods are shown in (Fig. 3.14). Although the runoff time series discontinues in 2002 remain the model results comparable.

The advantage of using a longer time series might thus outweigh the problem with the short runoff time series. Spring biomasses are mostly correlated with

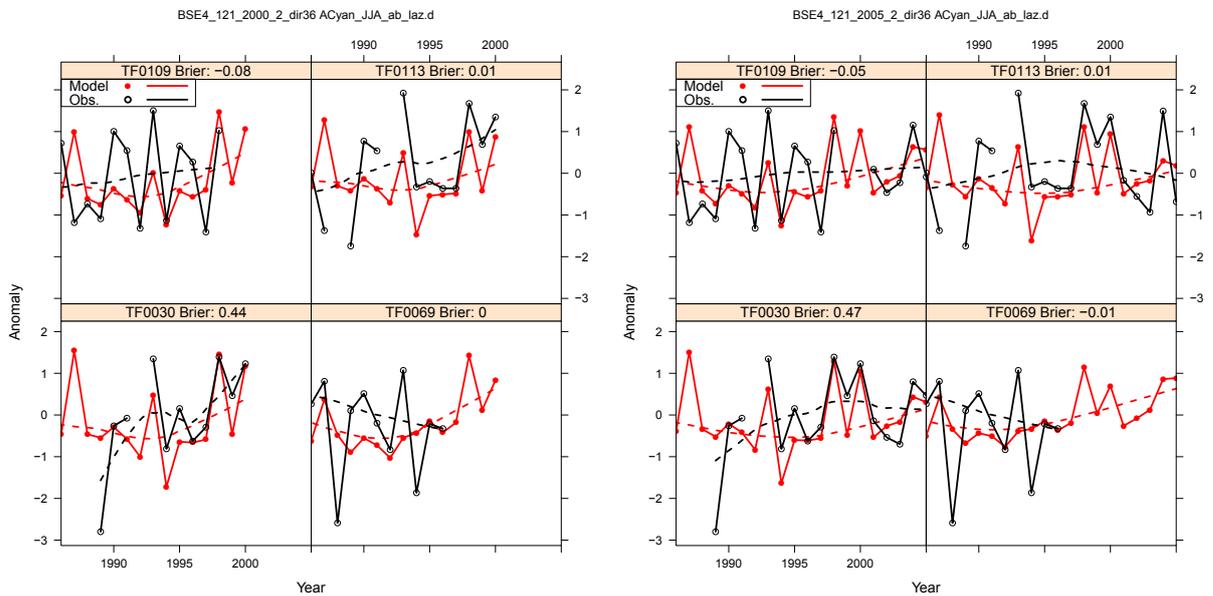


Figure 3.14: Cyanobacteria summer abundance. Observations (black) and model fit (red) for the station data. Dashed curves are the respective LOESS filtered series.

the January BSE in both modeled time spans. The BSE is 1-2-1-filtered, so it is actually the weighed winter BSE that is yielding the best correlations. For both modeled time spans usually the same lag of the BSE resulted in the best estimates.

The results indicate, that phytoplankton has some potential predictability, and that climate certainly has influence on phytoplankton, even though relationships can not be identified for all classes for all seasons and correlations are weak for most stations.

1979-2005		MAM			JJA			SON		
		Cor.	β	M	Cor.	β	M	Cor.	β	M
Bacillario.	ab	0.4	0.16	1	0.1	-0.02	7	0.48	0.23	10
	bm	0.34	0.11	3	0.2	0	6			
Chloro.	ab	0.09	-0.12	4	0.54	0.29	6			
	bm	0.32	0.02	4	0.31	0.09	6			
Chryso.	ab	0.4	-0.06	1	0.7	0.47	5			
	bm	0.56	0.3	1	0.17	-0.06	5			
Craspedo.	ab	0.59	0.28	4				0.28	-0.06	10
	bm	0.51	0.24	1				0.52	0.24	10
Crypto.	ab	0.2	-0.06	3	0.41	0.14	5	0.42	0.15	10
	bm				0.48	0.19	5			
Cyano.	ab	0.42	0.12	2	0.69	0.47	7	0.5	0.25	8
	bm	0.49	0.24	2	0.54	0.27	7			
Dinoph.	ab	0.4	0.16	2				0.56	0.31	11
	bm	0.63	0.36	4						
Eugleno.	ab				0.16	0	6	0.47	0.15	8
	bm				0.49	0.24	7			
Prasino.	ab	0.57	0.32	1	0.59	0.34	3			
	bm	0.58	0.33	3	0.34	0.09	6			
Total Flagel.	ab	0.27	0.05	1	0.46	0.2	5			
	bm	0	-0.13	1	0.37	0.13	3			
Total Phytopl.	ab	0.28	0.03	1	0.08	-0.08	2	0.51	0.26	8
	bm	0.11	-0.07	2	0.25	-0.01	2	0.17	-0.05	10
Ratio Bacil: Flagel.	ab	0.09	-0.05	1	0.14	-0.02	7	0.28	0.02	8
	bm				0.37	0.02	3	0.2	-0.03	11
Ratio Cyano.: Phytopl.	ab	0.43	0.11	3						
	bm									

Table 3.5: Maximum correlation coefficients and the maximum Brier-based score skill β between BSE and phytoplankton for the period 1979–2005. M denotes the month of the predictor. Only results with positive correlation or Brier-score are shown.

1979-2000		MAM			JJA			SON		
		Cor.	β	M	Cor.	β	M	Cor.	β	M
Bacillario.	ab	0.24	0.01	1	0.38	0.1	7	0.52	0.27	10
	bm	0.58	0.3	3	0.07	-0.11	6			
Chloro.	ab	0.6	0.29	4	0.49	0.15	6			
	bm	0.31	0.07	4	0.36	0.11	6			
Chryso.	ab	0.54	0.27	1	0.73	0.43	5			
	bm	0.52	0.25	1	0.67	0.45	5			
Craspedo.	ab							0.28	-0.05	10
	bm							0.63	0.28	10
Crypto.	ab	0.49	0.2	3	0.32	0.07	5	0.45	0.13	10
	bm				0.45	0.2	5			
Cyano.	ab	0.4	0.04	2	0.67	0.44	7	0.66	0.42	8
	bm	0.47	0.22	2	0.71	0.51	7			
Dinoph.	ab	0.3	0.06	2	0.23	0.04	4	0.44	0.19	11
	bm	0.5	0.25	4	0.38	0.14	3			
Eugleno.	ab							0.34	0.06	8
	bm				0.73	0.51	7			
Prasino.	ab	0.82	0.66	1	0.8	0.63	3	0.54	0.29	11
	bm	0.5	0.24	3	0.39	0.09	6			
Total Flagel.	ab	0.39	0.11	1	0.18	-0.02	5			
	bm	0.61	0.37	1	0.29	0.07	3			
Total Phytopl.	ab	0.68	0.46	1	0.1	-0.09	2	0.55	0.3	8
	bm				0.27	-0.05	2			
Ratio Bacil: Flagel.	ab	0.18	0.01	1	0.39	0.15	7	0.47	0.21	8
	bm				0.6	0.13	3			
Ratio Cyano.: Phytopl.	ab	0.47	0.15	3				0.08	0.01	11
	bm									

Table 3.6: Maximum correlation coefficients and the maximum Brier-based score skill β between BSE and phytoplankton for the period 1979–2000. M denotes the month of the predictor. Only results with positive correlation or Brier-score are shown

Chapter 4

Discussion

4.1 Occurrence of regime shifts

Parts of this Discussion have been published previously in Dippner et al. (2010) and (Junker et al. 2012).

Analysis and downscaling experiments with the North Sea macrozoobenthos data set indicate two biological regime shifts: 1989/1990 and 2001/2002. In the period 1978-1999, in which the 1989/1990 BRS occurred, statistical downscaling methods indicate a relatively high potential predictability of benthic macrofauna community structure (Dippner and Kröncke 2003). However, after ~ 2000 the correlations between NAO index and North Sea SST, macrofauna species number, abundance and biomass failed using linear statistical downscaling.

In order to increase predictive skill in presence of non-linear relations with the physical forcing, e.g. as in abrupt regime shifts, one can either search for a more suitable description of the predictor, that allows for a more linear relationship with the response variable, or use non-linear methods for prediction, that are able to reproduce the non-linear relationships, or a combination of both. Therefore, a relatively new kind of neural network algorithms were used, OPELM and OPKNN. They provide an alternative framework for predicting observations, if prior knowledge of the phenomenon is lacking or completely unknown. With fast learning times, both models provide a suitable framework for testing different combinations of predictors, and ranking provides insight which predictors work best together. This can be achieved in reasonable time

with many predictors.

Yet, even with these far more sophisticated methods than linear statistical downscaling, it was not possible to predict benthos biomass after 2000, if the learning set was restricted to the period before 2000. This confirms the conclusions from the experiment using linear downscaling, that an abrupt or discontinuous BRS occurred around 2000 or 2001/2002 where the system shifted to another basin of attraction. Allowing the algorithm to learn also from data past 2000, the prediction was greatly enhanced especially when a multivariate predictor is considered. This is in agreement with observations that the ecosystem has shifted to a new state where other factors determining how climate acts on the ecosystem have become dominant, e.g. switch from temperature to food availability limited growth.

The question still remains, where the differences in predictability and transition during the two BRSs originated.

As already mentioned, one source of unpredictability of the ecosystem is the occurrence of a highly non-linear relationship between the forcing, climate, and the response, macrozoobenthos in this case. This is typical for an abrupt or discontinuous shift according to the classification scheme for regime shift of Collie et al. (2004). On the other hand, when a time series remains predictable using linear downscaling despite significant shifts in the time series indicates this a continuously quasi-linear relationship with the predictor which is typical for a smooth shift.

In order to find the reason for the different types of shifts, a closer look at the changes in the atmosphere might be of help:

From 1976 to 2001, no change in trends occurred in global mean air temperature (Swanson and Tsonis 2009) or in SST in the North Atlantic and North Pacific. The AMO does not show any change in trend between 1976 and 2001 either (Dima and Lohmann 2007). Wavelet coherence analysis (Fig. 3.2) is supporting this, since significant coherence between the NAO index and the North Sea SST exists for this period in the ~ 16 year band. Both time series are in phase and the climate variability of the North Atlantic forces the SST in the southern North Sea. More striking is the strong phase coherence of AO and North Sea SST from the 1980s. This indicates a possible superiority of the AO as a predictor for the North Sea SST, which is also supported by (Deser

2000) who mentioned that the two station NAO index is not the optimal representation of the spatial pattern associated with it. Hurrell and Deser (2010) found by applying non-linear cluster analysis to winter daily SLP fields of the North Atlantic, an increase in the occurrence of the Atlantic Ridge and Atlantic Blocking patterns since 2001. This is supported by the composite winter SLP pattern for 2000-2009 (Fig. 3.3), which shows a strong meridional orientation of the SLP anomalies, typical for Atlantic blocking and ridge patterns. This increasing dominance of previously minor patterns is indicative for an increase in non-linearity in the climate system. This shift of the climate system towards higher non-linearity might be the reason for the unpredictability of the benthos time series, indicative for an abrupt regime shift.

In contrast, the regime shift in 1989/1990 is not accompanied with a CRS. No change in trends in physical properties occurred during this period, which is characterized by continuous warming. According to Hoerling et al. (2001), the continuous warming in 1976-2001 is caused by an increased heat transport from the tropics to the extra tropics, supporting a strong positive NAO, which is expressed in the well known warm winters and westerly winds in the study area. The continuously positive warming trend and strong NAO indicates a positive feedback mechanism according to the terminology of Suarez and Schopf (1988), and the macrozoobenthos responds in a quasi-linear way, typical for a smooth type BRS according to Collie et al. (2004).

To wrap up, the unpredictability of the time series during the shift around 2001/2002 indicate a non-linear shift, while the shift in 1989/1990 is a smooth shift. This shift is obviously triggered by large scale transitions in the atmosphere, i.e. a CRS, expressed in changes in the pressure field over the Northern Hemisphere and trend of temperature (Swanson and Tsonis 2009). These changes in pressure field resulted in a lower storm frequency in winter (Loewe 2009). The resulting calmer hydrodynamic conditions seem to favour tube building polychaetes (e.g. *Owenia fusiformis*) and burrowing amphipods (e.g. *Urothoe poseidonis*) and are resulting in exceptionally high abundance of juveniles of various species. While these species increased after 2000 were interface-feeding polychaetes, mobile amphipods and bivalves decreasing (Kröncke et al. 2013).

Another source of unpredictability is the insufficient description of the pre-

dicator. The improvement of the prediction by using a multivariate predictor shows the importance of local forcing on ecosystems (Dippner et al. 2012). The benthos remains unpredictable after the shift, even for the LOO-prediction, when considering only the AO, which proved to be a sufficient predictor for the period before 2000. However, if also local properties like precipitation and SST of the southern North Sea are considered, the trend and the dynamic even after 2000 are well reproduced (Fig.3.5). The prediction of the species number is showing this enhancement exceptionally clearly because the increase in species numbers seems to be related to the increasing SST (Fig. 3.6). Abundance was affected by a dramatic increase of juveniles in 2002, which was probably caused by exceptional high SST anomalies in the southern North Sea in the 1st quarter of 2002 resulting in early primary production and sufficient food availability for the larvae. But less precipitation since the shift seemed to have led to generally lower primary production (van Beusekom et al. 2009) and less food availability throughout the years, which caused the high mortality of the 2002 juveniles. Primary production or carbon flux might be missing factors in the data analyses and the reason for the weak correlation for abundance data.

Despite non-linear relationships between response and driver and an insufficient predictor, also data issues must be discussed here as a source of unpredictability. The low number of data points might have the biggest influence on the low predictability of the abundance data. Abundance is dependent on climate through several processes e.g. production of juveniles under good conditions and on the other hand death due to unfavourable conditions, which are acting on different time scales. The time series thus must have a sufficient length for the neural network to learn these relations. Our time series is not only short (28 years) but also characterised by several regime shifts that lead to different ecosystem states.

4.2 Enhancing Prediction: the Baltic Sea Index

Footing on the results of the study on macrozoobenthos off the island of Norderney in the North Sea the use of a multivariate predictor for the study on Baltic Sea ecosystem time series is certainly of advantage.

Therefore, the Baltic Sea environmental (BSE) index is developed. It aims

at better predictions of inter-annual and inter-decadal variability in environmental variables of the Baltic Sea. The index consists of four time series: the AO, the salinity between 120 and 200 m in the Gotland Sea, the integrated river runoff of all rivers draining into the Baltic Sea, and the relative vorticity of geostrophic wind over the Baltic Sea area. Each time series represents a specific forcing to the Baltic Sea in time and space and all relevant physical processes which are responsible for the forcing of inter-annual and inter-decadal variability of the Baltic Sea are incorporated. The AO index represents the northern hemisphere climate variability of both, the North Atlantic and the North Pacific. Therefore, the AO index is performing better than the NAO which considers only the climate variability of the North Atlantic. This is supported by the results of wavelet coherence of the AO and NAO with the physical parameters Landsort gauge, Gotland Sea SST and ice extent.

Major Baltic inflows strongly influence the salinity in the deeper layers of the Gotland Sea. The counterpart of major Baltic inflows is the hydrological cycle which is influencing the surface salinity. The effect of large-scale atmospheric blocking situations is considered by using the vorticity from the Chen index (Chen 2000). In the case of blocking, westerly winds disappear and meridional winds dominate. In such a case, the AO index and the NAO index do not contribute to the northern hemisphere climate variability. Therefore, the best way to consider the contribution of climate variability in case of blocking is the use of the relative vorticity of the geostrophic wind (Chen 2000). Each of the four considered time series contributes to the inter-annual and inter-decadal variability of the physical and biological system of the Baltic Sea.

Two unexpected results have been found: first, no meaningful combination has been found if Atlantic multi-decadal oscillation is considered as climate predictor for the physical response variables. This is rather surprising because various authors (e.g. Knight et al. 2006) have shown a strong correlation of Atlantic multi-decadal oscillation with the air temperature in central England and the precipitation over the Baltic Sea catchment area. This finding needs further investigations.

Second, correlation coefficients and skills indicate that a regional index like the Baltic Sea index or a multivariate regional index such as the BSE have a much better performance than a large-scale hemispheric index. This result

is a clear contradiction to the prediction paradox of Hallett et al. (2004). The prediction paradox says that large-scale climate indices seem to be better predictors of ecological processes than local climate. This might be perhaps the case in terrestrial ecosystems considered by Hallett et al. (2004). But, our results indicated that this is surely not the case in marine areas such as the North Sea (Dippner et al. 2010) and the Baltic Sea (Möllmann et al. 2009). Our results indicate that the multivariate BSE index combined of large- and regional-scale indices has an excellent performance and a high versatility much better than each single large-scale index only. Especially, the regime shift at the end of the 1980s and the extreme cold winters 1978/79 and 1995/96 are well reproduced.

Limitation

The development of the BSE index has some weak points. The most important is the shortness of the integrated river runoff time series. Unfortunately, this short time series also dominates the first EOF, and tests without it showed a clear performance decrease. Therefore, results of any downscaling experiments going beyond that year have to be interpreted with caution. The advantage of using a longer time series has to be weighed carefully against the disadvantage of lower interpretability of the results, and the results should always be compared with results of the shorter time series. This limitation to periods until 2002 also implies that it is not possible to study the regime shift in 2001/2002, which was the initial reason for the development of the BSE. However, the update of the integrated river runoff time series is currently under debate and there is hope that there will be an updated version in the near future.

4.3 Predictability of Phytoplankton

The overall result of the downscaling experiments with phytoplankton is, that although phytoplankton does have some potential predictability, it is very weak. While for most response variables with sufficient data a time series could be fitted for some station to observations that the variance of the error is lower than the variance of the observations, the fit is still insignificant and the overall results for all stations are poor, as the model skill judged by comparison of the

EOF time coefficients showed. Thus, considering all results with a significance level lower than 95% as insignificant, it can be summarized that predictions of phytoplankton, and even the fitting of phytoplankton time series, fail for the Arkona Sea as well as for the LUNG data.

While a successful model fit and prediction would have allowed for estimation of the extent to which phytoplankton might be predictable, the failure does not verify unpredictability of the system as such. Reasons for the low model skill are manifold but can be attributed to either the systems inherent dynamics or data issues.

The system itself may be governed by highly non-linear processes and interactions with the drivers that diminish predictive skill. An example are highly non-linear discontinuous regime shifts, where the ecosystem is in an unpredictable and sometimes chaotic state. But even for the weaker non-linear dynamics prediction skill decreases if the chosen statistical method is linear. It might be too much simplification to see the ecosystem as merely tracking the environment. It has been shown that the ecosystems processes amplify the stochastic physical forcing in a non-linear way (Hsieh et al. 2005). If the involved processes are so highly non-linear that they do not allow for linearisation of the relationship with a predictor by using e.g. a multivariate index like the BSE index, the statistical model must be non-linear too. It is not possible to resolve non-linear relationships with a simple linear statistical models. Further, spatial variability of a response variable might be so high, that small-scale local factors dominate, which is also reducing predictive skill. Time series length, sampling rate and error of course also influence predictive skill. An insufficient sampling rate might lead to aliasing, leading to false dynamical structures, or the long gaps between sampling might lead to simple missing of dynamical structures. Sampling strategy is also affecting skill if the strategy and statistical methods and requisites do not match. Further are sampling errors increasing noise but might also introduce some constant error or even erroneous shifts, sometimes seen in data as the "lab-effect" when e.g. the personnel in the counting laboratory changes.

The next paragraphs will see to which of the mentioned possible reasons apply in the studied case here.

4.3.1 System Dynamics

In theory the part of unpredictability that can be attributed can be found by looking at the speed at which prediction skill decreases with length of prediction time. The constant part of modelling error then belongs to noise. However, the overall skill of fit is already very poor and not significant so, does that mean that the data is just noise, or that phytoplankton dynamics are chaotic?

There are also other indicators for predictability of a dynamical system and the degree of non-linearity, e.g. by looking at its trajectories in phase space, the Poincaré map and autocorrelation function. For an unpredictable chaotic system, there are no clear structures present in phase space or Poincaré map and the autocorrelation function decays rapidly. However, also here do the short and sparse time series not allow to see any structures at all. The autocorrelation function decreases in fact so quickly, that it drops below the 95% significance mark already before lag 1. Figure 4.1 shows this for summer Cyanobacteria abundance.

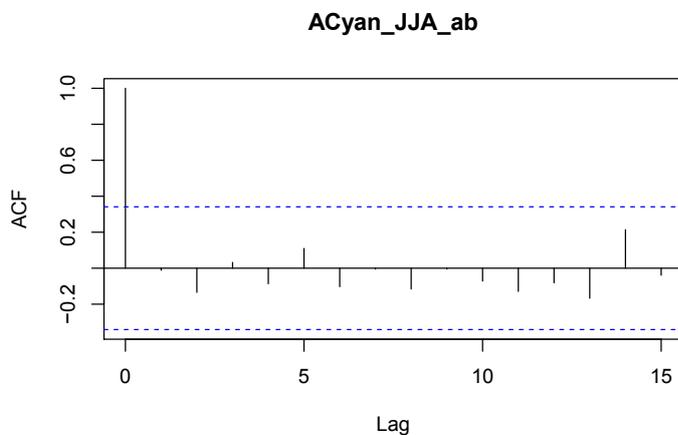


Figure 4.1: Autocorrelation function for summer Cyanobacteria abundance. Blue dashed line marks the 95% significance

The number of EOFs needed to characterise a system can also be seen as an indicator for the complexity of a system. The idea behind that is that with increasing complexity of a system also the number of needed EOFs to represent the systems dynamics increases. This can be seen in the amount of explained variance of the EOF patterns. In this study, the common picture

for most phytoplankton groups is that the 1st EOF accounts for more than half of the variance, the 2nd EOF adds another $\sim 20\%$, the 3rd a little less than 20% and the last has often $\sim 10\%$. Although it might seem quite a clear signal if only two EOFs are needed to explain more than $\sim 70\%$ variance, it has to be seen in relation to the number of stations. In other words, half of the possible EOF are needed to explain $\sim 70\%$ of the variance. The relatively low difference in the 2nd and 3rd EOF regarding explained variance and the still 10% in the last EOF are typical for systems with a notable amount of noise. If a system consists of only white noise, the amount of explained variance is the same in all EOF. Additionally, it is necessary to keep in mind that some stations discontinue, which distorts the pattern and overweights the signal of the station that is continued. So, the signal is not so clear after all and noise is certainly a factor decreasing model skill significantly.

Choice of Predictor

A reason for low predictive skill might be the use of the wrong predictor. It could be shown that the correlation with NAO stopped around 2000 (Chap. 3.2). In the experiments with zooplankton and physical parameters, the NAO index and the other considered predictors were clearly outperformed by the BSE index. In the experiment with phytoplankton, the first step in the downscaling procedure, where the best predictor/predictand combinations are searched for, has also shown that it gives the best overall performance. Yet, usage of the BSE index increases prediction skill, but it nevertheless does not allow for significant predictions of phytoplankton.

Because of the results of linear downscaling, a data mining experiment was performed in search of a better predictor using multivariate regression trees (Michelangeli et al. 1995; De'ath and Fabricius 2000) with a wide selection of seasonally averaged predictors ranging from locally measured abiotic factors such as nutrients, temperature and salinity to the well known climate indices. It was used "out-of-the-box", i.e. with mainly standard values as suggested by the R-software package "mvpart" (De'ath 2007). The results of this usually robust non-linear and non-parametric method were just as poor as the results from linear downscaling - no set of predictors were found that would allow for predictions of any phytoplankton group with a better Brier-based score

than 0.1. In fact, almost all scores were negative, the only positive score was found with the AO as predictor. The study was then discontinued due to the discouraging results. Therefore no results are shown here. However, the failure of this model experiment can be interpreted and it also points towards high levels of noise in the data.

4.3.2 Data Issues

Despite the uncertainties of the results of the previous section regarding the source of low model skill and degree of non-linearity, data issues account obviously for a substantial amount of error. The data issues can be separated into various possible sources of error which are discussed in the next paragraphs.

Spatial variability

Each station in the LUNG data set has distinct geographical properties that separates it clearly from the other stations, especially those of the near coastal zone. Therefore the reason for failure of prediction is here the high spatial heterogeneity of the area. Here, local factors dominate the ecosystem. Additionally the data set shows high temporal variability in the data set accompanied by a large number of absences. Also this can be attributed to local factors, such as the plume of the Oder river, which is not stationary. The high number of absences (zeros) further causes statistical problems since they cause a pronounced bimodal distribution. Statistical downscaling requires gaussian distribution, therefore the LUNG data set is excluded from the downscaling experiments.

Spatial variability as reason for low predictive skill is less obvious for the Arkona Sea data, since the geographical differences in station characteristics like depth, distance to land and river mouths and salinity are much lower. The stations are also all in the same basin, so that differences in basin characteristics do not play a role. The higher resemblance of stations than in the LUNG data set is indeed reflected in EOF analysis, however not as clearly as might be expected. The 1st EOF of the Arkona Sea data explains more variance than the corresponding number of EOFs (2 EOF for groups A and B, 1 EOF for C, 1-2 EOF for D) in the LUNG data set but the variance is still quite high in the

other EOF, which indicates high spatial variability is also in the Arkona data set. Possible reason to that is the non-uniform distribution of phytoplankton. If the station happens to lie in a plankton cloud, the abundance is overestimated. Satellite images of the cyanobacteria bloom in the Baltic Sea visualize this problem exceptionally clear.

Another source for spatial heterogeneity in the data might be the different sampling times at the different stations. Phytoplankton blooms on a time scale of days, if the sampling at the stations lie some days apart, might this introduce false spatial variability even if the phytoplankton blooms at the stations were perfectly synchronous.

Sampling issues

The above mentioned spatial heterogeneity in the data due to sampling is showing the importance of sampling strategy and timing on the interpretability of the data set.

Available ship time and financial restrictions make it necessary to reduce the number of field measurements to a minimum. Sampling is giving a snap-shot of phytoplankton abundance and biomass at a given time. The sampling strategy, i.e. the decision on sampling time, frequency and location, is determining the usability of these snap-shots for the study of the dynamics of the ecosystem. While the timely close sampling of several stations in the Baltic Sea during one campaign may determine the state of the Baltic Sea at time of sampling to a sufficient accuracy for the communication to the public, several of these snapshots do not necessarily allow to draw conclusions on the dynamics of the ecosystem. For that, the sampling frequency and time series length must match to resolve the systems dominating dynamic structures. It is impossible to resolve the systems inherent dynamics if a time series is too short or sampled too sparsely for the time scale of the dominating processes. That means for phytoplankton, if the sampling occurs with a monthly frequency always on the same date, this will allow for conclusions about the phytoplankton dynamics over the years for a specific time of year. But it will also inevitably lead to missing of phytoplankton peak abundances, that are governed by processes on a time scale of days, and thereby underestimate phytoplankton abundance.

The sampling frequency of the federal monitoring program conducted by the

IOW shows clearly the sparseness of the data. Figure 4.2 shows the total number of samples per month for the period 1994-2012 for the considered stations. It clearly shows a pattern where in some months no or only occasional sampling occurs, while others are overrepresented. Also the timing of the sampling has trends as figure 4.3 shows for station TF0113. However, statistical downscaling assumes stochastic data. This means for time series, that no patterns should be visible in the temporal distribution of the measurements. Resulting from the effort to describe the spring bloom as good as possible and to come the spring peak in abundance as close as possible, sampling is densest in spring, namely May. Much less data for the Arkona Sea is available in April since the standard sampling scheme does not include sampling then. This is probably also causing underestimation of phytoplankton abundance, since spring bloom occurs earlier in the Arkona Sea than in the rest of the Baltic Sea. February and March are also sampled more often than the summer and fall months. (The abundances are so low in February, that they are not included in the presented analysis and do not appear in the results.) To sample the summer cyanobacteria bloom, measurements are made in July and August, and the October and November are sampled again. Data for June, September, December and January are hardly available. The merely occasional sampling in April and the denser sampling in May, introduce errors in the seasonal means and make it necessary to analyse seasonal averages alone. Besides the problems that arise from variable sampling frequency and density, also a trend in the sampling times introduce errors in the data set.

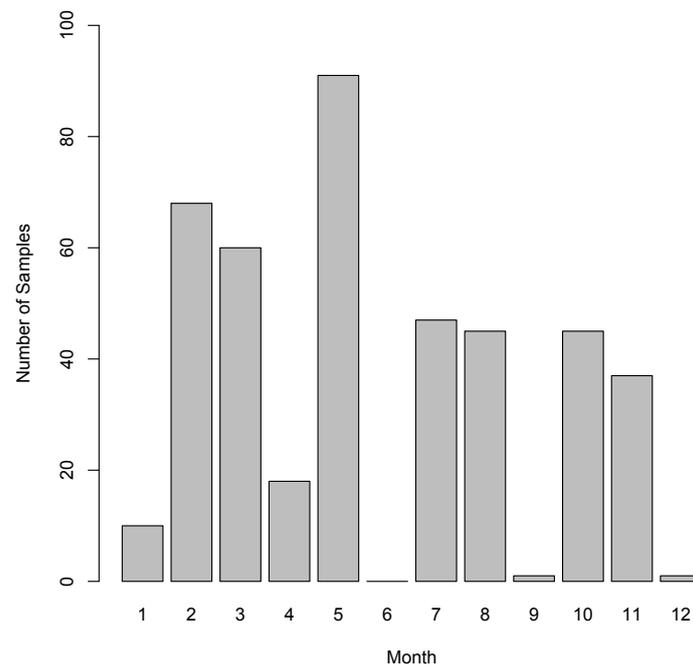


Figure 4.2: Sum of the number of phytoplankton samples at the 4 considered stations of the Arkona Sea per month. The basis is a complete extract of the database of the IOW for the 4 Stations and all phytoplankton. Time span: 1994-2012

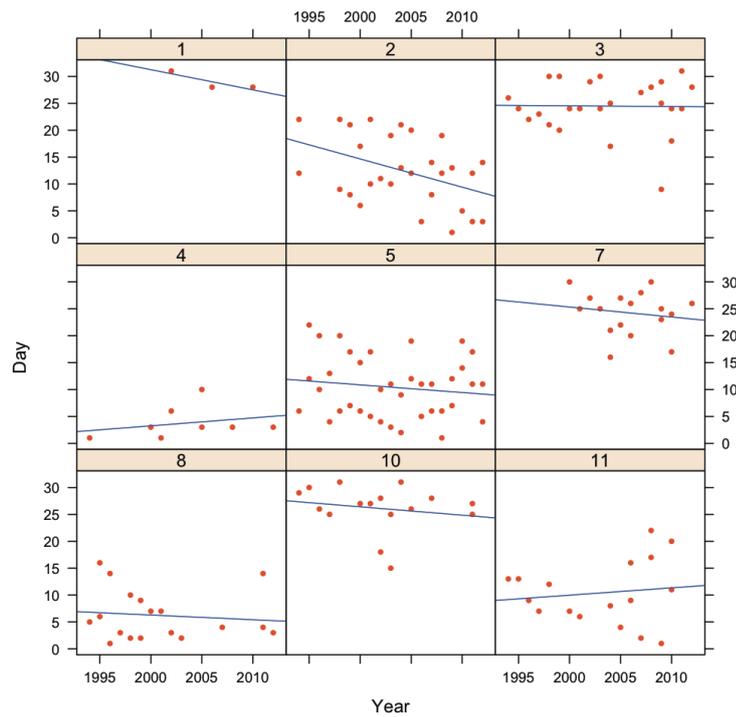


Figure 4.3: Sampling days of each month for station TF0113 (red dots) and a linear regression line (blue line). Only the linear fit for February is significant (95%). The basis is a complete extract of the database of the IOW for the Stations in the Arkona Sea and all phytoplankton. Months with no sampling are excluded. Time span: 1994-2012

Chapter 5

Conclusion

This study has shown that predictability of ecosystems depends on several aspects, ranging from internal non-linearity of the system in question, to data issues such as length of the series, sampling frequency and quality.

The analysis of the macrozoobenthos time series demonstrated that large scale transitions in the atmosphere like the regime shift in 2001/2002 can trigger a biological regime shift which renders a previously predictable time series unpredictable due to high non-linear response of the system. This indicates the necessity to further study the mechanisms causing BRSs and the relationship of marine biota with the physical environment. A better predictor, covering more relevant aspects of the physical environment might enhance predictability even in presence of CRSs. Likewise might other, non-parametric, non-linear methods be more appropriate for prediction of biological variables, especially in presence of regime shifts.

The necessity of using a multivariate descriptor inspite of using non-linear methods, like in the prediction experiment with benthic macrofauna in the southern North Sea, clearly demonstrate the importance of local forcing on ecosystems. This is supported by the downscaling experiment of physical characteristics and zooplankton time series in the Baltic Sea, where the use of the BSE index alone could significantly enhance prediction skill. This result contradicts the prediction paradox of Hallett et al. (2004). According to this paradox, the large-scale climate indices are better predictors of ecological processes than local climate. Hallett et al. (2004) claimed this paradox for terrestrial ecosystems. For the marine environment, this study clearly shows

a better performance of an index consisting of large-scale and regional-scale indices. That the BSE index is performing best for physical as well as biological data doesn't only make it a versatile tool in the study and prediction of environmental properties in the Baltic Sea, but it also shows that it covers a wide range of forcing factors for the dynamics of the Baltic Sea ecosystem. This is supported by the good predictability despite the presence of a regime shift at the end of the 1980s and the extremely cold winters 1978/1979 and 1995/1996.

Besides showing the importance of local factors in the prediction of ecosystems, the study with OPELM and OPKNN also demonstrated the feasibility of neural networks for studying and predicting ecosystems. In order to mitigate the changes of an ecosystem, it is necessary to identify the dominating factors and to find out whether these are mainly climate or of anthropogenic origin. Linear methods fail to grasp the changes occurring under abrupt regime shifts even after transformation of input data, since these shifts increase the non-linearity of the system. Non-linear multivariate statistical downscaling methods have an advantage over linear methods in this case. However, the validation and interpretation of these results has to occur with great care, as over fitting is common and results sometimes misleading. Linear methods have a great advantage here. They have proven to be robust, requiring less data and computing power and to be sufficient and efficient for the purpose of identifying statistical relationships between systems. Thus, it is advisable not to use one method for all statistical modelling requirements and scientific problems, but use non-linear methods in combination with linear methods in order to avoid misinterpretation and to improve efficiency.

Although the prediction skill of macrozoobenthos from the southern North Sea, and zooplankton and physical response variables of the Baltic Sea were greatly enhanced by usage of a multivariate index, the model skill of phytoplankton data of the coastal area as well as the Arkona Sea in the southern Baltic Sea was not increased to a level that would allow for prediction. The EOF analysis and the examination of the sampling sites, strategy and frequency point towards data issues and not the degree of non-linearity as the most probable source for low model skill. The sampling frequency does not allow to adequately characterize the phytoplankton development over time. A

denser and more continuous sampling is necessary to get a better representation of abundance peaks.

In order to be successfully manage and conserve marine ecosystems it is necessary to characterize and predict their current respective future status and dynamics. If it is not possible to estimate the result of management action, these actions are at risk to fail. Due to the many uncertainties associated with the prediction of the ecosystem, it is today often necessary to resort to the cautionary principle. Studies of predictability like this help to characterize relationships of the ecosystem with the physical environment to enhance prediction, which will help to optimally manage the marine environment.

Bibliography

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- Alheit, J., Bakun, A. (2010). Population synchronies within and between ocean basins: Apparent teleconnections and implications as to physical-biological linkage mechanisms. *Journal of Marine Systems*, 79(3-4):267–285. doi:doi:10.1016/j.jmarsys.2008.11.029.
- Alheit, J., Möllmann, C., Dutz, J., Kornilovs, G., Loewe, P., Mohrholz, V., Wasmund, N. (2005). Synchronous ecological regime shifts in the central Baltic and the North Sea in the late 1980s. *ICES Journal of Marine Science: Journal du Conseil*, 62(7):1205. doi:doi:10.1016/j.icesjms.2005.04.024.
- BACC Author Team (2008). *Assessment of climate change for the Baltic Sea basin*. Springer Verlag.
- Bakun, A. (2005). Regime shifts. In A. R. Robinson (ed.), *The global coastal ocean : multiscale interdisciplinary processes*, vol. 13 of *The Sea*. Harvard Univ. Press, Cambridge, Mass.
- Beaugrand, G. (2004). The North Sea regime shift: evidence, causes, mechanisms and consequences. *Progress in Oceanography*, 60(2-4):245–262. doi:10.1016/j.pocean.2004.02.018.
- Becker, G., Frey, H., Wegner, G. (1986). *Atlas der Temperatur an der Oberfläche der Nordsee: Wöchentliche und monatliche Mittelwerte für den Zeitraum 1971 bis 1980*. No. 17 in Dtsch. Hydrogr. Z., Ergebnis-Hefte. Deutsches Hydrographisches Institut.
- Becker, G. A., Pauly, M. (1996). Sea surface temperature changes in the North Sea and their causes. *ICES J. Mar. Sci.*, 53(6):887–898. doi:10.1006/jmsc.1996.0111.

- Bergström, S., Carlsson, B. (1994). River runoff to the Baltic Sea: 1950-1990. *Ambio. Stockholm*, 23(4):280–287.
- Bond, N. A., Overland, J. E., Spillane, M., Stabeno, P. (2003). Recent shifts in the state of the north pacific. *Geophysical Research Letters*, 30(23). doi:10.1029/2003GL018597.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85(7):1771–1789. Doi: 10.1890/03-9000.
- Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., Coloso, J., Hodgson, J. R., Kitchell, J. F., Seekell, D. A., Smith, L., Weidel, B. (2011). Early warnings of regime shifts: A whole-ecosystem experiment. *Science*, 332(6033):1079–1082. doi:10.1126/science.1203672.
- Casini, M., Bartolino, V., Molinero, J. C., Kornilovs, G. (2010). Linking fisheries, trophic interactions and climate: threshold dynamics drive herring clupea harengus growth in the central Baltic Sea. *Marine Ecology-Progress Series*, 413:241–252. doi:10.3354/meps08592.
- Casini, M., Hjelm, J., Molinero, J., Lövgren, J., Cardinale, M., Bartolino, V., Belgrano, A., Kornilovs, G. (2009). Trophic cascades promote threshold-like shifts in pelagic marine ecosystems. *Proceedings of the National Academy of Sciences*, 106(1):197.
- Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J. O., Jenouvrier, S., Stenseth, N. C. (2008). Wavelet analysis of ecological time series. *Oecologia*, 156(2):287–304. doi:10.1007/s00442-008-0993-2.
- Cazelles, B., Hales, S. (2006). Infectious diseases, climate influences, and nonstationarity. *PLoS Med*, 3(8):e328. doi:10.1371/journal.pmed.0030328.
- Chen, D. (2000). A monthly circulation climatology for Sweden and its application to a winter temperature case study. *International Journal of Climatology*, 20(10):1067–1076. doi:10.1002/1097-0088(200008)20:10<1067::AID-JOC528>3.0.CO;2-Q.
- Chen, D., Hellström, C. (1999). The influence of the north atlantic oscillation on the regional temperature variability in sweden: spatial and temporal variations. *Tellus A*, 51(4):505–516. doi:10.1034/j.1600-0870.1999.t01-4-00004.x.
- Collie, J. S., Richardson, K., Steele, J. H. (2004). Regime shifts: Can ecolog-

- ical theory illuminate the mechanisms? *Progress in Oceanography*, 60(2-4):281–302. doi:10.1016/j.pocean.2004.02.013.
- Daubechies, I. (1992). *Ten lectures on wavelets*. Society for Industrial Mathematics.
- De'ath, G. (2007). *mvpart: Multivariate partitioning, R package version 1.2-6*.
- De'ath, G., Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, 81(11):3178–3192. doi:10.1890/0012-9658(2000)081[3178:cartap]2.0.co;2.
- Deser, C. (2000). On the teleconnectivity of the "Arctic Oscillation". *Geophys. Res. Lett.*, 27(6):779–782. doi:10.1029/1999GL010945.
- Dima, M., Lohmann, G. (2007). A hemispheric mechanism for the Atlantic Multidecadal Oscillation. *Journal of Climate*, 20:2706–2719. doi:10.1175/JCLI4174.1.
- Dippner, J. W. (1997a). Recruitment success of different fish stocks in the North Sea in relation to climate variability. *Ocean Dynamics*, 49(2):277–293. doi:10.1007/BF02764039.
- Dippner, J. W. (1997b). SST anomalies in the north sea in relation to the North Atlantic Oscillation and the influence on the theoretical spawning time of fish. *Ocean Dynamics*, 49(2):267–275. doi:10.1007/BF02764038.
- Dippner, J. W., Hänninen, J., Kuosa, H., Vuorinen, I. (2001). The influence of climate variability on zooplankton abundance in the Northern Baltic Archipelago Sea (SW Finland). *ICES Journal of Marine Science*, 58(3):569–578. doi:10.1006/jmsc.2001.1048.
- Dippner, J. W., Ikauniece, A. (2001). Long-term zoobenthos variability in the Gulf of Riga in relation to climate variability. *Journal of Marine Systems*, 30(3-4):155–164. doi:10.1016/S0924-7963(01)00055-0.
- Dippner, J. W., Junker, K., Kröncke, I. (2010). Biological regime shifts and changes in predictability. *Geophys. Res. Lett.*, 37(24):L24701. doi:10.1029/2010gl045696.
- Dippner, J. W., Kornilovs, G., Junker, K. (2012). A multivariate Baltic Sea environmental index. *AMBIO*, 41:699–708. doi:10.1007/s13280-012-0260-y.
- Dippner, J. W., Kornilovs, G., Sidrevics, L. (2000). Long-term variability of mesozooplankton in the Central Baltic Sea. *Journal of Marine Systems*, 25(1):23–31. doi:10.1016/S0924-7963(00)00006-3.

- Dippner, J. W., Kröncke, I. (2003). Forecast of climate-induced change in macrozoobenthos in the southern North Sea in spring. *Climate Research*, 25(2):179–182. doi:10.3354/cr025179.
- Dippner, J. W., Ottersen, G. (2001). Cod and climate variability in the Barents Sea. *Climate Research*, 17(1):73–82.
- Donat, M. G., Renggli, D., Wild, S., Alexander, L. V., Leckebusch, G. C., Ulbrich, U. (2011). Reanalysis suggests long-term upward trends in european storminess since 1871. *Geophysical Research Letters*, 38. doi:10.1029/2011gl047995.
- Drinkwater, K. F. (2006). The regime shift of the 1920s and 1930s in the north atlantic. *Progress in Oceanography*, 68(2-4):134–151. doi:10.1016/j.pocean.2006.02.011.
- Duarte, C., Conley, D., Carstensen, J., Sánchez-Camacho, M. (2009). Return to neverland: shifting baselines affect eutrophication restoration targets. *Estuaries and Coasts*, 32(1):29–36. doi:10.1007/s12237-008-9111-2.
- Edwards, M., Richardson, A. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, 430(7002):881–884.
- EEA (2012). Climate change, impacts and vulnerability in Europe 2012: An indicator-based report. Tech. Rep. 12/2012.
- Efron, B., Hastie, T., Johnstone, I., Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, 32(2):407–499. doi:10.1214/009053604000000067.
- Enfield, D., Mestas-Nuñez, A., Trimble, P. (2001). The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental US. *Geophys. Res. Lett*, 28:2077–2080. doi:10.1029/2000GL012745.
- Feistel, R., Nausch, G., Wasmund, N. (2008). *State and Evolution of the Baltic Sea 1952-2005 – A detailed 50 year survey of meteorology and climate, physics, chemistry biology, and marine environment*. Wiley-Interscience, John Wiley & Sons Inc., Hoboken NJ.
- Fogarty, M. J., Murawski, S. A. (1998). Large-scale disturbance and the structure of marine systems: Fishery impacts on georges bank. *Ecological Applications*, 8:6–22.
- Gerlach, S. (1994). Oxygen conditions improve when the salinity in the Baltic Sea decreases. *Marine Pollution Bulletin*, 28(7):413–416. doi:10.1016/

0025-326X(94)90126-0.

- Graham, N. (1994). Decadal-scale climate variability in the tropical and North Pacific during the 1970s and 1980s: observations and model results. *Clim. Dyn.*, 10:135–162.
- Grinsted, A., Moore, J., Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11(5/6):561–566.
- Hagen, E., Feistel, R. (2005). Climatic turning points and regime shifts in the Baltic Sea region: the baltic winter index (WIBIX) 1659-2002. *Boreal Environment Research*, 10(3):211–224.
- Hallett, T. B., Coulson, T., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M., Grenfell, B. T. (2004). Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature*, 430(6995):71–75. doi:10.1038/nature02708.
- Hare, S. R., Mantua, N. J. (2000). Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography*, 47(2-4):103–145. doi:10.1016/S0079-6611(00)00033-1.
- Hasselmann, K. (1988). PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, 93(11):11015–11021.
- HELCOM (2008). Manual for marine monitoring in the COMBINE programme of HELCOM.
- HELCOM (2009). Hazardous substances of specific concern to the Baltic Sea - final report of the HAZARDOUS project. *HELCOM Balt. Sea Environ. Proc.*, 119.
- Hernroth, L. (ed.) (1985). *Recommendations on methods for marine biological studies in the Baltic Sea: Mesozooplankton Biomass Assessment*. 10.1034/j.1600-0870.1998.00012.x. Baltic Marine Biologists.
- Heyen, H., Dippner, J. W. (1998). Salinity variability in the german bight in relation to climate variability. *Tellus A*, 50(4):545–556.
- Heyen, H., Zorita, E., Von Storch, H. (1996). Statistical downscaling of monthly mean north atlantic air-pressure to sea level anomalies in the Baltic Sea. *Tellus A*, 48(2):312–323.
- Hoerling, M. P., Hurrell, J. W., Xu, T. (2001). Tropical origins for recent

- north atlantic climate change. *Science*, 292(5514):90–92. doi:10.1126/science.1058582.
- Hsieh, C.-h., Glaser, S. M., Lucas, A. J., Sugihara, G. (2005). Distinguishing random environmental fluctuations from ecological catastrophes for the north pacific ocean. *Nature*, 435(7040):336–340. doi:10.1038/nature03553.
- Huang, G., Zhu, Q., Siew, C. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3):489–501. doi:10.1016/j.neucom.2005.12.126.
- Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, 269(5224):676–679. doi:10.1126/science.269.5224.676.
- Hurrell, J. W., Deser, C. (2010). North Atlantic climate variability: The role of the North Atlantic Oscillation. *Journal of Marine Systems*, 79(3-4):231–244. doi:10.1016/j.jmarsys.2009.11.002.
- Hurvich, C., Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297.
- Huth, R. (1999). Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Climate Research*, 13:91–101.
- IPCC (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, and New York, NY, USA.
- Jaagus, J., Post, P., Tomingas, O. (2008). Changes in storminess on the western coast of estonia in relation to large-scale atmospheric circulation. *Climate Research*, 36(1):29–40. doi:10.3354/cr00725.
- Jevrejeva, S., Moore, J., Grinsted, A. (2003). Influence of the Arctic Oscillation and El Nino-Southern Oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach. *J. Geophys. Res*, 108(D21):4677. doi:10.1029/2003JD003417.
- Jung, T., Hilmer, M., Ruprecht, E., Kleppek, S., Gulev, S. K., Zolina, O. (2003). Characteristics of the recent eastward shift of interannual nao variability. *Journal of Climate*, 16(20):3371–3382. doi:10.1175/1520-0442(2003)016<3371:COTRES>2.0.CO;2.
- Junker, K., Dippner, J. W. (2011). A new Baltic Sea environmental index.

- Poster, ASLO Aquatic Sciences Confence, Puerto Rico.
- Junker, K., Sovilj, D., Kröncke, I., Dippner, J. W. (2012). Climate induced changes in benthic macrofauna - a non-linear model approach. *J. Mar. Syst.*, 96-97:90–94. doi:10.1016/j.jmarsys.2012.02.005.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471.
- Knight, J. R., Folland, C. K., Scaife, A. A. (2006). Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, 33(17):L17706. doi: 10.1029/2006gl026242.
- Kröncke, I., Dippner, J., Heyen, H., Zeiss, B. (1998). Long-term changes in macrofaunal communities off Norderney (East Frisia, Germany) in relation to climate variability. *Marine Ecology Progress Series*, 167:25–36.
- Kröncke, I., Reiss, H., Dippner, J. (2013). Effects of cold winters and biological regime shifts on the Norderney macrofauna communities in the southern North Sea. *Estuarine, Coastal and Shelf Science*, 119:79–90. doi:10.1016/j.ecss.2012.12.024.
- Kullback, S., Leibler, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.
- Lamb, H. H. (1950). Types and spells of weather around the year in the British Isles : Annual trends, seasonal structure of the year, singularities. *Quarterly Journal of the Royal Meteorological Society*, 76(330):393–429.
- Landesamt für Umwelt, Naturschutz und Geologie Mecklenburg-Vorpommern (2006). Gewässergütebericht Mecklenburg-Vorpommern 2003/2004/2005/2006: Ergebnisse der Güteüberwachung der Fließ-, Stand- und Küstengewässer und des Grundwassers in Mecklenburg-Vorpommern. Tech. Rep..
- Lau, K.-M., Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. *Bulletin of the American Meteorological Society*, 76(12):2391–2402. doi:10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.
- Lehmann, A., Getzlaff, K., Harlaß, J. (2011). Detailed assessment of climate variability in the Baltic Sea area for the period 1958 to 2009. *Climate*

- Research*, 46(2):185–196. doi:10.3354/cr00876.
- Lehmann, A., Krauss, W., Hinrichsen, H. H. (2002). Effects of remote and local atmospheric forcing on circulation and upwelling in the Baltic Sea. *Tellus A*, 54(3):299–316. doi:10.1034/j.1600-0870.2002.00289.x.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S. (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling*, 90(1):39–52. doi:10.1016/0304-3800(95)00142-5.
- Lek, S., Guégan, J. F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120(2-3):65–73. doi:10.1016/S0304-3800(99)00092-7.
- Levin, S. A. (1999). *Fragile dominion - Complexity and the commons*. Perseus Books.
- Lindgren, M., Möllmann, C., Nielsen, A., Brander, K., MacKenzie, B. R., Stenseth, N. C. (2010). Ecological forecasting under climate change: the case of Baltic cod. *Proceedings of the Royal Society B: Biological Sciences*. doi:10.1098/rspb.2010.0353.
- Livezey, R. (1995). The evaluation of forecast. In H. von Storch, A. Navarra (eds.), *Analysis of Climate Variability*, pp. 177–196. Springer-Verlag, Berlin.
- Loewe, P. (2009). Atmosphärenphysik. In *System Nordsee*. Bundesamt für Seeschifffahrt und Hydrographie.
- MacKenzie, B., Schiedek, D. (2007). Long-term sea surface temperature baselines - time series, spatial covariation and implications for biological processes. *Journal of Marine Systems*, 68(3-4):405–420. doi:10.1016/j.jmarsys.2007.01.003.
- Mantua, N. (2004). Methods for detecting regime shifts in large marine ecosystems: a review with approaches applied to north pacific data. *Progress in Oceanography*, 60(2-4):165–182. doi:10.1016/j.pocean.2004.02.016.
- Mantua, N., Hare, S., Zhang, Y., Wallace, J., Francis, R. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, 78:1069–1079.
- Matthäus, W., Franck, H. (1992). Characteristics of major baltic inflows—a statistical analysis. *Continental Shelf Research*, 12(12):1375–1400.

- Meier, H. E. M., Hordoir, R., Andersson, H. C., Dieterich, C., Eilola, K., Gustafsson, B. G., Høglund, A., Schimanke, S. (2012). Modeling the combined impact of changing climate and changing nutrient loads on the Baltic Sea environment in an ensemble of transient simulations for 1961-2099. *Climate Dynamics*, 39(9-10):2421–2441. doi:10.1007/s00382-012-1339-7.
- Mi, X., Ren, H., Ouyang, Z., Wei, W., Ma, K. (2005). The use of the Mexican Hat and the Morlet wavelets for detection of ecological patterns. *Plant Ecology*, 179(1):1–19. doi:10.1007/s11258-004-5089-4.
- Michaelsen, J. (1987). Cross-validation in statistical climate forecast models. *Journal of Applied Meteorology*, 26(11):1589–1600.
- Miche, Y., Sorjamaa, A., Lendasse, A. (2008). OP-ELM: Theory, experiments and a toolbox. In V. Kurková, R. Neruda, J. Koutník (eds.), *Artificial Neural Networks - ICANN 2008*, vol. 5163 of *Lecture Notes in Computer Science*, pp. 145–154. Springer Berlin / Heidelberg. doi:10.1007/978-3-540-87536-9_16.
- Michelangeli, P.-A., Vautard, R., Legras, B. (1995). Weather regimes: Recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, 52(8):1237–1256.
- Mikulski, Z. (1982). River inflow to the Baltic Sea 1921-1975. *Polish Academy of Science/Polish National Committee of the IHP UNESCO, Warsaw*.
- Minns, A., Hall, M. (1996). Artificial neural networks as rainfall-runoff models. *Hydrological Sciences Journal*, 41(3):399–417.
- Möllmann, C., Blenckner, T., Casini, M., Gardmark, A., Lindegren, M. (2011). Beauty is in the eye of the beholder: management of Baltic cod stock requires an ecosystem approach. *Marine Ecology Progress Series*, 431:293–297. doi:10.3354/meps09205.
- Möllmann, C., Diekmann, R., Müller-Karulis, B., Kornilovs, G., Plikshs, M., Axe, P. (2009). Reorganization of a large marine ecosystem due to atmospheric and anthropogenic pressure: a discontinuous regime shift in the Central Baltic Sea. *Global Change Biology*, 15(6):1377–1393. doi:10.1111/j.1365-2486.2008.01814.x.
- Möllmann, C., Kornilovs, G., Fetter, M., Köster, F. W., Hinrichsen, H. H. (2003). The marine copepod *Pseudocalanus elongatus* as a mediator between climate variability and fisheries in the Central Baltic Sea. *Fisheries*

- Oceanography*, 12(4/5):360. doi:10.1046/j.1365-2419.2003.00257.x.
- Omstedt, A., Elken, J., Lehmann, A., Piechura, J. (2004). Knowledge of the Baltic Sea physics gained during the BALTEX and related programmes. *Progress in Oceanography*, 63(1-2):1–28. doi:10.1016/j.pocean.2004.09.001.
- Peterson, W. T., Schwing, F. B. (2003a). A new climate regime in the Northeast Pacific ecosystems. *Geophys. Res. Lett.*, 30(17):1896. doi:10.1029/2003GL017528.
- Peterson, W. T., Schwing, F. B. (2003b). A new climate regime in the Northeast Pacific ecosystems. *Geophys. Res. Lett.*, 30(17):1896. doi:10.1029/2003GL017528.
- Preisendorfer, R. W., Mobley, C. D. (1988). *Principal component analysis in meteorology and oceanography*. Developments in atmospheric science ; 17. Elsevier, Amsterdam [u.a.].
- Reid, P. C., Borges, M. d. F., Svendsen, E. (2001). A regime shift in the North Sea circa 1988 linked to changes in the North Sea horse mackerel fishery. *Fisheries Research*, 50(1-2):163–171. doi:10.1016/S0165-7836(00)00249-6.
- Reid, P. C., Edwards, M., Hunt, H. G., Warner, A. J. (1998). Phytoplankton change in the North Atlantic. *Nature*, 391(6667):546–546. doi:10.1038/35290.
- Remane, A., Schlieper, C. (1971). *Biology of brackish water*, vol. 372. E. Schweizerbart'sche Verlagsbuchhandlung.
- Rott, E. (1981). Primary productivity and activity coefficients of the phytoplankton of a mesotrophic soft-water lake (Piburger See, Tirol, Austria). *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, 66(1):1–27.
- Rudnick, D., Davis, R. (2003). Red noise and regime shifts. *Deep-Sea Research Part I*, 50(6):691–699. doi:10.1016/S0967-0637(03)00053-0.
- Sagert, S., Rieling, T., Eggert, A., Schubert, H. (2008). Development of a phytoplankton indicator system for the ecological assessment of brackish coastal waters (German Baltic Sea coast). *Hydrobiologia*, 611(1):91–103. doi:10.1007/s10750-008-9456-3.
- Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596. doi:10.1038/35098000.

- Scheffer, M., Carpenter, S. R. (2003). Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in Ecology & Evolution*, 18(12):648–656. doi:10.1016/j.tree.2003.09.002.
- Skogen, M. D., Drinkwater, K., Hjollo, S. S., Schrum, C. (2011). North Sea sensitivity to atmospheric forcing. *Journal of Marine Systems*, 85(3-4):106–114. doi:10.1016/j.jmarsys.2010.12.008.
- Sommer, U., Adrian, R., Bauer, B., Winder, M. (2012). The response of temperate aquatic ecosystems to global warming: novel insights from a multidisciplinary project. *Marine Biology*, 159(11):2367–2377. doi:10.1007/s00227-012-2085-4.
- Sovilj, D., Sorjamaa, A., Yu, Q., Miche, Y., Séverin, E. (2010). OPELM and OPKNN in long-term prediction of time series using projected input data. *Neurocomputing*, 73(10-12):1976–1986. doi:10.1016/j.neucom.2009.11.033.
- Stenseth, N., Ottersen, G., Hurrell, J., Belgrano, A. (2004). *Marine ecosystems and climate variation: the North Atlantic: a comparative perspective*. Oxford University Press, USA.
- Stenseth, N. C., Mysterud, A. (2002). Climate, changing phenology, and other life history traits: nonlinearity and match-mismatch to the environment. *Proc Natl Acad Sci U S A*, 99(21):13379–13381. doi:10.1073/pnas.212519399.
- Stenseth, N. C., Mysterud, A., Ottersen, G., Hurrell, J. W. (2002). Ecological effects of climate fluctuations. *Science*, 297:1292–1296. doi:10.1126/science.1071281.
- Stige, L., Ottersen, G., Brander, K., Chan, K., Stenseth, N. (2006). Cod and climate: effect of the North Atlantic Oscillation on recruitment in the North Atlantic. *Marine Ecology Progress Series*, 325:227–241.
- Suarez, M., Schopf, P. (1988). A delayed action oscillator for ENSO. *Journal of the atmospheric Sciences*, 45(21):3283–3287.
- Sugihara, G., May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344(6268):734–741. 10.1038/344734a0.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, 7(1):13–26.

- Swanson, K. L., Tsonis, A. A. (2009). Has the climate recently shifted? *Geophys. Res. Lett.*, 36(6):L06711. doi:10.1029/2008gl037022.
- Taucher, J., Oeschies, A. (2011). Can we predict the direction of marine primary production change under global warming? *Geophys. Res. Lett.*, 38(2):L02603. doi:10.1029/2010gl045934.
- Thompson, D. W. J., Wallace, J. M. (1998). The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, 25(9):1297–1300.
- Torrence, C., Compo, G. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78.
- Torrence, C., Webster, P. J. (1999). Interdecadal changes in the ENSO-monsoon system. *Journal of Climate*, 12(8):2679–2690.
- Trenberth, K., Hurrell, J. (1994). Decadal atmospheric-ocean variations in the Pacific. *Clim. Dyn.*, 9:303–319.
- Tsonis, A., Elsner, J. (1992). Nonlinear prediction as a way of distinguishing chaos from random fractal sequences. *Nature*, 358:217–220.
- UNESCO (1968). *Zooplankton sampling*. United Nations Educational, Scientific and Cultural Organization (UNESCO).
- Utermöhl, H. (1958). Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *International Association of Theoretical and Applied Limnology Communications*, 9:1–38.
- van Beusekom, J., Bot, P., Carstensen, J., Goebel, J., Lenhart, H., Pätsch, J., Petenati, T., Raabe, T., Reise, K., Wetsteijn, B. (2009). Eutrophication. Thematic Report No. 6. In H. Marencic, J. d. Vlas (eds.), *Quality Status Report 2009*, Wadden Sea Ecosystem No. 25. Common Wadden Sea Secretariat, Trilateral Monitoring and Assessment Group, Wilhelmshaven, Germany.
- von Storch, H., Bruns, T., Fischer-Bruns, I., Hasselmann, K. (1988). Principal oscillation pattern analysis of the 30-to 60-day oscillation in a general circulation model equatorial troposphere. *J. Geophys. Res.*, 93(11):022–11.
- von Storch, H., Zorita, E., Cubasch, U. (1993). Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *Journal of Climate*, 6(6):1161–1171.
- Von Storch, H., Zwiers, F. (2001). *Statistical analysis in climate research*.

Cambridge Univ Press.

- Voss, M., Emeis, K., Hille, S., Neumann, T., Dippner, J. (2005). Nitrogen cycle of the Baltic Sea from an isotopic perspective. *Global Biogeochemical Cycles*, 19(3):GB3001. doi:10.1029/2004GB002338.
- Vuorinen, I., Hänninen, J., Rajasilta, M., Laine, P., Eklund, J., Montesino-Pouzols, F., Corona, F., Junker, K., Meier, H. M., Dippner, J. W. (2012). Horohalinicum revisited - alternative food webs due to zoogeographic migration. (under revision).
- Wales, D. J. (1991). Calculating the rate of loss of information from chaotic time series by forecasting. *Nature*, 350:485–488.
- Wang, G., Swanson, K. L., Tsonis, A. A. (2009). The pacemaker of major climate shifts. *Geophys. Res. Lett.*, 36:L07708. doi:10.1029/2008gl036874.
- Wasmund, N. (1994). Phytoplankton periodicity in a eutrophic coastal water of the Baltic Sea. *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, 79(2):259–285.
- Wilby, R., Wigley, T. (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4):530–548.
- Winder, M., Sommer, U. (2012). Phytoplankton response to a changing climate. *Hydrobiologia*, 698:5–16. doi:10.1007/s10750-012-1149-2.
- Winsor, P., Rodhe, J., Omstedt, A. (2001). Baltic Sea ocean climate: an analysis of 100 yr of hydrographic data with focus on the freshwater budget. erratum. *Clim Res*, 18:5–15.
- Winsor, P., Rodhe, J., Omstedt, A. (2003). Erratum: Baltic Sea ocean climate: an analysis of 100 yr of hydrographic data with focus on the freshwater budget. *Clim Res*, 25:183.
- Woodruff, S. D., Slutz, R. J., Jenne, R. L., Steurer, P. M. (1987). A comprehensive ocean-atmosphere data set. *Bulletin of the American Meteorological Society*, 68(10):1239–1250.
- Wulff, F., Rahm, L., Larsson, P. (2001). *A systems analysis of the Baltic Sea*. Springer Verlag.
- Yu, Q., Sorjamaa, A., Miche, Y., Lendasse, A., Sèverin, E., Guillen, A., Mateo, F. (2008). Optimal pruned K-nearest neighbors: OP-KNN application to financial modeling. In *HIS 2008, 8th International Conference on Hybrid*

Intelligent Systems, pp. 764–769. IEEE.

Zorita, E. (2009). Estimating climate change at regional scales. Lecture, Workshop on Time Series and Modelling of Environmental Data.

Zorita, E., Laine, A. (2000). Dependence of salinity and oxygen concentrations in the Baltic Sea on large-scale atmospheric circulation. *Climate Research*, 14(1):25–41.

Zorita, E., von Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12(8):2474–2489.

Appendix A

Abbreviations

AFDW	Ash-free dry weight
AIC	Akaike information criterion
AO	Arctic Oscillation
AMO	Atlantic Multidecadal Oscillation
BS	Baltic Sea
BSE	Baltic Sea Environmental index
BSH	German Federal Maritime and Hydrographic Agency
BSI	Baltic Sea Index
BRS	Biological Regime Shift
CCA	Canonical Correlation Analysis
COADS	Comprehensive Ocean-Atmosphere Data Set
CWT	Continuous Wavelet Transform
DJF	December-January-February mean
DWT	Discrete Wavelet Transform
ELM	Extreme Learning Machine
EOF	Empirical Orthogonal Function

ESRL	Earth System Research Laboratory
GS	Gotland Sea
HELCOM	Helsinki Commission
ICES	International Council for the Exploration of the Sea
IE	Baltic Sea ice extent
IOW	Leibniz Institute for Baltic Sea Research Warnemünde
JJA	June-July-August mean
LARS	least angle regression
LG	Landsort gauge
loess	local regression
LOO	Leave-one-out
LUNG	Landesamt für Umwelt, Naturschutz und Geologie of Mecklenburg-Western Pomerania
MAM	;March-April-May mean
NAO	North Atlantic Oscillation
NOAA	National Oceanic and Atmospheric Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centre of Environmental Predictions
OPELM	Optimally Pruned Extreme Learning Machine
OPKNN	Optimally Pruned K-Nearest Neighbours
SD	ICES subdivision
SMHI	Swedish Meteorological and Hydrological Institute
SLP	sea level pressure

SON	September-October-November mean
SST	sea surface temperature
WFT	windowed Fourier transform
WT	wavelet transform

Appendix B

Supporting Information

B.1 Arkona Sea EOF Analysis

	DJF				MAM				JJA				SON			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Chl-a					52	26	12	10	60	20	11	8	67	14	11	8
NO ₂	77	13	6	4	68	18	9	4	74	15	8	4	69	16	13	3
NO ₃	83	11	4	2	58	24	14	4	74	14	9	3	73	15	7	6
PO ₄	80	10	7	2	76	10	8	6	69	19	8	4	68	22	6	4
S	68	22	6	4	79	11	7	3	66	19	9	6	69	16	9	7
T	94	4	1		81	9	8	2	84	8	4	4	83	10	5	2

Table B.1: Explained variance of the EOFs of standardized abiotic data of the Arkona Sea in %. Only results for data with more than 50% presence is displayed.

Class/Group	Prop.	MAM				JJA				SON			
		1	2	3	4	1	2	3	4	1	2	3	4
Bacillariophyceae	ab	62	20	14	5	40	25	19	17	56	21	15	8
	bm	51	25	17	7	44	25	18	13	62	20	11	7
Chlorophyceae	ab	57	20	13	10	47	22	19	13	-	-	-	-
	bm	58	22	11	9	40	27	21	11	-	-	-	-
Chrysophyceae	ab	55	23	15	8	43	27	18	13	-	-	-	-
	bm	48	21	20	11	53	23	18	6	-	-	-	-
Craspedophyceae	ab	64	16	12	7	-	-	-	-	50	23	19	8
	bm	61	21	12	6	-	-	-	-	54	23	18	4
Cryptophyceae	ab	57	22	14	7	48	30	18	4	52	33	8	7
	bm	57	21	15	7	52	24	18	5	60	23	9	8
Cyanophyceae	ab	65	22	9	4	63	21	11	5	52	24	18	7
	bm	60	19	13	9	55	20	18	7	44	28	19	9
Dinophyceae	ab	45	22	18	15	54	24	14	9	61	22	13	4
	bm	58	18	15	9	53	24	16	7	61	20	15	4
Euglenophyceae	ab	52	24	16	8	43	32	13	12	49	23	19	10
	bm	68	16	12	4	54	28	12	6	63	19	11	6
Prasinophyceae	ab	63	21	11	5	39	32	17	11	64	18	11	7
	bm	69	15	8	8	45	31	16	8	62	17	12	8
Total Flagellates	ab	47	28	15	10	58	21	15	5	54	26	16	4
	bm	59	20	15	7	56	22	18	3	68	20	8	4
Total Phytoplankton	ab	43	32	16	9	59	20	16	5	53	26	17	4
	bm	50	27	15	8	56	23	19	2	66	20	11	4
Ratio Bacil.:Flagellates	ab	38	29	21	13	37	26	22	16	41	27	21	11
	bm	40	31	18	11	34	31	20	15	39	32	19	10

Table B.2: Explained variance of the EOFs of standardized phytoplankton data of the Arkona Sea in %. Only results for data with more than 50% presence is displayed. The ratio Bacillariophyceae:Flagellates is based on abundance and biomass data. ab: abundance, bm: biomass

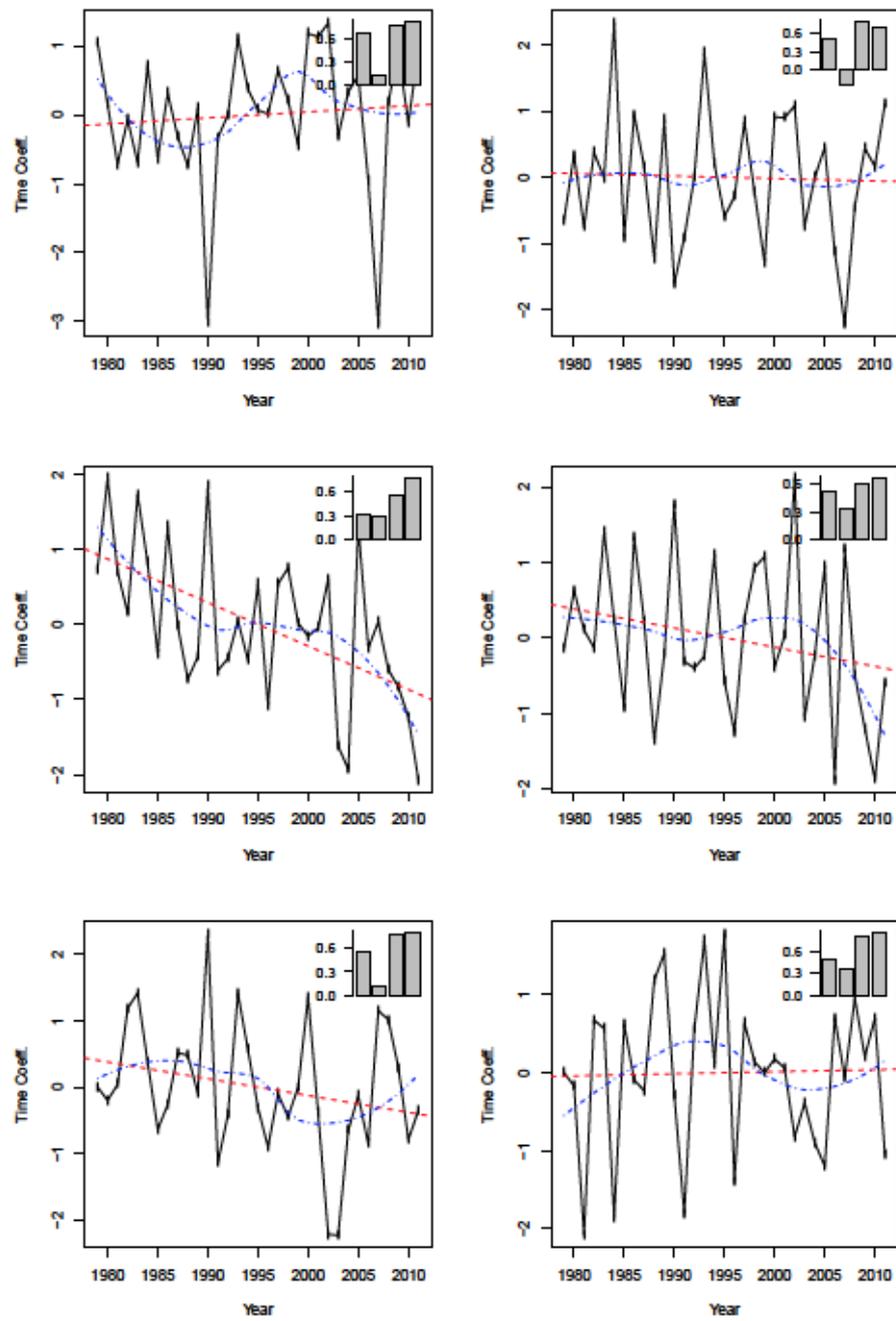


Figure B.1: Bacillariophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

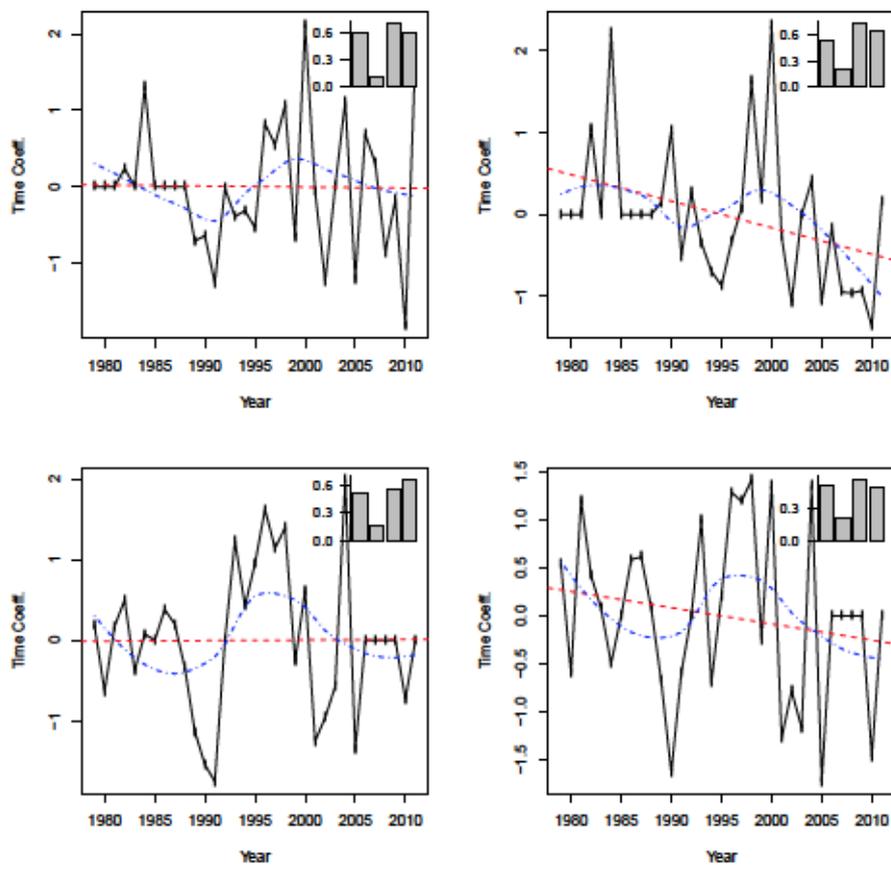


Figure B.2: Chlorophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

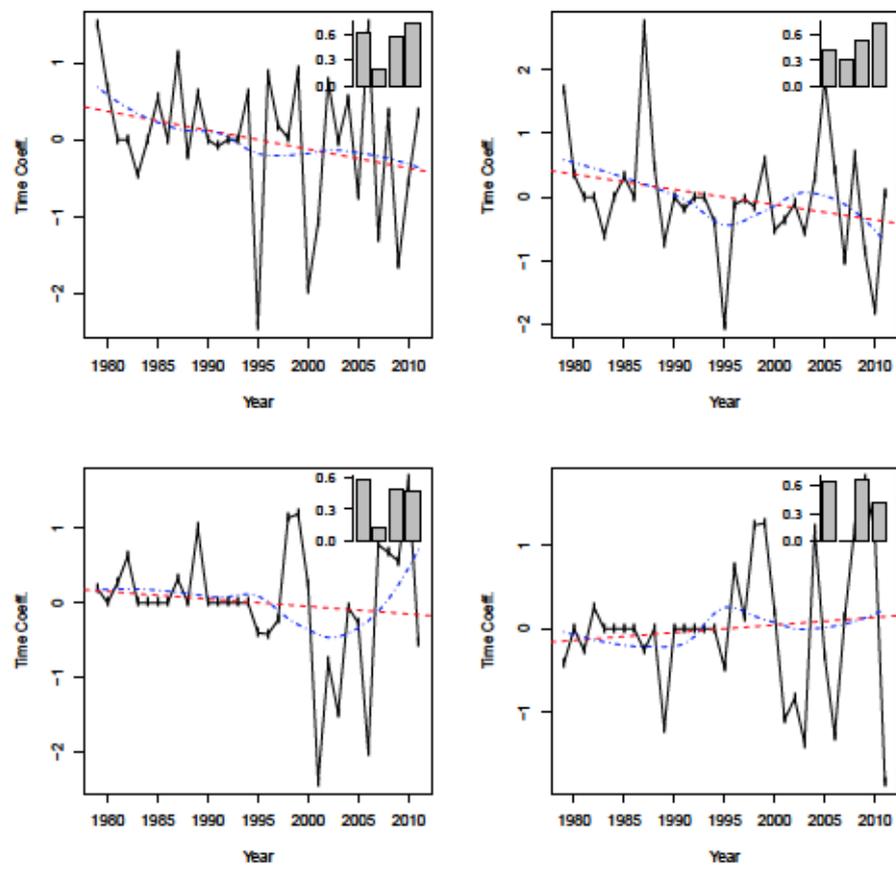


Figure B.3: Chrysophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

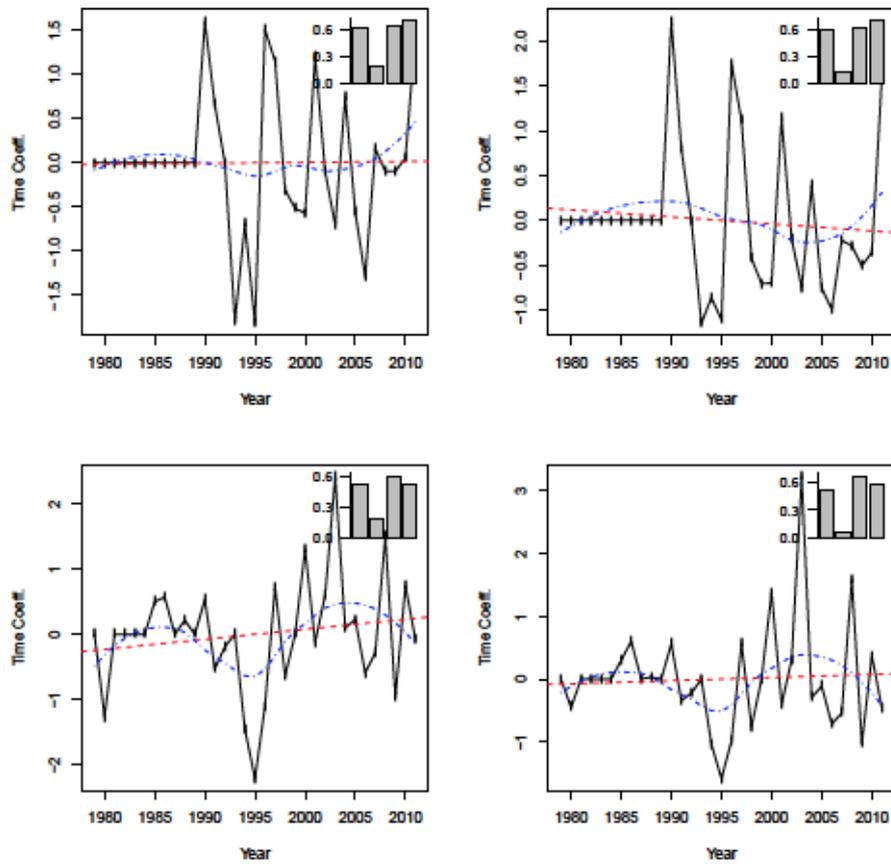


Figure B.4: Craspedophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), fall (SON)

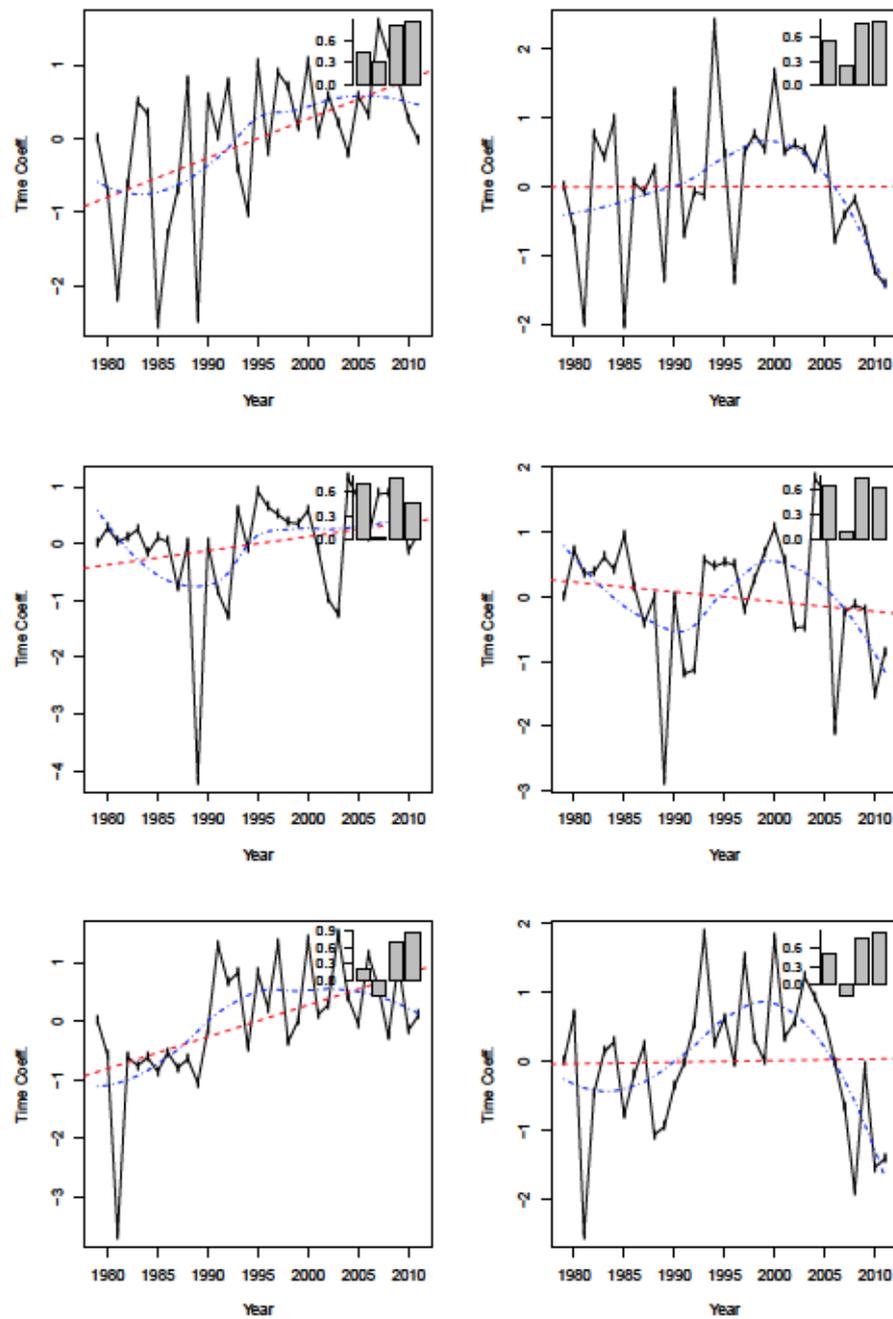


Figure B.5: Cryptophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

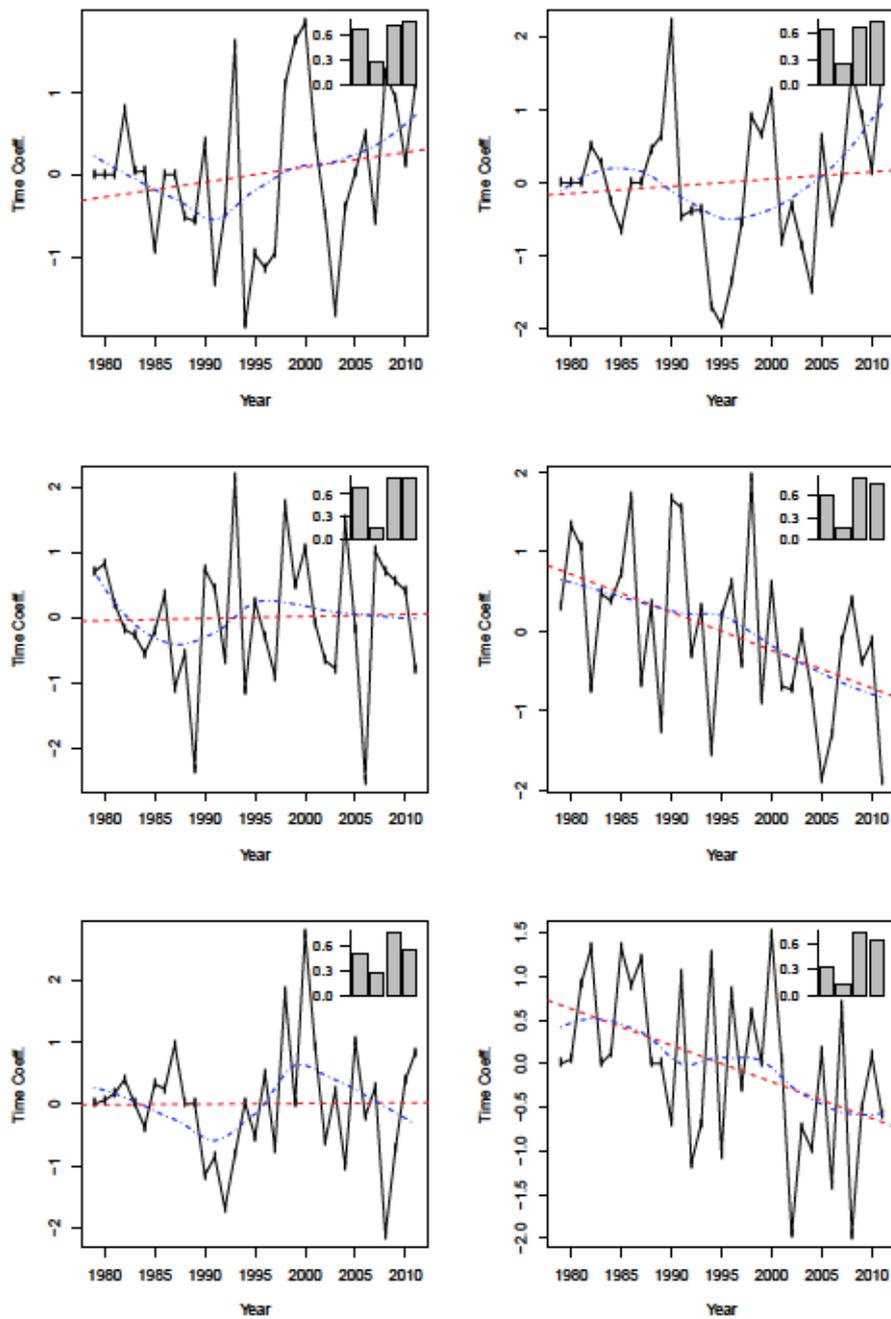


Figure B.6: Cyanophyceae (Cyanobacteria), time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

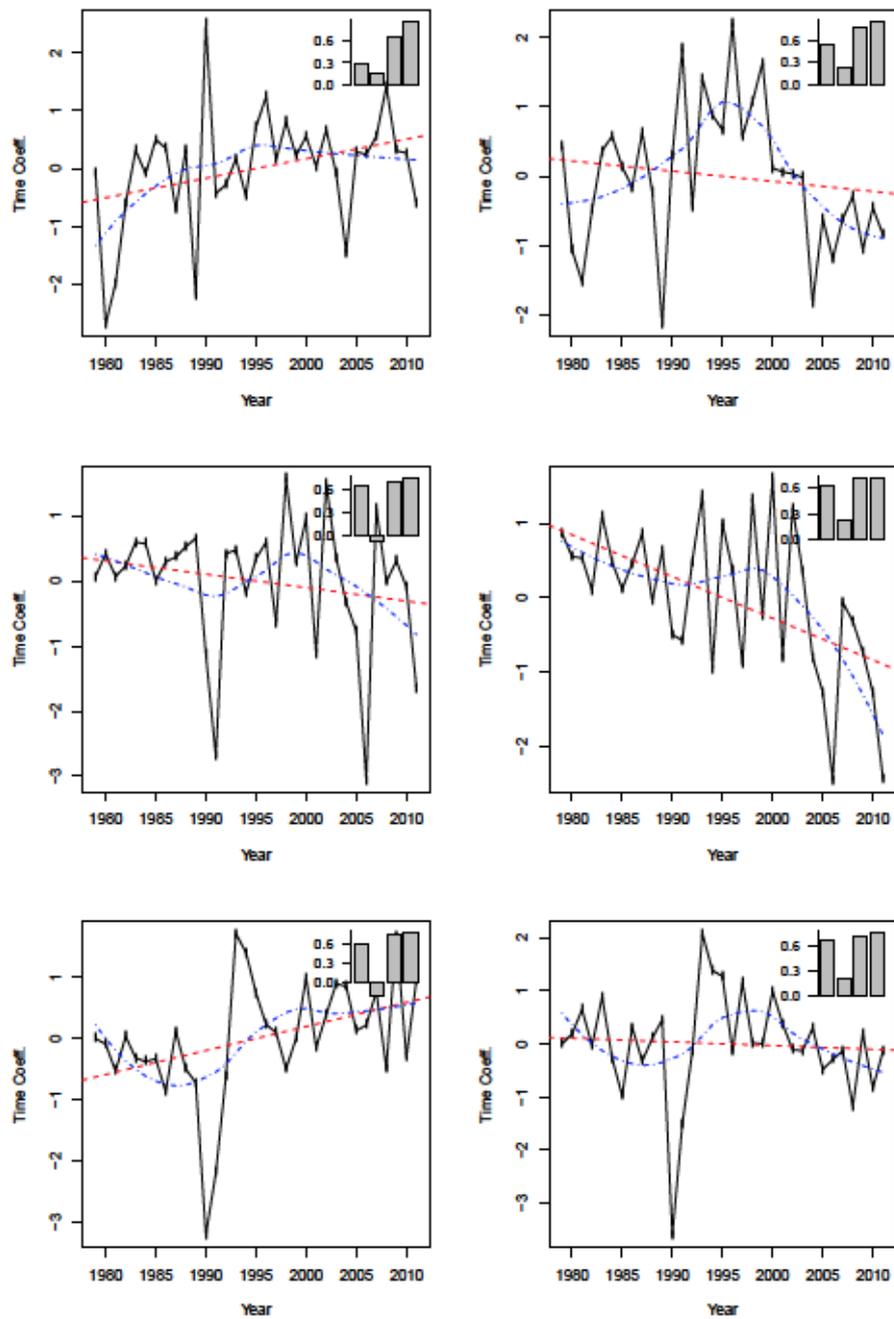


Figure B.7: Dinophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

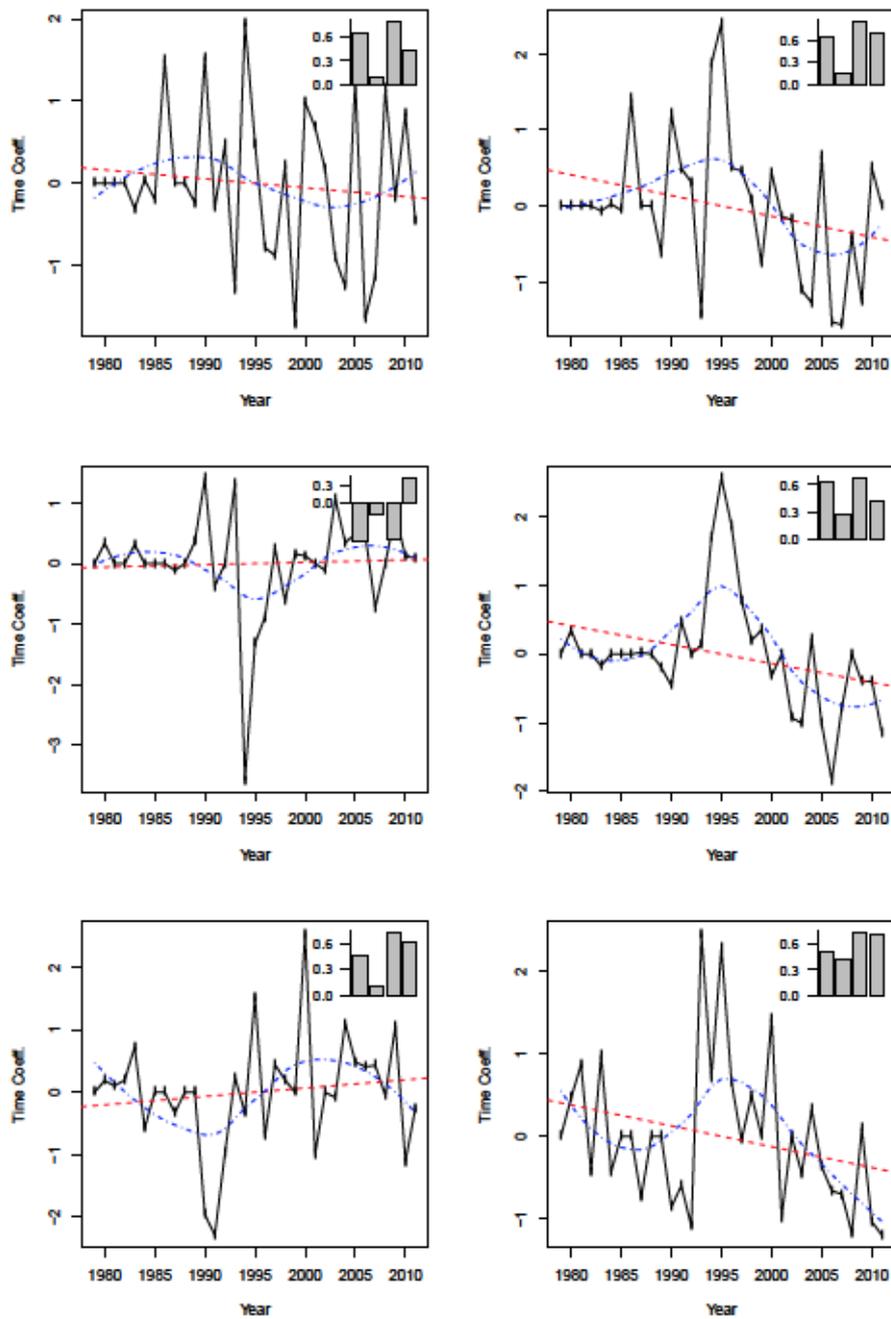


Figure B.8: Euglenophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

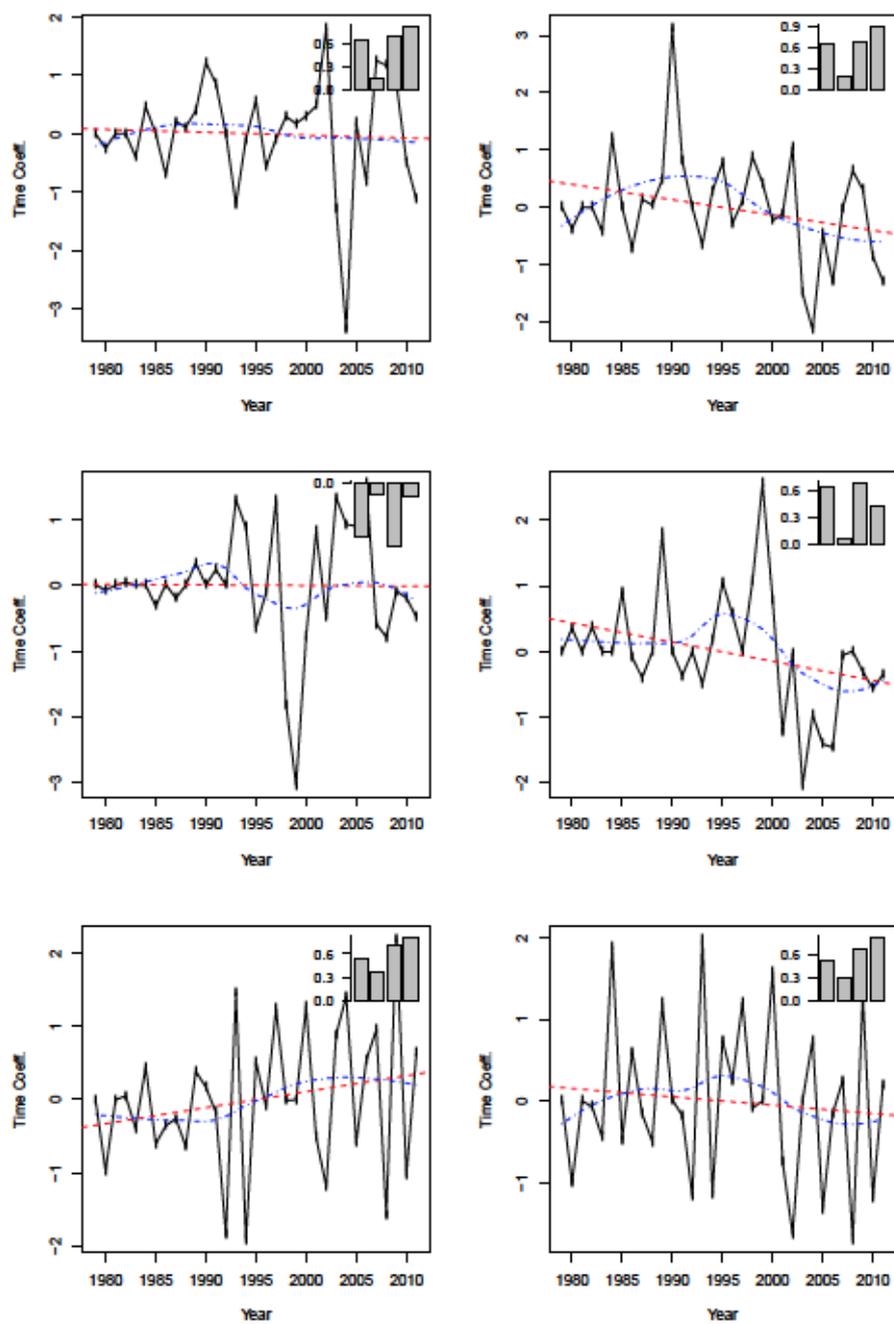


Figure B.9: Prasinophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

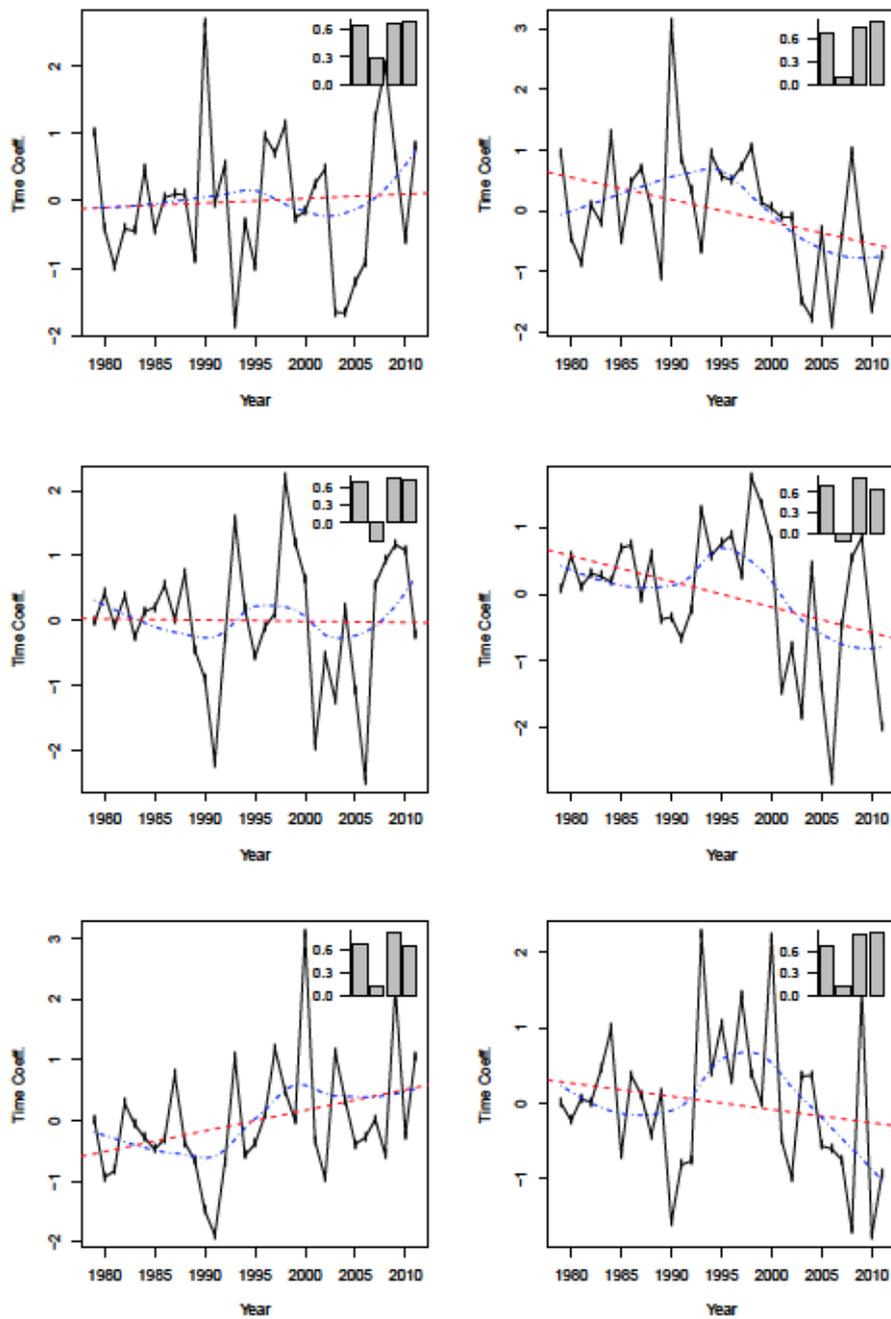


Figure B.10: Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

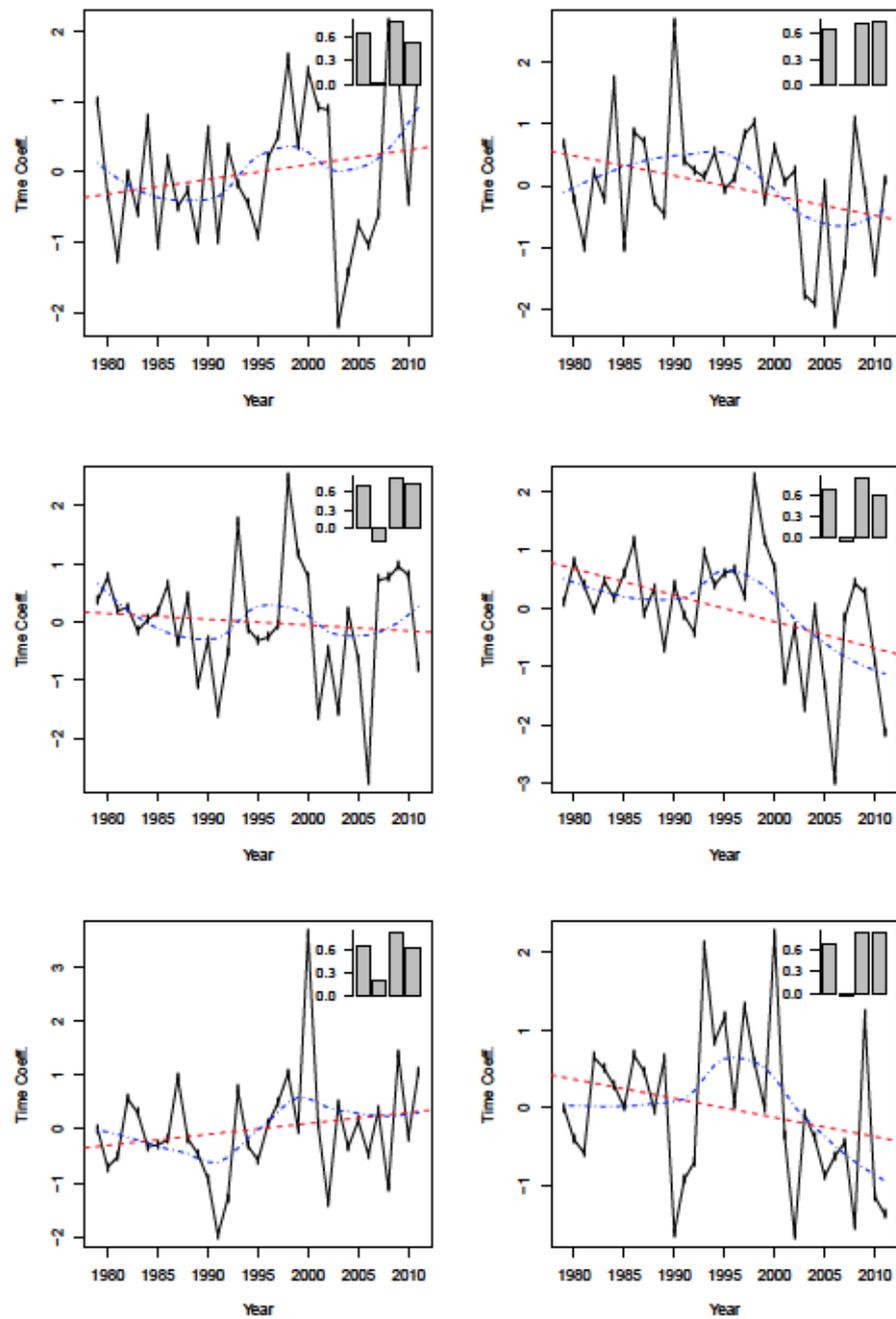


Figure B.11: Total Phytoplankton, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

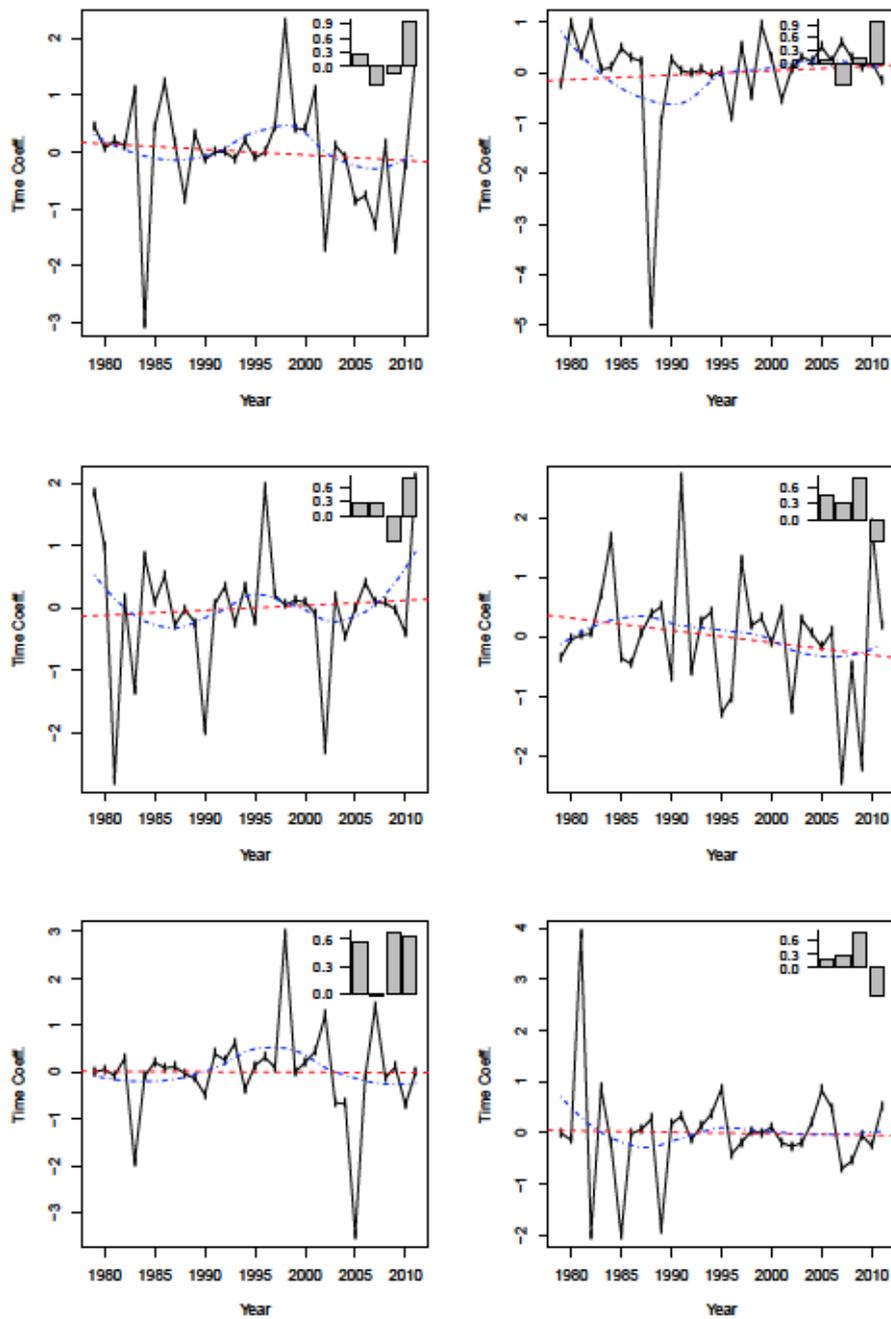


Figure B.12: Ratio Bacillariophyceae:Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

B.2 LUNG data EOF Analysis

B.2.1 All Stations

		MAM			JJA			SON		
Class/Group	Prop.	1	2	3	1	2	3	1	2	3
Bacillariophyceae	ab	32	17	14	29	25	17	34	24	11
	bv	23	19	15	23	22	17	31	22	15
Cryptophyceae	ab	30	18	13	27	17	15	26	23	13
	bv	30	21	18	28	17	14	31	19	16
Cyanophyceae	ab	-	-	-	30	26	15	-	-	-
	bv	-	-	-	30	24	19	-	-	-
Dinophyceae	ab	-	-	-	29	18	14	34	20	18
	bv	26	22	17	24	22	17	30	20	15
Total Flagellates	ab	23	21	18	22	20	16	26	24	14
	bv	27	16	13	26	19	15	27	16	14
Total phytoplankton	ab	21	19	17	26	18	16	33	17	12
	bv	25	20	15	29	19	13	27	21	14
Ratio Bacil.:Flagellates	ab	32	23	16	26	21	18	-	-	-
	bv	26	22	15	26	21	17	-	-	-

Table B.3: Explained variance of the first 3 EOFs of standardized phytoplankton data of the LUNG data set. All stations are considered (group A). Only results for data with more than 50% presence are displayed. ab:abundance, bv:biovolume

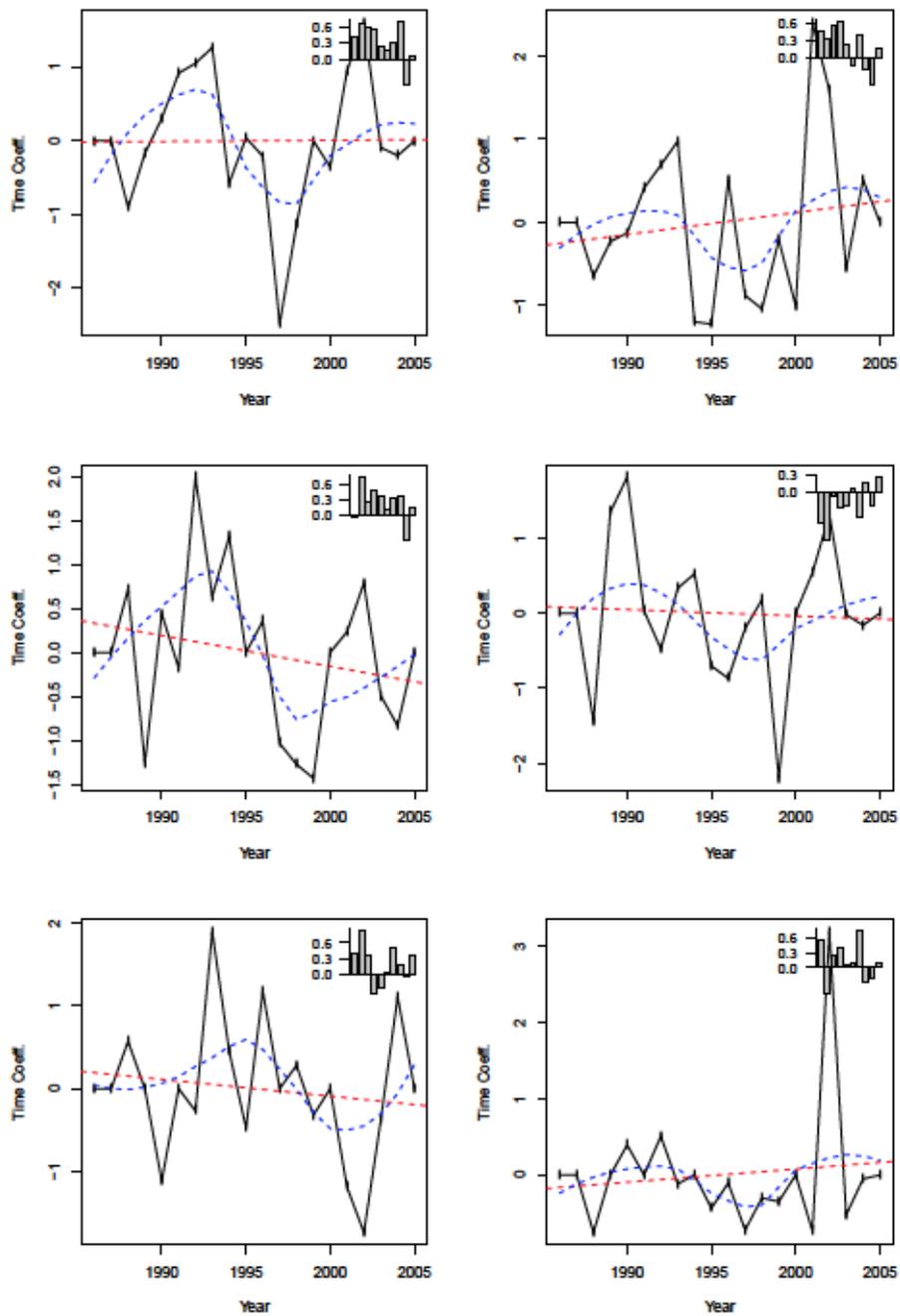


Figure B.13: Bacillariophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

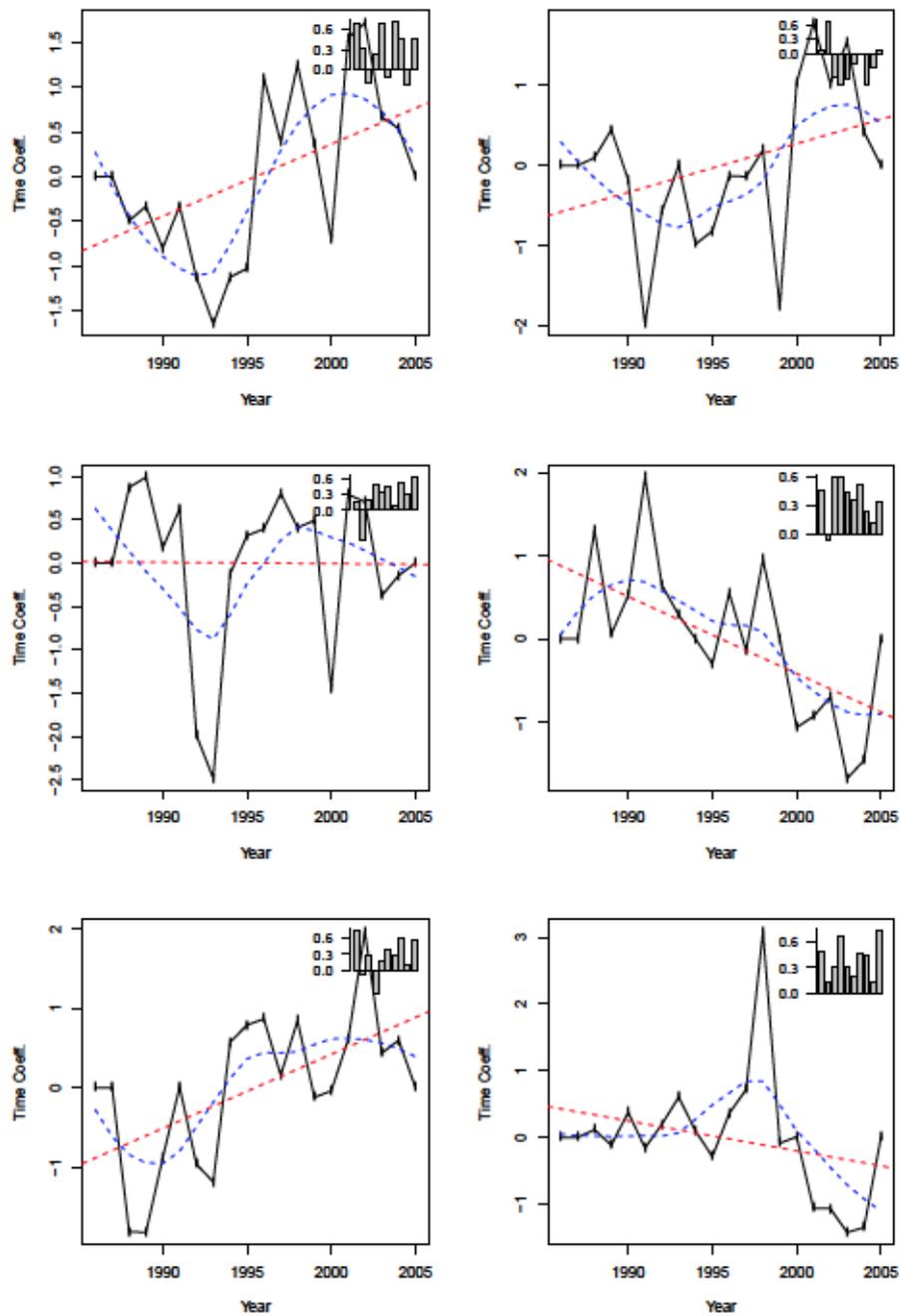


Figure B.14: Cryptophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

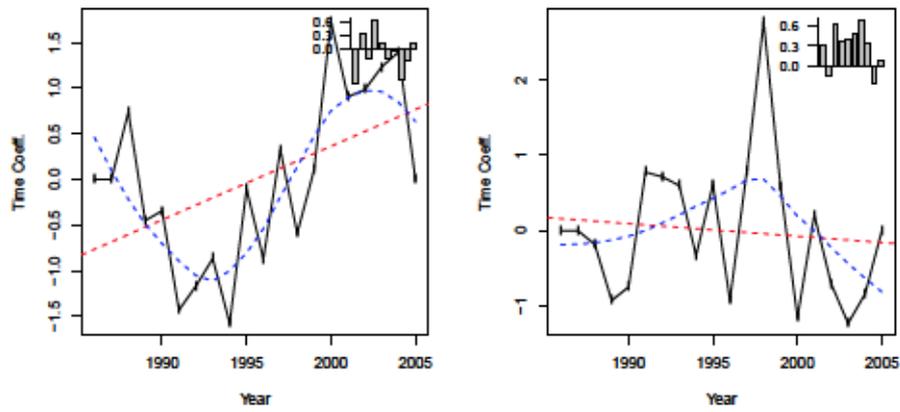


Figure B.15: Cyanophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; Only summer (JJA)

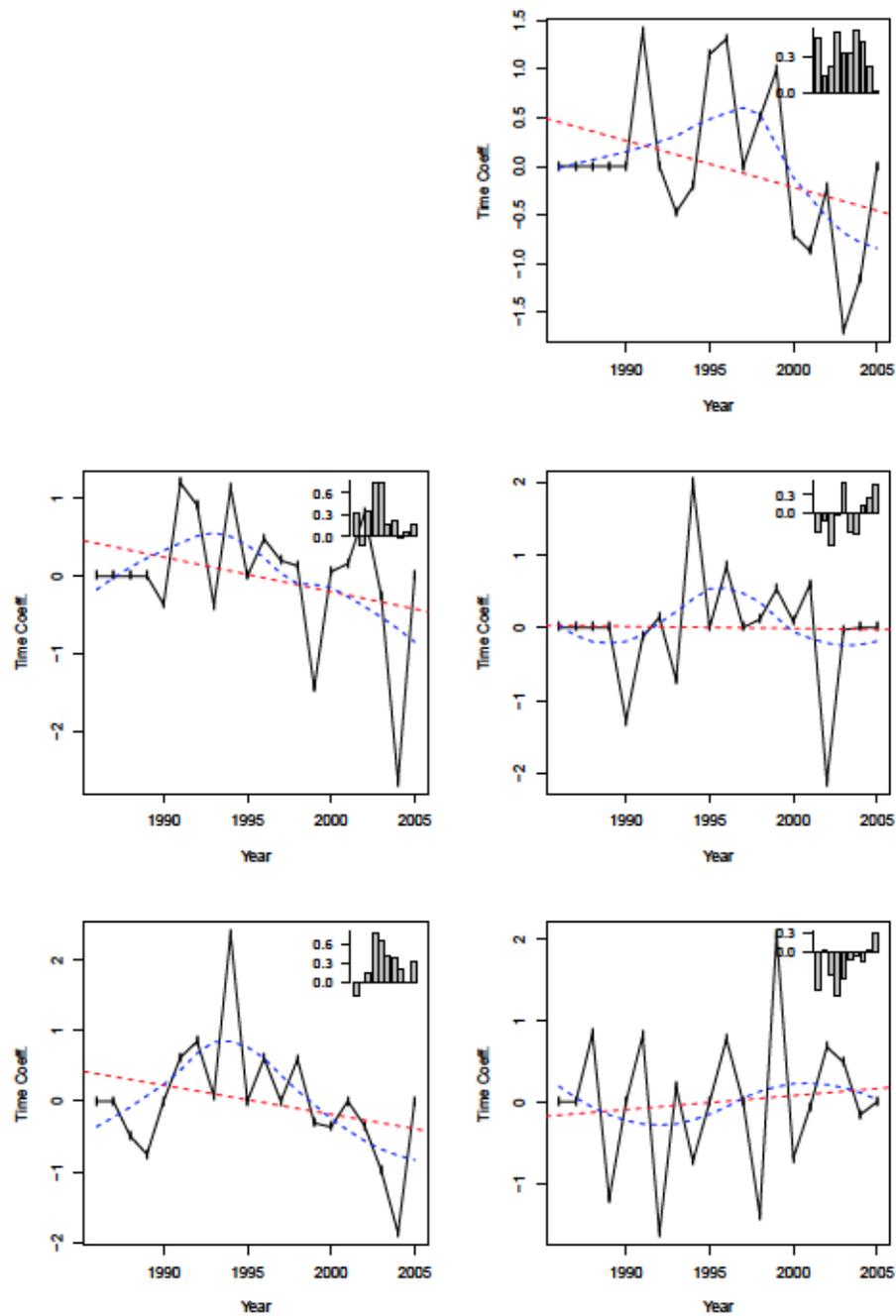


Figure B.16: Dinophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

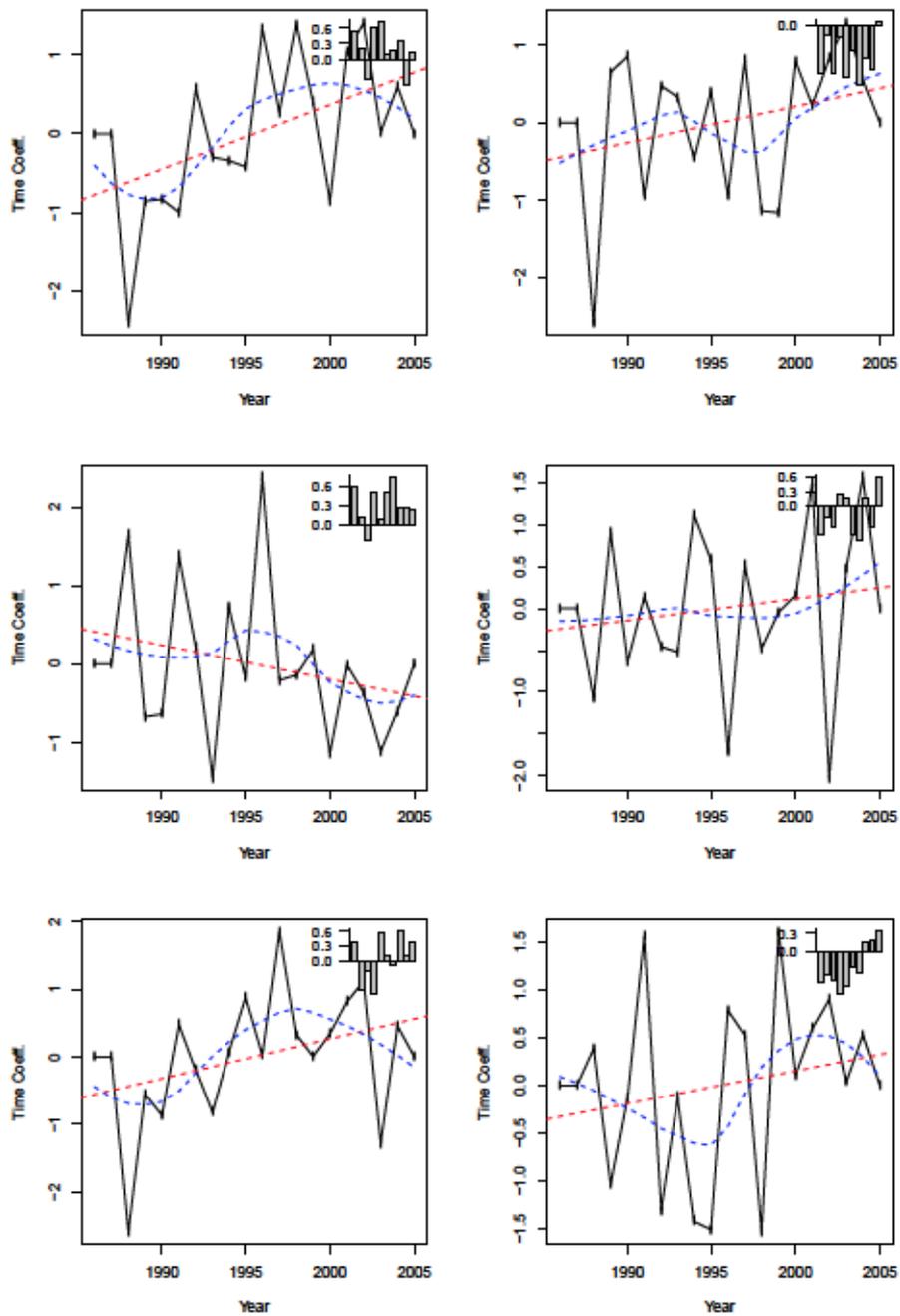


Figure B.17: Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

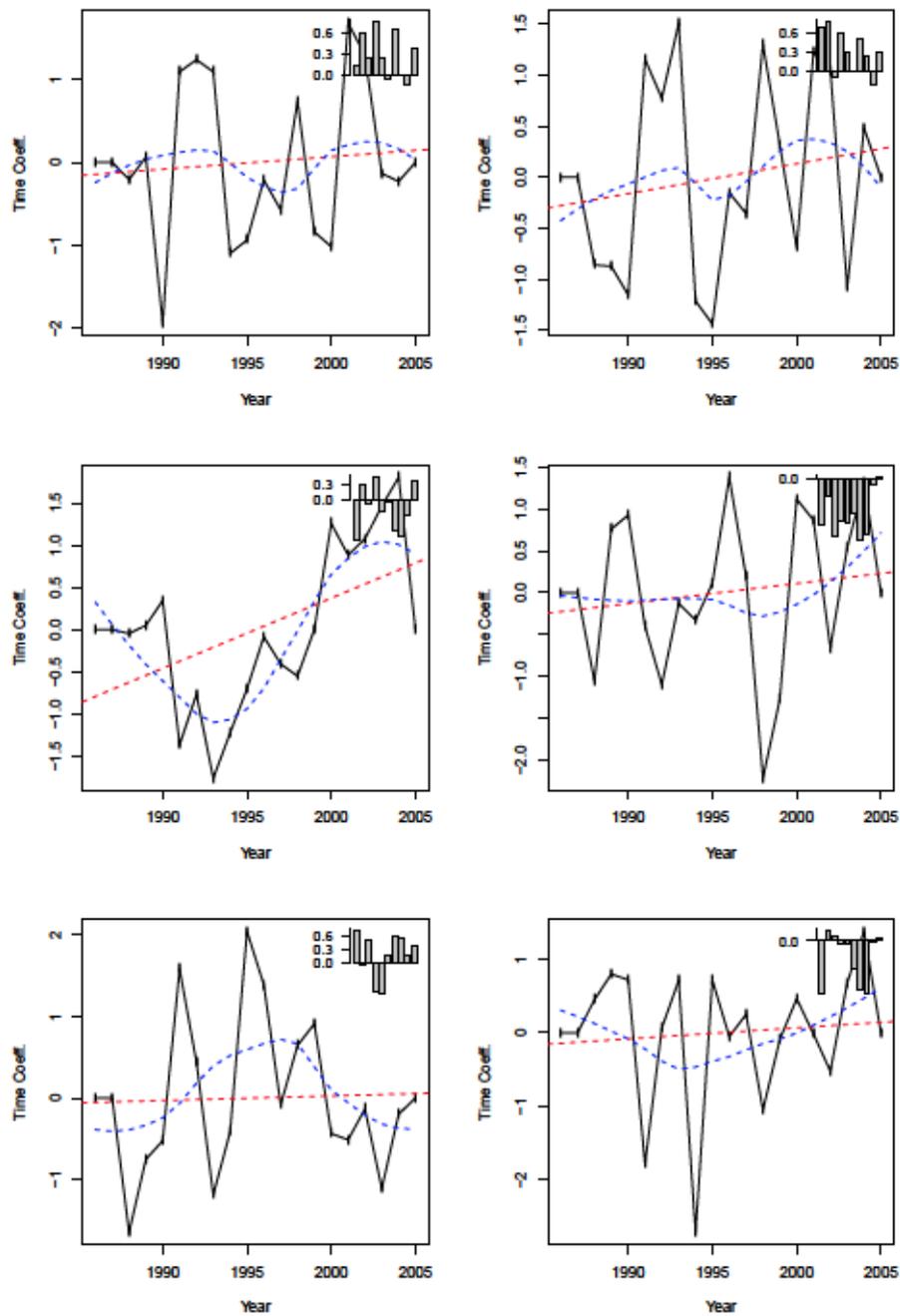


Figure B.18: Total Phytoplankton, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

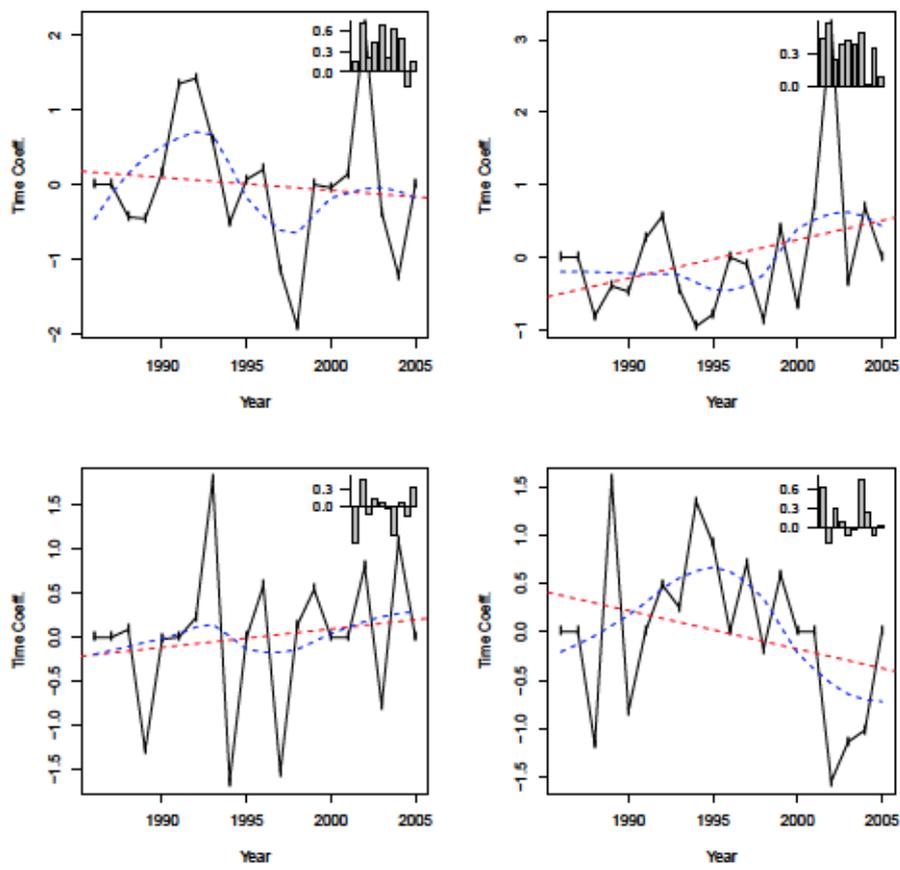


Figure B.19: Ratio Bacillariophyceae:Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

B.2.2 All Stations except S66 and KHM

		MAM			JJA			SON		
Class/Group	Prop.	1	2	3	1	2	3	1	2	3
Bacillariophyceae	ab	28	22	18	29	24	20	-	-	-
	bv	29	19	16	27	22	16	-	-	-
Cryptophyceae	ab	35	18	16	25	19	17	29	21	14
	bv	31	22	17	34	18	16	36	23	12
Cyanophyceae	ab	-	-	-	32	29	12	-	-	-
	bv	-	-	-	40	20	15	-	-	-
Dinophyceae	ab	-	-	-	32	20	16	39	21	16
	bv	28	26	17	26	23	19	35	19	17
Total Flagellates	ab	27	23	15	27	21	15	28	24	14
	bv	33	17	16	32	19	17	33	18	16
Total phytoplankton	ab	24	21	19	27	21	16	38	19	13
	bv	24	24	15	32	20	15	27	25	16
Ratio Bacil.:Flagellates	ab	31	29	16	-	-	-	-	-	-
	bv	32	29	17	-	-	-	-	-	-

Table B.4: Explained variance of the first 3 EOFs of standardized phytoplankton data of the LUNG data set. All stations are considered except for S66 and KHM(group B). Only results for data with more than 50% presence are displayed. ab:abundance, bv:biovolume

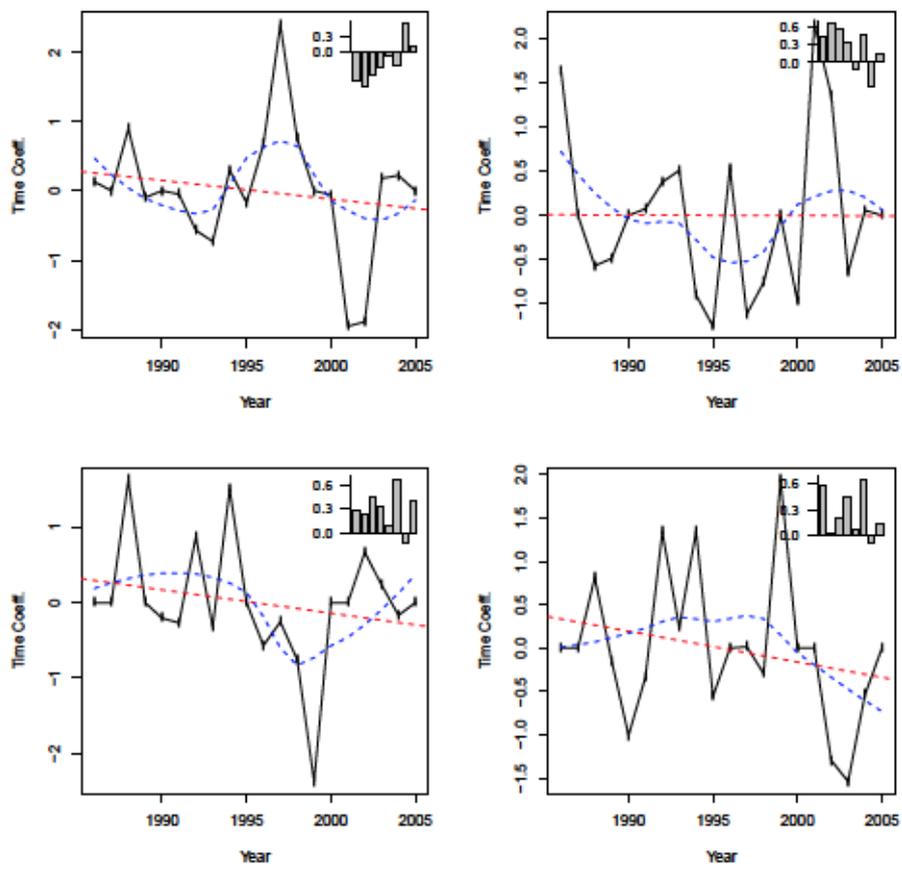


Figure B.20: Bacillariophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

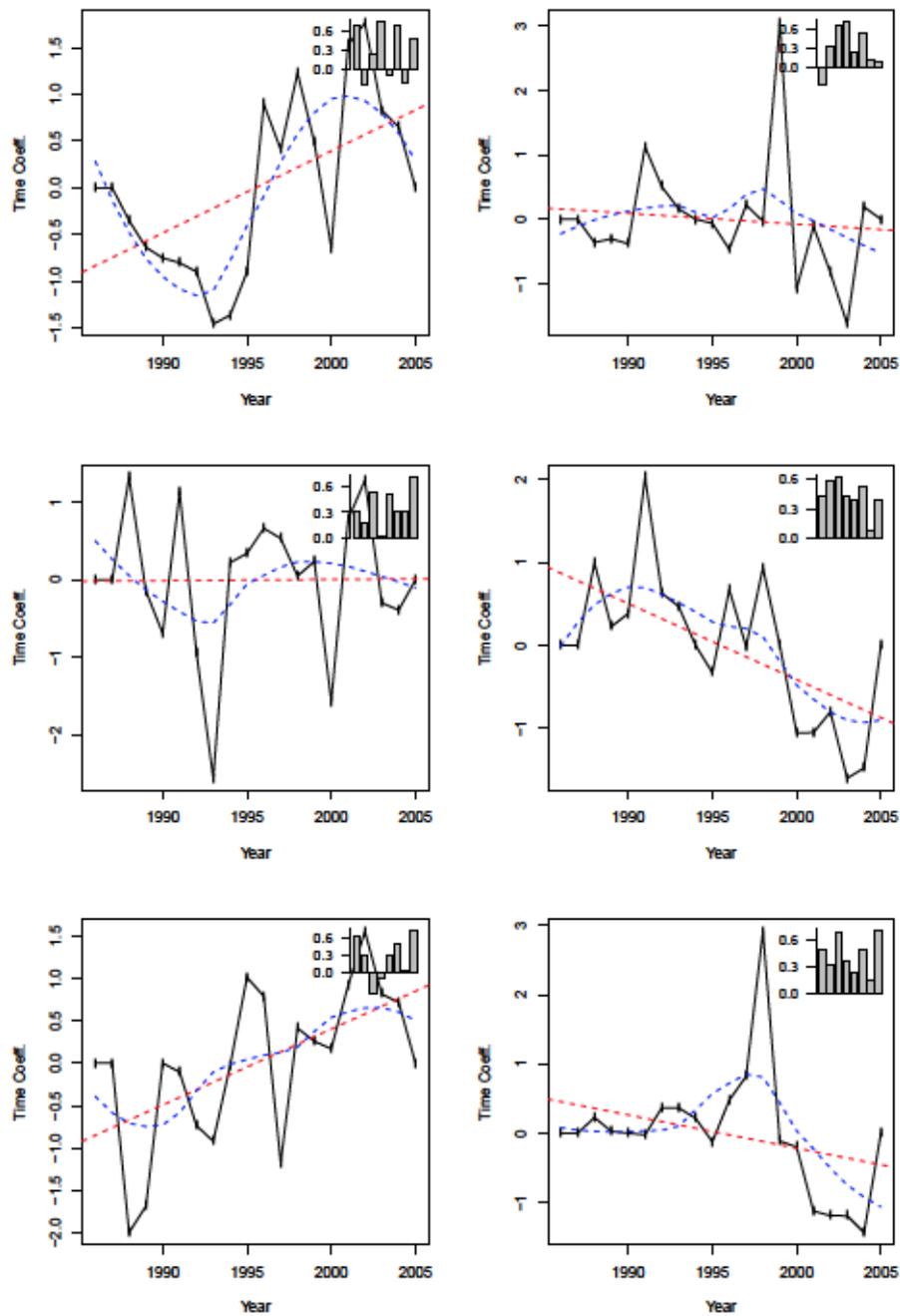


Figure B.21: Cryptophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

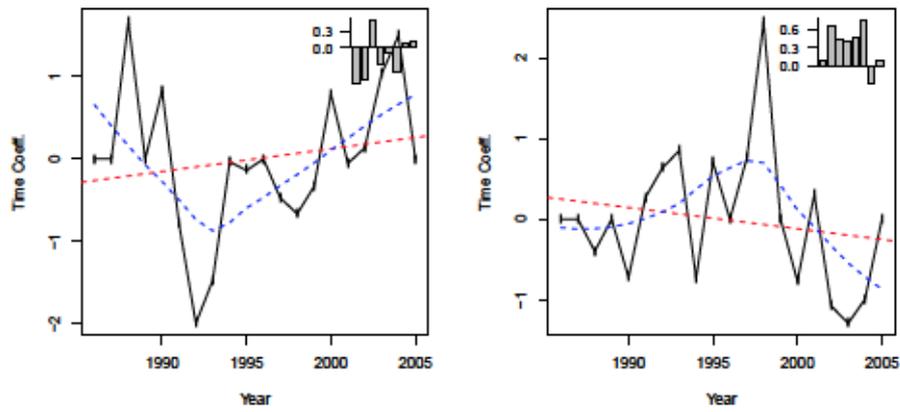


Figure B.22: Cyanophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; Only summer (JJA)

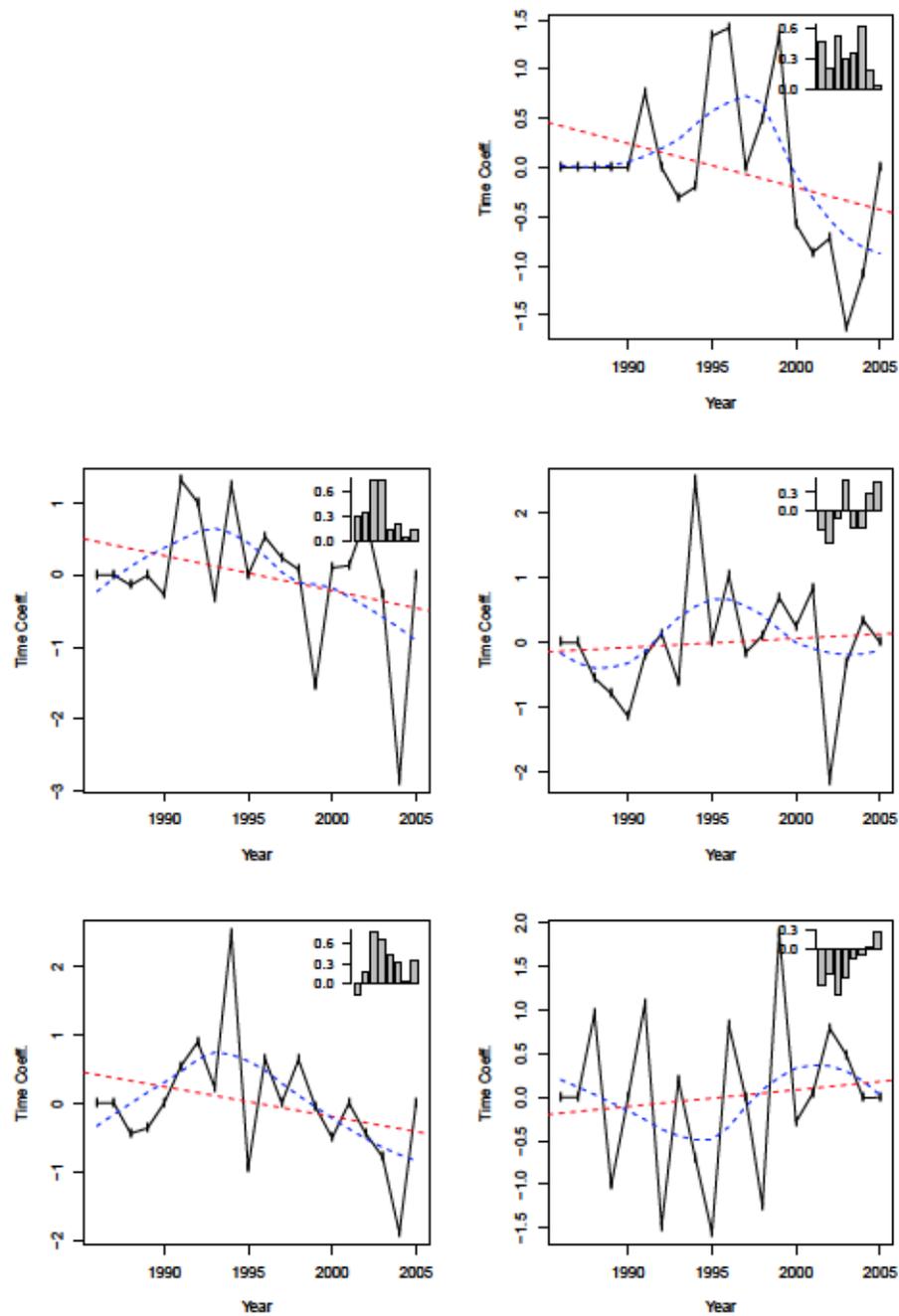


Figure B.23: Dinophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

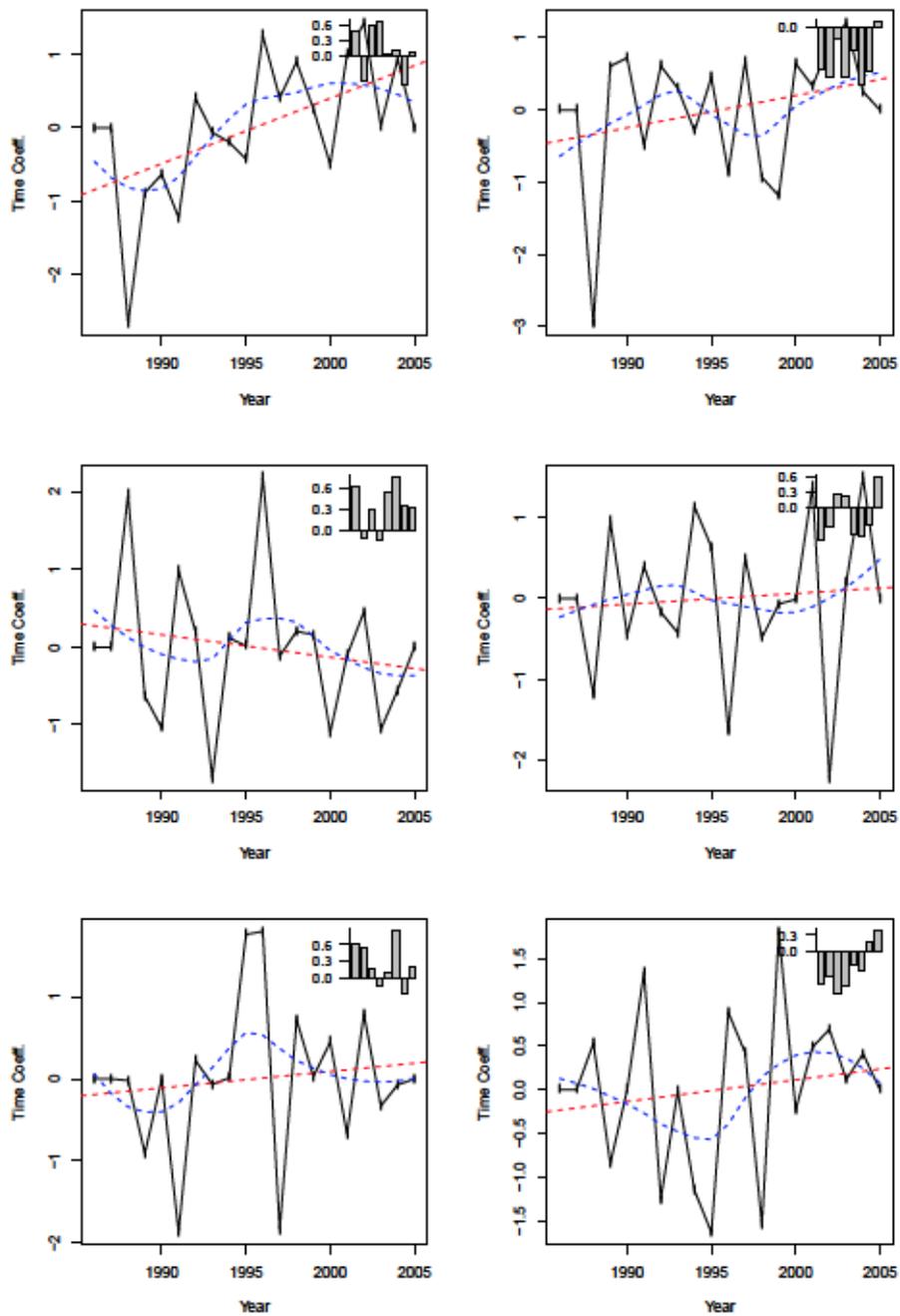


Figure B.24: Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

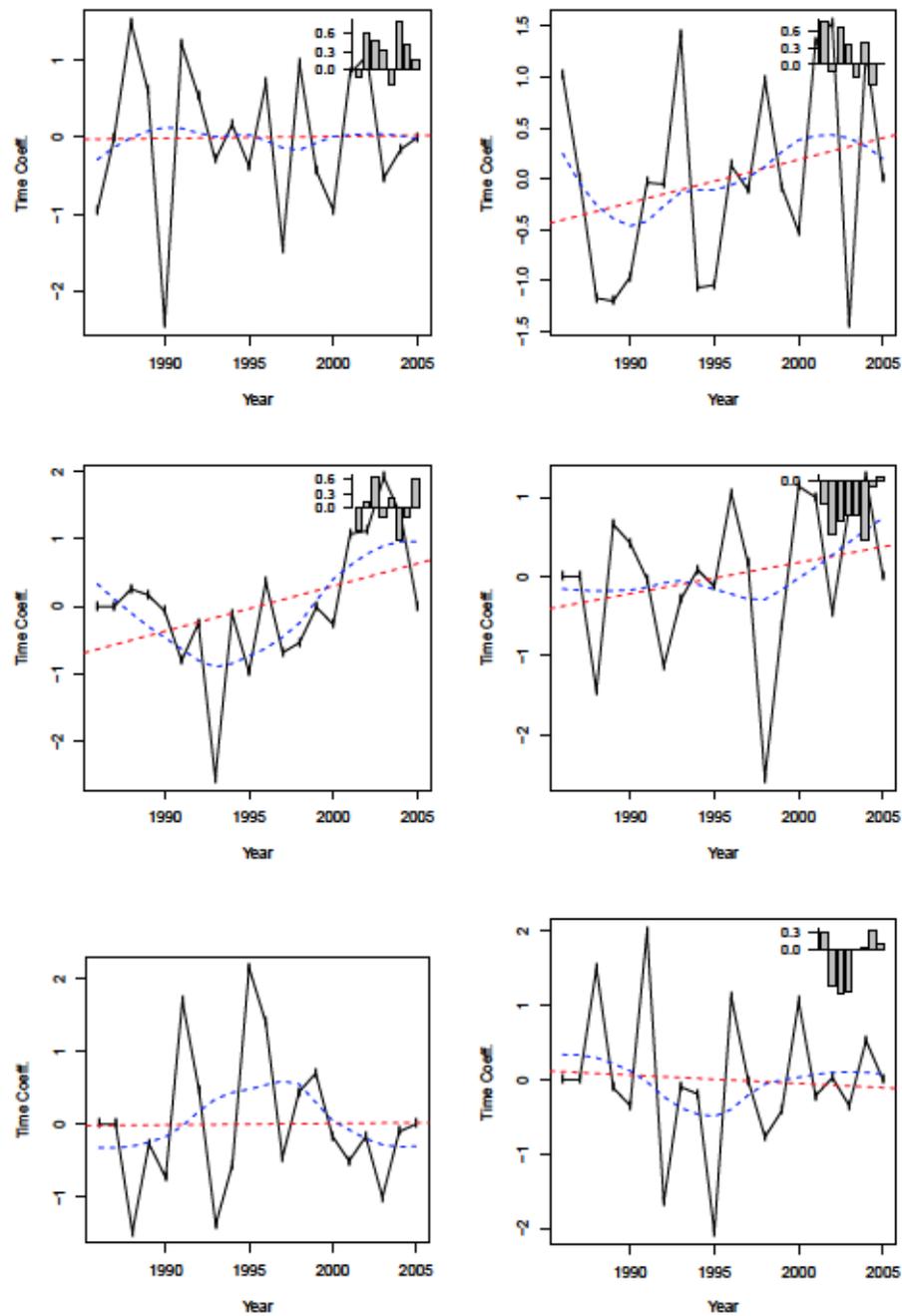


Figure B.25: Total Phytoplankton, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

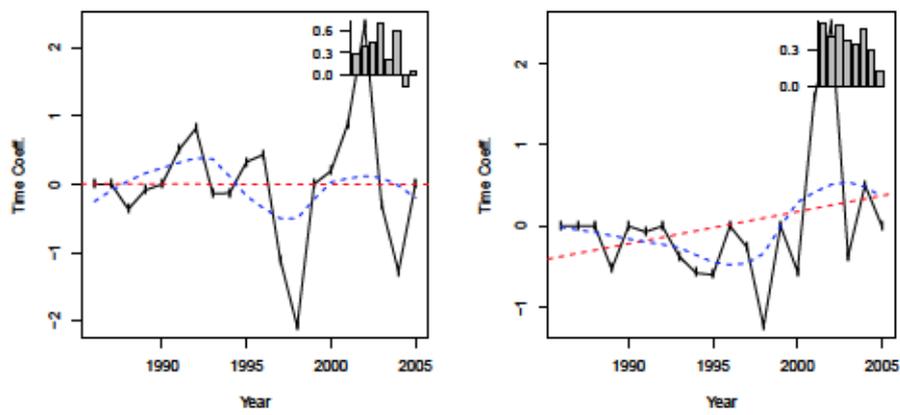


Figure B.26: Ratio Bacillariophyceae:Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

B.2.3 Only Outer Coastal Zone Stations

		MAM			JJA			SON		
Class/Group	Prop.	1	2	3	1	2	3	1	2	3
Bacillariophyceae	ab	50	22	17	-	-	-	-	-	-
	bv	48	27	18	-	-	-	-	-	-
Cryptophyceae	ab	37	35	17	36	28	26	39	23	23
	bv	61	19	14	57	23	14	51	20	15
Cyanophyceae	ab	-	-	-	42	30	17	-	-	-
	bv	-	-	-	59	25	13	-	-	-
Dinophyceae	ab	-	-	-	53	24	16	68	19	7
	bv	-	-	-	37	35	19	56	17	14
Total Flagellates	ab	40	31	18	39	29	23	38	24	21
	bv	43	26	18	34	33	21	53	26	15
Total Phytoplankton	ab	40	29	17	38	29	21	54	23	18
	bv	33	27	21	47	29	16	54	20	17
Ratio Bacil.:Flagellates	ab	50	27	15	-	-	-	-	-	-
	bv	-	-	-	-	-	-	-	-	-

Table B.5: Explained variance of the first 3 EOFs of standardized phytoplankton data of the LUNG data set. Only outer coastal stations are considered (group C). Only results for data with more than 50% presence are displayed. ab:abundance, bv:biovolume

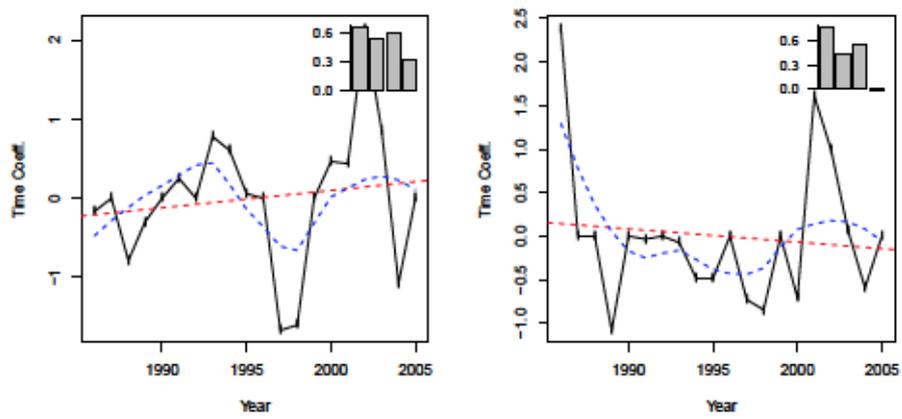


Figure B.27: Bacillariophyceae, time coefficients of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

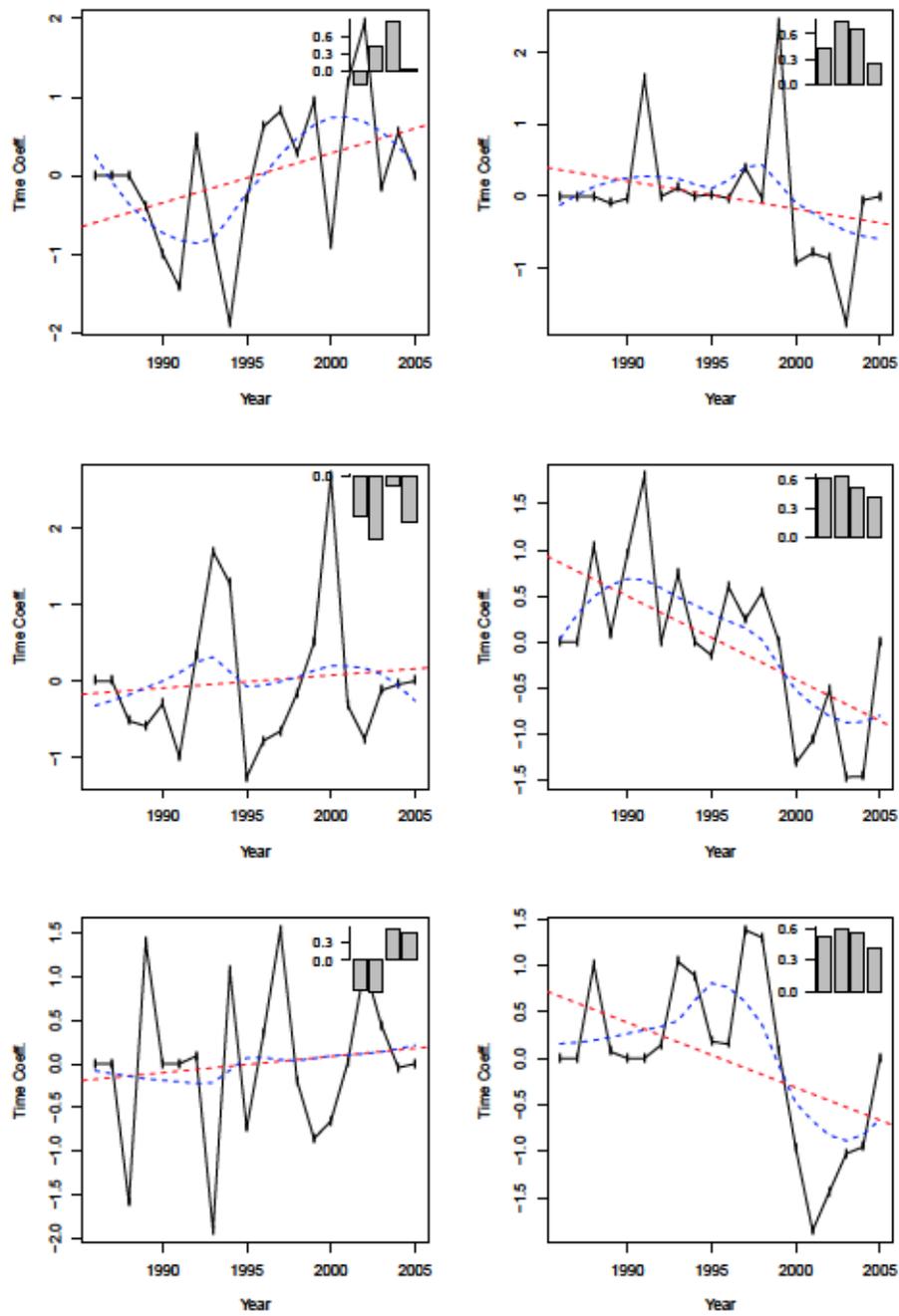


Figure B.28: Cryptophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

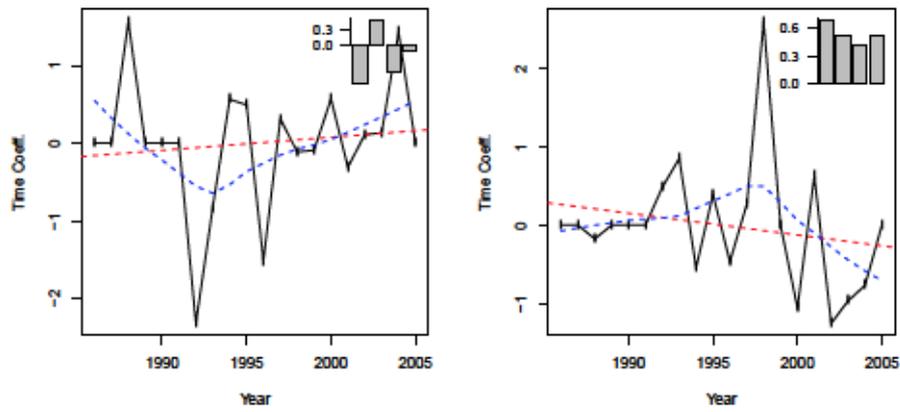


Figure B.29: Cyanophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; Only summer (JJA)

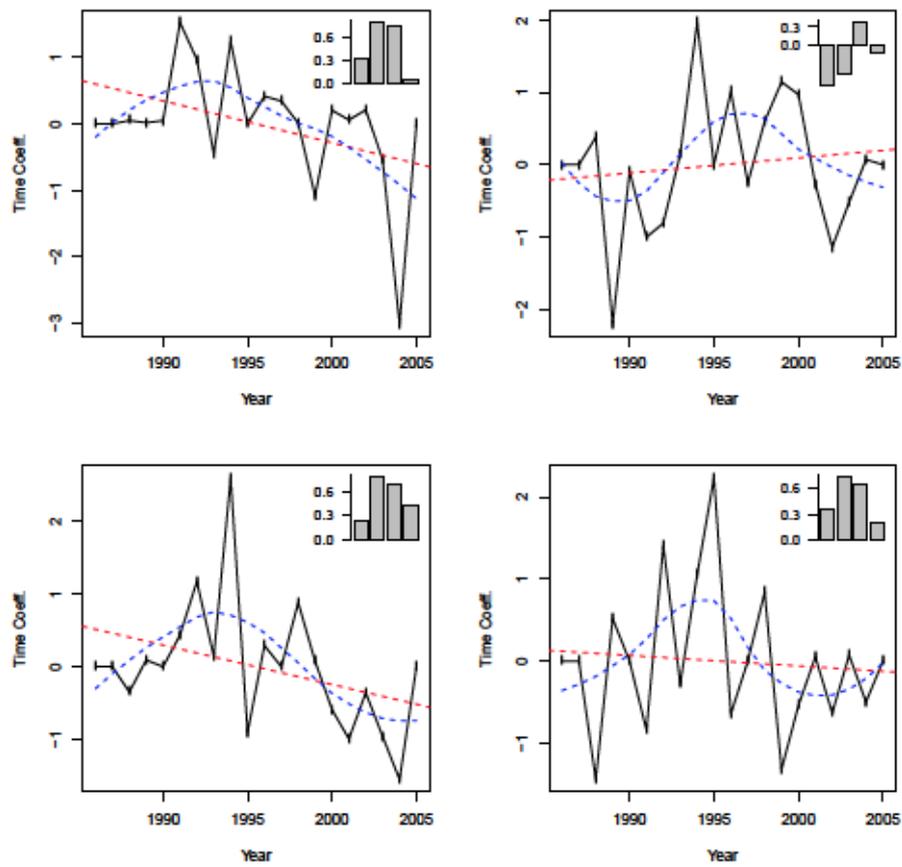


Figure B.30: Dinophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: summer (JJA), fall (SON)

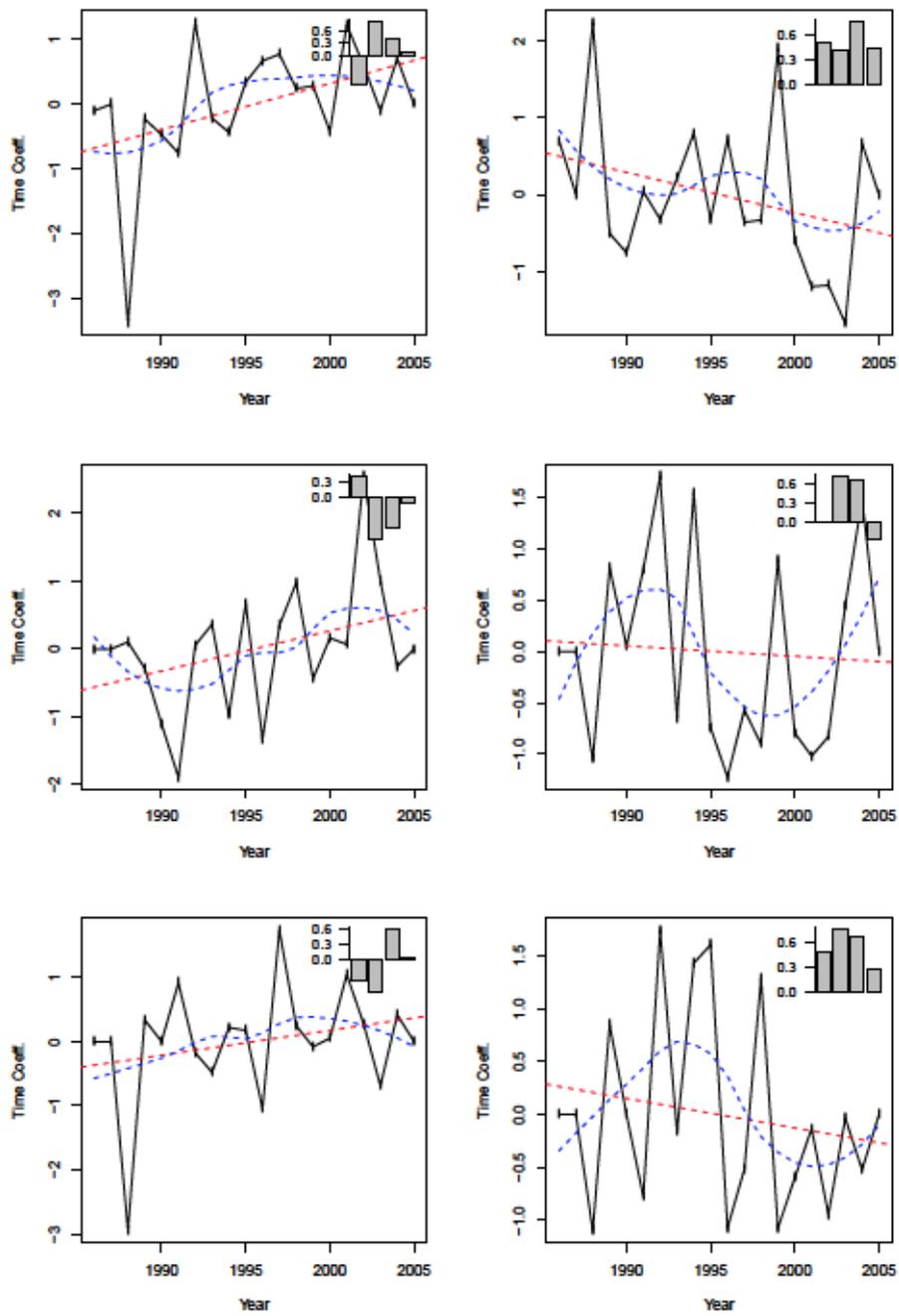


Figure B.31: Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

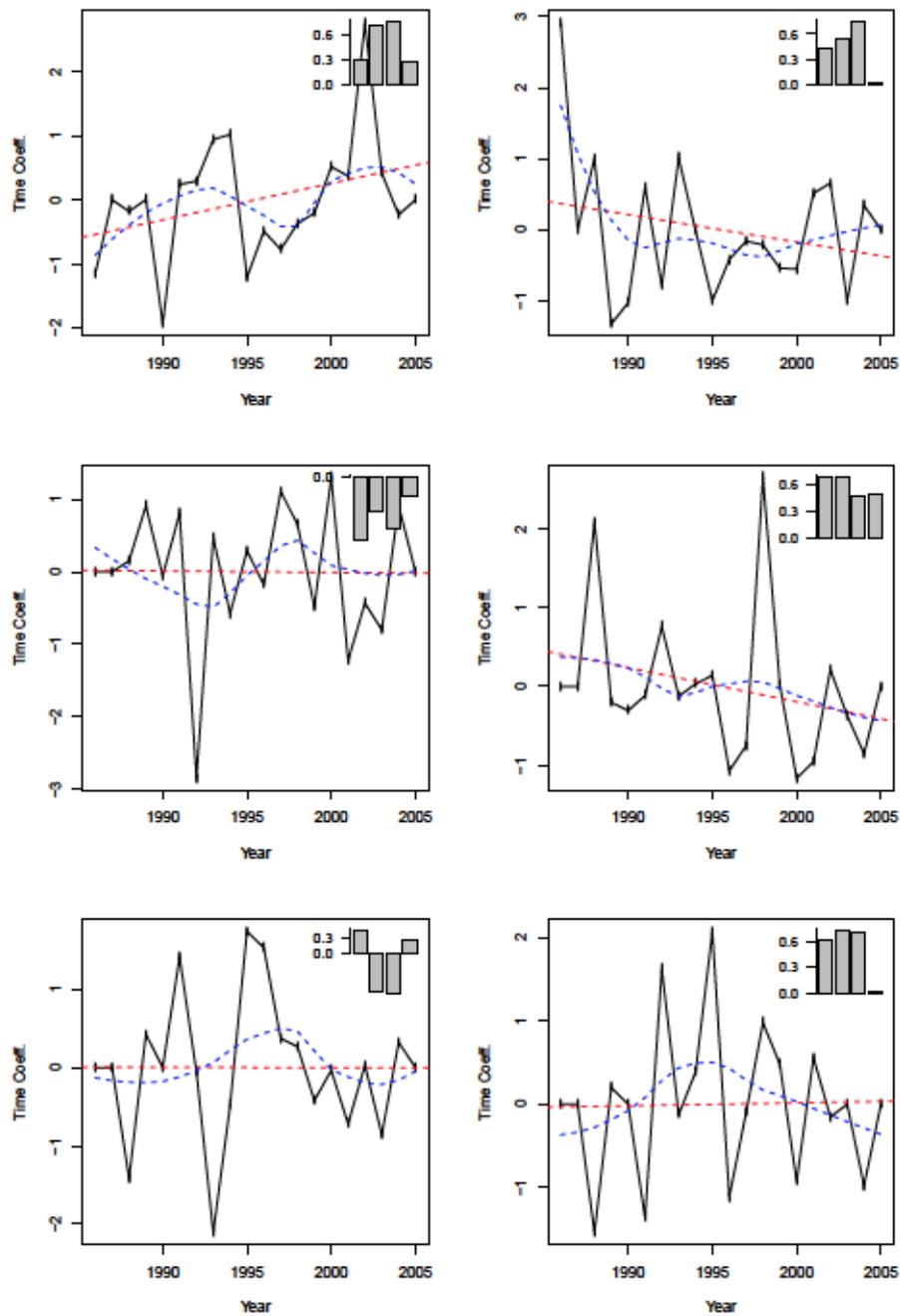


Figure B.32: Total Phytoplankton, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

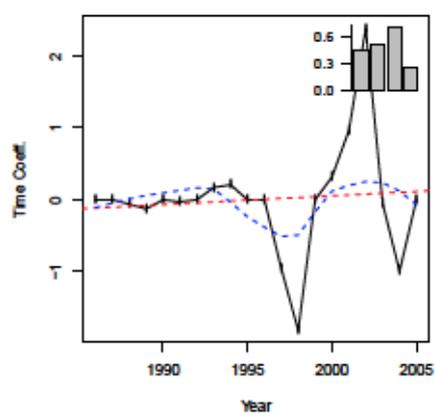


Figure B.33: Ratio Bacillariophyceae:Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, Spring (MAM)

B.2.4 Only Inner Coastal Zone Stations

		MAM			JJA			SON		
Class/Group	Prop.	1	2	3	1	2	3	1	2	3
Bacillariophyceae	ab	38	22	17	33	27	21	40	25	13
	bv	29	20	19	30	27	21	37	26	14
Cryptophyceae	ab	39	25	16	35	23	17	35	29	15
	bv	32	25	16	26	24	20	35	24	19
Cyanophyceae	ab	-	-	-	37	28	20	44	27	12
	bv	-	-	-	38	25	24	50	28	15
Dinophyceae	ab	-	-	-	-	-	-	-	-	-
	bv	30	25	19	-	-	-	-	-	-
total Flagellates	ab	31	21	17	29	26	17	35	25	17
	bv	32	23	17	33	24	20	25	23	19
total phytoplankton	ab	29	26	21	39	21	19	38	20	16
	bv	36	26	14	32	27	17	40	20	17
Ratio Bacil.:Flagellates	ab	34	25	22	35	22	20	33	24	15
	bv	29	25	15	32	28	17	39	23	16

Table B.6: Explained variance of the first 3 EOFs of standardized phytoplankton data of the LUNG data set. Only inner coastal stations are considered (group D). Only results for data with more than 50% presence are displayed. ab :abundance, bv:biovolume

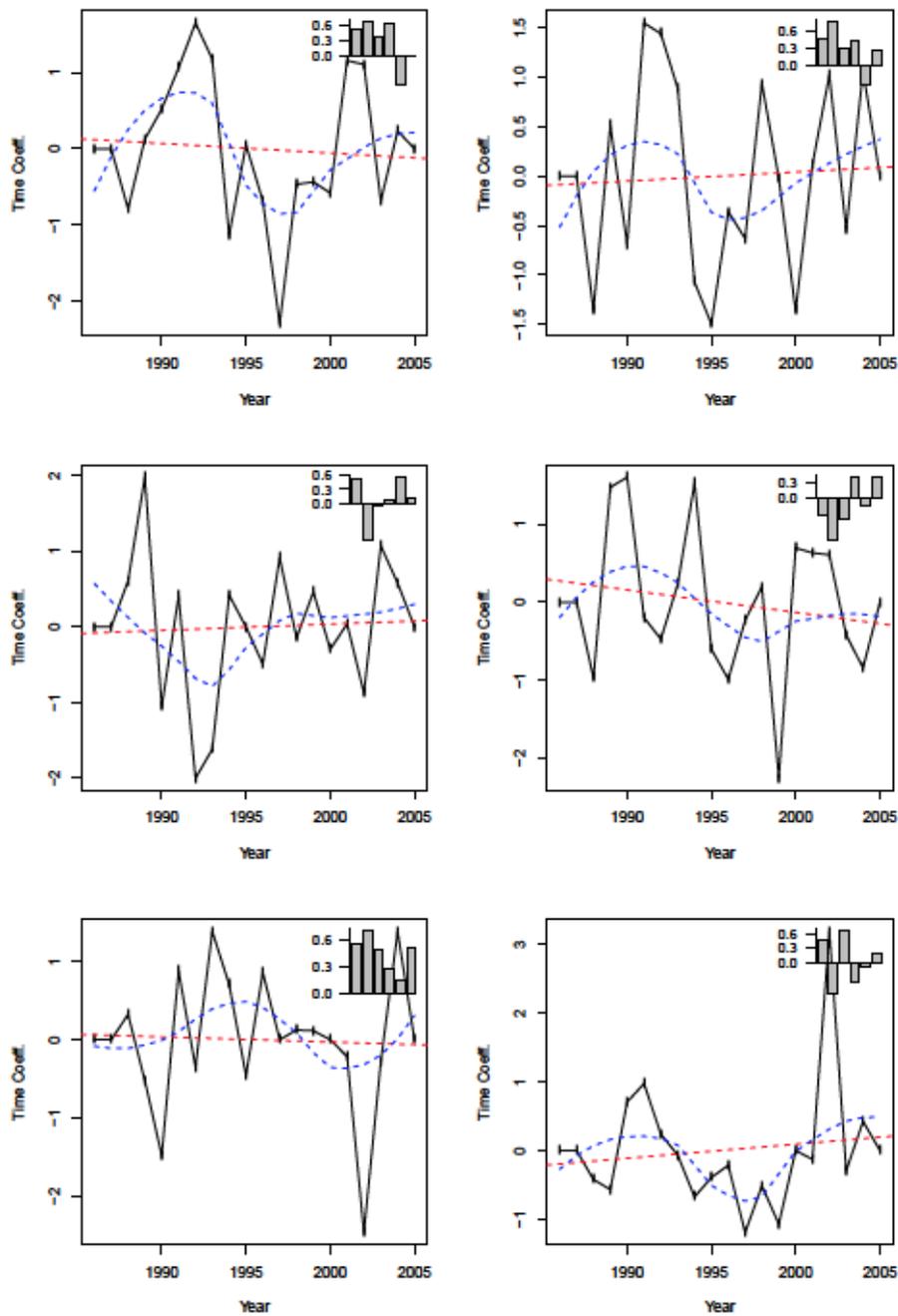


Figure B.34: Bacillariophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

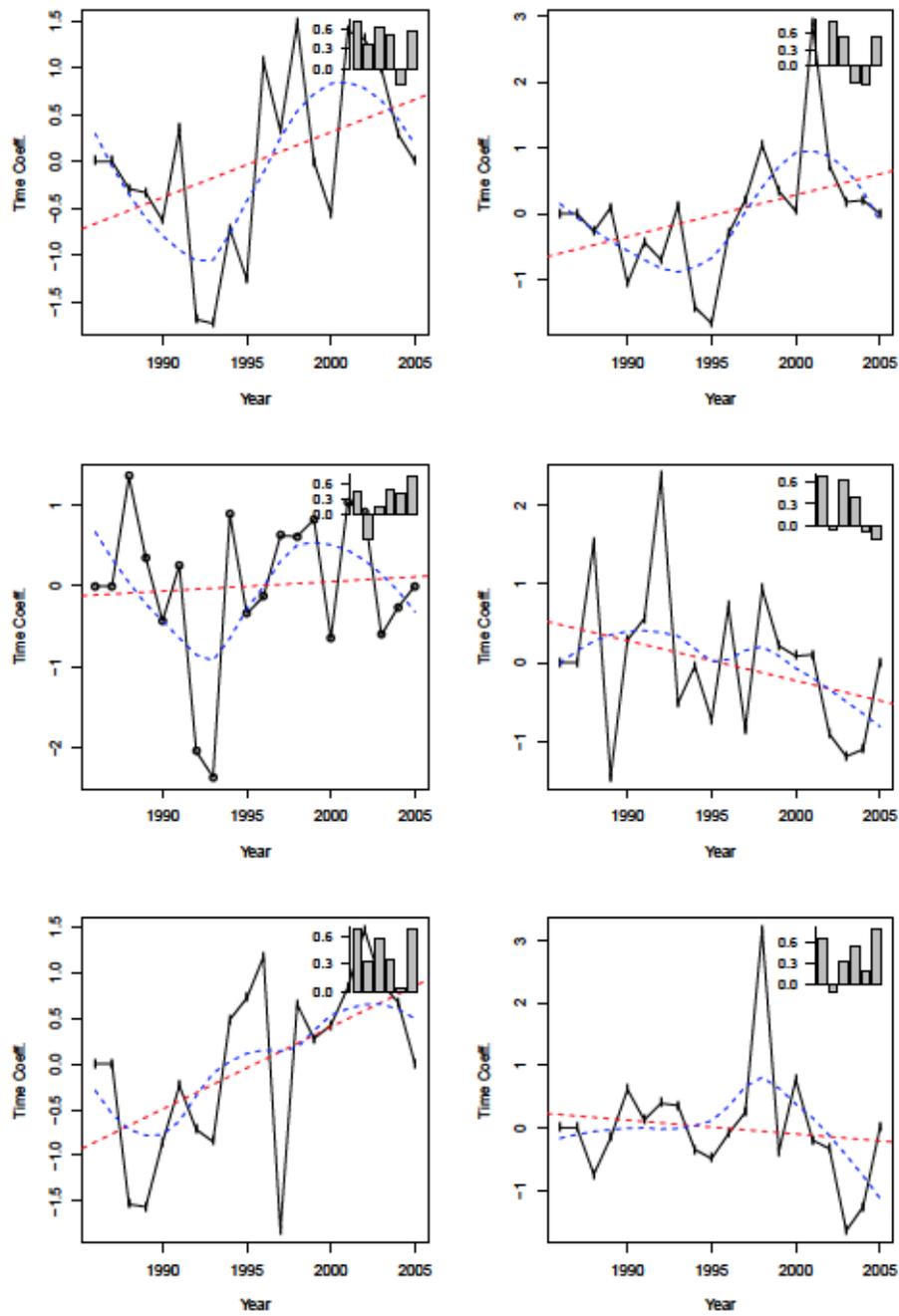


Figure B.35: Cryptophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

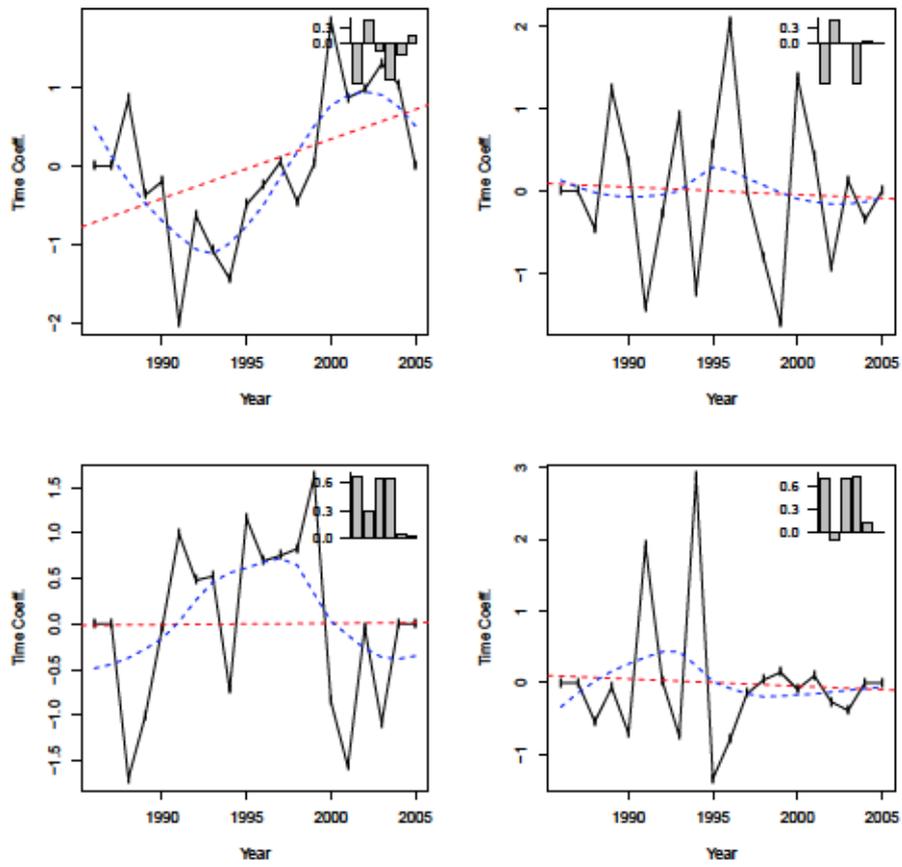


Figure B.36: Cyanophyceae, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Summer (JJA), fall (SON)

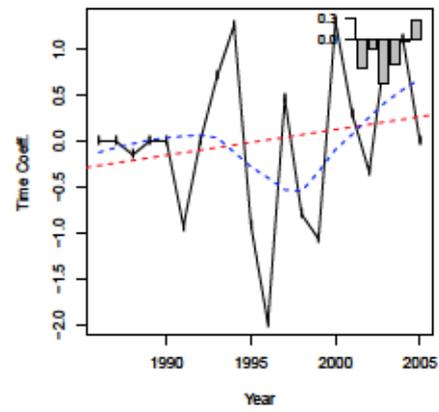


Figure B.37: Dinophyceae, time coefficients of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Biomass; Spring (MAM)

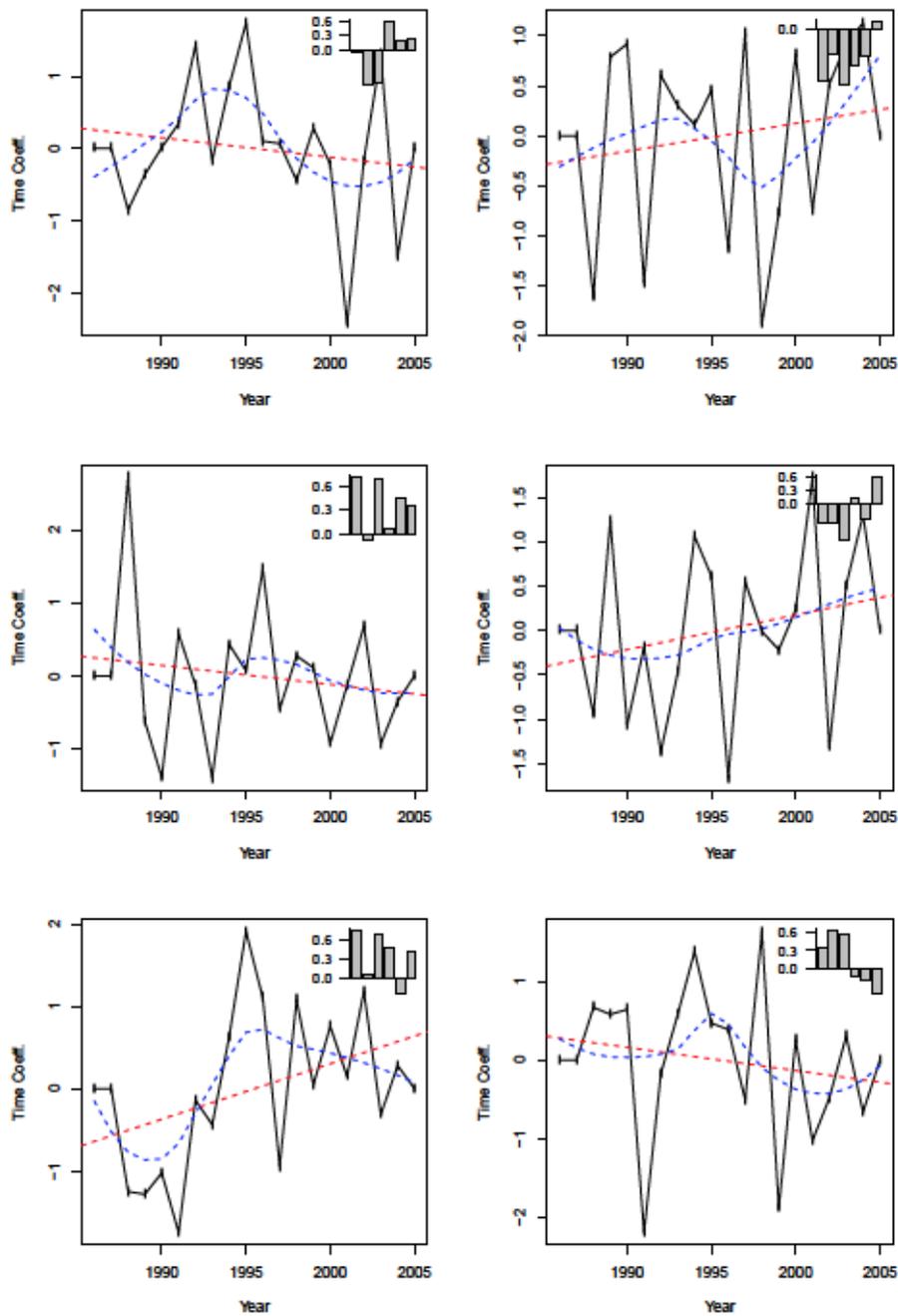


Figure B.38: Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

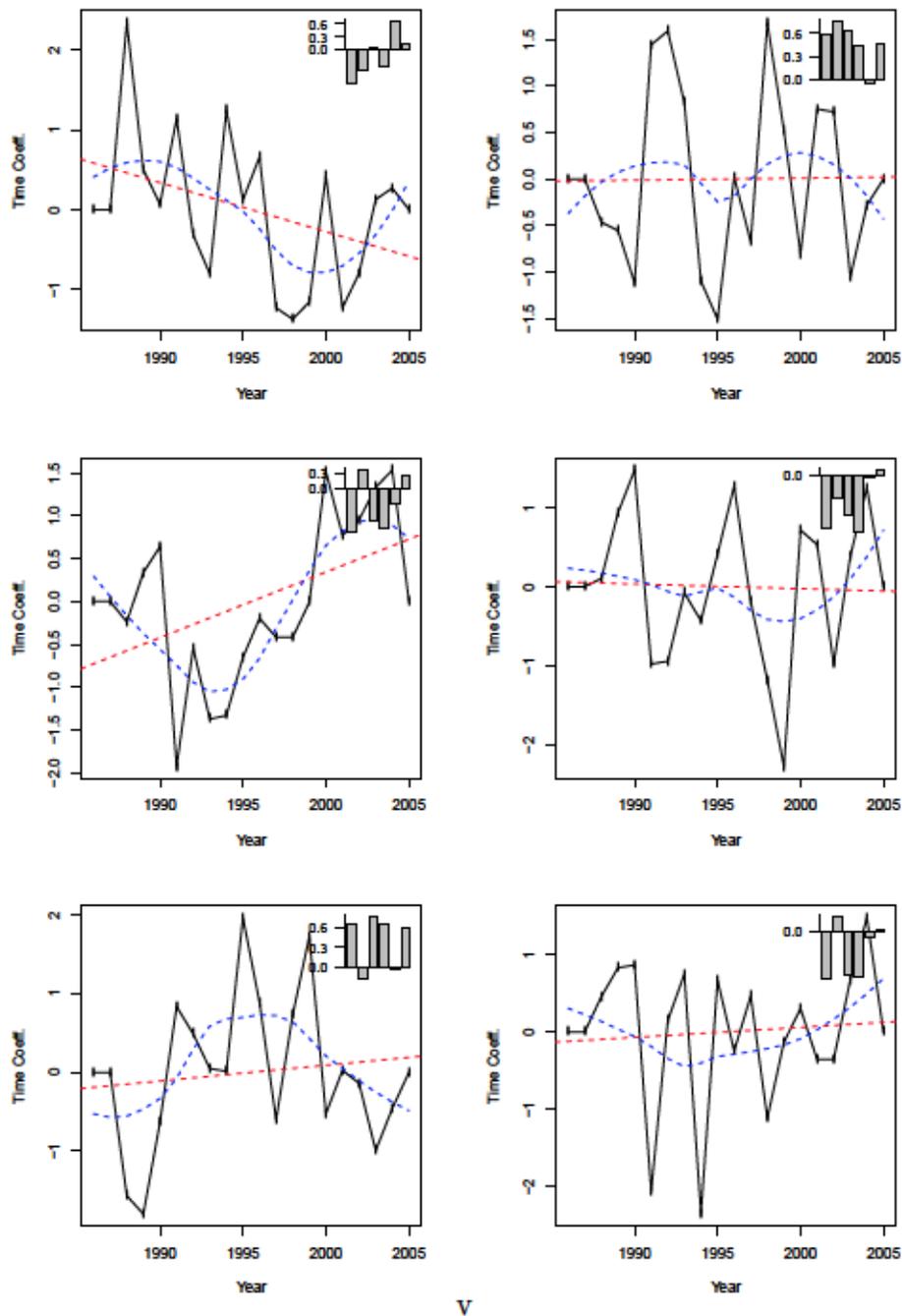


Figure B.39: Total Phytoplankton, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

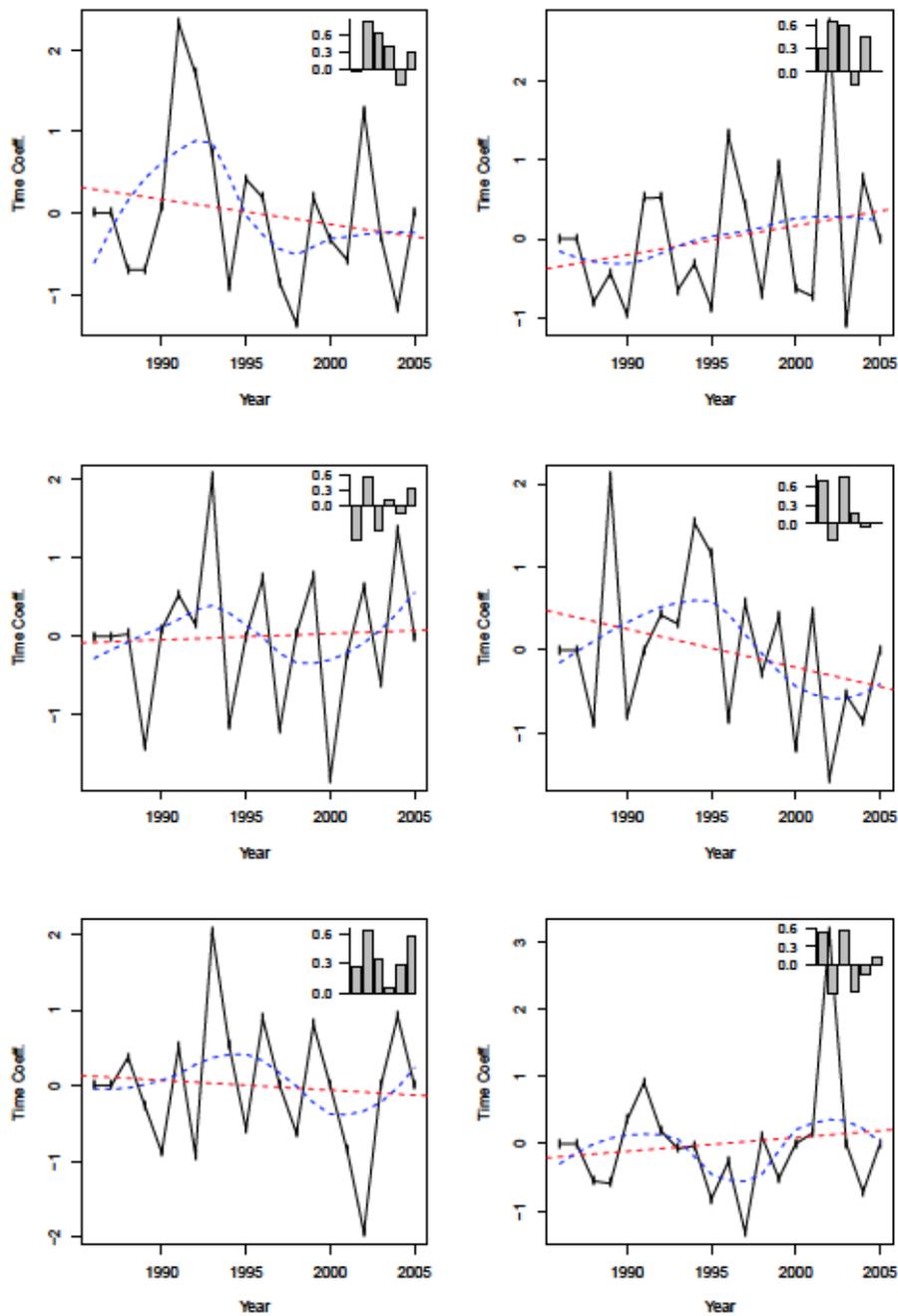


Figure B.40: Ratio Bacillariophyceae:Flagellates, time coefficients and patterns (Barplot) of the 1st EOF (black solid line), linear trend (red, dashed), loess smoother (blue, dash-dotted). Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA)

B.3 Arkona Sea Phytoplankton Downscaling 1979-2005

All results are for downscaling experiments with the 121-filtered BSE as predictor using all 4 EOF of the BSE and 2 EOF of the response variable. Downscaling period is 1979-2005.

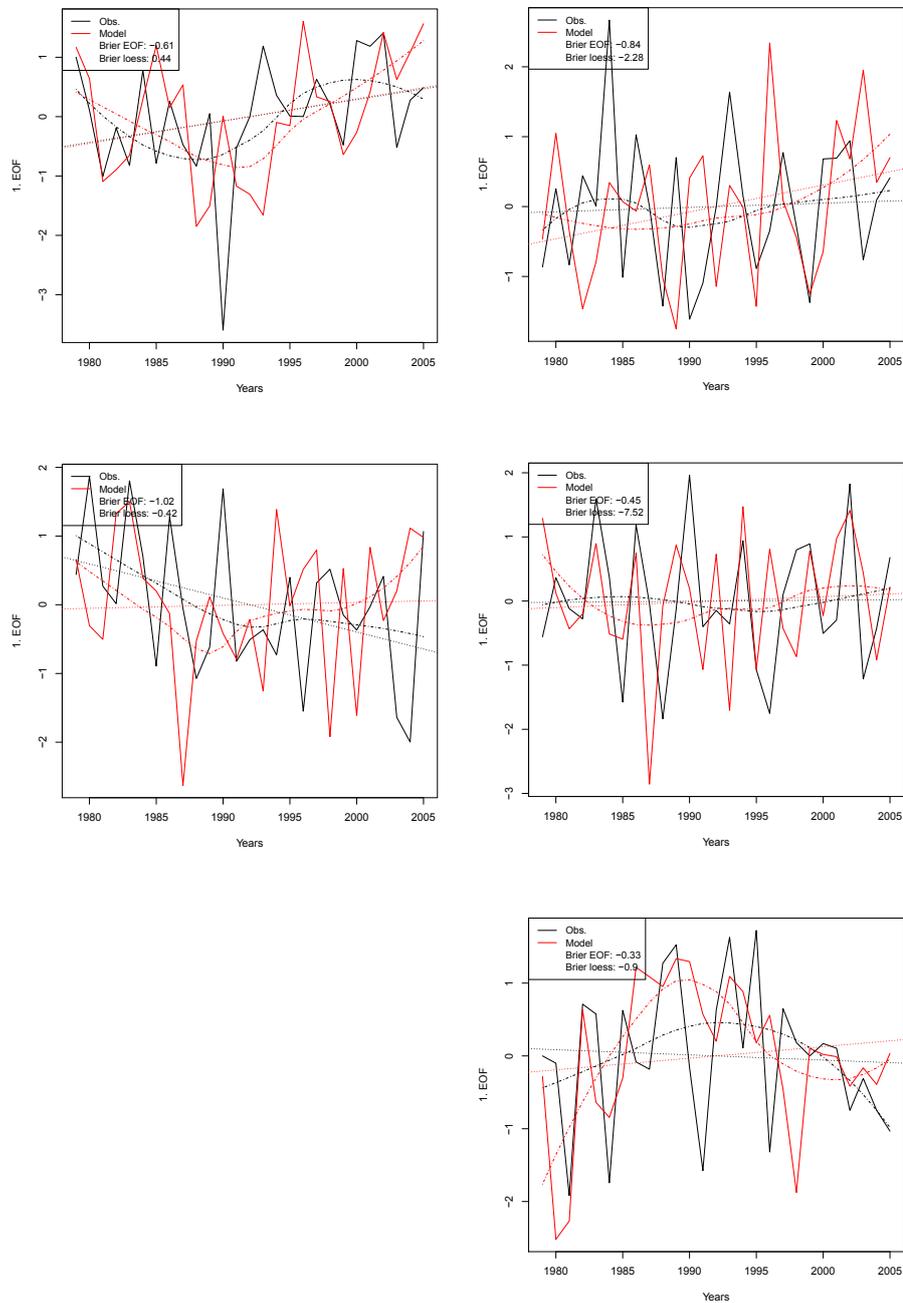


Figure B.41: Bacillariophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

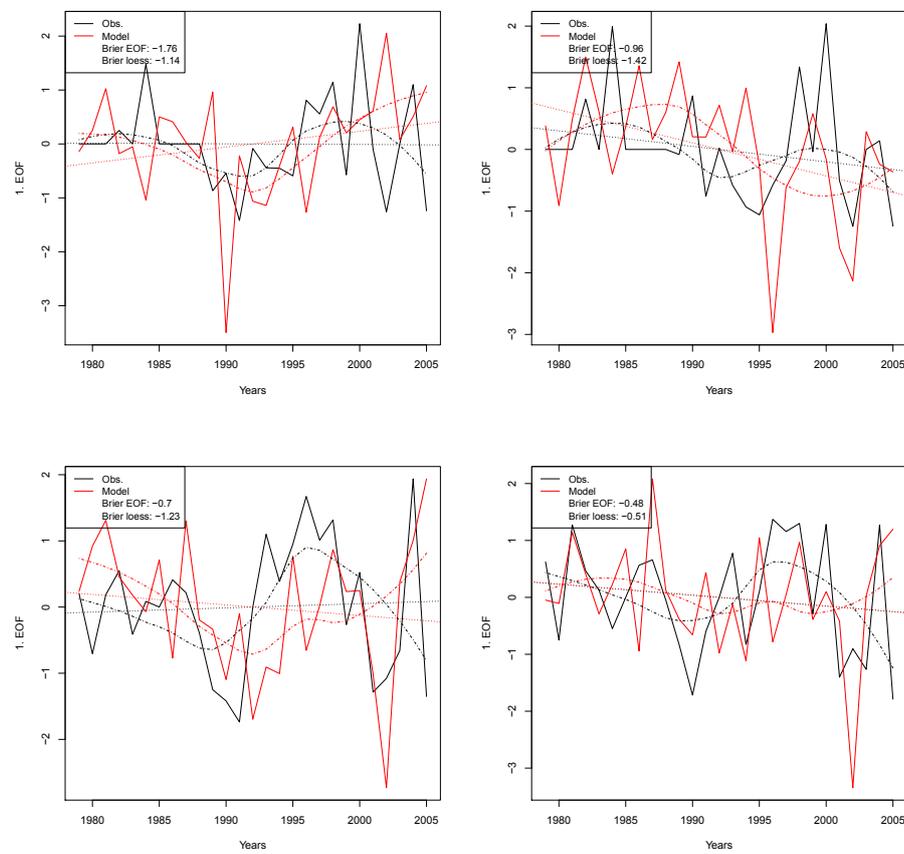


Figure B.42: Chlorophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

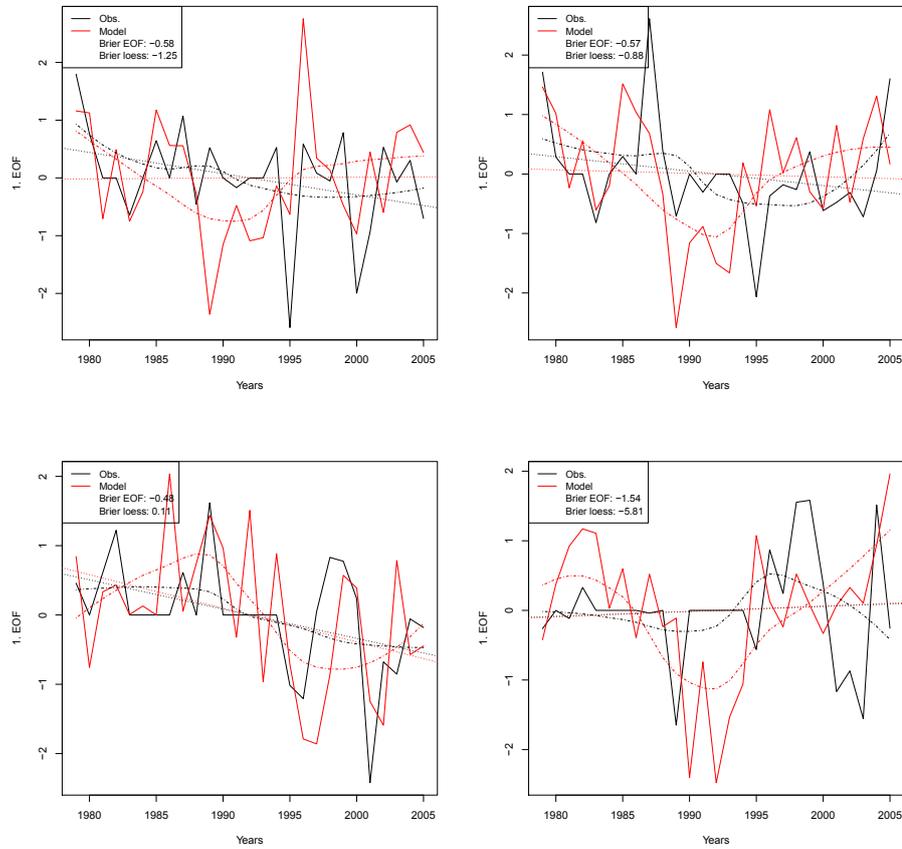


Figure B.43: Chrysophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

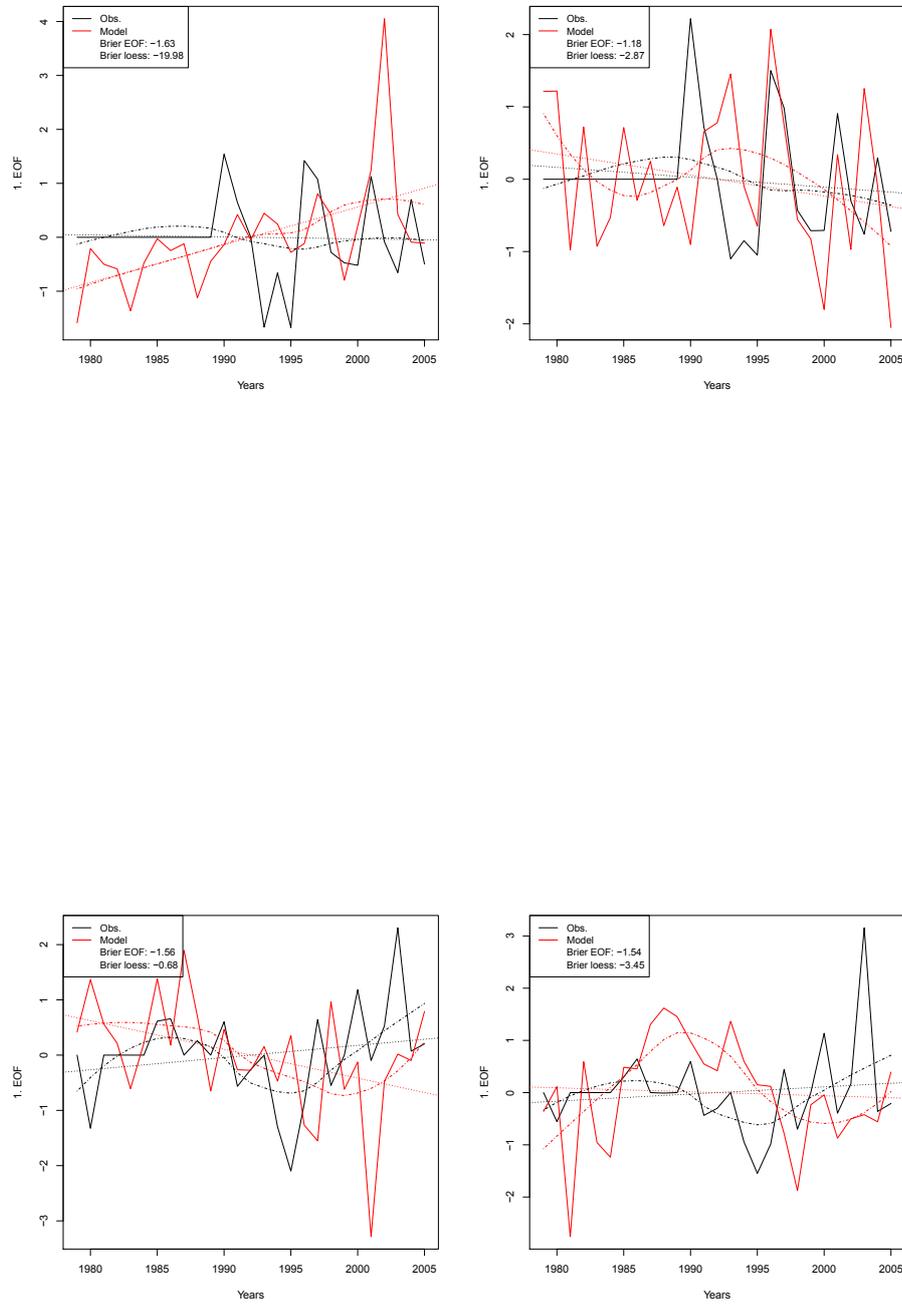


Figure B.44: Craspedophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

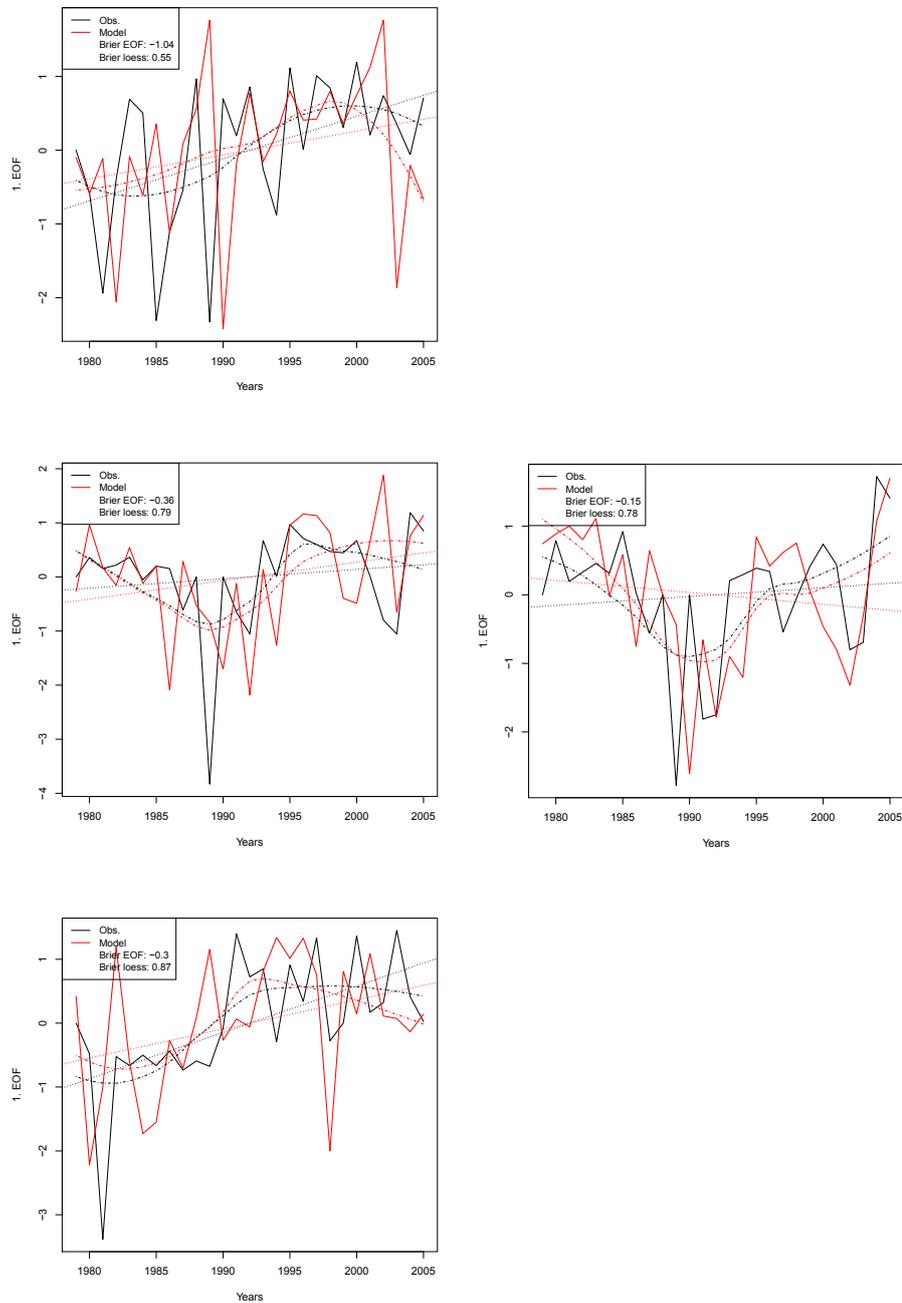


Figure B.45: Cryptophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

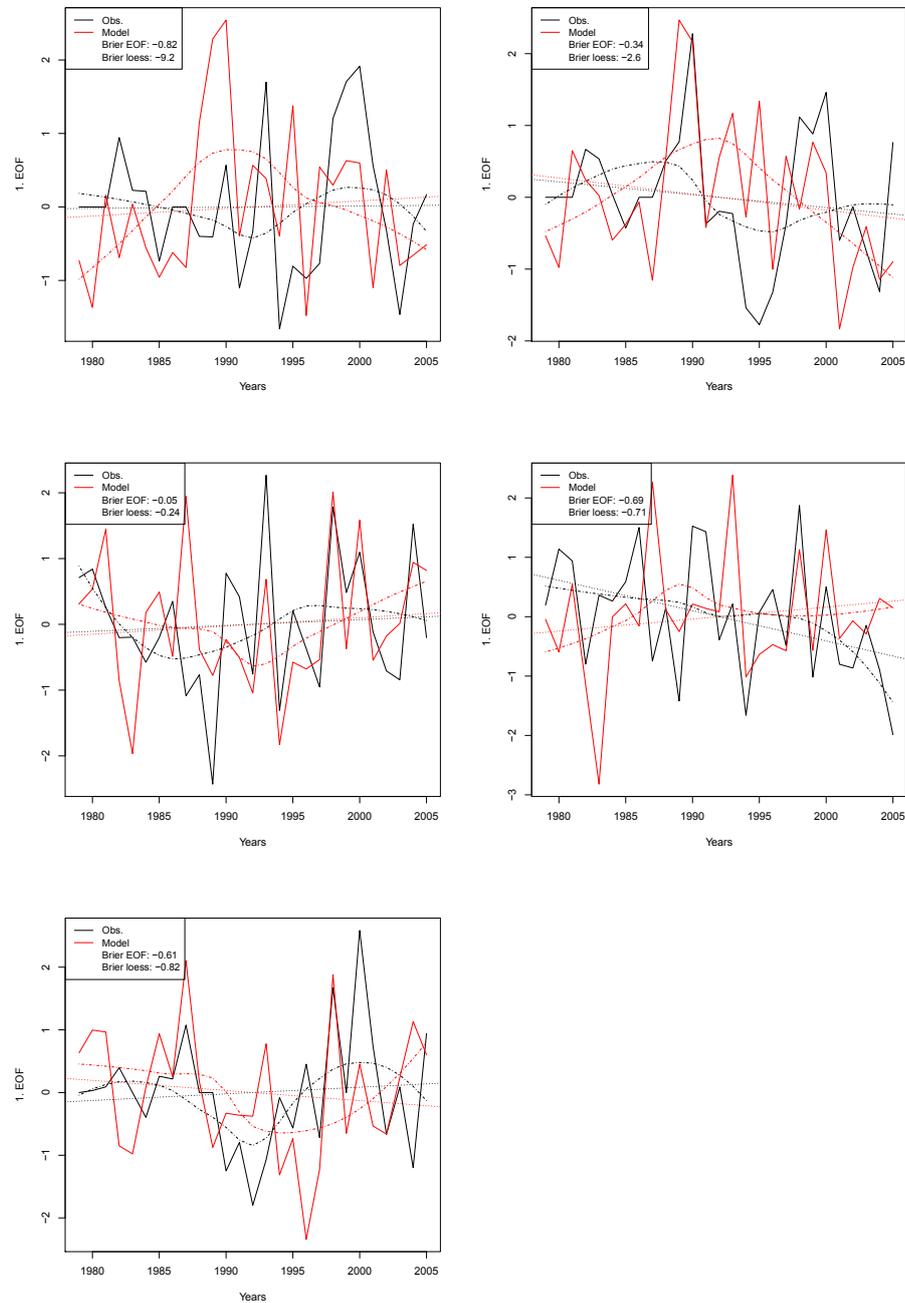


Figure B.46: Cyanophyceae (Cyanobacteria), time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

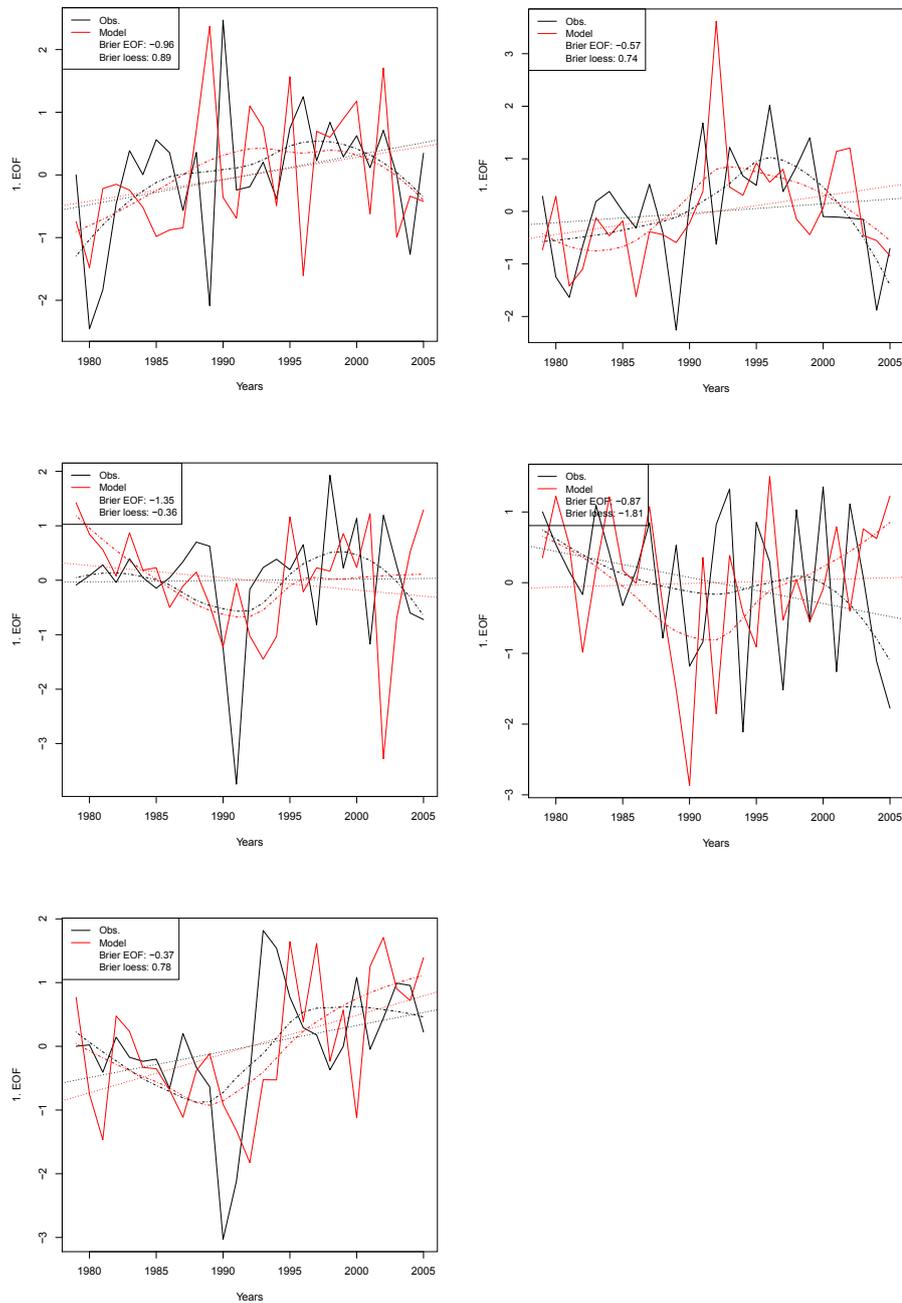


Figure B.47: Dinophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

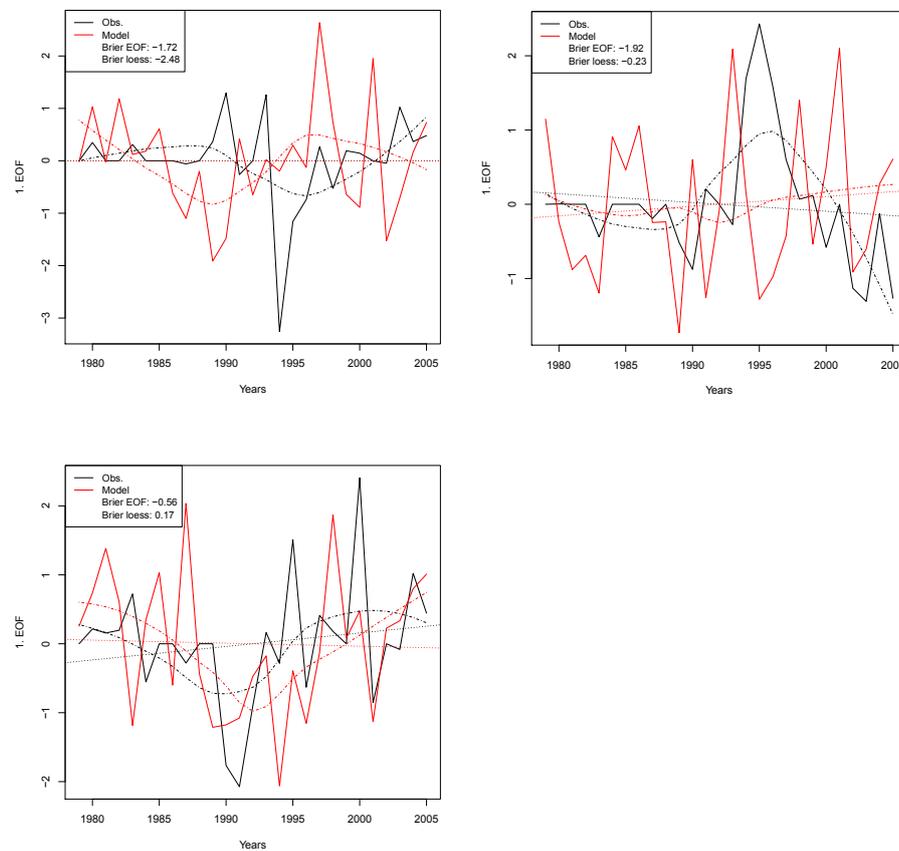


Figure B.48: Euglenophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

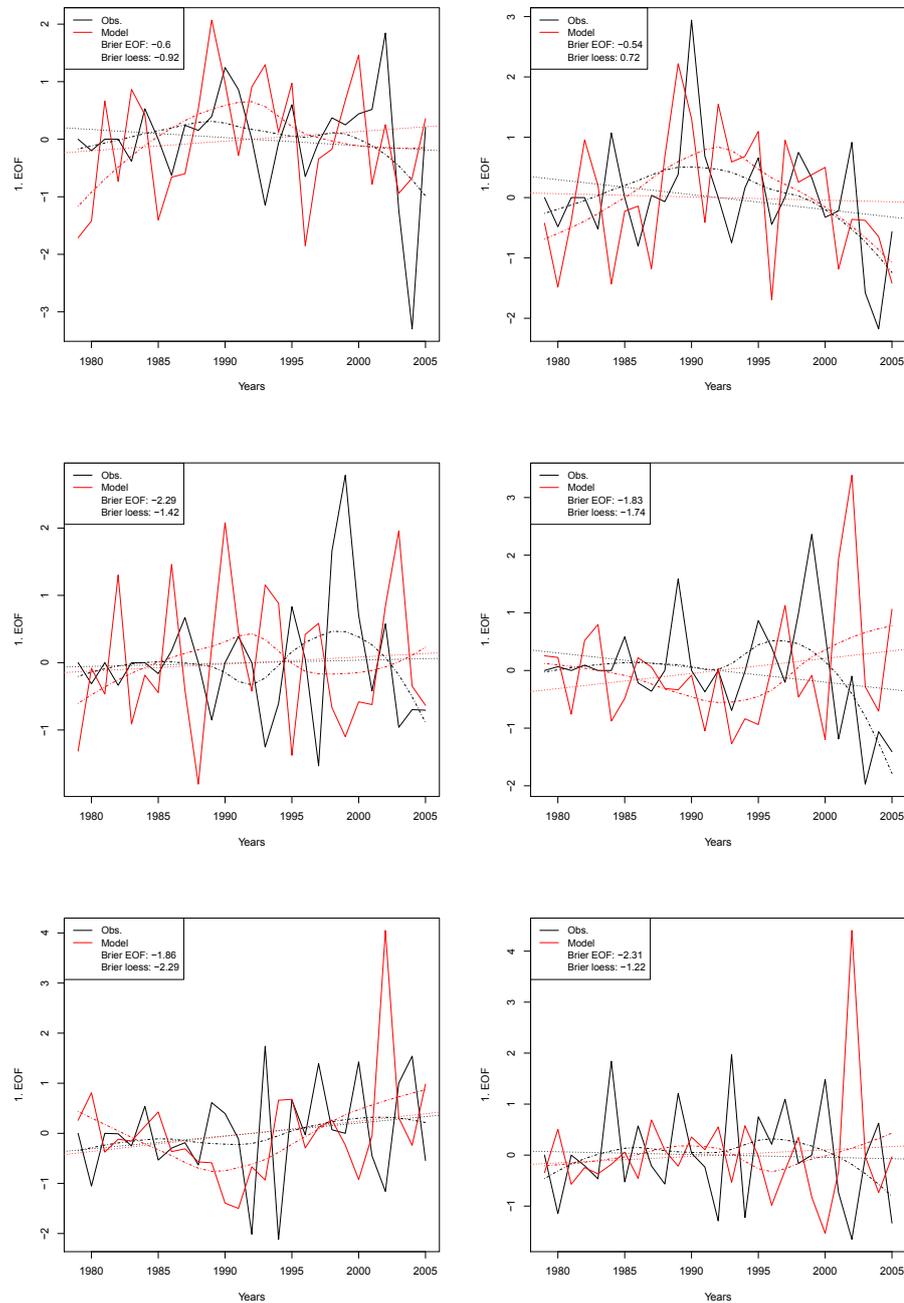


Figure B.49: Prasinophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

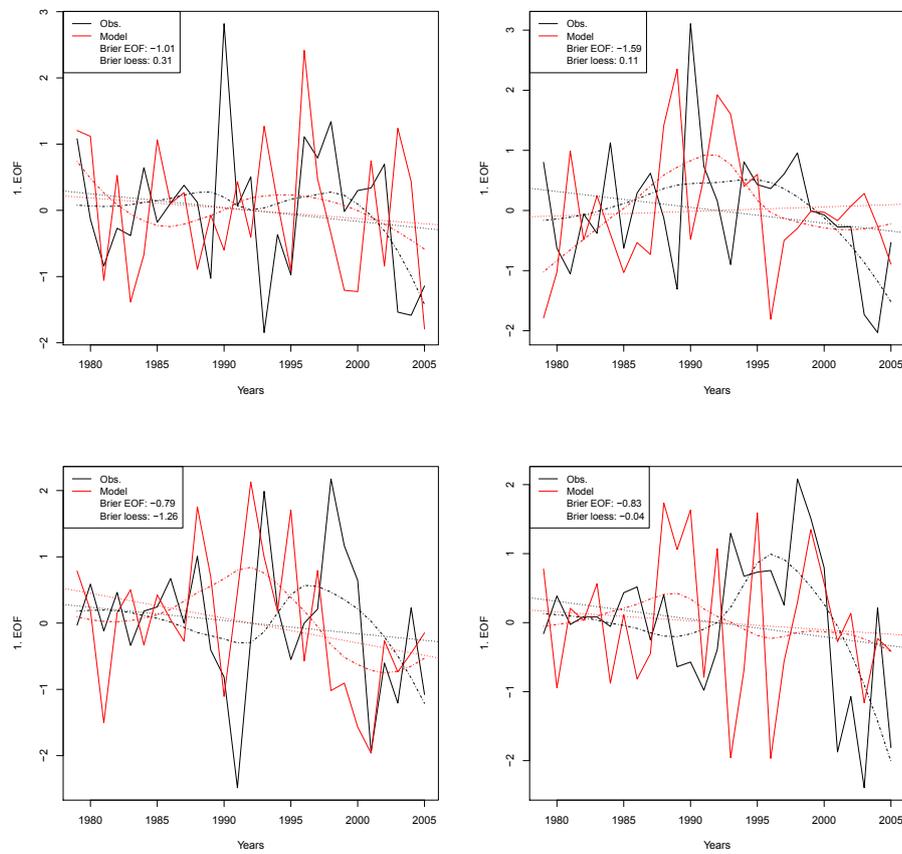


Figure B.50: Flagellates, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

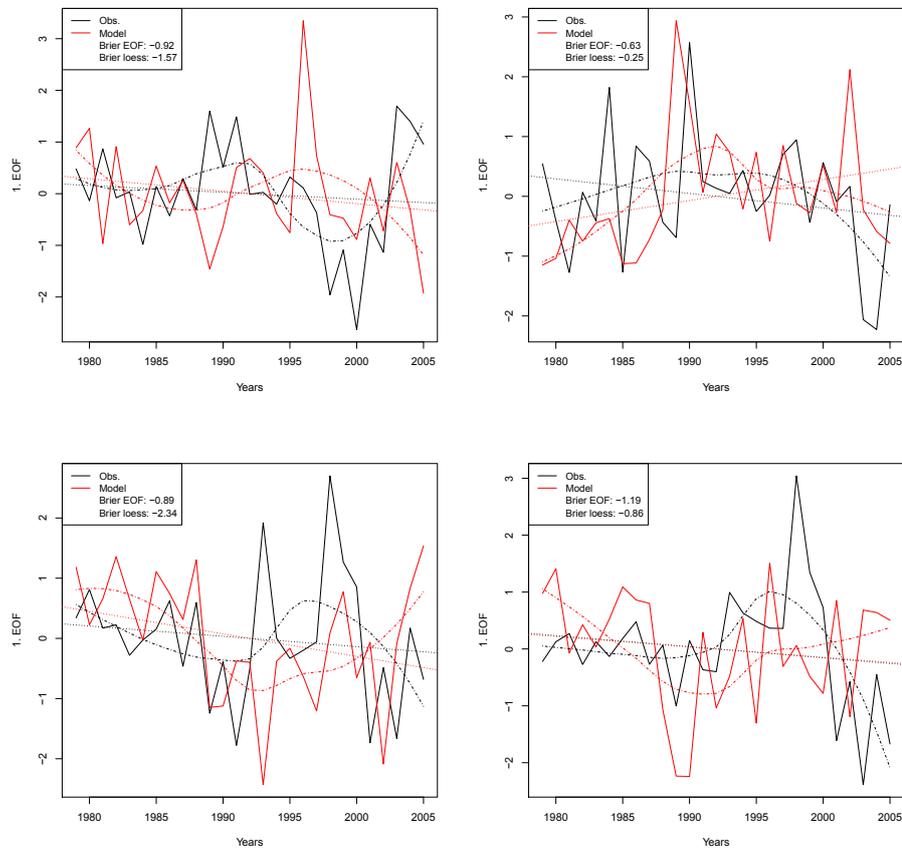


Figure B.51: Total Phytoplankton, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

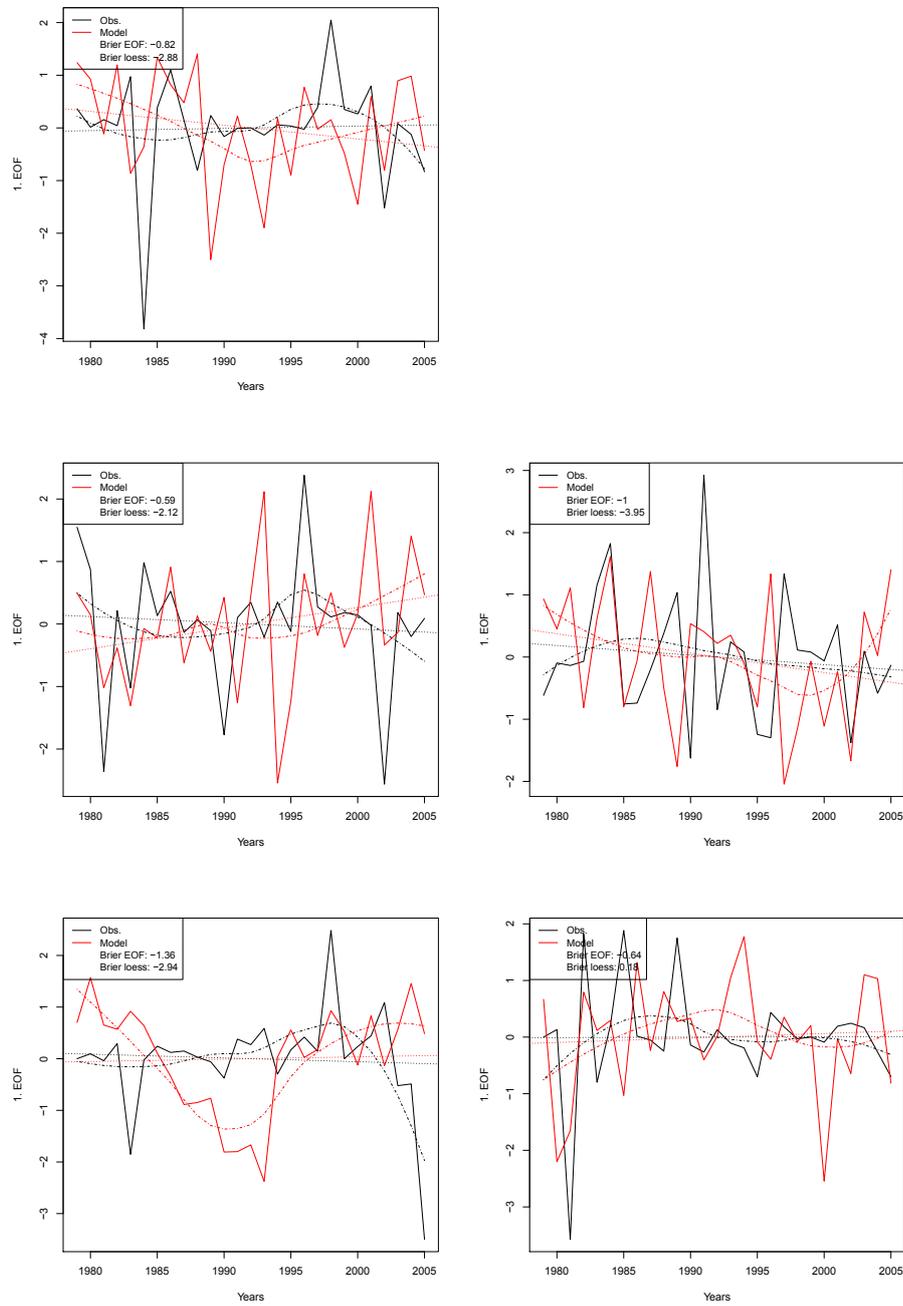


Figure B.52: Ratio Bacillariophyceae:Flagellates, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

B.4 Arkona Sea Phytoplankton Downscaling 1979-2000

All results are for downscaling experiments with the 121-filtered BSE as predictor using all 4 EOF of the BSE and 2 EOF of the response variable. Downscaling period is 1979-2000.

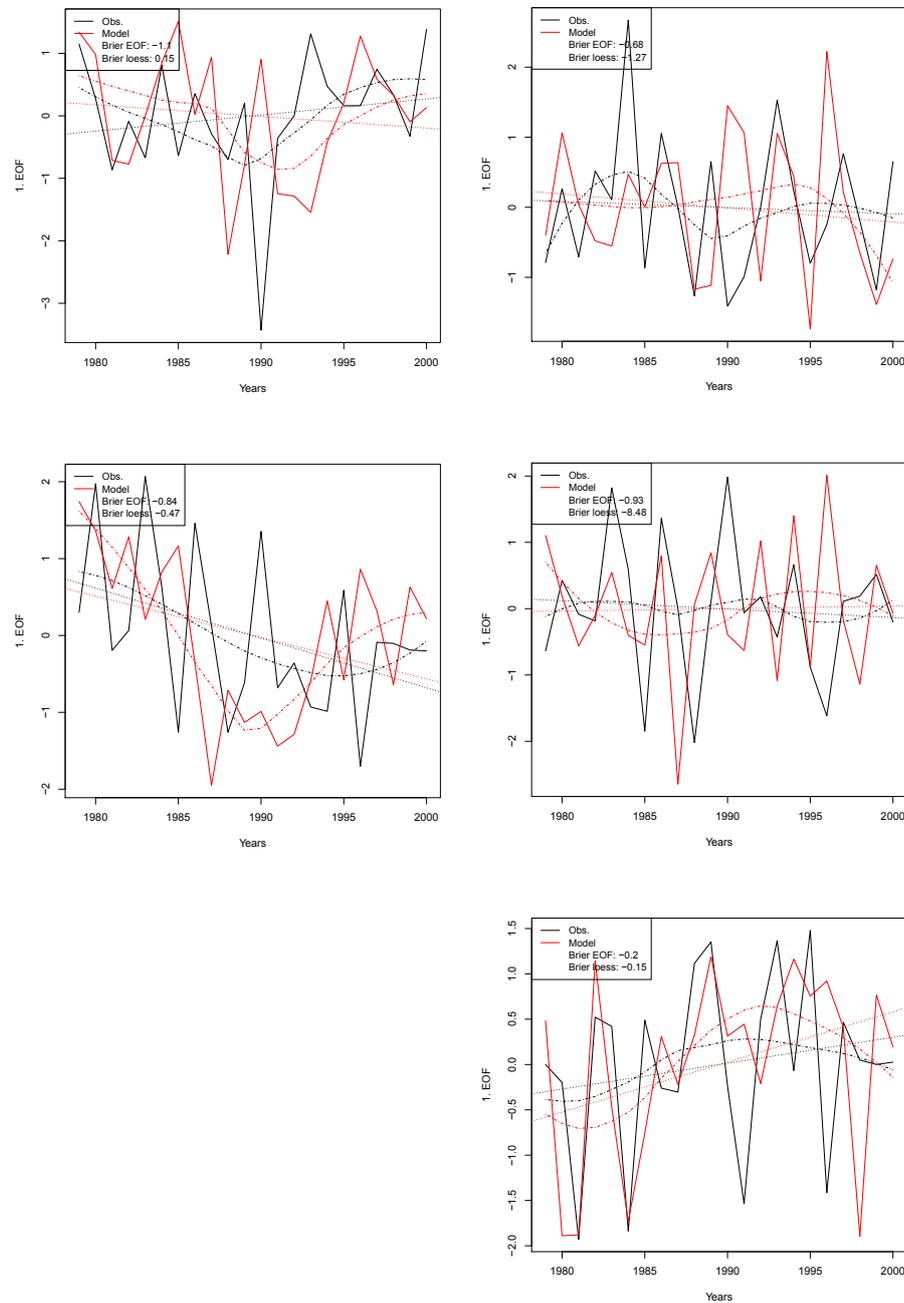


Figure B.53: Bacillariophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

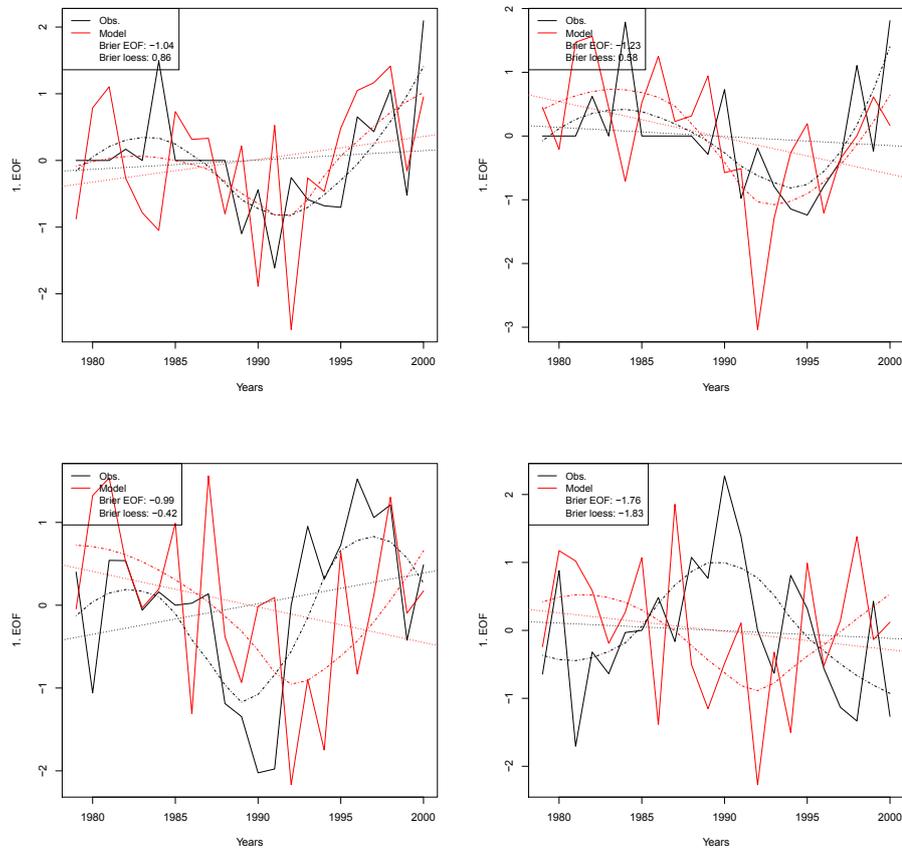


Figure B.54: Chlorophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

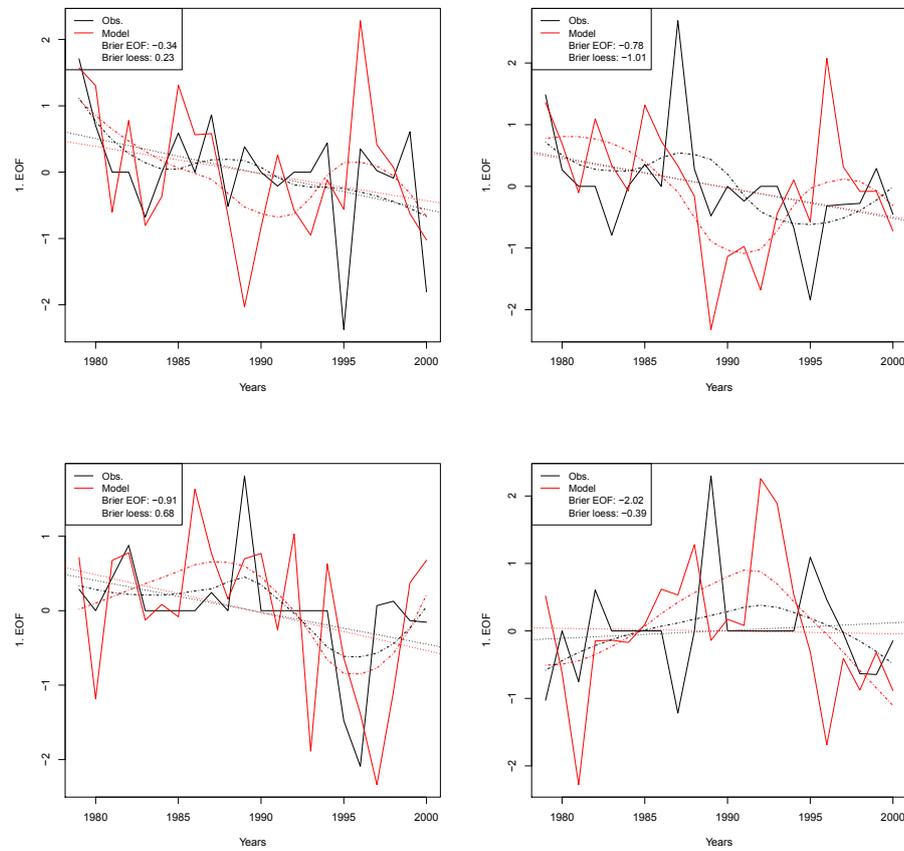


Figure B.55: Chrysophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

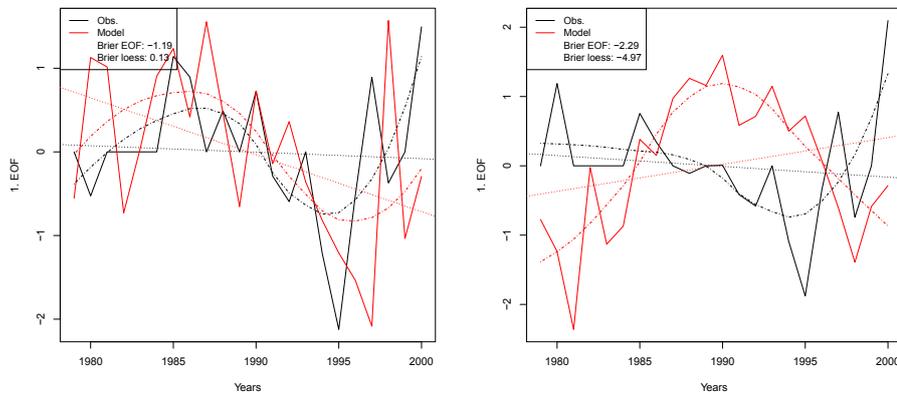


Figure B.56: Craspedophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

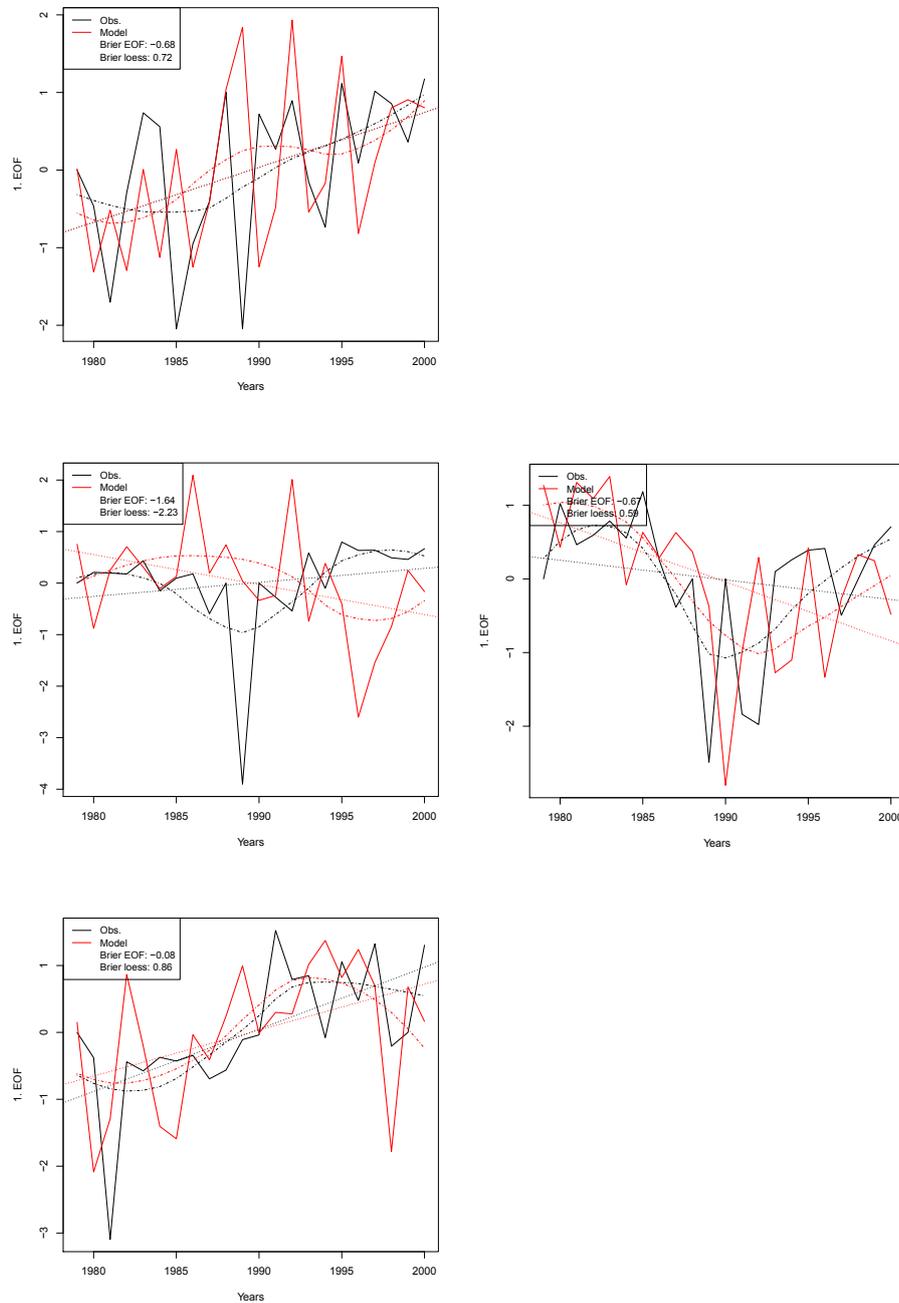


Figure B.57: Cryptophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

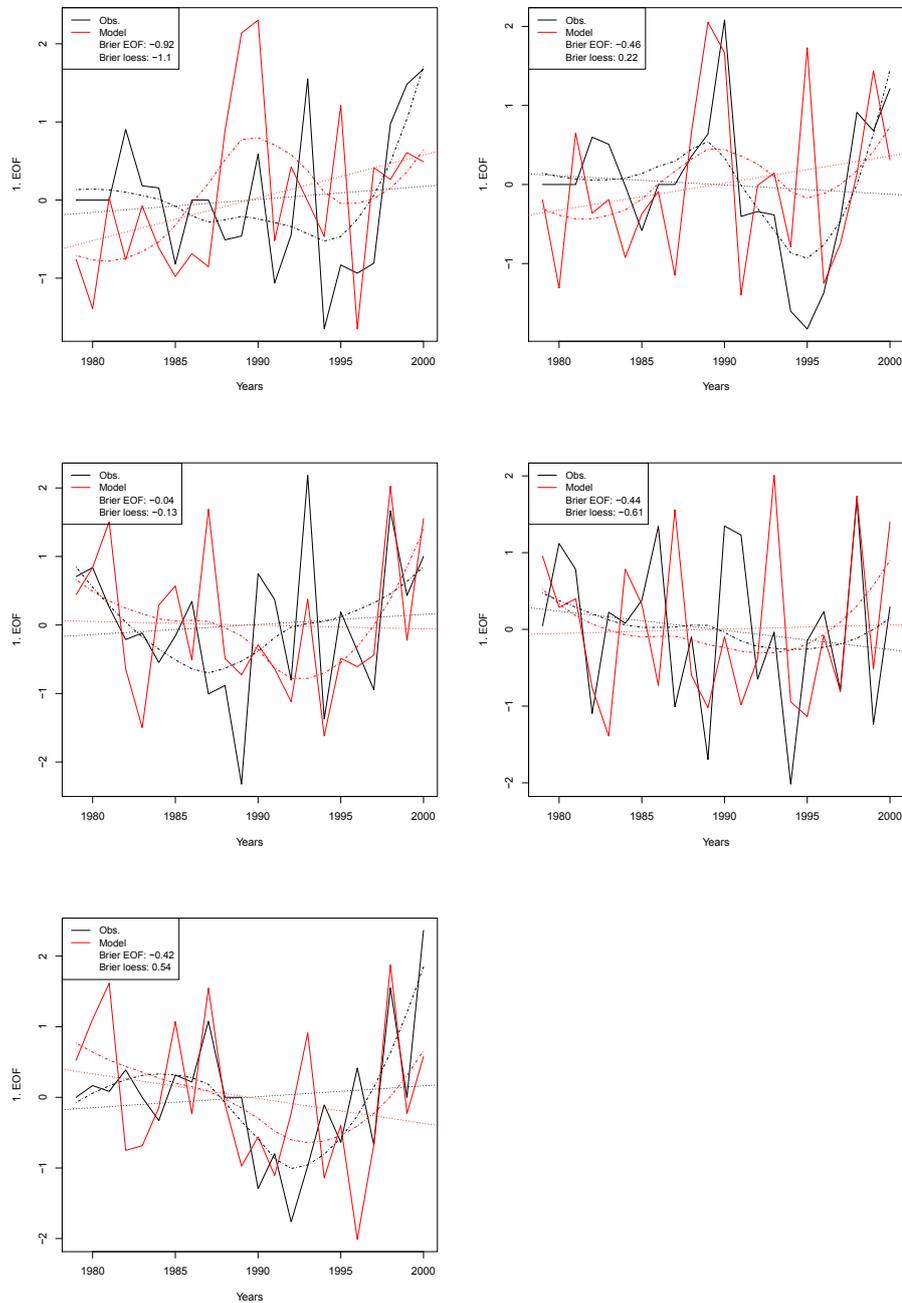


Figure B.58: Cyanophyceae (Cyanobacteria), time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

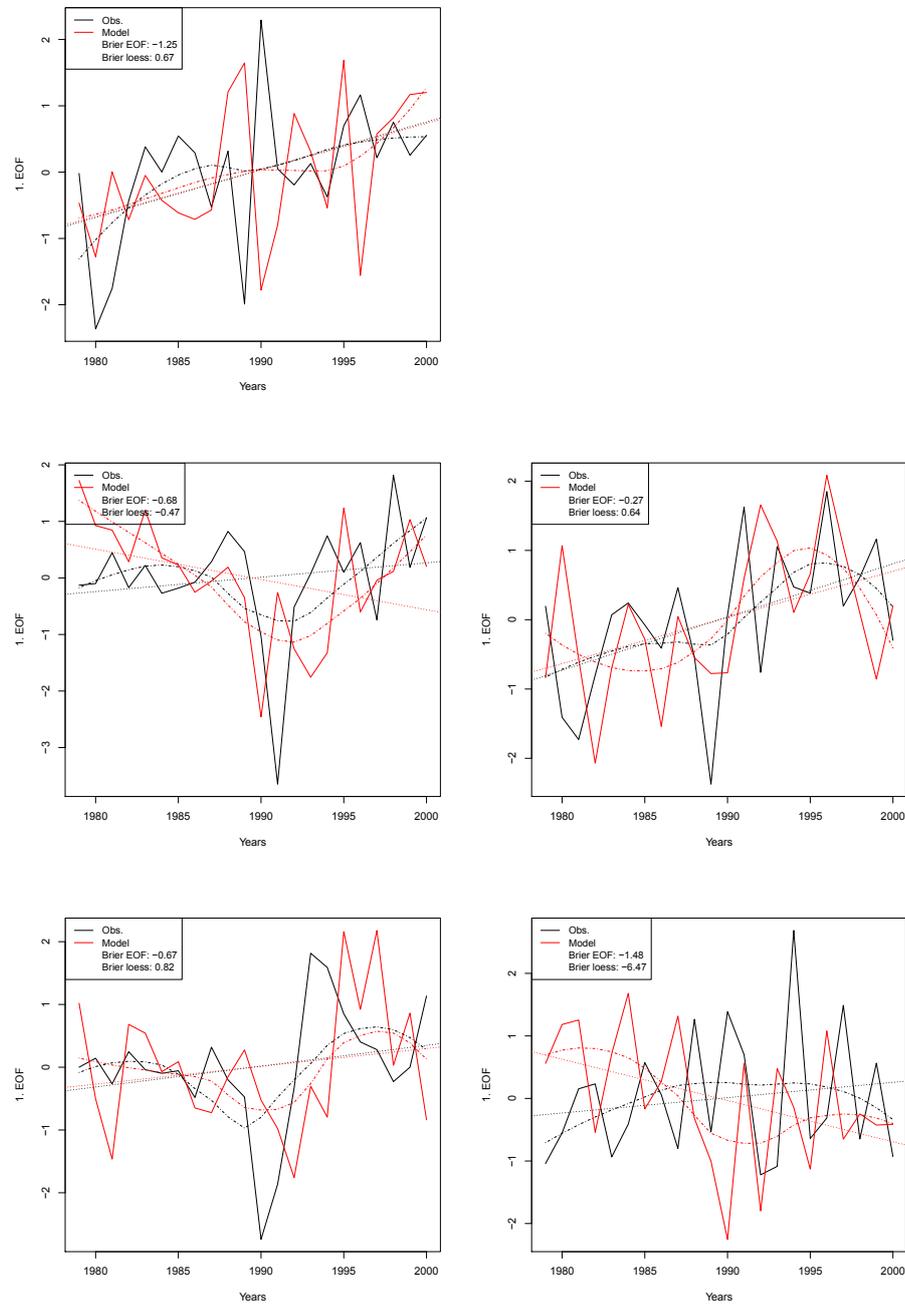


Figure B.59: Dinophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

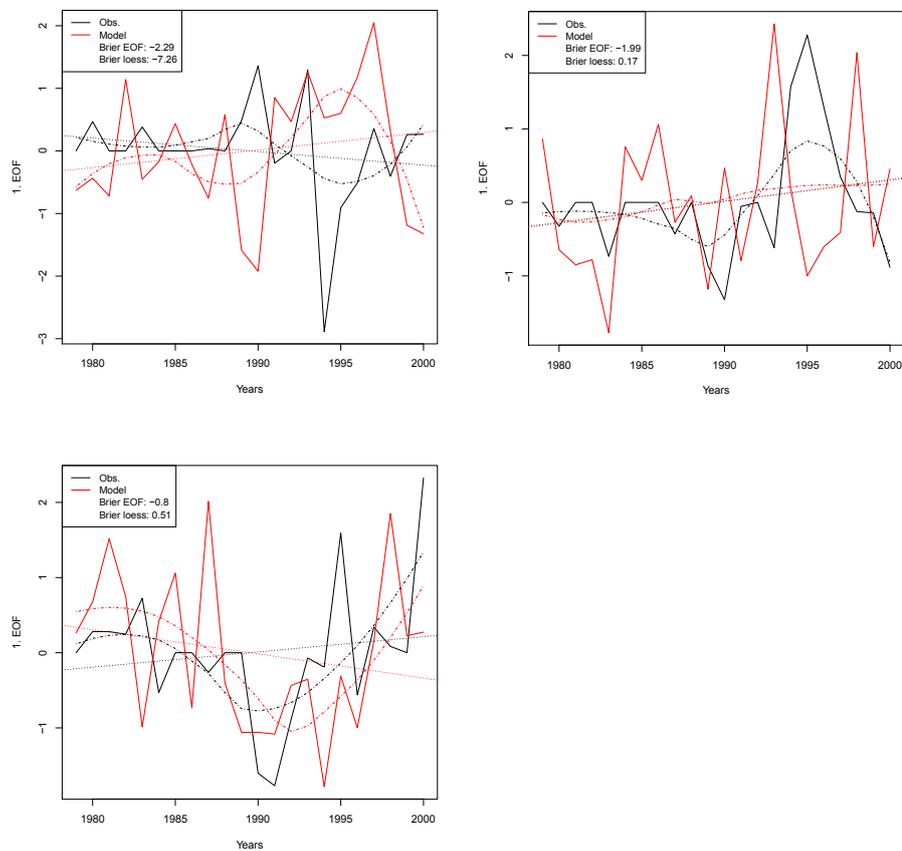


Figure B.60: Euglenophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

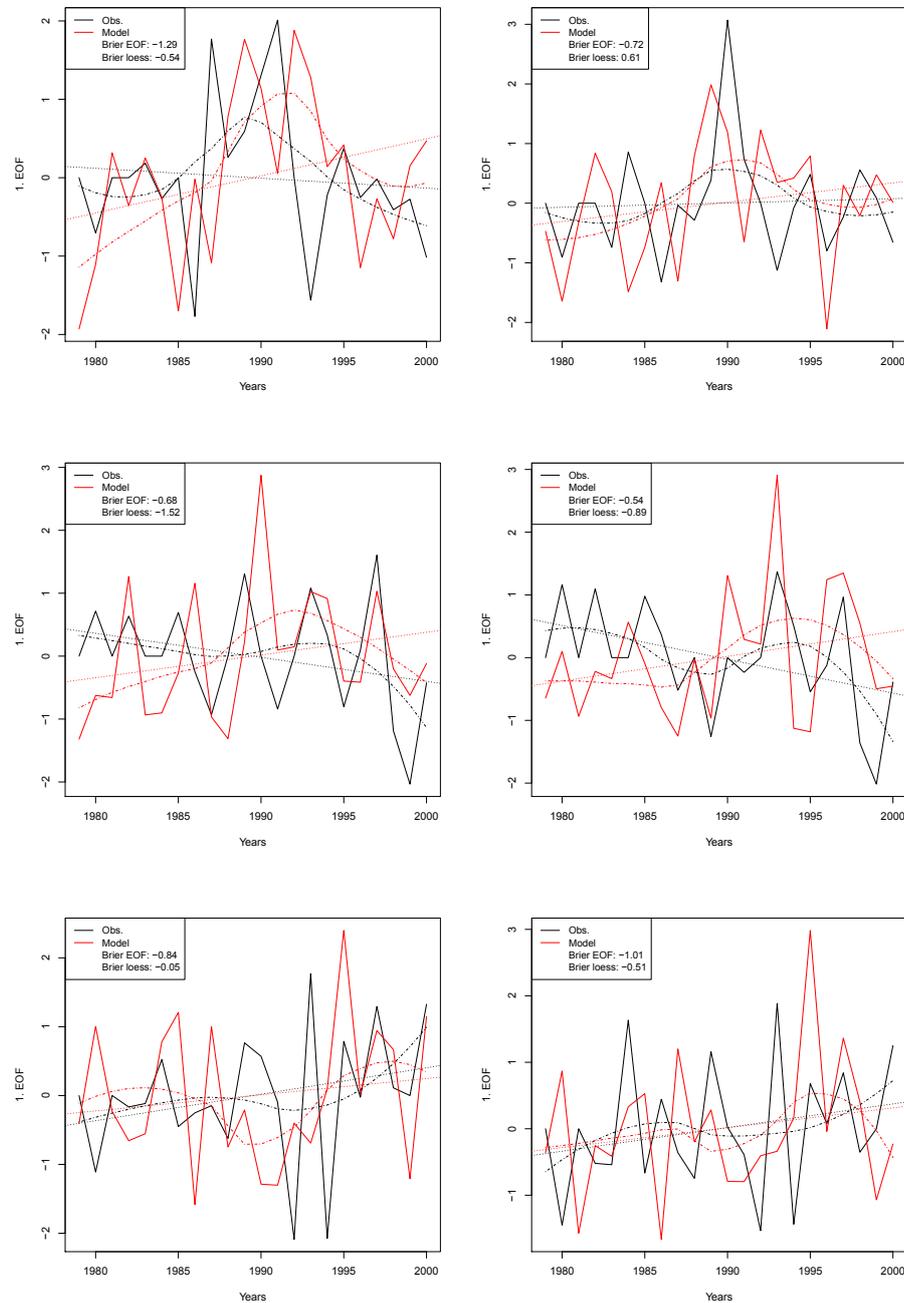


Figure B.61: Prasinophyceae, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

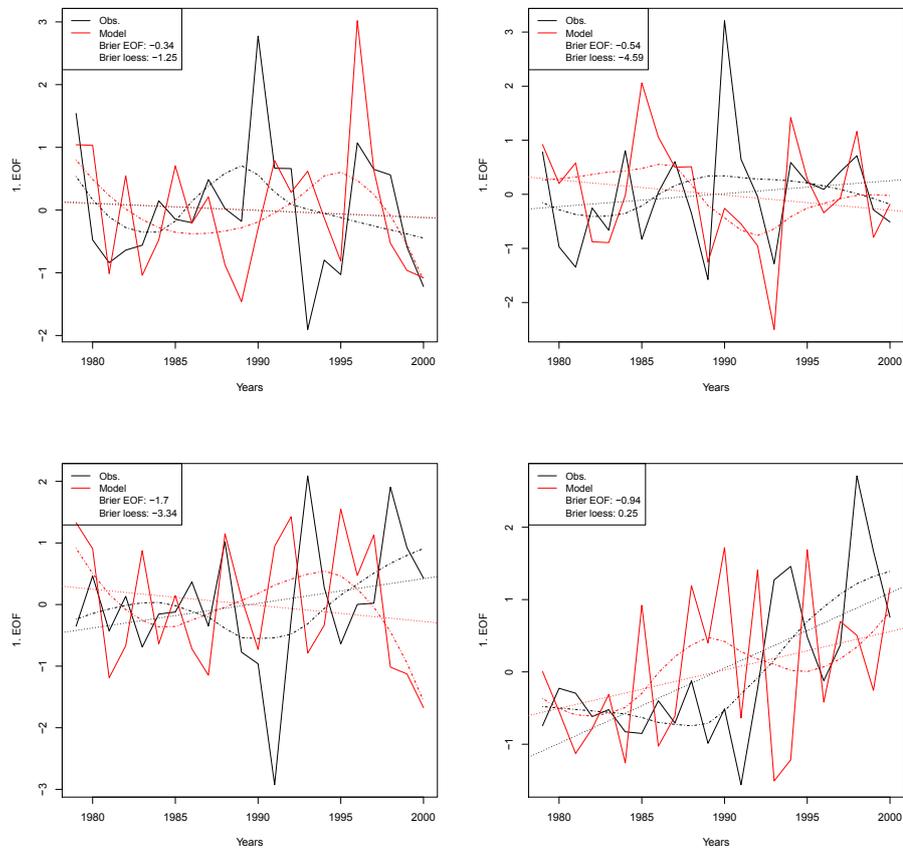


Figure B.62: Flagellates, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother (dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

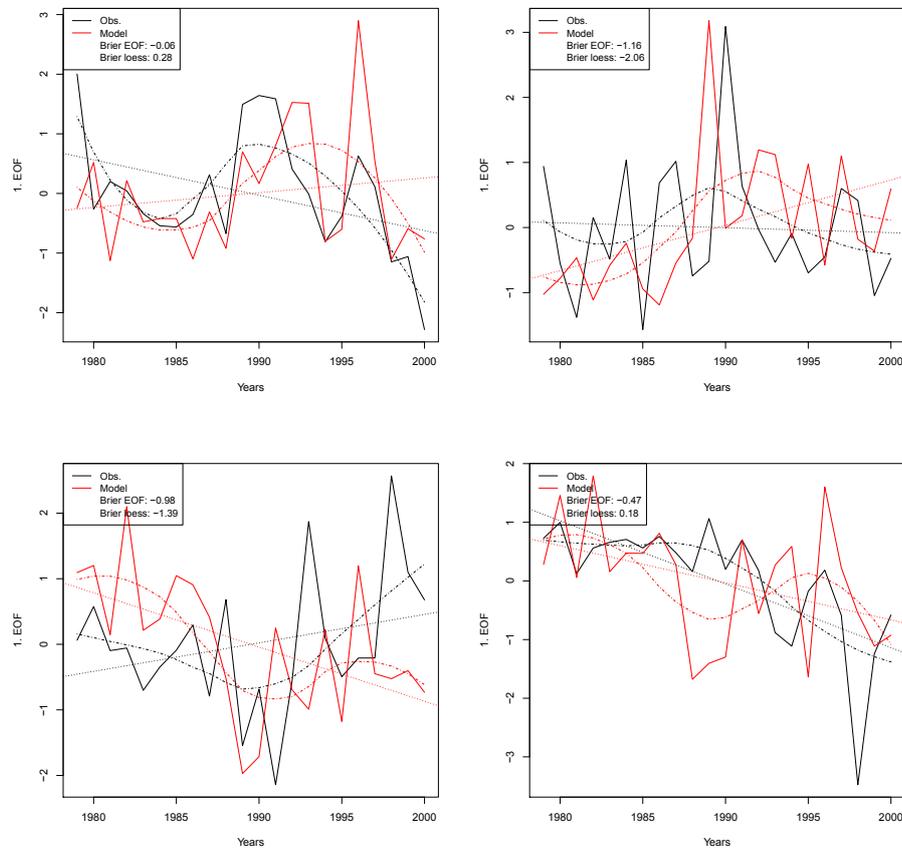


Figure B.63: Total Phytoplankton,time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)

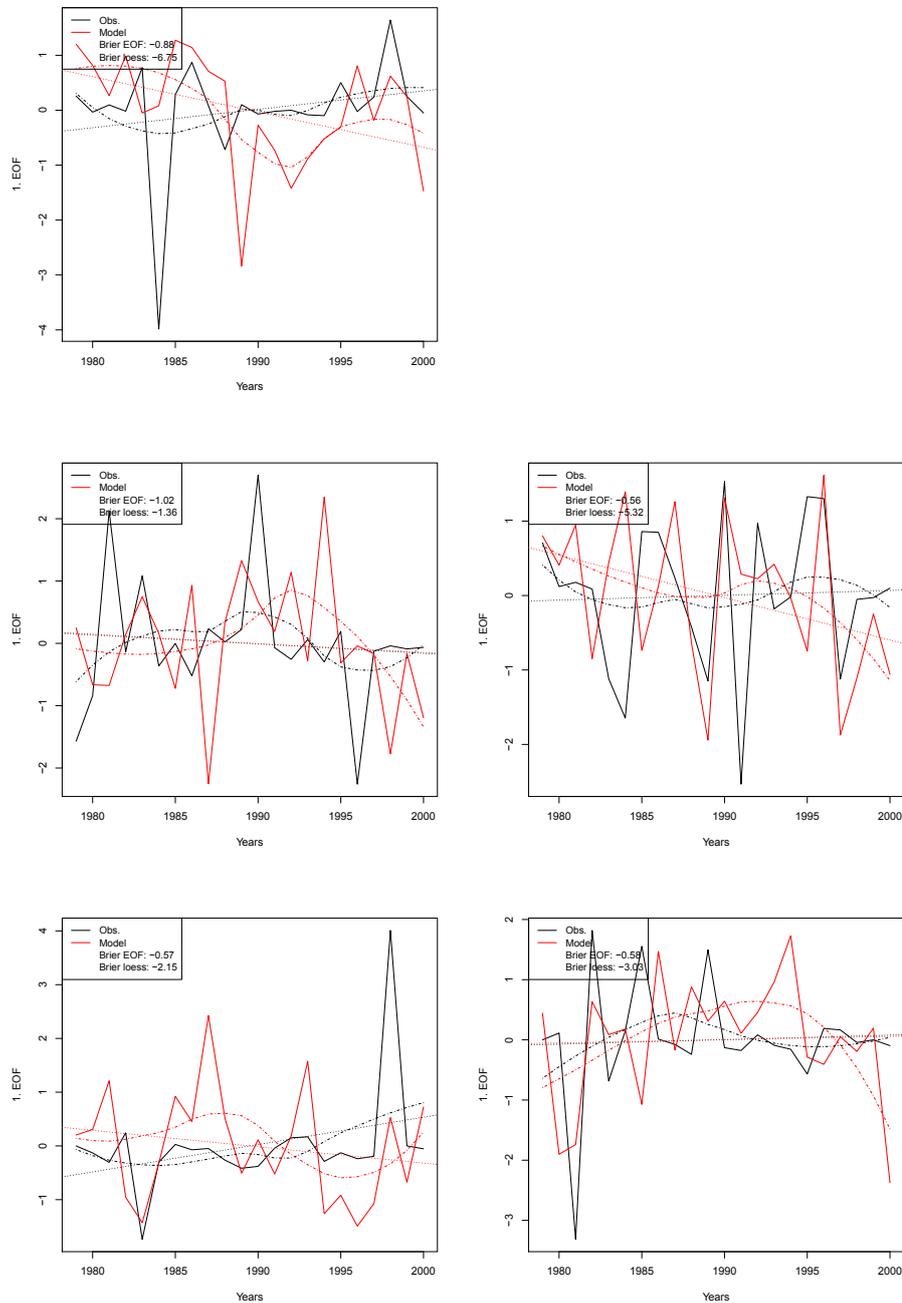


Figure B.64: Ratio Bacillariophyceae:Flagellates, time coefficients of the 1st EOF (solid line), linear trend (dotted), loess smoother(dash-dotted). Black: Observations, Red: Fitted. Left column: abundance, right column: biomass; From top: Spring (MAM), summer (JJA), fall (SON)