Interreg V A project „Wasserqualität - Waterkwaliteit"

Harmonisation of the Phytoplankton Assessment in the German and Dutch Wadden Sea
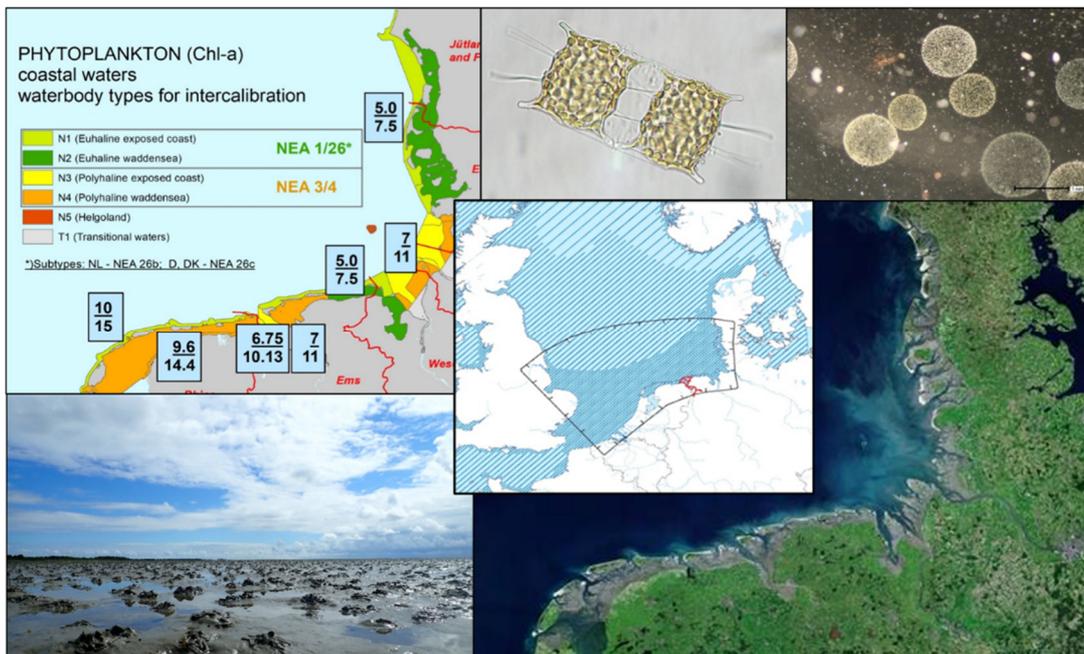
Harmonisation of the Phytoplankton Assessment in the German and Dutch Wadden Sea

Synthesis Report

Authors

Lena Rönn[1], Josie Antonucci di Carvalho[5], Anouk Blauw[4], Helmut Hillebrand[5], Onur Kerimoglu[7], Hermann Lenhart[6], Theo Prins[4], Gholamreza Shiravani[2], Laura Tack[3], Daniel Thewes[6], Tineke Troost[4]

Contributing authors:

Jürgen Knaack[1], Gerrit Niebeek[3], Michael Hanslik[1], Kerstin Kolbe[1], Lisa Schneider[4], Annelotte van der Linden[4], Lauriane Vilmin[4], Andreas Wurpts[2]

List of project partners

[1]Lead Partner: NLWKN – Betriebsstelle Brake-Oldenburg

Niedersächsischer Landesbetrieb für Wasserwirtschaft,
Küsten- und Naturschutz (NLWKN)
Betriebsstelle Brake-Oldenburg
Im Dreieck 12
26127 Oldenburg, DE

Project participants: Jürgen Knaack, Lena Rönn, Kerstin Kolbe, Michael Hanslik, Jan Witt

[2]NLWKN – Forschungsstelle Küste (FSK)

Niedersächsischer Landesbetrieb für Wasserwirtschaft,
Küsten- und Naturschutz (NLWKN)
Betriebsstelle Norden-Norderney
Jahnstr. 1
26506 Norden, DE

Project participants: Adreas Wurpts, Gholamreza Shiravani, Cordula Berkenbrink, Tea Behrends, Anne Ritzmann

[3]Project Partner: Rijkswaterstaat Noord-Nederland

Rijkswaterstaat Noord Nederland
Water, Verkeer en Leefomgeving
Zuiderwagenplein 2
8224 AD Lelystad, NL

Project participants: Gerrit Niebeek, Laura Tack, Leon van den Berg, Marcel van den Berg, Ellen Farwick, Raven Cammenga

[4]Subcontractor of Rijkswaterstaat: Deltares

Marine and Coastal Systems, Deltares
PO Box 177
2600 MH Delft, NL

Project participants: Theo Prins, Anouk Blauw, Tineke Troost, Lisa Schneider, Lauriane Vilmin, Annelotte van der Linden

[5]Project Partner: Helmholtz-Institut für Funktionelle Marine Biodiversität

Helmholtz-Institut für Funktionelle Marine Biodiversität (HIFMB)
an der Universität Oldenburg
Ammerländer Heerstraße 231
26129 Oldenburg, DE

Project participants: Helmut Hillebrand, Josie Antonucci di Carvalho, Ute Jacob

[6]Project Partner: Universität Hamburg

Arbeitsgruppe Wissenschaftliches Rechnen
Universität Hamburg
Min-Fakultät, Fachbereich Informatik
Bundesstraße 45a
20146 Hamburg, DE

Project participants: Hermann Lenhart, Daniel Thewes

[7]in Cooperation with: Universität Oldenburg

Citation: Rönn, L., Antonucci di Carvalho, J., Blauw, A., Hillebrand, H., Kerimoglu, O., Lenhart, H., Prins, T., Gholamreza, S., Tack, L., Thewes, D. & T. Troost (2023): Harmonisation of the Phytoplankton Assessment in the German and Dutch Wadden Sea. Interreg V A project "Wasserqualität - Waterkwaliteit" - Synthesis Report. Report prepared on behalf of NLWKN and Rijkswaterstaat, Oldenburg/Lelystad, 2023, 141 pp.

Cover: The cover shows a selection of photos from the Wadden Sea and WRRL-related figures. Map with current DE & NL WRRL thresholds for chlorophyll, © NLWKN (top left); Phytoplankton cells (*Odontella mobiliensis* & *Phaeocystis* sp.), © K. Kolbe (top middle & top right); Lower Saxony Wadden Sea, © S. Tyedmers (bottom left); Map of ecosystem-model domains, © Interreg project (middle); Satellite picture of the Wadden Sea, © QSR Wadden Sea – albedo39 Satellitenbildwerkstatt (bottom right).

Disclaimer: The designations employed and the presentation of the material contained in this report do not imply the expression of any opinion whatsoever on the part of NLWKN and Rijkswaterstaat concerning the legal status of any state, territory, city or area, or concerning the delimitations of their frontiers or boundaries. This report contains the views expressed by the authors and may not necessarily reflect the views of NLWKN or Rijkswaterstaat.

Table of contents

Executive summary

English

The undesirable accumulation of nutrients (eutrophication) remains a problematic issue in many European coastal waters, although measures taken in recent decades have already led to a decrease in nutrient river loads and coastal nutrient concentrations and thus to a reduction of impacts on coastal ecosystems such as the Wadden Sea UNESCO World Heritage Site.

Currently, phytoplankton biomass (measured as chlorophyll) serves as a key indicator of the extent of eutrophication. Since the previous assessment of phytoplankton in the German-Dutch coastal waters in the course of the Water Framework Directive (WFD) resulted in conflicting results, this Dutch/German Interreg research project was started to contribute to a more comprehensive understanding of the Wadden Sea ecosystem and a harmonized assessment of phytoplankton.

This project took an innovative perspective by adopting a multi-causal research approach and considered different parameters affecting phytoplankton and eutrophication in a transboundary analysis of long-term monitoring data and by ecosystem modelling. The latest scientific findings and results of the monitoring data analyses were combined and incorporated in the different ecosystem models to achieve the most realistic representation of the Wadden Sea system.

Data analyses and modelling results show that there are no major differences in chlorophyll concentrations between the German and Dutch Wadden Sea. The difference in levels of chlorophyll thresholds currently in place in Germany and the Netherlands are not supported by our scientific understanding of the natural conditions in the Wadden Sea. Further modelling results show that the required nitrogen reductions in the rivers to achieve the river management objective of annual mean total nitrogen concentrations of 2.8 mg/l at the limnic/marine border may not be enough to achieve good ecological status for phytoplankton in the water bodies of the Wadden Sea. Model results and data analyses indicate that chlorophyll does not react linearly to nitrogen reductions and that nitrogen is not the sole factor determining phytoplankton biomass in the coastal waters of the Wadden Sea.

Two ecosystem models were used to simulate pre-eutrophic, historic reference conditions as a basis to derive thresholds for the assessment of chlorophyll in the Wadden Sea. This approach (with more ecosystem models involved) was also used by OSPAR to derive harmonized chlorophyll thresholds for the eutrophication assessment for the Marine Strategy Framework Directive (MSFD) in the North-East Atlantic. The model results do not support the current differences in chlorophyll thresholds for the WFD Good/Moderate boundaries between the Netherlands and Germany. The thresholds calculated from this pre-eutrophic references, are, in all water bodies, higher than the current WFD thresholds. Differences in outcomes between the models reflect uncertainty in our understanding of the highly dynamic Wadden Sea system and hamper our ability to define accurate chlorophyll thresholds for the status assessment of phytoplankton.

The various analyses of the German and Dutch long-term plankton data have led to the new insight that the visible changes in the phytoplankton community are not exclusively based on changes in the eutrophication situation in the Wadden Sea, but also reflect a continuous, natural shift towards new communities. Changes in individual plankton parameters are very closely related to changes in environmental conditions, but since environmental conditions are in a constant state of flux (changes in nutrient levels, climate change), it is difficult to describe and assess a reference status for phytoplankton. The lack of a stable status quo therefore makes it difficult to set thresholds for different plankton parameters to assess its ecological status. However, parameters that describe community dynamics (such as multivariate biodiversity indicators) can be used to reveal trends in phytoplankton

composition and evaluate the influence of environmental factors. These should be taken into account in an extended and holistic approach to phytoplankton assessment within the WFD.

One of the main goals of this project was to develop an alternative assessment approach for phytoplankton in the Wadden Sea. During the course of the project, this turned out to be a complex task where we managed to address some basic principles. Nevertheless, this project provides an fundamental basis for a scientifically based system understanding of the Wadden Sea. The outcomes of this project contribute important findings to current assessment procedures on phytoplankton and eutrophication and are a good starting point for further discussions and developments at scientific and policy levels in the context of the WFD, MSFD and with regard to future work within OSPAR. Moreover, his project has strengthened the cooperation and exchange between the German and Dutch authorities and research institutions and promoted a common understanding of the cross-border issue of eutrophication in the Wadden Sea ecosystem.

Nederlands

De ongewenste accumulatie van nutriënten (eutrofiëring) blijft een problematisch probleem in veel Europese kustwateren, ondanks genomen maatregelen afgelopen decennia die al hebben geleid hebben tot een afname van de belasting van nutriënten in rivieren en aan de kust en daarmee tot een vermindering van de effecten op kustecosystemen zoals het UNESCO Werelderfgoed Waddenzee.

Momenteel dient de fytoplanktonbiomassa (gemeten als chlorofyl) als een belangrijke indicator voor de mate van eutrofiëring. Omdat de eerdere beoordeling van fytoplankton in de Duits-Nederlandse kustwateren in het kader van de Kaderrichtlijn Water (KRW) resulteerde in tegenstrijdige resultaten, is dit Nederlands/Duitse Interreg onderzoeksproject gestart om bij te dragen aan een beter begrip van het ecosysteem van de Waddenzee en een geharmoniseerde beoordeling van fytoplankton.

Dit project hanteerde een innovatief perspectief door te kiezen voor een multi-causale onderzoek aanpak met verschillende parameters die van invloed zijn op fytoplankton, door eutrofiëring te beschouwen in een grensoverschrijdende analyse van lang termijn monitoringgegevens en door ecosysteemmodellering. De nieuwste wetenschappelijke bevindingen en resultaten van de analyses van de monitoringgegevens werden gecombineerd en verwerkt in de verschillende ecosysteemmodellen om tot de meest realistische weergave van het Waddenzeesysteem te komen.

Gegevensanalyses en modelresultaten laten zien dat er geen grote verschillen zijn in chlorofylconcentraties tussen de Duitse en Nederlandse Waddenzee. Het verschil in niveaus van de chlorofyldrempelwaarden die momenteel in Duitsland en Nederland van kracht zijn, wordt niet ondersteund door ons wetenschappelijk begrip van de natuurlijke omstandigheden in de Waddenzee. Verdere modelresultaten laten zien dat de vereiste stikstofreducties in de rivieren om de doelstelling van een jaargemiddelde totale stikstofconcentraties van 2,8 mg/l aan de zoetwater/mariene grens te bereiken mogelijk niet voldoende zijn om een goede ecologische toestand voor fytoplankton in de wateren van de Waddenzee te bereiken. Modelresultaten en gegevensanalyses geven aan dat chlorofyl niet lineair reageert op stikstofreducties en dat stikstof niet de enige factor is die de fytoplanktonbiomassa in de kustwateren van de Waddenzee bepaalt.

Twee ecosysteemmodellen werden gebruikt om pre-eutrofische, historische referentieomstandigheden te simuleren als basis om drempelwaarden af te leiden voor de beoordeling van chlorofyl in de Waddenzee. Deze aanpak (met ecosysteemmodellen) werd ook gebruikt door OSPAR om geharmoniseerde chlorofyldrempelwaarden af te leiden voor de beoordeling van eutrofiëring voor de Kaderrichtlijn Mariene Strategie (KRM) in het noordoostelijk deel van de Atlantische Oceaan. De modelresultaten ondersteunen de huidige verschillen in

chlorofyldrempelwaarden voor de KRW Goed/Matig grenzen tussen Nederland en Duitsland niet. De drempelwaarden berekend op basis van deze pre-eutrofische referenties zijn in alle waterlichamen hoger dan de huidige KRW-drempelwaarden. Verschillen in uitkomsten tussen de modellen weerspiegelen de onzekerheid in ons begrip van het zeer dynamische Waddenzeesysteem en belemmeren ons vermogen om nauwkeurige chlorofyldrempelwaarden te definiëren voor de statusbeoordeling van fytoplankton.

De verschillende analyses van de Duitse en Nederlandse planktongegevens voor de lange termijn hebben geleid tot het nieuwe inzicht dat de zichtbare veranderingen in de fytoplanktongemeenschap niet uitsluitend gebaseerd zijn op veranderingen in de eutrofiëringssituatie in de Waddenzee, maar ook een continue, natuurlijke verschuiving naar nieuwe gemeenschappen weerspiegelen. Veranderingen in individuele planktonparameters zijn zeer nauw gerelateerd aan veranderingen in de milieuomstandigheden, maar omdat de milieuomstandigheden voortdurend in beweging zijn (veranderingen in nutriëntenlevels, klimaatverandering), is het moeilijk om een referentietoestand voor fytoplankton te beschrijven en te beoordelen. Het ontbreken van een stabiele status-quo maakt het daarom moeilijk om drempelwaarden vast te stellen voor verschillende planktonparameters om de ecologische toestand te beoordelen. Parameters die de dynamiek van de gemeenschap beschrijven (zoals multivariate biodiversiteitsindicatoren) kunnen echter wel gebruikt worden om trends in de samenstelling van het fytoplankton aan het licht te brengen en de invloed van milieufactoren te evalueren. Hiermee moet rekening worden gehouden in een uitgebreide en holistische benadering van de beoordeling van fytoplankton binnen de KRW.

Een van de belangrijkste doelen van dit project was het ontwikkelen van een alternatieve beoordelingsaanpak voor fytoplankton in de Waddenzee. In de loop van het project bleek dit een complexe taak te zijn waarbij we erin geslaagd zijn om een aantal basisprincipes aan te pakken. Desondanks biedt dit project een fundamentele basis voor een wetenschappelijk gefundeerd begrip van het systeem in de Waddenzee. De resultaten van dit project dragen bij aan belangrijke bevindingen voor de huidige beoordelingsprocedures voor fytoplankton en eutrofiëring en zijn een goed uitgangspunt voor verdere discussies en ontwikkelingen op wetenschappelijk- en beleidsniveau in de context van de KRW, KRM en met betrekking tot toekomstig werk binnen OSPAR. Bovendien heeft dit project de samenwerking en uitwisseling tussen de Duitse en Nederlandse autoriteiten en onderzoeksinstellingen versterkt en een gemeenschappelijk begrip van het grensoverschrijdende probleem van eutrofiëring in het ecosysteem van de Waddenzee bevorderd.

Deutsch

Die unerwünschte Anreicherung von Nährstoffen (Eutrophierung) ist in vielen europäischen Küstengewässern nach wie vor ein Problem, obwohl die in den letzten Jahrzehnten ergriffenen Maßnahmen bereits zu einem Rückgang der Nährstofffrachten aus den Flüssen und der Nährstoffkonzentrationen in den Küstengewässern und damit zu einer Verringerung der Auswirkungen auf Küstenökosysteme wie das UNESCO-Weltnaturerbe Wattenmeer geführt haben.

Derzeit dient die Phytoplankton-Biomasse (gemessen als Chlorophyll) als Schlüsselindikator für das Ausmaß der Eutrophierung. Da die bisherige Bewertung des Phytoplanktons in den deutsch-niederländischen Küstengewässern im Zuge der Wasserrahmenrichtlinie (WRRL) zu widersprüchlichen Ergebnissen führte, wurde dieses deutsch-niederländische Interreg-Forschungsprojekt gestartet, um zu einem umfassenderen Verständnis des Ökosystems Wattenmeer und einer harmonisierten Bewertung des Phytoplanktons beizutragen.

Dieses Projekt verfolgte eine innovative Herangehensweise ein, indem es einen multikausalen Forschungsansatz wählte und verschiedene Parameter, die das Phytoplankton und die Eutrophierung beeinflussen, in einer grenzüberschreitenden Analyse von Langzeitmonitoringdaten und durch Ökosystemmodellierung berücksichtigte. Die neuesten wissenschaftlichen Erkenntnisse und die Ergebnisse der Analysen der Monitoringdaten wurden kombiniert und in die verschiedenen Ökosystemmodelle integriert, um eine möglichst realistische Darstellung des Wattenmeersystems zu erreichen.

Datenanalysen und Modellierungsergebnisse zeigen, dass es keine wesentlichen Unterschiede in der Chlorophyllkonzentration zwischen dem deutschen und dem niederländischen Wattenmeer gibt. Die unterschiedlichen Chlorophyll-Grenzwerte, die derzeit in Deutschland und den Niederlanden gelten, werden durch unser wissenschaftliches Verständnis der natürlichen Bedingungen im Wattenmeer nicht gestützt. Weitere Modellierungsergebnisse zeigen, dass die erforderlichen Stickstoffreduzierungen in den Flüssen zur Erreichung des Bewirtschaftungsziels von jährlichen mittleren Gesamtstickstoffkonzentrationen von 2,8 mg/l an der limnischen/marinen Grenze möglicherweise nicht ausreichen, um einen guten ökologischen Zustand des Phytoplanktons in den Wasserkörpern des Wattenmeeres zu erreichen. Modellergebnisse und Datenanalysen deuten darauf hin, dass Chlorophyll nicht linear auf Stickstoffreduzierungen reagiert und dass Stickstoff nicht der einzige Faktor ist, der die Phytoplanktonbiomasse in den Küstengewässern des Wattenmeeres bestimmt.

Zwei Ökosystemmodelle wurden verwendet, um präeutrophe, historische Referenzbedingungen als Grundlage für die Ableitung von Grenzwerten für die Bewertung von Chlorophyll im Wattenmeer zu ermitteln. Dieser Ansatz (mit mehreren beteiligten Ökosystemmodellen) wurde auch von OSPAR verwendet, um harmonisierte Chlorophyll-Grenzwerte für die Eutrophierungsbewertung für die Meeresstrategie-Rahmenrichtlinie (MSRL) im Nordostatlantik abzuleiten. Die Modellergebnisse stützen die derzeitigen Unterschiede bei den Chlorophyll-Grenzwerten für die WRRL-Gut/Mäßig-Grenzen zwischen den Niederlanden und Deutschland nicht. Die Grenzwerte, die auf der Grundlage der präeutrophen Referenzwerte berechnet wurden, liegen in allen Wasserkörpern höher als die derzeitigen WRRL-Grenzwerte. Die unterschiedlichen Ergebnisse der Modelle spiegeln die Unsicherheit in unserem Verständnis des hochdynamischen Wattenmeersystems wider und erschweren es uns, genaue Chlorophyll-Grenzwerte für die Zustandsbewertung des Phytoplanktons festzulegen.

Die verschiedenen Analysen der deutschen und niederländischen Langzeit-Planktondaten haben zu der neuen Erkenntnis geführt, dass die sichtbaren Veränderungen in der Phytoplanktongemeinschaft nicht ausschließlich auf Veränderungen der Eutrophierungssituation im Wattenmeer beruhen, sondern auch eine kontinuierliche, natürliche Verschiebung hin zu neuen Gemeinschaften widerspiegeln. Veränderungen einzelner Planktonparameter stehen in engem Zusammenhang mit Veränderungen der Umweltbedingungen. Da die Umweltbedingungen jedoch einem ständigen Wandel unterworfen sind (Veränderungen der Nährstoffgehalte, Klimawandel), ist es schwierig, einen Referenzzustand für das Phytoplankton zu beschreiben und zu bewerten. Das Fehlen eines stabilen Status quo macht es daher schwierig, Grenzwerte für verschiedene Planktonparameter festzulegen, um den ökologischen Zustand zu bewerten. Parameter, die die Dynamik der Gemeinschaft beschreiben (z. B. multivariate Biodiversitätsindikatoren), können jedoch verwendet werden, um Trends in der Phytoplanktonzusammensetzung aufzuzeigen und den Einfluss von Umweltfaktoren zu bewerten. Diese sollten in einem erweiterten und ganzheitlichen Ansatz zur Bewertung des Phytoplanktons im Rahmen der WRRL berücksichtigt werden.

Eines der Hauptziele dieses Projekts war es, einen alternativen Bewertungsansatz für das Phytoplankton im Wattenmeer zu entwickeln. Im Laufe des Projekts stellte sich dies als eine sehr

komplexe Aufgabe heraus, bei der es uns gelang, nur einige grundlegende Prinzipien zu berücksichtigen. Nichtsdestotrotz bietet dieses Projekt eine grundlegende Basis für ein wissenschaftlich fundiertes Systemverständnis des Wattenmeeres. Die Ergebnisse dieses Projekts tragen wichtige Erkenntnisse zu den derzeitigen Bewertungsverfahren für Phytoplankton und Eutrophierung bei und sind ein guter Ausgangspunkt für weitere Diskussionen und Entwicklungen auf wissenschaftlicher und politischer Ebene im Zusammenhang mit der WRRL, der MSRL und im Hinblick auf die künftige Arbeit im Rahmen von OSPAR. Darüber hinaus hat das Projekt die Zusammenarbeit und den Austausch zwischen den deutschen und niederländischen Behörden und Forschungseinrichtungen gestärkt und ein gemeinsames Verständnis für das grenzüberschreitende Problem der Eutrophierung im Ökosystem Wattenmeer gefördert.

Acknowledgements

List of Abbreviations

ADTE	Advection-diffusion transport equations
BOCHTVWTM	Monitoring station: "Bocht van Watum"
BOOMKDP	Monitoring station: "Boomkensdiep"
Bork_W_1	Monitoring station: "Westerems, Emshörn Rinne"
C	Carbon
Chl, Chl-a	Chlorophyll a
CN	Carbon to nitrogen ratio
DANTZGT	Monitoring station: "Dantziggat"
DCSM	Deltares Continental Shelf Model
DE	Germany
DIN	Dissolved Inorganic Nitrogen
DIP	Dissolved Inorganic Phosphorous
DOOVBWT	Monitoring station: "Doove Balg west"
ENS	Effective Number of Species
EU	European Union
FSK	Forschungsstelle Küste of the NLWKN
GF/F	Glass fiber filter
GPM	Southern North Sea – Generalized Plankton Model
GROOTGND	Monitoring station: "Groote Gat noord"
HIFMB	Helmholtz Institute for Functional Marine Biodiversity
HPLC	High-performance liquid chromatography
HUIBGOT	Monitoring station: "Huibertgat oost"
ICBM	Institut für Chemie und Biologie des Meeres
ICC	Intra Class Correlation
ID	Identification
JaBu_W_1	Monitoring station: "Wilhelmshaven Mole"
LMM	Linear mixed effect model
LN	Natural logarithm
LOESS	Locally Weighted Least Squares Regression
MARSDND	Monitoring station: "Marsdiep noord"
MSFD	Marine Strategy Framework Directive
NA	missing values represented by the symbol NA (not available)
NB	An abbreviation for the Latin phrase nota bene, meaning "note well"
NL	The Netherlands
NLWKN	Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz
Nney_W_2	Monitoring station: "Norderney (HW)"
NP	Nitrogen to Phosphorus ratio
OSPAR	Oslo Paris Commission for the Protection of the North-East Atlantic
PH3	Pelagic Habitat Indicator - OSPAR Comission
PIE	Probability of Interspecific Encounters
POC	Particulate organic carbon
PSU	Practical Salinity Unit
RANS	Reynolds Averaged Navier Stokes

ROTTMPT3	Monitoring station: "Rottumerplaat 3 km uit de kust"
RWS	Rijkswaterstaat
S	Species richness
SEM	Structural equation model
SER	Species Exchange Ratio
Si	Silicon
SPM	Suspended particulate matter
TERSLG10	Monitoring station: "Terschelling 10 km uit de kust"
TERSLG4	Monitoring station: "Terschelling 4 km uit de kust"
TN	Total Nitrogen
TP	Total Phosphorus
VIF	Variance Inflation Factor
WeMu_W_1	Monitoring station: "Wesermündung "
WFD	Water Framework Directive
ZUIDOLWOT	Monitoring station: "Zuid Oost Lauwers oost"

1. Project introduction

1.1 Project background

The Netherlands, Germany and Denmark share a unique coastal ecosystem – the Wadden Sea – which has been designated as Natura2000 site under the European Birds and Habitats Directives (Directive 2009/147/EC; Directive 92/43/EEC) as well as a UNESCO World Heritage Site. The Wadden Sea stretches 500 km along the North Sea coast of these three countries and is a naturally productive as well as dynamic and complex ecosystem, which has a great ecological importance for many species such as fish populations, breeding and migrating sea birds and benthic invertebrates.

The Wadden Sea is nowadays simultaneously exposed to multiple anthropogenic stressors like fisheries, shipping, interventions related to energy production, climate change, pollution and tourism, which affect this sensitive ecosystem on different levels. Eutrophication – a process by which a water body of becomes enriched in dissolved nutrients – has been one of the biggest problems in European coastal waters, especially in the southern North Sea. In the 1970s and 1980s, large amounts of nutrients entered the coastal waters via rivers, from the atmosphere, land run-off or by direct discharges into the sea and caused undesirable disturbance of coastal ecosystems such as excessive growth of phytoplankton, increased turbidity and growth of opportunistic macroalgae but also a loss in biodiversity and a decline in seagrass and oxygen concentration in the water and bottom sediments (van Beusekom et al. 2017). Although measures implemented in the past decades have already led to a decrease in nutrient input to the marine environment and to a reduction in eutrophication effects, eutrophication still remains an ongoing and problematic issue in European coastal waters. In addition, cumulative effects related to climate change and global warming may worsen disturbances in biological communities as well as complicate efforts to demonstrate causal links and verify the effectiveness of measures.

A surplus of nutrient supply to marine ecosystems often leads to an increase in the amount of biomass of drifting unicellular algae, the phytoplankton. Phytoplankton biomass is also affected by water temperature, light availability as well as grazing by zooplankton or benthic filter feeders, but a significant correlation to riverine total nitrogen load has been demonstrated (van Beusekom et al. 2017). Therefore, the stage of eutrophication is commonly assessed via the photosynthetic pigment chlorophyll a (Chl) as an indicator of phytoplankton biomass.

The assessment of the ecological status of European waters takes place within the framework of the EU-Water Framework Directive (WFD, 2000/60/EC) for all coastal waters and the EU-Marine Strategy Framework Directive (MSFD, 2008/56/EC) for national marine waters adjacent to the WFD-areas. Both directives strive to reach a healthy status of European marine waters (WFD: "Good Ecological Status"; MSFD: "Good Environmental Status"). Within the WFD, phytoplankton is one of the quality elements that serves to assess the ecological status of coastal waters. For the classification of the ecological status, the WFD uses a system of five classes (high/good/moderate/poor/bad) defined by different thresholds of individual parameters that describe the current status with regard to a reference state. For the assessment of the status of phytoplankton, based on the WFD normative definitions, the required parameters are phytoplankton community, abundance and biomass as well as frequency and intensity of blooms. So far, the concentration of Chl – as a proxy for phytoplankton biomass – is one of the single parameters that has been widely used by European member states to assess the status of phytoplankton.

During a process called intercalibration, which is required by the WFD, the methods and parameters for the assessment of quality elements are harmonised between EU member states managing the same water body types. The intercalibration process is aimed at ensuring comparability of the classification results of the WFD assessment methods developed by the member states for the biological quality elements. For phytoplankton the intercalibration process between the Netherlands and Germany was formally completed in 2017 and the results were published in the EU Commission decision 2018/229 (EC 2018). Although Chl was intercalibrated successfully as the parameter to assess

the status of phytoplankton and eutrophication in most European countries, a satisfactory harmonization of applied Chl thresholds between Germany and the Netherlands was however not reached (Fig. 1.1, Table 1.1).
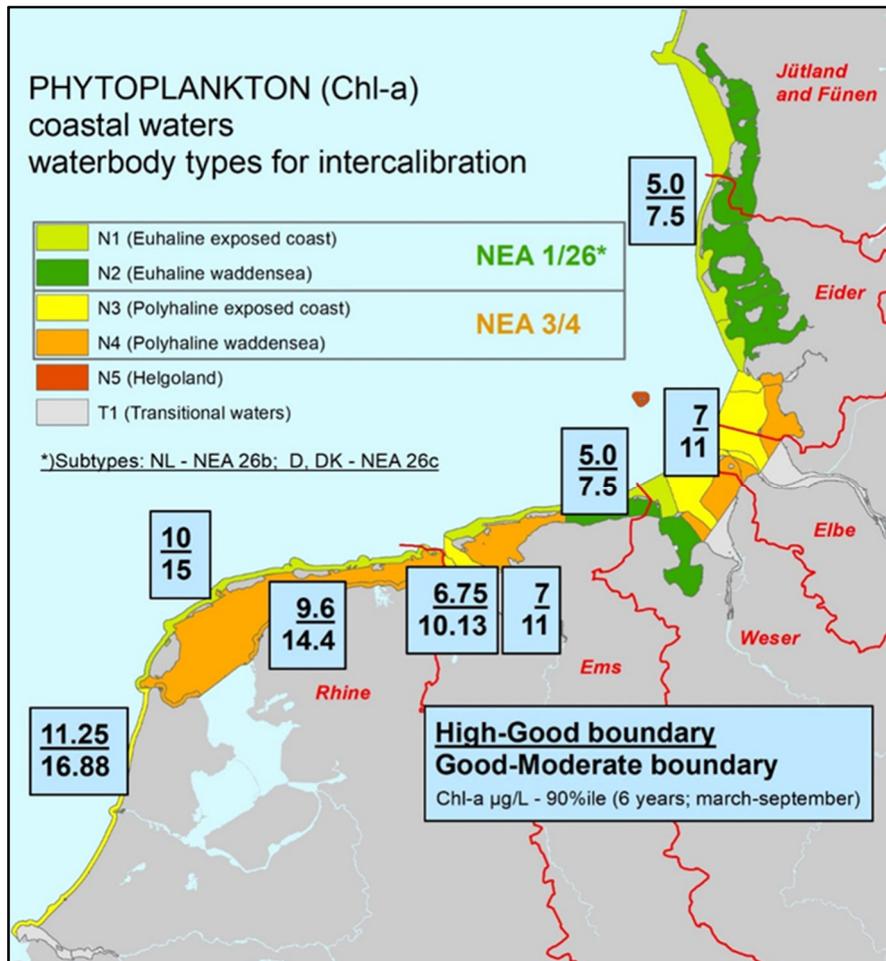
Figure 1.1. WFD-typology of coastal water bodies of the Netherlands and Germany with current chlorophyll *a* thresholds (EC 2018) for high/good and good/moderate boundaries, indicated as 90th percentile of chlorophyll *a* concentration (µg/l) of the growing season (March-September) over a six-year period.

Table 1.1: Current WFD chlorophyll *a* thresholds for high/good and good/moderate boundaries (EC 2018). Original threshold values indicated as 90th percentile of chlorophyll *a* concentration (µg/l) of the growing season (March-September) over a six-year period. Good/moderate threshold values are also indicated as mean values (highlighted column in blue) as compared to Table 7.2.

Country	WFD Water Body	Chl threshold high/good (µg/l; 90th percentile)	Chl threshold good/moderate (µg/l; 90th percentile)	Chl threshold good/moderate (µg/l; mean value)
DE	N1 (euhaline open coastal)	5	7,5	3,8
	N2 (euhaline Wadden Sea)	5	7,5	3,8
	N3 (polyhaline open coastal/ Ems Dollart)	7	11	5,5
	N4 (polyhaline Wadden Sea)	7	11	5,5
NL	N1 (Wadden Coast)	10	15	7,5
	N3 (Ems-Dollard Coast)	6,75	10,13	5,1
	N4 (Wadden Sea)	9,6	14,4	7,2

It needs to be pointed out, that for the Ems-N3 area, there are two Chl thresholds in use, which differ between the Dutch and the German part. For the Dutch water body, characterised as "Ems-Dollard coast", the mean value of 5.1 µg/l is applied, while for the German water body, indicated as "Ems-N3", a value of 5.5 µg/l is used. The water body 'Wadden Sea coast' is a narrow strip along the mainland coast consisting mainly of marshes and tidal flats. For this water body, the WFD assessment of nutrients and phytoplankton is based on the same monitoring and thresholds as the much larger water body 'Wadden Sea'. In this report, the water body 'Wadden Sea coast' is therefore not taken into consideration.

1.2 Project motivation

To solve this discrepancy and to harmonise the assessment of phytoplankton between the Netherlands and Germany in the shared water bodies of the Wadden Sea, the authorities of both countries agreed to aim for a solution in a newly established joint research project. The Interreg V A Program Deutschland-Nederland of the Ems-Dollart-Region was a suitable platform to approach this topic in a cross-border collaboration to strengthen the communication and cooperation between both countries on the scientific as well as on the administration and management level. It was commonly agreed that for the successful management of the coastal waters – especially at the border area between the Netherlands and Germany – a long-term agreement of neighbouring countries on this issue should be sought.

The current project *"Wasserqualität – Waterkwaliteit"* therefore aims at a common German-Dutch understanding of the magnitude of eutrophication in the coastal waters of the German-Dutch Wadden Sea and the adjacent North Sea and to gain a more comprehensive understanding of the role of phytoplankton in the Wadden Sea ecosystem that should lead to a harmonised assessment of phytoplankton in the context of the WFD.

During the project, the existing assessment of the quality component phytoplankton was reviewed, and a new, multi-causal research approach was developed that includes a collaborative ecosystem modelling approach in combination with the common analysis of German and Dutch phytoplankton long-time monitoring data series. This innovative approach examines the Wadden Sea ecosystem in a more holistic way than previously done in order to jointly achieve a coherent assessment of the phytoplankton and eutrophication status in the German-Dutch Wadden Sea and adjacent coastal waters.

This project has taken place in a voluntary framework and was intended, in addition to the monitoring programs and assessment procedures required by the WFD, to provide answers to open questions of relevance, which have not yet been adequately answered to further progress in this field. The project does not include new measurement and monitoring campaigns but is based solely on existing monitoring and modelling data from the project partners involved. The model periods were aligned as much as possible with the assessment intervals of the WFD, the MSFD and OSPAR commission to consider new developments and, conversely, to contribute to the current development and discussions within the WFD, MSFD and OSPAR commission.

The Wadden Sea UNESCO World Heritage Site is part of different river basins and catchment areas, which are managed by different EU member states. This bilateral project therefore supports the exchange and clustering of expertise of all project partners and improves and strengthens the communication and cooperation between German and Dutch governmental authorities and scientific research institutes.

A harmonized assessment of the degree of eutrophication in the Wadden Sea and a coherent assessment of phytoplankton based on sound and best available scientific knowledge – including the most recent ecosystem models – is of great importance for Germany and the Netherlands for a coherent management of the Dutch-German Wadden Sea. The results of this project therefore contribute to clarify the extent to which eutrophication in the German-Dutch Wadden Sea and adjacent coastal waters is problematic and what further measures (e.g., nutrient reductions) are needed to achieve a Wadden Sea without eutrophication effects i.e., to reach good ecological status according to the WFD. In addition, the derived assessment of phytoplankton will also support further developments and discussions within the framework of the MSFD and OSPAR commission.

1.3 Project objectives

The overall aim of this project is to promote a common German-Dutch understanding of the extent of eutrophication in the coastal waters of the Wadden Sea and adjacent North Sea and finally to determine and agree upon realistic reference conditions and common assessment levels for phytoplankton in the water bodies NEA 3/4 and NEA 1/26 to achieve the harmonisation of phytoplankton assessment between Germany and the Netherlands in the framework of the WFD.

Detailed objectives of this project are:

- To gain a more comprehensive understanding of the complex system of the Wadden Sea and to gain a holistic picture on phytoplankton conditions and eutrophication status;
- To realise an innovative, multi-causal research approach by considering different parameters concerning phytoplankton and eutrophication in a cross-national ecosystem modelling;
- To solve shortcomings of the intercalibration process with a stepwise approach using different ecosystem models and a comprehensive analysis of long-term monitoring data;
- To provide a reliable scientific background as a basis to harmonise the phytoplankton assessment in the Wadden Sea;
- To test several phytoplankton and eutrophication related parameters as possible criteria for phytoplankton assessment, including chlorophyll a;
- To test different scenarios of nutrient reduction and their effects on phytoplankton biomass and the eutrophication status in the Wadden Sea;
- To build a bridge between scientific analysis and operational management tools;
- To strengthen the bilateral cooperation between German and Dutch authorities and research institutions and to combine competence and expertise of all project partners to finally promote a common understanding of eutrophication in the Wadden Sea as a cross-border issue;
- Finally: to propose reference conditions and common assessment levels for phytoplankton in the coastal water bodies NEA 3/4 and NEA 1/26.

2. Joint system understanding

2.1 General description

The Wadden Sea ecosystem is a unique sandy-muddy tidal system, characterized by tidal flats that are bordered offshore by a chain of barrier islands and onshore by extensive salt marshes (CWSS 2017). The strong tidal influence and consequently the constant in- and outflow of water through the channels between the islands result in a continuous water exchange between Wadden Sea and adjacent North Sea and causes high turbidity, inflow of suspended matter and nutrients and extreme environmental conditions of alternated wetting and drying. Freshwater, nutrients and sediment are additionally introduced into the estuaries and coastal waters of the Wadden Sea by river run-off via the large rivers like Rhine, Meuse, Ems, Weser and Elbe.

Regarding its hydro- and morphodynamic properties, the Wadden Sea displays strong environmental gradients and different transitional zones. Due to these characteristics, a high variety of different habitats exist (including seagrass and mussel beds, reefs, salt marshes), which are home to a rich and diverse fauna and flora that additionally shape this highly productive ecosystem (CWSS 2017).

In this project, we aimed at improved system understanding through an analysis of existing data sets from monitoring programs of Germany and the Netherlands and through a common modelling exercise exploring relations between riverine nutrient loads and eutrophication in the Wadden Sea.

2.2 System understanding based on monitoring data

2.2.1 Monitoring stations

For this project, we analysed long time series data from German and Dutch monitoring stations in the Wadden Sea and adjacent coastal waters (Fig. 2.1). The stations TERSLG4 and ZUIDOLWOT were left out since the time series did not cover recent years (Table 2.1).
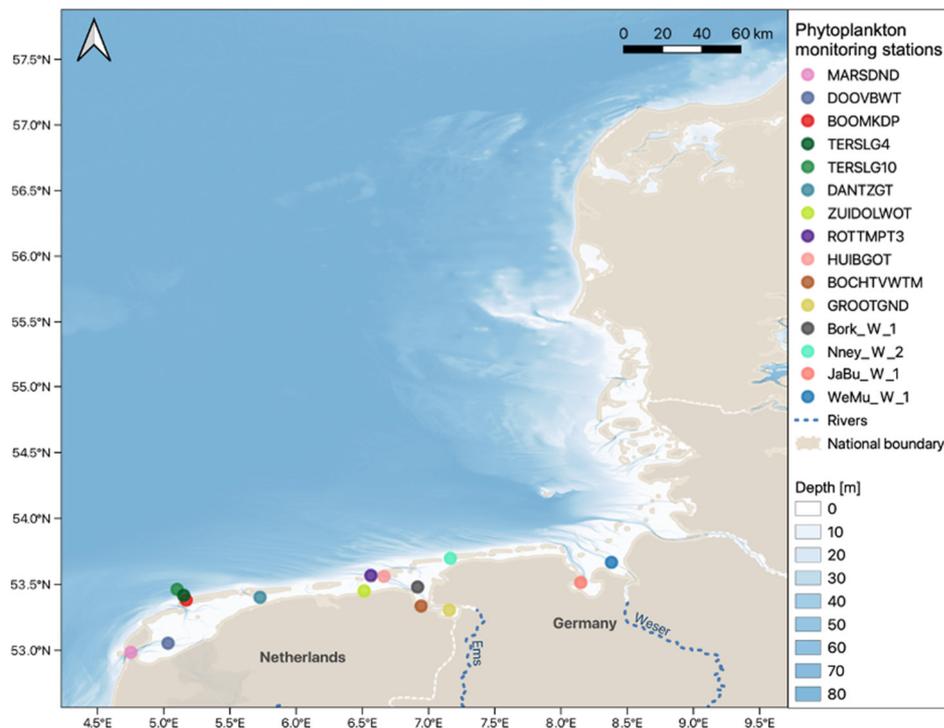
Figure 2.1. Map of the study area, including the phytoplankton monitoring stations in the Wadden Sea. Data is available for four German stations (Bork_W_1, Nney_W_2, JaBu_W_1 and WeMu_W_1) and 11 Dutch stations. Map created using the Free and Open Source QGIS (v.3.24.2). Bathymetry from EMODnet Bathymetry project (2020).

We refrain from reiterating the strategy behind these monitoring stations and the details of sampling, which have been laid out elsewhere (Hanslik et al. 1998, Prins et al. 2012). We used the original data sets but undertook a series of harmonization steps. First, we removed all species identified as purely heterotrophic. Second, we harmonized the species nomenclature between the two datasets, Dutch and German. Third, species-specific biomass was estimated from biovolume using the C-conversion equations described by Menden-Deuer and Lessard (2000), which for diatoms with biovolumes >3000 µm^3 is *0.288 * Volume$^{0.811}$ = pgC cell^{-1}*, for smaller diatoms and other groups: *0.216 * Volume$^{0.939}$ = pgC cell^{-1}*.

Table 2.1. Coastal monitoring stations in the Dutch and German monitoring programs.

Country	Station ID	Station name	Observation
Netherlands	MARSDND	Marsdiep noord	
Netherlands	DOOVBWT	Doove Balg west	
Netherlands	BOOMKDP	Boomkensdiep	
Netherlands	TERSLG4	Terschelling 4 km uit de kust	No data available after 2007
Netherlands	TERSLG10	Terschelling 10 km uit de kust	
Netherlands	DANTZGT	Dantziggat	
Netherlands	ZUIDOLWOT	Zuid Oost Lauwers oost	No data available after 2009
Netherlands	ROTTMPT3	Rottumerplaat 3 km uit de kust	
Netherlands	HUIBGOT	Huibertgat oost	
Netherlands	BOCHTVWTM	Bocht van Watum	
Netherlands	GROOTGND	Groote Gat noord	
Germany	Bork_W_1	Westerems, Emshörn Rinne	
Germany	Nney_W_2	Norderney (high tide)	
Germany	JaBu_W_1	Wilhelmshaven Mole	
Germany	WeMu_W_1	Wesermündung	No winter samples

2.2.2 Comparison between German and Dutch monitoring programs

a) Sampling frequency

Regarding the sampling frequency at each coastal station, we observed that the sampling effort varied considerably across stations and years, in both environmental (Fig. 2.2) and phytoplankton (Fig. 2.3) monitoring data sets. In Bork_W_1, the environmental parameters were only sampled in winter and autumn from 1994-2010 and sometimes in spring. From 2011 onwards, this station was sampled also in summer but then in a lower frequency in the winter months (Fig. 2.2). In this same station, phytoplankton samples were taken over the four seasons from 2007-2010, then only in spring, summer and autumn (Fig. 2.3). WeMu_W_1 was never sampled in winter. In some stations the sampling frequency increased over time (ex. In DOOVBWT), while in other stations it decreased over time (ex. GROOTGND and HUIBGOT). Also, the number of sampled months per season varied across stations and years. In general, the Dutch stations have longer time series data, especially for environmental variables, than the German stations.

Given the discrepancies in sampling and given that the main questions of this report address long-term trends in phytoplankton biomass and biodiversity, we calculated annual medians for each parameter. The median is superior to the mean as it is unaffected by extreme outliers and non-normal distributions of the data.
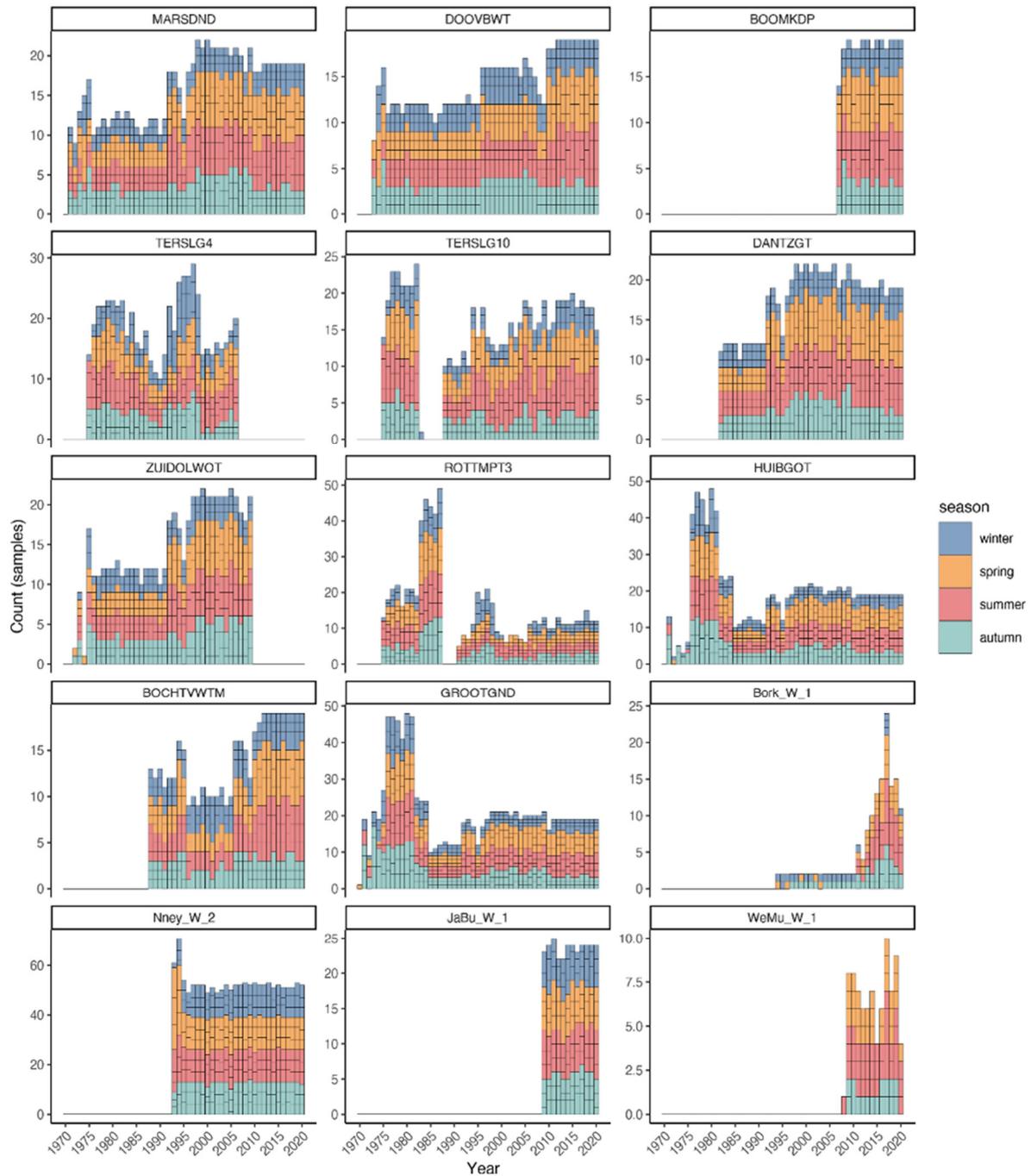
Figure 2.2. Sampling frequency for the water quality parameters of the Wadden Sea stations from 1970-2020. The bars represent the number of samples per year. Each segment of the bars represents one month sampled and the colours represent the season: *winter* = Dec, Jan, Feb; *spring* = Mar, Apr, May; *summer* = Jun, Jul, Aug; *autumn* = Sep, Oct, Nov.
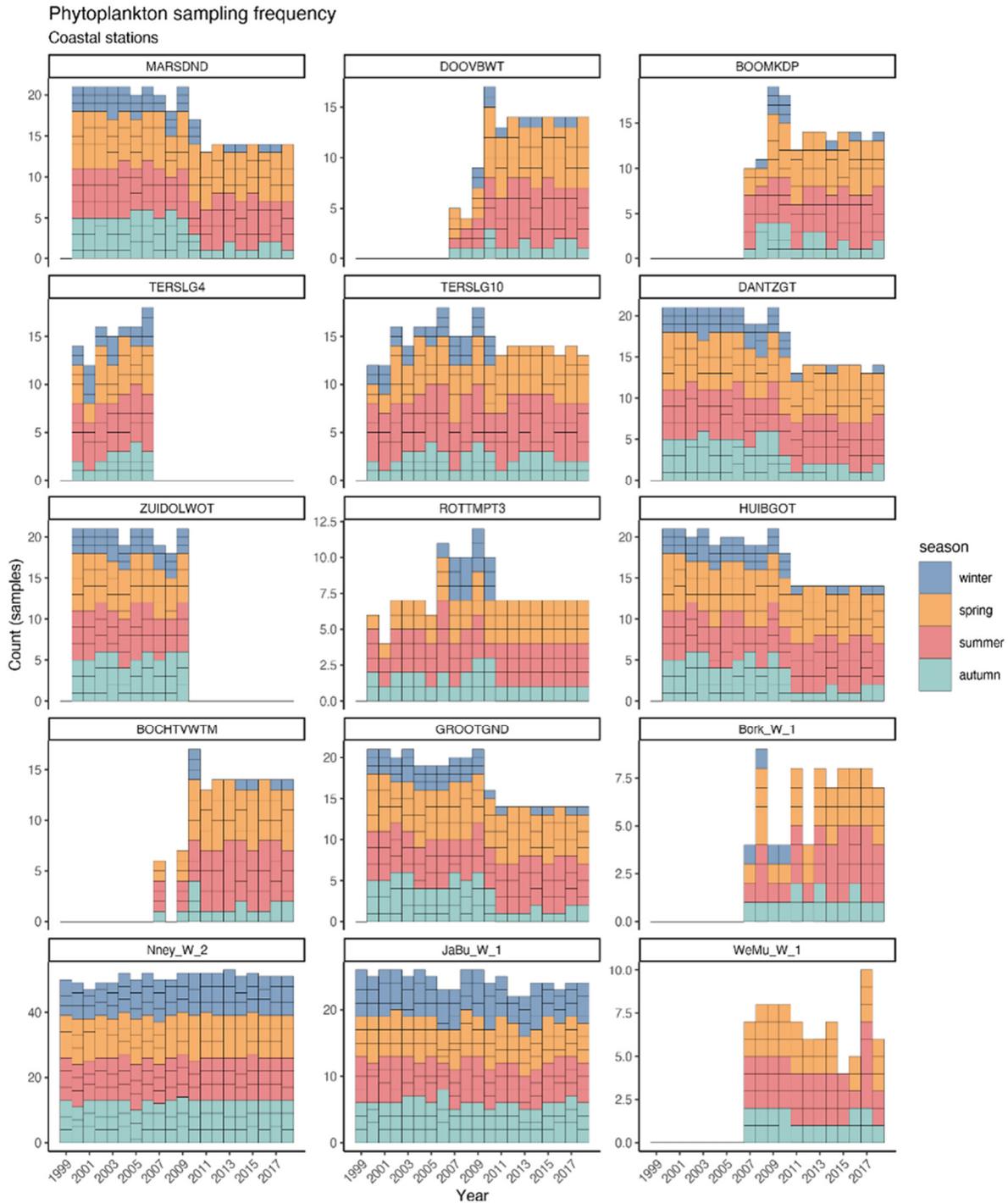
Figure 2.3. Sampling frequency for phytoplankton abundance of the Wadden Sea stations from 1999-2018. The bars represent the number of samples per year. Each segment of the bars represents one month sampled and the colours represent the season: *winter* = Dec, Jan, Feb; *spring* = Mar, Apr, May; *summer* = Jun, Jul, Aug; *autumn* = Sep, Oct, Nov.

b) *Biovolume estimates*

Regarding biovolume estimates, the Netherlands rely on literature values whereas Germany relies on per sample cell volume estimates. Phytoplankton cell size is an important trait that can provide insights on different morphological and physiological aspects of species and can be related to environmental changes and grazing (Hillebrand et al. 2022). Cell size analysis of the German Wadden Sea phytoplankton revealed that species are 30% smaller now than 15 years ago (Hillebrand et al. 2021).

This and further analyses can only be done when cell sizes are measured per sample. Based on these findings, we highlight the importance of measuring the cells from/in the samples instead of using standardised literature values, which are often overestimated and do not capture temporal changes.

c) Species composition

In order to compare the species composition between both data sets, we analysed nearby stations located in the Ems: the German station Bork_W_1 and the Dutch station BOCHTVWTM. Interestingly, only 40,8% of the identified species are shared between the stations. This difference in species composition can be related to the taxonomic expertise and effort of analysts involved in counting and identifying the cells. In long time series data, the microscopic taxonomy can be largely influenced by the change of taxonomists involved in the identification of phytoplankton species, but also by the development and improvement of analytical tools and sampling methods (Löder et al. 2012, Nohe et al. 2018). Temporal trends in carbon content also revealed a notable difference between the two stations, with consistently much higher values in the Dutch station.

d) Analytical methods to measure Chlorophyll

The methods to measure chlorophyll *a* (Chl) differ between the two countries, with photometry being used by Germany and High-performance liquid chromatography (HPLC) used by the Netherlands (Table 2.2).

While in the photometer the concentration of different pigments is given, the HPLC gives the measurement of Chl exclusively, which results in a more accurate value for Chl, which is lower than the concentration determined by the photometer (Baretta-Bekker et al. 2015).

In the photometer, the absorbance of the sample is first measured at 665 nm, then, after all Chl is converted to phaeophytin by hydrochloric acid, the sample is measurement again at 650 nm. The difference between the two measurements is used to obtain the proportion of active Chl and phaeophytin and to determine the total Chl content. As Chl and phaeophytin have different molar extinction coefficient, i. e. they absorb light to different extents, this is considered in the conversion formula. Therefore, three measurements are available in the German data set, *Total chlorophyll content, active chlorophyll and phaeophytin*, while in the Dutch data only c*hlorophyll-a* is available (Table 2.2).

Table 2.2. Chlorophyll measurements of the Dutch and German monitoring.

Country	Parameter	parameter description	unit	Method
The Netherlands	CHLFa	Chlorophyll-a	$\mu g\ l^{-1}$	HPLC
Germany	Chloro_gesamt	Total chlorophyll content	$\mu g\ l^{-1}$	Photometry
Germany	Chloro_aktiv	Active chlorophyll	$\mu g\ l^{-1}$	Photometry
Germany	Phaeophytin	Phaeophytin	$\mu g\ l^{-1}$	Photometry

In order to compare the differences in Chl concentrations between the countries, a test was planned whereby the German samples were to be analysed in parallel for a year in the same laboratory where the Dutch data are analysed. However, this comparison failed because the transport of samples got considerably delayed by the transport company and thus spoiled for further analyses.

Even though the countries use different analytical methods to measure Chl, it was agreed that the "Total chlorophyll content" in the German dataset is the best comparable metric to the "Chlorophyll-a" concentrations in the Dutch dataset. For future analyses, we recommend the harmonisation of the

analytical methods to measure Chl. For an accurate analysis of the chlorophyll-a pigment, HPLC is the preferred method.

2.2.3 Description of trends in environmental parameters and phytoplankton parameters

In order to show the temporal trends for each station as well as the trend for the entire data sets, we combined two different analytical approaches. First, we present the annual median per station and visualize the trend over time using a LOESS regression (Locally Weighted Least Squares Regression). The same is also done for German and Dutch data to test whether the temporal dynamics are different between countries. Second, the formal test for temporal changes relies on a linear mixed effect model (LMM), where the response variable is a function of year as fixed effect and (1|StationID) as random effect. The random effect allows for different intercepts for the stations but tests for a joint (common) slope with time. It therefore explicitly tests whether the response variable shows a joint and significant linear trend across all stations. The LMMs were performed in R using the lme4 package (Bates et al. 2015). To normalize data distribution, we calculated the natural log of the annual median of each environmental and biomass parameter except for temperature, pH and salinity.

a) *Environmental variables over time*

Total nitrogen (TN) and total phosphorus (TP) varied over several orders of magnitude between the most nutrient rich stations (GROOTGND and BOCHTVWTM) and the nutrient poorest (TERSLG10), but at each station both TN and TP significantly decreased over time (Fig. 2.4a, c). Consequently, we observed strong significant declines in TN and TP (Table 2.3), which did not differentiate much between NL and DE stations (Fig. 2.4b, d). This decrease, which encompassed almost an order of magnitude across the last 50 years, was proportionally larger for TP than for TN. Consequently, the molar N:P ratio significantly increased within and across stations (Fig. 2.4e, f, Table 2.3), exceeding previously reported high N:P ratios from the 1970s. The molar N:P ratio was constantly higher than the 16:1 Redfield ratio or the molar N:P ratio of 22:1 (Guildford & Hecky 2000), indicating a tendency towards increasing P-limitation close to the coast (Burson et al. 2016). For Si, we see a similar broad range of concentrations as for TN and TP between stations (Fig. 2.4g, h), but the temporal trend is much more subtle albeit overall negative (Table 2.3). Again, the two low salinity stations GROOTGND and BOCHTVWTM had the highest Si-concentrations, TERSLG10 the lowest.

Suspended particulate matter (SPM) varied among the stations, again with GROOTGND and BOCHTVWTM having highest and TERSLG10 lowest concentrations (Fig. 2.5a). The overall temporal decline in SPM (Table 2.3) was steeper in German than in Dutch stations (Figure 2.5b). Salinity increased with time (Table 2.3), which however was mainly visible in the Dutch stations and here the station with lowest salinity (GROOTGND) (Fig. 2.5c, d). The overall temperature increase of 0.4°C per decade indicates a strong warming effect (Fig. 2.5e, f, Table 2.3), which is significantly faster than for the open North Sea, which warmed by 1.3°C from 1969 to 2017 (UBA 2019). A significant but small increase in pH became visible across the Dutch stations, whereas the German stations rather declined in pH over time (Fig. 2.5g, h, Table 2.3).

When including the random effects, the LMM explained 35-92 % of the variation in environmental variables (conditional R^2 in Table 2.3). The temporal trends were strongest for TN, TP and their ratio as well as temperature (marginal R^2 in Table 2.3), whereas for SPM, Si and salinity the explained variance by the common linear term is <1%. Thus, in addition to direct human impacts on the Wadden Sea (fisheries, shipping, tourism), the ecosystem is characterized by massive multifactorial changes in the abiotic conditions, where current nutrient levels are lower than during the last 50 years, but temperatures and N:P ratios are higher than previously recorded.
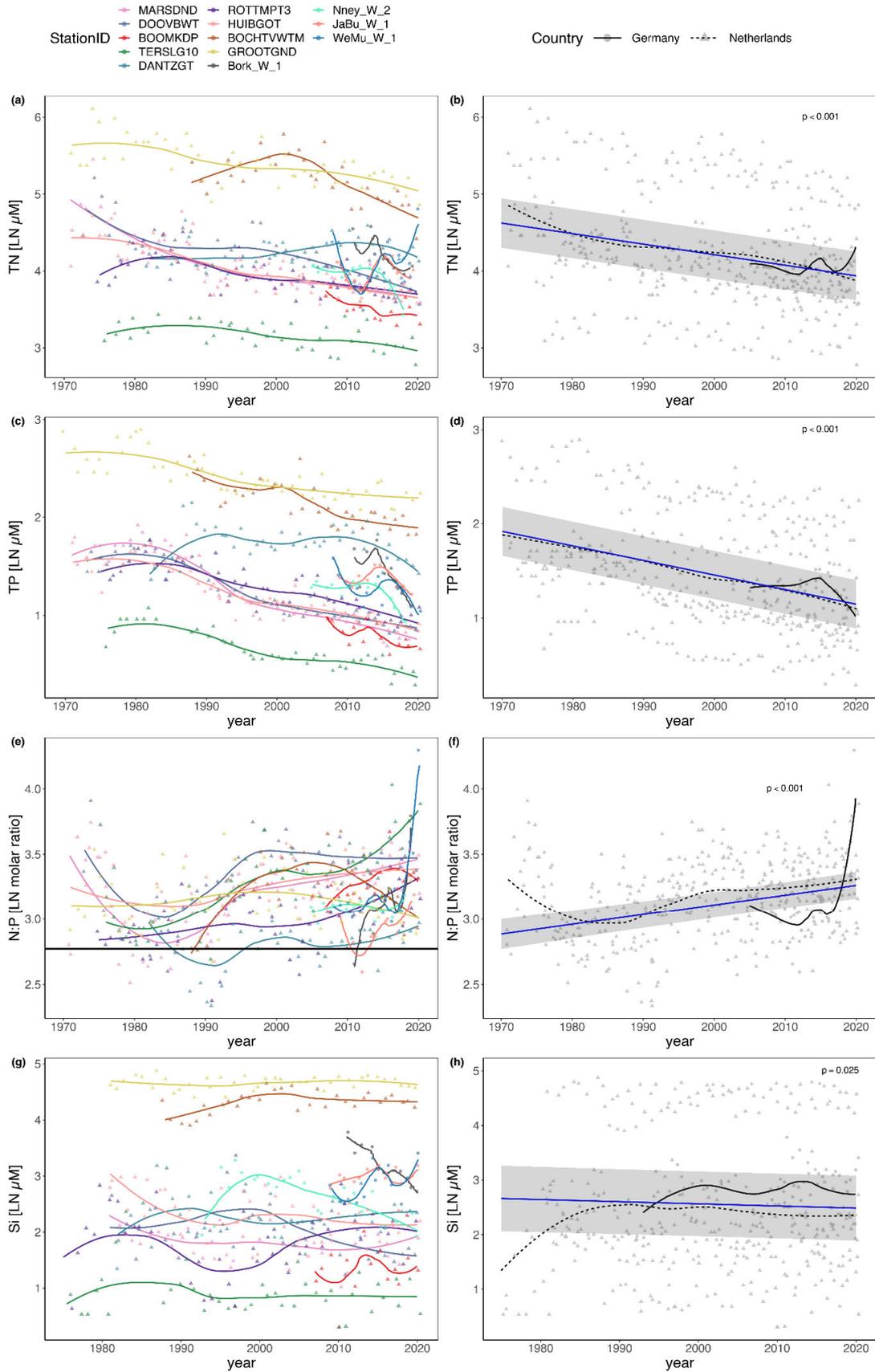
Figure 2.4. Temporal trend of nutrient concentrations at the Wadden Sea coastal stations for TN, TP, their ratio and Si. Left column: Annual means and LOESS trend lines coloured by station. Right column: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous line; NL: dashed line). *Data are LN transformed.*
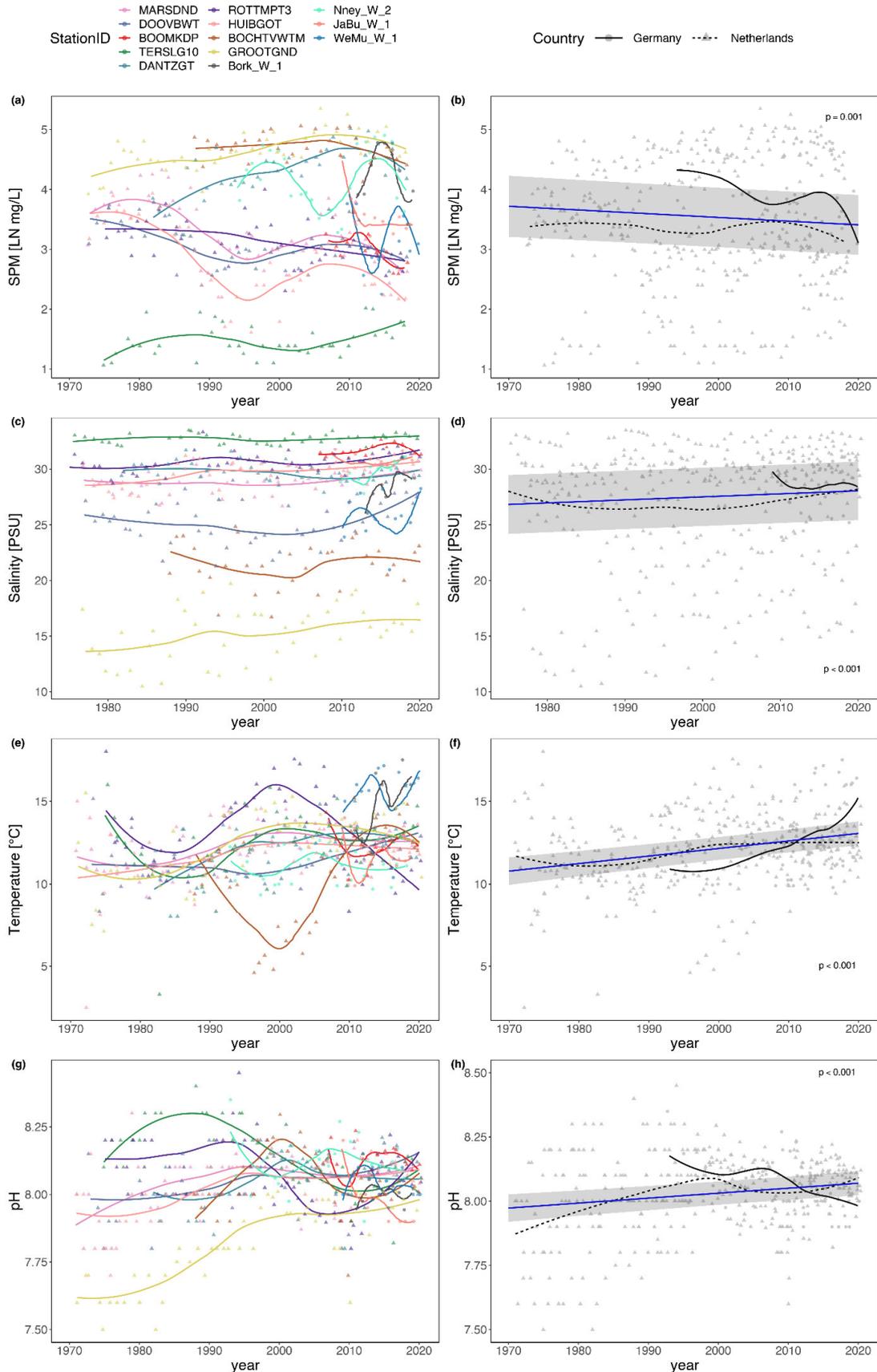
Figure 2.5. Temporal trend of environmental factors at the Wadden Sea coastal stations. Left column: Annual means and LOESS trend lines coloured by station. Right column: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous line; NL: dashed line). *Data of SPM is LN transformed.*

Table 2.3. Results of the linear mixed effect model (LMM) to analyse temporal trends of environmental parameters over the years, considering "StationID" as a random effect. For each response variable, we give estimates for intercept (year = 0) and slope (increase or decrease per year) as well as their significance as fixed effects. For random effects, we give theresidual variance (σ^2), the variance associated to the random terms (τ_{00}), the intra class correlation (ICC, how much of the overall variance is connected to the random term), and the number of stations (N). The full number of observations and the marginal and conditional R^2 values are given.

Predictors	LN.TN Estimates	p	LN.TP Estimates	p	LN.NP Estimates	p	LN.Si Estimates	p
(Intercept)	31.667	**<0.001**	32.264	**<0.001**	-11.938	**<0.001**	10.426	**0.003**
year	-0.014	**<0.001**	-0.015	**<0.001**	0.008	**<0.001**	-0.004	**0.025**
Random Effects								
σ^2	0.05		0.02		0.05		0.14	
τ_{00}	0.33 StationID		0.22 StationID		0.03 StationID		1.17 StationID	
ICC	0.86		0.91		0.36		0.89	
N	13 StationID		13 StationID		13 StationID		13 StationID	
Observations	407		411		407		392	
Marginal R^2/ Conditional R^2	0.086 / 0.877		0.160 / 0.921		0.115 / 0.429		0.002 / 0.890	

Predictors	LN.SPM Estimates	p	Salinity Estimates	p	Temperature Estimates	p	pH Estimates	p
(Intercept)	15.965	**<0.001**	-26.849	**0.035**	-78.822	**<0.001**	4.157	**<0.001**
year	-0.006	**0.001**	0.027	**<0.001**	0.045	**<0.001**	0.002	**<0.001**
Random Effects								
σ^2	0.19		2.07		3.45		0.01	
τ_{00}	0.82 StationID		22.94 StationID		1.44 StationID		0.01 StationID	
ICC	0.82		0.92		0.29		0.33	
N	13 StationID		13 StationID		13 StationID		13 StationID	
Observations	401		385		427		427	
Marginal R^2/ Conditional R^2	0.007 / 0.816		0.005 / 0.918		0.074 / 0.346		0.035 / 0.351	

b) Phytoplankton biomass over time

Different stations showed substantial differences in phytoplankton carbon biomass and different temporal patterns (Fig. 2.6a). Aggregating at the country level, we find that carbon biomass is two orders of magnitude higher in the Dutch than in the German data (Fig. 2.6b). The two countries also show different temporal dynamics with an order of magnitude increase in C biomass in the Netherlands between 1999 and 2014, followed by a slight decrease. In Germany, we find a decline in C biomass, if any trend. As more Dutch than German stations are monitored, the overall trend turns out to be significantly positive but with low explanatory power (Table 2.4).

Figure 2.6. Temporal trend of the phytoplankton biomass measured as carbon (a,b), chlorophyll (c,d) and the carbon to chlorophyll ratio (C:Chl) (e,f) at the Wadden Sea coastal stations. Left column: Annual means and LOESS trend lines coloured by station. Right column: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous line; NL: dashed line). *Data are LN transformed.*
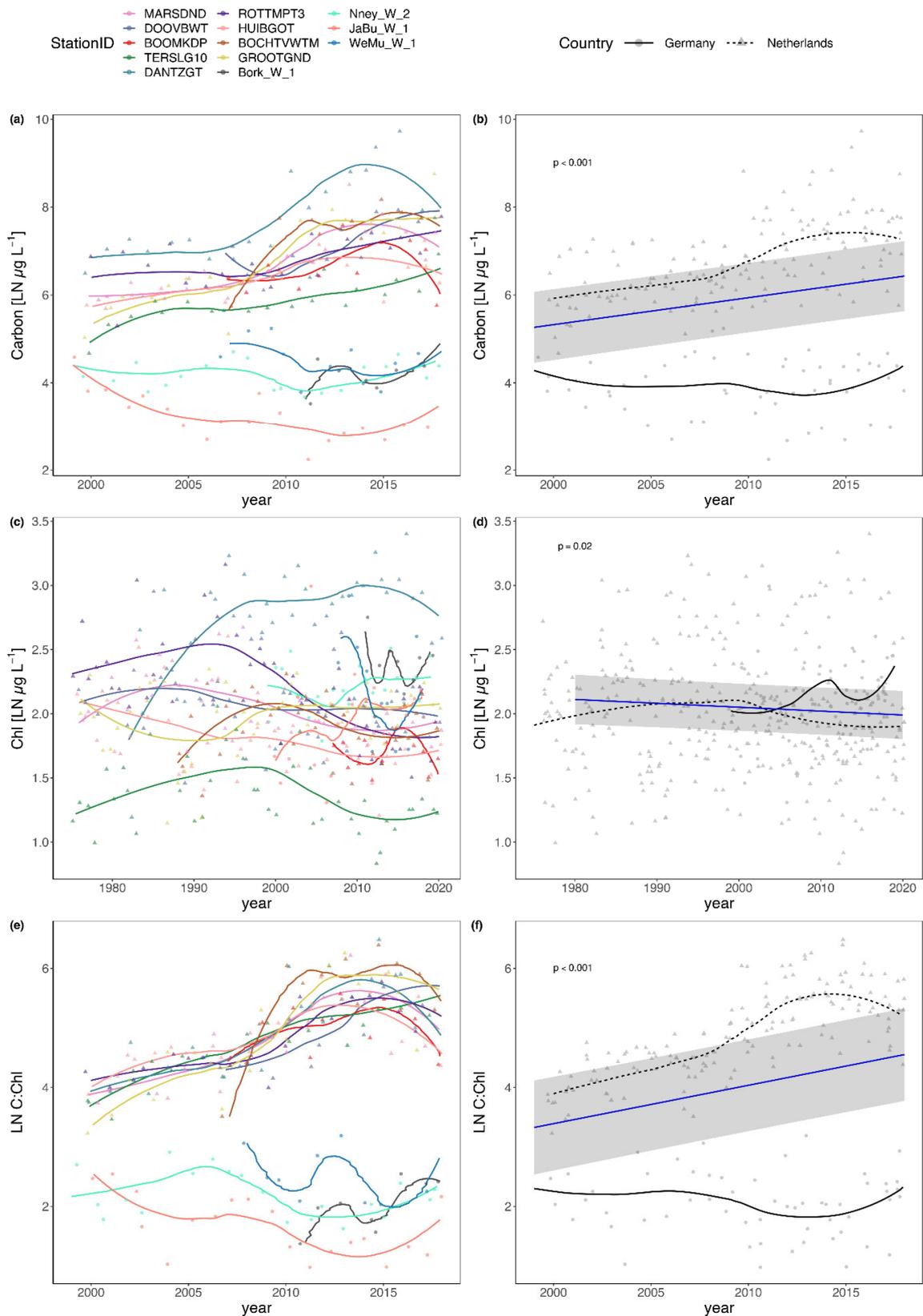
The temporal trend for Chl biomass is negative but even less prominent (Table 2.4), as single stations show little consistent variation, some declining since the late 1980s to early 1990s, others increasing or fluctuating (Fig. 2.6c, d, Table 2.4). In contrast to C-biomass, Chl didn't show strong differences between countries, but exhibited a lot of variance among the stations (Fig. 2.6c, d). The linear C:Chl ratio significantly increased over time (Fig. 2.6e, f, Table 2.4), solely based on the Dutch stations and thus reflecting the increase in carbon biomass (based on literature values) in the Netherlands.

Table 2.4. Results of the linear mixed effect mode, analysing the change in phytoplankton biomass over the years, considering "StationID" as a random effect. For each response variable, we give estimates for intercept (year = 0) and slope (increase or decrease per year) as well as their significance as fixed effects. For random effects, we give the residual variance (σ^2), the variance associated to the random terms (τ_{00}), the intra class correlation (ICC, how much of the overall variance is connected to the random term), and the number of stations (N). The full number of observations and the marginal and conditional R^2 values are given.

	LN Carbon L-1		LN Chl L-1		LN C:Chl	
Predictors	Estimates	p	Estimates	p	Estimates	p
(Intercept)	-117.473	**<0.001**	8.124	**0.002**	-125.592	**<0.001**
year	0.061	**<0.001**	-0.003	**0.021**	0.064	**<0.001**
Random Effects						
σ^2	0.33		0.09		0.30	
τ_{00}	2.06 StationID		0.11 StationID		1.96 StationID	
ICC	0.86		0.54		0.87	
N	13 StationID		13 StationID		13 StationID	
Observations	209		408		206	
Marginal R^2/ Conditional R^2	0.044 / 0.870		0.007 / 0.540		0.050 / 0.873	

c) Functional group biomass over time

Grouping species according to their functional traits can overcome the limitations of microscopy in identifying phytoplankton to species or genus level, and the difficulties of analysing data with many rare species (Cunningham & Lindenmayer 2005, Jansen et al. 2018). Functional group indicators have been shown to be relevant for describing community structure and biodiversity and are more comparable with other studies than species-based indicators (Mouillot et al. 2006).

In this project, five functional groups of phytoplankton were defined to be used in the ecosystem models of the Wadden Sea. These are diatoms, dinoflagellates, flagellates, cyanobacteria and *Phaeocystis*. Different species within a functional group share morphological or physiological characteristics, particularly the presence or absence of certain aspects such as silicate cell walls and flagella. The concept of classifying species into different functional groups based on their functional traits and characteristics serves to better describe their role in the ecosystem.

Functional groups are good at representing the major phylogenetic differences on which they are based. However, they have limited predictive power for what species actually do over time. Measured traits appear to be a nuanced and sensitive measure of functional diversity. As they are easily derived from the monitoring time series itself, it seems reasonable to at least include them in functional diversity.

We calculated the annual biomass of each functional group – diatoms, dinoflagellates, flagellates, cyanobacteria and *Phaeocystis* – as the sum of carbon biomass per sample, then calculated the annual median. Diatoms had the highest biomass over all functional groups and additionally showed the

clearest trend, with increasing biomass over the years at most of the Dutch stations (Fig. 2.7a, b, Table 2.5). Dinoflagellates showed no overall trend as their biomass varied across the stations. In some of the German stations it presented a similar pattern to the diversity measures, with a steep drop followed by an increase in biomass in some stations (Fig. 2.7c, d). Flagellates increased over time in most of the stations (Fig. 2.7e, f), whereas cyanobacteria significantly declined over time (Fig. 2.7g, h). *Phaeocystis* increased, consistently in the Dutch stations, whereas a more non-linear pattern prevailed in the German stations (Fig. 2.7i, j).
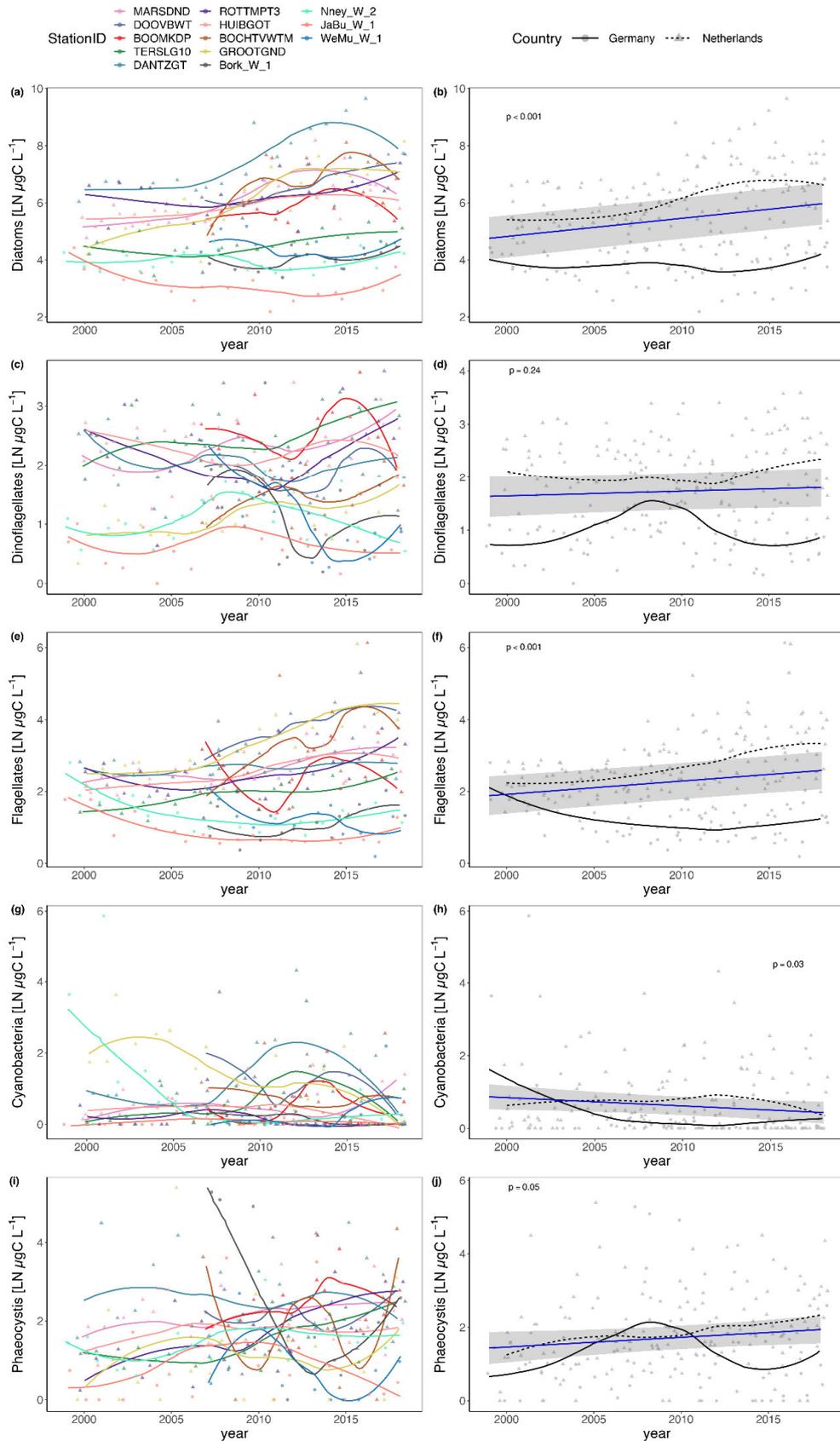
Figure 2.7. Temporal trend of the phytoplankton functional groups measured as the yearly biomass median of Diatoms (a,b), Dinoflagellates (c,d), Flagellates (e,f), Cyanobacteria (g,h) and *Phaeocystis* (i,j) at the Wadden Sea coastal stations. Left column: Annual means and LOESS trend lines coloured by station. Right columns: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous lines; NL: dashed lines). *Data input: annual median.*

Table 2.5. Results of the linear mixed effect model, analysing the change in phytoplankton functional groups' biomass (LN µgC L^{-1}) over years. Station ID is included as random effect. All details as in Table 2.3.

Predictors	Diatoms Estimates	p	Dinoflagellates Estimates	p	Flagellates Estimates	p	Cyanobacteria Estimates	p	Phaeocystis Estimates	p
(Intercept)	-122.741	**<0.001**	-16.266	0.288	-71.945	**<0.001**	46.554	**0.032**	-51.467	0.062
year	0.064	**<0.001**	0.009	0.239	0.037	**<0.001**	-0.023	**0.034**	0.026	0.054
Random Effects										
σ^2	0.45		0.32		0.39		0.64		1.04	
τ_{00}	1.69 StationID		0.38 StationID		0.82 StationID		0.17 StationID		0.26 StationID	
ICC	0.79		0.55		0.68		0.21		0.20	
N	13 StationID		13 StationID		13 StationID		13 StationID		13 StationID	
Observations	213		213		213		213		213	
MarginalR^2/ ConditionalR^2	0.052 / 0.802		0.003 / 0.546		0.031 / 0.688		0.018 / 0.221		0.015 / 0.213	

d) Biomass of dominant species over time

We also analysed the species contribution to biomass over time in the Wadden Sea stations. We first calculated the annual mean biomass of each taxon and analysed their relative biomass over years (stacked bars in Fig. 2.8). Taxa with less than 20% of relative biomass were coloured in grey, but still separated by the black horizontal lines on each bar. In the Dutch stations, most of the years were dominated by the diatom genus *Thalassiosira* sp., reaching up to 90% of the total biomass in some years and stations (e.g., in DANTZGT, year 2010 and ROTTMPT3, year 2011). In the German stations we observed a higher diversity of taxa contributing to biomass over the years, with a few years being dominated by *Phaeocystis* sp. (Fig. 2.8). Therefore, turnover seems to affect more the dominant species in the German stations than in the Dutch stations.
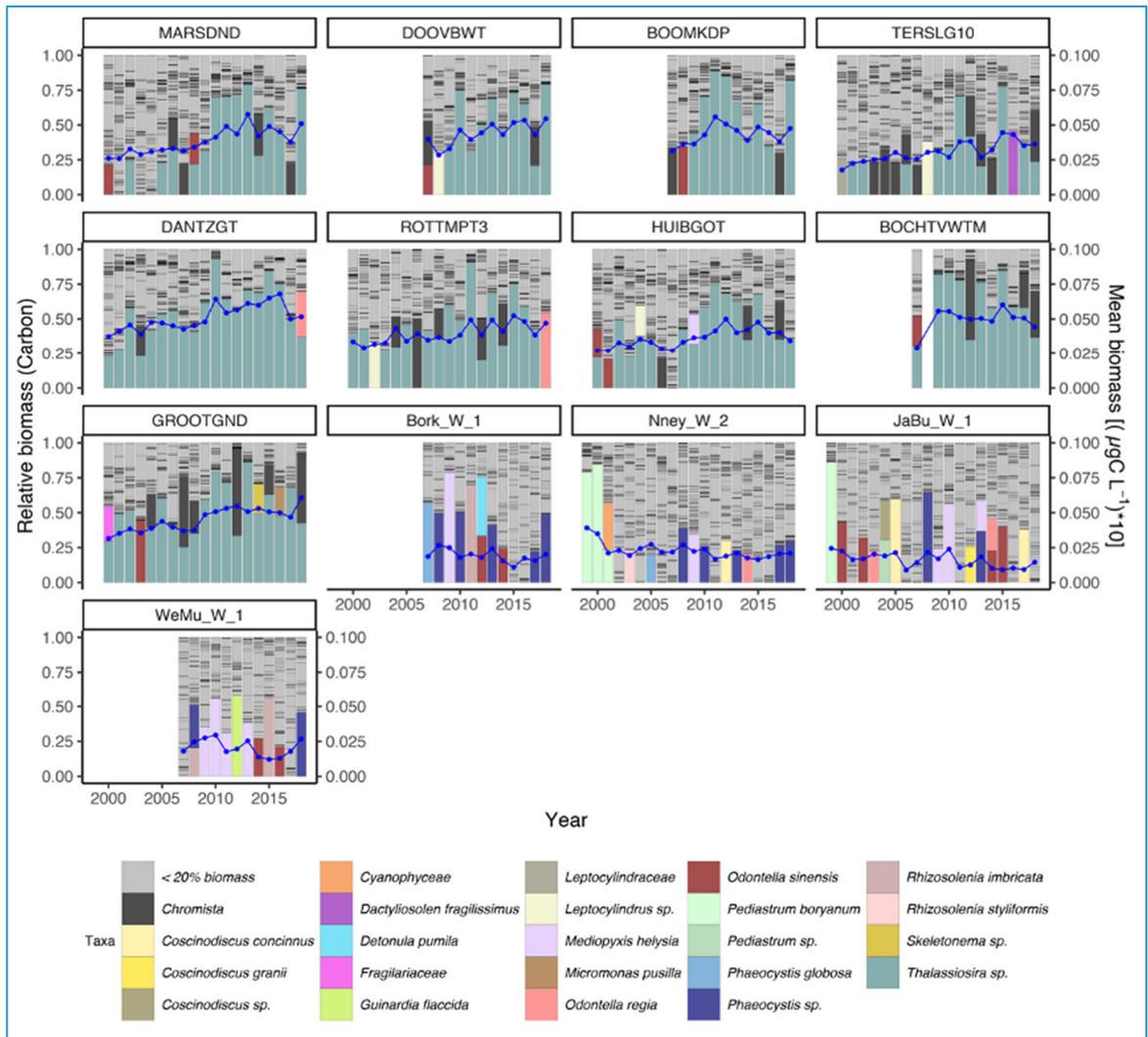
Figure 2.8. Relative carbon biomass of the phytoplankton taxa over time in the Wadden Sea stations (bars). Taxa with less than 20% of biomass contribution per year were grouped and coloured in grey. Annual mean biomass is shown on the right axis (blue line).

2.2.4 System understanding from combined data sets

2.2.4.1 Phytoplankton biomass and environmental conditions

Environmental factors explain a substantial amount of variance in the year-to-year variation in total phytoplankton biomass, 16-62% for C-biomass and 24-29% for Chl in the LMM (Table 2.5) and 58-60% in the SEM. Adding random intercepts for stations raised the conditional explained variance up to 92% (Table 2.6). Correlations (Fig. 2.9), LMM (Table 2.6, Figs. 2.10-2.11) and SEM (Figs. 2.12-2.13) often were in general agreement as to which factors were driving variance in biomass, but detailed differences between the two biomass measures and between countries were abundant. The difference between biomass measures partly reflects the differences in C estimates, although C- and Chl-biomass were positively correlated in both the Dutch and German data (Fig. 9, r=0.52 for NL, r = 0.36 for DE).
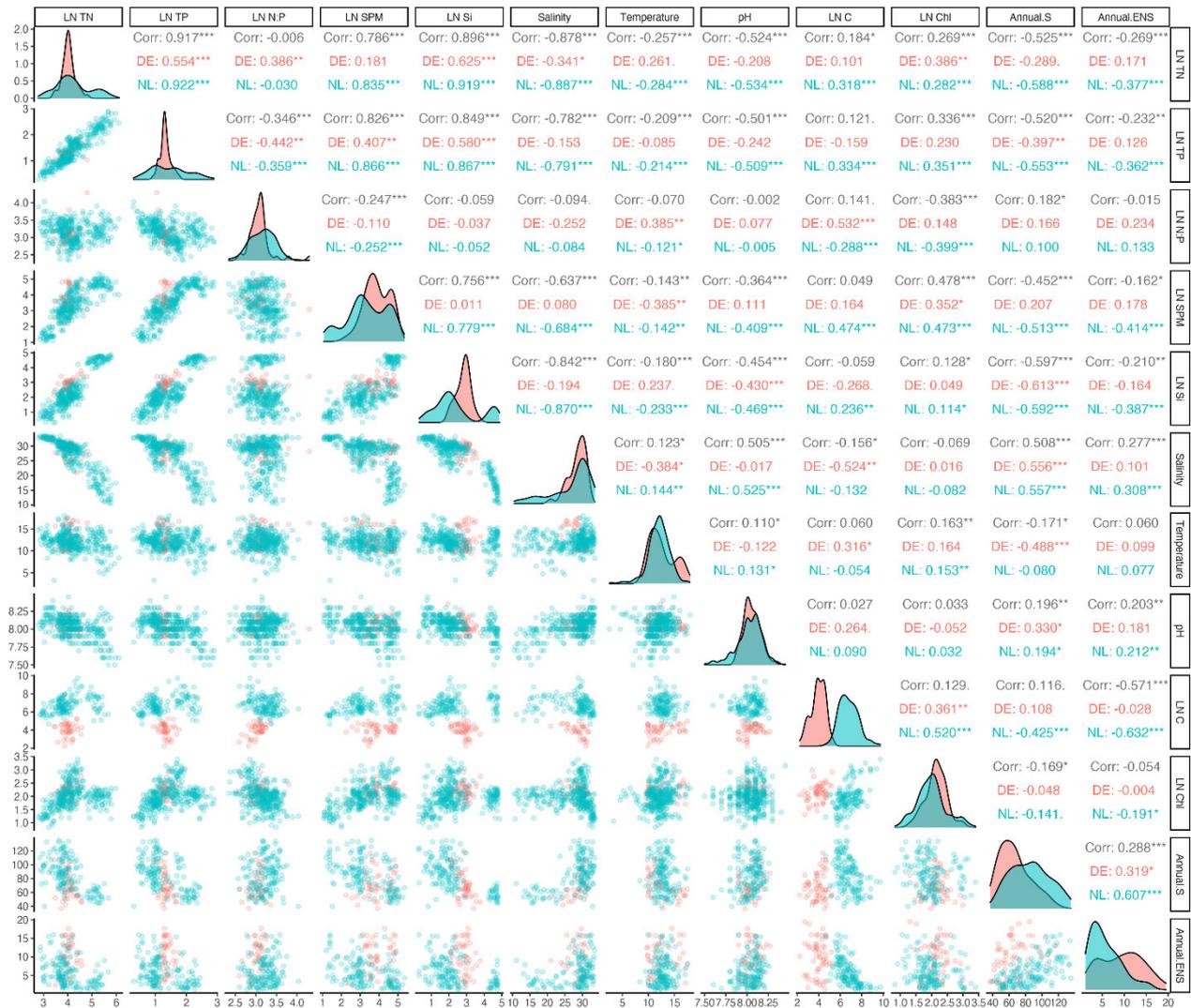
Figure 2.9. Correlation matrix among all environmental parameters and phytoplankton measures. The correlation coefficients are coloured according to country, NL in blue and DE in red. Asterisks next to correlation coefficient represent the significance level: *** *p<0.001*, **p<0.01, *p<0.05*. For more information on the variables, see table 2.5. *Data input: annual medians.*

We expected total biomass to increase with N and P availability. Indeed, we found positive pairwise correlations (Fig. 2.9) for both metrics (C and Chl) and both nutrients (N and P), which were stronger for Chl-biomass than for C-biomass and for Dutch than for German stations. LMMs detected the same positive association for TP (Table 2.6) and TN (Table A3.1 in the Annex) for Chl overall and in the Dutch data (Table A3.1 (Annex), Fig. 2.9). For C-biomass, effects were not significant as this relationship was partly covered by Si concentrations and salinity. Bivariate representation of biomass to nutrients shows a peak biomass appearing at ~90 µM TN and ~6 µM TP (Fig. 2.10a-d). The decline in biomass at higher nutrient levels however is strongly associated to the low-salinity stations GROOTGND and BOCHTVWTM. C-biomass further showed the previously described difference between the Dutch and German datasets, which might explain the overall lower consistency of the results for C compared to Chl. SEM found similar effects as TP was a significant driver of biomass for both C and Chl measurement, with an additional positive TN effect on Chl biomass (Figs. 2.10-2.11).

A clear separation between N or P as the main driver is difficult given the very high correlation between both nutrients. The SEM picks TP over TN for C-biomass and indeed the N:P ratios indicate rather a P- than a N-limitation (Fig. 2.10 e-f, see also below). Based on correlations, biomass significantly declined

with increasing N:P (indicative of more P-limitation) in Dutch but not German stations, where C-biomass even increased (Fig. 2.10-2.11). The LMM did not find significant slopes with N:P and the SEM could not encompass this (as it incorporates both TN and TP). Based on this evidence we lean towards concluding that P-limitation is the main state of the system at present. A more detailed answer on this would need a bioassay approach (see recommendations, Chapter 6).

Median dissolved silicate concentrations showed a positive bivariate correlation to C-biomass and Chl-biomass in the Dutch data (Figs. 2.10, 2.11 g-h), but no trend in the German data. In the multifactorial assessments this relationship turned consistently negative in the SEM (Figs. 2.12-2.13) and - for Chl-biomass – in the LMM (Table 2.6). This conversion of effects potentially reflects the high correlation between Si and N as well as P concentrations (Fig. 2.9), thus the general positive nutrient – biomass trend is already captured by TN and TP. After controlling for this general trend, years and stations with higher Si concentrations obviously tended to have lower biomass.

Any conclusion on nutrient limitation based on these analyses has the caveat that a potential light limitation cannot be assessed, which is a clear recommendation (see below). The only variable related to light is SPM, which however is partly reflecting biomass in itself as phytoplankton is a major part of the suspended particles. Consequently, a positive relationship emerges between SPM and biomass (both C and Chl) in the correlations (Fig. 2.9) and the LMM (Table 2.6), which again was stronger for Chl than C and for Dutch than German data. SPM could not be incorporated in the SEM.

The bivariate correlation between biomass and salinity tended to be negative, which mainly reflected that the station GROOTGND, which had the lowest salinity (<20 PSU), also showed exceptionally high nutrient concentrations and thus high C-biomass (Figs. 2.10-2.11). By contrast, towards full marine salinity (>27 PSU), both C- and Chl-biomass declined with salinity again. Controlling for the nutrient-salinity interaction in the LMM and SEM found negative salinity effects on biomass.

When significant, higher temperatures were consistently associated with higher biomass, in correlations (Fig. 2.10) and in the LMM (Table 2.6) for Chl-biomass in the Netherlands and C-biomass in Germany. Whether this is a direct temperature effect on algal growth or an indirect effect (e.g., via higher remineralisation) cannot be obtained from the data. The SEM did not detect any significant temperature effect except for weak negative effect on C-biomass. High Chl-biomass co-occurred with high pH in both LMM (overall and NL only) and SEM (Fig. 2.12-2.13). As algal photosynthesis affects the pH, the causality is not identifiable.

A major source of biomass variance in phytoplankton biomass remains elusive, as we have no direct information on the extent of zooplankton grazing and benthic filter-feeding on phytoplankton (see recommendations, Chapter 6).
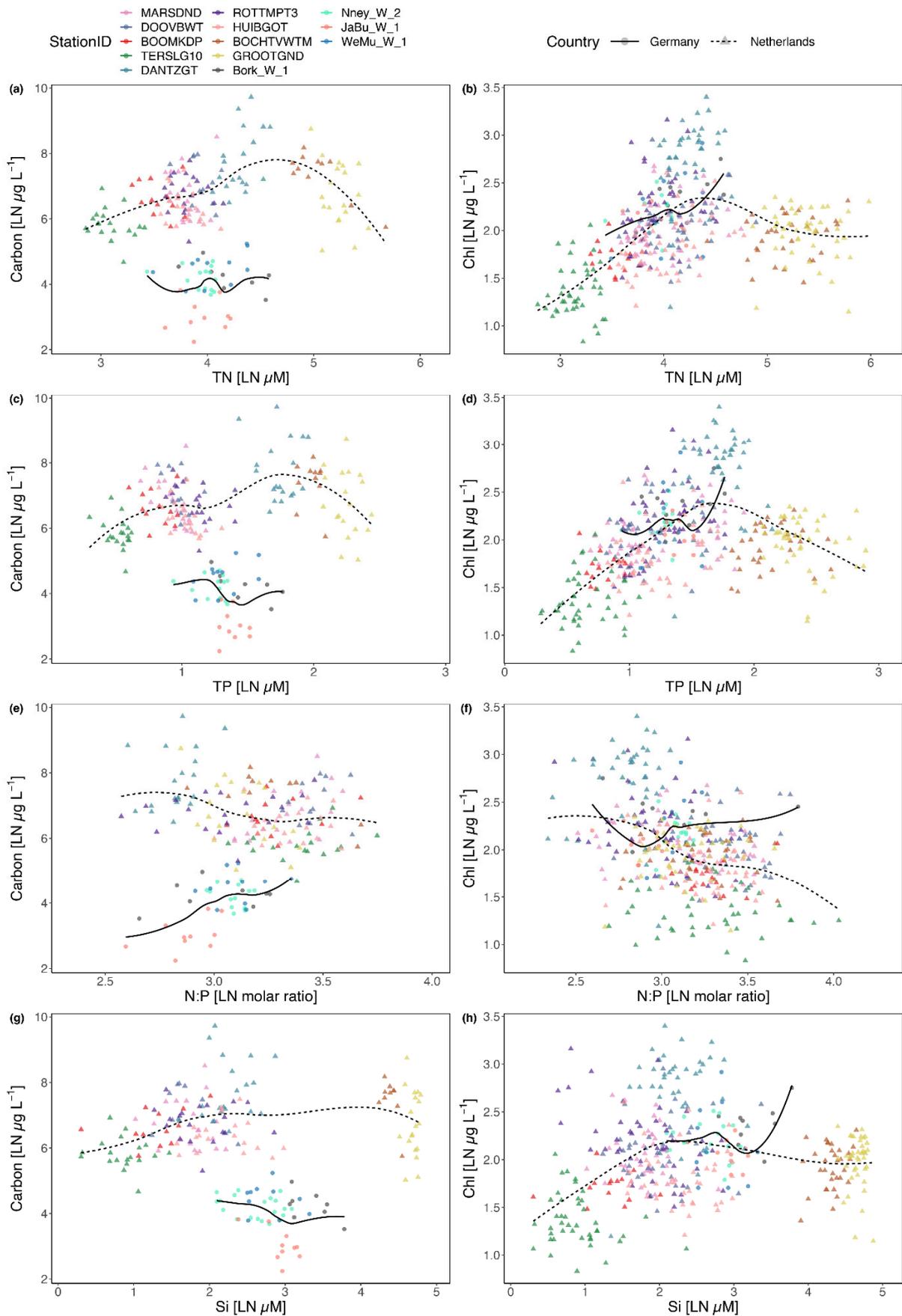
Figure 2.10. Carbon (left column) and chlorophyll (right column) as phytoplankton biomass estimates plotted against nutrients N, P, their ratio, and Si. Lines represent a loess fit for NL stations (dashed) and DE stations (continuous line). *Data input: annual scale.*
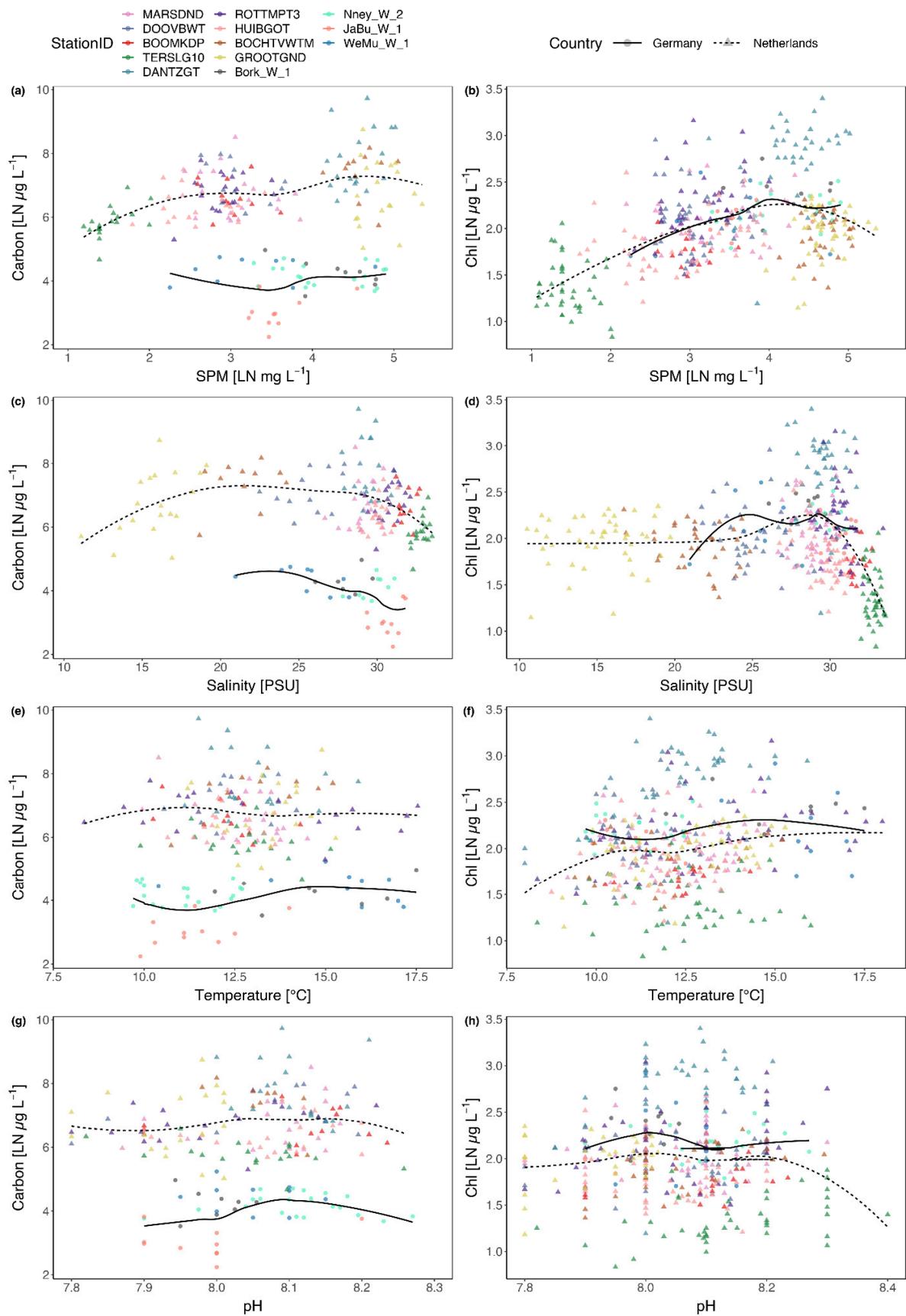
Figure 2.11. Carbon (left column) and chlorophyll (right column) as phytoplankton biomass estimates plotted against SPM, salinity, temperature and pH. Dashed line represents NL stations and continuous line, DE stations. *Data input: annual scale.*

Table 2.6. Results of the linear mixed effect model, analysing the effects of environmental factors on phytoplankton biomass (carbon and chlorophyll), considering "StationID" as a random effect. Bold numbers indicate significant predictors. The overall model for carbon biomass is highlighted as it is affected by the discrepancy in carbon estimates. When outputs differed between NL and DE, we highlighted the estimates in grey. Conditional R^2 for the last of the six models could not be obtained as no variance was associated to the random effects. *Data input: annual median.*

	Carbon (LN μgL^{-1})						Chl (LN μgL^{-1})					
	all		NL		DE		all		NL		DE	
Predictors	Estimates	p	Estimates	p	Estimates	p	Estimates	p	Estimates	p	Estimates	p
(Intercept)	1.170	0.801	-3.326	0.528	10.622	0.146	-4.986	<0.001	-5.539	<0.001	1.681	0.702
LN TP	-1.760	0.002	-1.131	0.063	-1.027	0.229	0.419	<0.001	0.453	<0.001	-0.025	0.958
LN NP	-0.159	0.652	-0.245	0.573	0.272	0.614	0.135	0.160	0.161	0.125	0.163	0.518
LN Si	0.110	0.497	0.360	0.037	-0.693	0.047	-0.131	0.002	-0.136	0.002	-0.025	0.896
LN SPM	0.420	0.007	0.712	<0.001	0.244	0.130	0.126	0.002	0.128	0.005	0.165	0.036
Salinity	0.096	0.035	0.075	0.073	-0.151	0.009	-0.008	0.497	-0.009	0.450	0.013	0.477
Temperature	-0.031	0.381	-0.034	0.450	0.118	0.011	0.046	<0.001	0.047	<0.001	0.031	0.138
pH	0.441	0.383	0.934	0.111	-0.295	0.731	0.691	<0.001	0.741	<0.001	-0.164	0.751
Random Effects												
σ^2	0.39		0.48		0.12		0.08		0.08		0.05	
τ$_{00}$	3.70 StationID		0.13 StationID		0.10 StationID		0.06 StationID		0.08 StationID		0.00 StationID	
ICC	0.91		0.22		0.45		0.46		0.51			
N	13 StationID		9 StationID		4 StationID		13 StationID		9 StationID		4 StationID	
Observations	182		148		34		338		303		35	
Marginal R^2 / Condit. R^2	0.163 / 0.921		0.226 / 0.395		0.620 / 0.791		0.282 / 0.613		0.289 / 0.655		0.244 / NA	
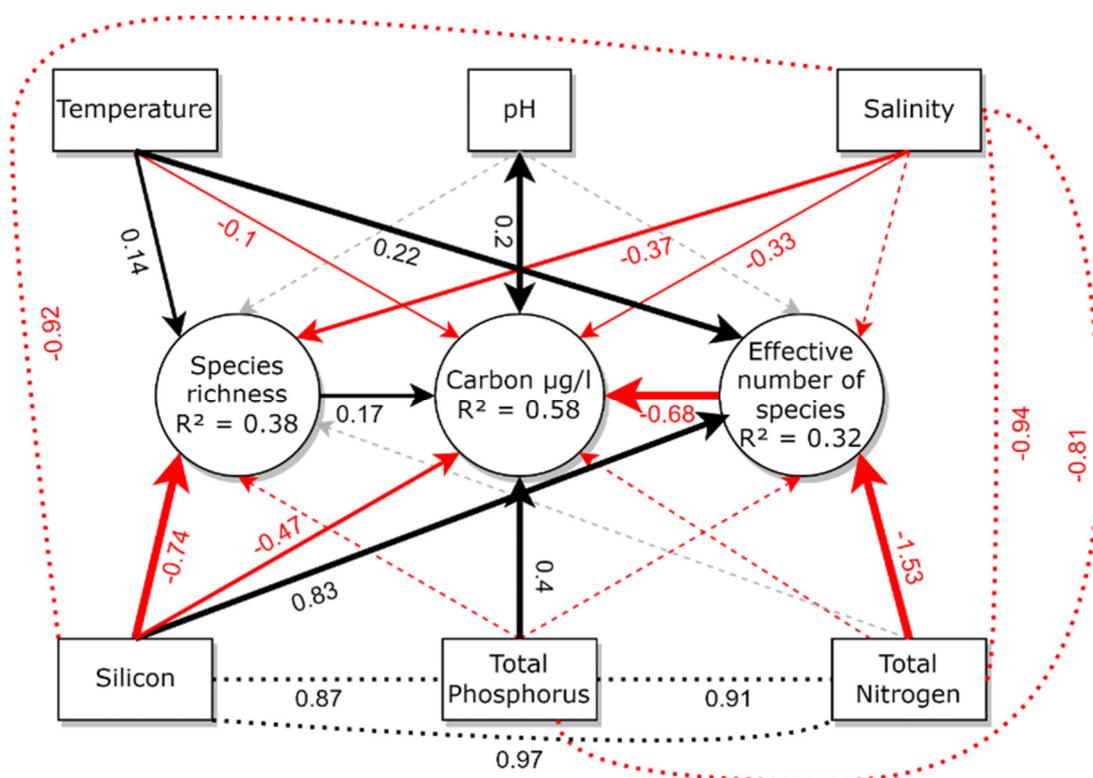
Figure 2.12. Analysis of annual data using structural equation model (SEM). Yearly average phytoplankton biomass as carbon and annual diversity (raw species richness and effective species number (analogous to evenness) modelled as a response to six environmental factors, biomass additionally affected by diversity. Black arrows = positive effects, red = negative, solid lines = significant effects, dotted lines not significant. Numbers are standardized path coefficients that can be interpreted as correlation coefficients.
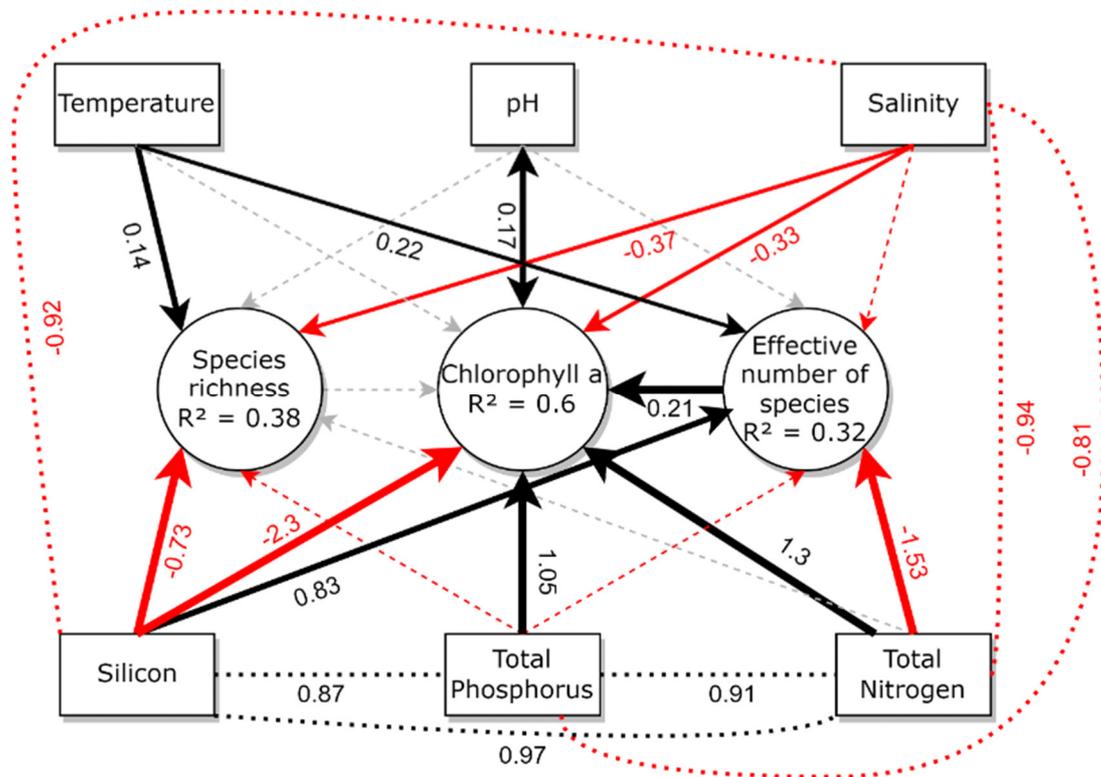
Figure 2.13. Analysis of annual data using structural equation model (SEM). Yearly average phytoplankton biomass as chlorophyll and annual diversity (raw species richness and effective species number (analogous to evenness) modelled as a response to six environmental factors, biomass additionally affected by diversity. Black arrows = positive effects, red = negative, solid lines = significant effects, dotted lines not significant. Numbers are standardized path coefficients that can be interpreted as correlation coefficients.

Summary: Phytoplankton biomass reflects changes in the Wadden Sea environment over time and between stations. Biomass generally increases with increasing nutrient concentrations, with N, P and Si contributing, but some evidence pointing towards a preponderance of P-limitation at the interannual scale. Thus, efforts to control phytoplankton biomass via nutrient reductions need to progress beyond reducing N alone (see different nutrient limitation periods as shown in Fig. 2.15). Biomass decreases towards more saline (farther away from land) and increases towards warmer conditions. So far, Chl seems to achieve more consistent results between countries and approaches, reflecting that C-biomass shows a strong difference between countries. However, Chl per cell is affected by light and thus part of the observed trends may derive from differences in irradiance. Light limitation and mortality via zooplankton grazing or benthic filter feeders are two potential constraints on phytoplankton biomass that are not assessed in the monitoring programs.

2.2.4.2 Phytoplankton dominance, functional groups and limiting factors

a) Dominant species as indicators of nutrient conditions

To investigate which species might be indicative of the nutrient status, we first selected the dominant taxa. For each species, we calculated its mean annual biomass, its mean proportional contribution to sample biomass and its frequency of occurrence. From 429 taxa in the data set, 106 were above median in all three categories. Of these we de-selected those that were rare (less than 1% of biomass across sampled) and infrequent (less than 75 occurrences total). From the 69 remaining species, we further reduced to 41 by focusing on species that were determined to species-level and by allowing only a few species from some of the dominant genera. The resulting set of potential indicator species was then related to total N, total P and their ratio as well as silicate.

For each of these species at each station, we obtained an average annual species biomass (carbon) and annual mean proportion of total C-biomass. The former can be considered the absolute response to nutrients, the latter a relative response in comparison to the rest of the community.

Using a LMM, we calculate the slope between (log-transformed) nutrients and (log-transformed) absolute or relative biomass for each species, always using StationID as a random effect. Only half of the selected species showed significant relations between their absolute or relative biomass and any of the four-nutrient axis (Fig. 2.14).
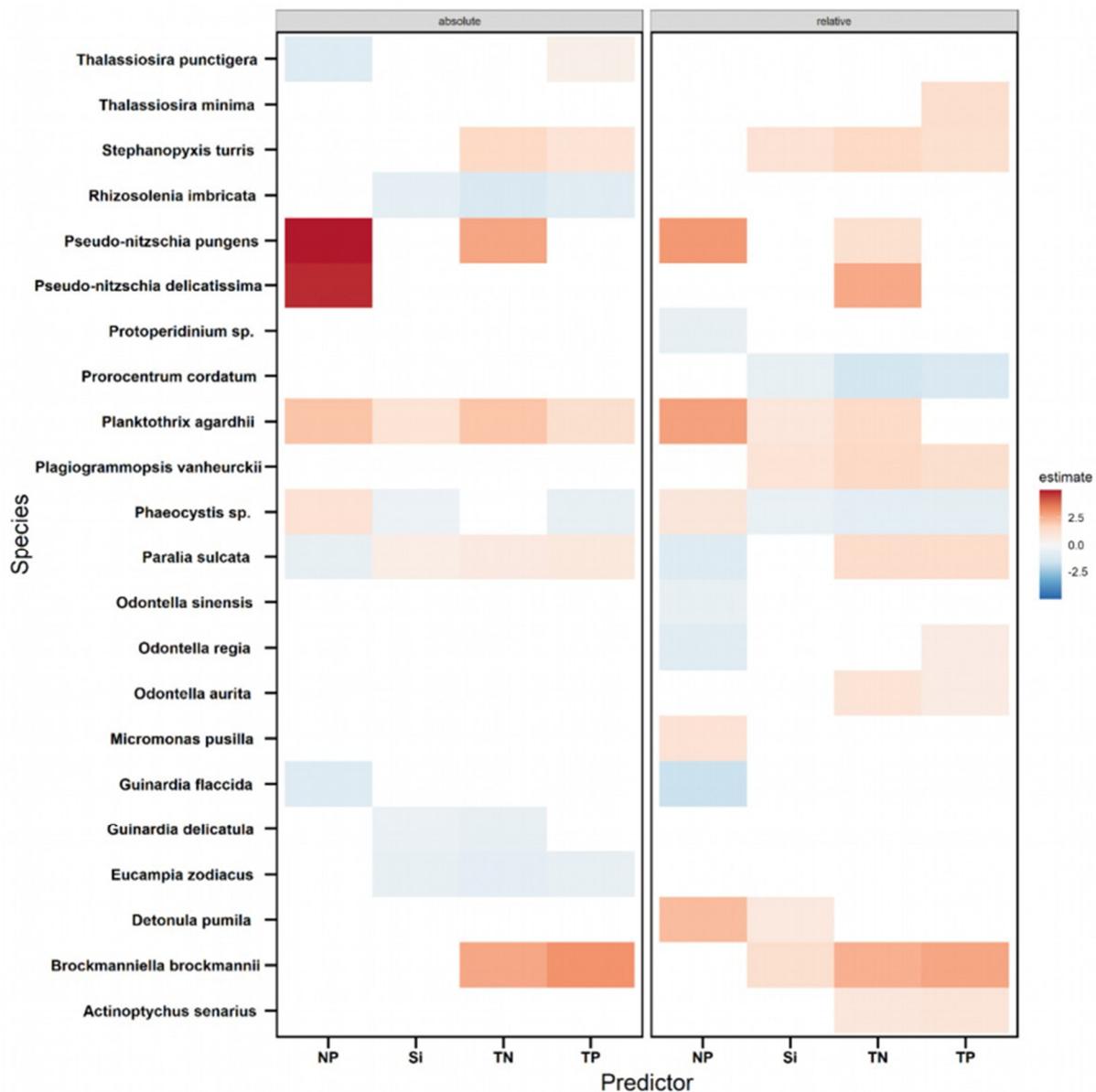
Figure 2.14. Slopes of species to nutrients and their ratio derived from univariate LMM with StationID as additional random factor. Slope estimates are colour coded for each regression where p <0.05, with red gradient indicating positive relationship and blue gradient negative relationship, blanks indicate non-significant regressions.

Among these, a few species stood out: The potentially toxic diatoms *Pseudo-nitzschia pungens* and *P. delicatissima* increased with increasing TN concentrations and N:P ratios, leading to a higher proportion at high N:P and TN. The diatoms *Guinardia flaccida* and *Eucampia zodiacus* declined with increasing TN. The diatom *Brockmanniella brockmanii* increased with both TN and TP in both absolute and relative terms. Most other relationships were weaker and confirmed well-known expectations such that a range of diatoms increased its relative share in biomass at high Si concentrations, whereas the proportion of *Prorocentrum cordatum* and *Phaeocystis* decreased.

Summary: Phytoplankton species respond to nutrient gradients, but rarely so consistent that they can serve as indicator species for nutrient conditions.

b)   Nutrient limitation

In contrast to the previous section on annual mean data, we moved to sample-based data to investigate if limitations by N and/or P are detectable over the course of the year. Both TN and TP show the expected seasonal pattern with high winter concentrations followed by a reduction towards summer and then increasing concentrations in late fall (Fig. 2.15a, b). This pattern was very regular for TN and more variable between stations for TP, reflecting the potential influence of rapid P-remineralization. The decline in nutrients coincides with increasing biomass in most of the algal functional groups (Fig. 2.16), which was especially pronounced for summer dinoflagellate blooms and spring *Phaeocystis* blooms. The overall most biomass-rich phytoplankton group, diatoms, showed a less clear seasonal pattern and was abundant throughout, but showed an early spring peak in most stations (NB: the scale in Fig. 2.15 is log-transformed, masking the differences in the most dominant group).

Indicative for the question of limitation is the N:P ratio, which again showed a very consistent seasonal pattern between years and stations (Fig. 2.15 c): N:P ratios peak in early spring, and decline towards a minimum around August, before they increase again. Difference between stations is much less than for the concentrations. Overall, N:P is higher than 22, a proposed indicator for P-limitation (Guildford & Hecky 2000), pointing towards P-limitation. Especially during the spring bloom, P-limitation is highly likely, whereas towards summer (and dinoflagellate dominance), N limitation is at least a possibility. In addition, silicate may be limiting for diatoms during parts of the growing season (Officer & Ryther 1980, Egge & Aksnes 1992, Prins et al. 2012). A test on limiting nutrients would be a bioassay approach to see the responses to nutrient spikes (see recommendations, Chapter 6). We reiterate that the conclusion of limitation is made in the absence of information on how limiting light is. We also point towards information from bioassays worldwide indicating that different species in an assemblage can be limited by different resources and co-limitation is the norm rather than an exception (Elser et al. 2007, Harpole et al. 2011).
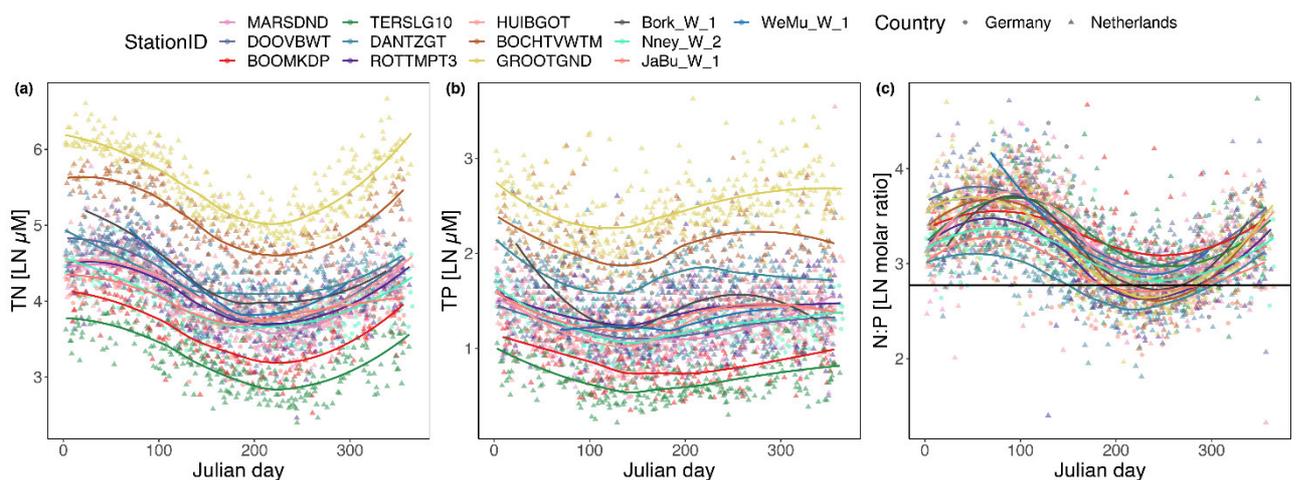
Figure 2.15. Seasonal trend of total nutrients at the Wadden Sea coastal stations. Horizontal line in panel c is N:P = 16. Coloured lines are loess fits per station. *Data input: Sample data.*
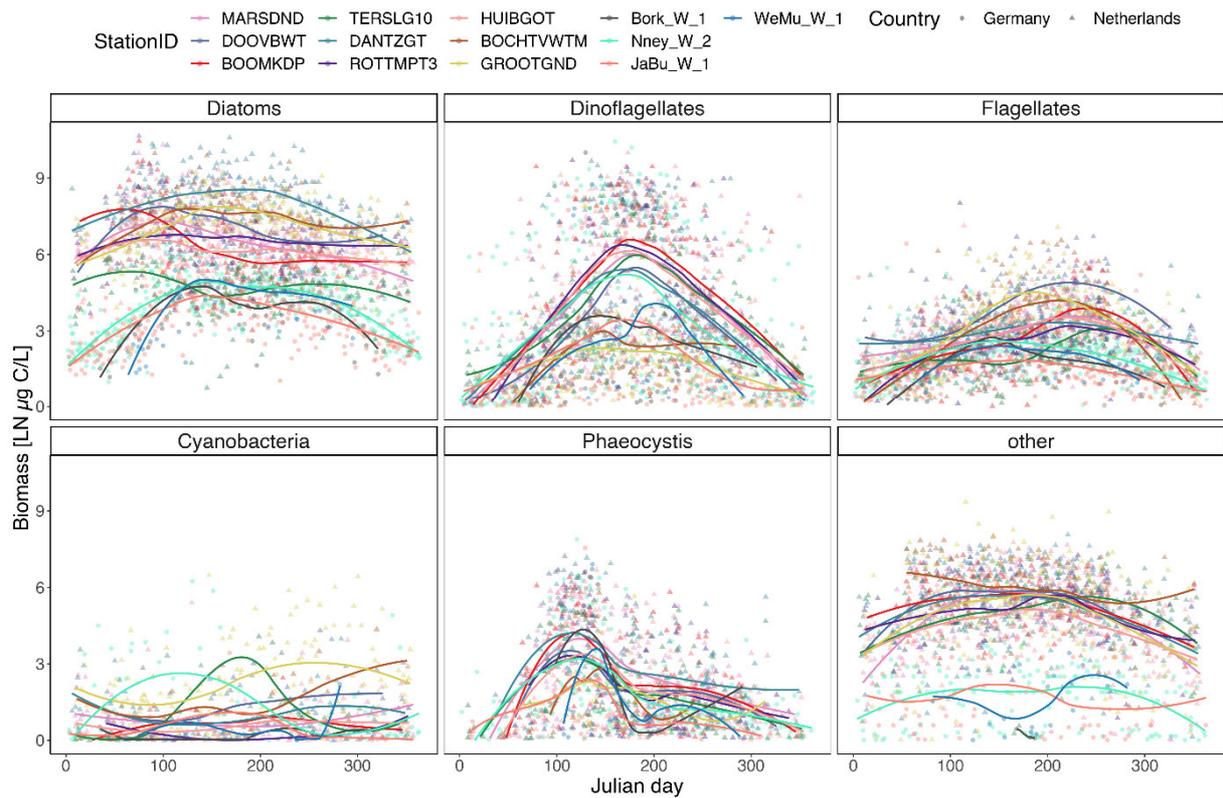
Figure 2.16. Seasonal trend of the biomass of functional groups at the Wadden Sea coastal stations. *Data input: Julian day median.*

Summary: N:P ratio strongly point towards a potential P-limitation, especially for the spring and early summer phases. Lower TN concentrations in late summer at least indicate a potential for N-limitation late in the growing season.

2.3 System understanding based on ecosystem modelling

2.3.1 Short model descriptions

a) Southern North Sea – Generalized Plankton Model (SNS-GPM)

The Southern North Sea – Generalized Plankton Model (SNS-GPM) is a 3D coupled ecosystem model. The configuration that was used in this project is developed based on the setup that was used in Lenhart et al. (2022), for the derivation of threshold values under the OSPAR convention (see Chapter 4.1). The hydrodynamical component is the General Estuarine Transport Model (GETM; Burchard & Bolding, 2002). The model domain is the southern North Sea, extending from 1°W to 9.5°E and 51°N to 55.5°N on a curvilinear grid (Fig. 2.17). The grid's resolution ranges from 1.5km in the Wadden Sea to 4km at the northern boundary. The vertical domain is divided into 20 equally spaced, depth following layers. GETM provides sea surface elevation, temperature, and salinity.
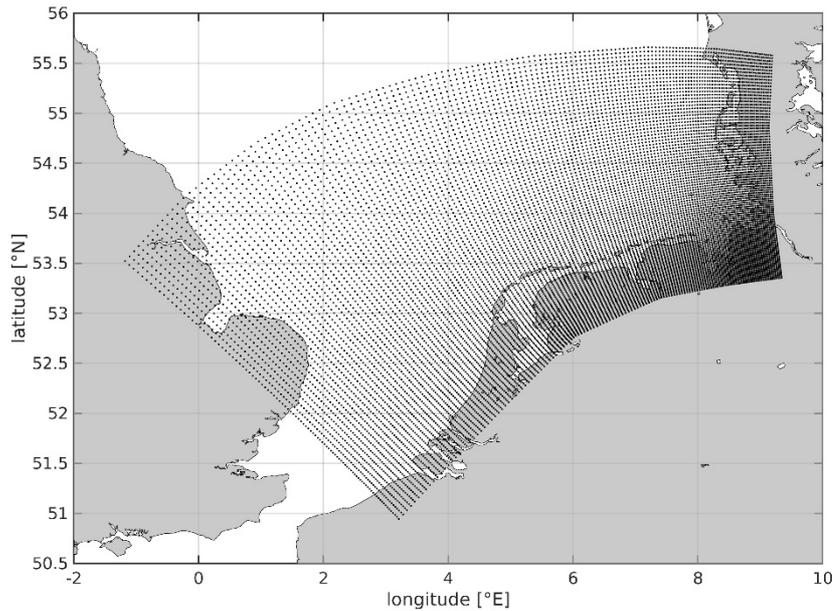
Figure 2.17. Map of the SNS-GPM model domain. Dots are grid points.

The biological module is the Generalized Plankton Model (GPM; Kerimoglu et al. 2017, Kerimoglu et al. 2020). In this most current version, the representation of silt specific attenuation was updated to feature satellite data (see Chapter 2.3.2). Further, two more functional plankton groups were introduced. These modifications made re-calibration necessary. All relevant parameters can be found in Chapter 2.3.2. The model can simulate the cycles of nitrogen, phosphorus, and silica, as well as organic carbon and dissolved oxygen. Detritus that sinks to the floor is remineralized in the sediment. In the current configuration, there are a total of six plankton groups, of which three are classed as purely autotrophic phytoplankton (diatoms, nanoflagellates and *Phaeocystis*), two as purely heterotrophic zooplankton (micro- and mesozooplankton), and one as mixotrophic plankton, being capable of both hetero- and autotrophy. *Phaeocystis* and mixotrophs were added in the harmonization stage during this project (Chapter 2.3.2).

The atmospheric boundary conditions are taken from the COSMO-CLM atmospheric hindcast at 0.22° resolution (Geyer 2014). This includes precipitation, total cloud cover, mean sea level pressure, relative humidity, air temperature at 2m above sea level, and zonal and meridional wind components at 10m above sea level. Evaporation is calculated by GETM, as was the shortwave radiation at the sea surface, using the total cloud cover, provided by COSMO-CLM, astronomical functions that were built in GETM, and seasonal changes in surface albedo according to Payne (1972). Longwave radiation was calculated following Clark et al. (1974). Momentum and heat fluxes were computed using bulk formulae by Kondo (1975). Atmospheric deposition of reduced and oxidized nitrogen, added to modelled NO_3 and NH_4, respectively, was taken from the European Monitoring and Evaluation Programme (EMEP). River input was given by the ICG-EMO data set for the current state and 2.8mg/l scenarios, as well as the two historic scenarios (see Chapter 4.1). Horizontal open boundary conditions at the south-western and northern boundary were given by the DCSM-FM model for temperature, salinity, nutrients and detritus, as well as TRIM-NP-2D (Weisse et al. 2015) for the surface elevation.

b) Ems estuary water quality model (WAQ)

The model of the Forschungsstelle Küste (FSK) is a three-dimensional Eulerian model, composed of 21 vertical sigma layers with a horizontal curvilinear grid of resolution from 10 m to 100 m. The FSK-model reproduces following hydrodynamic, morphodynamic and water quality parameters:

I.   hydrodynamic: flow velocity, water level, salinity, temperature
II.   morphodynamic: fractional sediment concentration, bed composition
III. by means of (I. and II.) water quality parameters: Chl-a concentration, concentration of four phytoplankton functional groups (diatoms, flagellates, dinoflagellates and *Phaeocystis*), total nitrogen (TN), total phosphorus (TP), total silica, and dissolved oxygen (OXY)
IV.   and through the newly developed modules for phosphate-remineralization in this project the rate of nutrients exchange between water column and sediment layer, particularly phosphate.

Model setup and further new developments were performed through the open-source numerical framework of (i) Delft3D-Flow coupled with (ii) D-Water Quality (DELWAQ). In this way, Delft3D-Flow was applied to model the hydrodynamic and morphodynamic parameters. For modelling the water quality parameters, D-Water Quality (DELWAQ) was applied. All numerical model computations were performed for the same bathymetry as is illustrated with the computational grid in Figure 2.18.
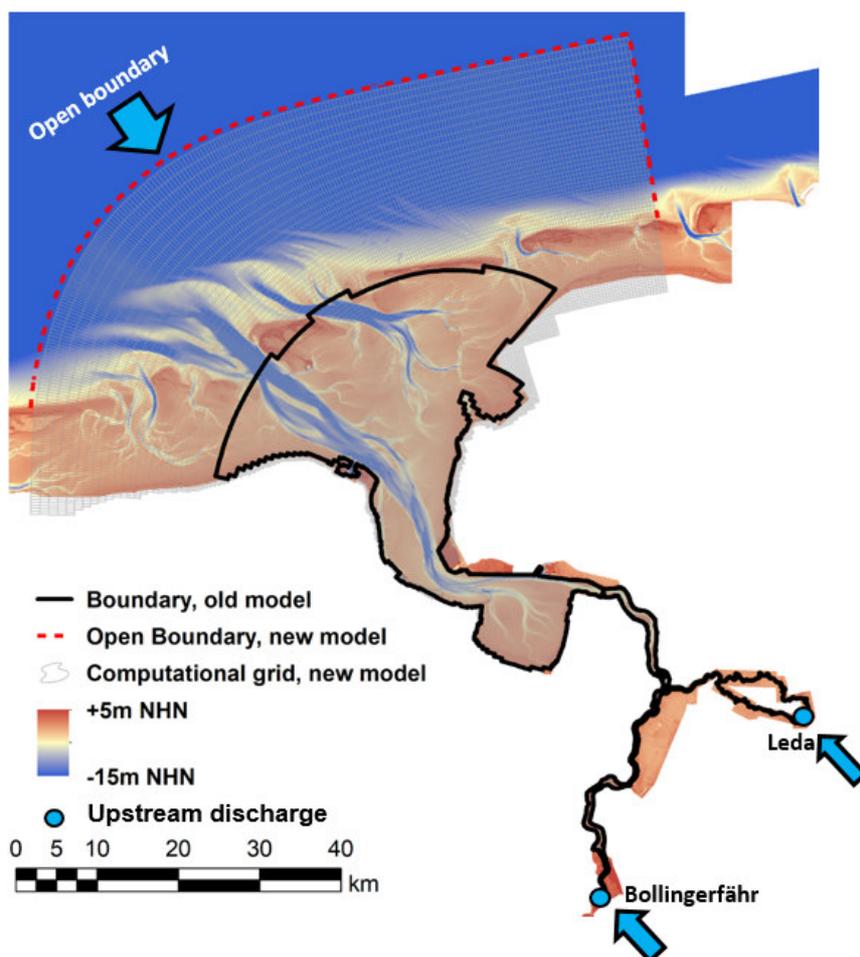
Figure 2.18. Bathymetry of the Ems estuary with the model open boundary and upstream discharge locations.

The hydrodynamic model setup is forced by boundary and initial conditions. Initial conditions are prescribed as the zero value for water level, current velocity and sediment concentration, but for temperature, and salinity, the averaged values of the measured data from gauges along the Ems estuary are interpolated. The boundary conditions for upstream are the Ems-river parameters for discharge (m^3/s), salinity (ppt), and temperature (°C) with temporal resolution of 15 min. These measurements have been performed by the Federal Waterways and Shipping Administration (WSA-Ems-Nordsee) for discharge (m^3/s) and salinity (ppt) at gauges Versen-Wehrdurchstich (52° 43' 58.74" N, 7° 14' 30.73") and Dreyschloot (53° 10' 40.79 " N, 7° 40' 8.63") and by the German Meteorological Service (Deutscher Wetterdienst, DWD) for temperature (°C) in Emden (53° 23' 17.16" N, 7° 13' 43.32" E).

The open boundary of the FSK-model is located in the North Sea and started from the German Wadden Sea of the Ems-estuary (Fig. 2.18). The hydrodynamic boundary conditions for the open boundary are provided by a nesting approach from the Continental Shelf Model of the FSK, which computes the hydrodynamic parameters for the southern North Sea including the German Bay.

The FSK-morphodynamic model setup applied in the Interreg project uses a bed composition of nine sediment fractions, where one fraction is cohesive fine sediment (mud) with grain size d < 0.063 mm, three fractions fine sand with 0.063 mm ≤ d < 0.2 mm, and five fractions coarse sand with 0.2 mm ≤ d < 0.8 mm. The FSK-hydro-morphodynamic model with the abovementioned model setup was run for the time frame of 01.01.2017 to 31.12.2017, and time step of 0.1 min (6 sec) to provide the steering hydro-morphodynamic communication data for the water quality model (DELWAQ).

The water quality and ecological model was set up with suitable boundary conditions with respect to the other models in the Interreg project i.e. SNS-GPM and the Deltares model, in order to provide harmonized model runs. The data for wind speed, wind direction and air pressure were provided by ICON-model of the German Meteorological Service and is read by the model in the corresponding time steps to include the effect of wind shear stress on the water surface and consequently its contribution to the water level change.

The water quality model is run using the open-source D-Water Quality (DELWAQ) numerical tool of Deltares. The numerical hydro-morphodynamic model results are provided to the water quality model by means of offline coupling. This particularly includes the fine sediment concentration, which plays an important role in light climate modelling for the growth of different phytoplankton functional groups.

Within the project's model harmonization runs, the boundary values and the modelled substances were taken from the overarching Deltares model. The FSK model in this regard has a unique feature by including the results of the sediment transport modelling for a direct light climate calculation. The approach allows a direct calculation of the visual light extinction induced by available sediment within the water column. Moreover, due to higher resolution and implementation of full Reynolds-Averaged-Navier-Stokes (RANS) transport equations, it can reproduce the mixing of water through estuarine baroclinic circulation, where the salty water of the North Sea mixes with the freshwater inflow of the Ems River.

The water quality model was further run for the current state (timeframe of whole 2017) and the corresponding boundary condition set consisting of reduced nitrogen concentrations according to the management objective of 2.8 mg TN/l.

The 3D Dutch Continental Shelf Model – Flexible Mesh (3D DCSM-FM) water quality and ecological model covers the entire Northwest European Continental Shelf, including the North Sea and adjacent shallow seas, such as the Wadden Sea (Fig. 2.19). The model uses D-Flow FM from the Delft3D FM software suite to simulate hydrodynamics and water quality and ecological processes (Zijl et al. 2021). It is used here in its most recent setup, used to derive threshold values for the newly defined assessment areas under the OSPAR convention (Lenhart et al. 2022). The model grid is coarser near the open boundaries and in deep waters and the resolution increases towards the shallower waters and in the Southern North Sea to 0.5x0.5 nautical miles. The water column is divided into a maximum of 50 vertical layers (20 layers of uniform thickness for the top 100 m or shallower areas, and a maximum of 30 deeper layers at fixed depths -- Z-sigma layer approach). 3D DCSM-FM computes water levels (tide and surge) as well as heat and salinity (Zijl et al. 2021).

Biogeochemical processes are simulated using the D-Water Quality module (Zijl et al. 2021). Processes were parameterised as in the GEM model (Blauw et al. 2009). The water quality model simulates the cycles of major nutrients (nitrogen, phosphorus and silica), organic carbon and dissolved oxygen. Phytoplankton dynamics are simulated using the BLOOM module (Los 2008). Four species groups are simulated: marine diatoms, flagellates, dinoflagellates and *Phaeocystis*. For each of these groups, three ecotypes are defined to account for adaptation to changing environmental conditions (i.e. nitrogen, phosphorus or energy limitation). Additionally, two groups of benthic filter-feeders (*Mytilus edulis* and *Ensis leei*) are represented using a Dynamic Energy Budget approach (Troost et al. 2010).

The model uses atmospheric forcing fields from the latest ERA5 ECMWF reanalysis product. Atmospheric deposition is included in the model as an extra source of DIN. Deposition rate is forced using the 2017 total (wet+dry) deposition fields of reduced and oxidized nitrogen from the Norwegian Meteorological Institute, calculated using the EMEP MSC-W chemical transport model (EMEP 2020). Offshore boundary conditions are defined based on the CMEMS global ocean biogeochemistry hindcast product as described in (Zijl et al. 2021). River inputs are defined at 319 locations using ICG-EMO data.

Suspended inorganic sediment concentrations are forced using a 2D time-varying field. This 2D field is derived from remote sensing-based winter concentrations from Nechad et al. (2010), transformed to a weekly suspended inorganic matter field using a cosine function. An additional correction was applied to the suspended inorganic matter field in shallower, less saline areas (e.g. the Wadden Sea area). This correction, calibrated on Wadden Sea measurements, is described in detail in Chapter 2.3.3. In the 3D DCSM-FM model it was applied to the entire domain, except for the Kattegat and Baltic Sea.

Yellow:	1/10° x 1/15°	~ 4.0 nm x 4.0 nm	
Green:	1/20° x 1/30°	~ 2.0 nm x 2.0 nm	→800m isobath
Blue:	1/40° x 1/60°	~ 1.0 nm x 1.0 nm	→200m isobath
Red:	0.75' x 0.5'	~ 0.5 nm x 0.5 nm	→ 50m isobath
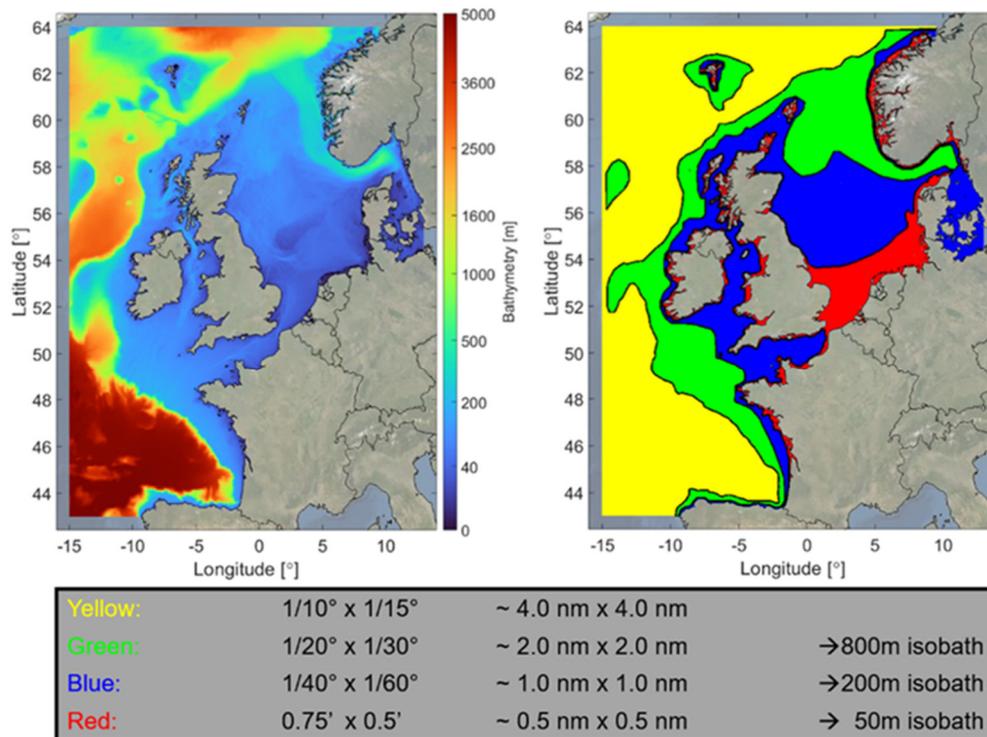
Figure 2.19. Bathymetry and grid cell size of the 3D DCSM-FM model.

2.3.2 Model harmonisation (development for the DCSM and SNS-GPM model)

The differences between the SNS-GPM model and DCSM-FM model were reduced step-by-step using the same approach and input data:

- Same river inputs and atmospheric deposition
- Same years: 2017 with 3 years spin-up
- Nested boundaries
- Comparison and convergence of phytoplankton model parameters
- Same model input on SPM concentrations

In this way the model results became more comparable. Both models used the same river inputs based on the river input database created by Sonja van Leeuwen (NIOZ) for the OSPAR working group for ecological modelling: ICG-EMO (see Chapter 4.1). The rates of atmospheric nutrient deposition were based on maps from the EMEP-model for 2017. Both models were run for the same years: 2014-2017; and model results evaluated and validated for 2017. The years 2014-2015 were used as spin-up years to let the concentration patterns in the model adapt to the updated model inputs. The SNS-model used concentrations of model variables at the model boundaries from the DCSM-FM model to make sure that differences in simulated concentrations do not arise from different concentrations in offshore waters along the edges of the SNS-GPM model.

Next, the SNS-GPM model adapted the definition of phytoplankton species groups in the model to match the phytoplankton species group definition in the DCSM-FM model. Both models now use four phytoplankton species groups: diatoms, flagellates, dinoflagellates and *Phaeocystis*. This choice was further supported by the data analysis of phytoplankton species composition in the Wadden Sea (Chapter 2.2.3c) showing that these are the 4 key phytoplankton species groups in this area. HIFMB also performed a literature review on the values of the phytoplankton model parameters for both models. Since the models use different model formulations, the set of model parameters differs and the parameter values could not be fully harmonized between the 2 models.

Figure 2.20 shows the interaction diagram for the SNS-GPM model, which illustrates the flow of matter through the ecosystem model. The different boxes indicate state variables or aggregation of state variables, which interact according to the arrows between the boxes. The circles within the boxes indicate the elements that are passed through, like C for carbon, N for nitrogen and P for phosphate.

The SNS-GPM model was extended by adding two more phytoplankton groups: *Phaeocystis* and mixotrophs. The mixotrophs group is functionally both autotroph and heterotroph and preys on small detritus, nanoflagellates and *Phaeocystis*. *Phaeocystis* can feed on the particulate organic phosphorus pool. The resulting structure is shown in Figure 2.20, wherein "B" stands for all pelagic phytoplankton groups (diatoms, nanoflagellates and *Phaeocystis*), and "M", framed in red, stands for mixotrophs.
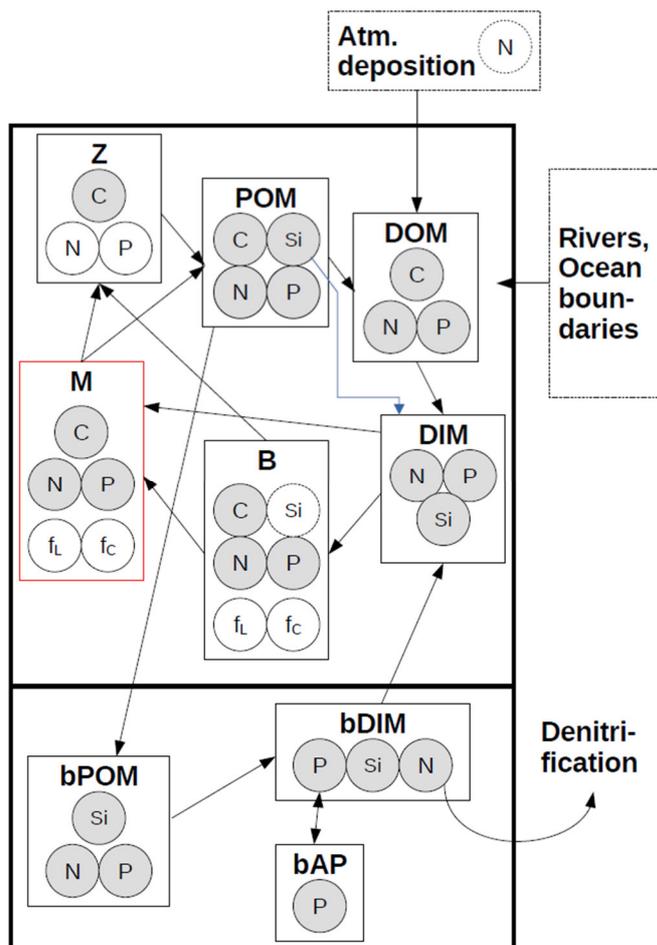
Figure 2.20. Flow chart diagram of the GPM model.

Following the above-mentioned literature review, the SNS-GPM model was extensively reparametrized, utilizing literature values wherever applicable or available. This process served to harmonize SNS-GPM with the DCSM model. Significant improvements were made in the seasonal dynamics of Chl. The relevant parameters can be found in Table 2.7. The concerning equations are found in Kerimoglu et al. (2017 & 2020).

Table 2.7. Parameters of the GPM model, after reparameterization. All other parameters may be found in Kerimoglu et al. 2017 & 2020.

Parameter	Diatoms	Nanofl.	Phaeoc.	Mixotr.
chlorophyll specific light ext. coefficient	0.0023	0.0023	0.0023	0.0023
sinking speed	4	-0.2	-0.2	0
linear mortality rate	0.1	0.1	0.1	0.05
quadratic mortality rate	0	0.001	0.0	0.04
max. P:C ratio	0.012	0.019	0.012	0.008
min. P:C ratio	0.004	0.002	0.008	0.007
max. N:C ratio	0.134	0.165	0.201	0.148
min. N:C ratio	0.052	0.029	0.146	0.117
molar C:Si ratio	8.4	-	-	-
max. C uptake rate	3.0	2.5	9.0	7.0
max. P uptake rate	0.05	0.0013	0.1	0.005
max. N uptake rate	0.9	0.6	1.0	0.7
half sat. for P uptake	0.04	0.04	0.5	0.007
half sat. for NO3 uptake	3.0	3.0	6.0	3.2
half sat. for NH4 uptake	1.0	1.0	8.0	8.0
half sat. for Si lim. Growth	3.0	-	-	-
init. slope of P-I-curve	6.0	3.0	6.0	3.0
max. CHL:C ratio	0.258	0.171	0.088	0.044

Finally, both models adopted the same approximation of the underwater light climate, based on satellite data of suspended particulate matter (Nechad et al. 2010, Fig. 2.21a), scaled in the Wadden Sea area to yield elevated turbidities, based on in-situ data of suspended particulate matter and their local bathymetry (Fig. 2.21b). The scaling factors applied depended on local salinity and water depth per grid cell (Fig. 2.21c). The difference between the corrected and uncorrected suspended particulate matter is shown in Figure 2.21d.
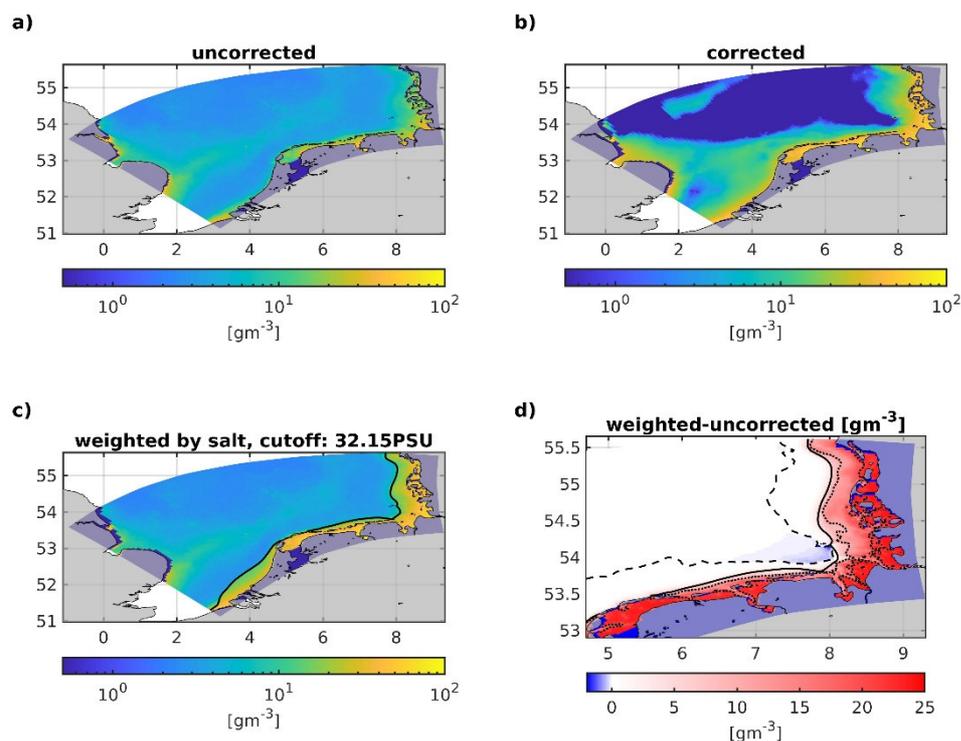
Figure 2.21. Suspended particulate matter, a) as given by the satellite algorithm (Nechad et al. 2010, here, interpolated onto the SNS-GPM model grid), b) the same as a) but with applied correction, c) the same as a) and b), but with a transition from the uncorrected (a) to the corrected (b), around the 32.15 PSU isohaline, and d) the difference between c) and a).

2.3.3 Further model improvements for the FSK model

The FSK-model used the same model formulations and parameter values as the DCSM-FM model, except for the grazers. The FSK model does not include benthic grazers, whereas the DCSM-FM model includes grazing by mussels. The FSK model uses a much finer model grid than the other two models and has a smaller model domain, focusing on the Ems estuary. The underwater light climate is based on silt simulations with the FSK model, rather than satellite data.

Another unique feature of the FSK-model is the phosphate remineralization (P-remineralization), which was developed and implemented within this project based on the analytical approach of Gypens et al. (2008). The developed module applies the sedimentation flux of the particulate organic carbon (POC), which is provided by means of the DELWAQ numerical computations. Comparing the deposited POC on the bed with the available POC in the topmost sediment layer determines the import of POC from the bed into the sediment layer or from the sediment layer into the bed layer. Then, the flux of POC is determined through the convection-diffusion transport equation for POC and due to the mass conservation consideration, the transported flux from the sediment layer into the seabed is reduced from the available POC in sediment layer and for the nutrients on the seabed. The transport equations were already published in the paper of Gypens et al. (2008) and the general form of the solutions are available. However, in order to limit the computational expenses, the transport equations were solved in a steady state condition. This is an important limitation of the Gypens et al. (2008) approach, where the temporal evolution of POC after import into the sediment layer within the calculational time step is neglected.

Moreover, the diffusion coefficient of POC in the transport equation is calculated by means of an empirical equation and the advection velocity inside the sediment layer is also an empirical function of the sedimentation flux of POC. Therefore, the development of the Gypens et al. (2008) approach in

D-Water Quality is considered as a first step in application of an analytical solution for P-remineralization in a numerical framework.

Furthermore, no measurements in the investigated catchment (Wadden Sea and Ems-estuary) were available to compare with the numerical results after the application of the new developments. Likewise, no measurement of nutrients inside the sediment layer were available to calibrate and validate the computed concentrations of POC and the resulted nutrients from the POC decomposition as well as transported nutrients from the seabed into the sediment layer.

Therefore, a new measurement campaign accompanied with the numerical investigations is encouraged to evaluate the accuracy of the newly developed module in D-Water Quality for P-remineralization. However, the analytical approach is significantly more efficient in comparison with the expensive numerical computations of nutrients inside the sediment layer, where around hundred sediment-sublayers (Gypens et al. 2008) have to be applied to achieve reasonable numerical results for nutrients and their exchange with the available nutrients on the seabed.

The application of the abovementioned developments for P-remineralization revealed that the contribution of additional nutrients from the sediment layer into the seabed lead to higher Chl concentration in tidal areas (e.g. BOCHTVWTM-station), slightly increasing in stations Nney_W_1 and Nney_W_2, and a conversely reduction of Chl in the deeper part of the estuary (HUIBGOT-station) (Fig. 2.22), which is confirmed by measured data and may be interpreted as a kind of indirect and qualitative validation of the approach. With the focus of this improvement on P-remineralization, one additional station of the NLWKN monitoring program (Nney_W_1) was taken into account. Nney_W_1 lies in close vicinity to station Nney_W_2 and represents the same water body (N1) but is sampled by helicopter.
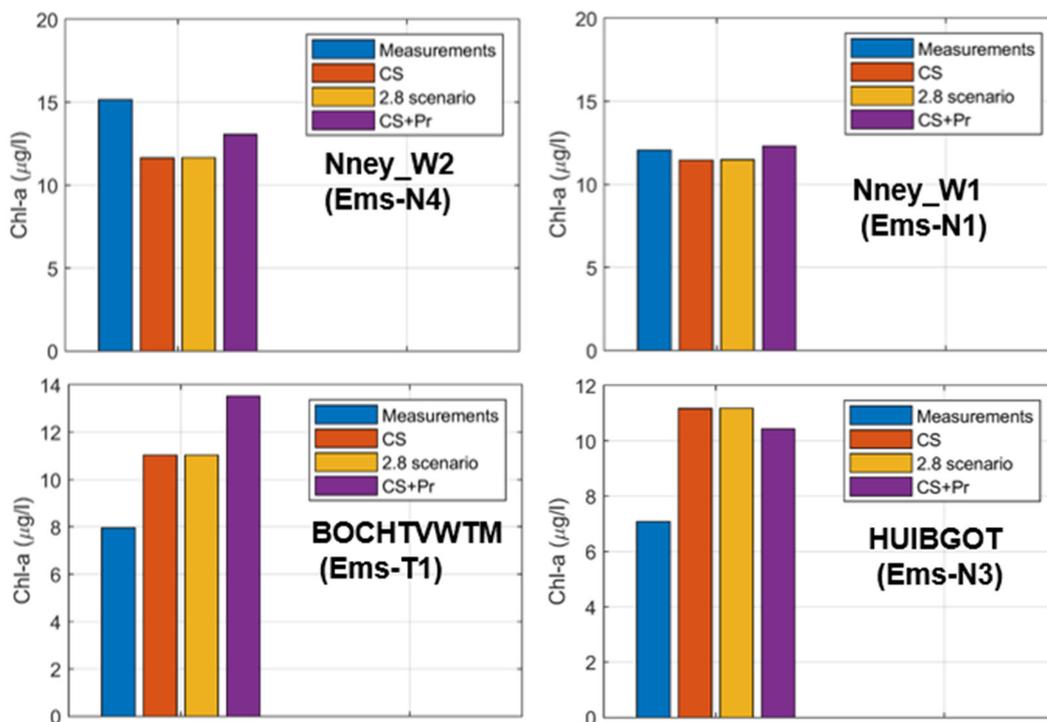
Figure 2.22: The P-remineralization contribution to the temporal and areal averaged chlorophyll concentration (CS+Pr) in comparison with measurements, current state (CS), and 2.8 (mg TN/l) reduction scenario.

In order to quantify the effect of P-remineralization on the overall budget, the model runs were performed for the current state as well as for the 2.8 mg TN/l scenario with and without inclusion of the process. This allows comparison of the model results with the previous model runs and evaluation of the role of P-remineralization with respect to its effect regarding Chl concentration.

A more detailed description of the implementation of P-remineralization in DELWAQ and the corresponding FSK-model results after model improvements is found in Annex 4.

2.3.4 Model validation

We have validated the model results of the Deltares model and the SNS model 1) by comparing spatial patterns in seasonal means of key variables with observed data and 2) by comparing time series and seasonal patterns of key variables with observed data.

a) Spatial patterns

Figure 2.23 shows the spatial patterns of annual mean salinity for 2017 for both models. Offshore the spatial patterns in salinity are fairly similar between both models. In the Wadden Sea salinities are higher in the SNS-model than in the DCSM-model, particularly along the coasts of the Netherlands and Niedersachsen. This indicates a stronger exchange of waters between the Wadden Sea and North Sea in the SNS-model, which may be explained by the coarser model grid. Since waters are always completely mixed within each model grid cell, a coarser grid enhances horizontal mixing of waters in the model.
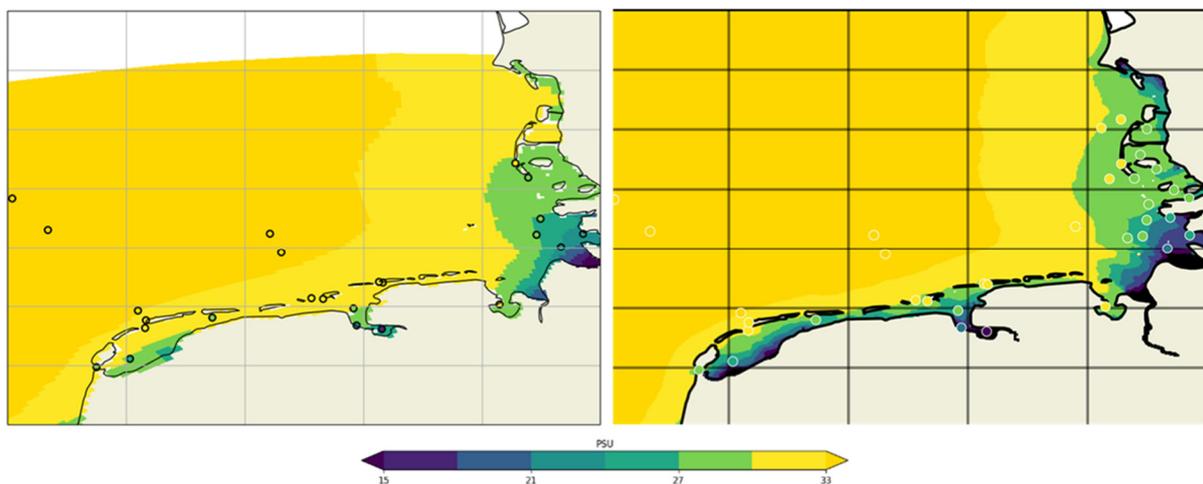
Figure 2.23. Spatial patterns in annual mean salinity in 2017 in the SNS-model (left) and DCSM-model (right). The colored circles indicate the annual mean salinities observed at monitoring locations.

Spatial patterns in dissolved inorganic nitrogen (DIN) are strongly controlled by the mixing of nutrient rich freshwater inputs and the North Sea (Fig. 2.24). Hence similar patterns are observed as for salinity. The SNS-model has a stronger exchange between the Wadden Sea and North Sea and the freshwater inputs from land are more strongly diluted with North Sea waters, leading to lower DIN concentrations in the Wadden Sea in the SNS-model compared to the DCSM-model. The winter DIN concentrations in the DCSM-model are more in line with the observed concentrations.
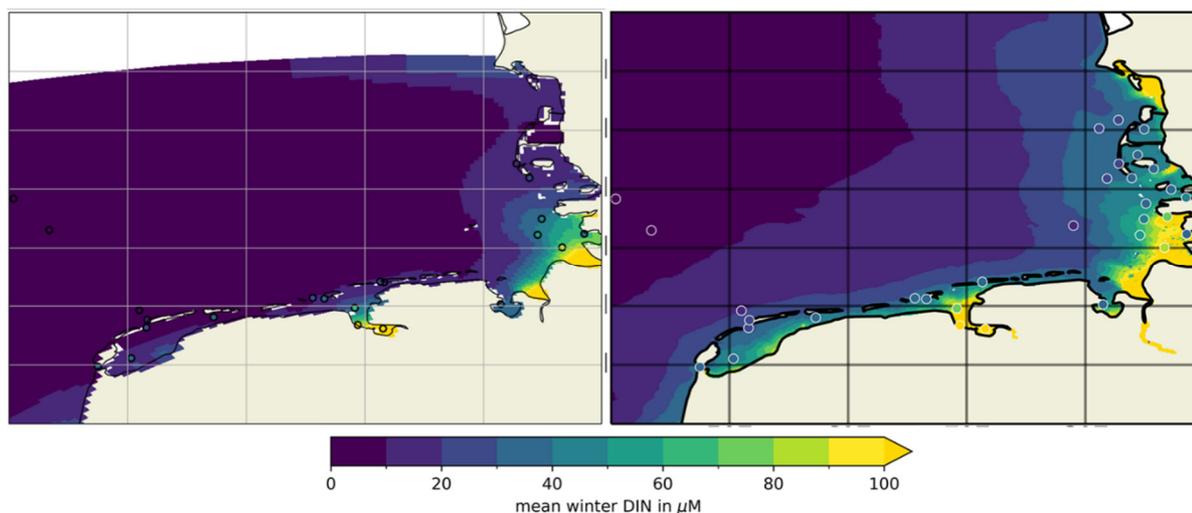
Figure 2.24. Spatial patterns in winter mean DIN in 2017 in the SNS-model (left) and DCSM-model (right). The colored circles indicate the winter mean concentrations observed at monitoring locations.

Spatial patterns in dissolved inorganic phosphorus (DIP) also show higher concentrations in the DCSM-FM model than in the SNS model (Fig. 2.25). In offshore waters in the German Bight the spatial patterns of winter mean DIP concentrations are more similar between the models than for DIN. The DCSM-FM model overestimates DIP concentrations in the western Dutch Wadden Sea. In the Ems estuary the SNS-model underestimates DIP concentrations and in the outer Ems estuary the DCSM-FM model overestimates DIP concentrations. In the eastern Wadden Sea both models are similarly close to the observations. Overall, the DIP concentrations in the Wadden Sea are higher in the DCSM-FM model than in the SNS-model, again indicating lower exchange of water with the North Sea in the 3D DCSM-FM model. This is in line with the results for salinity and DIN.
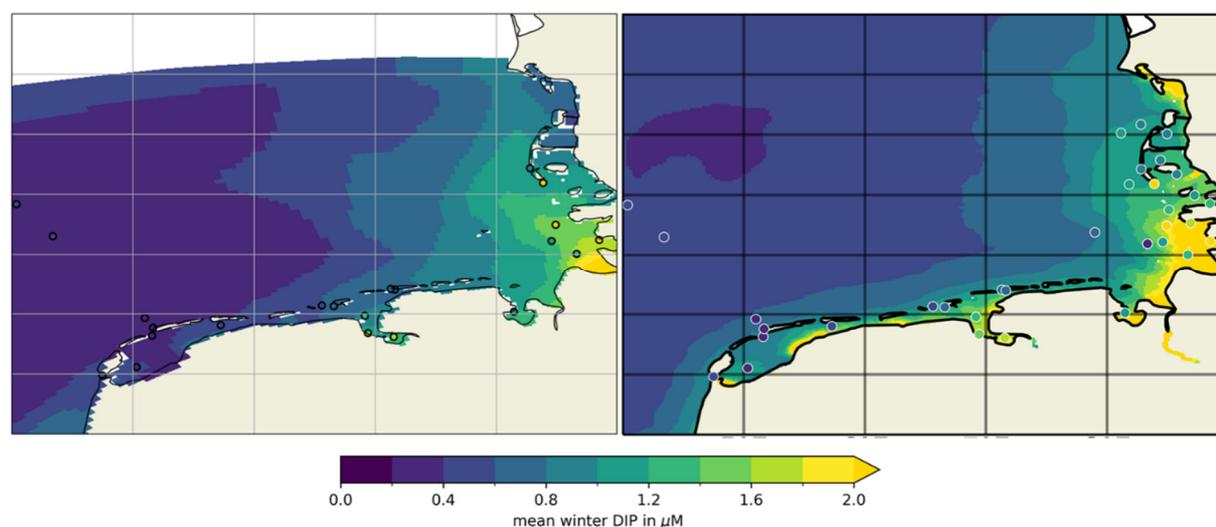
Figure 2.25. Spatial patterns in winter mean DIP in 2017 in the SNS-model (left) and DCSM-model (right). The colored circles indicate the winter mean concentrations observed at monitoring locations.

Despite the higher availability of nutrients in the DCSM-FM model compared to the SNS-model, it has lower Chl concentrations in coastal and offshore North Sea waters. Within the Wadden Sea however, nutrient and Chl concentrations in the DCSM-FM model are higher than those in the SNS-model. Furthermore, the DCSM-FM model shows much sharper spatial gradients than the SNS-model, with the deeper channels in the Wadden Sea having much lower Chl concentrations than the surrounding waters (Fig. 2.26). Satellite data of the Sentinel-2 based CMEMS product for coastal waters, also show

higher concentrations of Chl in shallow waters, compared to the deeper channels (Fig. 2.27). Just outside the Wadden Sea the DCSM-FM model shows strongly reduced Chl concentrations (dark blue area in Fig. 2.26). This is caused by light limitation in this area, after implementing the scaled silt concentration fields in the model input. The SNS-model uses the same underwater light climate but does not show reduced Chl concentrations in this area. These results indicate that Chl in the DCSM-FM model is much more affected by light limitation than in the SNS-model. Therefore, the joint approximation of the underwater light climate developed in this project seems inappropriate for the DCSM-FM model. Overall, the SNS-model slightly overestimates Chl concentrations in the Wadden Sea and offshore waters. For the DCSM-FM it is hard to assess whether the model results align with the observations. All monitoring locations are in deep water, where Chl concentrations are low, according to the model. There are no *in situ* observations available in shallow waters to assess whether the high concentrations simulated by the DCSM model in shallow waters are realistic.

Figure 2.26. Spatial patterns in growing season mean chlorophyll in 2017 in the SNS-model (left) and DCSM-model (right). The colored circles indicate the growing season mean concentrations observed in 2017 at monitoring locations.

Figure 2.27. Spatial patterns in growing season mean chlorophyll in 2020 - 2022 based on the CMEMS product for coastal waters.
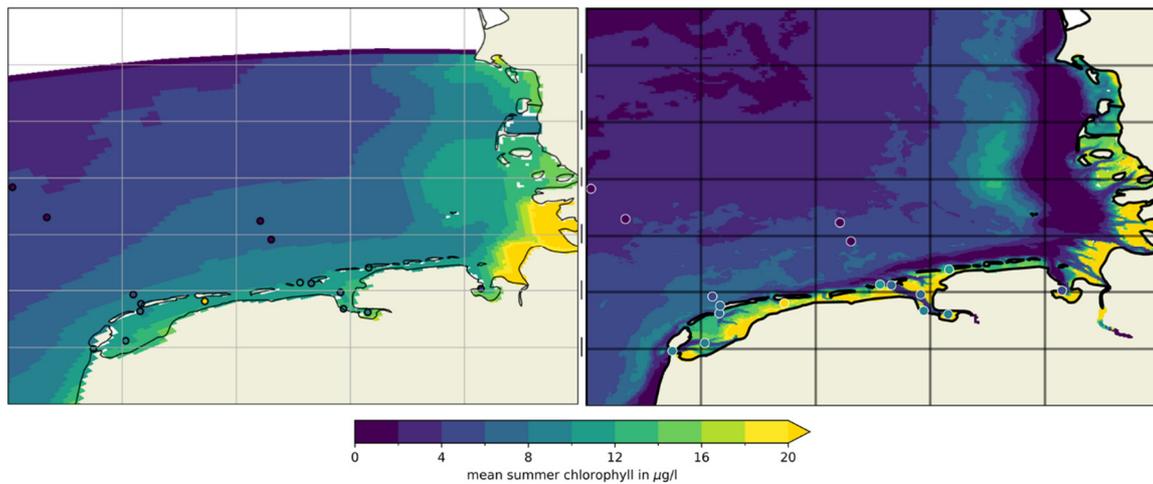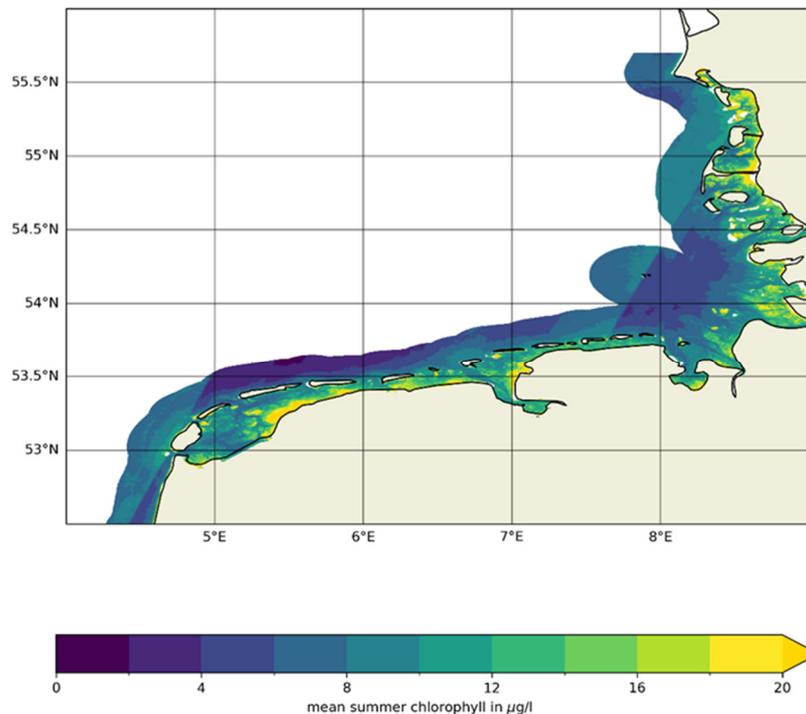
b) Seasonal patterns

Time series in the Dutch part of the Wadden Sea confirm that the DCSM-FM model underestimates salinity in the Wadden Sea and the SNS-model overestimates salinity (Fig. 2.28). This was also visible in the spatial patterns (Fig. 2.23). Interestingly, this pattern is not constant over the season in the western Dutch Wadden Sea (location DOOVBWT), which is strongly affected by freshwater inputs from Lake IJssel. In winter both models correspond well with observations. The DCSM shows lower salinities in summer than in winter, which is not visible in the observations and the SNS-model. Freshwater discharges from Lake Ijssel are generally lower in summer than in winter, leading to a higher salinity in the Wadden Sea. This is indeed observed in the river load database used in the modelling and in observations of salinity at DOOVBWT in 2015 and 2017, but less clearly in 2016.

DIN concentrations are much higher in freshwater than in the North Sea. Therefore, underestimation of salinity in the DCSM model leads to an overestimation of DIN concentrations which is indeed observed during the summer in the western Dutch Wadden Sea. Similarly, an overestimation of salinity in the SNS-model leads to underestimation of DIN concentrations in this area. In the eastern Dutch Wadden Sea (DANTZGT) the DCSM model simulates both salinity and DIN concentrations well. The SNS-model overestimates salinity and consequently underestimates DIN concentrations. Both models show a similar timing of decreasing DIN concentrations during the spring bloom, which roughly corresponds to the observations.

Seasonal variability of the other two nutrients (DIP and Si) is controlled by the same processes as for DIN: mixing between freshwater and sea water, uptake by phytoplankton and release from mineralization from organic matter. For silicate, the validation results indeed resemble those for DIN, with good correspondence in winter between both models and observations in both the western and eastern part. In summer the DCSM-model overestimates silicate concentrations in the western part but in the eastern part the DCSM-model corresponds to observation. However, in the SNS-model silicate concentrations are overestimated, despite overestimation of salinity. This indicates that the uptake of silicate by diatoms is underestimated in this model.

For DIP concentrations the seasonal concentration patterns are more complex than for the other two nutrients. After the decrease of DIP concentrations during the phytoplankton spring bloom, measured concentrations increase much faster during summer than those of DIN in both areas. This is correctly reproduced by the DCSM model. However, in late summer the concentrations decrease again in the observations, whereas in the DCSM model DIP concentrations keep increasing, leading to a strong overestimation of DIP concentrations. In the SNS model, DIP concentrations stay low during summer. In the eastern part of the Dutch Wadden Sea (DANTZGT), the DCSM model overestimates DIP concentrations in winter, although the salinity, DIN and silicate concentrations are simulated correctly. In summer the simulated DIP concentrations are more in line with observations. The SNS model underestimates DIP during most of the year, which is in line with results for salinity and DIN but the model simulates the winter concentrations correctly.

Chl concentrations in the SNS-model are much higher than those in the DCSM model in the western Dutch Wadden Sea. The DCSM model corresponds with observations during summer and most winters but underestimates the spring bloom peak. The SNS model reproduces the height of the spring bloom but overestimates concentrations particularly in winter. In the eastern Dutch Wadden Sea the DCSM model estimates summer concentrations correctly in most years. Only in 2016 observed concentrations were much higher than in other years, which was not reproduced by both models. In winter the DCSM model underestimates Chl concentrations and the SNS model slightly overestimates concentrations.
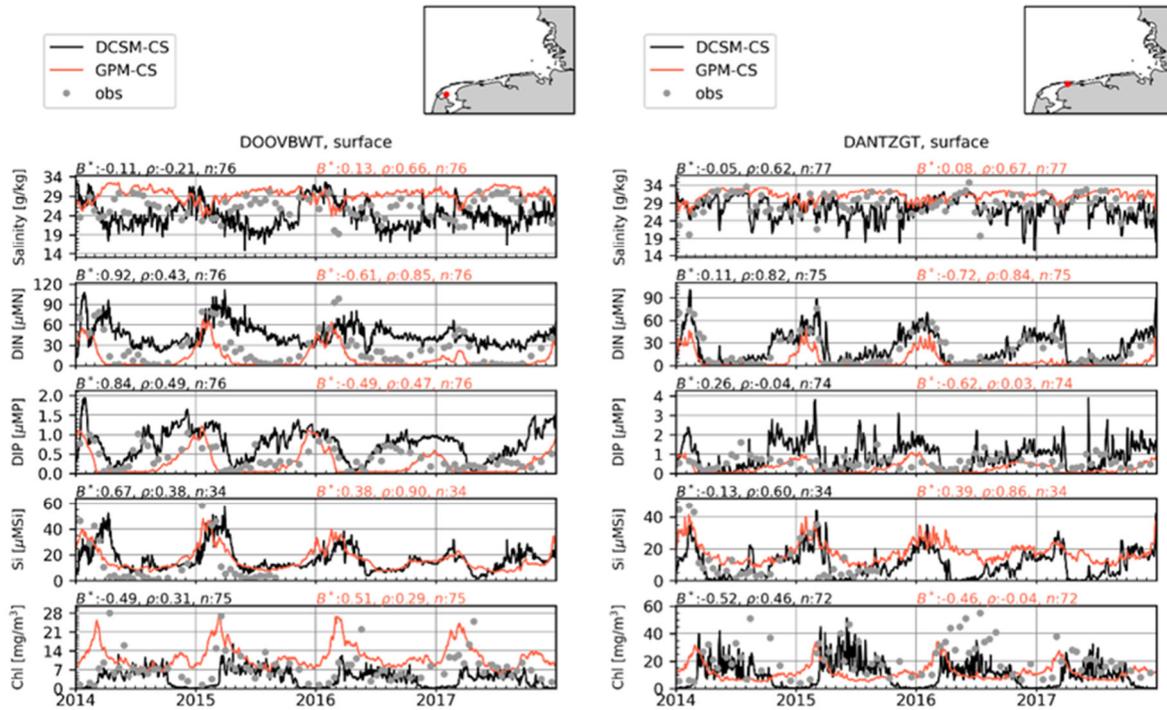
Figure 2.28. Seasonal patterns of salinity, nutrients and chlorophyll in 2014- 2017 in the western (left) and eastern (right) Dutch Wadden Sea in: in-situ observations (grey circles), SNS-model (red) and DCSM-model (black).

Figure 2.29 shows validation results for the Ems estuary: location Bocht van Watum (BOCHTVWTM) in the middle of the estuary and Huibertgat Oost (HUIBGOT) in the outer estuary. These show some similar patterns as described above for the Dutch Wadden Sea: the DCSM model underestimates salinity in summer and overestimates nutrient concentrations (DIN and Si) in summer. The SNS model overestimates salinity and underestimates nutrient concentrations (particularly DIN and DIP) during most of the year. The increase of DIP following the spring bloom is not captured by the SNS model and only partly by the DCSM model. In the estuary (BOCHTVWTM) both models overestimate Chl. Particularly, the DCSM model overestimates Chl concentrations in summer, due to too low suspended matter concentrations in the model. In the outer Ems estuary (HUIBGOT) a similar pattern is visible as for Chl at location DANTZGT in the eastern Dutch Wadden Sea: with the SNS model capturing the spring bloom height correctly but overestimating in the rest of the year and the DCSM model capturing summer concentrations correctly but underestimating in spring and winter.
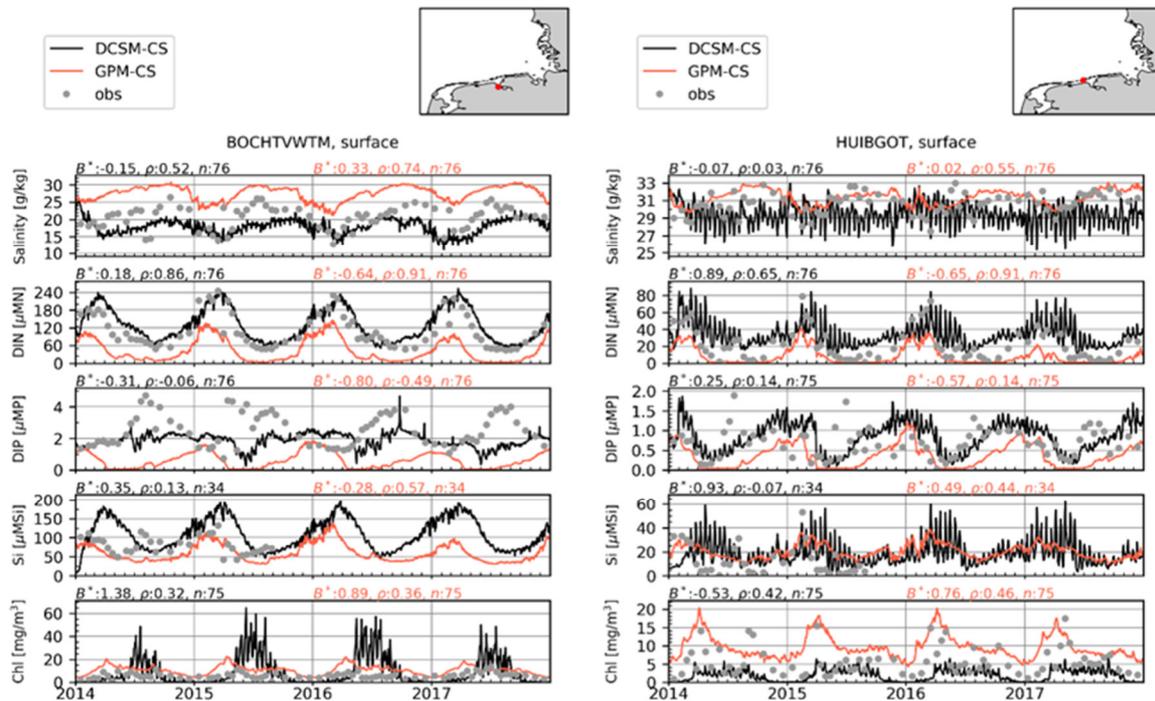
Figure 2.29. Seasonal patterns of salinity, nutrients and chlorophyll in 2014- 2017 in the inner (left) and outer (right) Ems estuary in: in-situ observations (grey circles), SNS-model (red) and DCSM-model (black).

In the western German Wadden Sea (Nney station) similar patterns are visible as in the eastern Dutch Wadden Sea (Fig. 2.30, left), with underestimated salinity in summer by the DCSM model and overestimated salinity by the SNS model. Also, the increase in DIP concentrations after the spring bloom is underestimated by both models. For silicate, winter values are correctly estimated by both models. In the SNS model there is insufficient update of silicate in summer, indicating an underestimation of diatom growth. However, validation results of diatom concentrations show that both models are fairly in line with observed diatom concentrations, or even overestimated (Fig. 2.31). In the DCSM model silicate concentrations are depleted in summer, but the uptake during spring comes too late, compared to the observations, indicating too high suspended matter concentrations in the model. Chl concentrations during summer are correctly estimated by both models but the spring bloom is missed by both models. In winter, the DCSM model underestimates Chl concentrations.

The Jadebusen station (JaBu) shows slightly different patterns than the previous locations (Fig. 2.30, right). Here, both models underestimate salinity in summer. Not only DIP but also DIN shows a fast increase following the spring bloom, which is not captured by the models. Furthermore, nutrient concentrations (DIN and DIP) are underestimated during spring, due to overestimated Chl concentrations in both models.
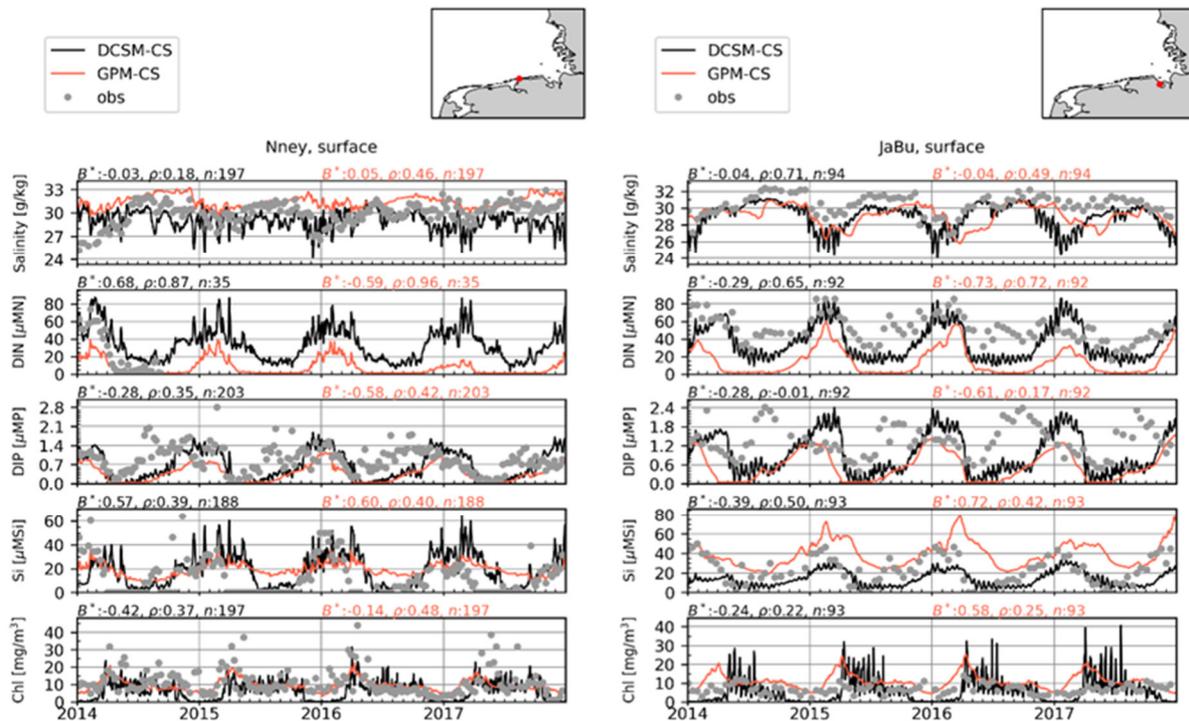
Figure 2.30. Seasonal patterns of salinity, nutrients and chlorophyll in 2014- 2017 in the western German Wadden Sea (left) and Jadebusen (right) in: in-situ observations (grey circles), SNS-model (red) and DCSM-model (black).

Summary and discussion of model validation results

Overall, the DCSM model generally underestimates salinity in the Wadden Sea and therefore overestimates nutrient availability, associated to freshwater inputs. In the SNS model, the opposite patterns are observed. Typical patterns for estuarine waters, such as fast increasing DIP (sometimes also DIN) concentrations after the spring bloom are hardly reproduced by the models. The DCSM model includes remineralization of organic matter, but this is insufficient to reproduce the observed high DIP concentrations in summer. Possibly, the accumulation of organic matter in the Wadden Sea sediments is underestimated or oxygen depletion in the sediment would need to be included in the models to capture the high release of DIP. A phytoplankton community dominated by diatoms, leading to high uptake of silicate during spring is also characteristic for estuarine waters. This is correctly simulated in the DCSM model but underestimated by the SNS model. However, validation results for diatoms do not show underestimation of diatom concentrations, which seems to contradict with the validation results for silicate (Fig. 2.31). The high uncertainty of the conversion from phytoplankton abundance (cells/L) to carbon content may play a role in this apparent contradiction. Chl concentrations during summer are generally correctly reproduced by both models, except in the inner Ems estuary. Winter concentrations are often underestimated by the DCSM model and sometimes overestimated by the SNS model.
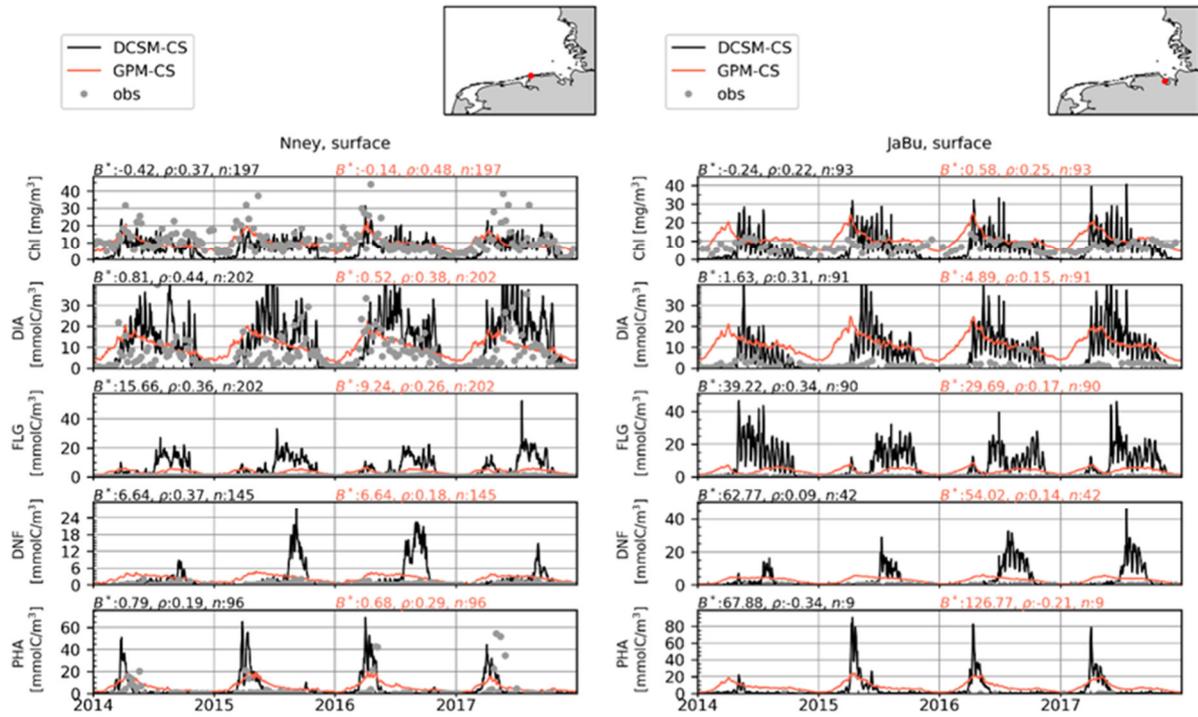
Figure 2.31. Seasonal patterns of chlorophyll and phytoplankton species groups (diatoms, flagellates, dinoflagellates and *Phaeocystis*) in 2014- 2017 in the western German Wadden Sea (left) and Jadebusen (right) in: in-situ observations (grey circles), SNS-model (red) and DCSM-model (black).

3. Effectiveness of current river management objectives

3.1 Description of the "2.8-scenario"

A model scenario was applied to study the effect of the current management objective for total nitrogen (TN) concentrations in the Dutch and German rivers discharging into the Wadden Sea and North Sea. Management objectives for annual average TN concentrations in the rivers at the limnic/marine border were derived by the Netherlands and Germany and implemented into national law/regulations; the German "Bewirtschaftungszielwert" was set to the annual mean 2.8 mg TN/l (BLMP 2011, Fischer et al. 2014), while the Dutch "streefwaarde" was set to the summer mean <2.5 mg TN/l for the large rivers (van der Molen et al. 2018, IKSR/CIPR/ICBR 2022). These objectives have been defined as the maximum concentration to ensure achieving Good ecological status for phytoplankton in coastal waters in the context of the WFD. Since the German annual average value of 2.8 mg TN/l is roughly equivalent to a summer average (April-September) of 2.5 mg TN/l applied in the Netherlands, both countries decided to use the value of 2.8 mg TN/l as a compliant value for the modelling activities in this project.

The following scenario will therefore be called the „2.8-scenario". Figure 3.1 provides an overview of the TN load reductions (compared to current loads) for the major Dutch and German rivers. Note that under this scenario, TP loads have not been changed.
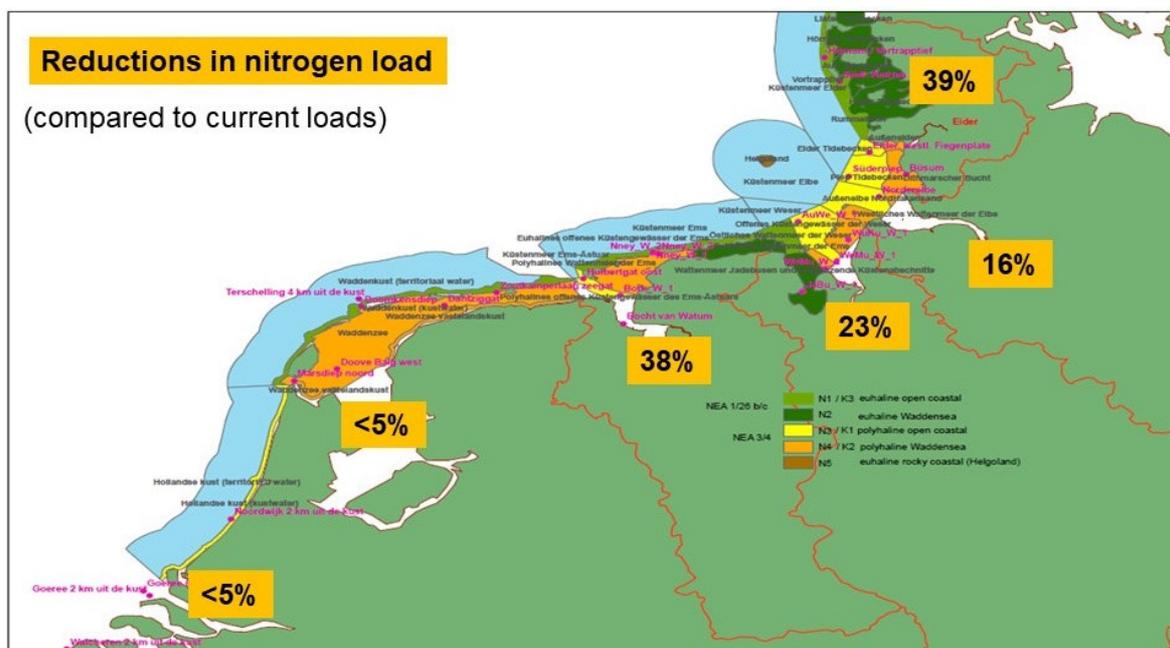
Figure 3.1. Overview of the TN load reduction compared to current loads, for the 2.8-scenario for major rivers. For more details, see Annex 1.

Within this Interreg project both countries have agreed to use this threshold value (for more details see Annex 1) for the scenario simulation. The aim of the model simulation is to provide a projection of the application of this threshold value for the nutrient load and the resulting concentrations of the eutrophication parameters, mainly Chl, in the transitional and coastal water bodies of the WFD in the Wadden Sea.

Annex 1 provides a detailed description of the definition of nutrient loads in this scenario.

3.2 Results of the model simulation for the "2.8-scenario"

In order to evaluate the effects of the 2.8-scenario, we compare the results from the model runs in this scenario to model runs for the current state, i.e. the years 2014-2017 with the actual riverine nutrient loads.

We first make a comparison of the spatial patterns in concentrations in this scenario to the current state. Then we look at the differences in time series between current state and the 2.8-scenario. Finally, we summarise the results for each WFD water body.

3.2.1 Spatial patterns

The results from the model simulations are shown as percent difference plots between the previously described Current State run (CS) (Chapter 2) and the new 2.8-scenario. Figure 3.2 shows the distribution for winter DIN and Figure 3.3 the related differences for summer Chl.
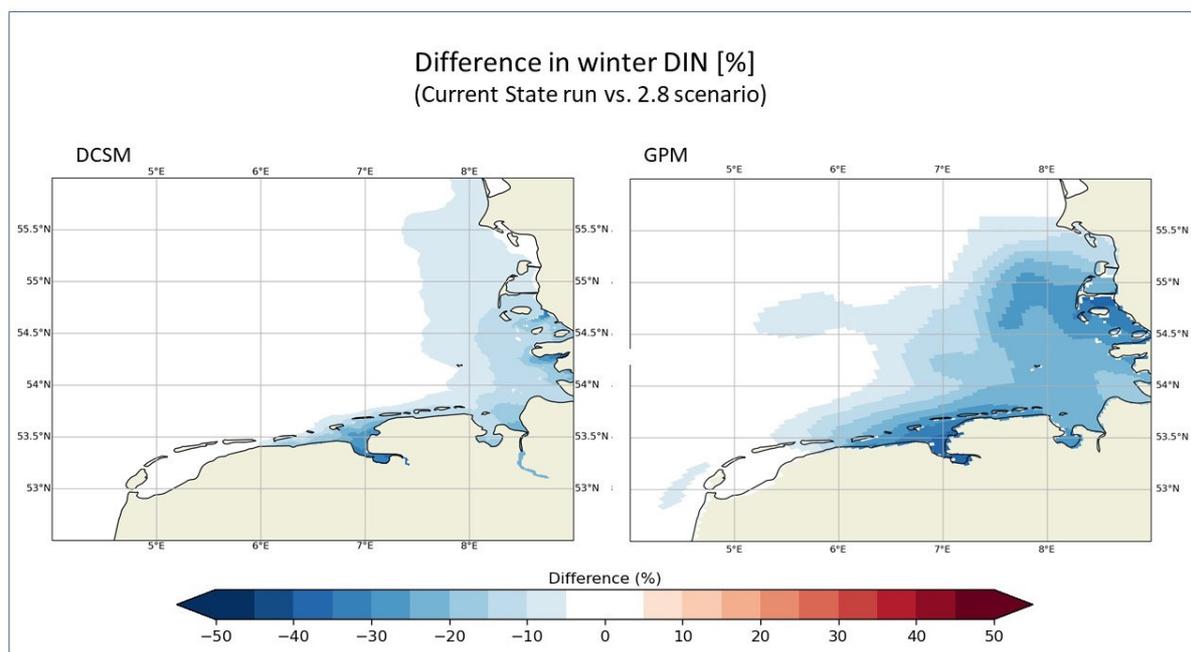
Figure 3.2. Horizontal distribution map of the percent difference between the CS simulation vs. the 2.8-scenario for DIN for the Southern North Sea. Blue colours indicate lower concentrations in the 2.8-scenario. Left: DCSM-FM model; right: SNS-GPM model.
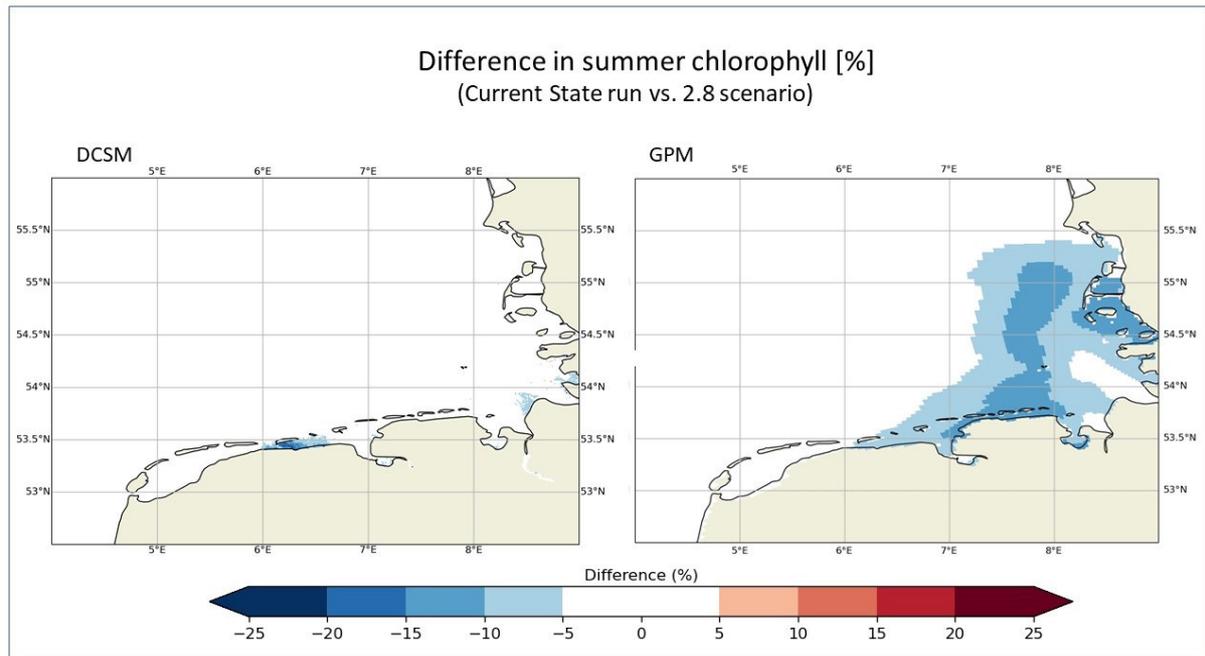
Figure 3.3. Horizontal distribution map of the percent difference between the CS simulation vs. the 2.8-scenario for chlorophyll for the Southern North Sea. Blue colours indicate lower concentrations in the 2.8-scenario. Left: DCSM-FM model; right: SNS-GPM model.

In Figure 3.2 the differences for DIN between current state and 2.8-scenario in the DCSM simulation are smaller and occur closer to the coast than in the GPM results. However, both models show the highest change of about 30 % within the Ems estuary and along the coast of Schleswig-Holstein. Both simulations agree in showing no difference for the major part of the Dutch Wadden Sea. Smaller differences are also shown for the inner part of the Elbe estuary. The spatial patterns reflect the different reduction levels in TN loads in the rivers (Fig. 3.1).

Figure 3.3 shows diverging results for Chl. While the DCSM model only reflects local reduction west of the Ems estuary and in the vicinity of the Ems estuary, the GPM model basically shows a response in Chl in the entire German Wadden Sea area, although with different intensity. It is interesting to note that the major reduction in Chl is not only restricted to the Wadden Sea area behind the Frisian islands, but spreads into the German Bight. In contrast, the direct river mouth of the Elbe shows no difference between current state and the 2.8-scenario, which is also the case for the DCSM model simulation.

3.2.2 Time series
For a detailed look into the effect of the 2.8-scenario for relevant eutrophication parameters a number of *in situ* stations are selected. They are different from the ones presented in chapter 2.3.3 for the model validation, since the focus here is on locations where differences between the two simulations are observed (Figs. 3.4, 3.5).

Figure 3.4 shows the timeseries for the station Borkum, with for both models the CS run (black) and the 2.8-scenario (red). As one would expect, there is no difference for the salinity time series. In contrast, the DIN timeseries represents a distinct difference between the two simulations, with lower values for the 2.8-scenario. This shows the impact of the 38 % lower nitrogen load by the Ems in the 2.8-scenario (Fig. 3.1) in both models. In the DCSM model the resulting Chl time series shows occasionally lower peak values in the 2.8-scenario while the overall time series does not change. These reduced Chl peak values go along with short term higher DIP concentration and longer periods of higher Si concentrations during the summer period. The GPM model shows the lower DIN values in the

2.8-scenario, resulting in a longer period of lower Chl concentrations during the summer period. The related difference in DIP can be observed in late summer with higher values in the 2.8-scenario, while in the Si time series, no differences can be observed.
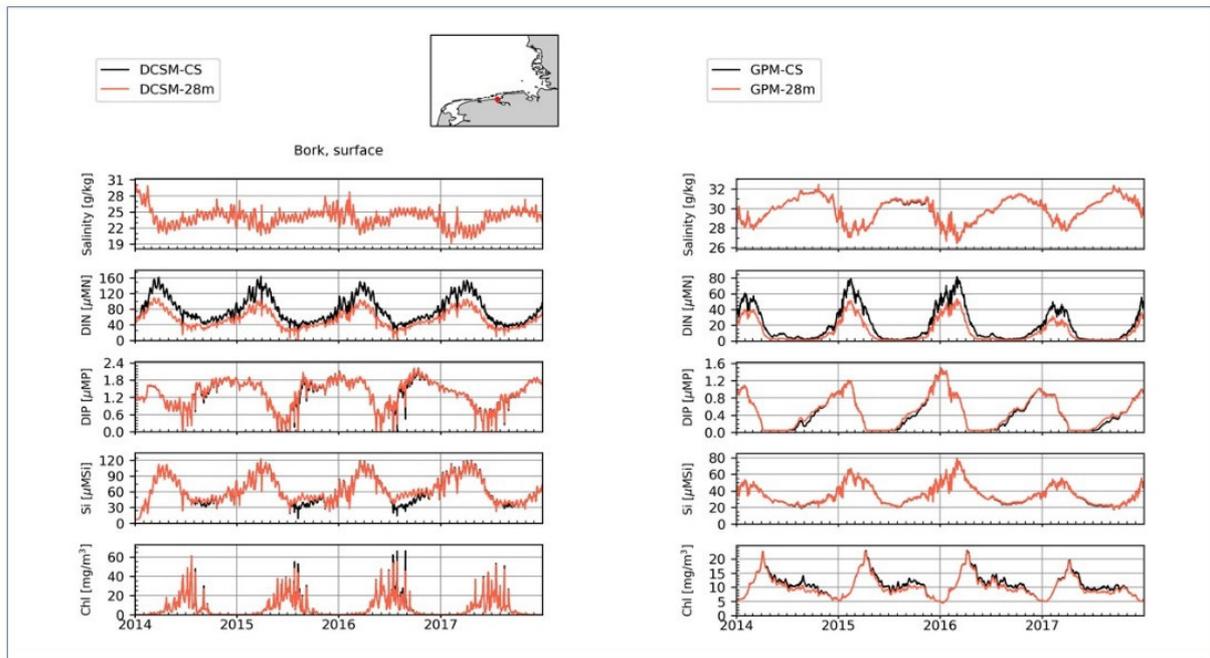
Figure 3.4. Timeseries at the station Borkum showing the difference between the CS simulation (black line) vs. the 2.8-scenario (red line) for the DCSM (left) and the GPM model (right) for the parameters salinity, DIN, DIP, Si and chlorophyll.
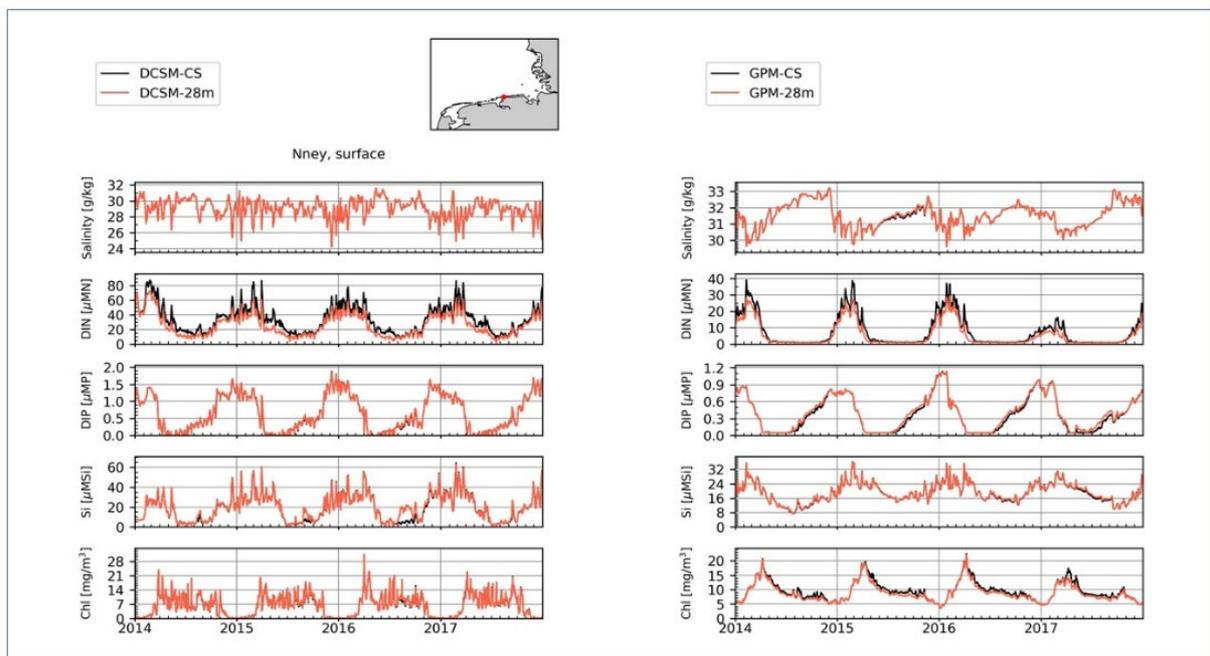
Figure 3.5. Timeseries at the station Norderney showing the difference between the CS simulation (black line) vs. the 2.8-scenario (red line) for the DCSM (left) and the GPM model (right) for the parameters salinity, DIN, DIP, Si and chlorophyll.

Figure 3.5 shows the timeseries at Norderney. The DIN time series in CS and 2.8-scenario are closer to each other than the ones for Borkum. For the DCSM model this results in only a very small deviation between the Chl values in the year 2016, accompanied by an observable difference in the silicate

values during the summer period. For DIP no difference can be observed in the DCSM model simulation. In the GPM model results for Chl there are differences during the summer periods with lower values in the 2.8-scenario. In contrast to the DCSM results, there are no changes in Si, but small changes for DIP during the remineralisation phase in late summer and autumn.
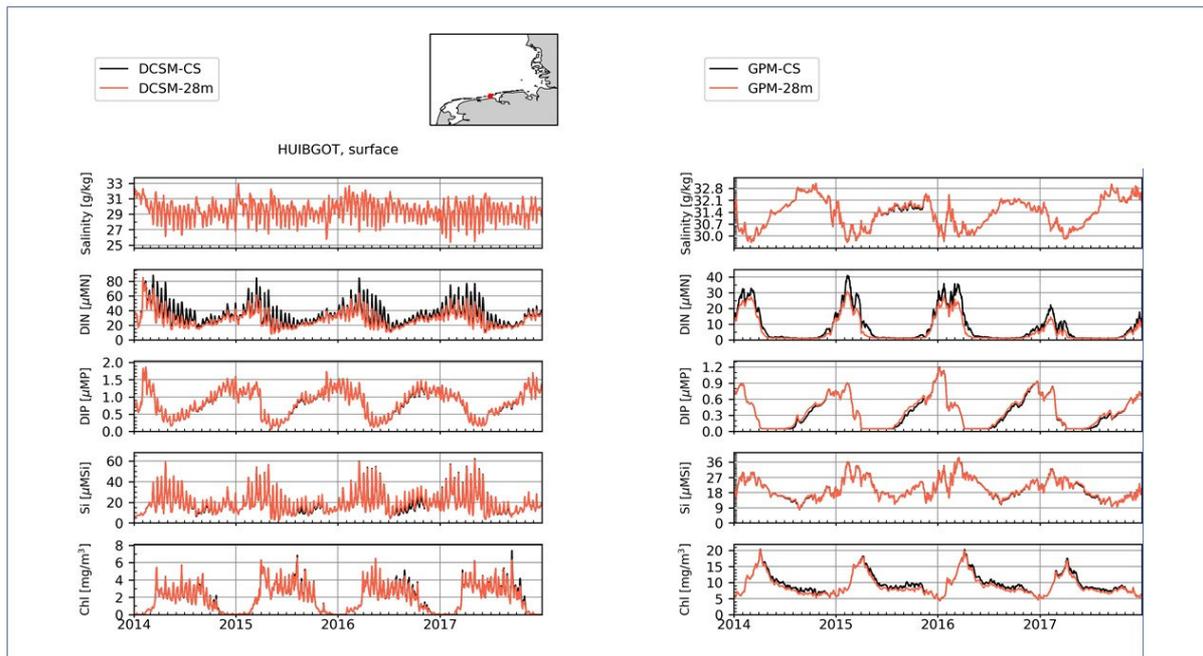
Figure 3.6. Timeseries at the station Huibgot showing the difference between the CS simulation (black) vs. the 2.8-scenario (red) for the DCSM (left) and the GPM model (right) for the parameter salinity, DIN, DIP, Si and chlorophyll.

Figure 3.6 shows the time series for the Dutch station Huibgot. The differences in the DIN timeseries are similar to the Norderney station (Fig. 3.5) and the response of the other parameters is very similar for both models, with lower peaks in the Chl values in the DCSM model and longer summer periods of lower Chl concentration in the GPM model.

3.2.3 Results aggregated per water body typology

Here, we present the results from the model simulations as bar plots for the different models. In comparison, the existing WFD threshold for Chl and the observed Chl concentrations (from measurements) are shown. The graphs (Figs. 3.7 – 3.10) all have the same structure, showing from left to right the existing WFD Good/Moderate boundary, the observed Chl concentration, the Chl concentration for the current state simulation (CS) as described in Chapter 2.3.3. and Chl concentrations in the 2.8-scenario. The bar plot for observed Chl shows the growing season mean for the year 2017 with the error bars showing the minimum and maximum values for the period 2014-2018. The results of the North Sea models are indicated as "GPM" (SNS-GPM, Universities Hamburg/Oldenburg) and "DCSM" (3D DCSM-FM, Deltares).

The various WFD water bodies are aggregated within their regional context, e.g. Dutch Wadden Sea (Fig. 3.7), Ems Wadden Sea (Fig. 3.8), Weser Wadden Sea (Fig. 3.9) and Elbe Wadden Sea (Fig. 3.10). For the Ems Wadden Sea (Fig. 3.8) also results from the FSK-model are integrated.

As already discussed in Chapter 2.3.3, the SNS-GPM model generally predicts higher concentrations of Chl than the 3D DCSM-FM model. In addition, Figure 3.7 shows that in both models, the 2.8-scenario

only results in slightly reduced Chl concentrations, compared to the current state. The models predict a very limited effect in the Dutch Wadden Sea, of reducing nutrient loads to the management objective. This is in line, however, with the limited reduction in TN loads in the river discharges that have the strongest impact on the Wadden Sea (e.g. Rhine, Lake Ijssel; see Fig. 3.1). For the Ems estuary, the 2.8-scenario represents a larger reduction in TN loads, but the effect on Chl is very small (also see Fig. 3.4). This can be explained by the overriding effect of light limitation in the Ems estuary.

The WFD assessment for phytoplankton of the Dutch Wadden Sea (Fig. 3.7A, C) is Moderate status (IenW 2022). The assessment results show that the coastal waters of the North Sea (Wadden coast, Fig. 3.7D) are in Good status (IenW 2022). As explained already in the context of Tab. 1.1 that for the Ems-N3 area, there are two Chl thresholds in use, which differ between the Dutch and the German part. For the Dutch water body, characterised as "Ems-Dollard coast", the mean value of 5.1 µg/l is applied, while for the German water body, indicated as "Ems-N3", a value of 5.5 µg/l is used.
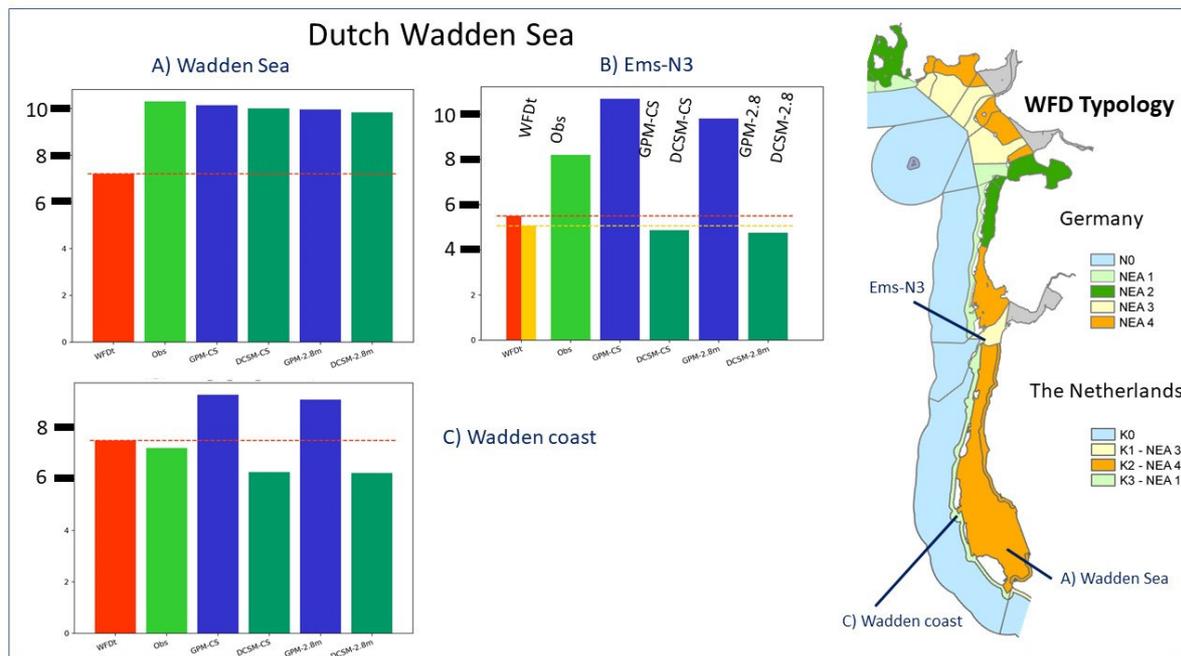
Figure 3.7. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the Dutch Wadden Sea. The results are presented for the Dutch WFD water bodies A) Wadden Sea, B) Ems N3, C) Wadden coast. For Ems N3 (B), two separate chlorophyll thresholds are currently used (red bar: 5.5 µg/l (DE); yellow bar: 5.1 µg/l (NL)). Shown chlorophyll observations represent NL measurements here.

The Ems region is presented in Figure 3.8. For the four water body types Ems N1 (Fig. 3.8A), Ems N2 (Fig. 3.8C), Ems N3 (Fig. 3.8B) and Ems N4 (Fig. 3.8D). The observations clearly show much higher concentrations than the current WFD threshold. Again, the simulation results for both the CS run and the 2.8-scenario show differences between the models. The results for the Ems N3 region (Fig. 3.8B) are identical to Figure 3.7B, except that Chl observations differ depending on country. The bar plot results for the Ems N3 region (Fig. 3.8B) and the Ems N4 region (Fig. 3.8D) include the model simulation from the fine scale FSK-model. Unfortunately, this model simulation is lacking a spin-up to reach equilibrium condition due to the long computation time. This is reflected in the extremely high Chl concentration in the CS run in comparison to the DCSM and GPM model results. All model results indicate that the effect of the 2.8-scenario for all water bod types is very limited.

The current WFD assessment for phytoplankton of the German water bodies Ems N1 and Ems N4 (Fig. 3.8A, D) is Bad status, while the water bodies Ems N2 and Ems N3 (Fig. 3.8B, C) are in Poor status (MU 2021).
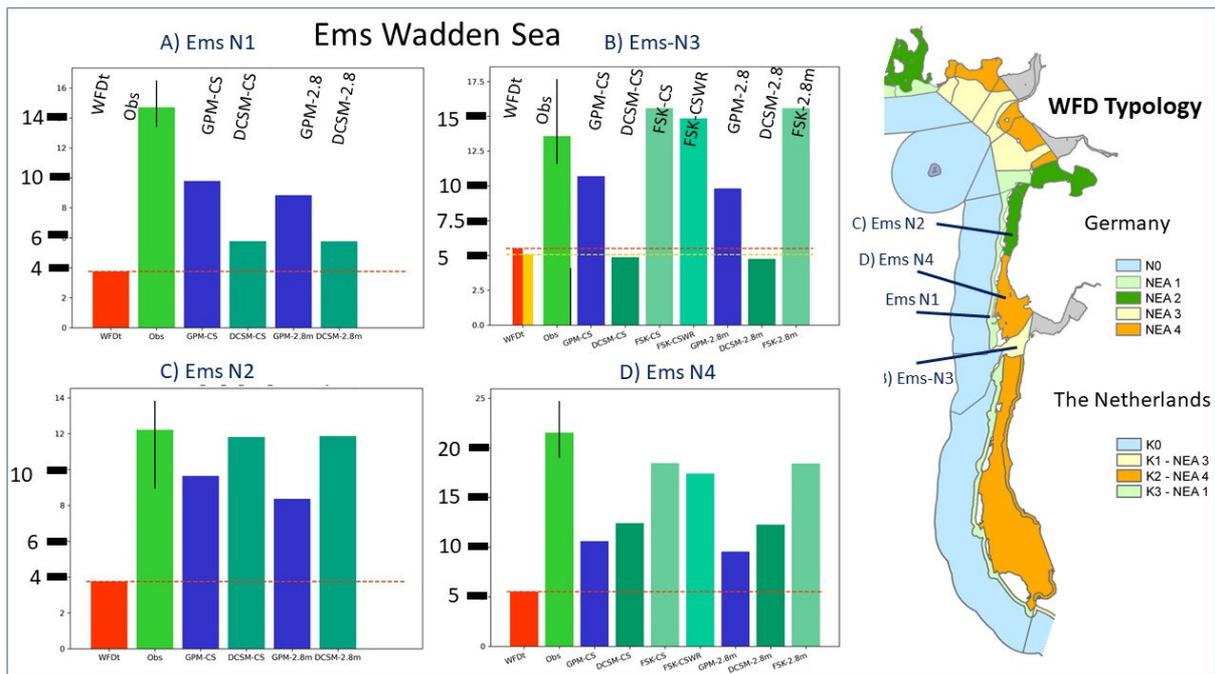
Figure 3.8. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Ems WFD water bodies A) Ems N1, B) Ems N3, C) Ems N2 and D) Ems N4. For B) Ems N3 and D) Ems N4 additional results from the FSK model are included. FSK-CS: current state, FSK-CSWR: current state with phosphate release. For Ems N3 (B), two separate chlorophyll thresholds are currently used (red bar: 5.5 µg/l (DE); yellow bar: 5.1 µg/l (NL)). Shown chlorophyll observations represent DE measurements here.

In Figure 3.9 the simulation results for the Weser water bodies are presented. Like in the Ems region, observed concentrations clearly exceed the WFD threshold. Again, there are differences between the models in the level of Chl concentrations. Both models predict very limited reductions in Chl concentrations in the 2.8-scenario. The SNS-GPM model shows slightly stronger effects of the reduction in TN loads in the 2.8-scenario than the 3D DCSM-FM model.

The current WFD assessment for phytoplankton of the German water bodies Weser N1-N4 (Fig. 3.9A-D) is Moderate status (MU 2021).
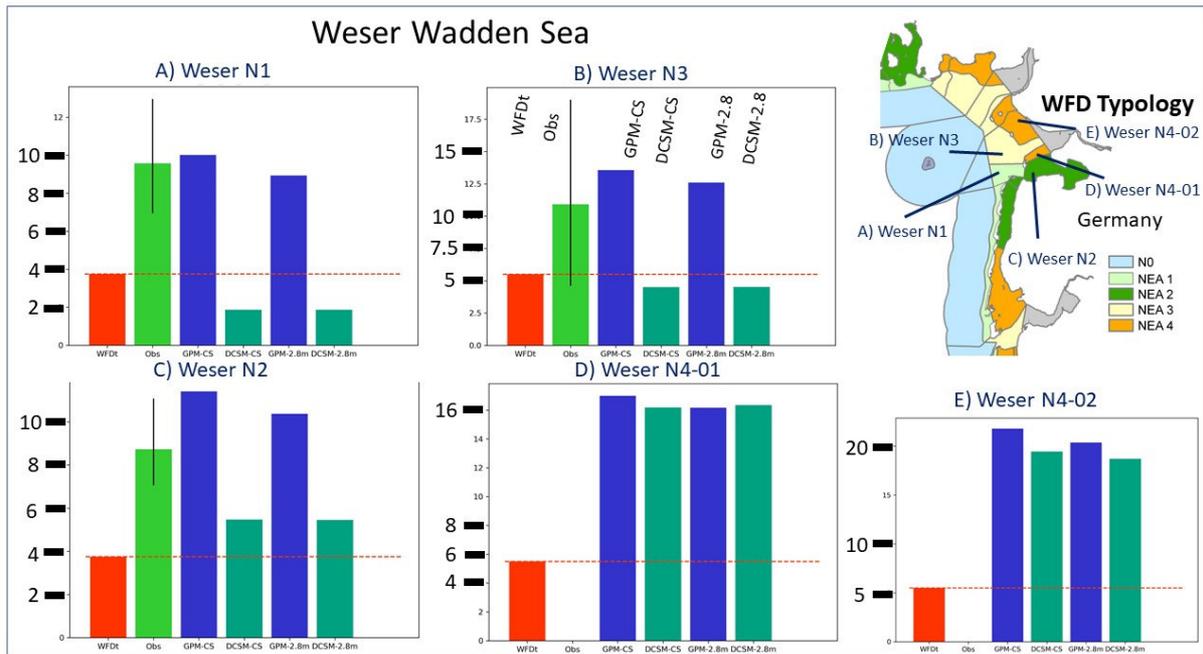
Figure 3.9. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Weser WFD water bodies A) Weser N1, B) Weser N3, C) Weser N2 and D) Weser N4.

Figure 3.10 shows the results of the simulation for the two Elbe WFD water body types Elbe N3 (Fig. 3.10A) and Elbe N4 (Fig. 3.10B). One can see the same pattern in the GPM and the DCSM model as shown for the Ems and Weser. In line with the reduction in TN loads in the Elbe River in the 2.8-scenario, which is more limited than in Ems and Weser (Fig. 3.1), the 2.8-scenario shows a very limited reduction in Chl concentrations in the 2.8-scenario in both models.
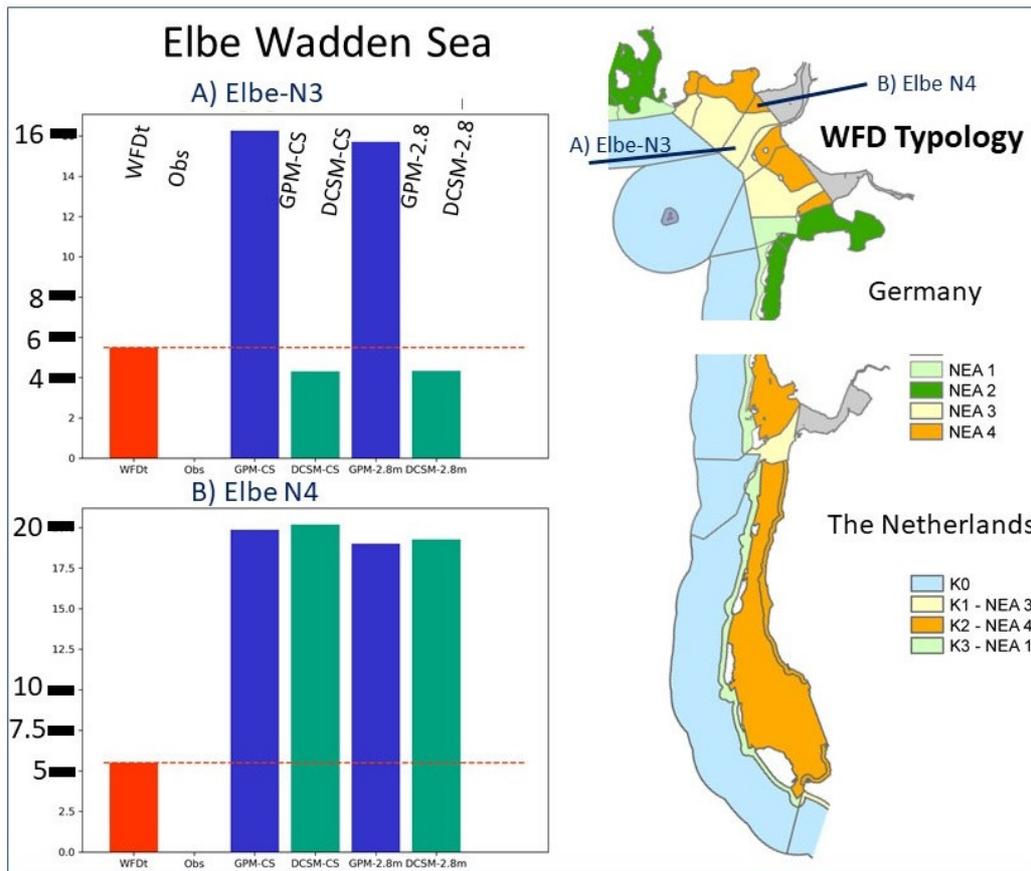
Figure 3.10. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Elbe WFD water bodies A) Elbe N3, B) Elbe N4.

As already discussed in Chapter 2.3.3, the 3D DCSM-FM model generally has lower salinity, higher nutrient concentrations and lower Chl concentrations in the Wadden Sea than the SNS-GPM model. This can, in most cases, be explained by the more coastal related characteristics of the SNS-GPM model, as can be seen in the salinity time series for a number of stations (Figs. 2.28, 2.29, 2.30). The pattern in the Ems N3 area is different, since the area is closer to the coast. One reason for the differences in the Chl concentration between the GPM and the DCSM model is the larger grid resolution of the SNS-GPM model for this area. Figure 3.11 shows the horizontal distribution of the seasonal mean Chl concentration for the Ems-Dollard and the adjacent area in the Wadden Sea. For the Ems estuary and the outflow region one can clearly see the differences between the two models. The DCSM model reflects the outflow channel of the Ems with lower Chl concentration. For the GPM model the lower spatial resolution (larger grid size) cannot capture this feature and represents this area with a much smoother Chl concentration.
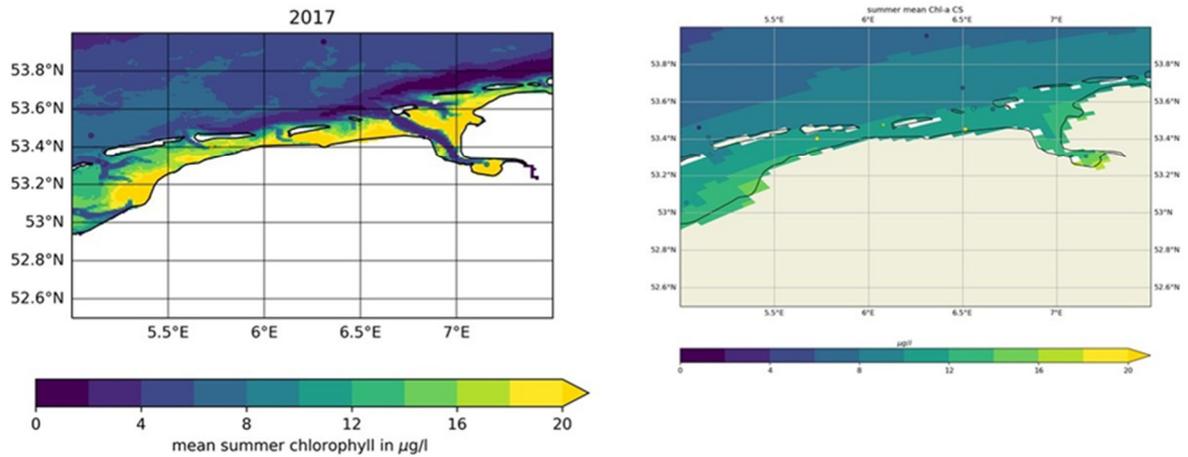
Figure 3.11. Comparison between the DCSM and the GPM model of the horizontal mean summer chlorophyll concentration from the CS simulation for 2017.

Finally, aggregated WFD water bodies are presented for the Weser and Ems region (Fig. 3.12) in comparison with the existing threshold as presented in Figure 3.12. The aggregated values were derived by averaging over all common areas weighted by surface area. It is interesting to note that the differences between the DCSM and the GPM model disappear when aggregated over these larger areas. This is especially true for the aggregated areas N1 and N2, both for the Weser and the Ems. While for these individual WFD typology areas the DCSM values were below the WFD threshold, they are above the threshold in the aggregated regional representation which in line with the GPM model results.
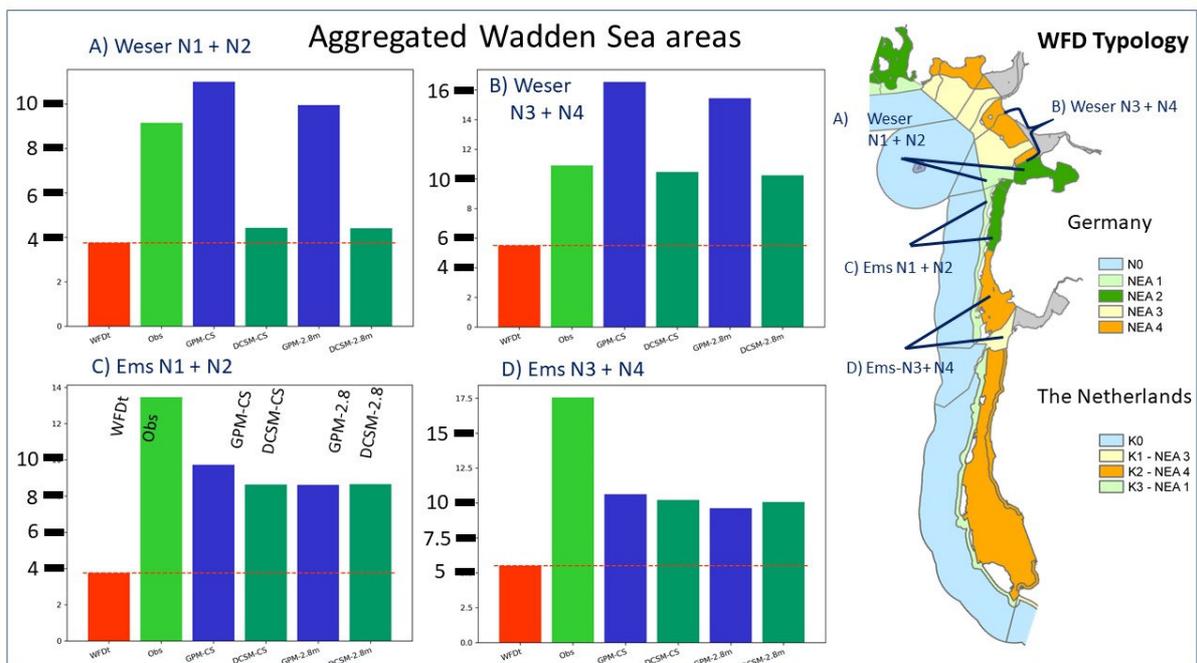
Figure 3.12. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the aggregated areas of the German Wadden Sea. The results are presented for the WFD water bodies A) Weser NEA 1/26, B) Weser NEA 3/4, C) Ems NEA 1/26 and D) Ems NEA 3/4. The chlorophyll observations in water body Ems NEA 3/4 (D) are combined German data only.

3.3 Effects of differences within 2.8-scenario calculation

The calculation of riverine TN loads with a 2.8 TN mg/l threshold is presented in Annex 1. This calculation is based on concentrations, determined by frequent monitoring during the year, which are averaged to get a yearly mean TN concentration. Based on this yearly mean, a reduction percentage needed to reduce to a yearly mean of 2.8 mg/l, was calculated.

In OSPAR ICG-EMO, a slightly different method was used (Annex 2). In the latter case, the calculation is based on yearly TN loads and yearly discharge volumes which are then used to calculate an average TN concentration and a resulting reduction percentage. The different results from both methods can be seen in Annex 2. The Annex 1 method generally results in slightly higher TN loads. It is important to notice that both methods follow the interannual variation in the loads for the different years (Fig. 3.13).

Figure 3.14 shows the results for the two calculation methods of the 2.8-scenario, as demonstrated in Figure 3.12, based on the GPM model application for the Ems region. For the selected areas, the differences in the calculation show that the slightly higher loads in the Annex 1 estimate result in slightly higher Chl concentrations compared to the Annex 2 estimate. However, the differences in the Chl concentrations are very small, which indicates that the impact of the difference between the two applied methods is negligible.
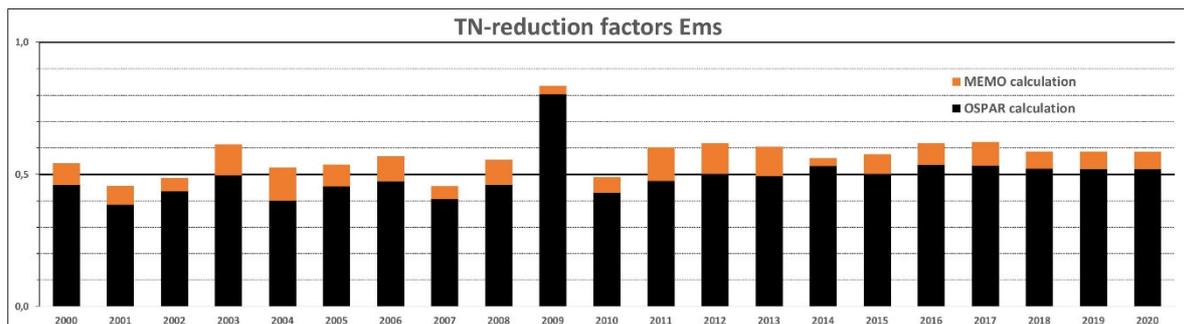
Figure 3.13. Comparison of TN loads in % of current loads based on the calculation methods described in Annex 1 ("MEMO calculation") and the "OSPAR calculation" described in this Annex 2, for the river Ems.
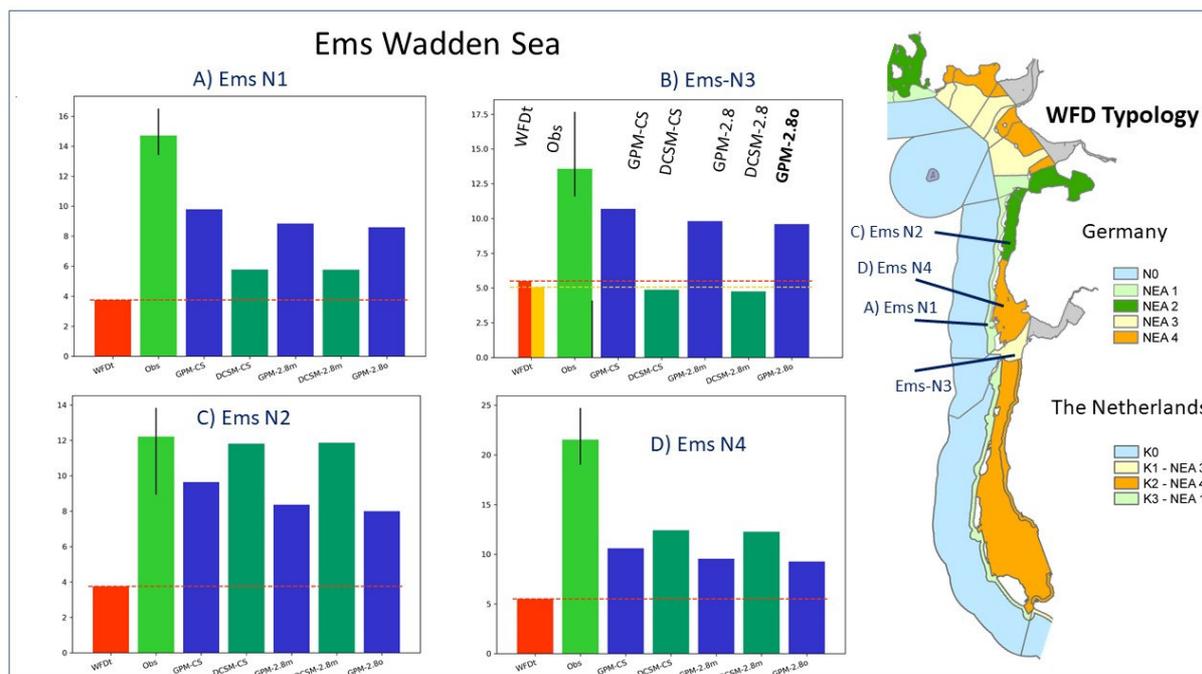
Figure 3.14. Comparison of the results of the 2.8-scenario, the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Ems WFD water bodies A) Ems N1, B) Ems N3, C) Ems N2 and D) Ems N4. In addition to the comparison between the DCMS and the GPM run based on the Annex 1 calculation (indicated by an "m" at the end), simulation results from the Annex 2 calculation are shown (indicated by an "o" at the end) for the GPM model. For Ems N3 (B), two separate chlorophyll thresholds are currently used (red bar: 5.5 µg/l (DE); yellow bar: 5.1 µg/l (NL)).

3.4 First synthesis from this approach

The "2.8-scenario" was applied to get a model estimate of the effect of reducing the riverine TN loads to a level where all rivers would comply with the management objective of an annual average TN concentration of 2.8 mg/l. In this scenario, TP and DIP river loads as well as atmospheric nitrogen deposition in the entire model domain, including the Dutch and German coastal region, were kept the same as in the model runs of the current state. TN and DIN loads of all rivers from other countries discharging into the North Sea were also kept at their current level, therefore the effect of nutrient reduction measures in other countries on the transboundary nutrient transport into the Wadden Sea area is not included here.

In this scenario, the reductions in TN loads from the rivers (compared to current loads) were different for each river, as in some rivers, TN concentrations are already close to the 2.8 mg/l threshold, while other rivers still have much higher TN concentrations. Consequently, the 2.8-scenario showed the strongest reduction in TN loads in the rivers Ems and Eider and only small reductions in the river Rhine and Lake IJssel (Fig. 3.1).

The model results showed that the modelled Chl concentrations in the 2.8-scenario, in general, were only slightly lower than the model estimates for the current state run. Reducing TN loads leads to proportional decreases in winter mean DIN concentrations in the Wadden Sea. But the reduction in Chl concentrations is much more limited (Table 3.1). Causes for the low response of Chl are the role of other co-limiting factors for phytoplankton growth, such as light limitation and P-limitation.

In most cases, current Chl concentrations in the Dutch and German part of the Wadden Sea exceed the WFD thresholds (Good/Moderate boundary), leading to a status classification as 'Moderate' or worse. The level of observed Chl concentrations indicates that in many cases a reduction of Chl concentrations of more than 50% is needed to meet the WFD threshold. Comparing observations to

current WFD thresholds shows that this is clearly the case in for example the Ems Wadden Sea area (Fig. 3.8), the Weser Wadden Sea area (Fig. 3.9) and the Elbe Wadden Sea area (Fig. 3.10).

The limited effect of the TN load reduction on the Chl concentration in the 2.8-scenario may indicate that achieving annual average TN concentrations of 2.8 mg/l in the rivers will not be enough to meet the WFD thresholds for Chl in the Wadden Sea, as the reduction in Chl that is predicted by the models (Table 3.1) is far less than the required >50% reduction estimated from observations.

Obviously, there is uncertainty in the model results as well as in the observations. The models have been shown to accurately reflect spatial gradients in nutrient and Chl concentrations (Figs. 2.23, 2.24, 2.25, 2.26): high Chl concentrations are estimated in areas with high nutrients, low Chl concentrations in areas with low nutrients. This gives confidence in the response of the models to changes in nutrient availability and, consequently, also gives confidence in the response of the models to changes in nutrient loads. Even though the absolute value of Chl concentrations estimated by the models sometimes deviates from the observations, the magnitude of the change in concentrations of DIN, DIP and Chl from the current state to the 2.8-scenario gives a reliable indication of the reduction that can be expected.

Table 3.1. Relative decrease in concentrations of DIN, DIP and chlorophyll in the 2.8-scenario, compared to the current state (CS model run). Changes are based on the average concentration per water body. Reduction is calculated as %reduction=(Concentration$_{2.8}$-Concentration$_{CS}$)/Concentration$_{CS}$*100%.

| WFD Water Body | Concentration reduction compared to current state | | | | | |
| | DIN | | DIP | | CHL | |
	DCSM-FM	SNS-GPM	DCSM-FM	SNS-GPM	DCSM-FM	SNS-GPM
NL Wadden Coast N1	-3%	-11%	0%	0%	-1%	-2%
NL/DE Ems N3*	-18%	-33%	0%	-1%	-2%	-8%
NL Wadden Sea N4	-2%	-6%	0%	0%	-2%	-2%
DE Ems N1	-10%	-27%	0%	-1%	0%	-10%
DE Ems N2	-13%	-27%	0%	-1%	0%	-13%
DE Ems N4	-20%	-33%	-1%	-1%	-1%	-10%
DE Weser N1	-9%	-25%	0%	-1%	0%	-11%
DE Weser N2	-13%	-25%	0%	-1%	0%	-9%
DE Weser N3	-14%	-24%	0%	-1%	0%	-7%
DE Weser N4-01	-18%	-25%	0%	-1%	1%	-5%
DE Weser N4-02	-15%	-23%	0%	-1%	-4%	-6%
DE Elbe N3	-12%	-19%	0%	-1%	1%	-3%
DE Elbe N4-01	-13%	-19%	0%	-1%	-5%	-4%
DE Elbe N4-02	-14%	-17%	-1%	-1%	1%	-2%

* For the Ems N3 area, model results are the same for the Duch water body "Ems-Dollard coast" and the German water body "Ems N3".

4. Derivation of threshold values using pre-eutrophic reference conditions

4.1 Description of the pre-eutrophic reference conditions

Under the WFD, the reference condition is "a description of the biological quality elements that exist, or would exist, at high status. That is, with no, or very minor disturbance from human activities." (EC 2003a). In the CIS Guidance a hierarchical approach is proposed to determine reference conditions:

1. An existing undisturbed site or a site with only very minor disturbance; or
2. historical data and information; or
3. models; or
4. expert judgement.

Similar to other coastal and transitional waters (EC 2003a), there are no existing undisturbed sites in the NE Atlantic in general and the Greater North Sea and the Wadden Sea in particular, that have comparable conditions in terms of e.g., freshwater input, light conditions, hydrodynamics. Consequently, there are no sites that could serve as a reference site. Historical data of sufficient quality that could provide information on the undisturbed state of the coastal and marine waters are also lacking. The next approach that could be used is modelling as described by EC (2003a): *"Type-specific biological reference conditions based on modelling may be derived using either predictive models or hindcasting methods. The methods shall use historical, palaeological and other available data and shall provide a sufficient level of confidence about the values for the reference conditions to ensure that the conditions so derived are consistent and valid for each surface water body type."*

In the recent exercise by OSPAR's Intersessional Correspondence Group on Ecosystem Modelling (ICG-EMO) to define coherent eutrophication thresholds for the 4th integrated eutrophication assessment, it was therefore decided to define thresholds based on a description of pre-eutrophic conditions in the marine environment (Lenhart et al. 2022). In this Interreg project we followed the same approach to determine WFD thresholds for the Wadden Sea and adjacent coastal water bodies.

The pre-eutrophic reference conditions defined by OSPAR were considered to be the background from which threshold concentrations for nutrients and Chl were derived by adding 50% to the background concentration. The threshold concentrations defined by OSPAR can be considered to be equivalent to the Good/Moderate boundary of the WFD (Fig. 4.1). Note that according to EC (2003b) *"Reference conditions (RC) do not equate necessarily to totally undisturbed, pristine conditions. They include very minor disturbance which means that human pressure is allowed as long as there are no or only very minor ecological effects"*.
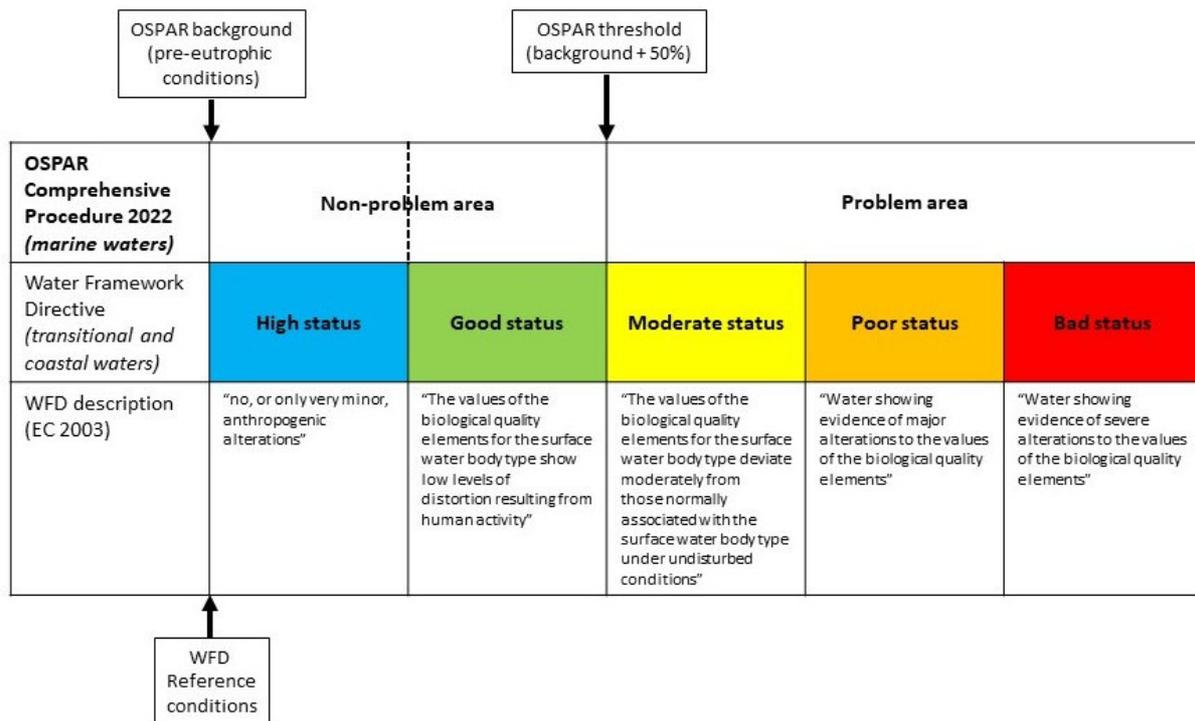
Figure 4.1. Scheme illustrating the relation between OSPAR threshold values for eutrophication assessment in marine waters and the WFD Good/Moderate boundary for the assessment in transitional and coastal waters.

The definition of the pre-eutrophication conditions was based on initial work of the JMP-EUNOSAT project (Blauw et al. 2019, Enserink et al. 2019). After the finalisation of the JMP EUNOSAT project, ICG EMO and the OSPAR Task group for the Comprehensive Procedure (TG COMP) made an effort to develop a common 'narrative' for a joint 'historic, pre-eutrophication' scenario based on the initial work of JMP Eunosat. The final approach is described in Lenhart et al. (2022) and Annex 7 of the OSPAR Comprehensive Procedure (OSPAR 2022).

Here, we summarize the main assumptions and steps, for a more detailed description we refer to Blauw et al. (2019) and Lenhart et al. (2022):

1) Pre-eutrophic period
   - The "pre-eutrophic" conditions were defined by describing conditions at the end of the 19th century. The describes an era with some level of anthropogenic disturbance (i.e. different from pristine conditions) but without the large and widespread use of nitrogen fertilizer, intensification of agriculture and emissions to surface waters (Galloway et al. 2003, Erisman et al. 2011, Sutton et al. 2011, Bouwman et al. 2013).
2) River loads
   - For those pre-eutrophic conditions, riverine nutrient loads to the sea were estimated using estimates from HYPE (Hydrological Predictions for the Environment), which is an integrated rainfall-runoff and nutrient transport model developed by SMHI under a Creative Commons open-source licence (Lindström et al. 2010). E-HYPE v.3.1.3 was used to simulate current day nutrient loads. The model input files were then modified to approximate conditions in 1900 and the model was executed without any further adjustment in the model parameters to generate the loads corresponding to 1900. Model input files with land use, agriculture practices such as irrigation and fertilization,

population-driven sources of pollution, and atmospheric deposition were modified based on available information.

3) Land use
   - Land use data were acquired from History Database of the Global Environment (HYDE) developed under the authority of the Netherlands Environmental Assessment Agency (Klein Goldewijk et al. 2010, Klein Goldewijk et al. 2011). HYDE presents time series of land use and population developed on a 5-minute grid (about 85 km^2 grid cell around the equator).

4) Human waste
   - HYDE data on urban and rural population in 1900 was used to estimate the number of people living in urban and rural settings in each catchment. The nutrient loads from the population were calculated using Population Equivalent (PE). The PE was estimated from diet and other factors relevant in 1900 (Schmid 2000, Smil 2000).

5) Agricultural practices
   - The amount of fertilizers applied to crops was adjusted as follows. The maximum application rate was set to 100 kg N per hectare and 50 kg P per hectare (Smil 2000, van Grinsven et al. 2015). The application rates were assumed to remain in the same relative proportion for different crops as in E-HYPE v.3.1.3.
   - The Haber-Bosch process that industrialized production of inorganic nitrogen fertilizers was first demonstrated in 1909 with a first industrial-level production starting a few years later. All nitrogen applied to fields was thus assumed to be in organic form (manure) in 1900. The phosphorus application rate was assumed to be applied to crops as manure for 80% of the application rate and as inorganic fertilizers for 20% of the application rate. This gives application rates in accordance with Kyllingsbæk (2005), Knudsen & Schnug (2016) and Vinther (2012). Any phosphorus applied on other land uses (e.g. pastures) was assumed to be in the organic form only.

6) Atmospheric deposition
   - Atmospheric deposition rates were calculated based on the estimations by Schöpp et al. (2003).

OSPAR TG-COMP decided to make some amendments to the estimates of reference nutrient loads from E-HYPE, as riverine P-loads in the E-Hype estimates were substantially higher than estimates from other catchment models for Germany (Moneris model) and Denmark (for details see Lenhart et al. 2022, van Leeuwen et al. 2023). Consequently, there were two scenarios with different estimates for historic phosphorous loads, while nitrogen loads were the same in both scenarios. Historic scenario HS1 is the scenario with the E-Hype results, historic scenario HS2 represents the scenario with adapted phosphorous loads.

Figure 4.2 provides an overview of the reduction level of the two historic scenarios HS1 and HS2 for the Dutch and German rivers compared to the current loads. One can clearly see that in the HS1 scenario the nitrogen load reduction is higher than the reduction in phosphorus loads. However, in the HS2 scenario, for a number of rivers like Rhine and Elbe, the phosphorus load reduction is relatively larger than the related nitrogen load reduction.
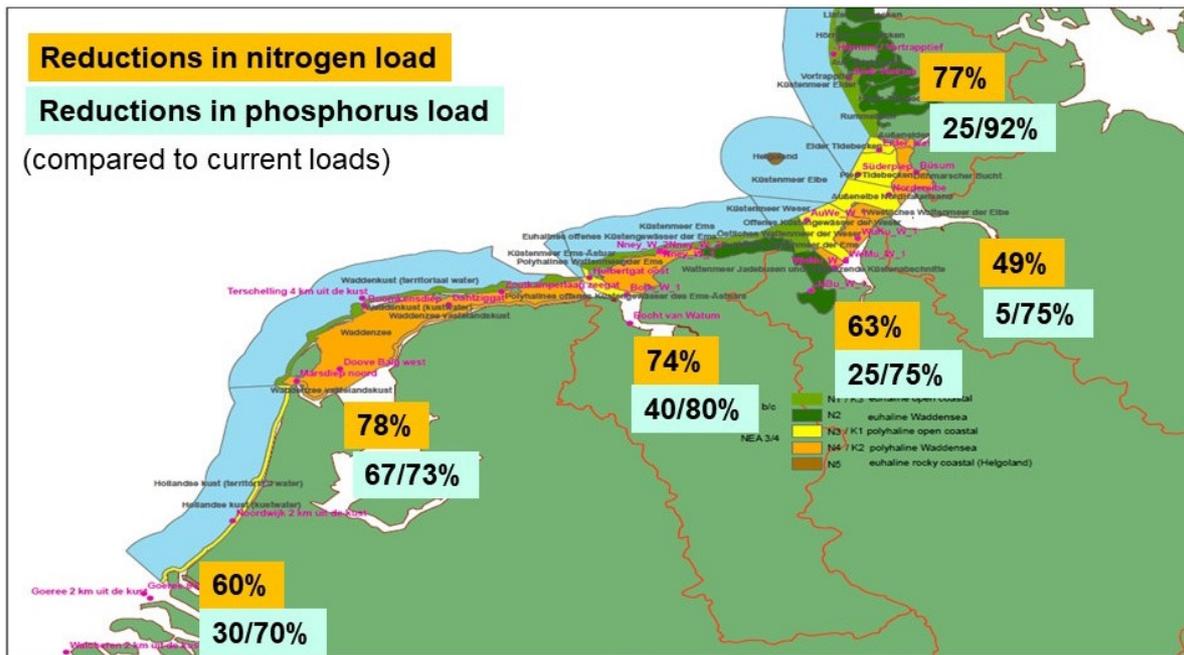
Figure 4.2. Overview of the load reduction compared to current loads (mean for 2009-2017) for the two pre-eutrophic historic scenarios (HS1 & HS2) in the major rivers, applied by OSPAR (Lenhart et al. 2022). Nitrogen loads are the same in both historic scenarios, phosphorus loads differ between the scenarios; the first number is scenario HS1, the second number is scenario HS2.

4.2 Evaluation of model runs

The focus of the evaluation of the historic OSPAR scenario is based on the simulation from the GPM model for the years 2014 – 2017. For comparison, the Chl distribution from the DCSM model will also be shown but based on the original OSPAR simulation period 2009-2014. Tests show that the reduction in the river loads between the years are the same in both applications, therefore only the differences between the years in the forcing need to be taken into account. For the final aggregation of the threshold values this should be negligible.

4.2.1 Spatial patterns

Figure 4.3 shows the difference in DIN concentration between the CS run and the HS1 and the HS2 scenario. One can clearly see high reduction levels in the DIN concentration in the HS scenarios along the coast. In the vicinity of the Elbe outflow the reduction level is reduced, which reflects the lower reduction in TN loads from the river Elbe in the HS scenarios (Fig. 4.2). Since there is no difference between the HS1 and the HS2 scenario, the patterns for the DIN distribution are the same for both scenarios. The increase in DIN concentration in the norther part of the GPM model domain (red area in Fig. 4.3) is related to boundary problems of the GPM model and a model artefact.
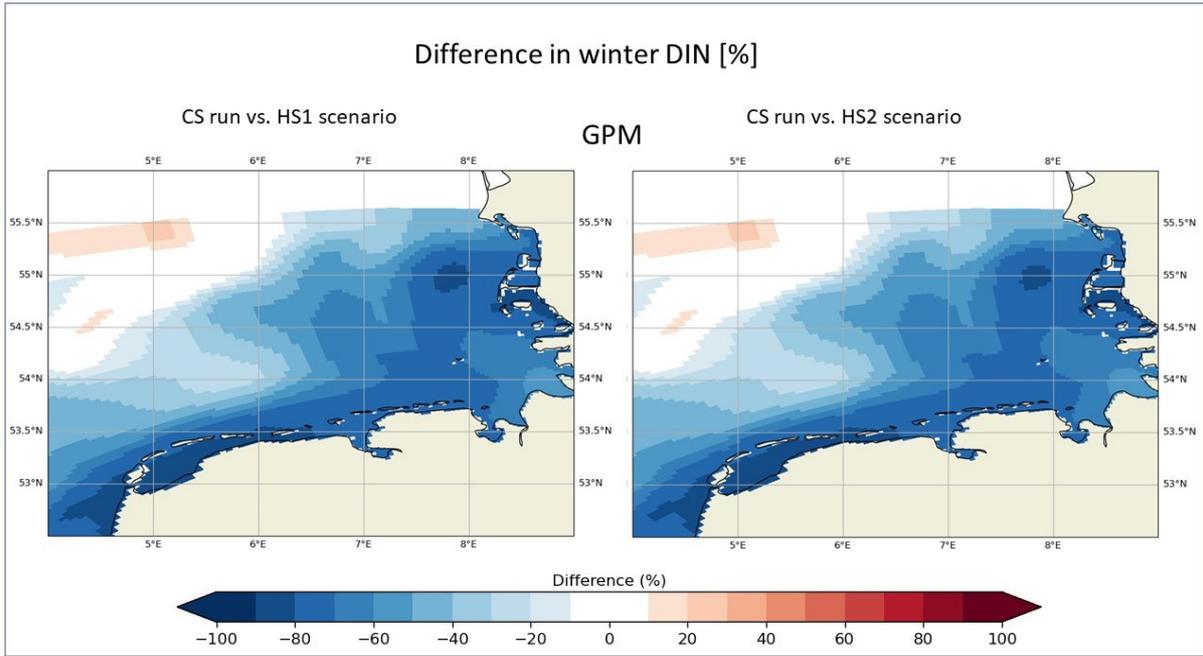
Figure 4.3. Horizontal distribution map of the percent difference between the CS simulation vs. the two historic OSPAR scenario for DIN for the Southern North Sea from the GPM model. Blue colours indicate lower concentrations in the historic scenarios.
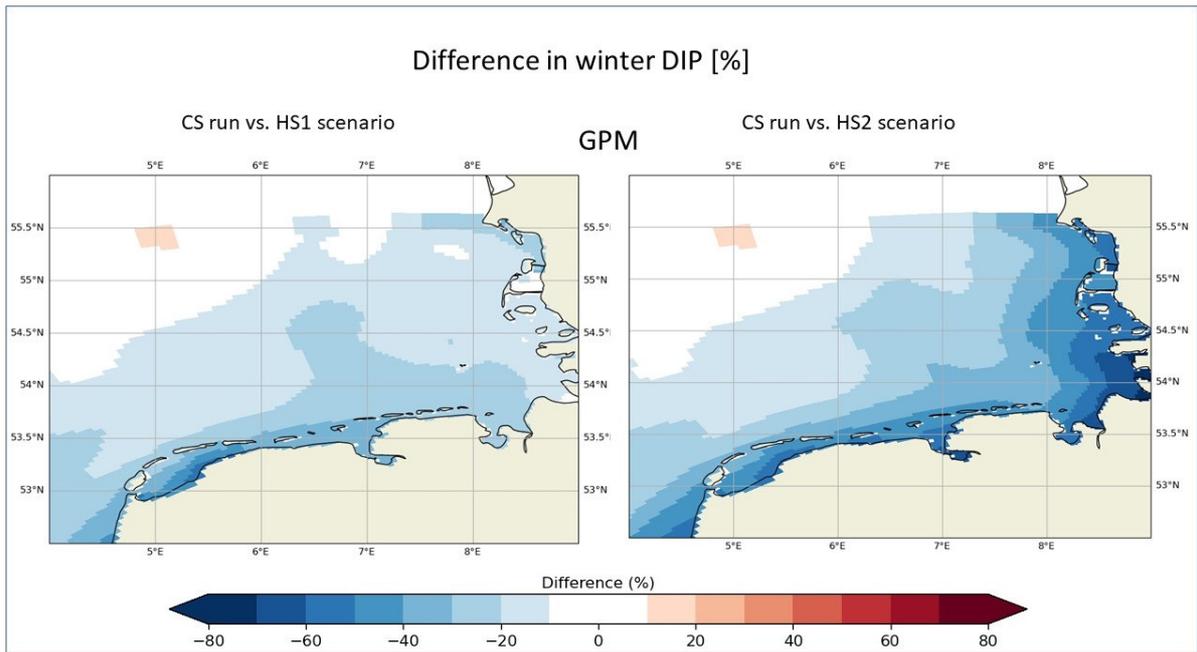
Figure 4.4. Horizontal distribution map of the percent difference between the CS simulation vs. the two historic OSPAR scenario for DIP for the Southern North Sea from the GPM model. Blue colours indicate lower concentrations in the historic scenarios.

Figure 4.4 shows the distribution of the DIP concentration difference between the CS run and the HS1 and the HS2 scenario. Since the two scenarios differ considerably in their phosphorus reduction (Fig. 4.2), the resulting distribution shows major differences. In the HS1 scenario the highest reduction is achieved within the western coastal region of the Duch Wadden Sea. Overall, the Wadden Sea areas of the Netherlands and Lower Saxony show reduction levels between 20 – 40 %. Near the Elbe region

the reduction is considerably smaller, which is related to the small reduction level of 5% in TP loads from the Elbe in this scenario. With a reduction level of 75 % in the historic TP loads in HS2, the Elbe region shows the maximum reduction of DIP concentrations, also affecting the Schleswig Holstein Wadden Sea region. For the Dutch and the Lower Saxony Wadden Sea the level of reduction in HS2 is higher than the HS1 scenario, but the horizontal distribution remains nearly the same.
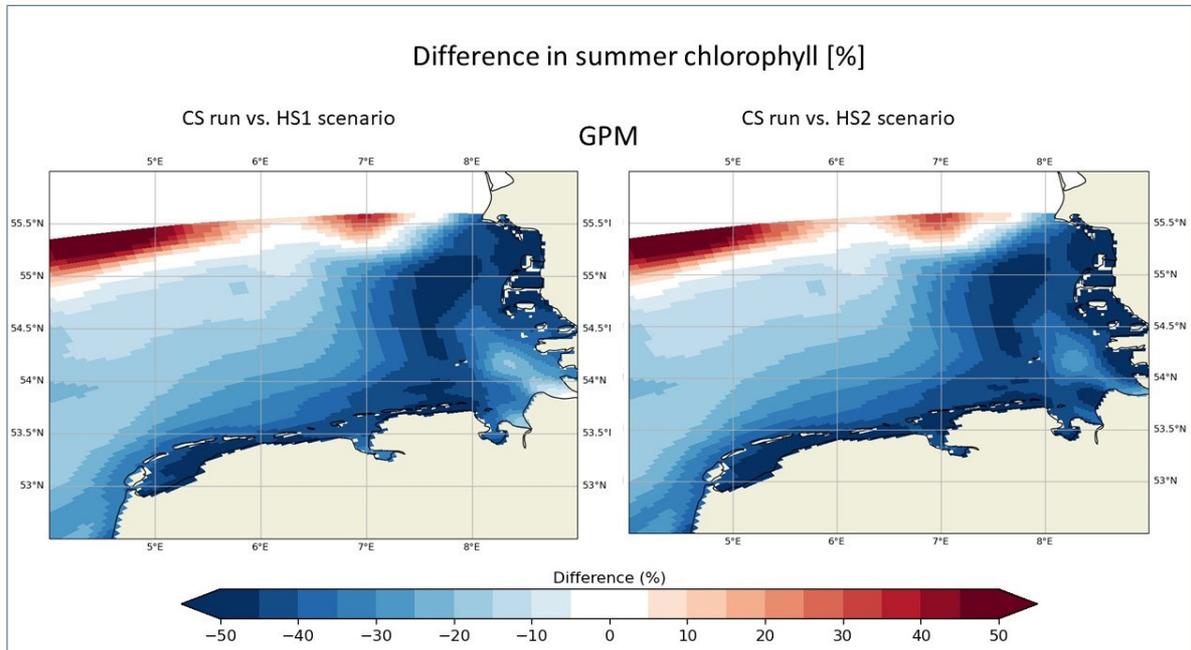
Figure 4.5. Horizontal distribution map of the percent difference between the CS simulation vs. the two historic OSPAR scenario for chlorophyll for the Southern North Sea from the GPM model. Blue colours indicate lower concentrations in the historic scenarios.

Figure 4.5 represents the difference in the Chl concentrations between the CS run and the HS1 and the HS2 scenario. In the HS1 scenario the highest Chl reduction occurs in the western Dutch Wadden Sea and west of the Elbe inlet in Lower Saxony. There is another local maximum within the German Bight, while the vicinity of the Elbe and Weser outflow shows limited response to the reduced nutrient loads. The results for the HS2 scenario are largely the same as for HS1 but the coastal region between the Weser and the Elbe reflects a higher reduction. The increase in Chl concentrations in the norther part of the GPM model domain (red area in Fig. 4.5) is related to boundary problems of the GPM model.
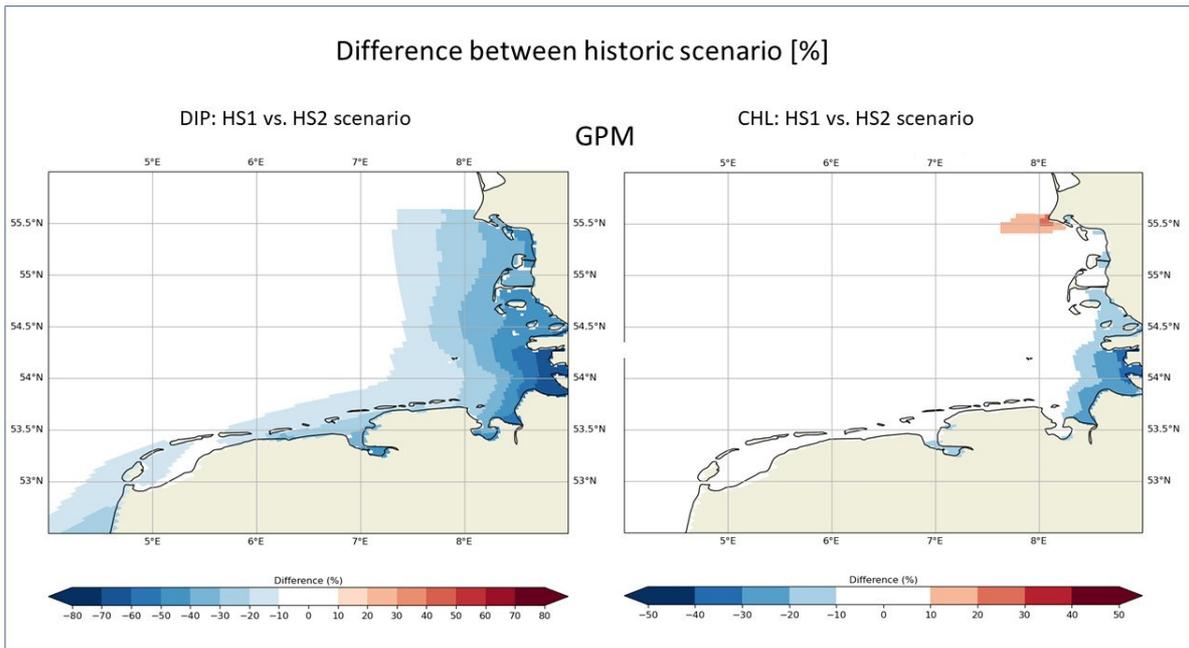
Figure 4.6. Horizontal distribution map of the percent difference between the HS1 vs. the HS2 historic OSPAR scenario for DIP and chlorophyll for the Southern North Sea from the GPM model. Blue colours indicate lower concentrations in scenario HS2.

Figure 4.6 shows the differences in DIP and Chl concentration between the HS1 and the HS2 scenario. One can clearly see the higher impact of the HS2 scenario for DIP in the Elbe mouth, along the coast of Schleswig-Holstein and near the mouth of the river Weser. The coastal region of the western Dutch Wadden Sea shows no difference at all, which might be related to the fact that there is a substantial distance towards the next major river source, like the Rhine. For Chl the differences between the two scenarios appear as a very local phenomenon. There is an increased reduction level in HS2 in the Elbe region and in the Ems region.
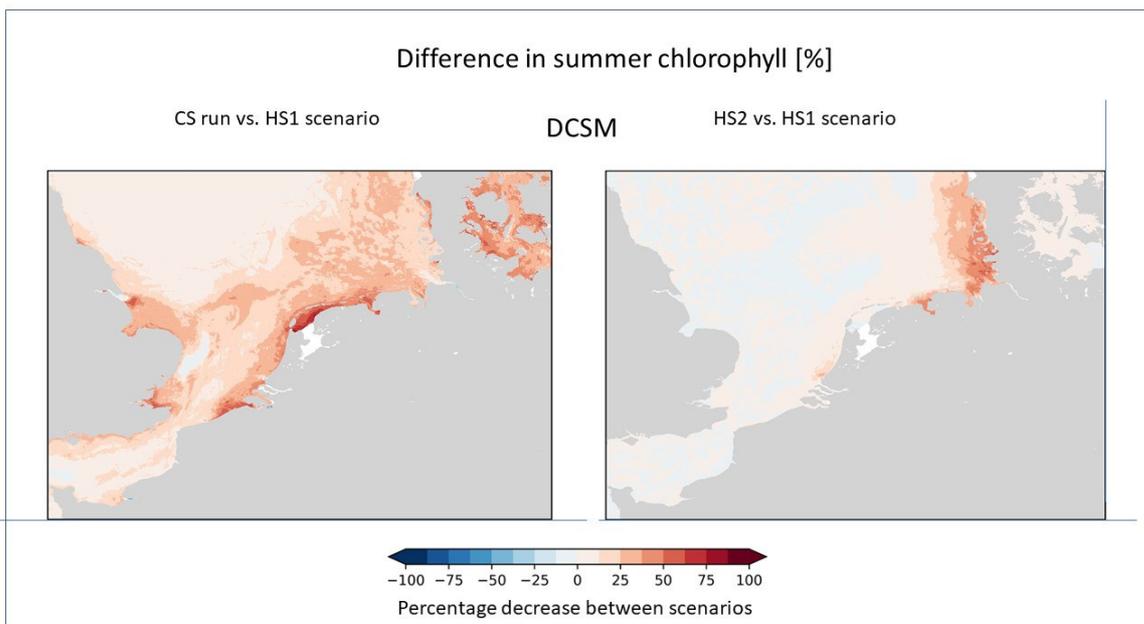
Figure 4.7. Horizontal distribution map of the percent difference between CS simulation vs. the historic HS1 scenario (left) and between the HS1 vs. the HS2 historic OSPAR scenario (right) for chlorophyll for the Southern North Sea from the DCSM model. Red colours indicate lower concentrations in scenario HS1 compared to CS (left) and in HS2 compared to HS1 (right).

Figure 4.7 shows the difference in Chl concentration between the CS run and the HS1 scenario and the difference between the HS1 and HS2 scenario (results DCSM model). The difference between HS1 and CS shows a pattern similar to the results of the GPM model shown in Figure 4.5. Both models also show similar differences between HS1 and HS2, but the DSMM model shows impacts further offshore than the GPM model shown in Figure 4.5.
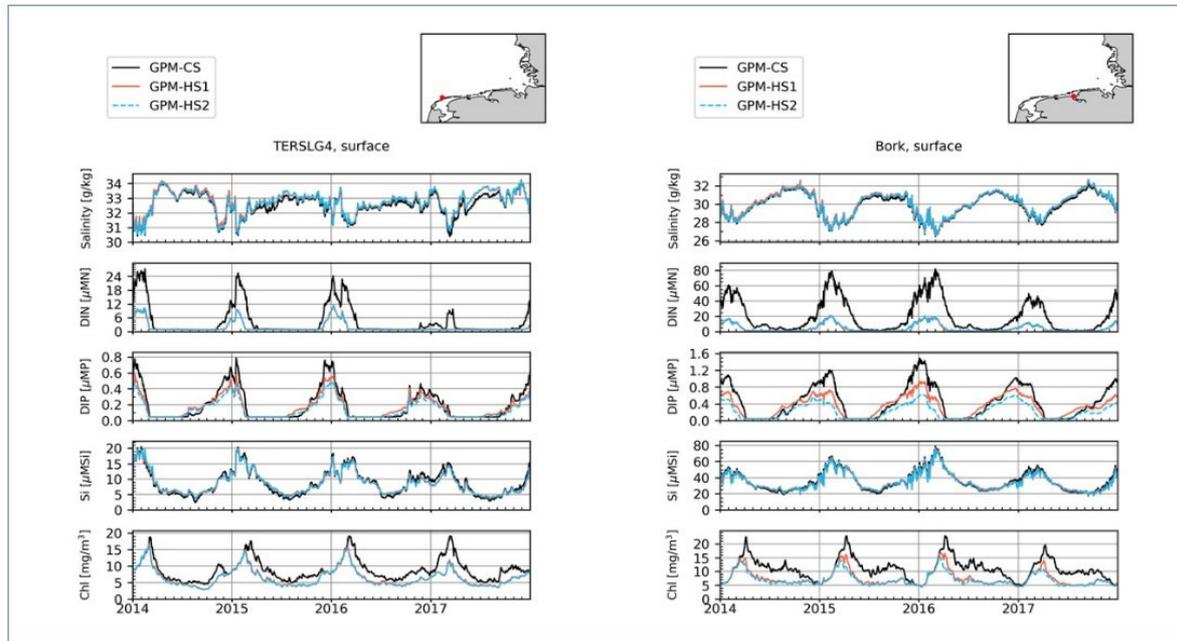
Figure 4.8. Time series of salinity, DIN, DIP, Si and chlorophyll at the stations Terschelling 4 km (left) and Borkum (right) showing the CS simulation (black), the HS1 run (red) and HS2 run (blue dotted) scenario from the GPM model.

4.2.2 Time series

Figure 4.8 shows the timeseries at Terschelling 4 and Borkum for salinity, nutrients and Chl from the historic scenarios and the CS run. At both stations, both nutrients DIN and DIP show a clear reduction in the concentration in the HS1 and HS2 time series. Since there is no difference in TN loads between HS1 and HS2, the time series for DIN are the same. The timeseries for DIP indicates much lower concentrations in the HS2 scenario, in line with the lower TP loads.

The Chl historic time series for Borkum (Fig. 4.8 right) show a considerably lower concentration compared to the CS run and in addition, small difference between the two scenarios. For Terschelling 4 (Fig. 4.8 left) there is no difference between the two historic scenarios, but both historic concentrations are consistently lower in comparison to the CS run.
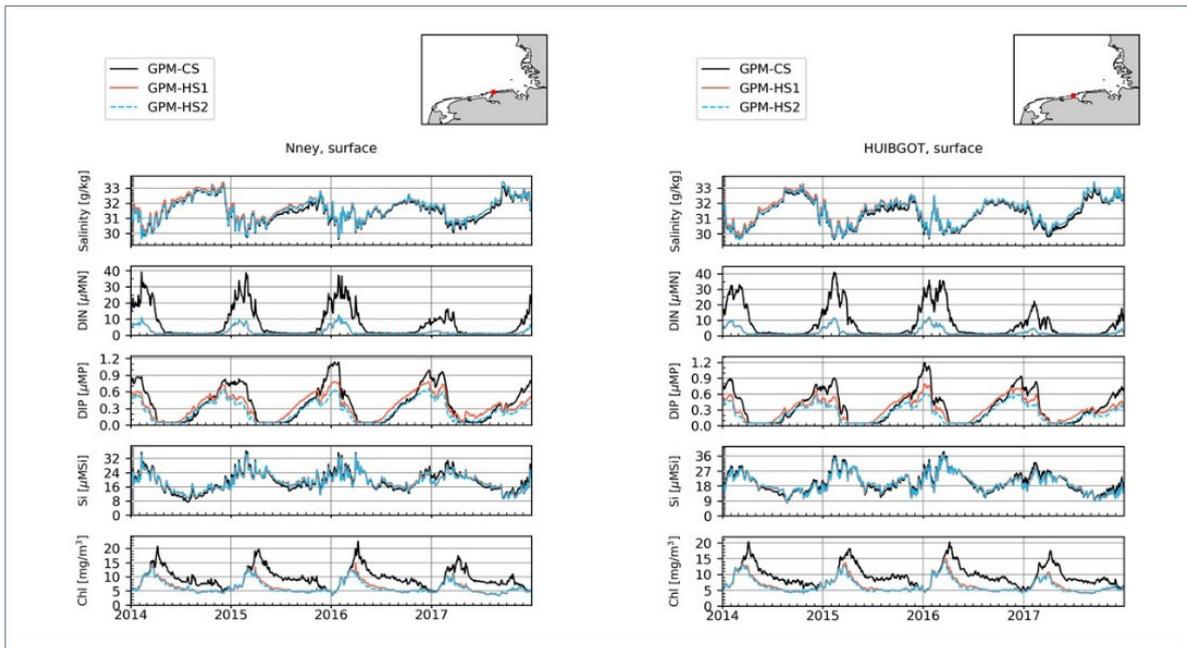
Figure 4.9. Time series of salinity, DIN, DIP, Si and chlorophyll at the stations Norderney (left) and Huibertgat oost (right) showing the CS simulation (black), the HS1 run (red) and HS2 run (blue dotted) from the GPM model.

Figure 4.9 shows the time series at Norderney and Huibertgat oost for salinity, nutrients and Chl. At both stations DIN time series of the two historic scenarios are lower than the CS run. There is no difference between the two historic scenarios. In the DIP timeseries one can identify a distinctly higher DIP concentration during the period of late summer and autumn for the HS1 scenario, even higher than in the CS run. When looking at the corresponding timeseries of the Chl concentration, one can see that the phytoplankton increase towards the spring bloom in the HS1 and HS2 scenario stops at a lower level than the CS run, and also summer concentrations are lower than in the CS run. Differences in Chl concentrations between the two historic timeseries are small but suggest that phytoplankton suffers from P-limitation in spring at both stations.
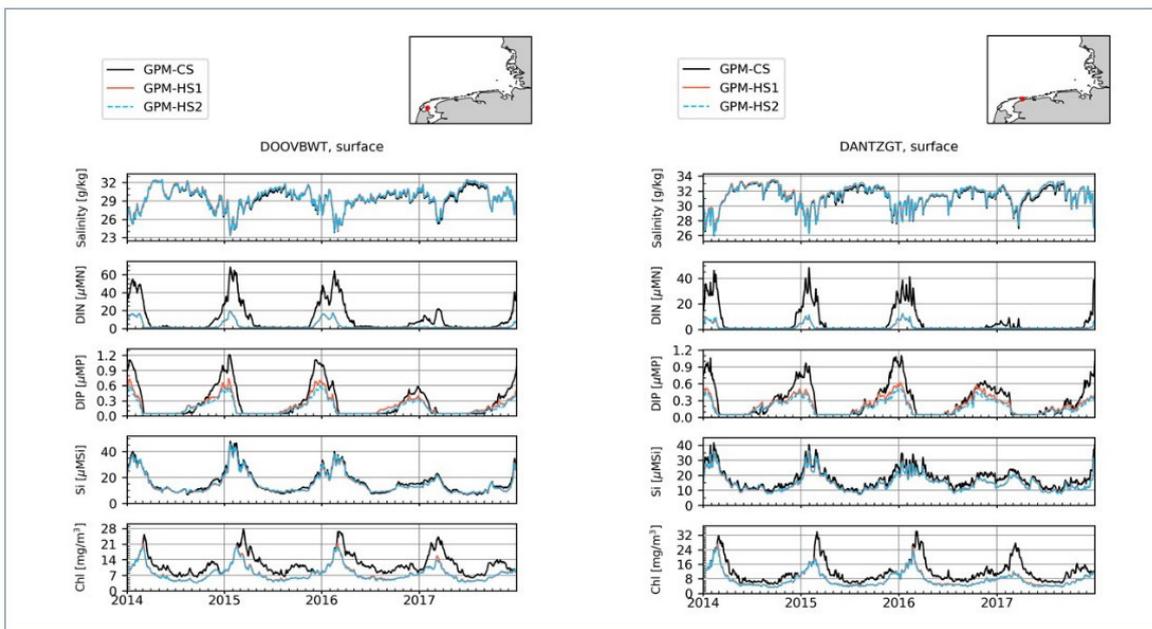
Figure 4.10a. Time series of salinity, DIN, DIP, Si and chlorophyll at the stations Doove Balg West (left) and Dantziggat (right) showing the CS simulation (black), the HS1 run (red) and HS2 run (blue dotted) from the GPM model.

Figure 4.10a shows the timeseries at Doove Balg and Danziggat for salinity, nutrients and Chl derived from the historic simulation from the GPM model in comparison to the CS run. Basically, the timeseries are very similar to the ones described for Norderney and Huibertgat oost in Figure 4.9. One small difference is the fact that the DIP concentration between the two historic runs do not differ much and do not exceed the concentration from the CS run. The Chl concentration show the same clipping of the spring bloom and a lower summer standing stock.
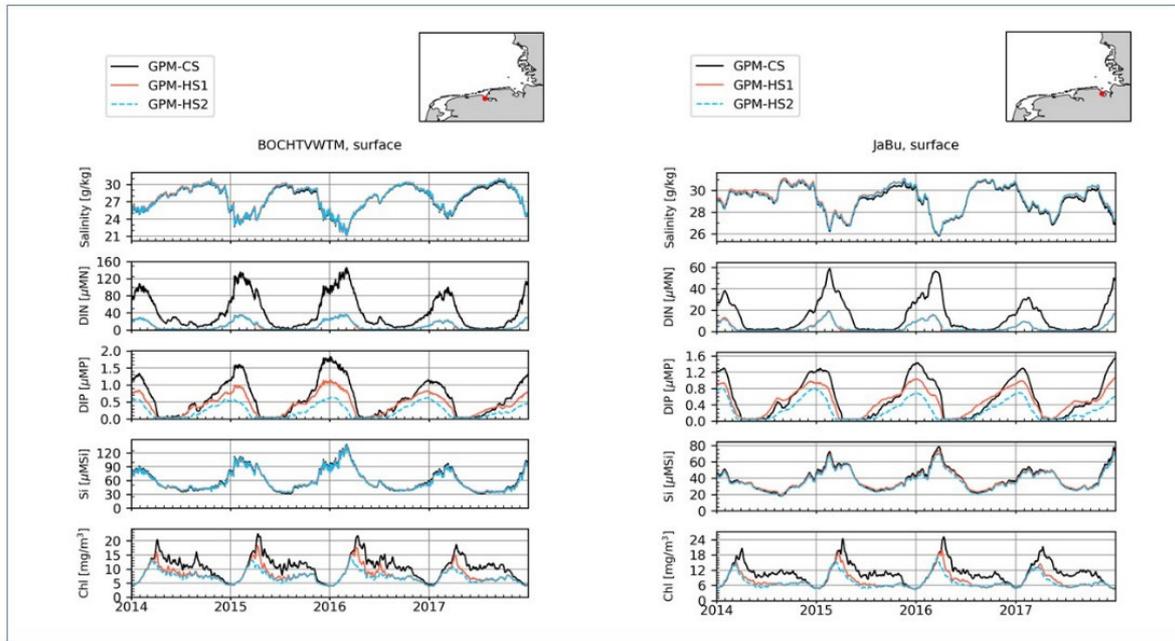
Figure 4.10b. Time series of salinity, DIN, DIP, Si and chlorophyll at the stations Bocht van Watum (left) and Jadebusen (right) showing the CS simulation (black), the HS1 run (red) and HS2 run (blue dotted) from the GPM model.

Figure 4.10b shows the time series at Bocht van Watum and Jadebusen for salinity, nutrients and CHL. At both stations the historic DIN concentrations are much lower than in the CS run. Also, in the DIP timeseries the historic concentrations are much lower than in the CS run, and in addition there is a strong difference between the two scenarios HS1 and HS2, with considerably lower concentrations in the HS2 scenario. The Chl concentration in the HS2 scenario is lower than the HS1 run in the beginning of the spring bloom until the decline towards the summer standing stock.

4.2.3 Results per water body type

The results from the historic scenarios HS1 and HS2 provided pre-eutrophic reference Chl concentrations, which were used to calculate average concentrations per water body type. From these average pre-eutrophic references, threshold concentrations of Chl were derived by increasing the concentrations with 50%. In the following figures, the derived thresholds are compared to the results from the Current State run (Chapter 2.3), the "2.8-scenario" (Chapter 3.2.3) and to the WFD threshold (Good/Moderate boundary) that is currently applicable for the assessment of phytoplankton in the WFD water bodies.
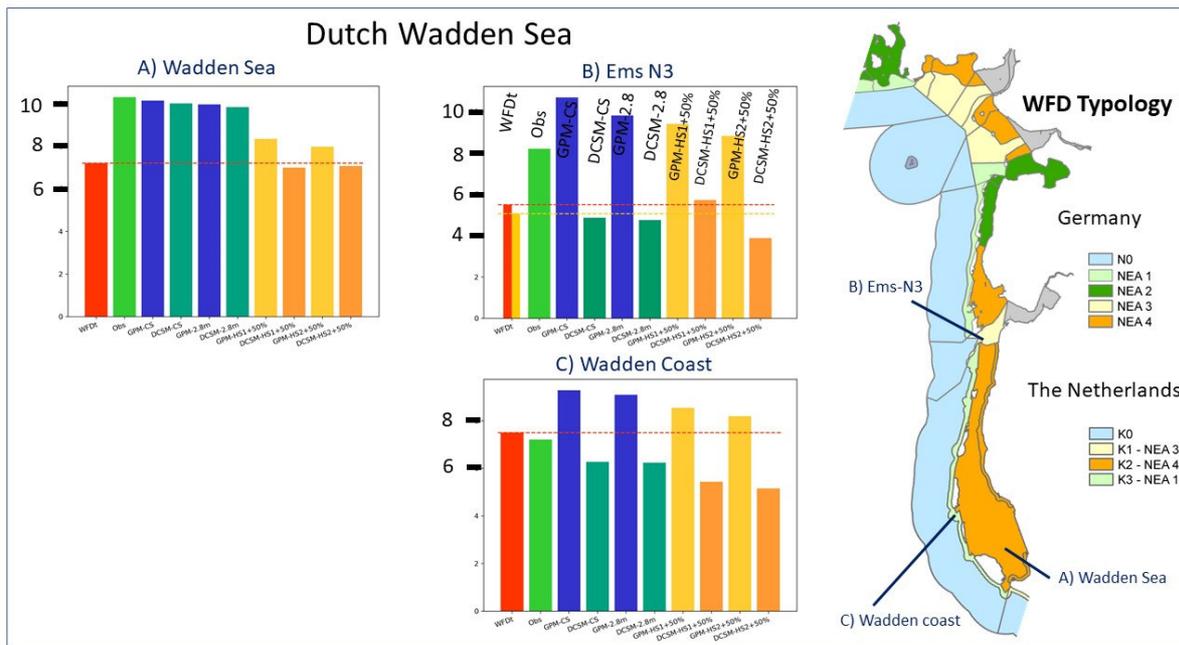
Figure 4.11. Chlorophyll thresholds (calculated from the two historic scenarios + 50 %), chlorophyll concentrations from observations and the CS simulation and the existing WFD threshold, for the Dutch Wadden Sea. The results are presented for the Dutch WFD water bodies A) Wadden Sea, B) Ems N3 (Ems-Dollard Coast), C) Wadden coast. For Ems N3 (B), two separate chlorophyll thresholds are currently used (red bar: 5.5 µg/l (DE); yellow bar: 5.1 µg/l (NL)). Shown chlorophyll observations represent NL measurements here.

Figure 4.11 provides an overview on the resulting Chl thresholds for the Dutch Wadden Sea area based on the historic scenario. In Figure 4.11A the level derived from the historic runs is clearly lower than the ones from the 2.8-scenario, for both model applications. While the GPM threshold for both historic scenario HS1 and HS2 are slightly above the existing threshold, the DCSM results for both historic runs are just below the threshold. Surprisingly, the Chl concentration from the DCSM HS2 scenario appears to be slightly higher in the HS2 scenario. In general, the low Chl concentration matches well with the strong response in the western Dutch Wadden Sea which can be observed in Figure 4.5. For the Ems N3 region (Fig. 4.11B) the mismatch between the two models in the Chl concentration is obvious. As already shown in Figure 3.11, the DCSM model has much lower Chl concentrations for both historic scenarios as well as for the 2.8-scenario in comparison to the GPM model. The DCSM model also shows a stronger effect of the HS2 scenario on Chl concentrations than the GPM model. The Chl thresholds from the historic GPM runs are above the existing WFD threshold (Good/Moderate boundary), while the DCSM model provides thresholds close to or below the WFD threshold. The Wadden coast shows similar differences between the two models as in Ems N3.
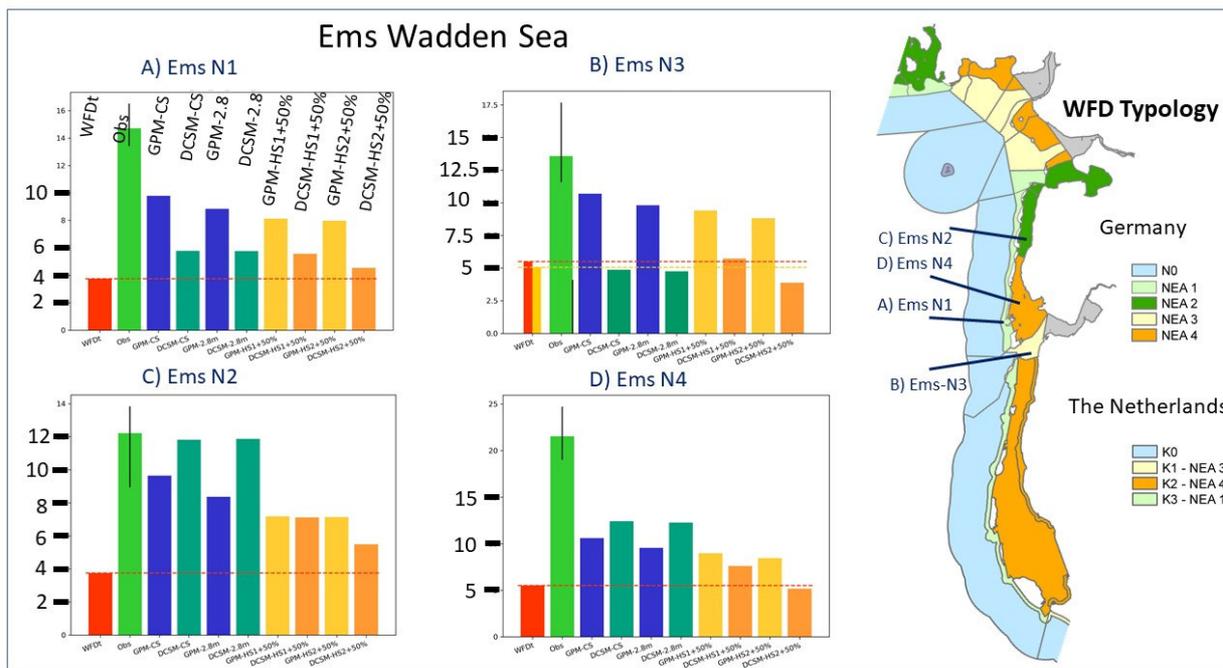
Figure 4.12. Chlorophyll thresholds (calculated from the two historic OSPAR scenarios +50 %), chlorophyll concentrations from observations and the CS simulation and the existing WFD threshold, for the German Wadden Sea. The results are presented for the Ems WFD water bodies A) Ems N1, B) Ems N3, C) Ems N2 and D) Ems N4. For Ems N3 (B), two separate chlorophyll thresholds are currently used (red bar: 5.5 µg/l (DE); yellow bar: 5.1 µg/l (NL)). Shown chlorophyll observations represent DE measurements here.

Figure 4.12 gives an overview of the Chl thresholds for the Ems Wadden Sea area based on the historic scenario, in comparison to observations, current state and 2.8-scenario runs. The results for the Ems N1 (Fig. 4.12A) and the Ems N2 region (Fig. 4.12C) show consistently higher Chl thresholds for all historic scenarios in comparison to the existing WFD threshold. The results for the Ems N3 region (Fig. 4.12B) are identical to Figure 4.11B, except that Chl observations differ depending on country. For the Ems N4 region (Fig. 4.12D) only the DCSM simulation for the HS2 scenario results in Chl thresholds below the existing WFD threshold value. The DCSM results show a larger effect of HS2 compared to HS1 than the GPM model.
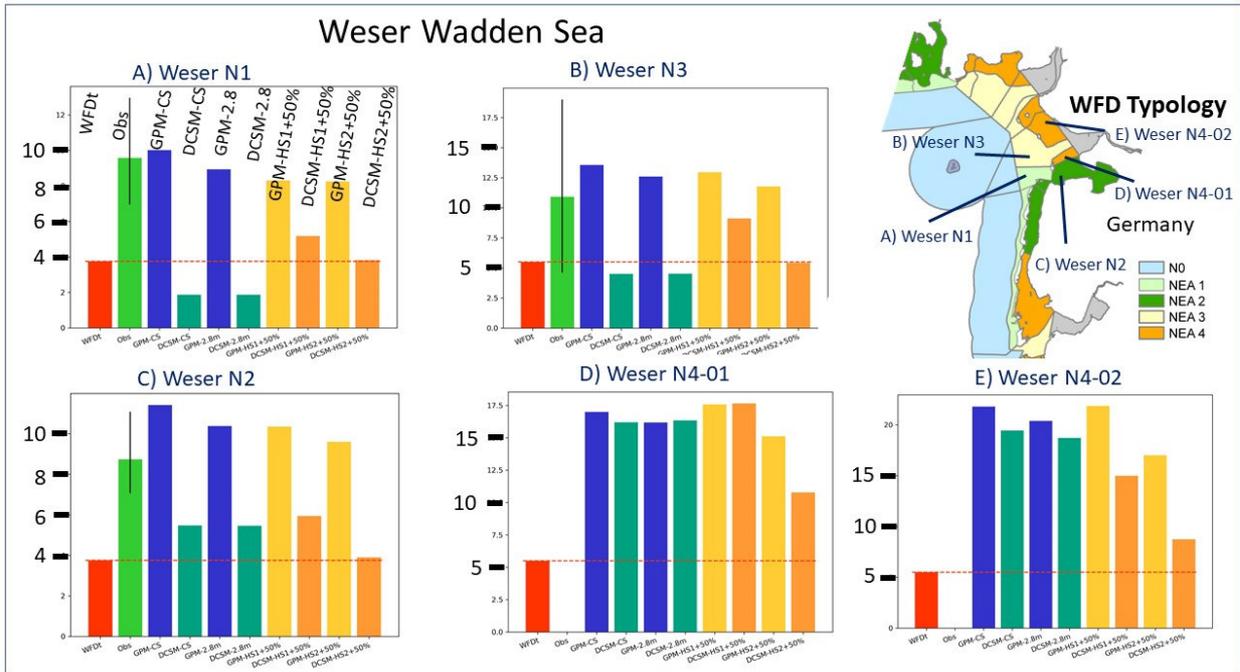
Figure 4.13. Chlorophyll thresholds (calculated from the two historic OSPAR scenarios +50 %), chlorophyll concentrations from observations and the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Weser WFD water bodies A) Weser N1, B) Weser N3, C) Weser N2 and D) Weser N4.

Figure 4.13 gives an overview on the resulting Chl thresholds for the Weser Wadden Sea area based on the historic scenario, in comparison to the observations, current state and 2.8-scenario runs. The DCSM results for the HS2 scenario result in a Chl threshold just around the existing WFD threshold for the Weser regions N1 (Fig. 4.13A), N2 (Fig. 4.13B) and N3 (Fig. 4.13C). All other simulation results for the historic scenario from both models show higher Chl thresholds than the existing WFD threshold. For Weser region N1and N3, the Chl concentrations in the CS run and 2.8-scenario of the DCSM simulation are lower than in the resulting HS2 result. For the Weser regions the Chl thresholds from the HS scenarios of both models exceed the existing WFD threshold.
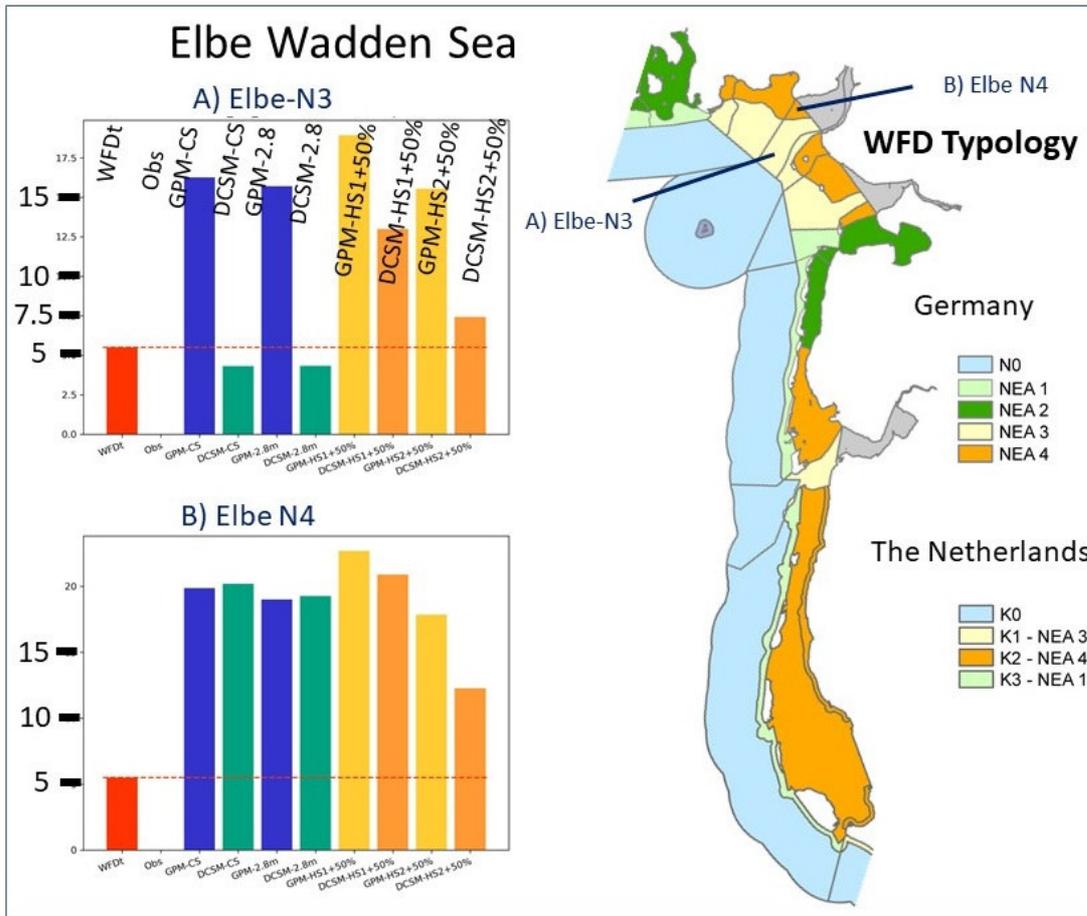
Figure 4.14. Chlorophyll thresholds (calculated from the two historic OSPAR scenarios + 50%), chlorophyll concentrations from observations and the CS simulation and the existing WFD threshold for the German Wadden Sea. The results are presented for the Elbe WFD water bodies A) Elbe N3, B) Elbe N4.

Figure 4.14 gives an overview on the resulting Chl thresholds based on the historic scenario, in comparison to the 2.8-scenario runs. For the Elbe region N3 (Fig. 4.14A) and Elbe N4 (Fig. 4.14B) all Chl concentrations from the historic scenarios exceed the existing WFD threshold. For both Elbe areas (Fig. 4.14A) the Chl concentrations in the CS run and the 2.8-scenario of both models are lower than the thresholds derived from HS1.

5. Towards an alternative assessment of phytoplankton via biodiversity metrics

5.1 Analysing biodiversity in monitoring data

Biodiversity assessments need a multivariate approach as no single variable captures even the most important aspects of community composition and change (Rombouts et al. 2019). Under the MSFD, biodiversity is considered one of the descriptors to assess the environmental quality status of an area. An ecosystem rich in biodiversity has environmental conditions that allow many different species to coexist, which increases community resilience against environmental change and extreme events, as well as the capability of ecosystems to provide nature's contributions to people (IPBES 2019). For the MSFD requirements, Rombouts et al. (2019) recommended a multivariate approach, which we in general follow with some modifications to reflect recent findings on statistical performance of these metrics (Chase & Knight 2013, Hillebrand et al. 2018).

In essence, our approach consists of a 2 x 2 combination of different assessment goals: First, we want to measure both the gross and net component of biodiversity change. Gross means that between time points, the composition changes by new species arriving (colonizations) or disappearing (extinctions) and species becoming rare or dominant. If colonizations = extinctions, the net outcome of this overall compositional change will be neutral, but if one prevails over the other, standing diversity increases or decreases. Thus, in line with recommendations for MSFD and the OSPAR indicator "Changes in Plankton Diversity" (PH3) (Rombouts et al. 2019), we combine the assessment of alpha (= standing) diversity with analyses of temporal beta (turnover) diversity.

Second, both net and gross biodiversity change can be measured based on species' presence and absence or based on species' relative proportion in the community. The latter mainly reflects the dominant species in a community, the former the change in the many rare species.
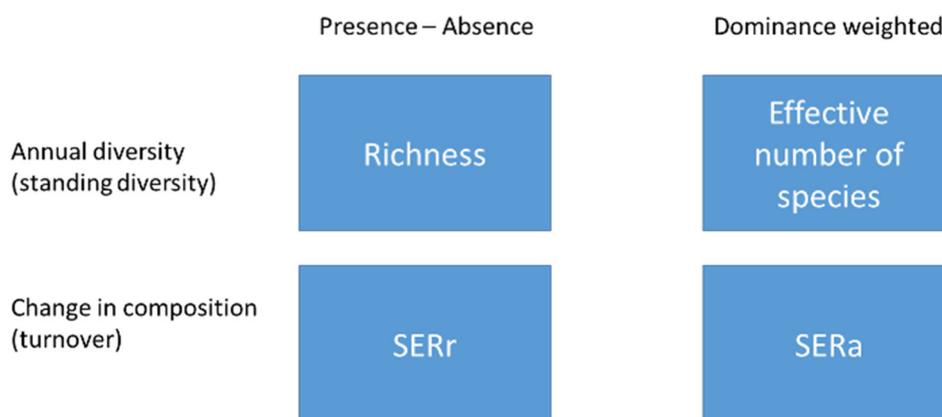
Figure 5.1. Biodiversity metrics analysed in this study.

This leads to the following metrics:

- Annual richness (S), the number of species occurring in a single year. We use S despite its use being limited by the fact that it is highly effort-dependent, i.e., more samples in a year, higher abundance of phytoplankton in a sample, more even dominance of species in a sample, more

complete assessment of a sample all alters the estimate for S substantially (Chase & Knight 2013, Hillebrand et al. 2018). However, it treats rare and dominant species equally and thus especially reflects changes in the presence or absence of rare species. Moreover, S is easy communicable as the number of species present is a metric not requiring any further explanation.

- As dominant-weighted measure of standing diversity, we used **ENS**, the effective number of species. It is related to the probability of interspecific encounters (PIE, the likelihood that two random individuals belong to different species). PIE is an entropy and related to the Hurlbert metric proposed by Rombouts et al. (2019), but ENS has been shown to be the most robust metric of sampling and abundances (Chase & Knight 2013). ENS equals the number of species you would encounter in an assemblage having the same entropy (PIE) but if all species were equally abundant. It can be envisioned as the number of species effectively taking part in the community. This analysis was based on the median biomass per year, i.e, all species occurring at least once during a year contributed to annual richness and ENS.

We follow Rombouts et al. (2019) in promoting that the analysis of standing diversity needs to be amended by an analysis of temporal turnover in community composition. However, we differentiate from Rombouts by pinpointing towards the fact that – in order to compare the gross to net changes in biodiversity – the employed metrics should weigh dominance in the same two different ways as S and ENS do. These measures thus correspond to richness and ENS, but in contrast to these measures of "annual diversity", SERr and SERa capture the difference in diversity between years.

- **SERr,** richness-based species exchange ratio, is a metric relying on presence and absence of species (as richness). It is identical to Jaccard dissimilarity that is often used in a spatial context and captures the proportion of the joint species from two time points that are NOT shared. Thus, if all species of time A also occur at time B, SERr = 0, if half of the species occur at one time point only, SERr = 0.5, and if time B has no species in common with time A, SERr = 1.0.
- **SERa** is the abundance-based species exchange ratio, thus the more dominant a species is, the more it influences turnover by going from rare (or absent) to dominant or from dominant to rare (or extinct). SERa also range from 0-1, with 1 = all dominant species exchanged. Like ENS, SERa weights dominance based on Simpson dominance (Hillebrand et al. 2018).

Both SERr and SERa turnover can be used for two different purposes: First, we measured **annual (immediate) turnover**, which compares consecutive years and thus reflects whether turnover from one year to the next becomes faster or slower with time. Second, we used **cumulative turnover**, which compares all samples to each other and relates this to temporal distance between the years, it thus reflects whether changes in composition continue (linear relationship between cumulative turnover and distance) or whether previous assemblages are found again (non-linear relationships returning to lower SERr or SERa at the end).
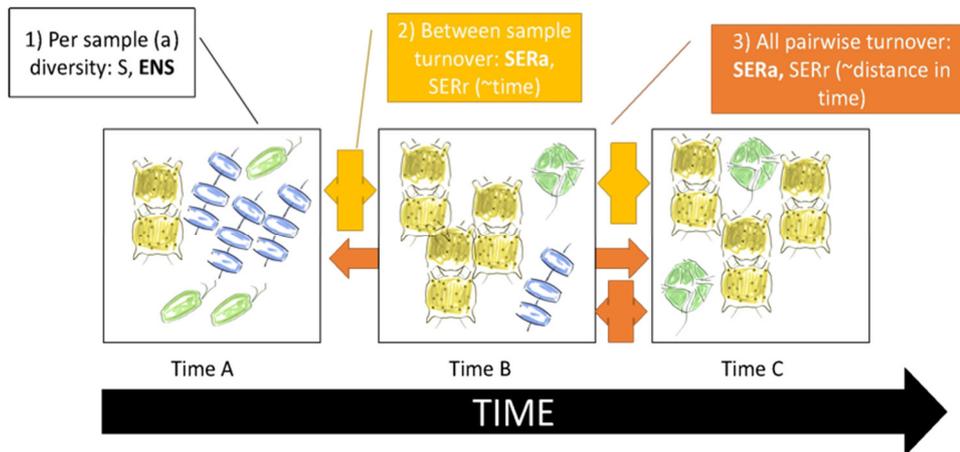
Figure 5.2. A schematic representation of how standing diversity (S and ENS) and turnover (SERa and SERr) are measured over time.

5.2 Phytoplankton diversity in the Wadden Sea

a) Phytoplankton biodiversity over time

Both metrics of standing diversity, richness (S) and effective number of species (ENS), decline with time across stations (Table 5.1), but this negative trend is only associated to the Dutch stations, which in general report more species, but lower ENS. However, this difference by country may be an artefact from German stations being sampled over a shorter period: in many stations across countries, richness and ENS peaks around 2007 to 2009, then declines until 2012, after which diversity increases again, also in the Dutch stations (Fig. 5.3a, b). However, a few stations (HUIBGOT and GROOTGND) show monotonic richness declines. Annual ENS varies across stations similarly to richness, but also shows an increase in most of the stations after 2012 (Fig. 5.3c, d).

Turnover between neighbouring years is variable over time but does neither speed up nor slow down (Fig. 5.4a-d, Table 5.1). When accumulating over time, a clear pattern of increasing turnover becomes visible (Fig. 5.4e-h). Turnover is larger in the German stations and consistently increases with increasing temporal distance for both *richness-based species exchange ratio* (SERr) and *abundance-based species exchange ratio* (SERa). The consistency of the pattern indicates that turnover does involve both shifts in species identity and dominance. It is also noteworthy, that the down right corner of the diagram is void of data, indicating that there is no "return" to a previous assemblage over long time scales, indicating a strongly directional compositional drift over time.
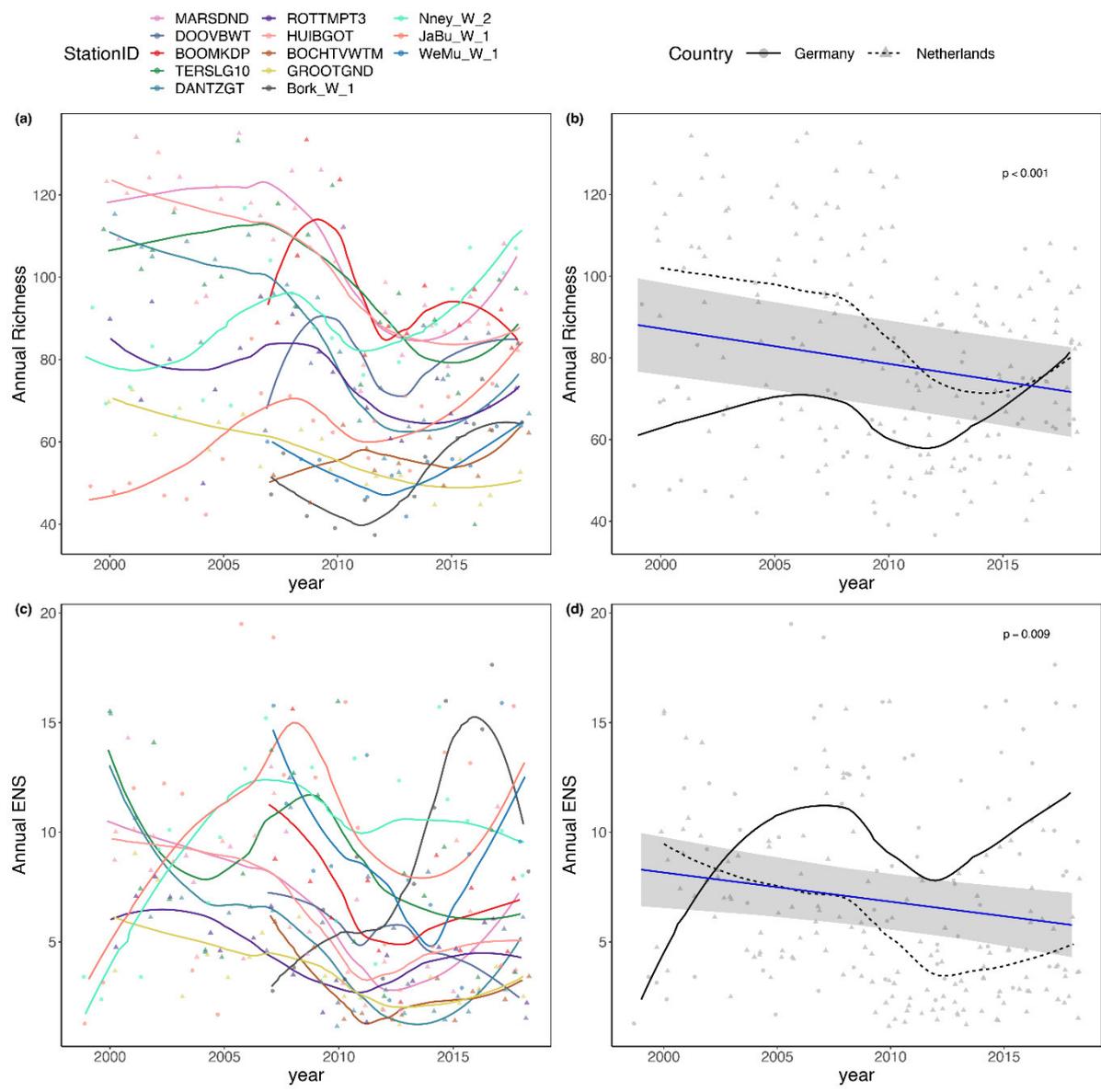
Figure 5.3. Temporal trend of the phytoplankton standing diversity: richness (a, b) and effective number of species (c,d) at the Wadden Sea coastal stations. Left column: Annual means and LOESS trend lines colored by station. Right columns: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous line; NL: dashed line). *Data are LN transformed.*
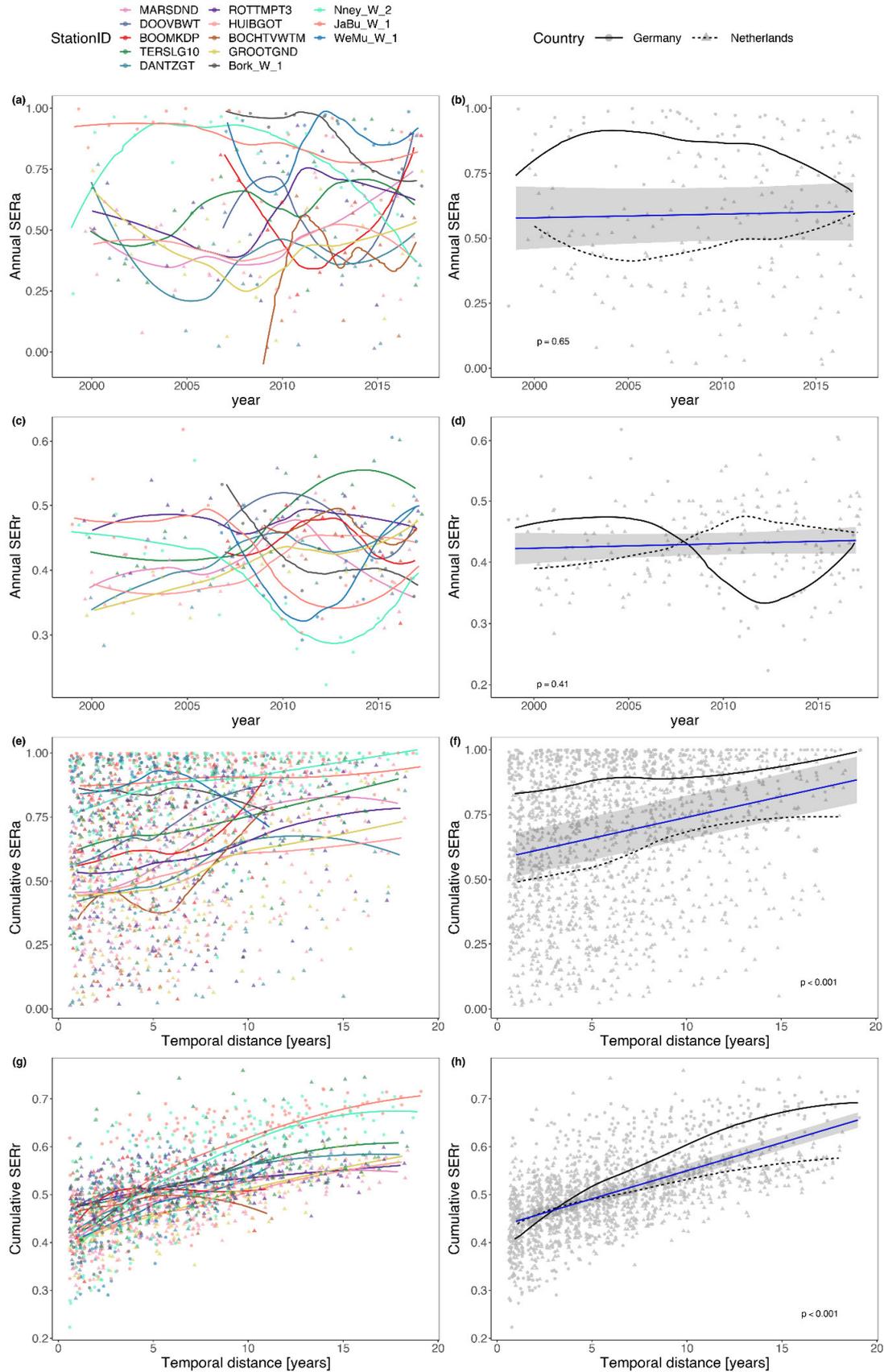
Figure 5.4. Temporal trend of the phytoplankton turnover (SERa and SERr) at the Wadden Sea coastal stations. Left column: Annual means and LOESS trend lines coloured by station. Right column: Overall predicted time effects from the LMM (blue line) with their confidence interval (grey shaded area) as well as separate LOESS trends for German and Dutch stations (DE: continuous line; NL: dashed line).

Table 5.1. Results of the linear mixed effect model, analysing the change in phytoplankton diversity and turnover over time (year) and the cumulative turnover between years (dist). Station ID is included as random effect. All details as in Table 2. Please note that for the cumulative turnover, the predictor is not year but temporal distance in years.

	Annual Richness		Annual ENS		SERa		SERr	
Predictors	Estimates	p	Estimates	p	Estimates	p	Estimates	p
(Intercept)	1815.494	**<0.001**	273.709	**0.007**	-2.271	0.721	-1.073	0.556
year	-0.864	**<0.001**	-0.133	**0.009**	0.001	0.653	0.001	0.410
Random Effects								
σ^2	207.12		13.94		0.05		0.00	
τ_{00}	360.31 StationID		4.32 StationID		0.03 StationID		0.00 StationID	
ICC	0.63		0.24		0.40		0.16	
N	13 StationID		13 StationID		13 StationID		13 StationID	
Observations	213		213		199		199	
Marginal R^2 / Conditional R^2	0.036 / 0.648		0.027 / 0.257		0.001 / 0.405		0.003 / 0.162	

	Cumulative SERa		Cumulative SERr	
Predictors	Estimates	p	Estimates	p
(Intercept)	0.579	**<0.001**	0.433	**<0.001**
Dist.	0.016	**<0.001**	0.012	**<0.001**
Random Effects				
σ^2	0.05		0.00	
τ_{00}	0.02 StationID		0.00 StationID	
ICC	0.34		0.15	
N	13 StationID		13 StationID	
Observations	1725		1725	
Marginal R^2 / Conditional R^2	0.064 / 0.380		0.382 / 0.472	

b) The impact of environmental factors on phytoplankton biodiversity

Analysis of biodiversity is scale and effort dependent, which prevents the use of absolute or threshold values. Still, the diversity trends with time were conclusive overall and their relation to environmental drivers was strong, as these explained 13-32% of the variance in biodiversity alone, which increased to up to 80% by including the random terms (Table 5.2). In the SEM, we found 44% of variance in S and 57% in ENS explained by the 6 abiotic variables. Thus, these multidimensional aspects offer a wide range of conclusions on the ongoing changes in biodiversity. As described above (Fig. 5.3), standing diversity (richness S and effective number of species ENS) declined in the Wadden Sea over the monitoring period analysed (2000 – 2019). This decline was mainly driven by the Dutch stations. Moreover, S and ENS were positively correlated (Fig. 2.9, chapter 2, r = 0.319 for DE and r = 0.607 for NL).

Nutrients were strongly negatively correlated to diversity. S and ENS consistently declined with Si within and across countries, independent whether analysed as correlations (Fig. 2.9, Chapter 2), LMM (Table 5.2, Fig. 5.5) or in the SEM (here only affecting S negatively, but ENS positively, Fig. 2.12-2.13, Chapter 2). Likewise, increasing TN was associated to lower S (correlation) and ENS (correlation, SEM), similar relationships appeared with TP but only in the correlations. The fact that TN and TP turned non-significant as predictors in the LMM is again due to the high correlation to Si and each other. According to LMM, S further declines with increasing N:P ratios, thus in P-limited situations (overall and in the Dutch stations).

Both S and ENS decrease with SPM (overall and in the Dutch stations, both correlations and LMM) and increase with salinity and pH in LMM and correlations (Fig. 2.9, Chapter 2 and Fig. 5.5, Table 5.2), but not in the SEM which only found a negative effect of salinity on S. By contrast, temperature effects were inconsistent as both ENS and S increased with temperature in the SEM but decreased in the correlations.

The SEM also finds biomass to be affected by diversity, with higher ENS associated to higher Chl biomass, but lower C-biomass, whereas higher species richness is connected to higher C-biomass (Figs. 2.12-2.13, Chapter 2.2.4).

Turnover accumulated over time, indicating a strong compositional drift over time (see Fig. 5.4g-h). These changes reflect that species composition continues to change even if some conditions are restored (such as reduction in eutrophication) as other ongoing changes (such as warming) and species immigration prevent compositional recovery.
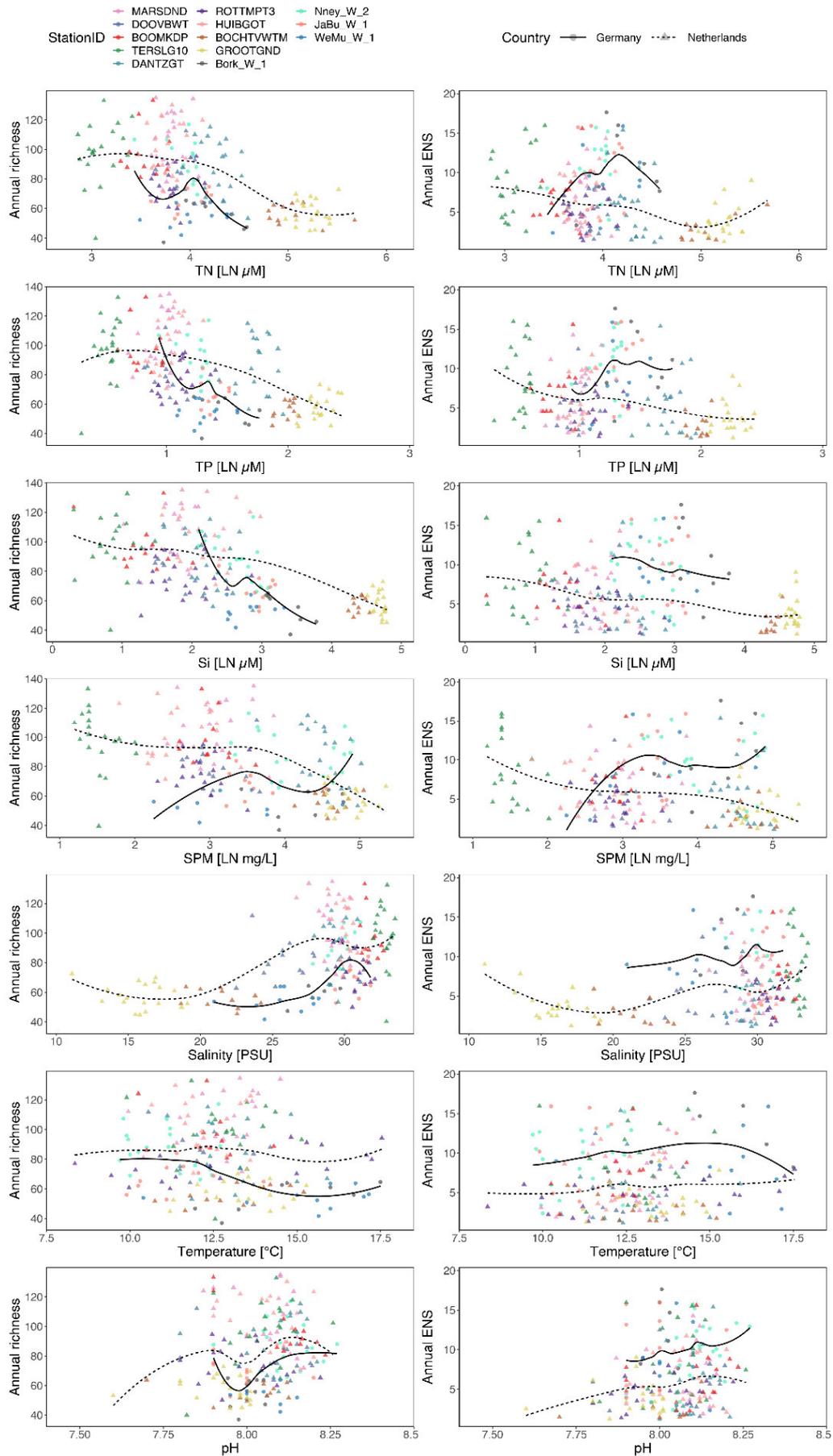
Figure 5.5. Standing diversity in relation to environmental parameters. *Data input: annual median.*

Table 5.2: Results of the linear mixed effect model, analysing the effects of environmental factors on phytoplankton standing diversity (annual richness and ENS), considering "station" as a random effect. When outputs differed between NL and DE, we highlighted the estimates in grey. Data input: *annual median.*

	Annual Richness (S)						Annual ENS					
	all		*NL*		*DE*		*all*		*NL*		*DE*	
Predictors	Estimates	p	Estimates	p	Estimates	p	Estimates	p	Estimates	p	Estimates	p
(Intercept)	0.180	0.454	0.474	0.043	-0.668	0.076	0.046	0.847	-0.173	0.363	0.794	0.236
TP	0.489	0.161	0.554	0.153	0.281	0.698	0.187	0.669	0.450	0.287	1.578	0.405
NP	-0.195	0.020	-0.223	0.018	0.115	0.558	-0.002	0.986	0.007	0.950	0.521	0.326
Si	-0.855	0.005	-0.972	0.002	-0.876	0.205	-0.885	0.011	-0.658	0.033	-2.270	0.219
SPM	-0.263	0.015	-0.381	0.011	0.020	0.860	0.076	0.577	-0.320	0.055	0.373	0.190
Salinity	-0.271	0.227	-0.451	0.087	0.745	0.051	-0.558	0.043	-0.494	0.069	0.410	0.521
Temp.	0.001	0.993	-0.034	0.689	-0.047	0.657	0.154	0.085	0.131	0.175	0.195	0.366
pH	-0.174	0.013	-0.178	0.026	-0.291	0.026	0.122	0.175	0.075	0.397	-0.030	0.928
Random Effects												
σ^2	0.36		0.40		0.12		0.60		0.52		0.94	
τ_{00}	0.58 StationID		0.35 StationID		0.31 StationID		0.45 StationID		0.17 StationID		0.27 StationID	
ICC	0.62		0.47		0.72		0.43		0.25		0.23	
N	13 StationID		9 StationID		4 StationID		13 StationID		9 StationID		4 StationID	
Obs	182		148		34		182		148		34	
Marg R² / Cond R²	0.223 / 0.703		0.315 / 0.637		0.283 / 0.802		0.133 / 0.505		0.143 / 0.356		0.210 / 0.389	

Summary: The diversity responses over time and with environmental factors were more consistent than the biomass response. Standing diversity declines with increasing nutrient availability, which emanated in different forms in the different statistical analysis. As diversity does not have an absolute scale, these conclusions could only be derived because of the extended time series. Thus, the information value of biodiversity lies in the trends, not in the absolute value. The temporal turnover indicates continued change in the phytoplankton composition and the absence of recovery of historic "pristine" communities.

6. Recommendations for future monitoring and assessment of phytoplankton

The two lead agencies providing the data, NLWKN and Rijkswaterstaat, are to be commended for the data-rich effort to monitor phytoplankton. Combining environmental, biomass and compositional data is key to a holistic view of the role of phytoplankton in the Wadden Sea and its response to potential risks for the good environmental status. Doing this across multiple stations and having offshore stations to compare to makes a very strong case for the assessment of phytoplankton.

While it would be helpful to have binary criteria for the status of the ecosystems, this is hardly possible. First, the monitoring datasets comprise a time period with massive environmental change, but mostly after the peak eutrophication in the 1980s. Thus, there is no information on a "pristine" status and the Wadden Sea has still higher concentrations of N and P than the North Sea despite reduced nutrient concentrations over time. Second, most changes over time and most relationships between environmental variables and phytoplankton descriptors are gradual. Therefore, fixed "threshold" values for biodiversity or biomass do not emerge from the analysis. Instead, gradual changes in drivers lead to gradual responses. Third, even if nutrients would be further reduced, the concomitant changes in other environmental factors such as temperature will not lead to a recovery of a previous species assemblage. The increase of cumulative turnover with time did not show any sign of return to a previous assemblage, which reflects that while nutrients decline, other factors change as well and lead to further changes in the composition.

In the light of these caveats, we recommend the following approaches based on a comparison of the reliability of relationships. We created a reliability metric ranging from -1 (clear consistent negative association) over 0 (no association) to +1 (fully consistent positive relationship). We derived this metric by a weighted vote count using the correlation coefficient (-1 to +1) or the path coefficients of the SEM as weights. For the LMM, we used -1 and +1 if the estimated was significant at $p<0.05$, ± 0.5 for $p< 0.1$, 0.25 for $p<0.2$ and 0.1 for $p<0.5$. Neutral relationships with $p>0.5$ in the LMM or no support in SEM were coded as 0.

Based on this assessment, the relationship between nutrients and Chl as well as SPM and biomass (both C and Chl) were the most predictable positive associations (Fig. 6.1). Negative effects of Si, Salinity and S on Chl biomass were weak or variable. Diversity (especially S) was negatively related to SPM and Si. pH and temperature had positive effects on ENS.
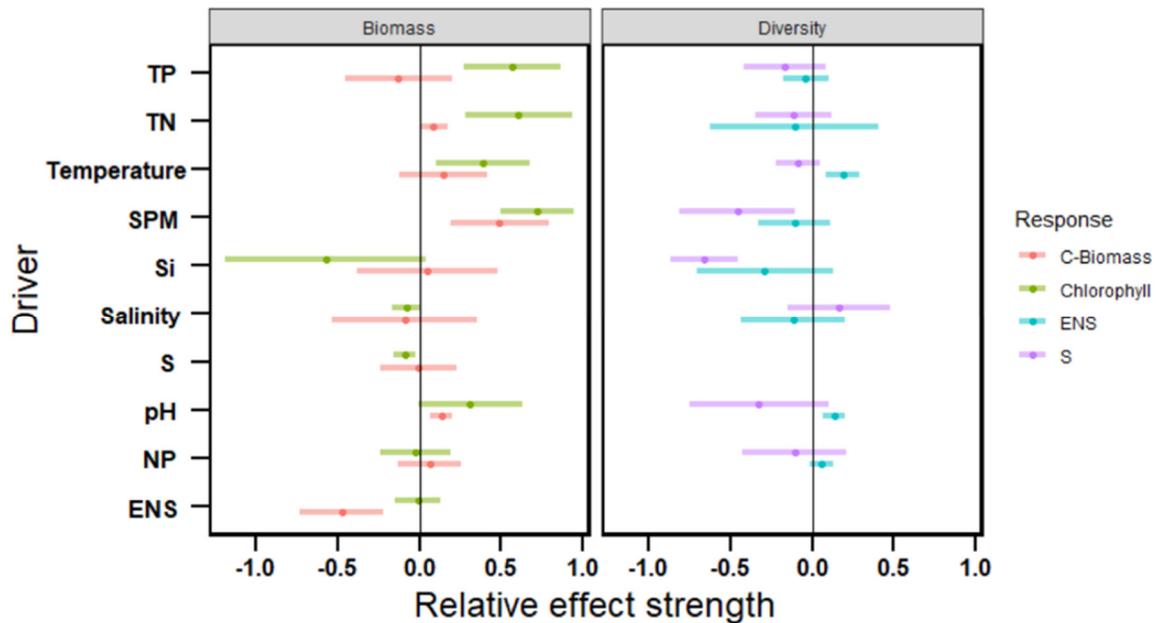
Figure 6.1. Relative strength of association (mean ± confidence interval) for drivers and responses (with different colours representing different response variable).

Instead of relying on thresholds for assessment, a focus on temporal trends seems more appropriate for this highly dynamics system. Amending the current analyses with new incoming data will allow a clearer picture of the ongoing developments. Less eutrophied situations are clearly linked to lower biomass and higher standing diversity of phytoplankton, the latter even more consistently related to nutrients (Figure 6.1). We especially consider the structural equation model (SEM) a very strong approach as it takes advantage of mechanistically well proven relationships including the diversity-biomass link. But also, a continued time series analysis can be used to inform the assessment. One can also use the information gathered so far to establish "internal baselines": Over the stations, a clear low biomass – high biodiversity optimum appears at TN = 25-30 µM and TP = 1.5 µM. Whether a further reduction in nutrients (towards 2.8 mg N/L from riverine sources) allows further declines in phytoplankton biomass and increases in diversity can only be observed if further actions are taken.

Summary: The analysis of phytoplankton biomass, diversity, and species composition are important tools for understanding the aquatic system and the environmental conditions. However, they do not serve as caveats to define "thresholds" values of a good or bad water quality status. The results of these analyses may reflect gradual changes occurring in the environment. In a dynamic system such as the Wadden Sea, the assessment of water quality should be based on temporal trends, which can only be analysed with continued monitoring programs. These programs should include not only regular and consistent phytoplankton sampling but also measurement of cell size (biovolume) and carbon content.

Recommendations for improvement

We include a few recommendations to further develop the assessment

1. Include light measurements: In the highly turbid waters of the Wadden Sea, light will potentially be frequently limiting algal growth and nutrient uptake. Including this information would strongly strengthen our ability to discuss limitation, understanding compositional change and derive mechanistic understanding of system behaviour. Light could be measured by optical sensors and should include surface values as well as values at 2-3 depth (e.g., 0.5, 1 and 2 m).

2. Align zooplankton measurements with phytoplankton measurements: NLWKN started 5 years ago to amend the phytoplankton time series with a zooplankton analysis. This is highly important given that without such information no change in top-down control of algal biomass can be detected. We thus recommend extending the analysis of zooplankton also to the Dutch stations. Moreover, we observed that zooplankton sampling occurred at different stations and different days than the sampling of the phytoplankton. This discrepancy should be remedied.

3. Publish harmonized data: This project has made major efforts to harmonize the data sets, which has resulted in the first cross-country high-resolution data set. We strongly advise to publish this status, preferably in a data repository which allows version control and thus continuous updating of the data. This has advantages for both assessment and science. As new data analysis methods emerge much faster than the agencies can adapt their approaches, open data access will allow leveraging the efforts of scientists worldwide that will use these data to calibrate their methods and propose advances. The assessment can directly profit from these insights as the data set is cited and its use can be tracked. For science, this does not only open a great data resource, but also avoids reinventing the wheel by starting from disparate excel files.

4. Cross-check taxonomy: We were unable to resolve some discrepancies in the composition of the NL and DE data. We recommend to actually exchange a few samples between laboratories in order to check their take on the species inventory. This will potentially enable further harmonization as the current small overlap in the taxon lists indicates some major discrepancies between the monitoring programs.

5. Bioassays: We recommend a pilot project with simple nutrient bioassays to check for the preponderance of nutrient limitation. In its easiest version, it would require additions of N, P and N+P to phytoplankton samples obtained during the normal monitoring campaigns for at least a subset of stations, measuring Chl after 24 hours. More comprising manipulations (including Si and light) are possible.

6. Measuring cell size from samples: Phytoplankton cell size is an important trait that can provide insights on different morphological and physiological aspects of species and can be related to environmental changes and grazing (Hillebrand et al. 2022). Cell size analysis of the German Wadden Sea phytoplankton revealed that species are 30% smaller now than 15 years ago (Hillebrand et al. 2021). This and further analyses can only be done when cell sizes are measured per sample. Based on these findings, we highlight the importance of measuring the cells from/in the samples instead of using standardised literature values, which are often overestimated and do not capture temporal changes.

7. Particulate organic carbon and pigments: We recommend amending the current sampling with two additional analyses. From the same sample that serves as basis for counting and Chl, two further subsamples shall be taken for a relevant number of stations and times. It would be sufficient to use 2-3 stations each in the Netherlands and in Germany for ca. 1 year. The first subsample shall be filtered on GF/F filters for measurement in a CN analyser (the additional

measurement of N is included and helpful for the limitation question) giving an independent total C measurement. These C-values can be compared to Chl and microscopy-based C estimates to identify congruence and discrepancies. The second subsample shall be used for a newly developed inexpensive way of estimating several pigments using photometers (Thrane et al. 2015). These could potentially suffice to identify major algal taxa, which could be compared to the counted data. The method is also comparable to the pigment-based analyses taken by ferry boxes and other high frequency sampling.

8. Continuation of the phytoplankton monitoring program: Our analysis suggests that the temporal trends of biotic and abiotic factors provide important information about the aquatic ecosystem. Therefore, a continuous and consistent monitoring program is beneficial for comprehending changes in environmental conditions and designing better water management plans.

7. Synthesis

7.1 What did the project achieve

This project started, as a follow-up on the intercalibration process of the Water Framework Directive (WFD), with the objective to address remaining issues related to the assessment of phytoplankton in the Dutch and German Wadden Sea. The project aims were:

- To gain a more comprehensive understanding of the complex system of the Wadden Sea and to gain a holistic picture on phytoplankton conditions and eutrophication status;
- To realise an innovative, multi-causal research approach by considering different parameters concerning phytoplankton and eutrophication in a cross-national ecosystem modelling;
- To solve shortcomings of the intercalibration process with a stepwise approach using different ecosystem models and a comprehensive analysis of long-term monitoring data;
- To provide a reliable scientific background as a basis to harmonise the phytoplankton assessment in the Wadden Sea;
- To test several phytoplankton and eutrophication related parameters as possible criteria for phytoplankton assessment, including chlorophyll a (Chl);
- To test different scenarios of nutrient reduction and their effects on phytoplankton biomass and the eutrophication status in the Wadden Sea;
- To build a bridge between scientific analysis and operational management tools;
- To strengthen the bilateral cooperation between German and Dutch authorities and research institutions and to combine competence and expertise of all project partners to finally promote a common understanding of eutrophication in the Wadden Sea as a cross-border issue;
- Finally: to propose reference conditions and common assessment levels for phytoplankton in the coastal water bodies NEA 3/4 and NEA 1/26.

To achieve these objectives, two approaches were followed to improve the harmonised assessment of the biological quality element Phytoplankton, as part of the assessment of the ecological status of WFD coastal water bodies in the Dutch and German Wadden Sea.

One approach was based on a detailed analysis of monitoring data of phytoplankton and environmental variables from Dutch and German monitoring programs. This analysis focused on a better understanding of the factors determining phytoplankton biomass and composition, in order to develop further methods for the assessment of composition of the phytoplankton community and abundance of phytoplankton taxa.

The second approach was based on the application of ecosystem models, to improve our systemic understanding of the relation between anthropogenic nutrient loads to the Wadden Sea and Chl concentrations as a proxy for phytoplankton biomass, including the effects of other environmental factors determining phytoplankton growth. The application of models made it possible to study the effects of different levels of nutrient loads on the Chl concentrations in the Wadden Sea and finally to provide a common scientific understanding as a basis for a harmonised definition of thresholds for the Chl assessment.

During this project, the intense exchange between modellers and phytoplankton experts from HIFMB made it possible to closely interlink the model development and improvement with analysed observations and literature data on phytoplankton as well as with newest results of the monitoring data analyses.

7.2 The assessment of composition and abundance of phytoplankton taxa via monitoring data

The Wadden Sea is a multi-pressure system that is nowadays affected by many overlapping stressors. It has changed dramatically over the last 50 years, as the monitoring data collected by Rijkswaterstaat (RWS) and the Lower Saxony Water Management, Coastal and Nature Protection Agency (NLWKN) show. A change in nutrient concentrations in an order of magnitude, an exceptionally strong warming and massive changes in the biomass, cell size and biodiversity of phytoplankton are clearly visible during this period.

The multifaceted plankton analyses have shown that monitoring data series are a good tool for detecting these long-term changes, both in environmental conditions and in phytoplankton parameters. Although changes in different plankton parameters are closely related to changes in environmental conditions, it is difficult to depict the one factor determining phytoplankton biomass, since nowadays multiple stressors affect the Wadden Sea ecosystem simultaneously and interact with each other (changes in nutrient pressures and ratios, climate change, interannual variability in growth conditions and other human pressures). Along with environmental conditions, phytoplankton communities are in a continuous state of change. The lack of a stable reference status makes it therefore hardly possible to identify fixed threshold values for different plankton parameters describing the quality status of the phytoplankton community, although trends can be identified via biodiversity parameters that describe phytoplankton dynamics (such as species richness, effective number of species, turnover, etc.).

The data analyses showed that phytoplankton biomass generally increases with increasing nutrient concentration, while phytoplankton biodiversity decreases. Hereby nitrogen as well as phosphorus and silicate play a role. There are clear indications that, especially in the spring and early summer, phosphorus limitation is predominant in the southern Wadden Sea, while lower nitrogen concentrations in late summer indicate at least potential nitrogen limitation in the late growing season.

But the observed changes in phytoplankton are not only due to a response to the peak of eutrophication in the 1980s and the subsequent recovery in the last decades with stronger decreases in P loadings than in N loadings and subsequent increases in N:P ratios. The changes in the phytoplankton community also reflect a continuous shift towards new communities. Consequently, it was neither possible to identify phytoplankton species nor single metrics describing phytoplankton diversity that can be used as indicators for nutrient enrichment.

The current threshold values of the existing phytoplankton assessment are based on the assumption that Chl, as an indicator of phytoplankton biomass, correlates well with nitrogen concentration. However, the results of this project show, both in the monitoring data as well as in the modelling results, that Chl does not linearly react to nitrogen load reductions. It could therefore be assumed that a further significant reduction in phytoplankton biomass cannot be achieved with a reduction in nitrogen alone, which is the main focus of the WFD measures also with the intention to align with the aim of restoring lower N:P ratios. Since it has been shown in this project that phosphorus limitation also plays a major role in phytoplankton growth in the Wadden Sea, efforts to regulate phytoplankton biomass by nutrient reduction must therefore go beyond the reduction of nitrogen alone. In addition to nutrient limitation, phytoplankton growth is also regulated by light availability and grazing pressure, two factors that should be considered in monitoring programs and future assessments.

Moreover, the response time of phytoplankton to reductions in nutrient input (after complete implementation of measures) and to achieve good ecological status is assumed to be 10-15 years in the Wadden Sea (LAWA 2019), which has to be taken into account when evaluating the efficiency of

measures, the effect they have on the marine environment and the status of quality elements such as phytoplankton.

Conclusion and recommendations

For the assessment of phytoplankton and the state of eutrophication, a holistic concept consisting of several different parameters and ecosystem elements is therefore proposed, that goes beyond the assessment of a single parameter alone. This would better reflect the complexity of both, the phytoplankton community and the dimension of eutrophication in the context of the ecological quality of an ecosystem, than the sole assessment via the biomass parameter Chl.

While it would be helpful to have binary criteria for the status of the ecosystem, this is hardly possible. First, the monitoring datasets comprise a time period with massive environmental change, but mostly after the peak eutrophication in the 1980s. Thus, there is no information on a "pristine" status and the Wadden Sea has still higher concentrations of N and P than the North Sea despite reduced nutrient concentrations over time. Second, most changes over time and most relationships between environmental variables and phytoplankton descriptors are gradual. Therefore, fixed "threshold" values for biodiversity or biomass do not emerge from the analysis. Instead, gradual changes in drivers lead to gradual responses. Third, even if nutrients would be further reduced, the concomitant changes in other environmental factors such as temperature will not lead to a recovery of a previous species assemblage. The increase of cumulative turnover with time did not show any sign of return to a previous assemblage, which reflects that while nutrients decline, other factors change as well and lead to further changes in the composition.

Instead of relying on thresholds for assessment, a focus on temporal trends seems more appropriate for this highly dynamics system. Amending the current analyses with new incoming data will allow a clearer picture of the ongoing developments. Less eutrophied situations are clearly linked to lower biomass and higher standing diversity of phytoplankton, the latter even more consistently related to nutrients. One can also use the information gathered so far to establish "internal baselines": over the stations, a clear low biomass – high biodiversity optimum for instance appears at TN = 25-30 µM and TP = 1.5 µM. Whether a further reduction in nutrients allows further declines in phytoplankton biomass and increases in diversity can only be observed if further actions are taken.

7.3 The assessment of chlorophyll via ecosystem modelling

One of the objectives of the project was to investigate the effects of nutrient reduction scenarios on phytoplankton biomass and to propose harmonised reference conditions and assessment levels (thresholds) for Chl. To fulfil these objectives, we followed two different modelling approaches in this project.

First, we conducted a nutrient reduction scenario, based on the existing management objective for total nitrogen (TN) of 2.8 mg/l at the limnic/marine border of German and Dutch rivers. The assumption was that, if rivers comply to this threshold concentration, good ecological status in coastal waters in accordance with the WFD will be achieved for the biological quality element phytoplankton.

In this case, the reduction requirement results from the difference between today's mean TN concentration in the rivers (based on the discharged loads) and the targeted average concentration of 2.8 mg TN/l, where both Germany and the Netherlands agreed on as a compliant scenario setup. The loads at an average of 2.8 mg TN/l then form the basis for the modelling of nutrient and Chl

concentrations in the coastal regions of the Wadden Sea. Since there is no textbook method for this calculation, possible variations in the calculation methodology were also considered. For this purpose, the standard "2.8-scenario" as described in Appendix 1, in which runoff is not explicitly taken into account, was supplemented by an additional calculation (Appendix 2), in which the daily discharge is included in the calculation. The second scenario served as a sensitivity test and was only applied by the GPM-model.

To explore the relations between nutrient loads and Chl concentrations in the Wadden Sea and adjacent coastal waters, the models simulated the current state (year 2017), with the scenario with reduced riverine TN loads that comply with the current management objective for TN of 2.8 mg/l ("2.8-scenario"). Two models covering the southern North Sea and the entire Wadden Sea (the Deltares model 3D DCSM-FM and the University of Hamburg model SNS-GPM) as well as one local model covering the Ems estuary and its adjacent Wadden Sea with significantly higher resolution (the Ems model of the Forschungsstelle Küste (FSK)) were used.

Second, the ecosystem models were applied to derive joint thresholds for Chl in the Dutch and German coastal water bodies based on the modelling of "pre-eutrophic" reference conditions in accordance with the method of threshold derivation applied in OSPAR for the wider North Sea. In this approach, two historic scenarios (HS1 & HS2), representing reference Chl concentrations in pre-eutrophic times, were modelled. As a first approach to align Chl assessment levels for coastal and offshore waters in the North Sea, thresholds were then calculated according to the standard method used so far within OSPAR and the WFD as "Chl reference-value + 50% = Chl threshold value".

7.3.1 Model application to investigate the effect of nitrogen reductions ("2.8-Scenario")

The "2.8-scenario" was applied to get a model estimate of the effect of reducing the riverine nitrogen loads to a level where all rivers would comply with the Dutch and German management objective of an annual average TN concentration of 2.8 mg/l. The reductions in TN loads from the rivers (compared to current loads) were different for each river, as in some rivers, TN concentrations are already close to the 2.8 mg/l target, while other rivers still have much higher TN concentrations. Consequently, the 2.8-scenario showed the strongest reduction in TN loads in the rivers Ems and Eider and only small reductions in the rivers like the Rhine (Nieuwe Waterweg) and the Maas (Haringvliet) (Tab. 7.1).

Table 7.1. Overview on TN loads and reductions for the 2.8-scenario for major Dutch and German rivers (selected from Tab. A1.1 in Annex 1).

Country	River	TN Load (kton)	TN Reduction (%)
DE	Eider	4	39 %
DE	Elbe	98	16 %
DE	Ems	16	38 %
DE	Weser	38	23 %
NL	Haringvliet	53	4 %
NL	Nieuwe Waterweg	128	1 %
NL	Noordzeekanal	6	0 %
NL	Scheldt	17	25 %

Reducing TN loads leads to proportional decreases in winter mean DIN concentrations in the Wadden Sea. But the reduction in Chl concentrations is much more limited. Causes for the low response of Chl are the role of other co-limiting factors for phytoplankton growth, such as light limitation and P-limitation. The resulting Chl concentrations are mostly still above the current threshold values of the WFD, as the overview Table 7.2 clearly shows (note that assessment periods of the latest WFD assessment of the two countries do not coincide). The limited effect of the TN load reduction on the Chl concentration in the 2.8-scenario may indicate that achieving annual average TN concentrations of 2.8 mg/l in the rivers will not be enough to meet the WFD thresholds for Chl in the Wadden Sea, as the reduction in Chl that is predicted by the models (Table 3.1) is far less than the required >50% reduction estimated from observations. However, it is assumed that it requires at least 10-15 years until the reduction of TN loads to 2.8 mg/l in the rivers leads to the desired Chl reduction in coastal waters (LAWA 2019).

Taking into account the model uncertainties, these results provide an indication of how effective the existing management objective is and whether the required nutrient reductions in the rivers (in accordance with this value) will lead to achieving good ecological status according to the WFD for the indicator Chl in the coastal waters. The analyses of the monitoring data also indicate that Chl does not react linearly to nitrogen reductions and that nitrogen is not the sole factor determining phytoplankton biomass in the coastal waters of the Wadden Sea. It is important to note, that nutrient loads (TN and DIN) of all rivers from other countries discharging into the North Sea were kept at their current level, therefore the effect of nutrient reduction measures in other countries on the transboundary nutrient transport into the Wadden Sea area is not included here. TP and DIP river loads as well as atmospheric nitrogen deposition in the entire model domain, including the Dutch and German coastal region, were also kept at current levels.

Table. 7.2. Overview of eutrophication status of DE and NL coastal water bodies related to the 2.8-scenario for chlorophyll, aggregated from Tab. 1.1 (threshold values) and Tab. 3.1 (reduction level of simulation runs) in the DCSM-FM (Deltares) and SNS-GPM (University of Hamburg) model and the interpretation of results from Fig. 3.7 – 3.10. **+** indicates resulting chlorophyll concentration above the threshold, **-** indicates resulting chlorophyll concentration below the threshold. Reduction is calculated as %reduction=(Concentration$_{2.8}$-Concentration$_{CS}$)/ Concentration$_{CS}$*100%. Colours indicate present status of phytoplankton in NL and DE water bodies: green=good, yellow=moderate, orange=poor, red=bad. DE assessment results are shown for 2014-2018, NL assessment results are shown for 2018-2020.

WFD Typology	WFD Water Body	Current WFD Chl threshold (µg/l)	Present status	DCSM-FM (% red.)	DCSM-FM (status vs. threshold)	SNS-GPM (% red.)	SNS-GPM (status vs. threshold)
NEA 1/2	NL Wadden Coast N1	7.5		-1 %	**-**	-2 %	**+**
NEA 3/4	NL Wadden Sea N4	7.2		-2 %	**+**	-2 %	**+**
NEA 1/2	DE Ems N1	3.8		0 %	**+**	-10 %	**+**
NEA 1/2	DE Ems N2	3.8		0 %	**+**	-13 %	**+**
NEA 3/4	NL/DE Ems N3*	5.1 / 5.5		-2 %	**-/-**	-8%	**+/+**
NEA 3/4	DE Ems N4	5.5		-1 %	**+**	-10 %	**+**
NEA 1/2	DE Weser N1	3.8		0 %	**-**	-11 %	**+**
NEA 1/2	DE Weser N2	3.8		0 %	**+**	-9 %	**+**
NEA 3/4	DE Weser N3	5.5		0 %	**-**	-7 %	**+**
NEA 3/4	DE Weser N4-01	5.5		1 %	**+**	-5 %	**+**
NEA 3/4	DE Weser N4-02	5.5		-4 %	**+**	- 6%	**+**
NEA 3/4	DE Elbe N3	5.5		1 %	**-**	-3 %	**+**
NEA 3/4	DE Elbe N4-01	5.5		-5 %	**+**	-4 %	**+**
NEA 3/4	DE Elbe N4-02	5.5		1 %	**+**	- 2%	**+**

* For the Ems-N3 area, there are two Chl thresholds in use, which differ between the Dutch and the German part. For the Duch water body, characterised as "Ems-Dollard coast", the mean value of 5.1 µg/l is applied (good status), while for the German water body, indicated as "Ems-N3", a value of 5.5 µg/l is used (poor status).

7.3.2 Model application to derive harmonised thresholds

The Chl concentrations in the model runs of the historic scenarios HS1 and HS2 are considered to represent the pre-eutrophic reference conditions. Chl thresholds representing the Good/Moderate boundary of the WFD are derived from those reference concentrations by adding 50%. This method is commonly used in the WFD and was also applied to derive thresholds for marine waters by OSPAR (Lenhart et al. 2022). Table 7.3 shows the thresholds derived by the two models for both historic scenarios.

In the OSPAR exercise to determine threshold values, an ensemble of eight models contributed to the aggregated result. A final proposal for threshold values was derived from the individual model results by a weighted averaging method based on the deviation of the model results from observed concentrations (Lenhart et al. 2022, van Leeuwen et al. 2023). In addition, in the final decision on thresholds for Chl several corrections to the proposed threshold values were made by the countries involved; this applies to two of the assessment areas directly bordering the Wadden Sea and coastal waters, i.e. Ems plume and Elbe plume where the proposed Chl thresholds (from model results for

HS2) were reduced by 30% 'to ensure a plausible gradient with WFD areas' (OSPAR 2022, Annex 6). For the other assessment area, Rhine plume, thresholds were based on the model results for HS1.

Here, we only had the results of two models, which makes the weighted approach used by OSPAR less applicable. Simply averaging the two model results is not a very suitable method either, as in some areas the concentrations estimated by the models deviate considerably from observations, pointing at under- or overestimation of concentrations by the models. In those cases, thresholds derived from the reference concentrations could be unrealistically low or high.

An alternative method to make a correction for the deviation of model results from *in situ* values was used by Schernewski et al. (2015) to determine WFD thresholds for the German part of the Baltic Sea. The relative difference between historic (reference) model data and current state model data is combined with *in situ* concentrations to derive a corrected reference concentration:

$$Chl_{reference, corrected} = Chl_{in\ situ} * Chl_{reference, model} / Chl_{current\ state, model}$$

If the model values for current state are lower than observed concentrations, the reference concentration is adjusted with a factor >1, if model values for current state are higher than observations, the adjustment factor is <1.

This correction method has also been applied by the UK in the OSPAR assessment for some specific cases where it was called the 'relative method' (OSPAR 2022).

The threshold values after correction by this "Schernewski" approach, are also included in Table 7.3. The correction in many cases leads to higher threshold values. However, the correction is based on observed *in situ* concentrations. It should be taken into account that the quality of the correction is therefore highly dependent on the quality of the observations. The "Schernewski" correction can only be carried out if there are sufficient reliable and accurate monitoring data available for a specific water body. For this reason, Table 7.3 now only shows results for a limited number of water bodies for which monitoring data were available. Figure 7.1 shows the thresholds from Table 7.3 without the correction.

Finally, for comparison, the current WFD threshold values (i.e. Good/Moderate boundary) are also shown in Table 7.3 and a comparison with the OSPAR thresholds in neighbouring waters is shown in Figure 7.1. Clearly, the thresholds reflect the gradient towards lower concentrations in offshore waters. Current WFD thresholds (EC 2018) show a much weaker coast-offshore gradient in concentrations than the thresholds derived from the model results, that have higher values in near-coastal waters in both Dutch and German waters.

It is clear that there are differences in the threshold values estimated from the two models. There are several reasons that explain why the model results differ. For the DCSM-FM model, computation times are long. The historic scenarios had been run before model harmonisation had been finalized. Due to limitations in time and computational capacity, it was not possible to re-run the historic scenarios with the DCSM-FM model after model harmonisation. The SNS-GPM has a coarser resolution than the DCSM-FM model, which leads to lower concentrations in areas with strong gradients in concentration. Finally, both models differ in the way light limitation is estimated. Consequently, the thresholds that are now estimated by the models are uncertain.

The model results also show that the two scenarios HS1 and HS2 lead to different estimates in thresholds. The differences between the two scenarios vary considerably between water bodies, however, and are relatively large in the Elbe area where the HS2 scenario has a substantially lower TP load in HS2. In the OSPAR decision on thresholds, the thresholds were based on scenario HS1 in Meuse and Rhine and on scenario HS2 in Ems, Weser and Elbe (OSPAR 2022, Annex 6).

Table 7.3. Threshold values (Good/Moderate boundary) of growing season mean (March-Sept.) chlorophyll per water body, resulting from the application of the two historic scenarios HS1 and HS2 in the DCSM-FM (Deltares) and SNS-GPM (University of Hamburg) model and corrected threshold values after application of the 'Schernewski' correction. For comparison, current WFD thresholds (Good/Moderate boundary) are shown as growing season mean values, calculated as 0.5*90th percentile. The G/M boundaries as 90th percentile values are shown in Table 1.1. Thresholds in bold show the results for HS1 in NL water bodies and for HS2 in DE water bodies, in line with the selection of pre-eutrophic scenarios in OSPAR (OSPAR 2022).

WFD Water body	DCSM-FM threshold HS1	DCSM-FM threshold HS2	DCSM-FM correction factor	DCSM-FM corrected threshold HS1	DCSM-FM corrected threshold HS2	SNS-GPM threshold HS1	SNS-GPM threshold HS2	SNS-GPM correction factor	SNS-GPM corrected threshold HS1	SNS-GPM corrected threshold HS2	Current WFD G/M boundary (mean)
NL Wadden Coast N1	**8.2**	7.7	2.8	23.0	21.8	**8.5**	8.2	1.9	16.3	15.6	7.5
NL Wadden Sea N4	**10.5**	10.6	1.8	18.8	19.0	**8.3**	8.0	1.8	14.7	14.1	7.2
NL/DE Ems N3*	**8.6**	5.8	3.2	27.9	18.9	**9.4**	8.8	1.5	13.9	13.0	5.1
DE Ems N1	8.3	**6.8**	2.5	21.3	17.4	8.1	**8.0**	1.5	12.2	12.0	3.8
DE Ems N2	10.7	**8.2**	1.0	11.0	8.5	7.2	**7.1**	1.3	9.1	9.0	3.8
Ems N4	11.4	**7.8**	1.7	19.8	13.4	9.0	**8.5**	2.0	18.2	17.2	5.5
Weser N1	7.8	**5.7**	5.1	39.9	29.4	8.3	**8.3**	1.0	7.9	7.9	3.8
Weser N2	8.9	**5.8**	1.6	14.2	9.3	10.3	**9.6**	0.8	7.9	7.3	3.8
Weser N3	13.7	**8.1**	2.4	33.1	19.7	13.0	**11.8**	0.8	10.4	9.5	5.5
Weser N4-01	26.5	**16.2**	-	-	-	17.6	**15.1**	-	-	-	5.5
Weser N4-02	22.5	**13.1**	-	-	-	21.9	**17.0**	-	-	-	5.5
Elbe N3	19.5	**11.1**	-	-	-	18.9	**15.5**	-	-	-	5.5
Elbe N4-01	31.3	**18.4**	-	-	-	22.7	**17.8**	-	-	-	5.5
Elbe N4-02	31.8	**18.7**	-	-	-	30.5	**21.8**	-	-	-	5.5

*For the Ems-N3 area, there are two Chl thresholds in use, which differ between the Dutch and the German part. For the Dutch water body, characterised as "Ems-Dollard coast", the mean value of 5.1 µg/l is applied (good status), while for the German water body, indicated as "Ems-N3", a value of 5.5 µg/l is used (poor status).
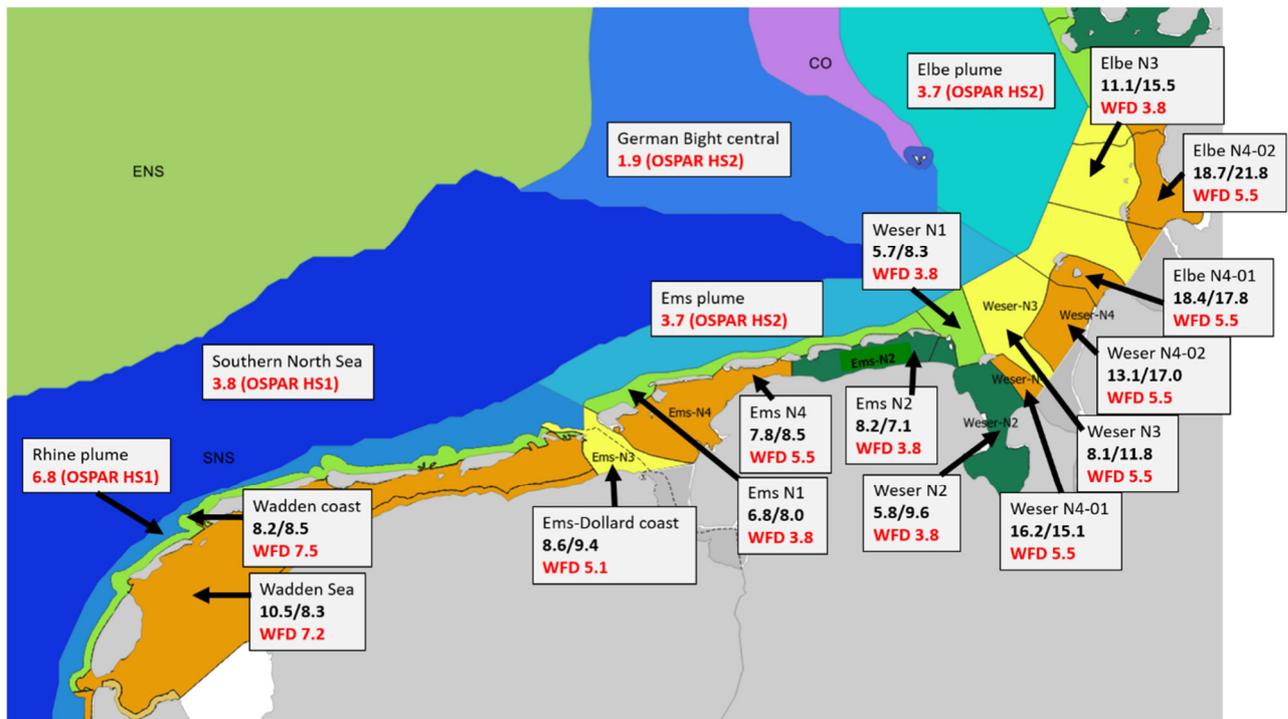
Figure 7.1. Map showing in red the current growing season mean chlorophyll thresholds for OSPAR (OSPAR 2022) and WFD (EC 2018). In black the thresholds for WFD water bodies resulting from the model approach in this project; the two values shown give the results from both models based on scenario HS1 for Dutch water bodies and scenario HS2 for German water bodies (bold values in Table 7.3). Note that the indicated thresholds are without the "Schernewski" correction.

7.3.3 Implications of the applied method of threshold setting

The results of the modelling of the pre-eutrophic reference conditions (historical scenarios) show that there are no major differences between the German and the Dutch Wadden Sea in the relation between nutrient enrichment and Chl concentrations as proxy for phytoplankton biomass. The difference in WFD threshold values for Chl in the shared water body type NEA 3/4, as currently exist in Germany and the Netherlands, are therefore not supported by natural conditions and scientific knowledge derived from the analysis of extensive monitoring data and the model results. The differences in the existing thresholds result from the fact that both countries have so far used different methods to derive threshold values. In this project, threshold values were calculated according to the OSPAR method "reference value+50% addition = threshold value". This leads to more harmonised thresholds for the Dutch and German Wadden Sea and helps to align the threshold values with the threshold defined for marine waters by OSPAR. There are still differences, as the estimates for thresholds for WFD waters are based on only two models, whereas for OSPAR eight models were included in the ensemble approach and the final OSPAR thresholds included corrections made to the model estimates (OSPAR 2022, Annex 6). Another systematic problem is that because Chl mean values are based on inhomogeneous measurements, which are often biased towards coastal stations, they cannot be directly compared against model means that are based on an area average from homogeneous information of the Chl concentration at every grid cell.

It should be noted that the method to derive threshold values for the WFD by adding 50% to a pre-eutrophic reference does not realistically reflect the individual environmental conditions in the different areas and water bodies of the Wadden Sea (or the wider North Sea, for that matter). First, simply adding 50% everywhere as the acceptable deviation from reference conditions does not reflect the spatial differences in the Wadden Sea with respect to freshwater influence, pressure from nutrient

loads and the influence of other environmental factors such as light that have an impact on the nutrient-Chl relationship. Second, using 50% as an acceptable deviation does not necessarily lead to the point where Chl concentrations are indicative of a change in the status of the ecosystem from "Good" to "Moderate".

In various discussions and ultimately through the available results it became clear, that this method requires a revision, since a simple addition of 50% on top of the reference value is not scientifically reasonable in all water bodies and thus the setting of threshold values is not adapted to the existing conditions. In the ICG-EMO report (Lenhart et al. 2022) the problem is classified as "by applying 50 % on top of reference condition for historic nutrient and chlorophyll concentration alike a linearity between nutrients and chlorophyll is introduced which has no scientific basis". This methodological discrepancy is an issue at present discussions within OSPAR expert groups and must be further addressed at higher levels in the corresponding scientific and administrative expert bodies (OSPAR, MSFD, EU).

For this reason, no concrete threshold values for the respective coastal water bodies are determined in this project, but for now only the results of the various models and scenarios are presented, which represent only an approximation and a recommendation for further discussion on the setting of threshold values.

Summarising the upper arguments, we come to these two major conclusions:

- The model results do not support the current differences in WFD Good/Moderate boundaries between the Netherlands and Germany and give no indications that thresholds in the Dutch part of the Wadden Sea should be different from the thresholds in the German part of the Wadden Sea.
- The thresholds calculated from the pre-eutrophic references, are, in all water bodies, higher than the current WFD thresholds. This is the case before and after the "Schernewski" correction.

7.3.4 How to deal with uncertainty

The Wadden Sea is a multi-pressure system that is now affected by many overlapping stressors. It has changed dramatically over the last 50 years, as the monitoring data collected by Rijkswaterstaat and NLWKN show. A change in nutrient concentrations in an order of magnitude, an exceptionally strong warming and massive changes in the biomass and biodiversity of phytoplankton are clearly visible during this period.

The classification of the ecological status of the Wadden Sea is the basis for identifying measures in those cases where the ecological status is Moderate or worse. Adequate confidence in the assessment result and the classification is therefore a prerequisite and this asks for recognition of sources of uncertainty in the assessment. Three types of uncertainty can be distinguished: incomplete knowledge, system unpredictability and ambiguity in the science-policy interface (Opdam et al. 2009). Incomplete knowledge comes from lack of data, imperfect cause-effect models, etc., which can to some extent be solved by scientific research. System unpredictability is the result of spatial and temporal variability and stochastic processes and is an inherent feature of ecosystems that can be addressed only to some extent by monitoring and data analysis. The third type of uncertainty, ambiguity in the science-policy interface is connected to the definition of objectives, indicators and thresholds.

In this project, we have tried to develop proposals for indicators and threshold levels that can be applied in the assessment of composition, abundance and biomass of phytoplankton.

The basis for a common assessment of the status of the Dutch and German Wadden Sea is a monitoring program that is suited to capture the information that is needed. The analysis of phytoplankton monitoring data showed that there are methodological differences between the monitoring programs of the Netherlands and Germany that complicate the comparability. Chapter 6 makes several recommendations for improvement. The availability of monitoring data, not only for phytoplankton composition but also for Chl as well as the temporal and spatial coverage of the monitoring program could be improved. This is particularly important as phytoplankton has a very high temporal and spatial variability, in particular in tidal dominated coastal systems, and monitoring with a low frequency results in underestimation of the phytoplankton biomass (Blauw et al. 2012, 2018, Fettweis et al. 2023). The use of harmonised thresholds for Chl is hindered also by the difference in analytical methods used by the Netherlands and Germany (HPLC versus photometer), that leads to different values for Chl concentrations. In the future, the use of satellite data in addition to *in situ* data could be a way to improve the accuracy of the assessment of Chl concentration.

An effective application of an alternative parameter for phytoplankton biomass – carbon content estimated from cell volume – was hampered by the fact that individual cell-size measurements were only available from one data set. Therefore, it is highly recommended to invest in measuring cell volume directly from the sample rather than using standardised literature values, which are often overestimated and do not capture temporal changes.

With our model approach we have simulated pre-eutrophic conditions, aligned with the approach that was used by OSPAR to set thresholds for Chl in marine waters. Despite extensive improvement and harmonisation of the models during this project, the models still show differences in the responses to nutrient input reduction. This is due on the one hand to the different spatial resolution of the models and on the other hand to their different sensitivity to light conditions. The Deltares model is more sensitive to changes in light climate and less sensitive to changes in nitrogen concentration, whereas the SNS-GPM model is less influenced by light climate but more influenced by changes in nitrogen concentrations. The FSK-model is directly coupled with a thoroughly validated morphodynamic model, which allows to compute a physically consistent SPM distribution and in turn have more realistic light climate information available in the water column.

The use of three models with harmonised forcings (e.g. nutrient loads, concentrations at the seaward boundaries) but with differences in model setup and formulations gives an impression of the uncertainty related to model estimates of nutrient and Chl concentrations in the Wadden Sea. Inevitably, while there is generally a good agreement between the model results, and between model results and observations (measurements), there are differences in model results on a water body level. Some of these differences can be explained by the characteristics of the models, but there will also be a remaining inherent uncertainty when more than one model is used. In OSPAR, the weighted result of an ensemble of eight models was assumed to lead to representative results. For the Dutch and German Wadden Sea, we used the two larger-scale models that were available from Deltares and the University of Hamburg. For those areas (Ems and adjacent Wadden Sea) also covered by the higher resolution FSK-model, we observed further improved results with respect to the available Chl measurements. It is left to further research to resolve to what extent the cause for this lies in the higher model resolution or the model formulation better adapted to local processes e.g. light climate, wetting/drying of tidal flats, improved nutrient exchange between bottom sediment and water column.

Finally, another source of considerable uncertainty is associated with the definition of the threshold that determines the boundary between Good and Moderate Ecological Status for phytoplankton. There are effects of eutrophication on the Wadden Sea ecosystem besides increased phytoplankton biomass, such as decline of seagrass, increased green alga coverage and anoxic spots (see van Beusekom et al. 2017). However, a quantitative link between the level of Chl concentrations and those eutrophication impacts has not yet been established. The lack of knowledge of a specific Chl concentration that may represent a 'tipping point' where the state of the ecosystem changes, is due to its general complexity and nonlinearity, but also incomplete knowledge. Moreover, it must be emphasized that, given the complexity of the dynamic system, model building always requires a trade-off between resolution and model complexity on the one hand and computational capability on the other. This applies even more when cross-scale applications are required, as in the present case. As a consequence, we have to resort to a method where the threshold is derived as an acceptable deviation from reference conditions. Consequently, the approach to use 50% deviation from reference to set the threshold clearly lacks a scientific basis and requires further discussion on the implementation of policy objectives.

We are facing considerable uncertainties in defining a method for the accurate assessment of the state of an ecosystem like the Wadden Sea. In addition, this assessment has to be applied to a system that is highly dynamic and shows large variability in time and space. At the same time, there is a need for an assessment that supports management in deciding on appropriate measures and prevents deterioration of the system. An iterative approach of revision of thresholds with the advancement of our understanding of the ecosystem could be a solution that also provides an opportunity to deal with changing conditions, such as the effects of climate change.

7.4 Final remarks

One of the main goals of this project was to develop an alternative assessment approach for phytoplankton in the Wadden Sea. During the course of the project, this turned out to be a complex task where we managed to address some basic principles. The project duration was too short to fully discuss the implications of the project results and to apply the further developments from the high-resolution model setup of FSK into the other two models. However, this project and the approaches that were developed are a good starting point for further discussion at scientific and policy levels, also with regard to future work within OSPAR.

The analysis of phytoplankton biomass, diversity, and species composition are important tools for understanding the aquatic system and the environmental conditions. However, they don't allow defining "thresholds" values of a good or bad water quality status. Rather, they quickly and reliably reflect the gradual changes occurring in the environment. In a dynamic system such as the Wadden Sea, the assessment of water quality should focus on such temporal trends, which can only be analysed with continued monitoring programs creating their own baselines for comparison. These programs should include not only regular and consistent phytoplankton sampling but also measurement of cell size (biovolume) and carbon content.

The implicit assumption underlying the current WFD thresholds for Chl, as an indicator of phytoplankton biomass, is that there is a strong correlation with nitrogen concentrations. However, the results of this project show, both in monitoring data as well as in modelling, that Chl concentrations do not react linearly to nitrogen inputs and that phosphorus limitation also plays a major role in phytoplankton growth in the Wadden Sea. The model scenario with river loads into the North Sea complying with the management objective of 2.8 mg TN/l, showed that the effect of reducing TN loads, with up to 20-40%, on Chl concentrations in the Wadden Sea is limited. The occurrence of phosphorus

limitation during part of the phytoplankton growing season and additionally the role of light limitation and grazing are causes for this limited effect of reduced nitrogen loads. While this could suggest that further reducing phosphorus loads may have a stronger effect on Chl on the short term, a further shift in the balance between nitrogen and phosphorus caused by further reductions in phosphorus loads may result in a larger excess of nitrogen elsewhere in the North Sea. Changes in the balance between nitrogen and phosphorus (N:P ratio) may potentially have impacts on growth, species composition and nutritional quality of marine phytoplankton as well, with knock-on effects on the marine food web (Burson et al. 2016, Grosse et al. 2017, Bi & Sommer 2020).

7.5 Summary of key recommendations

Research

o   Study the effect of light and light limitation on the development of phytoplankton biomass
o   Study the effect of potential phosphorous limitation in the Wadden Sea
o   Analyse the effects of changing (increasing) N:P ratios on all levels of the marine food web in the Wadden Sea and North Sea
o   Consider top-down control by zooplankton and benthic filter feeders on development of phytoplankton biomass

Ecosystem modelling

o   Continue to harmonize and align ecosystem models to achieve better outcomes
o   Implement the process of phosphate-release from the sediments into the Wadden Sea ecosystem models
o   Reduce the uncertainty of chlorophyll response on nutrient reduction
o   Include zooplankton and benthic filter feeders in all ecosystem models

Phytoplankton and eutrophication monitoring and assessment

o   Better align Dutch and German monitoring programs and methodology to ensure good comparability
o   Conduct parallel measurements of chlorophyll samples in Dutch and German laboratories to verify the comparability of methods
o   Include light measurements to understand algal growth limitations
o   Cross-check taxonomy to resolve discrepancies in species inventory
o   Measure phytoplankton cell size in samples for better insights in long-term changes
o   Align zooplankton measurements with phytoplankton measurements for better analysis
o   Conduct nutrient bioassays to assess nutrient limitation
o   Include particulate organic carbon and pigment analyses for additional information
o   Consider multiple parameters for the assessment of phytoplankton and eutrophication, instead of relying on chlorophyll alone
o   Instead of basing the assessment solely on absolute values like thresholds, also include the analysis of temporal trends in long-time series (e.g. for biodiversity parameters, biomass, cell size, etc.)
o   Publish harmonized data for improved assessment and scientific collaboration
o   Continue the phytoplankton monitoring programs for better understanding of ecosystem changes

Policy

- o Consider the Wadden Sea as a multi-stressor system in a holistic assessment
- o Continue the effort to support the dialog of harmonized science-based threshold setting
- o Revise method of threshold setting "reference+50%=threshold"
- o Consider that the river management objective of 2.8 mg TN/l might not be enough to achieve good ecological status for phytoplankton in coastal waters
- o Consider controlling phytoplankton biomass via nutrient reductions beyond reducing nitrogen alone
- o Continue the good cooperation and exchange between Dutch and German national authorities and scientific institutions
- o Support the process of alignment of methods and assessment approaches between WFD, MSFD and OSPAR

8. References

Baretta-Bekker, H., A. Sell, F. Marco-Rius, J. Wischnewski, L. Pamela Walsham, M. Mohlin, K. Wesslander, H. Ruiter, F. Gohin & L. Enserink (2015). The chlorophyll case study in the JMP NS/CS project. Document produced as part of the EU project: 'Towards joint Monitoring for the North Sea and Celtic Sea' (Ref: ENV/PP 2012/SEA). 72 pp. https://www.informatiehuismarien.nl/publish/pages/111449/chlorophyll_case_study_20150716_4602.pdf

Bates, D., M. Mächler, B. Bolker & S. Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67:48.

Bi, R. & U. Sommer (2020). Food Quantity and Quality Interactions at Phytoplankton–Zooplankton Interface: Chemical and Reproductive Responses in a Calanoid Copepod. *Frontiers in Marine Science* 7: 274.

Blauw, A., M. Eleveld, T. Prins, F. Zijl, J. Groenenboom, G. Winter, L. Kramer, T. Troost, A. Bartosova, J. Johansson, R. Capell, K. Eiola, A. Höglund, G.H. Tilstone, P.E. Land, V. Martinez-Vicente, S. Pardo & D. van der Zande (2019). Coherence in assessment framework of chlorophyll a and nutrients as part of the EU project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data. Ref: DG ENV/MSFD Second Cycle/2016. Activity 1 Report, 86 pp.

Blauw, A.N., H.F.J. Los, M. Bokhorst & P.L.A. Erftemeijer (2009). GEM: a generic ecological model for estuaries and coastal waters. *Hydrobiologia* 618: 175-198.

Blauw, A.N., Benincà, E., Laane, R.W.P.M., Greenwood, N., Huisman, J., 2012. Dancing with the tides: fluctuations of coastal phytoplankton orchestrated by different oscillatory modes of the tidal cycle. *PLoS ONE* 7 (11), e49319.

Blauw, A.N., Benincà, E., Laane, R.W.P.M., Greenwood, N., Huisman, J., 2018. Predictability and environmental drivers of chlorophyll fluctuations vary across different time scales and regions of the North Sea. *Progress in Oceanography* 161, 1−18.

BLMP (2011). Konzept zur Ableitung von Nährstoffreduzierungszielen in den Flussgebieten Ems, Weser, Elbe und Eider aufgrund von Anforderungen an den ökologischen Zustand der Küstengewässer gemäß Wasserrahmenrichtlinie. Ad-hoc-AG Nährstoffreduzierung des BLMP, Verabschiedet auf der 17. ARGE BLMP am 01.07.2011 und den Lenkungsausschuss der Expertengruppe Meer, 9. Sitzung am 24.3.11. Stand: 01.06.2011, 50 pp.

Bouwman, A.F., M.F.P. Bierkens, J. Griffioen, M.M. Hefting, J.J. Middelburg, H. Middelkoop & C.P. Slomp (2013). Nutrient dynamics, transfer and retention along the aquatic continuum from land to ocean: towards integration of ecological and biogeochemical models. *Biogeosciences* 10: 1–23.

Burchard H & Bolding Kristensen K (2002). GETM, a General Estuarine Transport Model. EUR 20253 EN. European Commission. JRC23237.

Burson, A., M. Stomp, L. Akil, C.P.D. Brussaard & J. Huisman (2016). Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the North Sea. *Limnology and Oceanography* 61: 869-888.

Chase, J. M. & T. M. Knight (2013). Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. *Ecology Letters* 16: 17-26.

Clark, N., Eber, L., Laurs, R., Renner, J., and J. Saur (1974). Heat exchange between ocean and atmosphere in the Eastern North Pacific for 1961–1971, Tech. rep., NOAA, Washington, DC, USA.

CMEMS: https://data.marine.copernicus.eu/product/OCEANCOLOUR_NWS_BGC_HR_L4_NRT_009_209/description

Cunningham, R. B., & D. B. Lindenmayer (2005). Modelling count data of rare species: some statistical issues. *Ecology*, 86(5), 1135-1142.

CWSS (2017) Introduction. In: Wadden Sea Quality Status Report. Eds.: Kloepper S. et al., Common Wadden Sea Secretariat, Wilhelmshaven, Germany. Last updated 01.03.2018.

EEC (1992). Council Directive 92 /43 /EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora.

EC (2009). Directive 2009/147/EC of the European Parliament and of the Council of 30 November 2009 on the conservation of wild birds.

EC (2000). Water Framework Directive (Directive 2000/60/EC) of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. European Commission, Brussels.

EC (2003a). Guidance Document No 5. Transitional and Coastal Waters – Typology, Reference Conditions and Classification Systems. *Common Implementation Strategy for the Water Framework Directive (2000/60/EC),* European Commission, Luxemburg, 107 pp.

EC (2003b). Guidance document no 10. River and lakes – Typology, reference conditions and classification systems. *Common Implementation Strategy for the Water Framework Directive (2000/60/EC),* European Commission, Luxemburg, 87 pp.

EC (2008). Marine Strategy Framework Directive (Directive 2008/56/EC) of the European Parliament and of the Council Directive establishing a framework for Community action in the field of marine environmental policy. European Commission, Brussels.

EC (2018). Commission Decision (EU) 2018/229 of 12 February 2018 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise and repealing Commission Decision 2013/480/EU.

Egge, J.K. & D.L. Aksnes (1992). Silicate as regulating nutrient in phytoplankton competition. *Marine Ecology Progress Series* 83: 281-289.

Elser, J. J., M. E. S. Bracken, E. E. Cleland, D. S. Gruner, W. S. Harpole, H. Hillebrand, J. T. Ngai, E. W. Seabloom, J. B. Shurin & J. E. Smith (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters* 0:1135-1142.

EMEP (2020). Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components, EMEP Status Report 1/2020 MSC-W & CCC & CEIP, Norwegian Meteorological Institute (EMEP/MSC-W), Oslo, Norway. https://www.diva-portal.org/smash/get/diva2:1527049/FULLTEXT01.pdf

Enserink, L., A. Blauw, D. van der Zande & S. Markager (2019). Summary report of the EU project 'Joint monitoring programme of the eutrophication of the North Sea with satellite data' (Ref: DG ENV/MSFD Second Cycle/2016). Rijkswaterstaat, Deltares, Royal Belgian Institute of Natural Sciences, Aarhus University, 21 pp.

Erisman, J.W., H. Van Grinsven, B. Grizzetti, F. Bouraoui, D. Powlson, M.A. Sutton, A. Bleeker & S. Reis (2011). The European nitrogen problem in a global perspective.In: *The European Nitrogen Assessment*; *Sutton, M.A., Howard, C.M., Erisman, J.W., Billen, G., Bleeker, A., Grennfelt, P., Van Grinsven, H. en Grizzetti, B.* (eds.). Cambridge University Press, Cambridge. p: 9-31. http://www.nine-esf.org/ENA-Book

Fettweis, M., R. Riethmüller, D. Van der Zande & X. Desmit (2023). Sample based water quality monitoring of coastal seas: How significant is the information loss in patchy time series compared to continuous ones? *Science of The Total Environment* 873: 162273.

Fischer, H., Fischer, M., Kuhn, U., Ulrich, A., Ollesch, G., Ritz, S., Trepel, M. & M. Venohr (2014). Empfehlung zur Übertragung flussbürtiger, meeres-ökologischer Reduzierungsziele ins Binnenland. Dresden, LAWA-AO Geschäftsstelle. Sächsisches Staatsministerium für Umwelt und Landwirtschaft, LAWA-Arbeitsprogramm Flussgebietsbewirtschaftung. Produktdatenblatt WRRL-2.4.7. Stand 18.06.2014, 17 pp.

Galloway, J.N., J.D. Aber, J.W. Erisman, S.P. Seitzinger, R.W. Howarth, E.B. Cowling & B.J. Cosby (2003). The Nitrogen Cascade. *BioScience* 53: 341-356.

Geyer, B (2014). High-resolution atmospheric reconstruction for Europe 1948–2012: coastDat2. *Earth Systems Science Data* 6: 147–164.

Grosse, J., A. Burson, M. Stomp, J. Huisman & H.T.S. Boschker (2017). From ecological stoichiometry to biochemical composition: Variation in N and P supply alters key biosynthetic rates in marine phytoplankton. *Frontiers in Microbiology* 8: 1299.

Guildford, S. J. & R. E. Hecky (2000). Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: Is there a common relationship? *Limnology and Oceanography* 45:1213-1223.

Gypens, N., Lancelot, C., & K. Soetaert (2008). Simple parameterisations for describing N and P diagenetic processes: Application in the North Sea. *Progress in Oceanography* 76(1), 89-110.

Hanslik M., Rahmel J., Bätje M., Knieriemen S., Schneider G. & S. Dick (1998). Der Jahresgang blütenbildender und toxischer Algen an der niedersächsischen Küste seit 1982. In: Umweltbundesamt Texte Bonn.

Harpole, W. S., J. T. Ngai, E. E. Cleland, E. W. Seabloom, E. T. Borer, M. E. S. Bracken, J. J. Elser, D. S. Gruner, H. Hillebrand, J. B. Shurin & J. E. Smith (2011). Nutrient co-limitation of primary producer communities. *Ecology Letters* 14:852-862.

Hillebrand, H., B. Blasius, E. T. Borer, J. M. Chase, J. A. Downing, B. K. Eriksson, C. T. Filstrup, W. S. Harpole, D. Hodapp, S. Larsen, A. M. Lewandowska, E. W. Seabloom, D. B. Van de Waal & A. B. Ryabov (2018). Biodiversity change is uncoupled from species richness trends: consequences for conservation and monitoring. *Journal of Applied Ecology* 55:169-184.

Hillebrand, H., Antonucci Di Carvalho, J., Dajka, J. C., Dürselen, C. D., Kerimoglu, O., Kuczynski, L., Rönn, L. & A. Ryabov (2021). Temporal declines in Wadden Sea phytoplankton cell volumes observed within and across species. *Limnology and Oceanography*, 67(2), 468-481.

Hillebrand, H., E. Acevedo-Trejos, S. D. Moorthi, A. Ryabov, M. Striebel, P. Thomas & M.-L. Schneider (2022). Cell size as driver and sentinel of phytoplankton community structure and functioning. *Functional Ecology* 36:276-293.

IenW (2022). Stroomgebiedbeheerplannen Rijn, Maas, Schelde en Eems 2022 – 2027. Den Haag, Ministerie van Infrastructuur en Waterstaat, Maart 2022.

IKSR/CIPR/ICBR (2022). Internationaal gecoördineerd stroomgebiedbeheerplan 2022-2027 van het internationaal stroomgebieddistrict Rijn / International koordinierter Bewirtschaftungsplan 2022-2027 für die internationale Flussgebietseinheit Rhein. Koblenz, IKSR/CIPR/ICBR, maart 2022, März 2022 https://www.iksr.org/fileadmin/user_upload/DKDM/Dokumente/BWP-HWRMP/DE/bwp_De_BWP_2021.pdf

IPBES (2019): Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. E. S. Brondizio, J. Settele, S. Díaz, & H. T. Ngo (editors). IPBES secretariat, Bonn, Germany. 1148 pp. https://doi.org/10.5281/zenodo.3831673

Jansen, J., Hill, N. A., Dunstan, P. K., Eléaume, M. P., & C. R. Johnson (2018). Taxonomic resolution, functional traits, and the influence of species groupings on mapping Antarctic seafloor biodiversity. *Frontiers in Ecology and Evolution* 6, 81.

Kerimoglu, O., R. Hofmeister, J. Maerz, R. Riethmüller & K.W. Wirtz (2017). The acclimative biogeochemical model of the southern North Sea. *Biogeosciences* 14: 4499-4531.

Kerimoglu, O., Voynova, Y. G., Chegini, F., Brix, H., Callies, U., Hofmeister, R., Klingbeil, K., Schrum, C. & J. E. E. van Beusekom (2020). Interactive impacts of meteorological and hydrological conditions on the physical and biogeochemical structure of a coastal system. *Biogeosciences* 17: 5097-5127.

Klein Goldewijk, K., A. Beusen & P. Janssen (2010). Long term dynamic modeling of global population and built-up area in a spatially explicit way, HYDE 3 .1. *The Holocene* 20: 565-573.

Klein Goldewijk, K., A. Beusen, M. de Vos & G. van Drecht (2011). The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years. *Global Ecology and Biogeography* 20: 73-86.

Knudsen, L. & E. Schnug (2016). Utilization of Phosphorus at Farm Level in Denmark. In: *Phosphorus in Agriculture: 100 % Zero*; *Schnug, E. en De Kok, L.J.* (eds.). Springer Netherlands, Dordrecht. p: 215-229. 10.1007/978-94-017-7612-7_11https://doi.org/10.1007/978-94-017-7612-7_11

Kondo, J. (1975). Air-sea bulk transfer coefficients in diabatic conditions, *Boundary-Layer Meteorology* 9, 91–112.

Kyllingsbæk, A. (2005). Nutrient balances and nutrient surpluses in Danish agriculture 1979–2002: nitrogen, phosphorus, potassium. Research Center Foulum, Danish Institute of Agricultural Sciences, DJF Report Nr. 116, 100 pp.

LAWA (2019). Empfehlung fur die Begründung von Fristverlängerungen auf Grund von „natürlichen Gegebenheiten" für die Ökologie. Ständiger Ausschuss „Oberirdische Gewässer und Küstengewässer" LAWA AO, 43 pp.

Lenhart, H., A. Blauw, X. Desmit, L. Fernand, R. Friedland, B. Heyden, O. Kerimoglu, G. Lacroix, A. van der Linden, J. van der Molen, M. Plus, T. Prins, I. Ruvalcaba Baroni, T. Silva, C. Stegert, D. Thewes, T. Troost, L. Vilmin & S. van Leeuwen (2022). ICG-EMO report on model comparison for historical scenarios as basis to derive new threshold values. OSPAR, London, Publication Number: 895/2022, 66 pp.

Lindström, G., C. Pers, J. Rosberg, J. Strömqvist & B. Arheimer (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research* 41: 295-319.

Löder, M. G. J., Kraberg, A. C., Aberle, N., Peters, S. & K.H. Wiltshire (2012). Dinoflagellates and ciliates at Helgoland Roads, North Sea. *Helgoland Marine Research* 66: 11-23.

Los, F.J., M.T. Villars & M.W.M. van der Tol (2008). A 3-dimensional primary production model (BLOOM/GEM) and its applications to the (southern) North Sea (coupled physical–chemical–ecological model). *Journal of Marine Systems,* 74: 259-294.

Menden-Deuer, S. & E.J. Lessard (2000). Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnology and Oceanography* 45(3): 569-579.

Mouillot, D., Spatharis, S., Reizopoulou, S., Laugier, T., Sabetta, L., Basset, A., & T. Chi (2006). Alternatives to taxonomic-based approaches to assess changes in transitional water communities. Aquatic Conservation: *Marine and Freshwater Ecosystems* 16(5), 469-482.

MU (2021). Niedersächsischer Beitrag zu den Bewirtschaftungsplänen 2021 bis 2027 der Flussgebiete Elbe, Weser, Ems und Rhein. Hannover, Niedersächsisches Ministerium für Umwelt, Energie, Bauen und Klimaschutz, Dezember 2021.

Nechad, B., Ruddick, K. G., & Y. Park (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sensing of Environment* 114(4), 854-866.

Nohe, A., Knockaert, C., Goffin, A., Dewitte, E., De Cauwer, K., Desmit, X., Vyverman, W., Tyberghein, L., Lagring, R. & K. Sabbe (2018). Marine phytoplankton community composition data from the Belgian part of the North Sea, 1968-2010. *Scientific Data* 5: 180126.

Officer, C.B. & J.H. Ryther (1980). The possible importance of silicon in marine eutrophication. *Marine Ecology Progress Series* 3: 83-91.

Opdam, P.F.M., M.E.A. Broekmeyer & F.H. Kistenkas (2009). Identifying uncertainties in judging the significance of human impacts on Natura 2000 sites. Environmental Science & Policy 12: 912-921.

OSPAR (2022). Common Procedure for the Identification of the Eutrophication Status of the OSPAR Maritime Area. OSPAR Agreement 2022-07. https://www.ospar.org/documents?v=49366

Payne, R. (1972). Albedo of the Sea Surface. *Journal of the Atmospheric Sciences* 29(5): 959–970.

Prins, T.C., Desmit, X. & J.G. Baretta-Bekker (2012). Phytoplankton composition in Dutch coastal waters responds to changes in riverine nutrient loads. *Journal of Sea Research* 73: 49-62.

Rombouts, I., N. Simon, A. Aubert, T. Cariou, E. Feunteun, L. Guérin, M. Hoebeke, A. McQuatters-Gollop, F. Rigaut-Jalabert & L. F. Artigas (2019). Changes in marine phytoplankton diversity: Assessment under the Marine Strategy Framework Directive. *Ecological Indicators* 102:265-277.

Schernewski, G., R. Friedland, M. Carstens, U. Hirt, W. Leujak, G. Nausch, T. Neumann, T. Petenati, S. Sagert, N. Wasmund & M. von Weber (2015). Implementation of European marine policy: New water quality targets for German Baltic waters. *Marine Policy* 51: 305-321.

Schmid, T. (2000). Food consumption and nitrogen flux. The impact of lifestyle in a long-term perspective. Paper presented at the workshop "How do households use their resources?" Department of Technology and Social Change, Linköping University, April 2000.

Schöpp, W., M. Posch, S. Mylona & M. Johansson (2003). Long-term development of acid deposition (1880–2030) in sensitive freshwater regions in Europe. *Hydrology and Earth System Science* 7: 436-446.

Smil, V. (2000). Phosphorus in the environment: Natural Flows and Human Interferences. *Annual Review of Energy and the Environment* 25: 53-88.

Sutton, M.A., C.M. Howard, J.W. Erisman, W.J. Bealey, G. Billen, A. Bleeker, A.F. Bouwman, P. Grennfelt, H. Van Grinsven & B. Grizzetti (2011). The challenge to integrate nitrogen science and policies: the European Nitrogen Assessment. In: *The European Nitrogen Assessment*; *Sutton, M.A., Howard, C.M., Erisman, J.W., Billen, G., Bleeker, A., Grennfelt, P., Van Grinsven, H. en Grizzetti, B.* (eds.). Cambridge University Press, Cambridge. p: 82-96. http://www.nine-esf.org/ENA-Book

Thrane, J-E., Kyle, M., Striebel, M., Haande, S., Grung, M., Rohrlack, T. & T. Andersen (2015). Spectrophotometric analysis of pigments: a critical assessment of a high-throughput method for analysis of algal pigment mixtures by spectral deconvolution. *PLoS ONE* 10(9): e0137645.

Troost, T.A., J.W.M. Wijsman, S. Saraiva & V. Freitas (2010). Modelling shellfish growth with dynamic energy budget models: An application for cockles and mussels in the Oosterschelde (southwest Netherlands). *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 3567–3577.

Umweltbundesamt (UBA) [Hg.] (2019). Monitoringbericht 2019 zur Deutschen Anpassungsstrategie an den Klimawandel. www.klivoportal.de/monitoringbericht2019

van Beusekom, J.E.E., Bot, P., Carstensen, J., Grage, A., Kolbe, K., Lenhart, H.-J., Pätsch, J., Petenati, T. & J. Rick (2017) *Eutrophication*. In: Wadden Sea Quality Status Report. Eds.: Kloepper S. et al., Common Wadden Sea Secretariat, Wilhelmshaven, Germany. Last updated 05.11.2018.

van der Molen, D.T., Pot, R., Evers, C.H.M., van Erpen, F.C.J. & L.L.J. van Nieuwerburgh (Eds.) (2018). *Referenties en maatlatten voor natuurlijke watertypen voor de Kaderrichtlijn Water 2021-   2027.* Amersfoort, STOWA, 489 pp.

van Grinsven, H.J.M., Bouwman, L., Cassman, K.G., van Es, H.M., McCrackin, M.L. & A.H.W. Beusen (2015). Losses of Ammonia and Nitrate from Agriculture and Their Effect on Nitrogen Recovery in the European Union and the United States between 1900 and 2050. *Journal of Environmental Quality* 44: 356-367.

van Leeuwen, S. M., H.-J. Lenhart, T. C. Prins, A. Blauw, X. Desmit, L. Fernand, R. Friedland, O. Kerimoglu, G. Lacroix, A. van der Linden, A. Lefebvre, J. van der Molen, M. Plus, I. Ruvalcaba Baroni, T. Silva, C. Stegert, T. A. Troost and L. Vilmin (2023). Deriving pre-eutrophic conditions from an ensemble model approach for the North-West European seas. *Frontiers in Marine Science* 10:1129951. doi: 10.3389/fmars.2023.1129951

Vinther, M. (2012). Oversigt over Landsforsøgene (annually). Videncentret for Landbrug. Planter & Miljø, Aarhus, Denmark.

Weisse, R., Bisling, P., Gaslikova, L., Geyer, B., Groll, N., Hortami, M., Matthias, V., Maneke, M., Meinke, I., Meyer, E.M.I., Schwichtenberg, F., Stempinski, F., Wiese, F. & K. Wöckner-Kluwe (2015). Climate services for marine applications in Europe. *Earth Perspectives* 2(3), 1-14.

Zijl, F., Laan, S. & J. Groenenboom (2021). Development of a 3D model for the NW European Shelf (3D DCSN-FM). Deltares, Delft, Deltares report 11205259-015-ZKS-0003.

Annexes

Annex 1: Reduction scenario for river concentrations of 2.8 mg TN/l at the limnic/marine border

A1.1 Introduction

A model scenario was proposed to determine the effects of the application of the management objective of total nitrogen (TN) concentration at the limnic-marine border in rivers (2.8 mg TN/l) in association to the Water Framework Directive (WFD). The impacts of this river concentration and the associated riverine TN loads on nutrient and chlorophyll (Chl) concentrations in the German and Dutch Wadden Sea and WFD coastal water bodies were estimated.

As discussed within the Interreg modelling group, the target concentration was applied to all river inlets within the Dutch and German coastal region. In the model runs, this means that a reduction in river loads was applied, to change the current loads for the years in the model simulations to loads fitting the target concentration.

Nutrient loads of all rivers from other countries were kept at their current level. Atmospheric nitrogen deposition as well as TP and DIP concentrations were also kept at current levels in the entire model domain including the Dutch and German coastal region.

A1.2 Objective

The aim is to provide an estimate of the ecological consequences of the application of the WFD measures in the Netherlands and Germany, on the marine environment in the German and Dutch Wadden Sea. The effect of this TN 2.8 mg/l scenario on nutrient and Chl concentrations in the Wadden Sea will be quantified through model estimates. This model application will provide insight in the impact of reducing TN concentrations in Dutch and German rivers to a level that meets the management objective for TN in the rivers. The results will show what levels of nutrient and Chl concentrations in the Wadden Sea and adjacent coastal waters will be reached in the "2.8 mg TN/l Scenario" and if such a scenario will result in achieving current WFD/MSFD thresholds for Chl in the Wadden Sea and North Sea coastal waters.

As a TN reduction is the focus of the WFD measures for both countries, TP loads should be kept on its current level. This is also necessary as a simultaneous change in TP loads would make it hard to interpret the effects of the WFD measures on TN loads.

A1.3 TN concentrations and necessary reduction in the main rivers

The target concentration of TN in German rivers is 2.8 mg/l for the annual average (BLMP 2011, Fischer et al. 2014). The same target concentration applies to the river Rhine (IKSR/CIPR/ICBR 2021). The annual average of 2.8 mg/l is roughly equivalent to a summer average (April-September) of 2.5 mg/l (van der Molen et al. 2018), which is used by the Netherlands as target for large rivers. The target concentration for TN represents the boundary between Good and Moderate status in compliance with the WFD. In this project, both countries agreed on a compliant approach using the target concentration of 2.8 mg/l for the model scenario for both German and Dutch rivers.

The database of the OSPAR modelling group ICG-EMO, maintained by Sonja van Leeuwen (NIOZ), provides daily loads and concentrations of TN for each river. Here, we used the data for the years 2009-2017, which are the years used in the modelling for OSPAR ICG-EMO and the Interreg project.

The annual average concentration was compared with the target TN concentration of 2.8 mg/l (annual average) to derive an annual reduction percentage for each river:

Required reduction TN in year *x*: Reduction$_{year\ x}$= (1-2.8/TN$_{year\ x,\ river\ y}$)*100%

where TN$_{year\ x,\ river\ y}$ is the annual average TN concentration in year *x* in river *y*.

This calculation gives an annual reduction percentage, that differs per year and per river. In combination with the daily load data from the ICG-EMO database, this provides estimates of the daily loads after reduction:

Scenario_TN-load$_{day\ z,\ year\ x}$ = Reduction$_{year\ x}$ * TN-load$_{day\ z,\ year\ x}$

This approach is similar to the approach suggested by Kerimoglu et al. (2018). A detailed analysis for major European rivers is provided in the technical report by Grosse & Lenhart (2015). This approach keeps the seasonal variation in concentrations and loads of TN, but lowers the levels to achieve an annual average concentration of 2.8 mg/l.

Table 1.4 shows the annual average river loads and concentrations of TN, the required reduction percentage and the annual average TN concentration after reduction, for all main rivers in the database. There are a few small rivers with a minor TN load to the North Sea (Arlau, Bongsieler Kanal, Miele). For the other rivers, the highest required reduction is for Ems and Eider, that have annual average TN concentrations in the range 3.3-5.6 mg/l, requiring reductions of 16-50% (average 39%) to achieve the 2.8 mg/l target.

The other rivers have lower reduction targets, but in some rivers the interannual variability is relatively large, e.g. Elbe 3-36% and Scheldt 5-35%.

In Haringvliet (Meuse river basin district), Nieuwe Waterweg, Noordzeekanaal and the two discharge points of Lake Ijssel (Rhine river basin district) annual average concentrations of TN are already below 2.8 mg/l in most years.

Figure A1.4 shows the current daily TN concentrations in the rivers and the concentrations after application of the reduction percentage. Figure A1.5 shows the current annual TN loads and the TN loads after reduction.

Table A1.1 gives a summary of the results by showing the average over the period 2009-2017 for the data in Table A1.3.

Table A1.1. Discharge (Q), TN & TP load and concentrations, average, minimum and maximum reduction for the 2.8 mg/l TN target and average concentration after reduction for the years 2009-2017. Data source: OSPAR ICG EMO database of European rivers (Lenhart et al. 2022).

	River	Q (m3/s)	TN load (kton)	TP load (ton)	TN (mg/l)	TP (mg/l)	Average reduction TN (%)	Min-Max reduction (%)		TN after reduction (mg/l)
DE	Arlau	3	1	24	4.52	0.20	38	31	39	2.80
DE	Bongsieler Kanal	10	2	68	4.52	0.20	38	31	39	2.80
DE	Eider	26	4	180	4.60	0.21	39	33	41	2.80
DE	Elbe	819	98	4155	3.37	0.17	16	3	36	2.80
DE	Ems	91	16	437	4.62	0.13	38	16	50	2.80
DE	Miele, Warwerorter Kanal	7	1	46	4.53	0.21	38	31	39	2.80
DE	Weser	292	38	1535	3.65	0.16	23	14	35	2.80
NL	Haringvliet	503	53	1423	2.91	0.09	4	0	17	2.79
NL	IJsselmeer-oost	230	19	490	2.43	0.06	0	0	4	2.42
NL	IJsselmeer-west	259	22	570	2.44	0.06	0	0	4	2.43
NL	Nieuwe Waterweg	1350	124	4784	2.74	0.11	1	0	9	2.70
NL	Noordzeekanaal	80	6	494	2.29	0.19	0	0	0	2.29
NL	Scheldt	123	17	1226	3.81	0.30	25	5	35	2.80

A1.4 TN concentrations and necessary reduction in small discharge points

In addition to the large rivers and discharge points mentioned above, there is a large number of small discharge points in the Netherlands and Germany (see Table A1.4). Those discharge points contribute <5% to the total N loads of all discharge points combined (Fig. A1.1). The largest load comes from the Cleveringsluizen in the north of the Netherlands, that contribute circa 50% to the total load of those small discharge points.

A calculation of reduction percentages to meet the 2.8 mg/l target for each of these individual discharge points was expected to be too inaccurate, due to the small size of the loads. Alternatively, the reduction percentages for these small discharges were derived from the reduction percentage of the nearest large river. Table A1.2 shows the link of the small discharge points to the larger rivers.
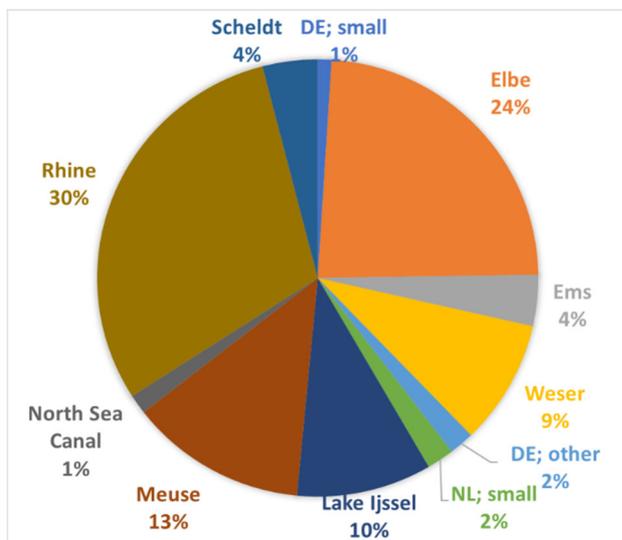
Figure A1.1. Relative contribution of the large rivers and small discharge points (DE;small and NL;small) to the total TN load (average for 2009-2017). The group DE;other are the other rivers from Table A1.1, not specified in this figure.

Table A1.2. Linking the small discharge points to the nearest main river.

River name	Nearest main river	River name	Nearest main river
Husumer_Au	Eider	Cleveringsluizen	
Tetenbuellspieker_Kanal		DeDrieDelfzijl	
Deichsiel_Suederhafen	Elbe	Duurswold	
Pinnau		Eemskanaal	
Accumersiel		Fiemel	
Bensersiel		Knock	Ems
Dangaster_Siel		Leybuchtsiel	
Eckwarder_Siel		NieuwStatenzijl	
Ems-Jade-Kanal		Noordpolderzijl	
Fedderwardersiel		Rozema	
Harlesiel	Weser	Spijksterpompen	
Jade-Wapeler_Siel		Harlingen	
Maade_Siel		Miedema	Lake IJssel East
Neuharlingersiel		Ropta	
Schleuse_Hooksiel		DeSchans	
Schweiburger_Siel		Dijkmanshuizen	
Vareler_Siel		Eierland	
Wanger_Siel		Helsdeur	Lake IJssel West
		Krassekeet	
		Oostoever	
		PrinsHendrik	
		Zandkes	

A1.5 TP concentrations and loads

Also shown in Table A1.1 are the average loads and concentrations of TP over the period 2009-2017. For the current scenario, no reduction in TP concentrations and loads is calculated, with the purpose of keeping the scenario simple by only changing one factor (TN loads). The increase of N:P ratios in Meuse (discharge point Haringvliet) and Rhine (discharge point Nieuwe Waterweg) shows that there has been a substantially larger decrease in TP loads than in TN loads (Fig. A1.2). However, in the other rivers (Scheldt and German rivers) there is no clear trend in N:P ratios, indicating that changes in TP and TN loads show similar magnitudes (Table A1.3).

Table A1.3. Reduction in TN and TP loads for the main discharge points, comparing the average load for 2009-2017 with the average load for 1990-1995. Increased loads are indicated in red.

		1990-1995	2009-2017	Reduction
Haringvliet	TN in kton	108	53	-51%
	TP in ton	5438	1423	-74%
Nieuwe Waterweg	TN in kton	212	124	-42%
	TP in ton	14996	4784	-68%
Scheldt	TN in kton	33	17	-48%
	TP in ton	2801	1226	-56%
Eider	TN in kton	4	4	14%
	TP in ton	155	180	16%
Elbe	TN in kton	143	98	-31%
	TP in ton	5847	4155	-29%
Ems	TN in kton	28	16	-43%
	TP in ton	794	437	-45%
Weser	TN in kton	73	38	-48%
	TP in ton	2607	1535	-41%
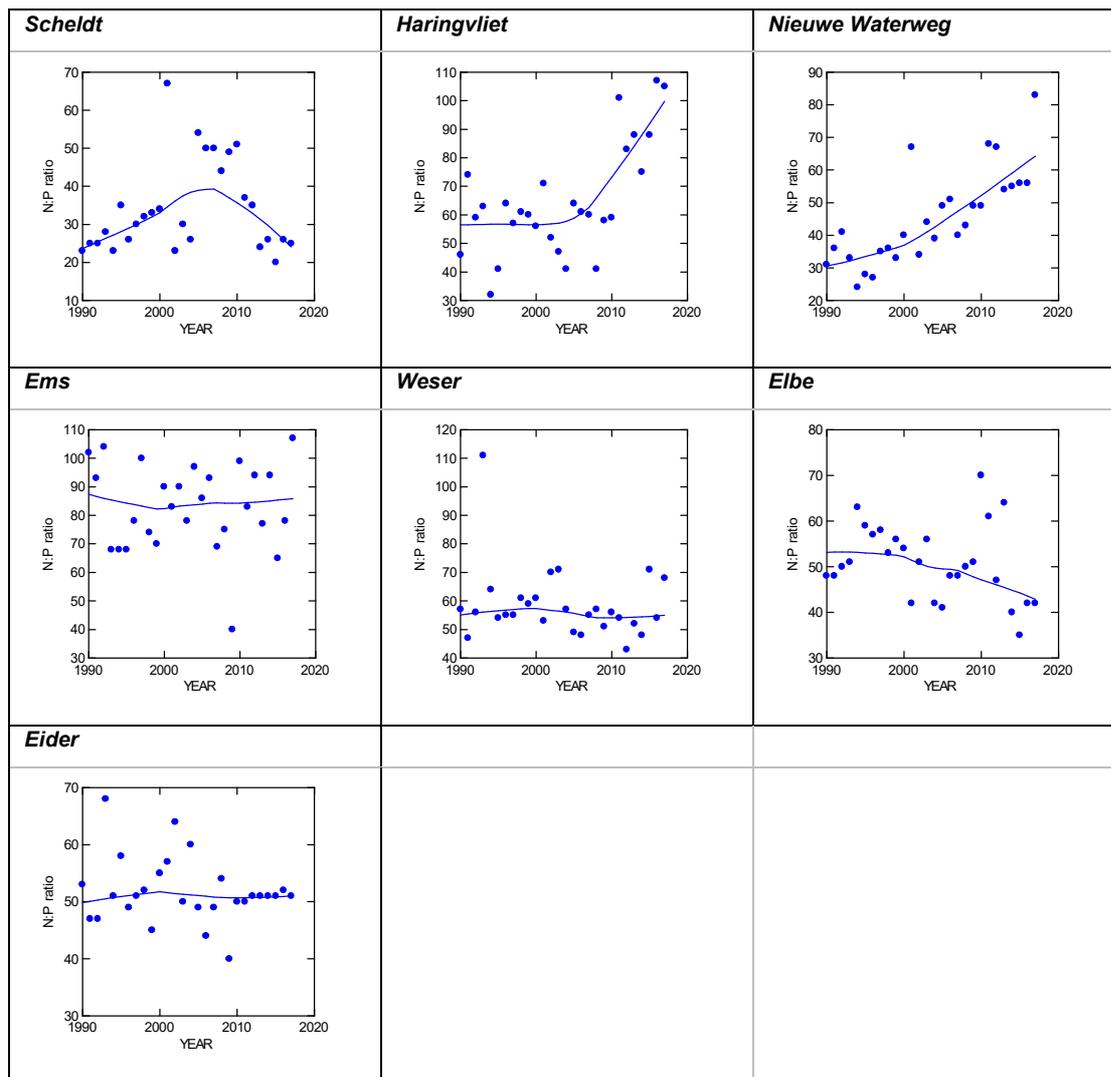
Figure A1.2. N:P ratios of the annual average river loads in the period 1990-2017.

A.1.6 Proposal for application of the 2.8 mg/l scenario

- Apply a specific reduction percentage per year per river (data shown in Table A1.4) to TN and DIN loads of Dutch and German rivers.
- If current concentrations are already below the target value of 2.8 mg/l, no reduction is applied, and current loads are used.
- For the river Scheldt we apply the reduction percentage from Table A1.4 as well.
    - formally the management objective for the river part of the Scheldt are not the competence of the Netherlands; the Belgium target of 2.5 mg/l for the summer mean is the same as the Dutch target (CIW 2016).
- Keep at current levels:
    - TN and DIN loads of all rivers from other countries entering the North Sea.
    - TP and DIP loads of all rivers.
    - Atmospheric nitrogen deposition.

A.1.7 Expected results

The model application with this 2.8-scenario will provide estimates of nutrient and Chl concentrations in the Wadden Sea and coastal waters of the North Sea after achieving the current management objective for TN in Dutch and German rivers.

Table A1.4. Annual average discharge (Q), TN load and concentration, required reduction for the 2.8 mg/l target and concentration after reduction for Dutch and German rivers in the years 2009-2017.

Country	River	Year	Q (m³/s)	TN load (kton)	TN (mg/l)	Reduction (% TN)	TN (mg/l) after reduction
DE	Arlau	2009	2	0.3	4.06	31	2.80
DE	Arlau	2010	2	0.4	4.56	39	2.80
DE	Arlau	2011	2	0.4	4.58	39	2.80
DE	Arlau	2012	4	0.7	4.59	39	2.80
DE	Arlau	2013	4	0.7	4.58	39	2.80
DE	Arlau	2014	4	0.7	4.58	39	2.80
DE	Arlau	2015	4	0.7	4.58	39	2.80
DE	Arlau	2016	3	0.5	4.59	39	2.80
DE	Arlau	2017	4	0.7	4.58	39	2.80
DE	Bongsieler Kanal	2009	7	0.9	4.04	31	2.80
DE	Bongsieler Kanal	2010	7	1.1	4.55	38	2.80
DE	Bongsieler Kanal	2011	7	1.1	4.57	39	2.80
DE	Bongsieler Kanal	2012	11	1.7	4.59	39	2.80
DE	Bongsieler Kanal	2013	13	2.0	4.58	39	2.80
DE	Bongsieler Kanal	2014	13	2.0	4.58	39	2.80
DE	Bongsieler Kanal	2015	13	2.0	4.58	39	2.80
DE	Bongsieler Kanal	2016	9	1.4	4.59	39	2.80
DE	Bongsieler Kanal	2017	11	1.7	4.58	39	2.80
DE	Eider	2009	17	2.3	4.16	33	2.80
DE	Eider	2010	17	2.6	4.66	40	2.80
DE	Eider	2011	17	2.6	4.68	40	2.80
DE	Eider	2012	30	4.8	4.70	40	2.80
DE	Eider	2013	33	5.2	4.59	39	2.80
DE	Eider	2014	33	5.2	4.59	39	2.80
DE	Eider	2015	33	5.2	4.59	39	2.80
DE	Eider	2016	22	3.7	4.75	41	2.80
DE	Eider	2017	30	4.8	4.70	40	2.80
DE	Elbe	2009	770	90.3	3.34	16	2.80
DE	Elbe	2010	1193	168.0	4.39	36	2.80
DE	Elbe	2011	1008	134.7	3.51	20	2.80
DE	Elbe	2012	770	80.8	2.93	4	2.80
DE	Elbe	2013	1190	147.4	3.74	25	2.80
DE	Elbe	2014	575	57.9	3.07	9	2.80

Country	River	Year	Q (m³/s)	TN load (kton)	TN (mg/l)	Reduction (% TN)	TN (mg/l) after reduction
DE	Elbe	2015	584	58.9	2.89	3	2.80
DE	Elbe	2016	591	65.1	3.17	12	2.80
DE	Elbe	2017	693	77.1	3.29	15	2.80
DE	Ems	2009	82	9.0	3.34	16	2.80
DE	Ems	2010	108	21.9	5.58	50	2.80
DE	Ems	2011	88	16.4	4.64	40	2.80
DE	Ems	2012	87	15.4	4.53	38	2.80
DE	Ems	2013	77	13.7	4.63	40	2.80
DE	Ems	2014	81	13.4	4.98	44	2.80
DE	Ems	2015	108	19.1	4.85	42	2.80
DE	Ems	2016	102	16.9	4.53	38	2.80
DE	Ems	2017	88	14.6	4.49	38	2.80
DE	Miele, Warwerorter Kanal	2009	6	0.7	4.06	31	2.80
DE	Miele, Warwerorter Kanal	2010	6	0.9	4.56	39	2.80
DE	Miele, Warwerorter Kanal	2011	6	0.9	4.58	39	2.80
DE	Miele, Warwerorter Kanal	2012	7	1.1	4.60	39	2.80
DE	Miele, Warwerorter Kanal	2013	8	1.2	4.58	39	2.80
DE	Miele, Warwerorter Kanal	2014	8	1.2	4.58	39	2.80
DE	Miele, Warwerorter Kanal	2015	8	1.2	4.58	39	2.80
DE	Miele, Warwerorter Kanal	2016	6	1.0	4.59	39	2.80
DE	Miele, Warwerorter Kanal	2017	7	1.1	4.59	39	2.80
DE	Weser	2009	281	36.9	3.73	25	2.80
DE	Weser	2010	366	55.0	4.30	35	2.80
DE	Weser	2011	288	37.1	3.29	15	2.80
DE	Weser	2012	265	30.4	3.25	14	2.80
DE	Weser	2013	332	46.3	4.07	31	2.80
DE	Weser	2014	231	25.5	3.38	17	2.80
DE	Weser	2015	276	38.2	3.82	27	2.80
DE	Weser	2016	268	32.6	3.32	16	2.80
DE	Weser	2017	319	39.6	3.71	25	2.80
NL	Haringvliet	2009	351	37.2	2.83	1	2.80
NL	Haringvliet	2010	531	56.2	2.89	3	2.80
NL	Haringvliet	2011	394	44.7	2.90	4	2.80
NL	Haringvliet	2012	547	54.1	2.78	0	2.78
NL	Haringvliet	2013	821	81.8	2.93	4	2.80
NL	Haringvliet	2014	374	34.2	2.75	0	2.75
NL	Haringvliet	2015	434	45.5	2.84	1	2.80
NL	Haringvliet	2016	749	87.6	3.37	17	2.80
NL	Haringvliet	2017	329	39.2	2.96	5	2.80
NL	IJsselmeer-oost	2009	171	12.7	2.22	0	2.22

Country	River	Year	Q (m³/s)	TN load (kton)	TN (mg/l)	Reduction (% TN)	TN (mg/l) after reduction
NL	IJsselmeer-oost	2010	251	24.2	2.91	4	2.80
NL	IJsselmeer-oost	2011	203	19.0	2.76	0	2.76
NL	IJsselmeer-oost	2012	251	19.5	2.36	0	2.36
NL	IJsselmeer-oost	2013	270	23.2	2.56	0	2.56
NL	IJsselmeer-oost	2014	229	16.5	2.21	0	2.21
NL	IJsselmeer-oost	2015	257	23.5	2.67	0	2.67
NL	IJsselmeer-oost	2016	244	20.1	2.32	0	2.32
NL	IJsselmeer-oost	2017	197	13.1	1.89	0	1.89
NL	IJsselmeer-west	2009	246	18.4	2.22	0	2.22
NL	IJsselmeer-west	2010	290	27.8	2.92	4	2.80
NL	IJsselmeer-west	2011	236	22.9	2.75	0	2.75
NL	IJsselmeer-west	2012	275	21.4	2.35	0	2.35
NL	IJsselmeer-west	2013	298	25.8	2.59	0	2.59
NL	IJsselmeer-west	2014	246	17.6	2.20	0	2.20
NL	IJsselmeer-west	2015	267	24.5	2.61	0	2.61
NL	IJsselmeer-west	2016	231	20.6	2.43	0	2.43
NL	IJsselmeer-west	2017	241	16.4	1.90	0	1.90
NL	Nieuwe Waterweg	2009	1417	135.1	2.89	3	2.80
NL	Nieuwe Waterweg	2010	1554	157.0	3.06	9	2.80
NL	Nieuwe Waterweg	2011	1260	120.9	2.75	0	2.75
NL	Nieuwe Waterweg	2012	1261	110.2	2.60	0	2.60
NL	Nieuwe Waterweg	2013	1532	137.8	2.72	0	2.72
NL	Nieuwe Waterweg	2014	1352	112.7	2.56	0	2.56
NL	Nieuwe Waterweg	2015	1271	110.6	2.55	0	2.55
NL	Nieuwe Waterweg	2016	1319	118.3	2.73	0	2.73
NL	Nieuwe Waterweg	2017	1188	111.0	2.83	1	2.80
NL	Noordzeekanaal	2009	78	5.6	2.20	0	2.20
NL	Noordzeekanaal	2010	95	7.2	2.40	0	2.40
NL	Noordzeekanaal	2011	85	6.5	2.41	0	2.41
NL	Noordzeekanaal	2012	82	5.8	2.23	0	2.23
NL	Noordzeekanaal	2013	79	5.6	2.21	0	2.21
NL	Noordzeekanaal	2014	82	5.5	2.08	0	2.08
NL	Noordzeekanaal	2015	85	5.8	2.10	0	2.10
NL	Noordzeekanaal	2016	57	5.2	2.60	0	2.60
NL	Noordzeekanaal	2017	81	6.3	2.39	0	2.39
NL	Scheldt	2009	115	16.4	3.98	30	2.80
NL	Scheldt	2010	129	18.9	4.30	35	2.80
NL	Scheldt	2011	106	14.8	3.90	28	2.80
NL	Scheldt	2012	137	19.5	4.12	32	2.80
NL	Scheldt	2013	143	21.1	4.23	34	2.80

Country	River	Year	Q (m³/s)	TN load (kton)	TN (mg/l)	Reduction (% TN)	TN (mg/l) after reduction
NL	Scheldt	2014	122	14.8	3.54	21	2.80
NL	Scheldt	2015	115	13.7	2.94	5	2.80
NL	Scheldt	2016	143	19.3	3.95	29	2.80
NL	Scheldt	2017	94	10.9	3.30	15	2.80
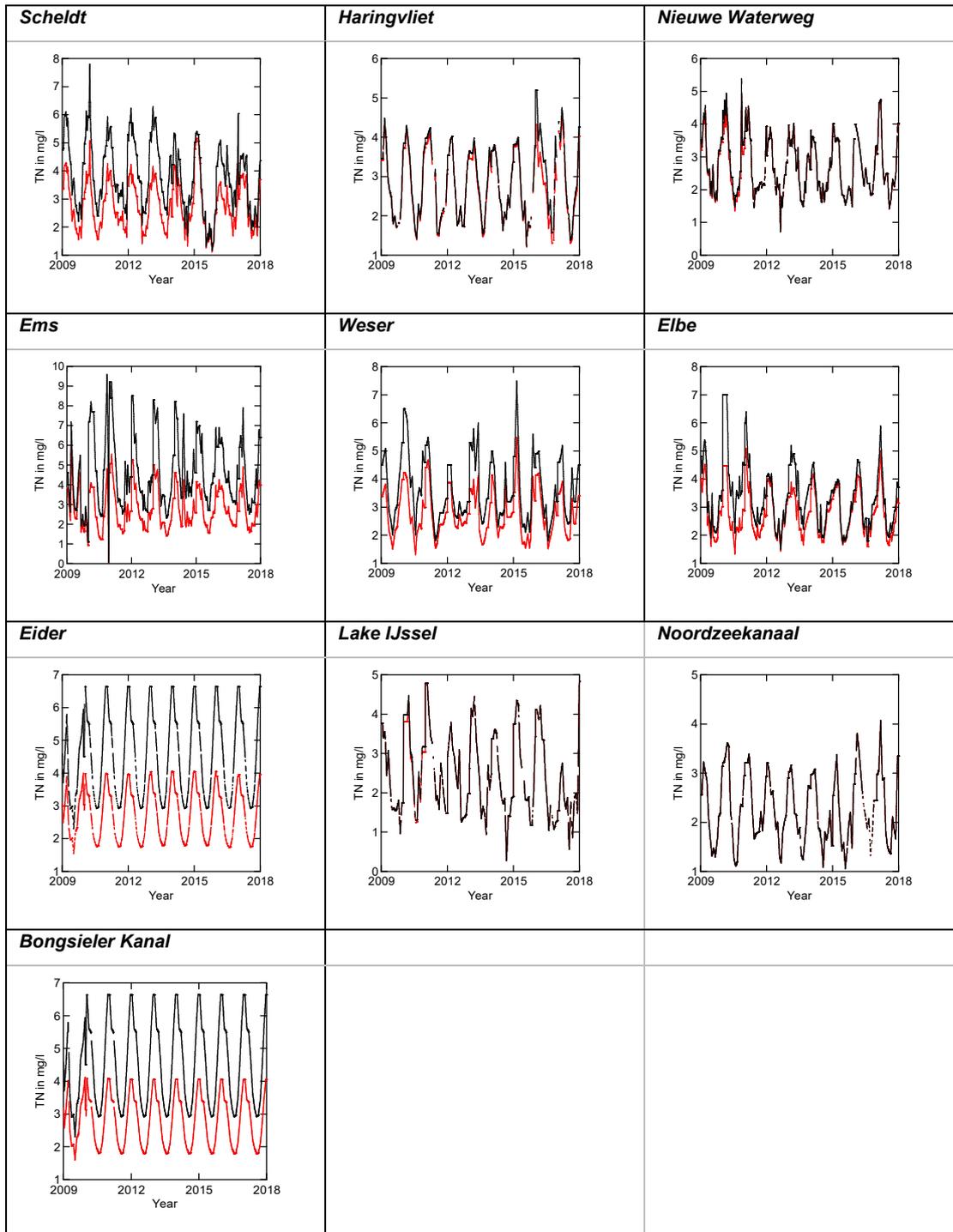
Figure A1.3. TN concentrations in the twelve largest Dutch and German rivers. Current concentrations are shown in black and concentrations after reduction to the annual average 2.8 mg/l are shown in red.
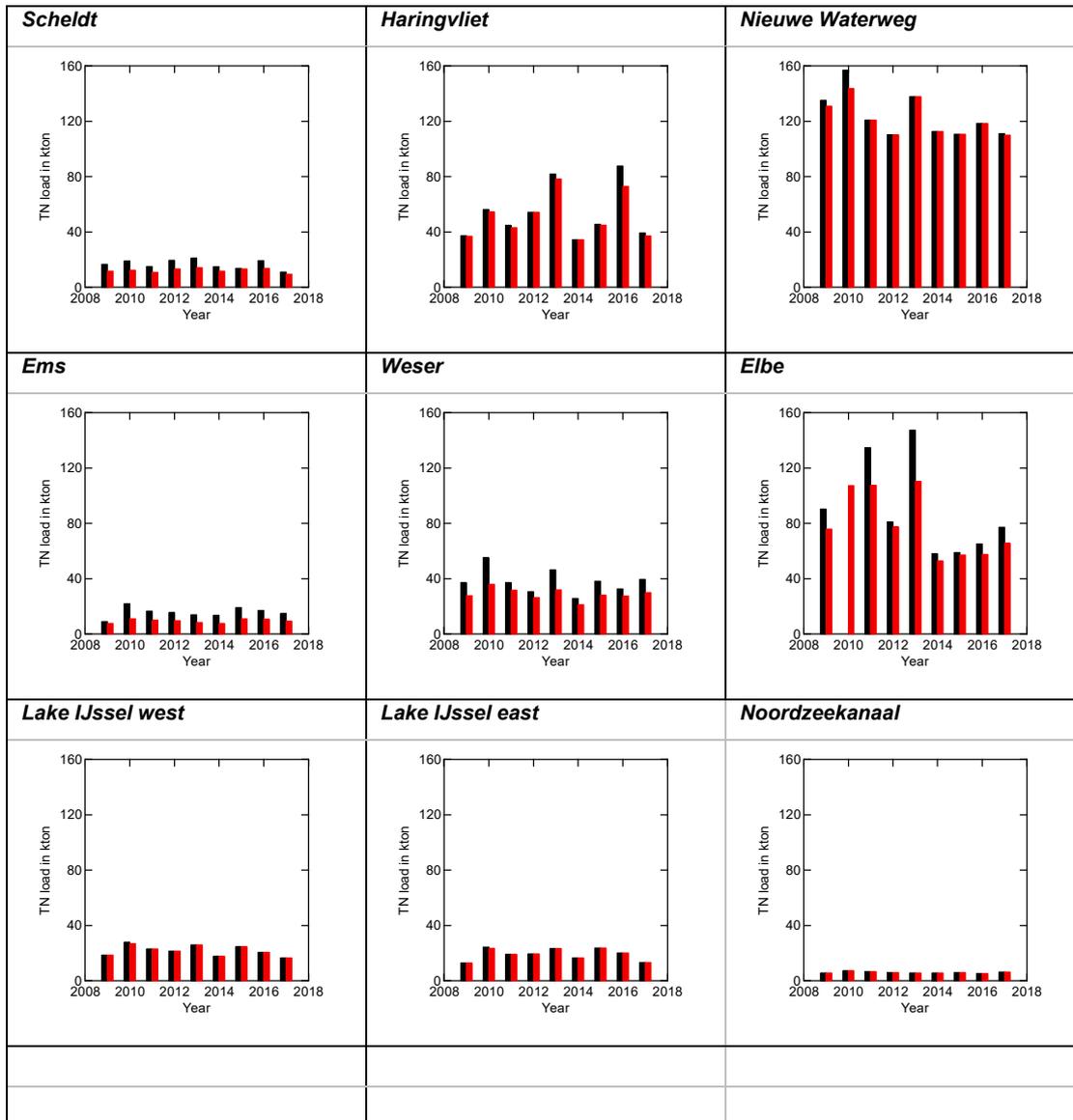
Figure A1.4. Annual TN loads in the nine largest Dutch and German discharge points. Current loads are shown in black and loads after reduction to the annual average 2.8 mg/l are shown in red.

Figure A1.5. Map showing the location of the small discharge points discussed in section A1.4 (red) and the large rivers and discharge points from section A1.3 (yellow).

Annex 2: Calculation of the annual average TN concentration (OSPAR calculation)

In comparison to the previous (Annex 1) calculation for the 2.8 TN mg/l scenario there is an alternative way to calculate the reduction factors for the riverine nutrient loads needed for the simulation. This method is based on the yearly TN load related to yearly discharge volume. The calculation was used to derive estimates of eutrophication target concentrations for OSPAR ICG-EMO, therefore the method is termed "OSPAR calculation" and is described in detail below.

In order to translate the TN target concentration of 2.8 mg/l into the corresponding TN load reduction for the different rivers, the time series of actual TN loads is translated into time series of TN concentrations using Equation (1):

$$CONC(X) = \frac{LOAD(X)}{Q} \qquad (1)$$

where *CONC* and *LOAD* represent the concentration and load of the considered quantity. Q represents the corresponding freshwater discharge.

The river data files provide daily data of freshwater discharge Q and loads of TN. In the case of missing Q, the concentration cannot be calculated. As the time series of the different loads may contain data gaps or even lack completely, a simple gap filling is applied.

Subsequently, the discharge-weighted annual average TN concentration is calculated for each year of the period of interest. According to Kerimoglu et al. (2018), the reduction factor for each year is then calculated as:

$$f_{red}(Y) = \frac{\overline{CONC_{TN}(Y)}}{CONC_{TN,target}} \qquad (2)$$

where $CONC_{TN}$ describes the discharge-weighted annually averaged TN concentration for year Y, and $CONC_{TN, target}$ represents the TN target concentration of 2.8 mg/l. The time series of reduced TN concentrations and loads is then calculated by multiplying the original TN time series with the annual reduction factor f_{red}.

In contrast, the "MEMO calculation" based on the annual average TN concentration uses only the known TN concentrations, see Equation (3).

$$\overline{CONC_{TN}(Y)} = \frac{\sum_1^x CONC_{TN}(X)}{X} \qquad (3)$$

Figure A2.1 provides an overview of the resulting loads (in % of current loads) based on the different methods for Rhine, Weser and Elbe. One can see that the Annex 1 method (orange bars) generally results in higher loads than the OSPAR based calculation of this Annex 2 (black bars). Note that both methods follow the interannual variation in the loads.

Table A2.1 gives more detailed information for a selected number of rivers is presented for the mean TN concentration and the TN load for the Current state (CS) simulation and the two calculation methods for the 2.8-scenario.
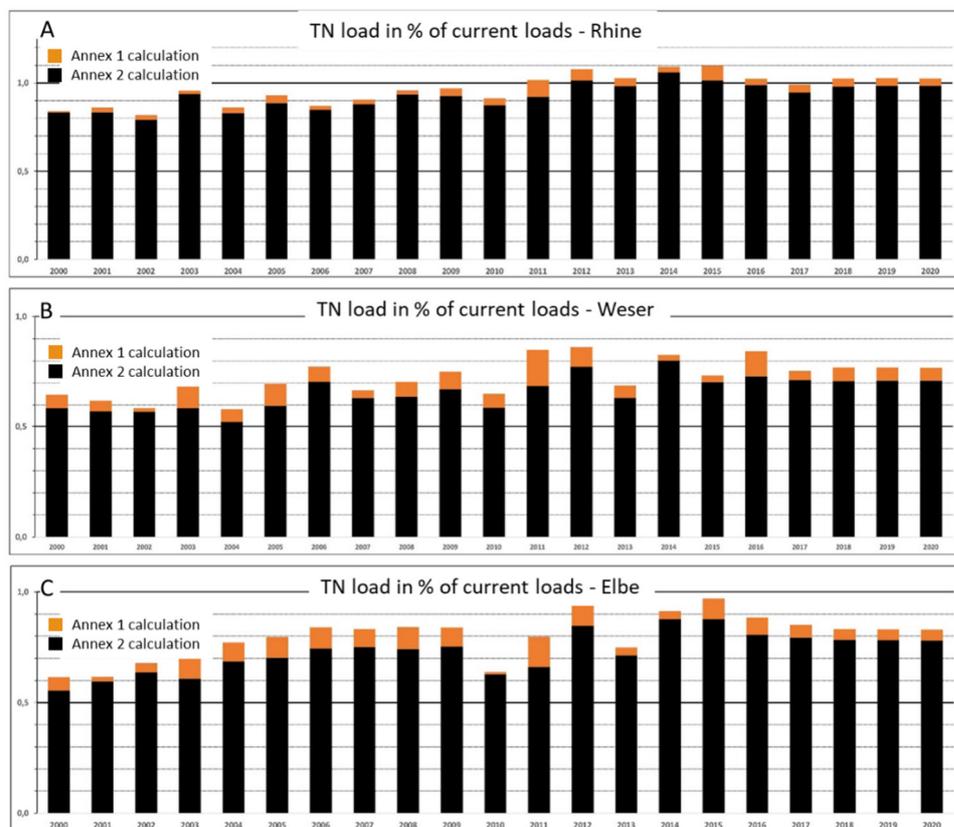
Figure A2.1. Comparison of TN loads in % of current (2017) loads based on the calculation methods described in Annex 1 and the "OSPAR calculation" described in this Annex, for the rivers A) Rhine, B) Weser and C) Elbe.

Table A2.1. Comparison of TN concentrations and loads in the year 2017 and TN loads in the 2.8 mg/l scenario according to the calculation methods of Annex 1 and Annex 2.

River	Current state (2017)		WFD threshold TN (mg/l)	TN 2.8 mg/l scenario			
	TN concentration (mg/l)	TN Load (kton/year)		TN Load method Annex 1 (kton/year)	TN Load method Annex 1 In % of current loads	TN Load method Annex 2 (kton/year)	TN Load method Annex 2 In % of current loads
Scheldt	3.30	10.9	2.8	9.3	85	8.3	76
Rhine	2.83	111.0	2.8	109.9	99	105.4	95
Ems	4.50	14.6	2.8	9.1	62	7.8	53
Weser	3.71	39.6	2.8	29.7	75	28.2	71
Elbe	3.29	77.1	2.8	65.6	85	61.2	79
Eider	4.70	4.8	2.8	2.9	59	2.7	56

Annex 3: Phytoplankton monitoring data

Table A3.1. Results of the linear mixed effect model, analysing the effects of environmental factors on phytoplankton biomass (carbon and chlorophyll *a*), considering "station" as a random effect. The different outputs between NL and DE have been highlighted in grey. Data input: *annual median.*

	Carbon (LN µgL⁻¹)						Chl (LN µgL⁻¹)					
	all		NL		DE		all		NL		DE	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	-1.782	0.721	-6.522	0.239	10.353	0.167	-5.381	**<0.001**	-6.044	**<0.001**	1.644	0.706
LN TN	-0.264	0.501	0.089	0.843	-0.122	0.832	0.462	**<0.001**	0.517	**<0.001**	0.002	0.995
LN NP	0.448	0.173	0.216	0.547	0.366	0.587	-0.121	0.072	-0.116	0.100	0.167	0.519
LN Si	-0.004	0.981	0.180	0.278	-0.912	**0.023**	-0.157	**<0.001**	-0.162	**<0.001**	-0.033	0.862
LN SPM	0.235	0.114	0.519	**0.003**	0.133	0.376	0.117	**0.003**	0.110	**0.013**	0.163	**0.025**
Salinity	0.128	**0.009**	0.105	**0.015**	-0.137	**0.020**	0.007	0.560	0.008	0.508	0.014	0.462
T	-0.046	0.206	-0.043	0.347	0.116	**0.016**	0.050	**<0.001**	0.052	**<0.001**	0.031	0.138
pH	0.449	0.392	0.967	0.103	-0.323	0.713	0.632	**<0.001**	0.671	**<0.001**	-0.163	0.752
Random Effects												
σ²	0.41		0.50		0.13		0.07		0.08		0.05	
τ00	3.16 StationID		0.09 StationID		0.12 StationID		0.05 StationID		0.07 StationID		0.00 StationID	
ICC	0.88		0.16		0.49		0.43		0.47			
N	13 StationID		9 StationID		4 StationID		13 StationID		9 StationID		4 StationID	
Obs	182		148		34		338		303		35	
Marg. R² / Cond. R²	0.117 / 0.897		0.265 / 0.382		0.560 / 0.775		0.309 / 0.605		0.321 / 0.640		0.244 / NA	

Table A3.2. Results of the linear mixed effect model, analysing the effects of environmental factors on phytoplankton standing diversity (annual richness and ENS), considering "station" as a random effect. The different outputs between NL and DE have been highlighted in grey. Data input: *annual median (scaled values).*

	Annual Richness						Annual ENS					
	all		NL		DE		all		NL		DE	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	0.148	0.515	0.412	0.055	-0.482	0.178	0.124	0.580	-0.137	0.434	0.678	0.339
TN	0.450	0.120	0.386	0.228	1.220	0.126	0.553	0.133	0.737	**0.038**	1.298	0.553
NP	-0.274	**0.001**	-0.300	**0.001**	-0.089	0.692	-0.064	0.536	-0.090	0.358	0.292	0.641
Si	-0.875	**0.004**	-0.898	**0.003**	-1.634	**0.042**	-1.068	**0.002**	-0.751	**0.010**	-2.398	0.262
SPM	-0.242	**0.017**	-0.339	**0.015**	-0.032	0.756	0.040	0.753	-0.351	**0.020**	0.446	0.108
Salinity	-0.220	0.340	-0.391	0.156	0.834	**0.021**	-0.464	0.101	-0.355	0.206	0.264	0.705
T	0.018	0.798	-0.006	0.940	-0.049	0.633	0.166	0.060	0.165	0.085	0.194	0.399
pH	-0.178	**0.011**	-0.183	**0.023**	-0.299	**0.017**	0.119	0.184	0.067	0.451	0.013	0.969
Random Effects												
σ²	0.36		0.40		0.11		0.60		0.51		0.92	
τ00	0.56 StationID		0.33 StationID		0.30 StationID		0.47 StationID		0.17 StationID		0.45 StationID	
ICC	0.61		0.45		0.73		0.44		0.25		0.33	
N	13 StationID		9 StationID		4 StationID		13 StationID		9 StationID		4 StationID	
Obs	182		148		34		182		148		34	
Marg. R² / Cond. R²	0.258 / 0.713		0.346 / 0.639		0.341 / 0.825		0.152 / 0.528		0.169 / 0.380		0.222 / 0.476	

# Annex 4: Implementation of Phosphate remineralization in the FSK-model

Phosphorus burial into the sediment is a natural process. It can on the one hand reduce the bioavailable phosphorus in the water column (Asmala et al. 2017), but on the other hand, can change the bed-sediment from a phosphorus sink to a new phosphorus source by means of remineralization of buried organic matter. As an example, the reduction of land-based phosphorus inputs into the Scheldt estuary did not lead to a significant reduction of phytoplankton biomass as a consequence of nutrients reduction (Di Pane et al. 2022, Gypens et al. 2008). Therefore, benthic phosphate-remineralization (P-remineralization) in sediment and its transport into the water column could be represented as an important nutrient source, which provides the sustaining need of phytoplankton growth (Gypens et al. 2008). As a result, understanding the biogeochemical processes associated with the P-remineralisation has attracted the interest of researchers in recent decades.

To assess and predict the eutrophication problem, the remineralization process and the nutrient exchange between water column and sediment layer, it is necessary to have a reasonable estimation of the available nutrients. Since oxygen is involved in these processes, special attention has to be paid to the Wadden Sea tidal flats, which become fully aerated twice a day. However, due to the complexity of nutrient exchange between water and sediment, the majority of existing biogeochemical models do not include these additional processes (Gypens et al. 2008).

In the Interreg research project, the FSK-model, due to its higher spatial resolution, detailed wetting and drying of tidal flats, and capability of sediment transport modelling was selected to investigate the phosphate remineralization for the German and Dutch Wadden Sea with focus on the Ems estuary. To this end, the analytical approach of Gypens et al. (2008) was implemented into the open source Delft3D-Water Quality (DELWAQ) numerical model framework.

Model approach

The main source for the buried phosphorus in the implemented process in the FSK-model is the deposited particulate organic matter (POM) on the sediment bed. The transport of the deposited organic matter on the sediment surface into the sediment layer is modelled through steady state as advection-diffusion transport equations (ADTE) described by Gypens et al. (2008). The decomposition of the buried organic matter within the sediment layer is represented by means of a sink term for organic matter in its transport equation and as a source term in the transport equation of other processes, apart from oxygen. The modelled substances by means of the steady-state ADTE are POM, oxygen (O_2), nitrate (NO_3), ammonium (NH_4), phosphate (PO_4) and iron (Fe)-bounded phosphate as illustrated in Figure A4.1.
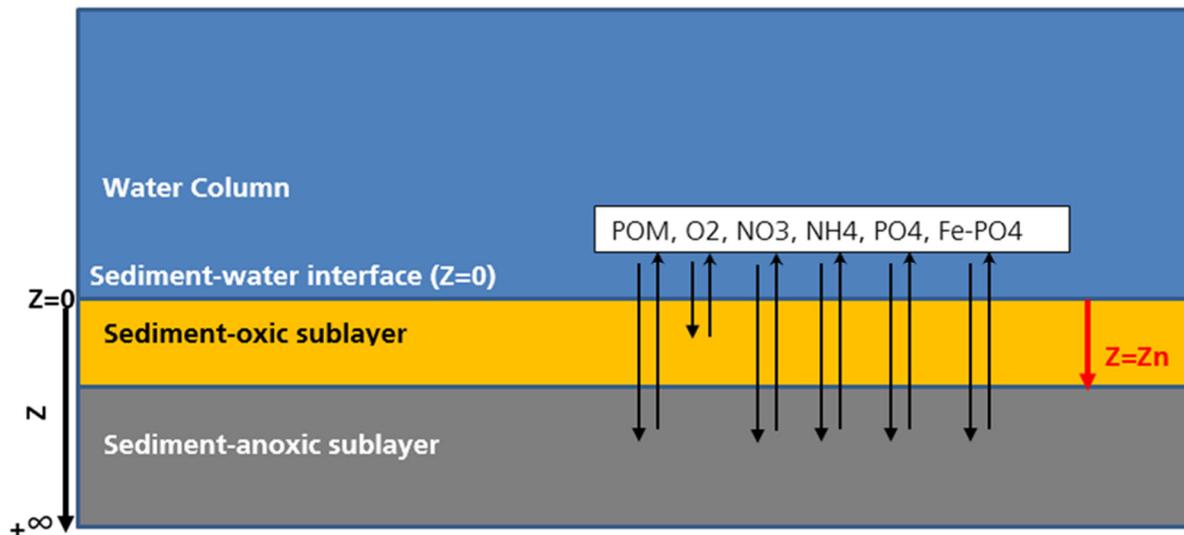
Figure A 4.1. The transported substances in the FSK-model are: POM, O_2, NO_3, NH_4, PO_4 and formed Fe-PO_4. $0 < Z \leq Z_n$ is the oxic layer with $[O_2] > 0$ and $Z_n < Z < +\infty$ is the anoxic layer with $[O_2] = 0$ ([ ] means concentration).

The sediment bed in the developed module formulation for phosphate remineralization is composed of two sublayers neighboring at depth z_n. The transport equations within these differ with regard to the available substances in these sublayers, especially oxygen. The upper sublayer starting at the sediment-water (S-W) interface is the oxic sublayer and the lower one is considered anoxic. Therefore, finding the oxic-anoxic boundary, which is referred to in literature as penetration depth (z_n) is the deciding task for determining the substances concentration in the sediment sublayers.

Gypens et al. (2008) studied the aerobic respiration, nitrification and re-oxidation of oxygen demand units within the oxic layer, which result in the anoxic sublayer and are transported into the oxic layer. These processes include the oxygen consumption, and therefore the oxygen transport equation includes the sinks corresponding to these processes.

The penetration depth is determined through the conservation equation for the oxygen flux at the vicinity of the oxic-anoxic interface. However, the consumption of oxygen through nitrification couples the nitrate transport equation with the oxygen transport equation. Therefore, finding the penetration depth requires nitrate concentration. This coupling is solved using an iteration loop, which results in the determination of the penetration depth z_n.

The diffusion coefficient for organic matter is calculated based on an empirical relationship provided by Gypens et al. (2008), where the mineralized organic carbon is multiplied by a value of 25.4. Likewise, the sedimentation rate in the transport equations is empirically calculated based on the found diffusion coefficient.

After determination of the penetration depth, the oxygen concentration profile for the oxic layer is solved and the nitrate, ammonium, pore water phosphate and iron (Fe)-bounded phosphate concentration for both, oxic and anoxic layers through the analytical equations are calculated.

Finally, the exchange between the water column and the sediment layer is calculated based on the ADTE, where the transport flux is determined by comparing the substance concentrations at the interface (z=0) from the analytical equations solved from the differential equations provided by Gypens et al.(2008) and the corresponding concentrations modelled by DELWAQ in the lowest water layer.

The implemented module in DELWAQ has the possibility to account for the type of bottom sediment. This is achieved by means of an empirical modification of the diffusion coefficients based on the sediment porosity and tortuosity as was presented by Gypens et al. (2008).

Model results

1DV model results

To understand the conceptual behavior of the analytically solved partial differential equations of Gypens et al. (2008), the equations were tested by means of a 1DV model test in MATLAB to evaluate the vertical profile of substances inside the sediment layer. Figure A4.2 shows the vertical profile concentration for modelled substances of a test case with organic matter flux of 0.088 mmol C cm^{-2} yr^{-1}, nitrate and oxygen concentration at the S-W interface of 200 and 20 mmol/m^3, respectively. As can be seen in Figure A4.2, the organic matter concentration is reduced exponentially from the S-W interface (z=0) to the deeper parts of the sediment layer. For this test case, the organic matter reaches the value of approximately zero at the sediment depth of smaller than 2 cm (~1.65 cm). This means, that usually the settled organic matter is present with its greatest concentration at the S-W interface and at the deeper parts of the sediment layer the transported organic matter is significantly reduced.

Nitrate is found from the analytical solution of ADTE equation provided by Gypens et al. (2008). As is shown in Figure A4.2, the nitrate vertical profile concentration behaves differently in the oxic and anoxic sediment sublayers. In the oxic-sublayer ($z<z_n$), the nitrate concentration due to the nitrification of ammonium, which resulted from the decomposition of the transported organic matter in sediment is increased, but within the anoxic layer due to the denitrification of nitrate its concentration is decreased.
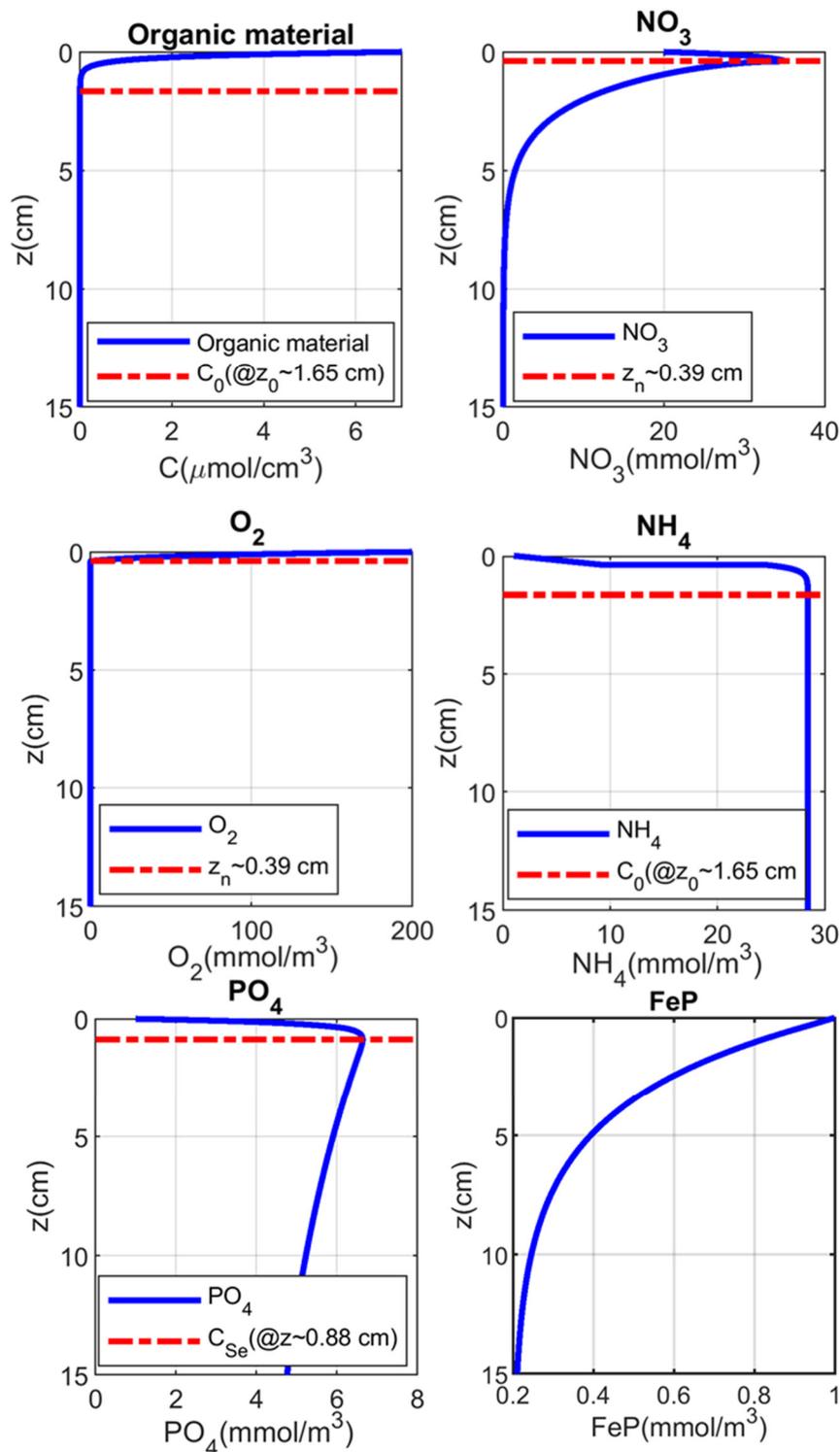
Figure A4.2. Vertical profile of 1DV model results for organic material, nitrate, oxygen, ammonium, pore water phosphate and Fe-bounded phosphate concentration of an organic matter flux of 0.088 mmol C cm^{-2} yr^{-1}. The bottom water oxygen concentration is set to 200 mmol/m^3, and nitrate concentration to 20 mmol/m^3. z_0 is where the organic matter concentration is zero (corresponding to the c_0), z_n is the penetration depth, and C_{Se} is where the sorption of pore water phosphate initiated.

The oxygen concentration profile is found by means of the analytical solution of the differential equation for transport of oxygen in sediment layer. It shows a reduced concentration profile inside the sediment layer and defines the anoxic sublayer, where the concentration reaches zero. The reduction of oxygen is explained through its consumption for aerobic decomposition of organic matter and nitrification of ammonium produced after remineralization of organic matter. In the nitrification

process, one mol of produced ammonium needs two mol of oxygen for nitrification, therefore, the consumption of oxygen happens with higher gradient compared to the profile gradient of other substances. The oxygen penetration depth for this test case is smaller than 1 cm (~0.39 cm). Therefore, the major part of the sediment layer is usually located in the anoxic condition and the oxic part of sediment layer is only a short length of upper sediment layer close to the S-W interface.

Ammonium concentration is increased inside the oxic layer due to the aerobic decomposition of organic matter. However, this increase is not continued, due to the reduction of organic matter concentration exponentially inside the sediment layer. Therefore, high increase of ammonium concentration inside the oxic sublayer is the result of available organic matter and oxygen in this sublayer.

Pore water phosphate concentration inside the oxic sediment sublayer is also increased due to the aerobic decomposition of organic matter and due to the limitation of sorption concentration. When the phosphate concentration is higher than the equilibrium concentration for sorption, its increase stops. This concentration change occurred for this test case at a depth smaller than 1 cm (~0.88 cm). This means again, that the maximum concentration of phosphate is usually located inside the oxic sediment sublayer. The reduction of the phosphate concentration due to concentration exceedance from the equilibrium concentration inside the oxic sediment sublayer is also smaller compared to aerobic decomposition, where enough organic matter and oxygen is available. However, the reduction of the organic matter concentration within the anoxic sublayer is greater than the decomposition of organic matter.

The Fe-bound phosphate concentration is described as for organic matter using a decreasing concentration profile inside the oxic sediment layer, because it does not have a source term like pore water phosphate, where the decomposition of organic matter causes an increasing part in its concentration profile. The transport equation of Fe-bound phosphate can be increased, when the pore water phosphate concentration is greater than the equilibrium. However, this increment is modified by multiplication to the sediment porosity and the rate constant for kinetic sorption of pore water phosphate concentration. If this value is greater than the advection and diffusion transported Fe-bound phosphate concentration, then the Fe-bound phosphate concentration like the pore water phosphate concentration also shows an increase before decreasing to the boundary value for Fe-bound phosphate at the deepest part of sediment layer.

3D model results

After evaluation of the processes of phosphate remineralization using the 1DV model, the analytical equations resulted from the solution of the differential equations provided by Gypens et al. (2008) were implemented into the open source Delft3D water Quality (DELWAQ). Before discussing the model results, it should be noted that the analytical solutions imply a steady-state solution within every time step. Moreover, the transport of substances inside the sediment layer due to advection and diffusion is neglected.

The model compares the concentration of substances at the vicinity of the water-sediment interface with the available concentrations inside the sediment layer, which determines the direction of the transport between water and sediment. Figure A4.3 illustrates the model concept for the exchange of substances between water and sediment. The increased substance concentration within the water layer close to the bed is transported by advection and diffusion inside the water column, and therefore

the transport of substances between water and sediment contributes to the concentration of substances in the water body of the model domain.
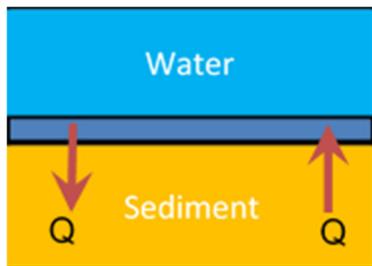
Figure A4.3. Principle sketch of the exchange of substances between sediment and water column. Q is the flux of the substances, which are exchanged between the last layer of the water column and the sediment layer.

Model runs with and without phosphate remineralization were performed. The comparison of these results give a good understanding of the principal effect of the additional phosphate source/sink, though it has to be stressed, that relevant limitations (steady state assumption on time step level, Gypens's empirical coefficients) had to be introduced for the sake of computational power savings. Therefore, we still consider these results as an approximation, which may be improved in the future and which have to be validated by further data. It is though the first time that this formulation for phosphate remineralization is included in this type of water quality model.

Furthermore, to include the formation/release of Fe-bound phosphate (vivianite-P and apatite-P) into the water column and the bed, the module for apatite (APATP) and vivianite (VIVP) were added to the common processes data of the DELWAQ model. The formation of apatite and vivianite in DELWAQ is formulated as a function of available oxygen and dissolved phosphate concentration in the water column. These are coupled with the Fe-bound phosphate equations from the new implementation of the phosphate remineralization.

Figure A4.4 shows the time series of depth-averaged numerical model results for three observation points: Norderney_W_2 (Nney_W_2), Borkum_W_1 (Bork_W_1) and Rottomplate3 (ROTTMPT3) for the time span between April to September 2017. The behavior of the numerical results with inclusion of the phosphate remineralization differs between these three monitoring stations. An increase in the Chl values occurred for Nney_W_2 between around 15.05.2017 and 15.06.2017 However, the concentration for the other times is mostly similar to the model results of the case without inclusion of the phosphate remineralization and sometimes the Chl concentration in the case with phosphate remineralization is smaller.

For Bork_W_1, an increase of the Chl concentration occurred between 01.06.2017 to 15.06.2017 and before this time span, the Chl concentration is almost the same as in the other case. In contrast, the Chl concentration is reduced after inclusion of phosphate remineralization from August to September. The interpretation for this could be the exchange of organic matter between water and sediment. Due to the dependency of the phosphate concentration in the sediment layer on the transported organic matter, the gradient of phosphate release to the water column changes with the rate of transported organic matter into the sediment layer. Moreover, the transport of dissolved phosphate through diffusion and advection from the water column into the sediment layer could lead to the reduction of phosphate in the water column. This process was observed in nature, where the buried phosphate in the sediment layer caused a reduction of phosphate concentration in the water column (Asmala et al. 2017).

At the ROTTMPT3 station, no increase of Chl concentration was observed, where the Chl concentration before July is almost as in the case without a phosphate module. Interestingly, the Chl concentration is reduced after July by inclusion the phosphate remineralization module.

To evaluate the general performance of the implementation of the phosphate remineralization module and interpret the different Chl concentration behavior, concentration of organic matter and phosphate at the vicinity of the S-W interface (z=0 in the analytical solutions of the vertical profile concentration) for abovementioned stations are compared as illustrated in Figure 4.5. As can be seen, all three stations show positive values and an increasing trend of the accumulation of organic matter inside the sediment layer.
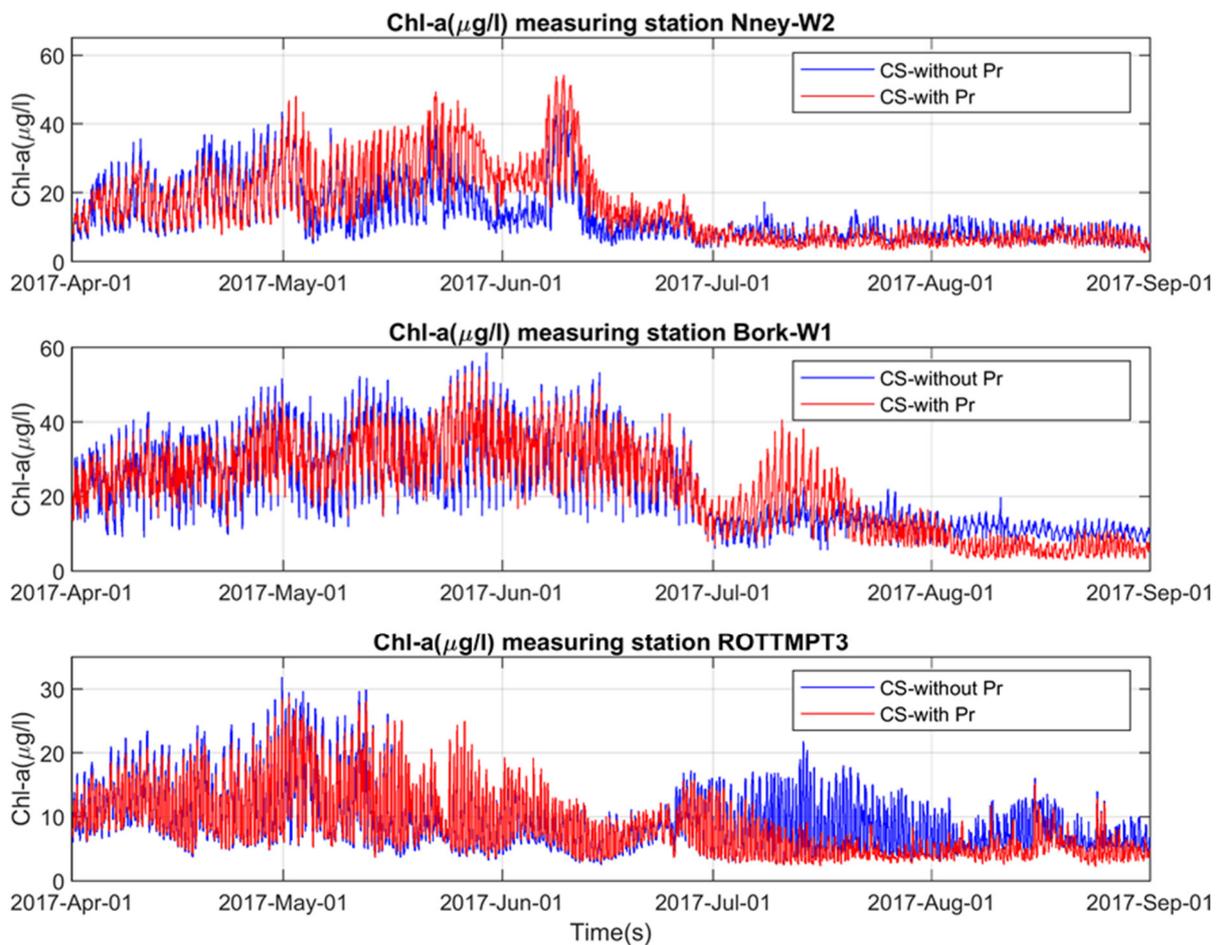
Figure A4.4. Time series of model results for Chl concentration (µg/l) with and without inclusion of the phosphate remineralization (Pr) implementation for the current state (CS: year 2017).
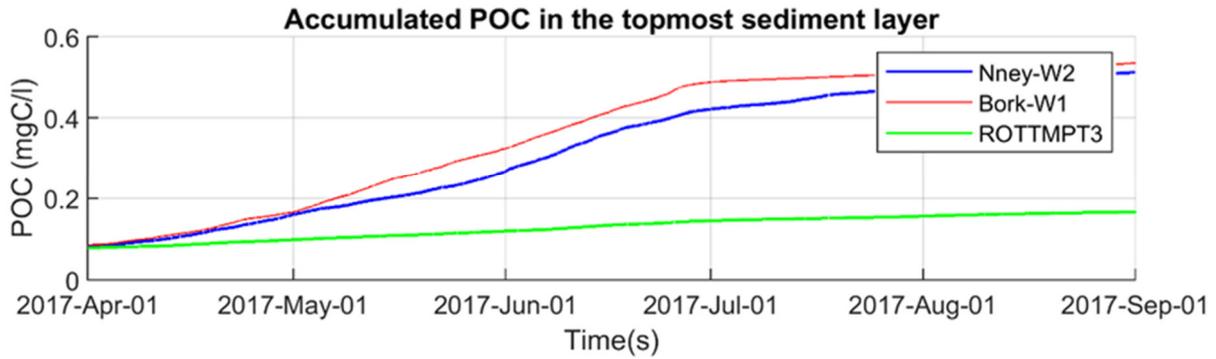
Figure A4.5. Time series of accumulated organic matter in the topmost sediment layer (z=0 in the analytical equations) for Nney_W_2, Bork_W_1, and ROTTMPT3.
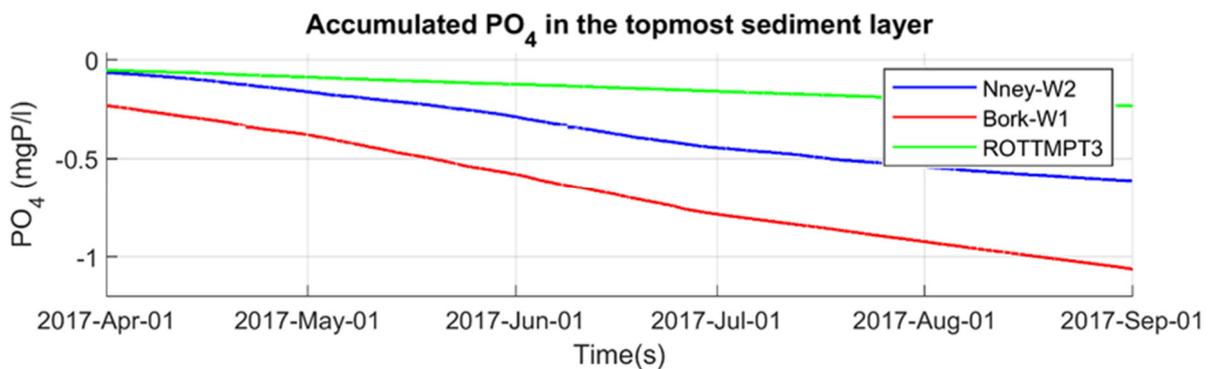
Figure A4.6. Time series of accumulated phosphate in the topmost sediment layer (z=0 in the analytical equations) for Nney_W_2, Bork_W_1, and ROTTMPT3.

Therefore, it is expected that high phosphate concentration due to the remineralization of phosphate in the vicinity of the S-W interface is produced, so that the phosphate concentration in this part of sediment is higher than the phosphate concentration in the water layer close to the bed. Therefore, the direction of phosphate exchange is in this condition always from the sediment layer toward the water column. The negative sign of the phosphate concentration in Figure A4.6 represents the transport direction from sediment layer to the water body. This is also in agreement with 1DV model results in Figure A4.2, where the phosphate concentration in the oxic layer has an increasing trend.

ROTTMPT3 in Figure A4.5 has the smallest organic matter concentration compared with Nney_W_2 and Bork_W_1. This is reflected in the accumulated phosphate concentration in Figure A4.6. Therefore, it is estimated that the high part of available phosphate concentration at ROTTMPT3 is from the advection-diffusion of pore water phosphate from the water body into the sediment layer, which leads to the reduction of phosphate concentration in the water layer close to the sediment layer in the case of inclusion of phosphate remineralization.

To sum up, the implemented analytical approach in the 3D-regional scale of DELWAQ revealed that the exchange of phosphate between water and sediment can alter the Chl concentration. The influence of phosphate exchange does not always lead to increasing the phosphate and CHL concentration but can also reduce the phosphate concentration compared to the model cases without inclusion of the phosphate remineralization. This reduction effect of inclusion the phosphate exchange module is a matter of more discussions and requires further investigations. However, the transport of phosphate from water to the sediment layer, which induces the reduction of phosphate concentration in the water column, was also observed and discussed in literature (e.g. Asmala et al. 2017).

References Annex

Asmala, E., Carstensen, J., Conley, D. J., Slomp, C. P., Stadmark, J., & M. Voss (2017). Efficiency of the coastal filter: Nitrogen and phosphorus removal in the Baltic Sea. *Limnology and Oceanography*, *62*(S1), S222-S238.

BLMP (2011). Konzept zur Ableitung von Nährstoffreduzierungszielen in den Flussgebieten Ems, Weser, Elbe und Eider aufgrund von Anforderungen an den ökologischen Zustand der Küstengewässer gemäß Wasserrahmenrichtlinie. Ad-hoc-AG Nährstoffreduzierung des BLMP, Verabschiedet auf der 17. ARGE BLMP am 01.07.2011 und den Lenkungsausschuss der Expertengruppe Meer, 9. Sitzung am 24.3.11. Stand: 01.06.2011, 50 pp. http://www.blmp-online.de/PDF/WRRL/WRRL_Papier_Naehrstoffe.pdf

CIW (2016). Stroomgebiedbeheerplan voor de Schelde 2016-2021. Bekkenspecifiek deel Benedenscheldebekken. Coördinatiecommissie Integraal Waterbeleid - CIW, pp. https://www.vlaanderen.be/publicaties/stroomgebiedbeheerplan-voor-de-schelde-2016-2021-bekkenspecifiek-deel-benedenscheldebekken

Di Pane, J., Wiltshire, K. H., McLean, M., Boersma, M., & C.L. Meunier (2022). Environmentally induced functional shifts in phytoplankton and their potential consequences for ecosystem functioning. *Global Change Biology*, *28*(8), 2804-2819.

Fischer, H., Fischer, M., Kuhn, U., Ulrich, A., Ollesch, G., Ritz, S., Trepel, M. & M. Venohr (2014). Empfehlung zur Übertragung flussbürtiger, meeres-ökologischer Reduzierungsziele ins Binnenland LAWA-AO Geschäftsstelle. Sächsisches Staatsministerium für Umwelt und Landwirtschaft, Dresden, LAWA-Arbeitsprogramm Flussgebietsbewirtschaftung. Produktdatenblatt WRRL-2.4.7. Stand 18.06.2014, 17 pp. http://www.wasserblick.net/servlet/is/142651/WRRL_2.4.7_Uebertragung_Reduzierungsziele_Gesamtstickstoff.pdf?command=downloadContent&filename=WRRL_2.4.7_Uebertragung_Reduzierungsziele_Gesamtstickstoff.pdf

Grosse, F. & H.-J. Lenhart (2015). Derivation of WFD reduction scenarios for nitrogen. Memo 30-10-2015, University of Hamburg.

Gypens, N., Lancelot, C., & K. Soetaert (2008). Simple parameterizations for describing N and P diagenetic processes: Application in the North Sea. *Progress in Oceanography*, *76*(1), 89-110.

IKSR/CIPR/ICBR (2021). Internationaal gecoördineerd stroomgebiedbeheerplan 2022-2027 van het internationaal stroomgebieddistrict Rijn/International koordinierter Bewirtschaftungsplan 2022-2027 für die internationale Flussgebietseinheit Rhein. IKSR/CIPR/ICBR, Koblenz, Concept/Entwurf.

Kerimoglu, O., Große, F., Kreus, M., Lenhart, H.-J. & J.E.E. van Beusekom (2018). A model-based projection of historical state of a coastal ecosystem: Relevance of phytoplankton stoichiometry. *Science of The Total Environment* 639: 1311-1323.

van der Molen, D.T., Pot, R., Evers, C.H.M., van Erpen, F.C.J. & L.L.J. van Nieuwerburgh, Eds. (2018). *Referenties en maatlatten voor natuurlijke watertypen voor de Kaderrichtlijn Water 2021-2027*. STOWA, Amersfoort, Rapportnummer 2018-49, 489 pp.

Project participants

NLWKN Brake-Oldenburg: Jürgen Knaack, Lena Rönn, Kerstin Kolbe, Michael Hanslik, Jan Witt

NLWKN Forschungsstelle Küste (FSK): Andreas Wurpts, Gholamreza Shiravani, Cordula Berkenbrink, Tea Behrends, Anne Ritzmann

Rijkswaterstaat: Gerrit Niebeek, Dick As, Laura Tack, Leon van den Berg, Marcel van den Berg, Ellen Farwick, Raven Cammenga

Deltares: Theo Prins, Anouk Blauw, Tineke Troost, Lauriane Vilmin, Lisa Schneider, Annelotte van der Linden

HIFMB: Helmut Hillebrand, Josie Antonucci di Carvalho, Ute Jacob

University of Hamburg: Hermann Lenhart, Daniel Thewes

**Group picture of project participants with stakeholders and guests at the
"Final Project Event" in Groningen October 2022**