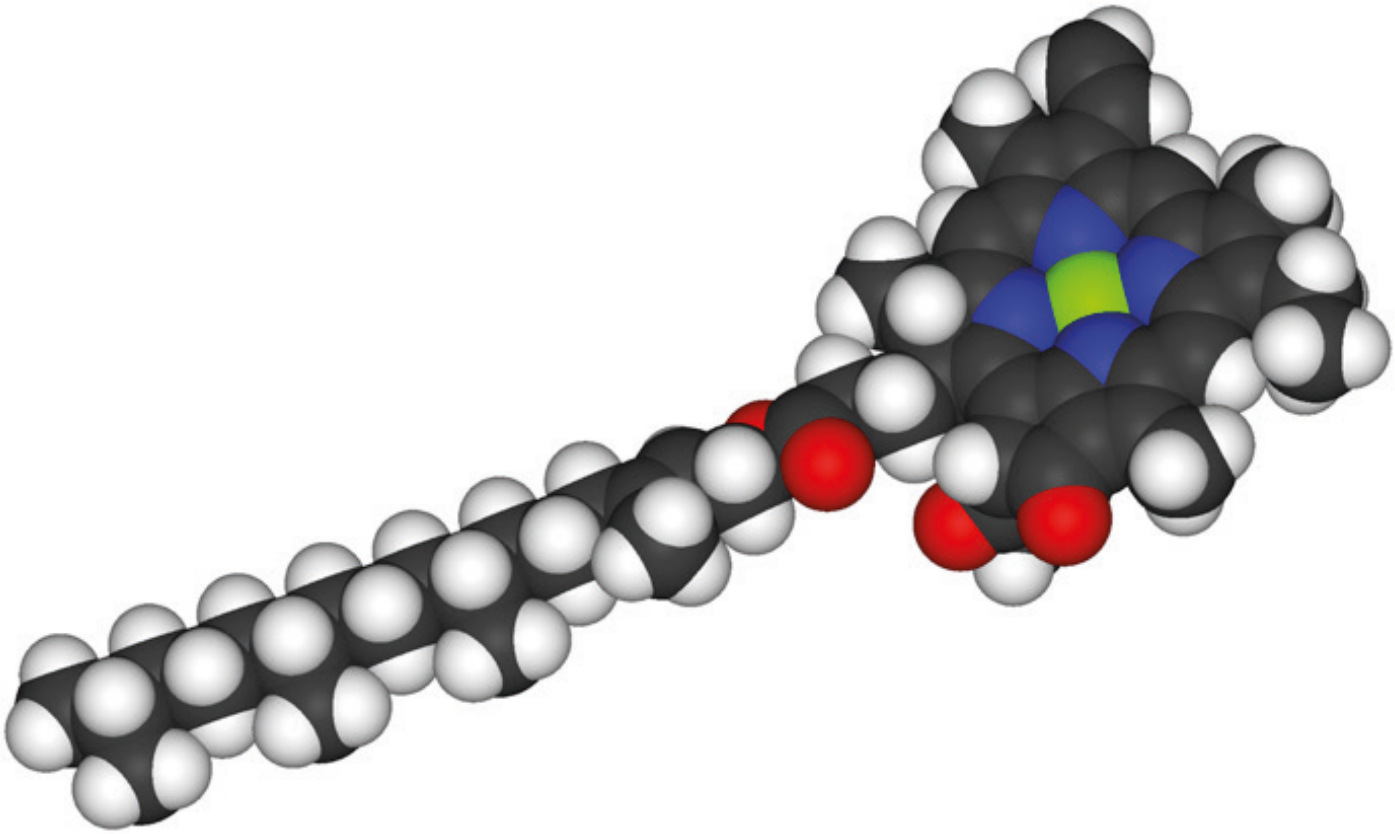
RAPPORT L.NR. 7167-2017 DK7

NIVA Denmark

AARHUS UNIVERSITY
DEPARTMENT OF BIOSCIENCE

DHI

Statistical analyses of chlorophyll-a data sampled by a Ferrybox on the Oslo-Kiel ferry and by the NOVANA programme



NIVA Denmark Water Research

REPORT

Main Office
Gaustadalléen 21
NO-0349 Oslo, Norway
Phone (47) 22 18 51 00
Telefax (47) 22 18 52 00
Internet: www.niva.no

NIVA Region South
Jon Lilletuns vei 3
NO-4879 Grimstad, Norway
Phone (47) 22 18 51 00
Telefax (47) 37 04 45 13

NIVA Region East
Sandvikaveien 59
NO-2312 Ottestad, Norway
Phone (47) 22 18 51 00
Telefax (47) 62 57 66 53

NIVA Region West
Thormøhlensgate 53 D
NO-5006 Bergen Norway
Phone (47) 22 18 51 00
Telefax (47) 55 31 22 14

NIVA Denmark
Ørestads Boulevard 73
DK-2300 Copenhagen, Denmark
Phone (45) 88 96 96 70
www.niva-danmark.dk

Title	Serial number	Date
Statistical analyses of chlorophyll-*a* data sampled by a Ferrybox on the Oslo-Kiel ferry and by the NOVANA programme	7167-2017 DK7	30 January 2017
Author(s) Jesper H. Andersen, NIVA Denmark Water Research Therese Harvey, NIVA Denmark Water Research Emilie Kallenbach, NIVA Denmark Water Research Ciarán Murray, NIVA Denmark Water Research Dag Ø. Hjermann, Norwegian Institute for Water Research (NIVA) Trond Kristiansen, Norwegian Institute for Water Research (NIVA) Kai Sørensen, Norwegian Institute for Water Research (NIVA) Jacob Carstensen, Aarhus University Anders Erichsen, DHI Lars Boye Hansen, DHI	Topic group Eutrophication	Distribution MST + public
	Geographical area Danish Straits, Skagerrak and Kattegat	Printed NIVA

Client(s)	Client's reference
Danish Environmental Protection Agency (MST) (formerly Danish Agency for Water and Nature Management (SVANA))	CFNIE

Summary

This project has analysed a large chlorophyll-*a* data set sampled by the Ferrybox on the ferry between Oslo and Kiel and compared it with data sampled in the Danish NOVANA programme. A partnership consisting of NIVA Denmark Water Research, Norwegian Institute for Water (NIVA), Aarhus University and DHI has: (1) Collated relevant data (chlorophyll-*a*, salinity and temperature) from various sources (i.e. the Ferrybox on the Oslo-Kiel ferry, from the NOVANA programme, from satellite and from modelling activities), (2) assessed uncertainty for temperature, salinity and chlorophyll-a, and (3) transformed Ferrybox-based data to a data product aligned with chlorophyll-*a* data sampled under the NOVANA programme. The derived data product has been quality assured and submitted to the Danish EPA for their use regarding specific activities, e.g. Initial Assessments under the Marine Strategy Framework Directive and Water Framework Directive, Danish reporting to HELCOM and OSPAR and for reporting of NOVANA.

Fire emneord	Four keywords
1. Klorofyl-a	1. Chlorophyll-a
2. Eutrofiering	2. Eutrophication
3. Overvågning og tilstandsvurdering	3. Monitoring and status assessment
4. Ferrybox	4. Ferrybox

Jesper H. Andersen
Chief Scientist

Nikolai Friberg
Research Director

ISBN 978-82-577-6902-4
NIVA Report ISSN 1894-7948

Statistical analyses of chlorophyll-*a* data sampled by a Ferrybox on the Oslo-Kiel ferry and by the NOVANA programme

Cover image from https://commons.wikimedia.org/wiki/ File:Chlorophyll-a-3D-vdW.png and illustrates a space-filling model of the chlorophyll-*a* molecule in 3D. The model shows chlorophyll-*a*'s porphyrin ring with its four nitrogen atoms (blue) surrounding a magnesium atom (green) and with attached side chains. In addition, the model shows the long hydrocarbon tail of chlorophyll-*a*.

Preface

The project has been funded by the Danish Environmental Protection Agency (formerly Agency for Water and Nature Management (SVANA)) with an overarching aim of developing a chlorophyll-*a* data product, based on Ferrybox measurements.

The specific purpose of this project has been to analyse a large chlorophyll-*a* data set (2008-2015) sampled by the Ferrybox on the Oslo-Kiel ferry and to compare this data with data sampled in the Danish NOVANA programme.

As an initial step, a project group consisting of NIVA Denmark Water Research, Norwegian Institute for Water research (NIVA), Aarhus University and DHI has:

- Collated relevant data (chlorophyll-*a*, salinity and temperature) from various sources (i.e. the Ferrybox on the Oslo-Kiel ferry, from the NOVANA programme, from satellite and from modelling activities) and assessed uncertainty for temperature, salinity and chlorophyll-*a*.
- Transformed Ferrybox-data to a data product aligned with chlorophyll-*a* data sampled under the NOVANA programme.

The data product developed by this study has been quality assured and submitted to SVANA for their use regarding specific activities, e.g. Initial Assessments under the Marine Strategy Framework Directive and Water Framework Directive, Danish reporting to HELCOM and OSPAR and for reporting within the NOVANA programme.

Copenhagen, 30 January 2017

Jesper H. Andersen

Table of contents

1 Introduction

The Danish marine waters have been consistently monitored since the mid-70's. The Belt Project (1974-1978) is considered the beginning of the monitoring of the open parts of the Danish Straits (Ærtebjerg et al.1981). Focus was on hydrochemistry (nutrients, salinity, oxygen) and phytoplankton (chlorophyll-*a*, primary production, etc.) and many stations in the current monitoring activities were originally designated as part of the Belt Project.

While the Belt Project focused on the open parts of the Danish Straits, the compliance monitoring carried out by the Regional Authorities (counties) in accordance with environmental protection legislation focused on coastal waters and a broader range of indicators. In addition to hydrochemistry and phytoplankton, a substantial effort was put into monitoring of benthic communities, i.e. benthic macroinvertebrates and submerged aquatic vegetation.

However, the combination of the compliance monitoring carried out by the counties and the monitoring of the open parts of the Danish Straits by the Belt Project and its successor, the National Pollution Monitoring Programme (Ærtebjerg 1981; ATV 1990), did not give complete geographical coverage of the Danish marine environment. To ensure nation-wide marine monitoring, the Danish Aquatic Monitoring Programme 1989-1992, was established as part of the Danish Action Plan on the Aquatic Environment I (VMP I) from 1987 (see Miljøstyrelsen (1989) and Andersen (2012) for details).

New knowledge and improved understanding of key ecological processes led to a revision of the marine monitoring programme. A key improvement of the revised programme, Danish Aquatic Monitoring Programme 1993-1997 (VMP II) (Miljøstyrelsen 1993), was the upgrading of selected hydrochemistry stations to so-called intensive stations with sampling frequencies of 40-52 times per year.

The programme for 1993-1997 was evaluated in 1996 and the subsequent re-designed programme, Danish Aquatic and Assessment Programme (NOVA-2003) (Miljøstyrelsen 2000), was considered a leap forward. Not only were number of stations and frequencies increased, a broad range of new ecological indicators were included, e.g. nutrient pools and fluxes in sediments, and hazardous substances (heavy metals and POPs) in sediments and biota. Modelling was also included in the programme. In a few coastal waters (6 fjords) water and nutrient budgets were modelled. In the open water, transports of nutrients were modelled quarterly. Importantly, the outcome of the new activities, especially the modelling, was included in the annual reporting. For details about NOVA-2003, please see Miljøstyrelsen (2000).

The NOVANA 2004-2007 programme (Svendsen et al. 2005) emerged from an evaluation of NOVA-2003 and the need for prioritization between sub-programmes, in particular more focus on biodiversity and less focus on surface waters. 'Biodiversity' was included in the revised programme, anchored in the EC Habitats and Birds Directives, but at the expense of modelling and monitoring of eutrophication and hazardous substances. Thus, the marine sub-programme was for the first time ever faced with reductions.

In 2006, a mid-term revision was carried out resulting in significant reductions in a broad range of marine monitoring activities from 2007. The types of reductions were in spatial coverage (number of stations) and frequencies (many intensive stations had frequencies reduced to 1992 levels) (Bilj et al. 2007).

A revised programme, National Monitoring and Assessment Programme for the Aquatic and Terrestrial. Environments (NOVANA) came into effect in 2011 covering the period 2011-2015 (Naturstyrelsen 2011). This programme included the same themes and indicators/parameters as its predecessor (NOVANA 2004-2009), but the activity level was again slightly reduced (Figure 1.1).

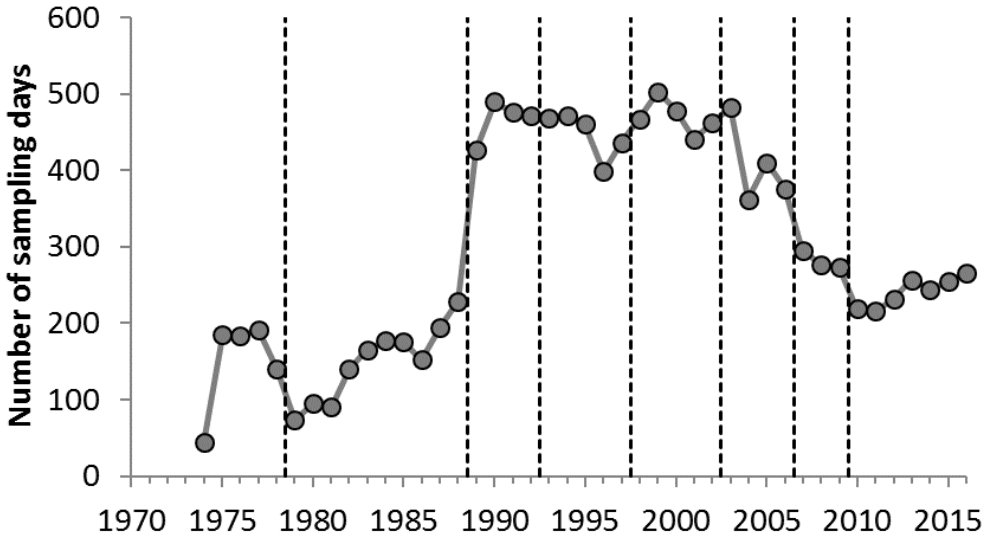


Figure 1.1 *Number of sampling days with regard to chlorophyll-a in the Inner Danish Waters 1974-2015.*

Since the mid-term revision of NOVANA 2004-2009, there has been a tendency towards reduction in the Danish marine monitoring activities. These well-known and well-documented reductions have resulted in a reduced spatial and temporal coverage of the open parts of the Inner Danish waters compared to the monitoring activities in the 90's.

Despite the above-mentioned reductions, the results of the Danish marine monitoring activities are often fed into international assessments of environmental status, e.g. by HELCOM (HELCOM 2009, Fleming-Lehtinen et al. 2015) and OSPAR (Andersen et al. 2016 and OSPAR 2017).

The Danish National Aquatic Monitoring and Assessment Program (DNAMAP[1]) does still include all relevant aspects (eutrophication, hazardous substances and biodiversity). In addition, new aspects emerging from the EU Marine Strategy Framework Directive (e.g. noise and marine litter) have been initiated with a long-term aim of integration in DNAMAP. Hence, discussions on how to improve the gaps in time and space, especially in open marine waters have been taking place for some time. A provisional answer on how to improve the current situation has been to consider other data sources, e.g. modelling and ships of opportunity[2].

The overarching aim of this study has been to develop and test a simple methodology for transforming fluorescence-based estimates of chlorophyll-*a* concentrations into a data product comparable with *in situ* measured chlorophyll-*a* values, taking temporal and spatial variations into account. This data product can be combined with *in situ* measurements to give a significantly larger data pool, with better spatial and temporal coverage compared than the NOVANA data set alone.

[1] The abbreviation DNAMAP covers the following monitoring programmes: VMP I, VMP II, NOVA-2003, NOVANA 2004-2009 NOVANA 2011-2015.
[2] Volunteering ships with environmental monitoring equipment that routinely passes specific waters.

2 Methods

2.1 Study area

This data used in this study was collected in German, Danish and Norwegian marine waters. The M/S Color Fantasy route starts in the Skagerrak from Oslo, goes through the Kattegat and Great Belt and ends in Kiel (Figure 2.2).

Oslo fjord
The Oslo fjord (Danish: Oslofjorden) is an inlet in the south-east of Norway, opening into the Skagerrak. It is, approximately, bounded to the south by the 59°N latitude and stretches to Oslo in the north.

Skagerrak
The Skagerrak is a strait between the south-east coast of Norway, the south-west coast of Sweden, and the Jutland peninsula, connecting the North Sea and the Kattegat. The Skagerrak is 240 km long and between 80 and 140 km wide, with an area of 47,000 km^2. It deepens toward the Norwegian coast, reaching over 700 m depth in the Norwegian Trench.

Kattegat
The Kattegat is a 30 000 km^2 sea area bounded by the Jutland peninsula to the west, the islands of the Danish Straits to the south and the provinces of Bohuslän, Västergötland, Halland and Skåne (Scania) in Sweden to the east. The Kattegat is connected to the Baltic Sea in the south, via the Danish straits, and to the Skagerrak in the North.

Great Belt
The Great Belt, (Danish: Storebælt), is the largest of the three Danish Straits (Little Belt, Great Belt, and the Sound) that connect the Baltic Sea to the Kattegat. It is located between the major islands of Zealand (Danish: Sjælland) and Funen (Danish: Fyn) in Denmark. The Great Belt is 60 km long and varies in width from 16 to 32 km.

Kiel Bay
The Kiel Bay (Danish: Kielerbugten) or Bay of Kiel is a bay in the southwestern Baltic Sea, bounded by the coastline of Schleswig-Holstein (Germany) in the south and by the Danish islands Als, Ærø, and Langeland in the north. It is connected to the Bay of Mecklenburg in the east and to two of the Danish straits in the north (Little Belt, Great Belt).

2.2 Data sources

2.2.1 Ferrybox

The ships of opportunity presently operating in Norway are part of an informal network of ships with autonomous physical-chemical-biological seawater sensor systems (termed a Ferrybox). The Norwegian Ferrybox concept was built upon an initial Norwegian Research Council-funded project (2001-2002) and developed through the EU Ferrybox project (www.FerryBox.no). Ferrybox systems are presently running on five ships: M/S Color Fantasy operating between Oslo, Norway and Kiel, Germany, M/S Trollfjord operating between Bergen, Norway and Kirkenes, Norway, M/S Norbjørn operating between Tromsø, Norway and Longyearbyen, Svalbard, M/S Norrøna operating between Hirtshals, Denmark, Tórshavn, Faroe Islands and Seyðisfjörður, Iceland and the M/S Oslofjord operating a high frequency route in the mouth of the Oslofjord. New lines are being planned, e.g. in the North Sea.

The Ferrybox is a set of environmental sensors on-board a ship of opportunity. The sensors are placed in a seawater stream that is pumped (2L/min) from ~4 meters depth from surface (Figure 2.1). The core sensors include, at a minimum, temperature and salinity. Some Ferrybox systems contain upgrades that

include sensors to detect chlorophyll-*a* fluorescence, cyano-bacteria fluorescence, turbidity, and oxygen. More recently, with financing from programs like the EU-funded JERICO, JERICO-NEXT, and NeXOS projects, in addition to NIVA institutional funding, new experimental and prototype chemical and biological sensors have been added. Some ships also have prototype systems that are interfaced with sensors on-deck that measure atmospheric variables and optical signals from the sea surface. Some systems include an automatic water sampler that can take water samples, on demand, to be stored for calibration after more advanced analysis in laboratory. Data are transferred via internet communication (when available) to a shore-based database where preliminary quality control is performed prior to being sent to the CMEMS data portal (www.myocean.eu).

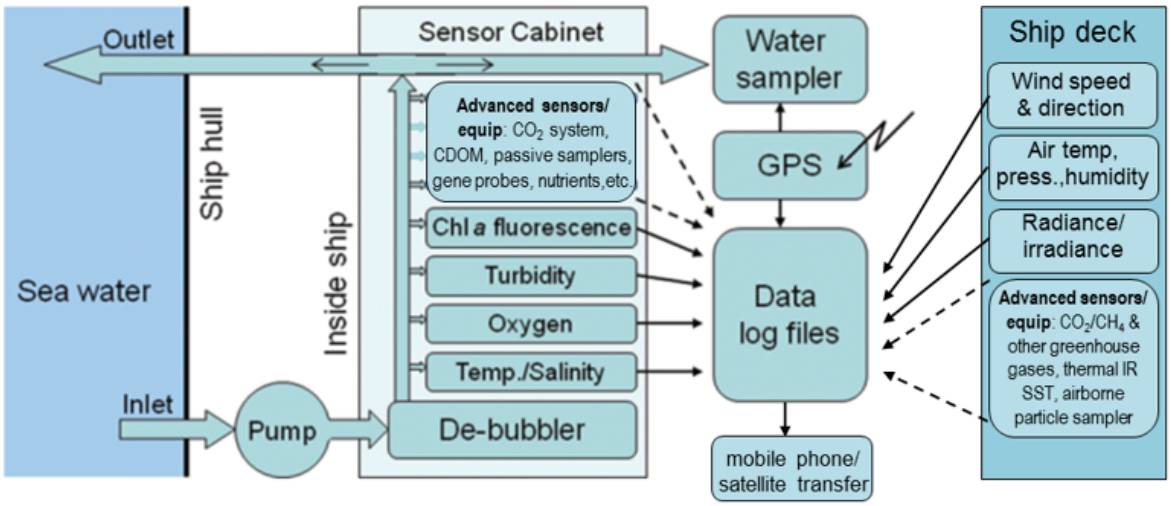


Figure 2.1 *An example of an autonomous Ferrybox system that includes a seawater pump, pipes, sensors, water sampler, and data logging hardware/software. Figure is adapted from Volent et al. (2011).*

2.2.2 The Ferrybox used in present study

The Ferrybox line between Kiel and Oslo covering the Kattegat region is operated daily by M/S Color Fantasy. It has been in operation since 2008, departing at 14:00 and arriving at 10:00 the following day. Every second day it arrives in Oslo or Kiel. Most of central Kattegat is covered during night time, which is important for interpretation of the chlorophyll-*a* fluorescence data (Chl-a_Fl, see below).

Chl-a_Fl is measured using a TriOS Microflu Chl-*a* sensor. The sensor works by exploiting the fact that chlorophyll-a, when excited by an external light source, re-emits a small portion of the absorbed energy as fluorescence at around 680 nm. Inside the sensor, the algal Chl-*a* is excited by short pulses of blue light, and the sensor measures the resulting fluorescence from the algal Chl-*a*. The sensor has been calibrated using an algal culture *of Skeletonema costatum* as a first primary laboratory calibration. At selected stations along the route, water samples are collected, typically on a biweekly basis. These samples are later filtered for Chl-a analysis and additional calibration. Due to the large diversity of chloroplast pigmentation and cell morphology among phytoplankton (Johnsen & Sakshaug 1997, Suggett et al. 2009), the ratio between Chl-a_Fl and Chl-*a* varies with species composition and community structure. The selection of stations used for this calibration step has attempted to minimize this effect by using stations in both the Skagerrak, Kattegat and the Baltic Sea. In addition, some fixed stations in the Skagerrak have been used. The Chl-*a* analysis is done by spectrophotometric analysis after a 100% Methanol extraction or a HPLC detection depending on the projects involved (NOVANA has used ethanol extraction, but the two methods should yield very similar results, as long as the correct conversion factors are used for calculating Chl-*a*).

After a quality control of the Chl-a_Fl data for e.g. biofouling, the ratios of Chl-a_Fl/Chl-*a* are used to prepare a yearly calibration based on the water samples. The ratio will vary with season and phytoplankton species. Also, the diurnal variation in this ratio is significant. At night-time, the phytoplankton is in a dark-

regulated state and the fluorescence of a certain mass of Chl-*a* is quite predictable. In contrast, under high light conditions during daytime, the algae use several mechanisms (non-photochemical quenching) to protect the photosynthetic apparatus from surplus energy from the absorbed light (Rodríguez et al. 2006, Huot & Babin 2010). This has the effect of decreasing fluorescence below the theoretical expected value during daytime, and may be more unpredictable because it is affected for instance when a cloud passes in front of the sun. For this reason, nocturnal fluorescence measurements give a more reliable indication of Chl-*a* (Huot & Babin 2010). We have not yet introduced operational correction for varying light levels during night-time (i.e., that it is slightly more light at the start and end of the night), but this research is ongoing.

2.2.3 *In situ* measured data from NOVANA and other sources

Data from the Danish and Swedish marine monitoring programs were extracted from stations close to the ferry transect between Oslo and Kiel, passing through Skagerrak, Kattegat, Great Belt, and Kiel Bight (Figure 2.2). The Danish open-water monitoring program is coordinated with Sweden, such that several of the stations are shared and monitoring data are exchanged between the two countries. The extracted data included discrete water samples at standard depths (1, 5, 10, 15, 20, 25 m etc.), where chlorophyll-*a* was measured according to the Danish standard methods (Markager and Fossing, 2014). Water samples were filtered and chlorophyll-*a* was extracted from the filter using ethanol and measured spectrophotometrically. All monitoring data had positions for the actual sampling as well as date and time of sampling.

Many stations had a low sampling frequency, with approximately 3 samples per year over the 9-year period (Table 2.1). Data were available for all years for all stations, except the Swedish station Å17, where 2016 data were not yet available. Four stations: Anholt (413) and Ålborg Bugt (409) Gniben (925) and Great Belt (6700053) were monitored approximately biweekly throughout the period.

Table 2.1 *Monitoring stations and their depths as well as the period with observations, number of sampling days and the levels for chlorophyll-a and fluorescence. Stations are listed from north to south.*

Station	Depth	Period	Chlorophyll-*a*		Fluorescence	
	(m)		#sampling days	Average (µg/l)	#sampling days	Average
Å17	348	2008-2015	79	1.56	0	
1007	47	2008-2016	32	2.04	28	0.96
905	77	2008-2016	128	2.38	26	0.95
409	14.5	2008-2016	204	2.40	191	0.97
413	55	2008-2016	198	2.32	25	0.94
415	16	2008-2016	24	2.35	27	1.11
925	43	2008-2016	210	2.12	196	2.19
935	48	2008-2016	33	2.42	28	2.97
6700009	23	2008-2016	0		69	3.04
6700053	32	2008-2016	245	2.77	244	2.99
939	38	2008-2016	33	2.10	31	2.80
6700051	40	2008-2016	0		54	3.00
443	36	2008-2016	32	2.75	26	1.48
450	32	2008-2016	32	2.88	27	1.54
N3	19	2008-2016	29	3.45	29	4.03

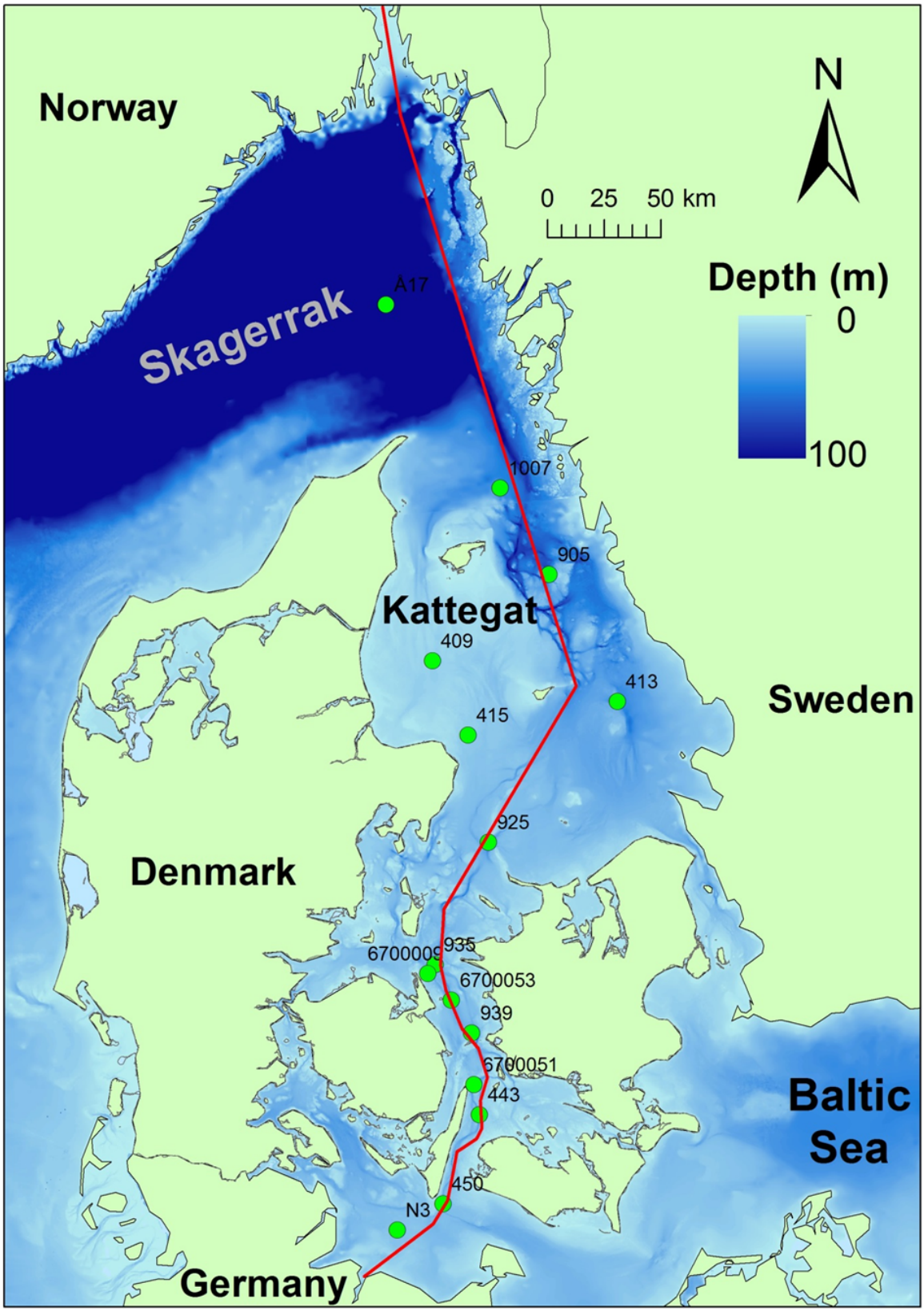


Figure 2.2 *Location of open-water monitoring stations along the Oslo-Kiel ferry transect.*

2.2.4 Data from a 3D biogeochemical model

As part of the preparation of the second generation of the Danish River Basin Management Plans (RBMP), the Nature Protection Agency (Naturstyrelsen) now the Danish Environmental Protection Agency (Miljøstyrelsen) initiated a research project in 2014 aimed at the development of models for implementation in the national water management program. A part of that model development was built on mechanistic 3D biogeochemical models. These models make several assumptions about cause-effects relations between e.g. water transport, nutrient loadings and algal biomass.

One of the developed models, the model covering the inner Danish waters (the IDW model), covers all marine waters from Skagerrak to the northern part of the Baltic Sea, but with the highest resolution in the coastal Danish waterbodies i.e. the areas covered by the Water Framework Directive (WFD). The IDW model comprises 3 sub-models:

1. A hydrodynamic model (MIKE3FM). Based on atmospheric pressure, wind speed and direction, air temperature and freshwater input this model calculates the water movements (currents) within the inner Danish waters as well as the different densities derived from salinity and water temperature. Hence, this model also describes potential separations between surface and bottom waters due to these density differences.
2. An advection and dispersion model (AD). This model is used for calculations of transport of any substance included in the model. Based on the model results from this sub-model the concentrations of the different substances are estimated based on dilution and transport alone.
3. A biogeochemical model (ECO Lab). Based on a number of differential equations this model calculates the changes in concentrations of the different substances based on e.g. water temperature and solar radiation. This could e.g. be the change in nitrate and chlorophyll-*a* concentrations based on primary production.

The IDW model – and hence, the model from which we have extracted data for this project – includes the following pelagic components: Phytoplankton (diatoms, flagellates and cyano-bacteria), chlorophyll-*a*, zoo-plankton (micro- and meso-zooplankton), particulate organic matter, dissolved organic matter, inorganic nutrients and dissolved oxygen. In addition, this model includes interactions between the pelagic and the benthic phase, including organic and inorganic pools of nutrients in the sediment as well as the biomass of eelgrass, macroalgae and micro-benthic algae. As this project focuses mostly on the possibilities for monitoring the open – and mostly deeper – parts of the Danish waters the biomass of eelgrass, macroalgae and micro-benthic algae are of less importance here and hence not included when using the model.

The IDW model was optimised to fit the chlorophyll-*a* concentration and water transparency (K_d) in the coastal parts of the Danish water bodies, and therefore for this project we used the model directly without any further optimisation.

For this project, we have extracted daily chlorophyll-*a* model data in locations corresponding to six NOVANA stations corresponding to stations that are also included as part of the Ferrybox analysis. The NOVANA stations are 1007, 925, 935, 6700053, 939 and N3.

The IDW model has a vertical resolution of 1 m within Kattegat, Great Belt and the Fehmarn Belt, and model data have been extracted from 2 m depth.

The model is based on the years 2002-2011. As analyses carried out in present study regards the period 2008-2015, the period 2008-2011 is the only overlap between the two. Hence, only data from this period is presented in this analysis. In Figure 2.3, comparisons between modelled chlorophyll-*a* data and NOVANA observations are exemplified.

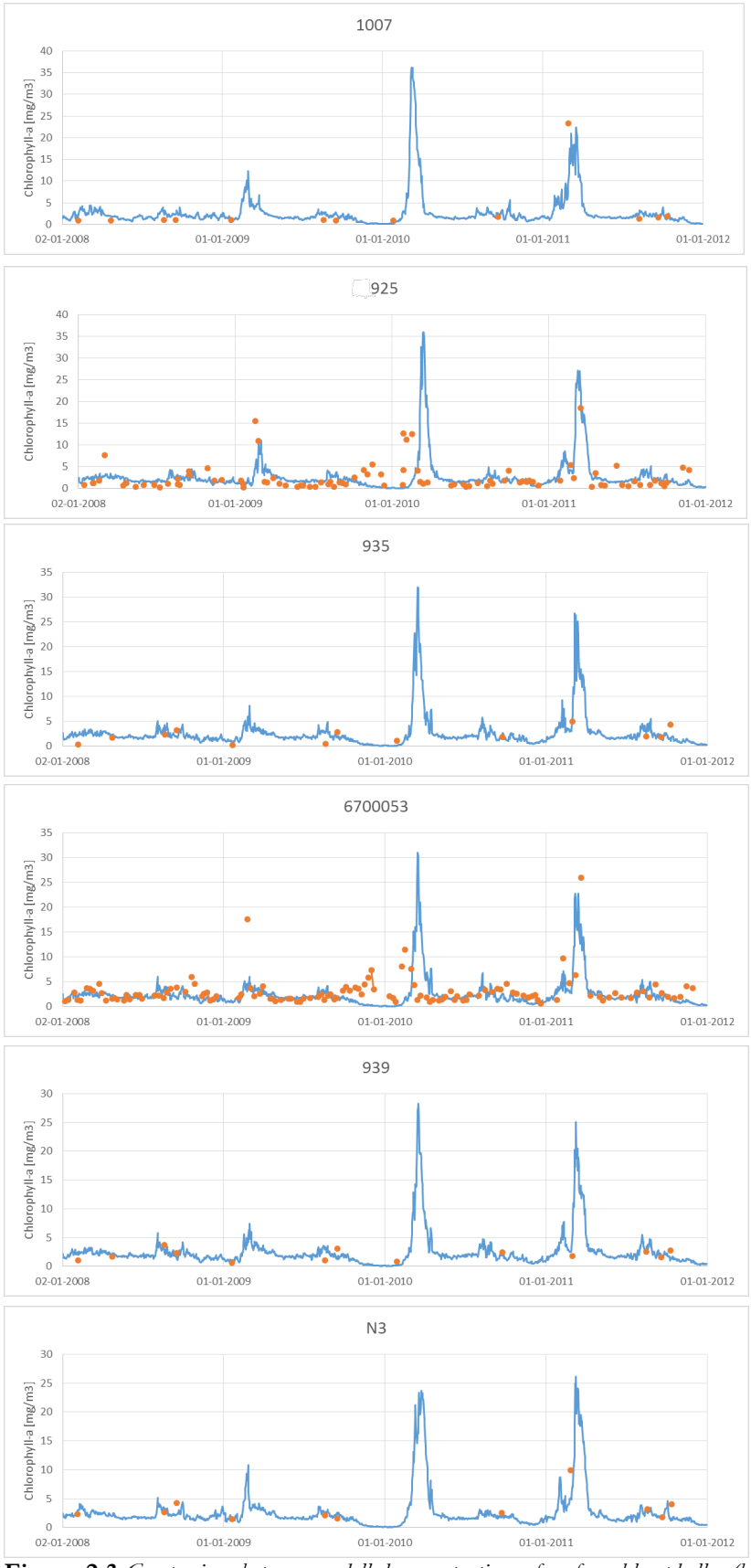


Figure 2.3 *Comparison between modelled concentrations of surface chlorophyll-a (blue line) and measured chlorophyll-a values averaged over the top 3 m (orange dots) from north (top figures) towards south (bottom figures).*

2.3 Statistical analyses

For this study, we used data from the southbound Ferrybox route, even though the northbound route in some cases passed closer to oceanographic stations. For most of the Danish waters, fluorescence data is most reliable, i.e. less impacted by the physiological status of the phytoplankton, during nighttime. The southbound route passes the northernmost station (1007) at around 8 p.m., passes then the other stations in late evening and early night until the southernmost station (939) is passed between 3 and 4 a.m.

For each day, we calculated the minimum distance between the Ferrybox and each station (Table 2.2).

Table 2.2 *NOVANA station data (listed from north to south) and distances to the Ferrybox transect (in km). 'Distance to Ferrybox' is the median distance to the Ferrybox line when heading south. The table also shows the total number of measurements of chlorophyll-a at 5 m depth (N), and the number of months for which data exist for at least three years of chlorophyll-a measurements at 5 m depth (N months ≥ 3 yrs.). Stations <20 km from the Ferrybox line and with year-round sampling are shown in bold.*

Station	Latitude	Depth (m)	Distance to Ferrybox (km)	N	N months ≥ 3 years
1007	57.53	47.0	1.3	32	4
905	**57.18**	**77.0**	**4.2**	**128**	**12**
409	56.86	14.5	57.9	321	12
413	**56.67**	**55.0**	**19.0**	**199**	**12**
415	56.56	16.0	33.2	24	4
925	**56.13**	**43.0**	**1.4**	**307**	**12**
935	55.65	48.0	1.5	33	4
6700053	**55.51**	**32.0**	**0.5**	**247**	**12**
939	55.38	38.0	1.4	33	4
443	55.05	36.0	1.4	32	4
450	54.70	32.0	0.6	32	4

Seven stations were <2 km from the Ferrybox route; however, many of these stations have only data from selected months (Table 2.3).

Table 2.3 *NOVANA station data, data coverage per month. For each month, the number of years in the period 2008-2016 (station 1007: 2008-2015) with measurements of chlorophyll-a at 5 m depth is shown.*

	Month												
Station	1	2	3	4	5	6	7	8	9	10	11	12	Lat
1007	2	6	1	1	0	0	0	8	8	5	0	0	57.5
409	9	8	9	9	9	9	9	9	8	9	8	9	56.9
413	7	8	8	8	5	7	8	9	9	9	8	8	56.7
415	2	6	0	0	0	1	0	5	6	4	0	0	56.6
443	2	7	0	1	0	0	0	6	9	6	0	0	55.1
450	2	7	0	1	0	0	0	6	9	6	0	0	54.7
6700053	9	9	9	9	9	9	9	9	9	9	9	7	55.5
905	7	8	8	8	5	7	8	9	9	9	8	8	57.2
925	9	8	9	8	9	8	8	8	8	8	9	7	56.1
935	2	5	2	1	0	0	0	8	9	6	0	0	55.7
939	2	6	1	1	0	0	0	8	9	6	0	0	55.4

For each station, we selected all Ferrybox data within a threshold distance of *r* km from the position of the station. We then took the median of the chlorophyll-*a* estimate for the Ferrybox passage closest in time to the station data (allowing a maximum of 48 hours between station measurements and Ferrybox data, assuming that station measurements were taken at noon). The median value was applied to avoid bias from potential outliers. The selection of the threshold distance *r* has a large effect on the results: if *r* is too small, very few Ferrybox data are selected and the resulting median chlorophyll-*a* was highly variable due to small-scale variation in phytoplankton abundance. If *r* is too large, the Ferrybox sample includes

too large an area and the sample is not representative for local conditions. To achieve the best selection of *r*, we did a simple regression between station-based chlorophyll-*a* measurements at 5 m (+/- 0.5 m) depth and median Ferrybox chlorophyll-*a* estimates for each of the four focus stations (905, 413, 925, 6700053). For each station, we used only the days which had at least one Ferrybox observation within the shortest threshold distance tested, in order to use the same days to test different threshold distances (otherwise the R-squared values would not be comparable). The value of *r* corresponding to the best fit was used for selecting Ferrybox data. For each day and station, we extracted the median values of chlorophyll-*a* as well as temperature and salinity within a radius of *r* from the station.

While the Ferrybox samples water at a depth of approximately 4 m and the Ferrybox is mounted close to the stern, the ship creates some mixing of water and the sampled water may therefore originate also from more shallow water depths. To investigate to which degree the water sampled by the Ferrybox is mixed, we used salinity data from 0-1 m and from 5 m from three stations (905, 925 and 6700053) that are close to the Ferrybox route and have year-round data. We performed two analyses on these data. In the first analysis, we analyzed Ferrybox salinity (i.e., median value for each passing of the station, as described above) as a function of station measurements of salinity at 0-1 m and 5 m, using the following statistical equation:

$$Salinity_{Ferrybox,i} = b_1 \cdot Salinity_{0-1\,m,i} + b2 \cdot Salinity_{5\,m,i} + \varepsilon_i$$

i.e., a linear regression with intercept fixed to be zero. If the sampled Ferrybox water is a mixture of x % of 0-1 m water and y % of 5 m water, we expect the estimates to be b_1 = x % and b_2 = y %, with $b_1 + b_2$ = 100%. In the second analysis of the salinity data, we used the same stations, but picked only days with a substantial difference by depth, i.e. when salinity at 5 m minus salinity at 0-1 m is ≥1. With these data, we compared the absolute difference between Ferrybox salinity and (a) salinity at 0-1 m, (b) salinity at 5 m, and (c) mean salinity at 0-1 and 5 m.

For chlorophyll-*a* from station measurements and the corresponding Ferrybox estimates, we performed a statistical analysis of station-measured chlorophyll-*a* as a function of Ferrybox-based chlorophyll-*a*. Depth of the station data was selected based on the analysis of salinity (see previous paragraph). As plots of the data did not indicate evidence of non-linear relationships on the logarithmic scale, we used simple linear models of log-transformed data. We fitted both models for each station separately, and to all data combined. For the models fitted to each station, we considered three linear models of different complexity, Models 1-3 (Table 2.4). We used AICc (Akaike's Model Criterion adjusted for small sample size, Hurvich and Tsai 1989) to select the optimal model (i.e., the best compromise between model complexity and goodness-of-fit) among models 1-3. As models 1-3 cannot be used to adjust Ferrybox data between stations, we also fitted a set of models (models 4-7) for all data combined. Here, we let latitude alone represent space, since longitude varies little along the ferry route. The larger sample size allowed us to use non-linear models of the type GAM (Generalized additive models) (Hastie & Tibshirani, 1990). However, in order to avoid excessive complexity and based on plots, we assumed that the basic relationship between Ferrybox-based and station-based chlorophyll-*a* measurements was always linear; non-linear effects were only allowed for additional adjustment for season and latitude. Again, we considered models of different complexity (Table 2.4). In model 4, we assume the same relationship between Ferrybox and station data for all areas and times of year. In model 5 we let adjustment of Ferrybox data depend on the time of year using the non-linear (spline) function *f*. In model 6, adjustment of Ferrybox data depends both on the time of year (function *f*) and on the latitude (function *g*); however, we assume that the seasonal pattern (i.e., function *f*) is the same for all areas. In model 7, we let the seasonal pattern also change with area, so the effects of season and latitude are non-additive (i.e., season and latitude interact). Again, we compared models 4-7 using AICc. All models were fitted and evaluated based on data over 4 years, 2012-2015, as calibration of Ferrybox chlorophyll-*a* estimates was expected to be most accurate for these years.

The resulting models were used to adjust the Ferrybox estimates for all years.

Table 2.4 *Models for station-based chlorophyll-a measurements (Y) as a function of Ferrybox chlorophyll-a estimates (X) as well as month. Both variables were log-transformed using c = 0.1. In the GAM models (model 6-8),*

Model	Formula	Interpretation
Separate linear models for each station		
Model 1	$log(Y + c) = a + b * log(X + c)$	Relationship same for all months
Model 2	$log(Y + c) = a_{Month} + b * log(X + c)$	Slope same for all months, intercept varies
Model 3	$log(Y + c) = a_{Month} + b_{Month} * log(X + c)$	Slope and intercept varies among months
Common non-linear model for all stations		
Model 4	$log(Y + c) = a + b * log(X + c)$	Relationship same for all stations, months
Model 5	$log(Y + c) = a + b * log(X + c) + f(month)$	Relationship does not differ with latitude, but non-linear effect of time of year
Model 6	$log(Y + c) = a + b * log(X + c) + f(latitude)$	Non-linear, additive effects of time of year and latitude
Model 7	$log(Y + c) = a + b * log(X + c) + f(month) + g(latitude()$	Non-linear, additive effects of time of year and latitude
Model 8	$log(Y + c) = a + b * log(X + c) + f(month, latitude)$	Relationship affected by non-linear, non-additive effects of time of year and latitude

3 Results

The threshold distance of *r* km from each oceanographic station, used for selection of Ferrybox data, was selected based on regression between station-based chlorophyll-*a* measurements at 5 m (+/- 0.5 m) depth and median Ferrybox chlorophyll-*a* estimates for each station.

The results of these regressions (Figure 3.1) show a clear optimum *r* for most stations, varying from 1 km for station 6700053 in the Great Belt (which lies only 0.5 km away from the Ferrybox route) to 24 km for station 413 (which lies 19 km away). The larger the threshold distance, the more Ferrybox data are included in the comparison. This is probably the main reason that the correlation between the two types of estimates is unexpectedly lower for low threshold distances: if the threshold distance is so short that the ferry route goes inside the threshold distance for just a few minutes, the Ferrybox sample may be more affected by randomness inside the small sample.

These values of optimum *r* were used to extract Ferrybox data for further analysis. The p-values of the regression analysis of salinity (Table 3.1) showed that salinities measured at both 0-1 m and 5 m depths contribute to explaining the salinity measured in the Ferrybox.

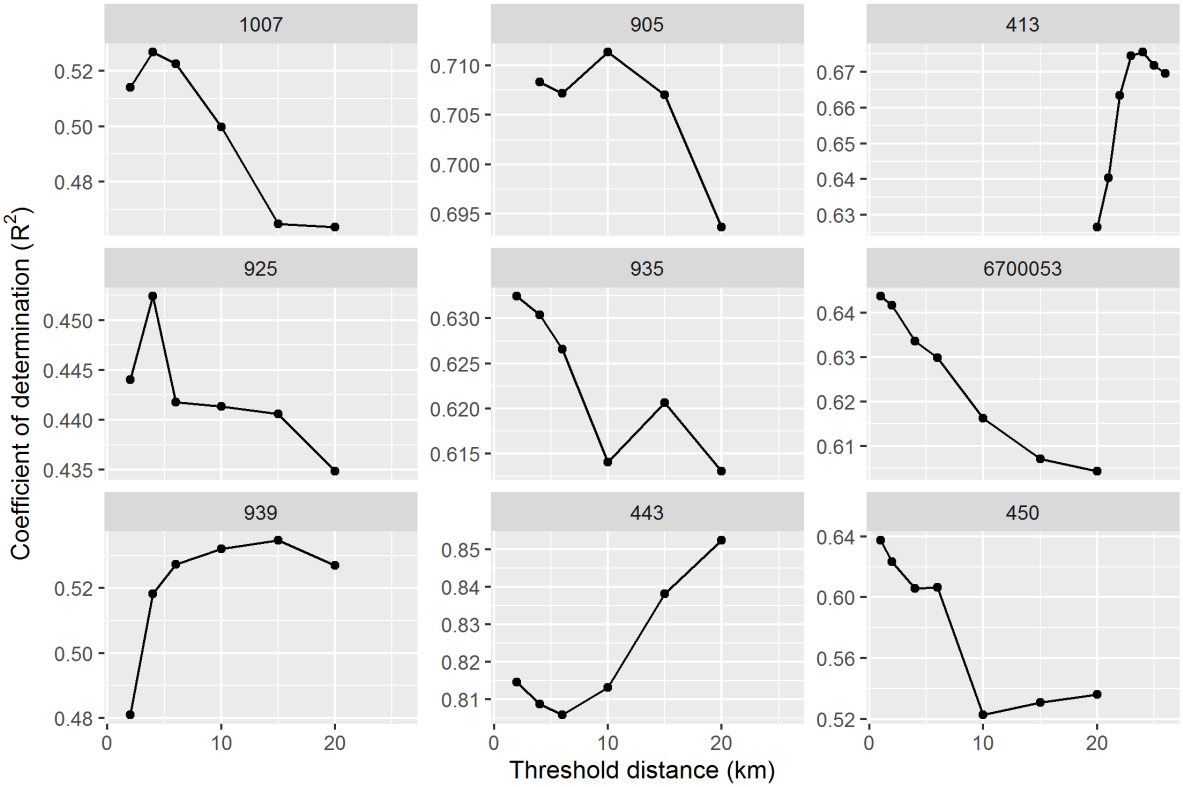


Figure 3.1 *Selecting optimal threshold distance (r) for selecting Ferrybox data close to oceanographic stations (sorted from north to south). The graph shows the fit of a regression between station-based chlorophyll-a measurements and median Ferrybox chlorophyll-a estimates, as measured by the coefficient of determination (R-square).*

The sum of the estimates is very close to 1 for each station, indicating that Ferrybox salinity indeed can be expressed as a weighted mean of salinities at 0-1 m and 5 m depths. For instance, for station 905, the estimates of 0.64 and 0.35 indicates that water measured by the sensor is equivalent to a mix of 64% water from 0-1 m depth and 35% water from 5 m depth. (We say "equivalent to", as the water intake of course also gets water from between 1-4 m depth.) In general, the estimates for $Salinity_{0\text{-}1\ m}$ are somewhat above

0.5, indicating that the Ferrybox salinity is somewhat closer to the 0-1 m salinity than to 5 m salinity (Figure 3.2). However, the standard error does not indicate that the estimates deviates from 0.5 (Figure 3.2). Thus, Ferrybox salinity corresponds to the unweighted mean of salinities measured at both 0-1 m and 5 m depths. This is supported by plots and analysis of days where salinities at 0-1 m and 5 m depth differed by at least 1. For two of the stations (905 and 6700053), the median of absolute difference is smallest for the mean salinity at 0-1 and 5 m depth. For the last station (925), Ferrybox salinity is closest to 0-1 m salinity. The difference between stations was not significant ($P > 0.2$), but the absolute difference was significantly smaller for mean salinity than for either 0-1 or 5 m depth ($P = 0.001$).

Table 3.1 *Regression of Ferrybox salinity as a function of station salinity at 0-1 and 5 m depth. N = number of observations; Sum estimates = $Salinity_{0\text{-}1\ m}$ estimate + $Salinity_{0\text{-}1\ m}$ estimate.*

Station	$Salinity_{0\text{-}1\ m}$ Estimate	P-value	$Salinity_{5\ m}$ Estimate	P-value	N	Sum estimates
6700053	0.57	0.001	0.41	0.017	88	0.98
905	0.64	0.000	0.35	0.001	82	0.99
925	0.76	0.000	0.21	0.051	122	0.97

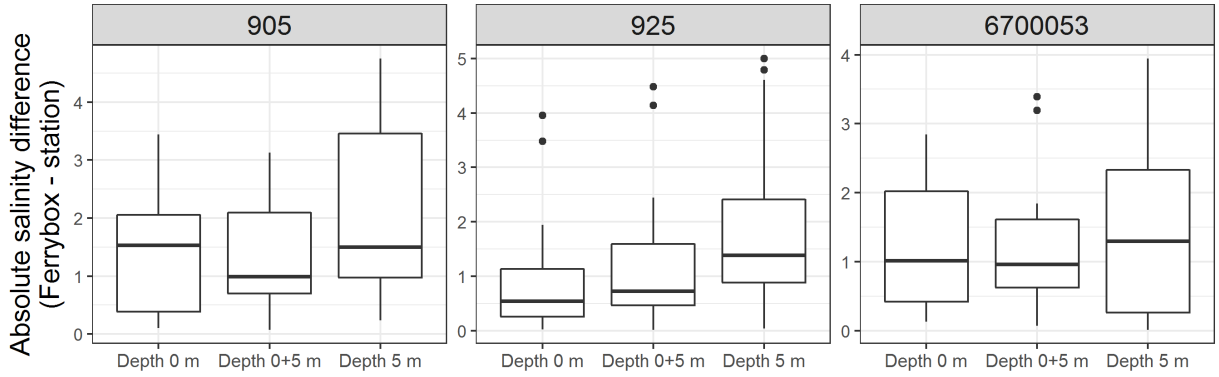


Figure 3.2 *Absolute difference between Ferrybox salinity and station salinity at 0-1 m, 5 m and the mean salinity of 0-1 and 5 m. The figures show the first and third quartile of the data (box), the median value (horizontal line inside box), the approximate 95% confidence interval of the median (the ends of the vertical lines), and outlier observations outside the confidence interval (dots).*

Since the analysis of salinity indicated that the water sampled by the Ferrybox is approximately equivalent to a 1:1 mix of 0-1 and 5 m water (the estimates for $Salinity_{0\text{-}1\ m}$ was not significantly different from 0.5 and not significantly different among stations). Therefore, for each station, we used mean values for 0-1 and 5 m depth for station data of chlorophyll-*a*. When we plot these data before statistical adjustment, the relationship between the Ferrybox estimates and the station-based chlorophyll-*a* deviates from a 1:1 relationship for some stations (Figure 3.3). As an example, for station 6700053 in the Great Belt, the Ferrybox-based estimates consistently underestimate the actual chlorophyll-*a*, while at station 413 the Ferrybox-based estimates are unbiased for lower levels of chlorophyll-*a* but give underestimates at higher chlorophyll-*a* levels. This indicates, that the calibration based on bottle samples can, to some degree, estimate the actual chlorophyll-*a*, but that some extra adjustment is necessary. Also, it indicates that the optimal adjustment may depend on the area. While the Ferrybox-based estimates are quite unbiased for the stations in the open Kattegat (from north: stations 1007, 905, 413, 925), the Ferrybox-estimates tend to be too low in the Great Belt (stations 935, 670053, 939, 443 and 450).

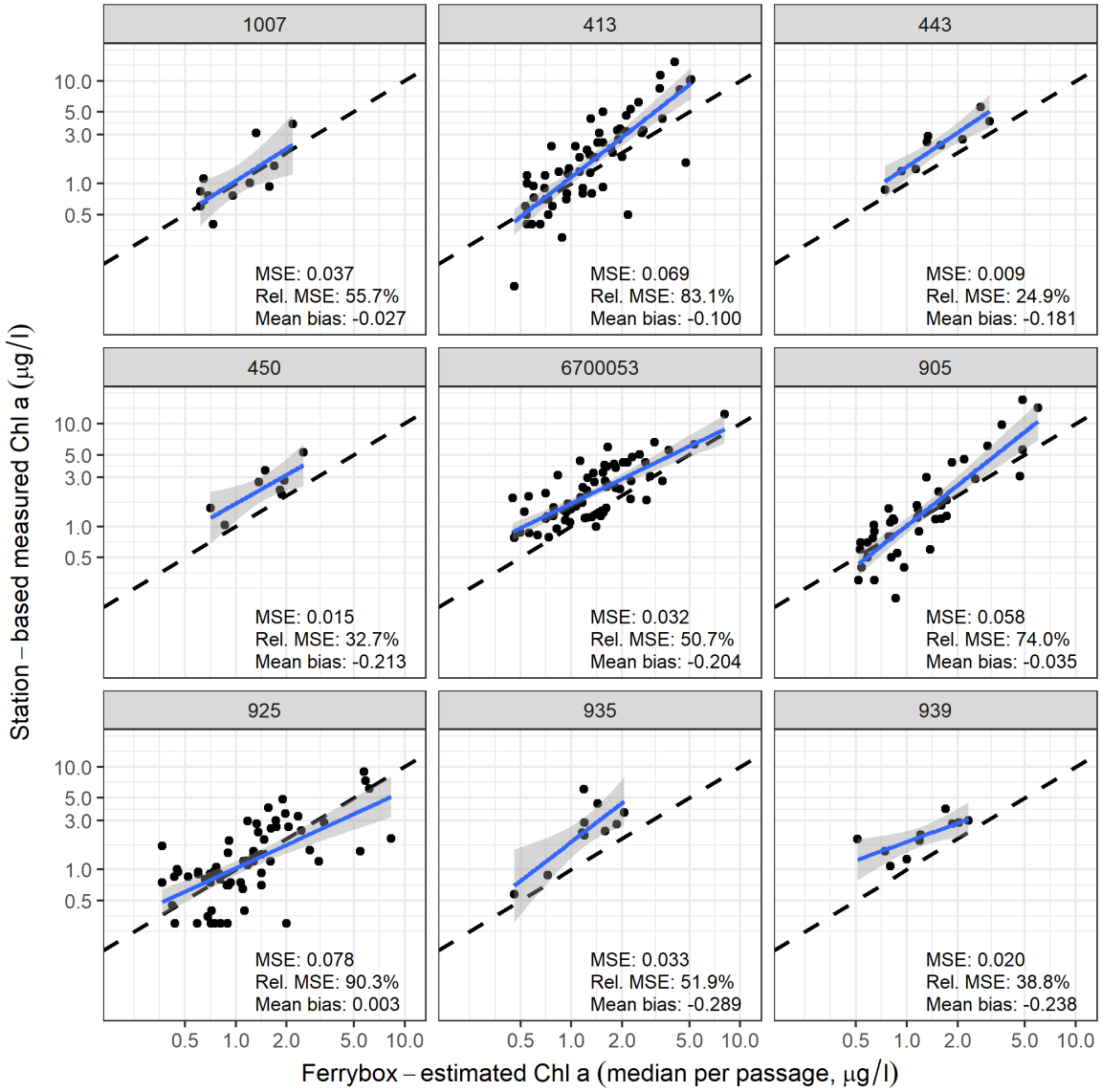


Figure 3.3 *Match-ups between station-based chlorophyll-a measurements (vertical axis) and Ferrybox chlorophyll-a estimates (horizontal axis) for each station. Each point represents one station-based chlorophyll-a measurement at 5 m depth as well as the median of Ferrybox-based estimates of chlorophyll-a within a given radius around the station. The blue line is a linear regression fit (with confidence intervals), while the dashed line is the 1:1 relationship. The figure also shows MSE (mean square error) for log_{10} (Ferrybox estimate) as a function of log_{10} (station-based estimates), the back-transformed relative MSE in percent, and the mean bias (Ferrybox minus station) on log_{10} scale. The data are from the years 2012-2015.*

The results for models fitted separately to each station (Model 1-3; Table 3.2) indicates that for 4 stations, there was no significant effect of month (Model 1). The three stations with most data (413, 925 and 6700053) all showed that including month as an additive effect improves Ferrybox estimates (Model 3). For two stations, there was evidence for a changing slope among months (Model 3), but it should be noted that these stations have little data and this result should be treated with caution. When these models were applied as an adjustment, the resulting fit was reasonably good (Figure 3.4). This adjustment was applied to all data 2008-2015, resulting in one time series for every station (Annex 1 Figure A1.1 & A1.2.).

Table 3.2 *Model selection for each station. The table shows dAICc, defined as the AICc value minus the minimum AICc value for that station. AICc is AIC (Akaike's Model Criterion) adjusted for small sample size (Hurvich and Tsai 1989). For each station, the optimal model is the one with dAICc = 0. Model 1: no significant effect of month, Model 2: additive effect of month (log-scale), Model 3: changing slope among months.*

	Model 1		Model 2		Model 3	
Station	**dAICc**	**d.f.**	**dAICc**	**d.f.**	**dAICc**	**d.f.**
1007	0.00	3	28.92	7	Inf	10
905	0.00	3	16.25	14	72.48	24
413	20.26	3	0.00	14	28.79	25
925	39.80	3	0.00	14	17.45	25
935	242.82	3	289.68	7	0.00	10
6700053	8.91	3	0.00	14	29.12	25
939	0.00	3	28.87	6	Inf	9
443	0.00	3	41.57	6	Inf	8
450	154.72	3	227.69	6	0.00	8

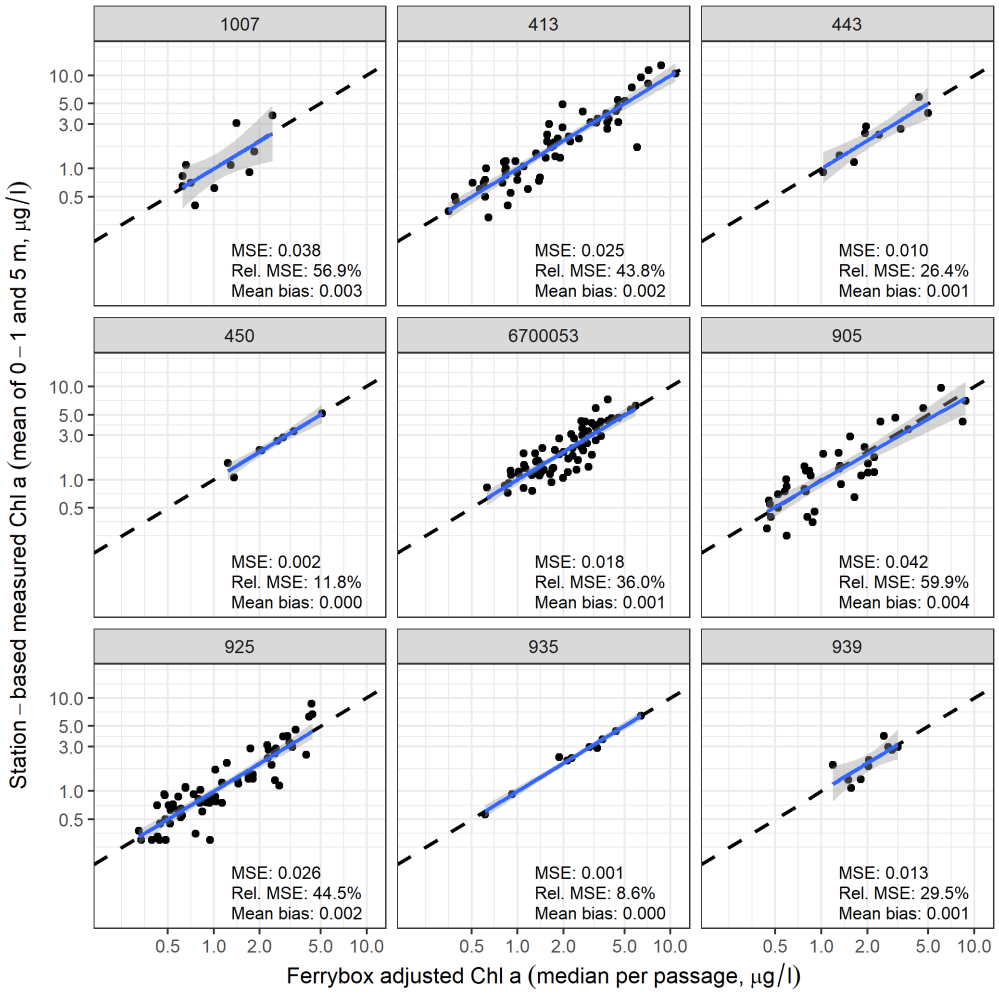


Figure 3.4 *Relationship between station-based chlorophyll-a and Ferrybox-based chlorophyll-a adjusted for local and seasonal effects. The data are from the years 2012-2015. Each point represents one adjusted station-based chlorophyll-a measurement at 5 m depth as well as the median of Ferrybox-based estimates of chlorophyll-a within a given radius around the station. The blue line is a linear regression fit (with confidence intervals), while the dashed line is the 1:1 relationship. The data are from the years 2012-2015.*

For models 4-8, fitted to the entire data set, the models that did not take into account season or latitude (Model 4) or only one of these factors (Models 5 and 6) were strongly inferior to the models that took both season and latitude into account (Models 7 and 8). While model 8 had the lowest AICc value, Models 7 and 8 can be considered to be approximately equally good according to dAICc (a rule of thumb is that models with dAICc < 2 are approximately equally good; (Burnham & Anderson, 2004)). In other words, the higher complexity of Model 8 compared to Model 7 (model has approximately 3.5 higher degrees of freedom) is more or less compensated by its better fit to the data. Models 7 and 8 are visualized in Fig. 3.5. For Model 7, it is clear that Ferrybox-based values need more adjustment upwards in winter than in summer (Fig. 3.5b) for areas in the Great Belt (ca. 55.0-55.5 latitude; Fig. 3.5c). Model 8 shows that the seasonal effect on adjustment should be stronger in Kattegat (Fig. 3.5f) than in the Great Belt (Fig. 3.5.e). A technical detail that should be noted is that we were able to specify that the Month effect should be a cyclic spine in the case of Model 7, but not in Model 8. Specifying Month as a cyclic spine ensures that there is no discontinuity between December and January. However, this constraint is technically difficult to implement in Model 8, where Month is part of an interaction effect; however, the resulting December-January discontinuity (Fig. 3.5e-f) appears to be small. The predicted values produced by adjusting Ferrybox-based measurements using Models 7 and 8 appears to show a quite good fit to the actual time series (Fig. 3.6), although the model (as statistical models usually do) tend to produce lower "peaks" and higher "troughs" compared to the actual data.

Table 3.3 *Model selection for models fitted to the full data set (across stations). The table shows dAICc, defined as the AICc value minus the minimum AICc value for that station. AICc is AIC (Akaike's Model Criterion) adjusted for small sample size (Hurvich and Tsai 1989). The optimal model is the one with dAICc = 0. Model 1: no significant effect of month, Model 2: additive effect of month (log-scale), Model 3: changing slope among months.*

	Model 4	Model 5	Model 6	Model 7	Model 8
d.f.	3,00	5,40	7,13	9,73	13,20
dAICc	121,99	43,70	98,12	1,47	0,00

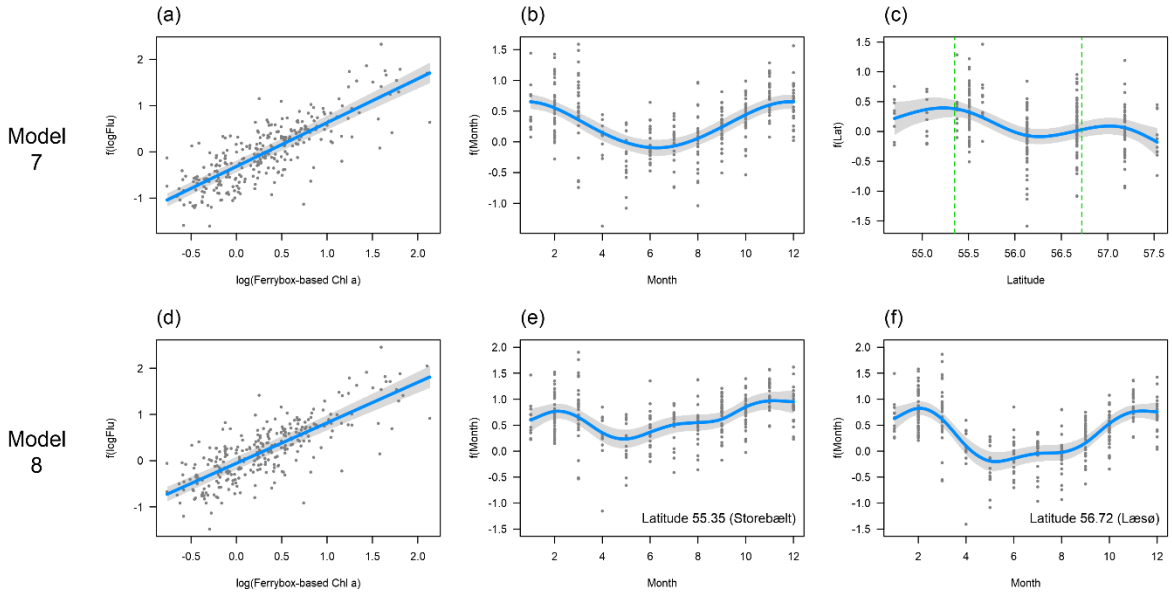


Figure 3.5 *Effects of each factor in Model 7 (a-c) and Model 8 (d-f). For Model 7, each subplot shows the partial effect of each factor (the effect of that factor after the effect of the two other factors have been removed). For instance, the y values in subplot (a) shows the predicted log(chlorophyll a) values when Ferrybox-measured chlorophyll-a varies, given median values of Month and Latitude. For Model 8, plot (d) shows the partial effect of Ferrybox-measured chlorophyll-a, while plots (e) and (f) shows the partial effect of Month at two given values of Latitude. I.e., it shows two cross-sections of the 3-dimensional surface that defines the (Month, Latitude) effect. If the curve was plotted for intermediate values of Latitude, it would show a gradual change from plot (e) to plot (f). The Latitude values used in (e) and (f) are shown as dotted vertical lines in (c).*

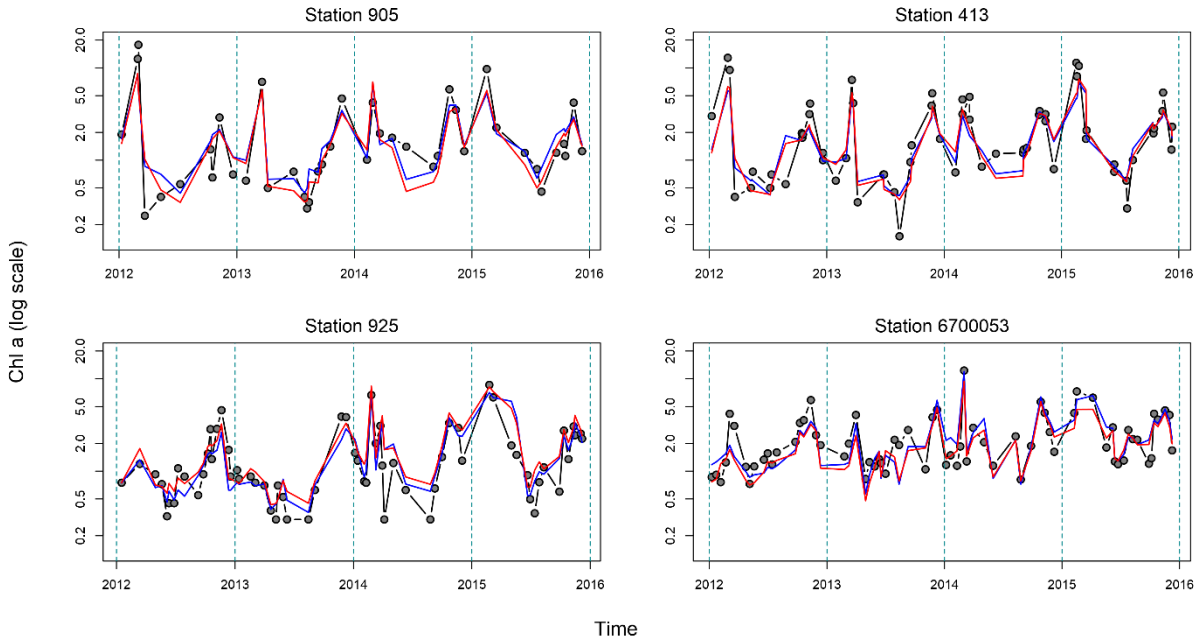


Figure 3.6 *Actual and predicted values of chlorophyll-a at the four stations with most data. The actual values (grey dots connected by black lines) show chlorophyll-a measured at the stations (given as the average at 0-1 m depth and 5 m depth). The colored lines show the predicted values using Model 7 (blue) and Model 8 (red).*

4 Discussion and conclusions

This study has been initiated by the Danish Environmental Protection Agency to clarify whether the Ferrybox measurements of chlorophyll-*a* can be adjusted for local and seasonal effects in a way where the adjusted data set potentially can be combined with NOVANA data in order to expand the total pool of chlorophyll-*a* measurement for the Danish Straits. While the adjustment done in this report are done on a station-by-station basis (Table 3.2), there appear to be general spatial and seasonal patterns which may be used to adjust the Ferrybox data also between stations and through the main bloom seasons.

Based on the analyses and results of this study (see Chapter 3), we take stock of progress by answering the questions (1) where are we now? and (2) where are we going? And based on the answers to these two questions we draw up conclusions and perspectives of this study.

4.1 Where are we now?

The open Danish marine waters are currently not monitored as intensively as the coastal waters. Hence, approaches to improve both temporal and spatial coverage of the open waters would be beneficial to better characterise these waters and to improve the assessment of the environmental status in a reliable way.

The monitoring and management of Danish marine waters have for decades focused on the Danish Straits. These are monitored on a daily basis via the Ferrybox on-board the Oslo-Kiel ferry. The ferrybox is operated by the Norwegian Institute for Water Research (NIVA). The data collected by this system represent valuable information for inclusion in the Danish marine assessment activities.

Statistical analyses of the chlorophyll-*a* measurements from the Oslo-Kiel ferry in combination with temperature and salinity data at different depths has been conducted to match up Ferrybox data and NOVANA. The correlations between the adjusted Ferrybox data and NOVANA data (Figure 3.4) indicated that the data product is comparable with *in situ* measurements from the NOVANA programme.

4.2 Where are we going?

We believe, based on this study, that composite products where data from various sources are combined into new product is a way forward. So far, we suggest combining *in situ* measurements from the NOVANA programme with measurements from a Ferrybox on-board the Oslo-Kiel ferry.

A potential and interesting next step would be the development of so-called aggregate products, where data from multiple sources and monitoring programs, sampled with different methods are combined to a single data product.

Therefore, we have included a brief comparison between NOVANA observations and satellite data from the sensors MERIS (the MEdium Resolution Imaging Spectrometer) and MODIS Aqua (the Moderate Resolution Imaging Spectroradiometer) based on data processed and analysed by DHI. For the period 2009 - April 2012, DHI used full-resolution data from MERIS. MERIS was designed and optimized for observations of fluorescence in water. MERIS stopped working in April 2012 and therefore data from MODIS were used for 2013 – 2016 instead.

MODIS Aqua data has a horizontal resolution between 750-1000 m, and hence, has an optimal resolution for monitoring larger areas as e.g. the sub-regions and their sub-divisions associated with the Marine Strategy Framework Directive (MSFD).

For each *in situ* measurement location, DHI extracted 3x3 pixels from time-series of satellite-derived chlorophyll-*a* concentrations. Then the median value of the nine pixels for each location was calculated. The

median was only used if all nine pixels were correctly extracted, because missing data indicates cloud coverage or other atmospheric artefacts in data, which could hamper the chlorophyll-*a* retrieval. Satellite data collected during winter months at high latitudes has a higher uncertainty because there is often not enough light available for the satellite sensor to take reliable measurements. Therefore, all the data collected from November to February were removed.

Since the waters around Denmark are optically complex and bio-optical water constituents derived from universal satellite data processing algorithms are often not able to represent this complexity, the MERIS data was tuned to local conditions and optimized to better capture higher chlorophyll-*a* values and algal blooms. The chlorophyll-*a* concentrations were retrieved using the algorithm: MERIS Case 2 Regional (C2R version 1.6.2) (Doerffer & Schiller 2007; Doerffer & Schiller 2008). Prior to the retrieval, a separate NN module performs atmospheric correction, using MERIS L1b geometric and reflectance data. A local adaptation derived by DHI GRAS based on calibration factors derived from NOVANA *in situ* measurements was applied to the original algorithm.

The MODIS data has undergone standard processing for case 2 waters using the Seadas version 7.3 with OC3 algorithm with MUMM correction and MUMM NIR calculation (Ruddick et al., 2000), an atmospheric correction procedure for turbid waters. No calibration to local conditions was carried out. In Annex 1 Figure A1.3, examples of the comparison between the satellite data and NOVANA observations are included. In Figure 4.1 MERIS and MODIS are plotted against *in situ* and fluorescence measures with the associated r-square value for station 925 and 6700053. It should be noted, that since coinciding measurements derived from satellites and *in situ*/fluorescence measures are rare, satellite data are matched with *in situ*/fluorescence data within a time range of one week before and after the satellite derived measure. Time series plots comparing measured and satellite based chlorophyll-*a* values are shown in figure A1.3.

As illustrated in this project, several different data sources and methods exist (e.g. traditional bottle measurements, Ferrybox data, Satellite data and to some degree also mechanistic models) creating new possibilities to be included in the national monitoring of the open (and to some degree also coastal) waters. Not all sources and methods can be applied for monitoring alone, but methods for extracting the information contained in the different data and combining these different sources into aggregated products give huge potential for significantly strengthening the monitoring of the MSFD sub-regions and their subdivisions, at minimum additional cost.

All data presented in this report represents a value, but the precision of the single observation is different and there are large differences in what the observations represent horizontally and vertically:

- Traditional bottle measurements are typically very precise, but the observation only represents the collected water volume in the bottle. The collected water volume does not always represent an entire water body as is evident in the Figure 4.2, but in most samples, we do not know how representative the observation is. On the other hand, the traditional bottle measurements are often extracted in many depths beside the surface samples, and analysed for several parameters beside chlorophyll-*a*. This contrasts with most of the other observations included in this report.
- Ferrybox measurements are also precise, but the methodology used differs from the *in situ* method. Temporal resolution is high compared to traditional ship-based measurements. The spatial resolution is also high along the route of the ferry, and using a spatially explicit adjustment that takes latitude into account, chlorophyll-*a* can be reasonably well estimated throughout its route through Danish waters. A drawback is that these measurements are taken from a single depth and that the translation between the different monitoring stations are different from location-to-location.
- The advantage of satellite data is the simultaneous spatial coverage. Here satellite data provides time series of surface chlorophyll-*a* concentrations covering large geographical areas, and hence, documents spatial distributions and variations. A disadvantage, however, is that it can be difficult to get accurate chlorophyll-*a* retrievals in complex coastal waters, such as the Danish Straits, with a high

content of humic substances and CDOM which influence the optical properties of the water. However, where the absolute values are associated with some uncertainties the relative differences and spatial patterns are considered much more precise.

- Also, we included mechanistic models as a supplement to monitoring. Models allows for estimated 3D time series of many parameters measured in NOVANA. Models, however, build on assumptions, and hence, are simple, limited and pragmatic reproductions of reality. Especially, when the reality 'behaves' surprisingly – e.g. the large winter/spring bloom between 2009 and 2010 or the *Pseudochattonella* sp. bloom in spring 2011 – models have difficulty in predicting these events and are not optimised to do this, as we do not understand why such events occur. These limitations are exacerbated by the highly dynamic nature of the oceanographic conditions at the interface between the Baltic Sea and the Kattegat. The strength of applying models, lies within the models' capabilities of predicting the effects on e.g. chlorophyll-*a* concentrations of reducing nutrient loadings or e.g. when evaluating the importance of grazing pressure from zooplankton. Hence, there is a difference between how direct observations and model results can be integrated as part of the monitoring.

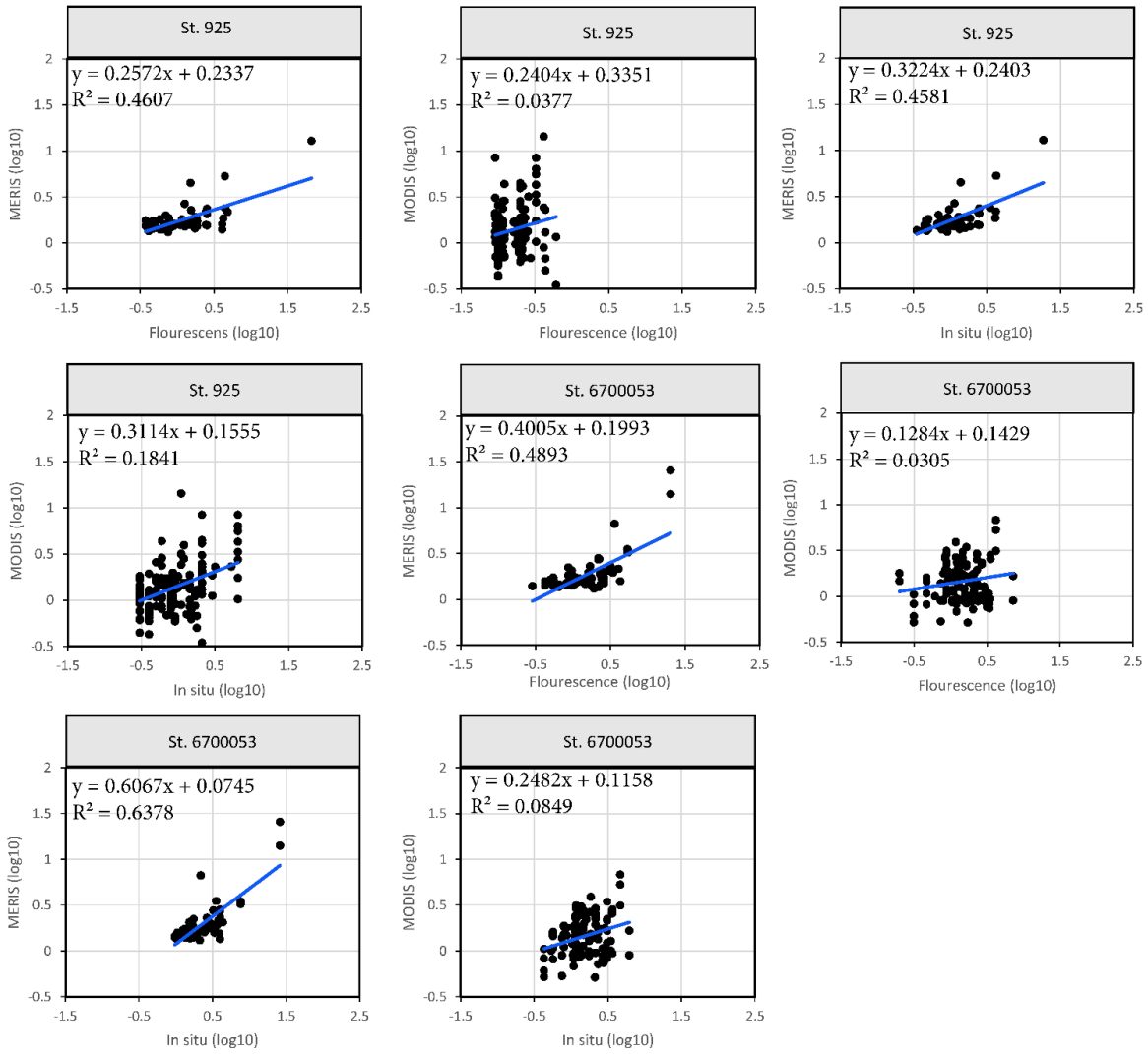


Figure 4.1 *Relationship between satellite derived log-transformed chlorophyll-a concentrations (MODIS and MERIS) and measured log-transformed chlorophyll-a values and fluorescence for the two stations 925 and 6700053. R-square values are given together with the regression. The blue line is a linear regression fit.*

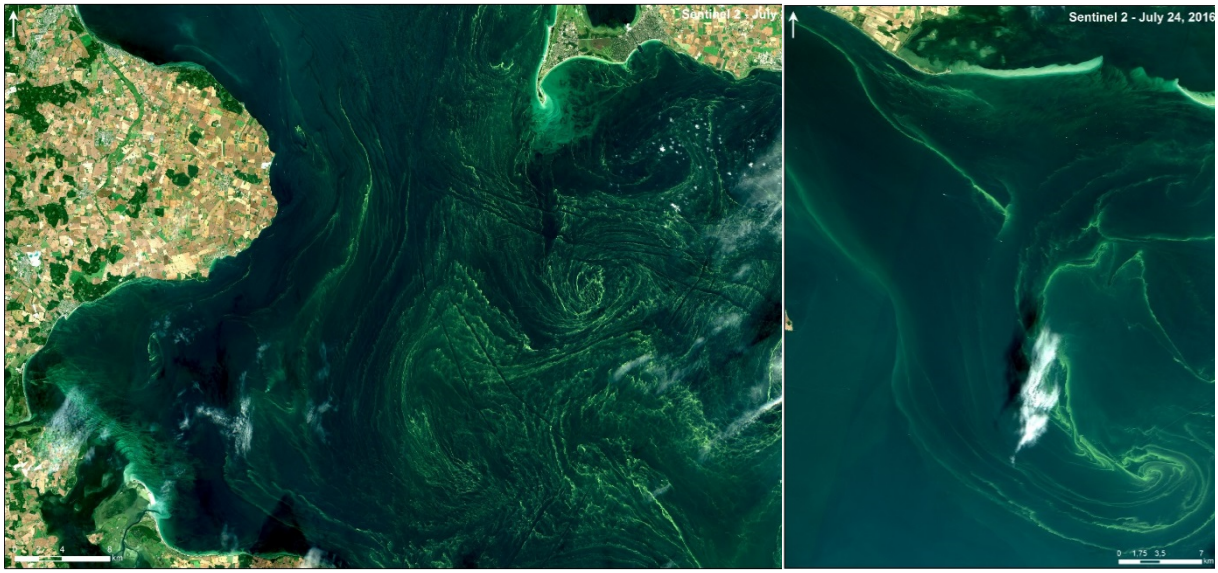


Figure 4.2 *An example on algae-patterns forming in Fakse Bay (left figure) and south of Rødsand (right figure). Both images are from the 24th of July 2016 and are derived from the Sentinel 2 satellite (© ESA, DHI GRAS).*

Looking ahead, many opportunities exist combining the different data sources and/or models. In addition, technical improvements comprise opportunities as e.g. the Sentinel 2 satellite. From 2017, we can expect the possibility of Sentinel 2 data every 5 days. With such data, we expect to be able to supplement the chlorophyll-*a* data with other and detailed observations on both turbidity, coast stability and some marine habitats. Also, the enhancement of the Ferrybox network covering more transects, and potentially more parameters, will support the traditional monitoring.

Finally, a great deal of work integrating data and models (data assimilation) is being carried out worldwide, and here mechanistic models will be an additional data source and method for the ongoing monitoring.

4.3 Conclusions

Based on the analysis of Ferrybox data and NOVANA data, we conclude the following:

1. the temporal resolution of the data set from the Oslo-Kiel ferry, sampled by the Ferrybox system, is much higher than sampling frequencies of NOVANA activities
2. an added value of the Ferrybox data set is the high spatial resolution along the Oslo-Kiel transect
3. the adjustment of the Ferrybox data set for spatial and seasonal effects leads to a data product that can be directly combined with NOVANA data

Hence, we suggest that the Ferrybox data sets and the data product based on the Ferrybox data sets are a useful and important addition to the NOVANA activities. Combining the data from these two monitoring activities, e.g. by merging the Ferrybox-based data product with the NOVANA data is straightforward, seen from a technical perspective.

In a wider perspective, the development of so-called composite products, i.e. a chlorophyll-*a* data set or map based on chlorophyll-*a* measurements from different sources would be useful in the context of an ecosystem-based approach to management of coastal and marine waters. A suite of activities, e.g. development, testing and verification is required, but the outcomes of this study clearly indicate that this path seem worthwhile to follow.

5 References

Ærtebjerg, G.N., T.S. Jacobsen, E. Gargas & E. Buch (1981): *The Belt Project. Evaluation of the Physical, Chemical and Biological Measurements*. The National Agency of Environmental Protection, Denmark. 122 pp.

Andersen, J. H. (2012): *Ecosystem-based management of coastal eutrophication. Connecting science, policy and society. Ph.D. thesis*. University of Copenhagen https://doi.org/10.13140/RG.2.1.1018.5444.

Andersen, J. H., E. Kallenbach, C. Murray & A.B. Ledang (2016): *Eutrophication in the Danish parts of the North Sea, Skagerrak and Kattegat 2006-2014. A literature-based status assessment. 43*. Norsk institutt for vannforskning. Retrieved from https://brage.bibsys.no/xmlui/handle/11250/2406499

ATV (1990): Vandmiljøplanens tilblivelse og iværksættelse. Akademiet for de tekniske Videnskaber. ISBN 87-88699-74-9. 152 pp. (in Danish)

Bijl, L. van der, S. Boutrup & P.N. Jensen (red.) (2007): NOVANA. *Det nationale program for overvågning af vandmiljøet og naturen. Programbeskrivelse 2007-09 – del 2*. Danmarks Miljøundersøgelser, Aarhus Universitet. 120 s. – Faglig rapport fra DMU nr. 615. http://www.dmu.dk/Pub/FR615.pdf. (in Danish)

Burnham, K.P., & D.R. Anderson (2004): Multimodel Inference. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Doerffer, R., & H. Schiller (2007): The MERIS case 2 water algorithm. *International Journal of Remote Sensing*, 28(3-4), 517–535.

Doerffer, R., & H. Schiller (2008): *MERIS Regional Coastal and Lake Case 2 Water Project - Atmospheric Correction ATBD*. 21502 Geesthacht, Germany, Version 1.0 18. May 2008: GKSS Research Center.

Fleming-Lehtinen, V., J.H. Andersen, J. Carstensen, E. Łysiak-Pastuszak, C. Murray, M. Pyhälä & M. Laamanen (2015): Recent developments in assessment methodology reveal that the Baltic Sea eutrophication problem is expanding. *Ecological Indicators*, *48*, 380–388. https://doi.org/10.1016/j.ecolind.2014.08.022

Hastie, T., & R. Tibshirani (1990): *Generalized additive models*. Chapman and Hall. Retrieved from https://www.crcpress.com/Generalized-Additive-Models/Hastie-Tibshirani/p/book/9780412343902

HELCOM (2009): *Eutrophication in the Baltic Sea – An integrated thematic assessment of the effects of nutrient enrichment and eutrophication in the Baltic Sea region. Baltic Sea Environmental Prooceedings No. 115B.* Retrieved from https://www.eea.europa.eu/data-and-maps/indicators/nutrients-in-transitional-coastal-and-3/helcom-2009.-eutrophication-in-the

Huot, Y., & M. Babin (2010): Overview of Fluorescence Protocols: Theory, Basic Concepts, and Practice. In *Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications* (pp. 31–74). Dordrecht: Springer Netherlands https://doi.org/10.1007/978-90-481-9268-73

Hurvich, C.M., & C.L. Tsai (1989): Regression and time series model selection in small samples. *Biometrika*, *76*, 297–307.

Johnsen, G., & E. Sakshaug (2007): Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry. *Journal of Phycology*, *43*(6), 1236–1251. https://doi.org/10.1111/j.1529-8817.2007.00422.x

Markager, S. & H. Fossing (2014): *Klorofyl a koncentration*. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi. Teknisk anvisning fra Det Marine Fagdatacenter, DCE, nr. M07 ver.2. http://pure.au.dk/portal/files/69209669/TA_M07_Klorofyl_a_VERSION_2_.pdf (in Danish)

Miljøministeriet (1989): *Vandmiljøplanens overvågningsprogram*. Miljøprojekt 115. Miljøstyrelsen, Miljøministeriet, København. (in Danish)

Miljøstyrelsen (2000): *NOVA-2003. Programbeskrivelse for det nationale program for overvågning af vandmiljøet 1998-2003. Redegørelse nr.1*. Miljøstyrelsen, Miljø- og Energiministeriet, København. (in Danish)

Miljøstyrelsen (1993): *Vandmiljøplanens overvågningsprogram 1993-1997. - Redegørelse fra Miljøstyrelsen nr. 2/1993*, Miljøstyrelsen, København. (in Danish)

Naturstyrelsen (2011): *Det Nationale Overvågningsprogram for Vand og Natur. NOVANA 2011-2015. Programbeskrivelse - del 1.* Miljøministeriet, København. (in Danish)

OSPAR (2017): *Eutrophication Status of the OSPAR Maritime Area. Third OSPAR Integrated Report on the Eutrophication Status of the OSPAR Maritime Area, 2006-2014*. Retrieved from https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/pressures-human-activities/eutrophication/third-comp-summary-eutrophication/

Rodríguez, F., M. Chauton, G. Johnsen, K. Andresen, L.M. Olsen & M. Zapata (2006): Photoacclimation in phytoplankton: implications for biomass estimates, pigment functionality and chemotaxonomy. *Marine Biology*, *148*(5), 963–971. https://doi.org/10.1007/s00227-005-0138-7

Ruddick, K., De Cauwer, V., Park, Y. -J., & G. Moore (2006): Seaborne measurements of near infrared water - leaving reflectance: The similarity spectrum for turbid waters. *Limnol. Oceanogr.* 51, 1167–1179

Svendsen, L.M., L. van der Bijl, S. Boutrup & B. Norup (Eds.) (2005): *NOVANA. Det nationale program for overvågning af vandmiljøet og naturen. Programbeskrivelse – del 1.* Danmarks Miljøundersøgelser. 48 s. – Faglig rapport fra DMU nr. 495. (in Danish)

Suggett, D., C. Moore, A. Hickman & R. Geider (2009): Interpretation of fast repetition rate (FRR) fluorescence: signatures of phytoplankton community structure versus physiological state. *Marine Ecology Progress Series*, *376*, 1–19. https://doi.org/10.3354/meps07830

Volent, Z., G. Johnsen, E.K. Hovland, A. Folkestad, L.M. Olsen, K. Tangen & K. Sørensen (2011): Improved monitoring of phytoplankton bloom dynamics in a Norwegian fjord by integrating satellite data, pigment analysis, and Ferrybox data with a coastal observation network. *Journal of Applied Remote Sensing, 5*(1), 53561. https://doi.org/10.1117/1.3658032

Annex 1: Supplementary material

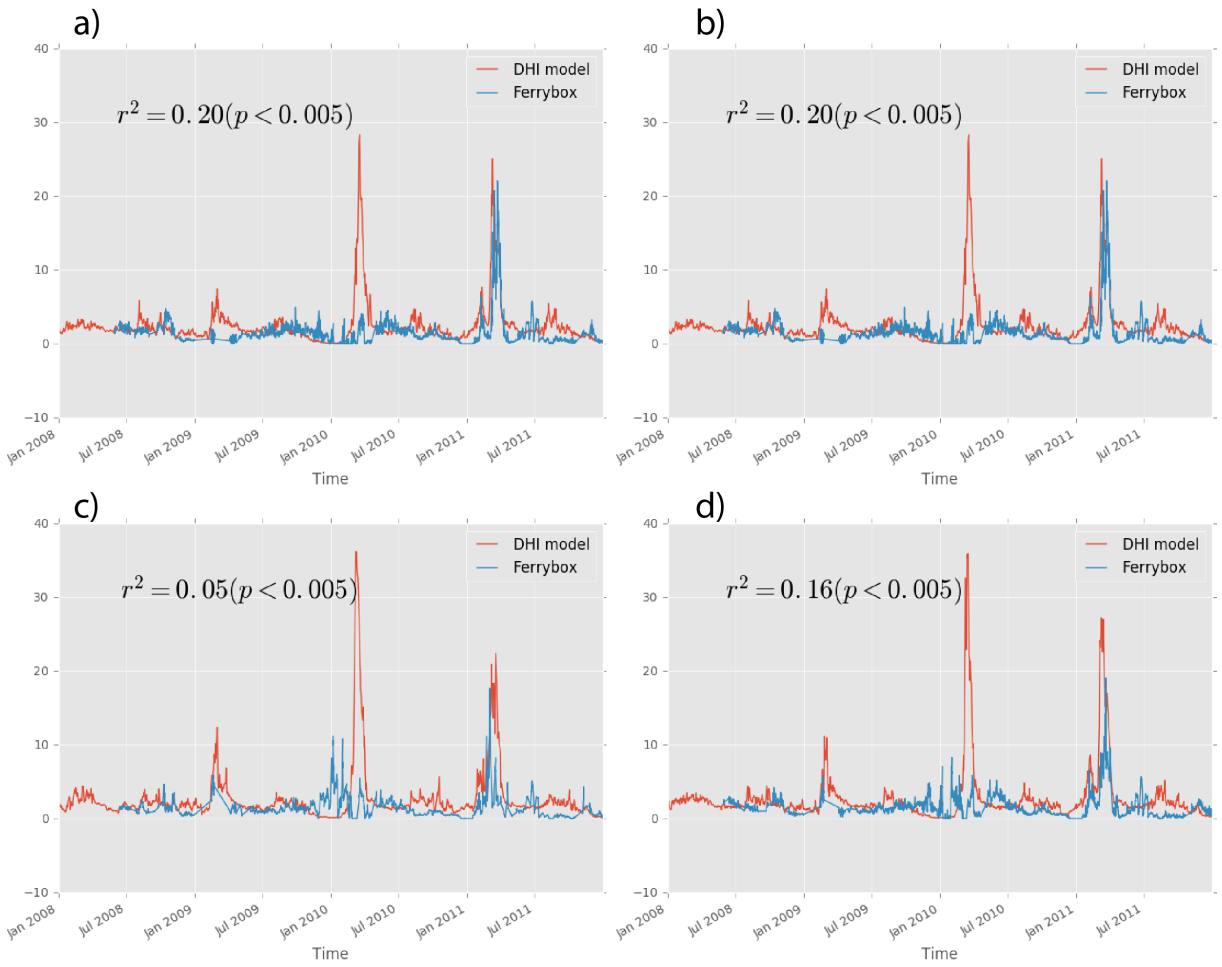


Figure A1.1 Comparison between chlorophyll-a from Ferrybox observations and ocean model (DHI) from selected stations a) 935, b) 939 c) 1007, and d) 925. Ferrybox data measured within a 5-km radius of the geographical position for each station were extracted and plotted. Correlations between model and Ferrybox were calculated using Pearsons linear correlation algorithm. Data from station N3 and 6700053 for 5 km radius was not available.

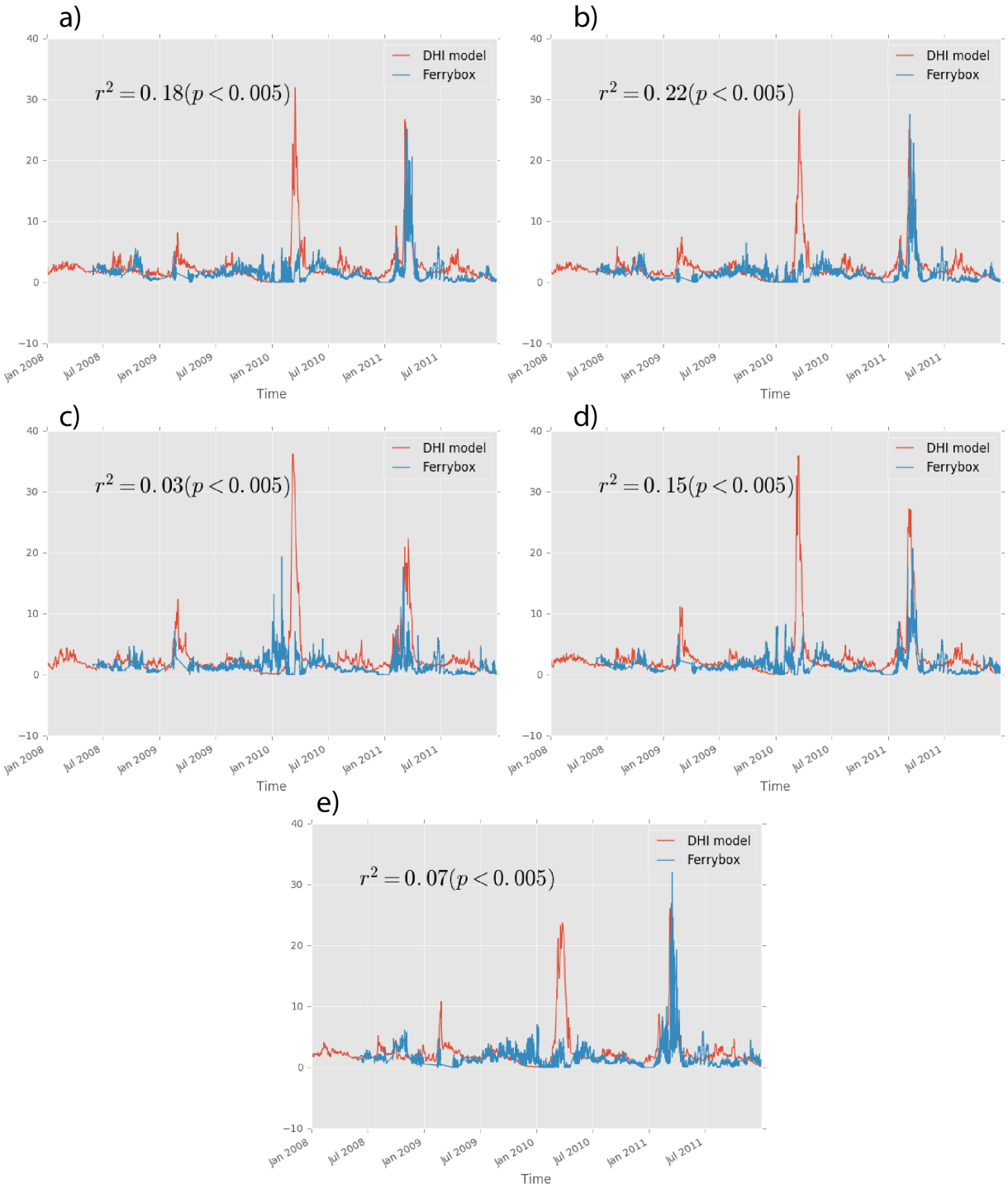
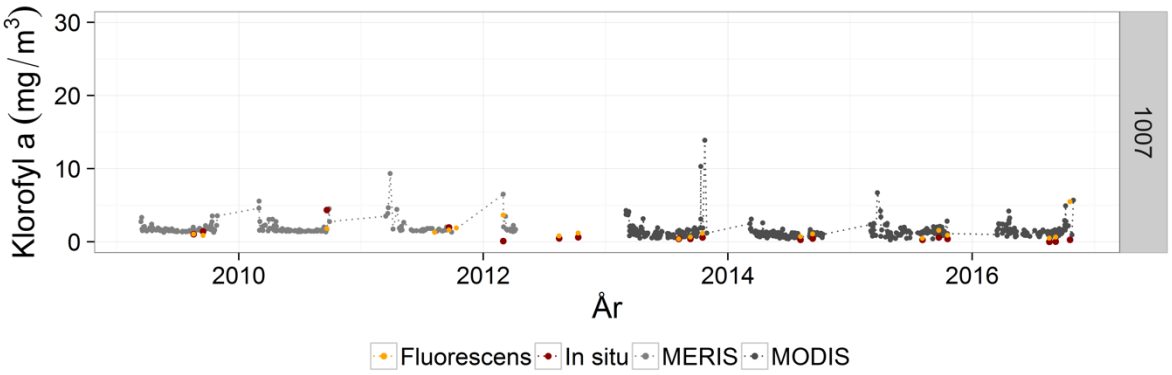
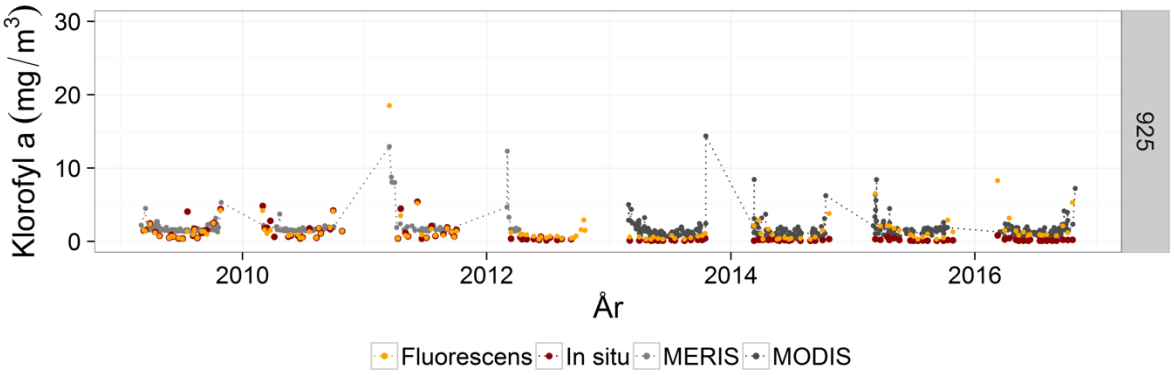
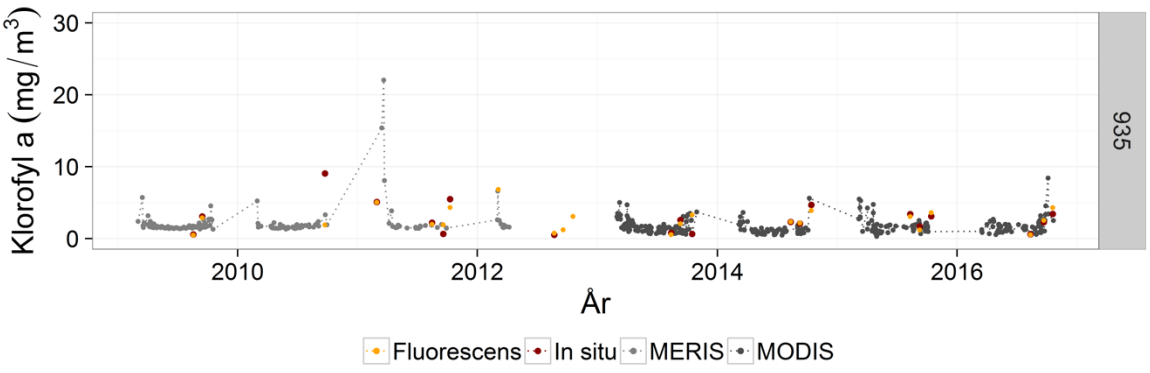
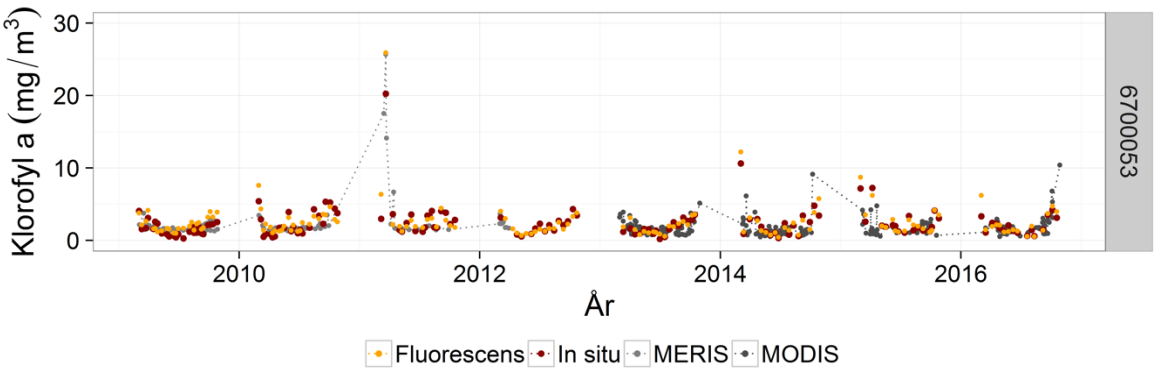


Figure A1.2 Comparison between chlorophyll-a from Ferrybox observations and ocean model (DHI) from selected stations a) 935, b) 939 c) 1007, d) 925 and e) N3. Ferrybox data measured within a 20-km radius of the geographical position for each station were extracted and plotted. Correlations between model and Ferrybox were calculated using Pearsons linear correlation algorithm. Data from station 6700053 for 25 km radius was not available.


Klorofyl a (mg/m3)
30
20
10
0
2010
2012
2014
2016
År
1007
Fluorescens
In situ
MERIS
MODIS


Klorofyl a (mg/m3)
30
20
10
0
2010
2012
2014
2016
År
925
Fluorescens
In situ
MERIS
MODIS


Klorofyl a (mg/m3)
30
20
10
0
2010
2012
2014
2016
År
935
Fluorescens
In situ
MERIS
MODIS


Klorofyl a (mg/m3)
30
20
10
0
2010
2012
2014
2016
År
6700053
Fluorescens
In situ
MERIS
MODIS

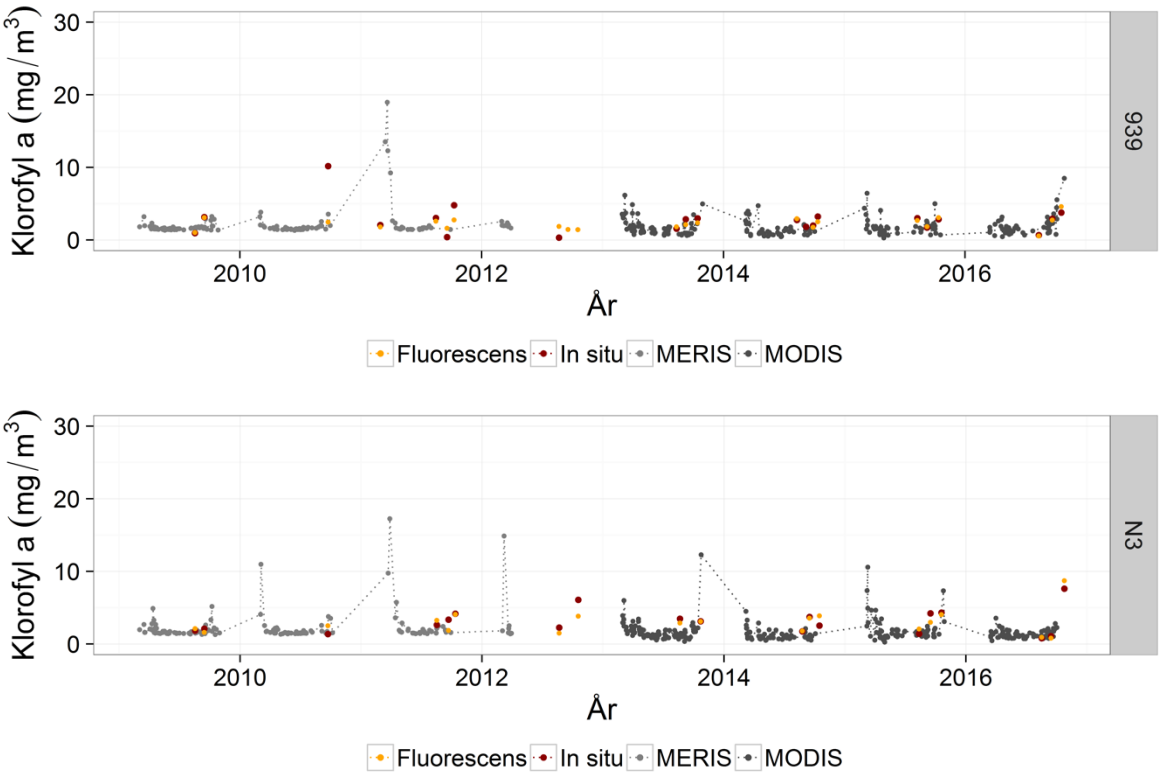


Figure A1.3 Comparison between satellite derived chlorophyll-a concentrations (grey dots and dotted lines) and measured chlorophyll-a values averaged over the top 3 m (orange dots) from north (top figures) towards south (bottom figures).

NIVA Denmark is the name, water is our game

NIVA Denmark Water Research is a regional office of the Norwegian Institute for Water Research (NIVA) and has just recently been established to resolve environmental issues concerning the freshwater and marine systems that relate to Denmark.

NIVA Denmark has primary focus on research-based implementation of a number of EU's directives inter alia the Water Framework Directive and the Marine Strategy Framework Directive together with international conventions (HELCOM, OSPAR, BDC). We occasionally provide consultancy to authorities and small and medium-sized companies.

NIVA Denmark is a a place for practice, observation, testing and synthesis. Key research and test areas include eutrophication, hazardous substances, biodiversity, and ecosystem health as well as the implications of multiple human activities in marine waters and in streams, rivers and lakes. We develop indicators, monitoring methods and tools to assess the state of an ecosystem in order to carry out analyses and contribute to evidence based and sustainable solutions to the challenges we and the environment face.

NIVA Denmark, as a regional office to NIVA has thus the backing of more than 200 dedicated researchers and experts.

NIVA Denmark

NIVA Denmark
Winghouse
Ørestads Boulevard 73
DK-2300 Copenhagen
Telephone: +45 8896 9670
E-mail: post@niva-denmark.dk
www.niva-denmark.dk