



KONSOLIDERING AF PLANKTONDATA I NOVANA

Fytoplanktondata

Teknisk rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 287

2023



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

KONSOLIDERING AF PLANKTONDATA I NOVANA

Fytoplanktondata

Teknisk rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 287

2023

Hans Jakobsen
David Rytter
Per Andersen
Andreas Holbach

Aarhus Universitet, Department of Ecoscience



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

Serietitel og nummer:	Teknisk rapport fra DCE - Nationalt Center for Miljø og Energi nr. 287
Kategori:	Rådgivningsrapporter
Titel:	Konsolidering af planktondata i NOVANA
Undertitel:	Fytoplanktondata
Forfatter(e): Institution(er):	Hans Jakobsen, David Rytter, Per Andersen, Andreas Holbach Aarhus Universitet, Department of Ecoscience
Udgiver: URL:	Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi © http://dce.au.dk
Udgivelsesår: Redaktion afsluttet:	Oktober 2023 Oktober 2023
Faglig kommentering: Kvalitetssikring, DCE: Sproglig kvalitetssikring:	Peter Henriksen Lars Moeslund Svendsen Else Vihlborg Staalsen
Ekstern kommentering:	Miljøstyrelsen (MST). Kommentarerne findes her: http://dce2.au.dk/pub/komm/TR287_komm.pdf
Finansiell støtte:	Denne rapport er finansieret af en allonge fra miljøstyrelsen.
Bedes citeret:	Jakobsen, H; Rytter, D; Andersen, P; Holbach, A. 2023. Konsolidering af planktondata i NOVANA. Fytoplanktondata. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 52 s. Teknisk rapport nr. 287
	Gengivelse tilladt med tydelig kildeangivelse
Sammenfatning:	Denne rapport beskriver arbejdet med at indlæse og kvalitetssikre NOVANA planktondata, indsamlet under de tidligere amter, samt data fra Miljøstyrelsens STOQ-database. Data er blevet gennemløbet adskillige iterationer i kvalitetssikringsprocessen, og der er blevet foretaget en del korrektioner i konverteringskonstanterne. Der er udført to typer kvalitetssikring: 1. multivariat kontrol samt 2. kvalitetssikring efter metoden udviklet af AU. Det endelige godkendte datasæt omfatter 15.809 unikke optalte prøver af fyto- og heterotroft protist plankton
Emneord:	kvalitetssikring, fytoplankton, heterotroft protist plankton, NOVANA, STOQ, overvågning
Foto forside:	Forskellige kiselager” Fortografen er Jakobsen HH
ISBN: ISSN (elektronisk):	978-87-7156-810-3 2244-9991
Sideantal:	52

Indhold

Datablad	2
Sammenfatning	5
Summary	6
Introduktion	7
1 Metoder	9
1.1 Metodebeskrivelse	9
1.2 MST Faglig kvalitetssikring	11
1.3 FDC-faglig kvalitetssikring	14
2 Resultat	17
2.1 MST faglig kvalitetssikring	18
2.2 FDC faglig kvalitetssikring	26
3 Afsluttende bemærkninger og konklusion	34
4 Referencer	36
5 Bilag	38
Bilag B1 Tabel over stationer, der er samlet og struktureret i ODA	38
Bilag B2 Oversigt over de heterotrofe protister	43
Bilag B3 Udvalgte stationer som eksempel på variationer	47
Bilag B4 NOVANA stationer med heterotrofe protister	48
Bilag B5 QA_Dublet_klassifikation	49
Bilag B6 QA_Dublet_klassifikation_20021024	50

Sammenfatning

Denne rapport beskriver kvalitetssikringen af historiske NOVANA planktondata. I første del blev en række databaser fra de nedlagte amters miljøovervågning samt data fra Miljøstyrelsens STOQ-database indlæst. Disse data inkluderede information om monitoreringen af fyto- og mikrozooplankton. Denne proces involverede omfattende arbejde med at identificere dubletter på tværs af de forskellige databaser, som til sidst blev samlet til et sammenhængende datasæt i ODA-format.

I samarbejde med Miljøstyrelsen blev 46 prøvetagningsstationer prioriteret og inkluderet i den endelige kvalitetssikring. En delmængde af disse stationer (11 i alt) blev overvåget for heterotroft protist-plankton i kortere eller længere perioder. I alt blev der indlæst 16.377 optalte prøver, der blev identificeret som unikke. Derefter blev der udført kvalitetssikring i henhold til standarderne beskrevet i Dansk Standard 16269-4:2010 (ISO 16269-4:2010(E)). Denne omfattede anvendelse af to statistiske metoder til multivariat kontrol af afvigelser, baseret på (1) Mahalanobis-afstand og (2) Cooks-afstand.

Herefter blev hele datasættet vurderet i overensstemmelse med retningslinjerne for fytoplanktondata udviklet af AU, som er implementeret i de årlige NOVANA-afrapporteringer for marine områder.

Kvalitetssikringen blev udført iterativt i flere trin. I den første kørsel blev der identificeret betydelige afvigelser. Det blev konstateret, at der manglede bio- og plasmavolumener for mange af de ældre data på artsniveau, og at der ikke var overensstemmelse mellem kulfaktorerne og de tekniske anvisninger. Efter erstatning af manglende bio- og plasmavolumener med standardværdier og opdatering af kulfaktorerne blev de anvendte algoritmer til afvigelseskontrol markant forbedret. Efter den endelige kvalitetssikring blev der godkendt i alt 15.809 prøver af planktondata.

Summary

This report describes the quality assurance process for historical NOVANA plankton data. In the first step, a series of databases from the discontinued county (DK: AMT) environmental monitoring and data from the Danish Environmental Protection Agency's STOQ database were loaded. These data included information on the monitoring of phyto- and microzooplankton. This process involved extensive work to identify duplicates across the different databases, which were eventually compiled into a cohesive dataset in ODA format.

In collaboration with the Danish Environmental Protection Agency, 46 sampling stations were prioritized and included in the final quality assurance process. A subset of these stations (11 in total) was monitored for heterotrophic protist plankton for varying periods. In total, 16,377 samples were loaded and identified as unique. Subsequently, quality assurance was conducted according to the standards described in Danish Standard 16269-4:2010 (ISO 16269-4:2010(E)). This involved the application of two statistical methods for multivariate control of deviations based on (1) Mahalanobis distance and (2) Cook's distance.

Following this, the entire dataset was evaluated in accordance with the phytoplankton data guidelines developed by Aarhus University, which are implemented in the annual NOVANA reports for marine areas. The quality assurance process was carried out iteratively in multiple steps. In the initial run, significant deviations were identified. It was observed that there were missing bio- and plasma-volumes for many of the older data at the species level, and there was inconsistency between the carbon conversion factors and the technical instructions. After replacing missing bio- and plasma-volumes with standard values and updating the carbon factors, the algorithms used for deviation control were significantly improved. Following the final quality assurance, a total of 15,809 samples of plankton data were approved.

Introduktion

DCE har indgået en kontrakt med Miljøstyrelsen (MST) omkring konsolidering af marine fyto- og mikrozooplanktondata (herefter kaldet "planktondata") på baggrund af en projektbeskrivelse fra juli 2022 "Tilbud på konsolidering af planktondata". Formålet med projektet var at sikre historiske planktondata, der ikke allerede var blevet overført til fagsystemet STOQ og/eller DCE's overfladevandsdatabase ODA. Herudover var formålet at det samlede datasæt blev harmoniseret og indlagt i ODA og blev kvalitetssikret (KS) til niveau MST fagligt KS og fagdatacenter fagligt kvalitetssikret efter retningslinjerne i aftalen mellem Miljøministeriet og Aarhus Universitet (Høgslund m.fl., 2019). Data vil så være klar til på et senere tidspunkt i et separat projekt til at kunne overføres til VanDa.

Før nærværende projekt blev der i 2021 gennemført et andet projekt for MST, der primært omfattede opgaver i relation til udvikling af funktionaliteter for plankton i relation til VanDa databasen (fælleoffentlige overfladevandsdatabase som ligger hos Danmarks Miljøportal) og at få tilvejebragt og modtaget data fra historiske databaser hos bl.a. WSP.

Planktondata er blevet indsamlet under NOVANA (Det Nationale Overvågningsprogram for Vandmiljø og Natur) og fra forudgående overvågningsprogrammer indsamlet af de tidligere amter, og som har været gemt i flere fagsystemer. En del planktondata lå i fagsystemet STOQ¹, men herudover findes der planktondata i andre databaser:

Tre versioner af fagsystemet Planktonsys

Miljøbiologisk Laboratoriums ved Kirsten Olrik

DCE's ODA.

Der har tidligere været gjort forskellige forsøg på at sikre de historiske databaser, og der har været eksempler på, at de forskellige backupper efterfølgende er blevet ændret uafhængigt af hinanden. En stor del af disse data er ikke blevet kvalitetssikret af hverken MST eller af DCE, og der udestod derfor en væsentlig opgave med at kvalitetssikre planktondata, der i nogle tilfælde er over fyrré år gamle.

I projektbeskrivelse angav DCE at "Alle tidligere kvalitetsmærker i STOQ er slettet, og dertil er omfang af data, der kan overføres til berigelse af ODA ukendt. Det betyder at omfanget af opgaven til faglig og FDC faglig kvalitets sikring (KS) er ukendt, og DCE tager derfor forbehold for dette i det endelige produkt." De kopier DCE modtog af nogle af de gamle databaser fulgte ikke den struktur og database format som i STOQ, som var forudsat i DCE's tilbud, og det krævede mange iteration at få fremskaffet alle nødvendige data og metadata som har medført at DCE har måtte lægge en del egne ressourcer i

¹ DCE har en fuld kopi af STOQ fra det tidspunkt den blev nedlagt, som er anvendt i dette projekt. Denne STOQ-version indeholder planktondata ud over hvad der har været overført til VanDa

opgaveløsningen. Nærværende notat beskriver det arbejde som det har været muligt at gennemføre.

På trods af de omtalte udfordringer vurderes plankton data til overordnet at være i god forfatning, da datatypen er oparbejdet efter et fælles metodegrundlag (Edler 1979, Olrik 1991, CEN 2011, 2015), og opbevaringen af data i et struktureret databaseformat har fulgt internationale standarder/anbefalinger (Zingone m.fl., 2015). Dertil har en solid taksonomisk forskning centreret omkring universiteterne i Danmark samt et konsulentmiljø med begrænset udskiftning i bemanningen, sikret en generel konsistens i data gennem årene.

Fyto- og mikrozooplankton blev til og med 2015 rapporteret i den årlige havrapport "Marine områder 2015 NOVANA" (Hansen m.fl., 2015). MST KS er vel-dokumenteret siden overgangen til STOQ-databasesystemet, men der var usikkerhed på omfanget af KS historisk, der tidligere er fundet sted i regi af de gamle amter, da DCE har ikke kunnet finde en metodebeskrivelse for tidlige gennemførte KS.

Usikkerheden har betydet, at DCE frem til og med 2015 har gennemført beregningen af kulstofbiomasse i forbindelse med i de årlige havrapporter. Siden 2015 har planktondata ikke været rapporteret i forbindelse med den årlige havrapport, og der er ligeledes ikke udført FDC KS af planktondata.

Projektet har omfattet tre trin:

1. Aggregering af datakilder til en fælles datastruktur samt identificering af mulige dubletter.
2. MST KS svarende til nedenstående definitioner.
3. Fagdatacentret faglige kvalitetssikring (FDC KS).

Efter gennemførelsen af dataaggregeringen er der blevet etableret en generel strukturering af alle data til en fælles ODA databasestruktur. Det betyder, at information som tidspunkter, enheder, koder, størrelser, omregningskonstanter, stadier osv. nu er blevet harmoniseret. Det skal også bemærkes at begreberne MST og FDC KS rummer en del overlap da begge procedurer har indbyrdes afhængigheder. Ligeledes er MST KS ikke identisk med kvalitetssikring, der er gennemført af MST under og efter indlæsning i STOQ, men omfatter i dette projekt en række ISO certificerede statistiske kvalitetssikringsmetoder.

Det endelige dataprodukt vil være afstemt efter de tekniske anvisninger for fytoplankton TA09 og TA10 (Fossing & Jakobsen 2015, Jakobsen & Fossing 2017). Det færdige dataprodukt var tilgængelig i ODA databasen fra 31. december 2022, og ved gennemførelsen af FDC KS vil data dermed i princippet være godkendt af AU til at kunne overføres til MSTs' VanDa database, når det aftales.

Projektet er igangsat 22. august 2022 efter endelig godkendelse af MST og er afsluttet 31. december 2022 og en rapportering af projektet i 2023.

1 Metoder

1.1 Metodebeskrivelse

Konsolidering og klargøring til en mulig fremtidig overførsel til VanDA af planktondata omfatter tre trin:

1. Aggregering af datakilder til en fælles datastruktur samt identificering af mulige dubletter.
2. MST KS svarende til nedenstående definitioner.
3. Fagdatacentret faglige kvalitetssikring (FDC KS).

Aggregerede historiske data i ODA

Alle planktondata blev indlagt i ODA fra tre versioner af planktonsys-databasen (som ligger hos firmaet WSP, det tidligere ORBICON, Ringkøbing og Aarhus amter), data fra firmaet Miljøbiologisk Laboratorium og fra MST (fra STOQ-databasen) (tabel 1.1).

Databasestrukturer på tværs af datakilderne (tabel 1.1), blev afklaret gennem et omfattende udredningsarbejde, så data kunne indlægges i en fælles datastruktur i ODA.

I den oprindelige aftale med MST var det forudsætningen for dette udredningsarbejde, at datastrukturen var ens på tværs af datakilder. Data i STOQ foreligger i et fast struktureret format, der ikke har ændret sig siden i ibrugtagning. Derimod er de øvrige datakilder ikke løbende vedligeholdt, og datastrukturen er forskellig mellem disse. Harmoniseringen af disse data viste sig at være meget ressourcekrævende, da struktur, koder og enheder har været afvigende i de individuelle datakilder angivet i tabel 1.1.

Tabel 1.1. Oversigt over antallet af plankton (fyto og microzooplankton) prøver AU modtog fra de forskellige kilder

Alias	Beskrivelse	Antal prøver
MBL	Miljøbiologisk Laboratorium	1.717
OCP	ORBICON	12.052
RKBP	Ringkøbing amt	6.812
AAP	Aarhus amt	3.601
STOQP	STOQ fytoplankton	11.742

Slutproduktet efter dataaggregeringen blev en generel strukturering af alle data til en fælles ODA databasestruktur. Det betyder, at information som tidspunkter, enheder, koder, størrelser, omregningskonstanter, stadier osv. er blevet harmoniseret.

Der viste sig at være et stort overlap mellem planktondata på tværs af de databaser, som DCE modtog kopi af, samt en række forkerte stationsangivelser, hvilket betød, at der var en stort mængde redundante data (dubletter), som skulle udredes, før data kunne kvalitetssikres i en kopi af ODA.

Dublet håndtering

Efter samling af alle data i ODA blev der med udgangspunkt i dato udtrukket en liste med prøver, hvor der var mere end én prøve, der repræsenterer potentielle dubletter. Station og dybde blev ikke medtaget, da der viste sig at være dubletter på tværs af stationer og dybder.

På potentielle dubletter registreredes nøgledata, så som ID'er, antal arter, antal dimensioner, artsprofiler², dimensionsprofiler, cellekoncentration, cellevolumen og evt. andre parametre. For at sikre sammenlignelighed blev en generel konvertering af arter i artsprofilen foretaget, da artkoder kan være konverteret forskelligt på tværs af datakilderne i tabel 1.1. Herfra genereredes en liste med kombinationer af potentielle dubletter, som kunne sammenlignes parvist. Derved kunne der identificeres i størrelsesordenen 170.000 potentielle dubletprøver. Dette blev efterfulgt af en sammenligning af prøveprofilerne for de potentielle dubletter på baggrund af station, prøvetager, analytiker, tidspunkt, dybder, antal arter, antal dimensioner, artsprofiler, dimensionsprofiler, parametre og summeringer som samlet resulterede i en dublet score. På baggrund af denne score og en manuel vurdering blev et bud på dubletter givet station for station (svarende til +200 station kombinationer). Denne vurdering blev givet til faglig ekspert for en endelig faglig vurdering.

Følgende prioriterings rækkefølge anvendes:

1. Unikke prøver fik tildelt kvalitetsflaget "*elektronisk godkendt*" i ODA.
2. Tilsyneladende dubletter, der med sikkerhed identificeredes automatisk som unikke fik alle kvalitetsflaget "*elektronisk godkendt*".
3. Dublet, der med sikkerhed kunne identificeres automatisk som kilde, fik kvalitetsflaget elektronisk godkendt, resten i dublet gruppen fik kvalitetsflaget "*elektronisk forkastet*" (omkring 30 % af data blev identificeret som dubletter), men opbevares fortsat i ODA.

Det endelige datasæt (elektronisk godkendte data) efter aggregering og dublet-håndteringen indeholder en lang række historiske stationer oplistet i bilag B1.1. Stationer markeret med X er stationer som blev udvalgt i forbindelse med dette projekt i samarbejde med MST-medarbejdere med solidt kendskab til planktonovervågningen. De repræsenterer som første prioritet, stationer som fortsat er aktive i forbindelse under NOVANA. Som anden prioritet blev der udvalgt stationer, som ikke længere er aktive fx som en del af NOVANA, men som tidligere har været del af overvågningsprogrammer gennemført som del af den landsdækkende vandmiljøovervågning, eller som en del af "amts"-overvågning og hvor der foreligger kortere eller længere tidsserier. Disse prioriterede stationer blev videreført til MST og FDC KS. Stationer uden for prioriteringen blev rensset for dubletter og findes i struktureret form i ODA og vil på et senere tidspunkt kunne gennemgå MST og FDC KS.

Der kan være dubletter, som efterfølgende manuelt bliver afklaret. I alt blev der indlæst 377.834 artsregistreringer fordelt på 46 stationer.

I beregningen af kulstofbiomasser af fytoplanktondata indgår en række variable så som biovolumen (cellevolumen) og plasmavolumen (anvendes kun for kiselalger), celle koncentrationer osv. Alle nødvendige værdier for disse

² Artsprofil er profilen, der dannes af alle arter i en given prøve

variable blev gennemgået, og der blev tilføjet gennemsnitsværdier for biovolumen og plasmavolumen (anvendes kun for kiselalger), hvor disse manglede.

1.2 MST kvalitetssikring

Der blev udviklet et værktøj i statistikprogrammet R til MST faglig KS, som blev udført i tre trin. Kort beskrevet afslørede det første trin inkonsistente registreringer for arter, individer og morfologi, det andet trin identificerede statistisk signifikante univariate afvigelser på artsniveau, mens man i det tredje trin så på afvigelser på univariate og bivariate afvigelser på stationsniveau. De væsentlige parametre for beregningen af biovolumen og kulstofbiomasse blev taget i betragtning

Det blev aftalt mellem MST og DCE, at kvalitetssikringen af data bestod af en række kvalitetstrin. Nogle af disse blev kaldt MST faglig KS, mens de sidste trin i kæden blev kaldt FDC faglig KS. I dette projekt havde DCE til opgave at udføre begge dele.

Metoden til MST faglig KS blev udviklet særskilt for disse data under projektet og tog udgangspunkt i statistiske metoder beskrevet i Dansk Standard 16269-4:2010 (Dansk-Standard 2010) samt metoderne beskrevet i Cano *m.fl.* (2015).

MST KS udviklede sig til en iterativ proces. Afvigelser blev så vidt muligt rettet i databasen igennem harmonisering og genberegning, hvor der blev udpeget nogle systematiske fejl i datasættet (f.eks. manglende kulstofbiomasse ved eksisterende biovolumen og inkonsistente kulfaktorer for arterne). Til sidst blev data mærket som MST fagligt godkendt eller ikke fagligt godkendt i ODA.

Data til den MST KS blev indlæst efter dublethåndtering fra ODA databasen med de følgende restriktioner:

- Data godkendt efter dublettjek
- Celleantal pr. L > 0,
- Klasse skulle ikke være en del af: *Maxillopoda*, *Eurotatoria*, *Gastropoda*, *Hydrozoa*, *Bivalvia*, *Branchiopoda*, *Appendicularia*, *Malacostraca*, *Polychaeta*
- PlanktonSystemKode skulle være en del af databaserne: MBL, OCP, RKBP, AAP, STOQP (tabel 1.1)

Til MST KS blev data indlæst fra de 46 prioriterede stationer, som omfatter omkring 50 % af alle individregistreringer. Det skal noteres, at der findes en lang række artobservationer, der er markeret med "X", og som påviser, at en given art er identificeret, men som ikke indgår som en kvantitativ observation. Dette gælder typisk arter, der er fundet ved overordnet gennemgang af f.eks. planktonprøvens netprøve.

Trin 1: Overordnede fagligt kvalitetstjek

Ufuldstændige artsregistreringer

For artsregistreringer blev det undersøgt, om der efter genberegning af biovolumener og kulstofbiomasse stadig manglede valide værdier > 0 for enten biovolumen pr. celle, kulstofindehold pr. celle eller biovolumen pr. L. Alle fundne ufuldstændige artsregistreringer blev sat til under kontrol (UK) (figur 1.1).

Inkonsistente individregistreringer

Der blev undersøgt for inkonsistente individregistreringer, hvor der enten manglede værdier for biovolumen pr. celle > 0 celler per L^{-1} og hvor der var registreret målte celledimensioner, eller hvor de målte celledimensioner var 0. Alle fundne inkonsistente individregistreringer blev sat til UK (under kontrol) figur 1.1). De tilhørende artsregistreringer blev ført videre i kvalitetssikringsprocessen, da de relevante variabler blev genberegnet på artsniveau.

Inkonsistente morfologi registreringer

Valide værdier for biovolumen pr. L og kulstofbiomasse pr. L kan kun beregnes ved det samme morfologigrundlag for både opmåling og optælling. Derfor blev der undersøgt for registreringer, hvor der var inkonsistens mellem morfologi for optælling og opmåling. Alle fundne inkonsistente morfologiregistreringer blev sat til UK (figur 1.1).

Der blev udført to typer af afvigerkontrol. Vi begyndte med kontrol af enkeltcellers opmålinger med fokus på afgivelse af formelparametre, der indgik i volumenberegningerne (individniveau). Derefter fortsatte vi med afvigerkontrol af den endelige gennemsnitlige cellevolumen, beregnet ud fra opmålinger af alle enkelte individer (artsniveau). Det var ikke alle prøver, hvor der var noteret værdier på individniveau, så her blev der anvendt gennemsnitsværdier på artniveau.

Trin 2: Univariat kontrol på arts- og individniveau

Univariat kontrol³ på artsniveau blev udført for alle registrerede arter i data-sættet inden for de prioriterede stationer (bilag B1.1). Ved univariat kvalitetskontrol blev der ledt efter enkelte ekstreme værdier, der afveg statistisk signifikant fra hovedfordelingen af den undersøgte variabel. For hver art (Art-Kode) blev der kontrolleret for signifikante univariate afvigelser i de følgende variabler:

- Celleantal pr. L: Inden afvigerkontrol blev der filtreret for registreringer af enkeltceller. Ved andre optalte morfologier er det ikke muligt at sammenligne Celleantal pr. L, da agglomerater kan sammensættes af meget forskellige antal af enkeltceller. Andre morfologier end enkeltceller er dermed ikke blevet kvalitetssikret for denne variabel.
- Biovolumen pr. celle (individniveau): Inden afvigerkontrol blev der filtreret for registreringer af enkeltceller, dvs. der undersøgte på niveau af enkelte individopmålinger, hvor disse er registreret. Ved andre morfologier end enkeltceller, så som kolonier og kæder (agglomerater), er det ikke muligt at sammenligne biovolumen pr. celle, da agglomerater kan sammensættes af meget forskellige antal af enkeltceller, og derfor varierer meget. Andre morfologier end enkeltceller er dermed ikke blevet kvalitetssikret for denne variabel.
- Kulstofindhold pr. celle (individniveau): Inden afvigerkontrol blev der filtreret for registreringer af enkeltceller (se ovenfor).
- Celledimensioner: Inden afvigerkontrol blev der filtreret for registreringer af enkeltceller (individniveau - se ovenfor). Dimensionerne er derefter opdelt op i delmængder for alle forekommende dimensionsparametre.
- Biovolumen pr. L (genberegnet på artsniveau).
- Kulstofbiomasse pr. L (genberegnet på artsniveau).

³ Univariat kontrol er kontrol af en enkelt variabel

Til afvigelseskontrol blev der anvendt en 'generalized extreme studentized deviate (GESD) many-outlier' procedure (ISO 16269-4:2010(E)). Denne metode blev valgt, da den er velegnet til at håndtere større mængder af afvigelser og kan automatiseres. Det er dog vigtigt at bemærke, at GESD-testen kræver en nogenlunde normalfordeling af data. Det er ikke muligt at undersøge og transformere manuelt alle fordelinger for alle arter og variable. Derfor blev der i første omgang udført en Box-Cox transformation (ISO 16269-4:2010(E)) for at normalisere hvert datasæt.

Det maksimale antal af tilladte afvigelser i hvert datasæt blev defineret på forhånd af GESD-testen og blev sat til højst 5 % af alle anvendte datapunkter. Signifikansniveauet blev fastsat til $p = 0,05$. Derudover blev GESD-testen kun udført, når der var mere end 30 observationer til rådighed, da færre end 30 observationer kun har minimal betydning for antallet af arter og den samlede biomasse, men kan have en stor variation.

Trin 3: Uni- og bivariat kontrol af afvigelser på stationsniveau

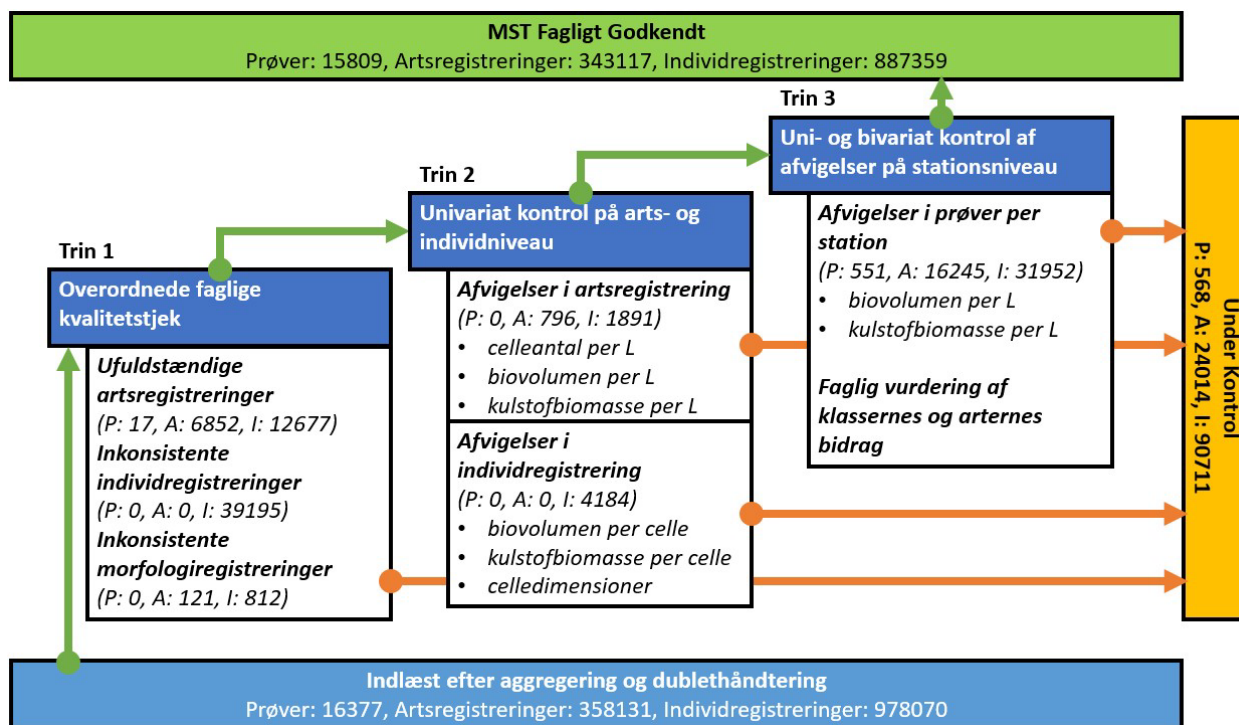
På stationsniveau blev der foretaget afvigelseskontrol efter beregning af den samlede biovolumen pr. L og kulstofbiomasse pr. L for hver prøve. Kun artsregistreringer, der ikke tidligere var blevet klassificeret som ukendte (UK), blev inkluderet i denne kontrol. Der blev anvendt GESD-proceduren (i henhold til ISO 16269-4:2010(E), som beskrevet tidligere) til univariate analyser af begge variable på stationsniveau.

Derudover blev der foretaget bivariate afvigelseskontroller ved at sammenligne den samlede biovolumen pr. L og kulstofbiomasse pr. L for alle prøver fra en enkelt station. Ved bivariate eller multivariate afvigelseskontroller blev hvert værdisæt fra to eller flere tilfældige variable undersøgt for afvigelser fra hovedfordelingen. Eksempler på resultaterne findes i resultatafsnittet. For eksempel vil en bivariat afvigelse blive bemærket som et punkt, der signifikant afviger fra resten af punktskyen i et x-y scatterplot.

I Dansk Standard 16269-4:2010 (ISO 16269-4:2010(E)) er der beskrevet to statistiske metoder til multivariat kontrol af afvigelser, baseret på (1) Mahalanobis distance og (2) Cooks distance.

- Mahalanobis distance beregner for hvert punkt afstanden mellem dets værdier og centrum af datasættet under hensyntagen til covariansen mellem variable. De nødvendige estimater for middelværdi og kovariansmatrix blev beregnet ved hjælp af den robuste 'minimum co-variance estimate' (MCE) metode, som er specificeret i ISO 16269-4:2010(E). Hvis den beregnede robuste Mahalanobis distance overstiger en kritisk tærskelværdi (se definitionen i ISO 16269-4:2010(E)), betragtes punktet som en potentiel afvigelse.
- Cooks distance beskriver et punkts indflydelse på en lineær regressionsmodel i det multivariate datasæt. Dette indebærer en sammenligning af regressionsmodellens parametre med og uden inklusion af det pågældende punkt. Cooks distance angiver størrelsen af afvigelsen mellem disse parametre. En almindelig tærskelværdi for at identificere potentielle afvigelser er en Cooks distance større end $4/n$, hvor n er antallet af datapunkter, der indgår i beregningen.

Vi implementerede en kombination af begge metoder og udpegede de mulige afvigende datapunkter, der lå over tærskelværdien, for begge tests. Hvis en prøve for en station blev udpeget som en mulig afvigelse, udviklede vi en rutine for at undersøge de bagvedliggende data.



Figur 1.1. Skematisk oversigt over forløbet i MST KS. P: antal prøveregistreringer, A: antal artsregistreringer, I: antal individregistreringer. De specifikke registreringer for trin 1, 2, og 3 er vedlagt som EXCEL filen " MST KS figur 1.1. XLSX"

Trin 1 i denne rutine var en sammenstilling af de dominerende fytoplanktonklasser for den pågældende station og fordelingen af deres relative andel (≥ 5 % af kulstofbiomassen eller biovolumen i prøven) med de relative andele i den udpegede prøve. Derefter blev klasserne undersøgt i trin 2 for at identificere dominerende arter. Slutteligt fulgte undersøgelser i trin 3 af de enkelte artsregistreringer for at sammenligne celleantal pr. L, biovolumen pr. L, kulstofbiomasse pr. L, biovolumen pr. celle og dimensionsparametrene med fordelingerne for den pågældende art på tværs af hele datasættet. Denne tilgang skabte et fagligt grundlag for at træffe beslutninger om, hvorvidt den undersøgte artsregistrering i prøven skulle forkastes, klassificeres som ukendt (UK) eller accepteres. Disse beslutninger blev markeret i ODA med kvalitetssikringsflag.

1.3 FDC kvalitetssikring

FDC KS omfattede data, der havde gennemgået MST KS. Processen involverede en undersøgelse af biomassens måneds- og årsgennemsnit. Stationsplaceringerne blev vist på figur 1.2 med angivelse af indsamlingstidspunkt og omfang i tabel 1.2. Data blev opdelt i tre sæt. Et sæt repræsenterede data for fjorde og kystnære områder, et datasæt dækkede indre åbne farvande, og et datasæt dækkede stationer i Nordsøen (ydre farvande, se figur 1.2). Der blev ikke beregnet månes or årsgennemsnit for Nordsøen da datasætte er for begrænset til at denne type analyse ikke giver faglig mening. Derudover blev der udtrukket stationer til individuel vurdering, som beskrevet af Fossing m.fl. (2018).

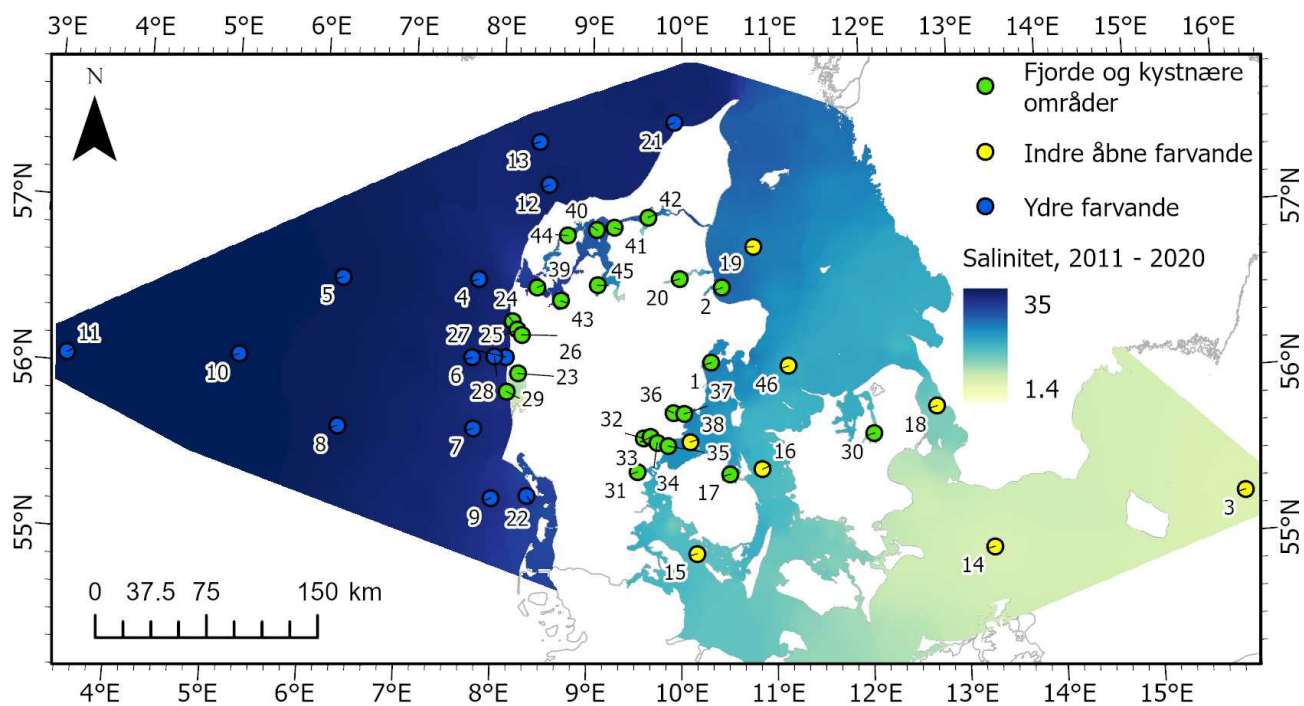
Data for gennemsnitsbiomasse varierede med sæson og år og havde en fordeling, som ikke umiddelbart kunne beskrives som normalfordelt. Dette betød, at aritmetiske middelværdier overvurderede årsgennemsnittet.

Tabel 1.2. De 46 prioriterede stationer med angivelse af kildestationsnavn (station), antal år, der er indsamlet data, samt perioden med første og sidste år der findes data for.

Nummer på kort	Station	Antal år	Periode (første - sidste år)
1	ARH170006	35	1978-2021
2	ARH190004	20	1989-2009
3	BRKBMPK2	23	1999-2021
4	DMU1023	3	2019-2021
5	DMU1026	4	2018-2021
6	DMU1035	3	2018-2021
7	DMU1042	3	2018-2021
8	DMU1046	2	2020-2021
9	DMU1052	3	2018-2021
10	DMU1075	4	2018-2021
11	DMU1077	4	2018-2021
12	DMU1130	4	2018-2021
13	DMU1133	4	2018-2021
14	DMU444	6	2016-2021
15	FYN6300043	34	1985-2021
16	FYN6700053	9	1989-2011
17	FYN6900017	23	1998-2021
18	KBH431	22	1998-2021
19	NOR409	24	1998-2021
20	NOR5503	30	1991-2021
21	NOR7715	8	2005-2021
22	RIB1510007	32	1990-2021
23	RKB1	34	1986-2021
24	RKB21	12	1985-1998
25	RKB22	12	1998-2010
26	RKB23	18	1985-2004
27	RKB41	9	1988-1997
28	RKB42	1	1988-1988
29	RKB43	11	1989-2000
30	ROS60	33	1987-2021
31	VEJ0003350	1	2004-2004
32	VEJ0003772	2	1988-1989
33	VEJ0004273	21	1988-2009
34	VEJ0004669	2	1988-1989
35	VEJ0005367	2	1988-1989
36	VEJ0005790	21	1989-2010
37	VEJ0006489	1	1992-1992
38	VEJ0006870	25	1992-2021
39	VIB3702	25	1983-2009
40	VIB3708	38	1983-2021
41	VIB3709	3	1997-2009
42	VIB3711	13	1983-1997
43	VIB3720	14	1983-2014
44	VIB3723	12	1983-1997
45	VIB3727	38	1983-2021
46	VSJ20925	23	1998-2021

Et kompromis var at beregne årsgennemsnittet som medianværdier ved hjælp af en generel lineær model, som beskrevet i bilag 2 i Hansen m.fl. (2020). De

endelige data svarede til data, der blev rapporteret i den årlige NOVANA rapport.



Figur 1.2. Placeringen af de 46 prioriterede stationer. Nummereringen referer til tabel 1.1.

2 Resultat

Biovolumen, plasmavolumen og gennemsnitsværdier

Gennemsnitsværdier (standardværdier) for biovolumen og plasmavolumen blev beregnet for arter der er opmålt for alle specifikke arter. Hvor der mangede enten biovolumen eller plasmavolumen, blev en standardværdi tilføjet. Standardværdierne blev beregnet som middelværdien af logtransformerede volumener for alle artsobservationer, der efter beregning blev tilbage transformerede. Denne algoritme blev anvendt da cellestørrelser følger en lognormal fordeling (Jakobsen m.fl., 2015). De beregnede værdier er vedlagt i EXCEL filen "MST KS figur 1.1. XLSX

Tabel 2.1. Omfanget af opdateringer af kulfaktorer.

KulFaktor (gammel)	KulFaktor(ny)	Antal opdateret
Mangler	0.00198	279
0	0.00198	12
0.0017	0.00198	701
0.002	0.00198	31
0.003	0.00198	71
0.11	0.00198	203
Mangler	0.11	63812
0	0.11	4589
0.0017	0.11	4
0.002	0.11	7
0.007	0.11	1
0.12	0.11	1095
0.13	0.11	586
1	0.11	1
Mangler	0.13	15563
0	0.13	1759
0.11	0.13	22468
Mangler	0.2	2769
0	0.2	673
0.11	0.2	55755
0.12	0.2	17833
Ikke opdaterede/eksisterende kulfaktorer er korrekt		
0.11	0.11	405373
0.13	0.13	110116

I mange tilfælde blev der angivet identiske værdier for plasmavolumen og cellevolumen for kiselalger. Det var ikke muligt at afgøre, hvilken af de to værdier der var korrekt, alternativt blev kun cellevolumen angivet. I den aktuelle analyse blev de angivne værdier for plasmavolumen derfor anvendt, selvom de var identiske med cellevolumen, i beregningen af kulstofbiomassen for kiselalger. Hvis der blev angivet en cellevolumen, men hverken plasmavolumen eller kulstofbiomasse for kiselalger på artsniveau, blev cellevolumen brugt som erstatning for plasmavolumen i beregningen af kulstofbiomassen. Dette valg blev truffet efter flere iterative kørsler i FDC KS, der viste urealistisk høj kulstofbiomasse, især for kiselalger, og at standardplasmavolumener i nogle tilfælde oversteg den registrerede cellevolumen. Det var ikke en overraskende udfordring, da cellevolumen og dermed plasmavolumener for kiselalger er karakteriseret ved meget variable celle størrelser, som følge af deres livscyklus

(Harrison m.fl., 2015, Jakobsen m.fl., 2015). Den prioriterede algoritme, der blev anvendt til at estimere volumener, blev benyttet i det videre arbejde med kvalitetssikring, herunder genberegningen af kulstofbiomasser på arts- og prøve-niveau. Det er derfor forbundet med en vis usikkerhed at substituere manglende plasmavolumener for kiselalger med standardværdier.

Biovolumen til kulstofbiomasse

Omregningen fra cellevolumen til kulstofbiomasse skete ved at korrigere med en kulstoffaktor, der beregnede kulstofbiomassen af en enkelt celle, og dette blev derefter multipliceret cellevolumen med en konstant "kulfaktor". Der fandtes en række konstante "kulfaktorer," der blev brugt i konverteringen af biovolu-men til kulstofbiomasse pr. celle. Som udgangspunkt blev værdierne fra TA09 anvendt (Fossing & Jakobsen 2015, Jakobsen & Fossing 2017). Der opstod imid-lertid afvigelser fra de faste kulfaktorer i datasættet, og disse blev opdateret (ændret) til de faste og korrekte konstanter. Omfanget af opdateringer af kul-faktorerne kan findes i tabel 2.1.

Det skal bemærkes, at i den nyere litteratur, herunder i HELCOM PEG-grup-pen, anbefales anvendelse af en eksponentiel funktion efter Menden-Deuer og Lessard (2000). Ved opdatering af alle manglende biovolumener pr. celle er det i fremtiden muligt at implementere denne metode. Dette har også den fordel, at der korrigeres for manglende plasmavolumen i beregningen af kiselalgebio-massen på artsniveau i de tilfælde, hvor kun cellevolumen er angivet.

I den nyeste HELCOM COMBINE-manual er der angivet to modeller til beregning af kulstof. En model anvendes for flagellater, og en model anvendes for kiselalger. Modellen, der foreslås i HELCOM COMBINE-manualen til beregning af volumen:kulstof af kiselalger, er hentet fra et studie af Menden-Deuer og Lessard (2000). Modellen i HELCOM COMBINE-manualen er ifølge den op-rindelige reference gældende for celler $>3000 \mu\text{m}^3$. Når HELCOM COMBINE-manualen forslår at anvende beregningen for celler $<3000 \mu\text{m}^3$ påfører en ukendt usikkerhed for mindre celler. Derudover er kildematerialet i for kiselalger $> 3000 \mu\text{m}^3$ ganske begrænset og omfatter få observationer af større kiselalger. Der er ingen perfekt løsning på problemstillingen, eller noget rigtigt eller forkert svar, og derfor anbefalede DCE, at man fastholdt den nuværende beregningsmetode, som man havde anvendt siden begyndelsen af halvfemserne. Det skal alligevel noteres at det samlede datasæt i fremtiden kan implementeres til Menden-Deuer og Lessard (2000) metoden. Dette har også den fordel, at der korrigeres for manglende plasmavolumen i beregningen af kiselalgebiomassen på artsniveau i de tilfælde, hvor kun cellevolumen er angivet. Det skal noteres, at DCE har opdateret kulfaktoren for ciliater. Der havde tidligere været anvendt forskellige værdier i NOVANA-programmet, og DCE havde derfor valgt at anvende 0,2 (Putt & Stoecker 1989), som typisk blev anvendt i den peer-reviewede litteratur. For furealgen *Noctiluca* spp. havde man anvendt $1,98 \times 10^{-3}$ og $C \mu\text{m}^3$ celle efter Tada m.fl. (2000).

2.1 MST kvalitetssikring

Overordnede kvalitetstjek

For de 46 prioriterede stationer blev der i alt indlæst data fra 16.377 prøver, som indeholdt i alt 358.131 artsregistreringer og 978.070 individregistreringer (tabel 2.2). Efter de overordnede kvalitetstjek blev det resterende datasæt reduceret med i alt 17 prøver, 6.973 artsregistreringer og 52.684 individregistreringer, da

nogle af registreringerne blev markeret som ukorrekte (UK). Metoden for de forskellige overordnede kvalitetstjek er beskrevet i de følgende afsnit.

Ufuldstændige artsregistreringer

Efter genberegningen af data på artsniveau blev der fundet 6.852 ufuldstændige artsregistreringer, hvilket udgjorde ca. 1,9 % af alle indlæste artsregistreringer.

For disse 6.852 artsregistreringer (trin 1, figur 1.1) manglede der kulstofbiomasse pr. L. Desuden var der heller ikke angivet plasmavolumen pr. individ for disse registreringer. Artsregistreringerne tilhørte følgende klasser af kiselalger: Bacillariophyceae, Bacillariophyceae/Fragilariophyceae, Fragilariophyceae, Mediophyceae, Coscinodiscophyceae. Derudover var der registreringer for klasserne Chlorophyceae, Chrysophyceae, Cyanophyceae, Dictyochophyceae, Dinophyceae, Oligohymenophorea, Oligotrichea, Pyramimonadophyceae, Zoomastigophora.

Disse artsregistreringer blev markeret som ukorrekte (UK), men de blev ikke inddraget i den videre kvalitetssikringsproces, da det ville forårsage ubalance mellem kulstofbiomasse og biovolumen. Der var også 12.677 tilhørende individregistreringer, der blev markeret som under kontrol.

Tabel 2.2. Oversigt over den indlæste datamængde og antal registreringer, der er blevet sat til UK igennem de 3 faglige kvalitetssikringstrin (Figur 1.1 viser flow i processen). Trin 3 er ikke endnu afsluttet (NA), derfor angives her kun det foreløbige antal prøver, der blev identificeret som mulige afvigelser. Størstedelen af disse prøver vil dog sandsynligvis accepteres, efter individuel gennemgang af prøverne.

	Indlæst	Trin 1	Trin 2	Trin 3
Prøver	16.377	17	NA	(551)
Artsreg.	358.131	6.973	796	NA
Individreg.	978.070	52.684(39.195+12.677+812)	4.184	NA

Inkonsistente individregistreringer

Der blev fundet 39.195 inkonsistente individregistreringer, hvor kun 56 af dem skyldtes, at biovolumen pr. celle var ≤ 0 . For de resterende 39.139 individregistreringer manglede der biovolumen pr. celle, selvom dimensioner var blevet opmålt. Vi vurderede, at der i de fleste tilfælde var uoverensstemmelse mellem de angivne dimensionsparametre og de krævede parametre i formelen til biovolumenberegning.

På artsniveau blev inkonsistente individregistreringer erstattet med standardvolumener og kulstofbiomasser pr. celle. Her er et eksempel: I en prøve fra station ROS60 dateret den 12.02.2002 var der en artsregistrering (Specie_ID: 32240) for Dinophyceae, hvor der blev registreret to dimensionsparametre: 'd' = 15 μm og 'h' = 20 μm . Den tilknyttede formel i databasen ($\pi/4dl-2l$) kræver dog dimensionsparametrene 'dl' og 'l' i stedet. Derfor kunne der ikke beregnes biovolumen pr. celle for denne registrering. Det er indlysende, at 'd' svarer til 'dl' og 'h' svarer til 'l', men det var kun muligt at gennemføre en delvis automatiseret rettelse, da de enkelte rettelser skulle kontrolleres manuelt. Det var ikke muligt at gennemgå alle disse inden for projektets rammer.

De tilhørende artsregistreringer blev ikke sorteret fra i den videregående kvalitetssikringsproces, da biovolumen og kulstofbiomasse på artsniveau blev

genberegnet for alle arter som standardværdier. Individregistreringer blev dog markeret som ukorrekte (UK), fordi mange af de fejlagtige dimensionsparametre muligvis kunne rettes på et senere tidspunkt.

Inkonsistente morfologiregistreringer

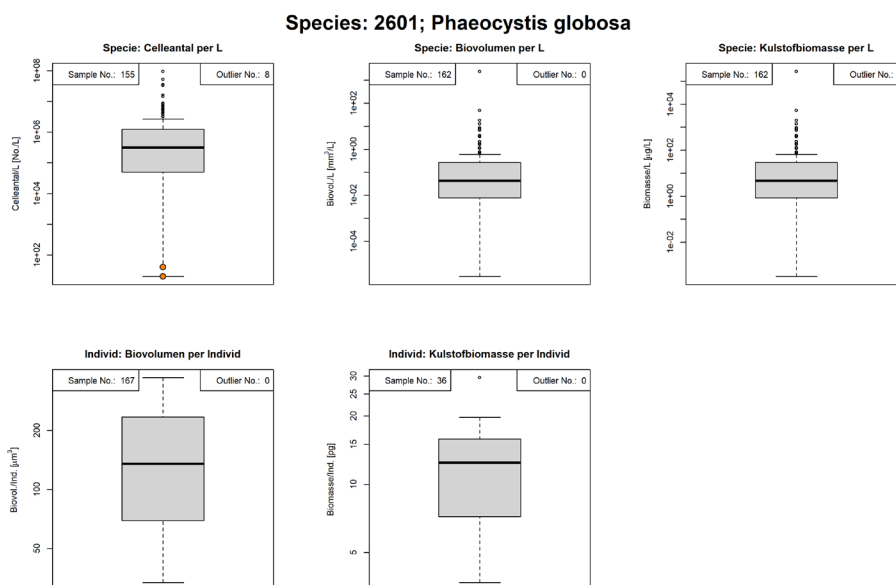
Der blev fundet 121 artsregistreringer med inkonsistente morfologiangivelser i databasen. Disse artsregistreringer blev markeret som UK, fordi mange af dem muligvis kan harmoniseres på et senere tidspunkt. Der var 812 tilhørende individregistreringer, som i samme omgang blev markeret som UK.

- Eksempel: I en prøve fra station KBH431 d. 11.09.2012 er der en artsregistrering (Specie_ID: 264645) for *Calothrix*, hvor det er angivet, at arten foreligger som 'tråd' og er blevet talt som 'cyste', men målt som 'tråd'.

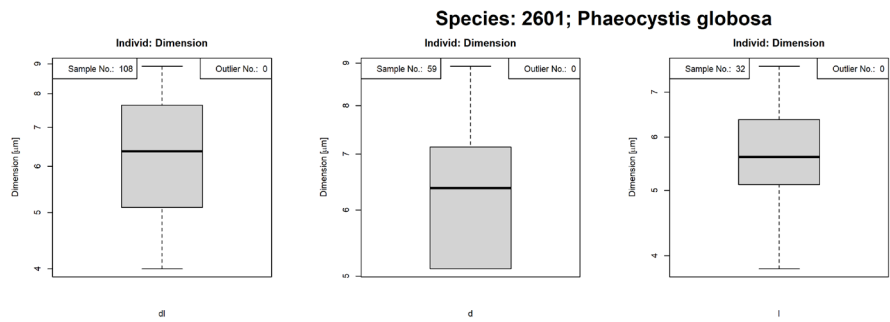
Univariat kontrol på artsniveau

Samlet set blev der gennemført univariat kvalitetskontrol for 722 arter (defineret af den tilknyttede ArtKode). Dette inkluderede alle arter, der blev fundet på de 46 prioriterede stationer. Som beskrevet tidligere blev variablerne undersøgt både på arts- og individniveau. For hver art blev der genereret et sæt grafer (se eksempler i figur 2.1 og figur 2.2 for *Phaeocystis globosa*), der viste afvigelser i forhold til variabelens generelle fordeling på tværs af hele datasættet. For eksempel blev der identificeret 8 artsregistreringer for *Phaeocystis globosa* med et meget lavt antal celler pr. L som potentielle afvigelser. Det er værd at bemærke, at antallet af registreringer, der indgår i beregningen, kunne variere for hver variabel. Det gav f.eks. kun mening at sammenligne celleantal pr. L for enkeltceller, mens morfologien for biovolumen pr. L og kulstofbiomasse pr. L ikke havde nogen betydning.

Figur 2.1. Eksempel på den univariate kvalitetskontrol for *Phaeocystis globosa* og de fem variabler til bestemmelse af kulstofbiomasse og biovolumen. Orange punkter markerer potentielle afvigelser.



Figur 2.2. Eksempel af den univariate kvalitetskontrol for *Phaeocystis globosa* og dimensioner. Ingen afvigelser detekteret.

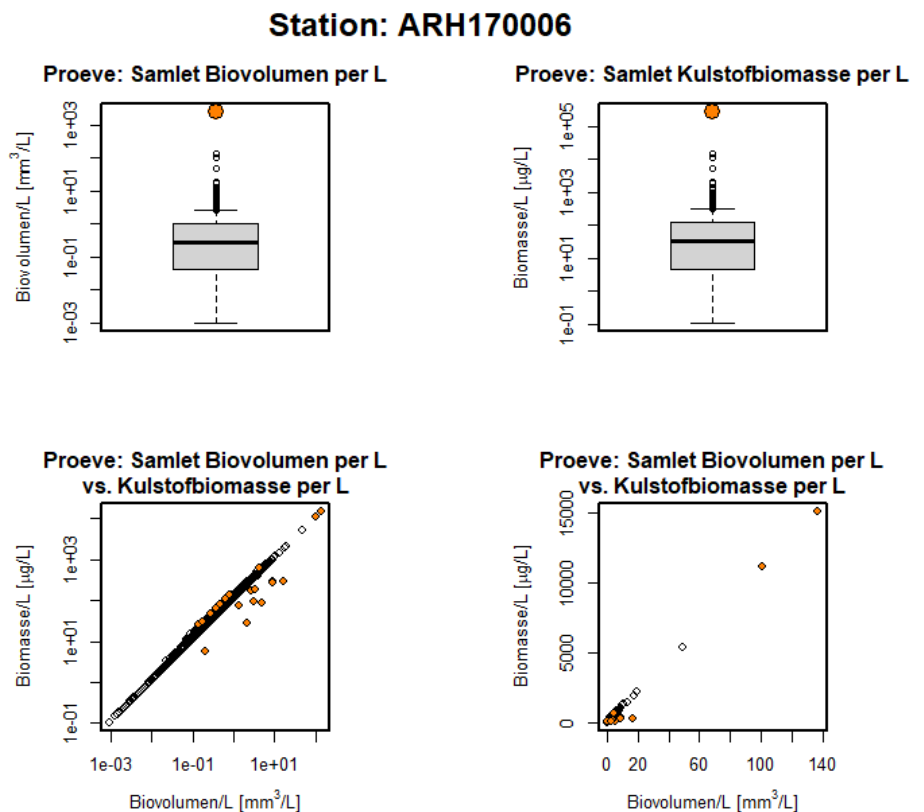


I alt blev der fundet potentielle afvigelser for 796 artsregistreringer og 4.184 individregistreringer (tabel 2.2.)

Trin 3: Uni- og bivariat kontrol af på prøve- og stationsniveau

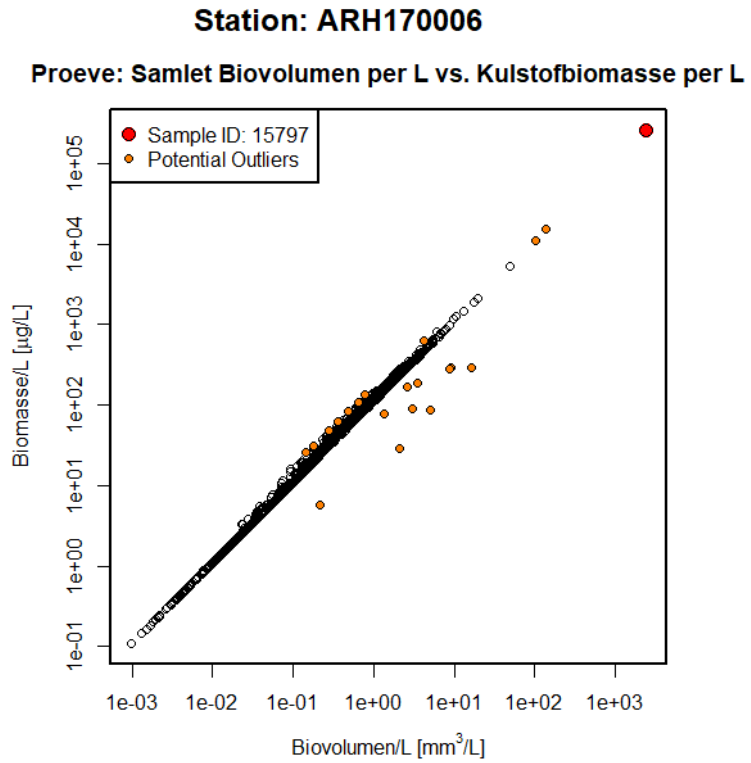
I trin tre undersøgte vi for afvigelser baseret på den samlede biovolumen pr. L og kulstofbiomasse pr. L på prøve x stationsniveau (se eksempel for station ARH170006 i Figur 5). Disse afvigelser kunne gennemgås enkeltvis (se eksempel i Figur 2.4-2.12) for at undersøge de underliggende registreringer, der lå bag en afvigende prøve. I alt blev der udpeget 553 prøver med potentielle afvigelser, som derefter skulle analyseres nærmere. De følgende to eksempler viser fremgangsmåden i R-værktøjet, som leverede grundlaget for at træffe beslutninger om prøvernes videre håndtering⁴. Vi vurderede, at de fleste afvigende prøver, ligesom de to følgende eksempler, kunne forklares som atypiske biologiske afvigelser og derfor ikke skyldtes fejl i data.

Figur 2.3. Eksempel på uni- og bivariat kvalitetskontrol for alle prøver fra station ARH170006. For denne station er udpeget i alt 21 prøver som potentielle afvigelser, 1 ekstrem værdi for biovolumen pr. L og kulstofbiomasse pr. L, og 20 værdier, der faldt udenfor den bivariante analyse ved sammenstilling af biovolumen pr. L og kulstofbiomasse pr. L. Bemærk ekstremværdier udpeget i det første univariate trin indgår ikke i den bivariante analyse. Bivariate plots vises både på log- og normal-skala. Orange punkter markerer potentielle afvigelser.



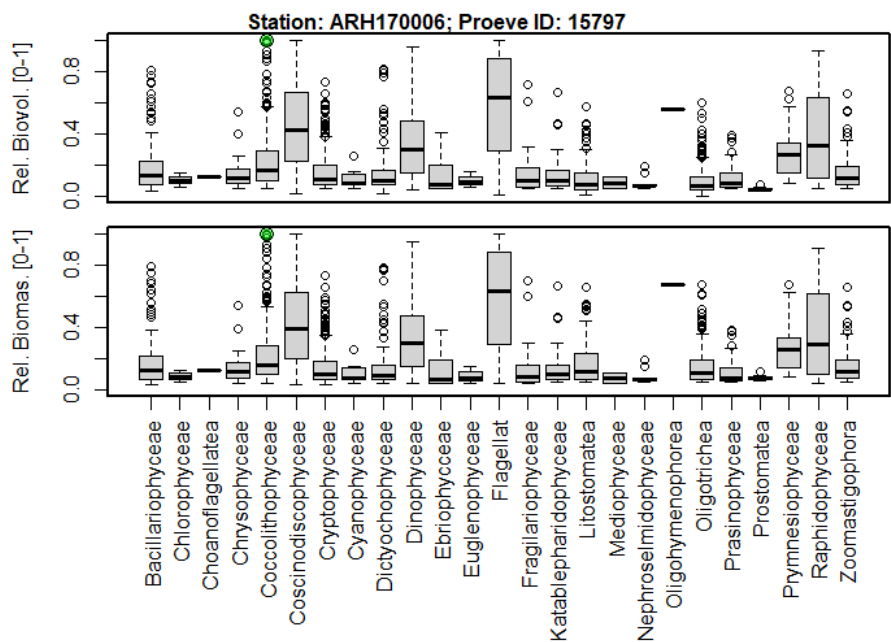
⁴ Begge prøver godkendes efterfølgende, så der udpeges i alt 551 prøver

Figur 2.4. Eksempel på en specifik afvigelse, der bliver undersøgt nærmere i de følgende figurer. Afgigelsen tilhører prøven med ID 15797 fra d. 06.04.1995. Orange punkter markerer potentielle afvigelser.



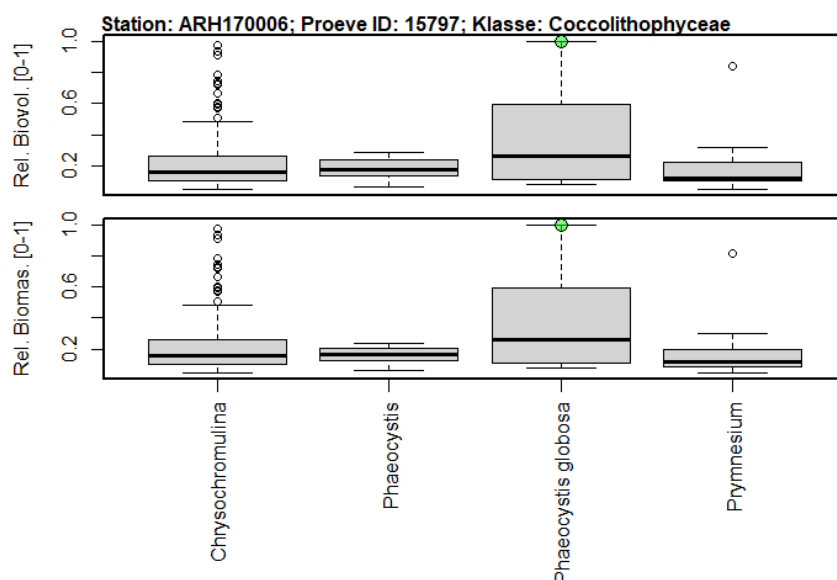
Efter der blev udpeget en potentiel afvigelse, blev der foretaget en sammenstilling af fytoplanktonklasser, der udgjorde mindst 5 % af kulstofbiomassen i en af prøverne fra den pågældende station. For hver fytoplanktonklasse blev der vist fordelingen af deres relative andel i samtlige prøver fra stationen (boxplots), mens værdierne for den undersøgte prøve blev lagt ovenpå (grønne punkter). Et eksempel for prøven med ID 15797 ses i figur 2.5, hvor der kun blev registreret en enkelt fytoplanktonklasse (Coccolithophyceae) med betydelig kulstofbiomasse, som udgjorde 99,9 % af både den samlede biovolumen pr. L og kulstofbiomasse pr. L i denne prøve. Dette var også de højeste værdier registreret for denne klasse på station ARH170006.

Figur 2.5. Eksempel på fordelingen af relative andele for alle fytoplanktonklasser, der udgjorde mindst 5 % af den samlede biovolumen eller kulstofbiomasse i en af prøverne fra station ARH170006 (boxplots). Grønne punkter viser værdier for den undersøgte prøve med ID 15.797 fra d. 6. april 1995. Der er kun en enkelt klasse (*Coccolithophyceae*) med betydelig kulstofbiomasse registreret i denne prøve. Grønne punkter markerer potentielle afvigelser, der identificeres i biovolumen og i biomassen.



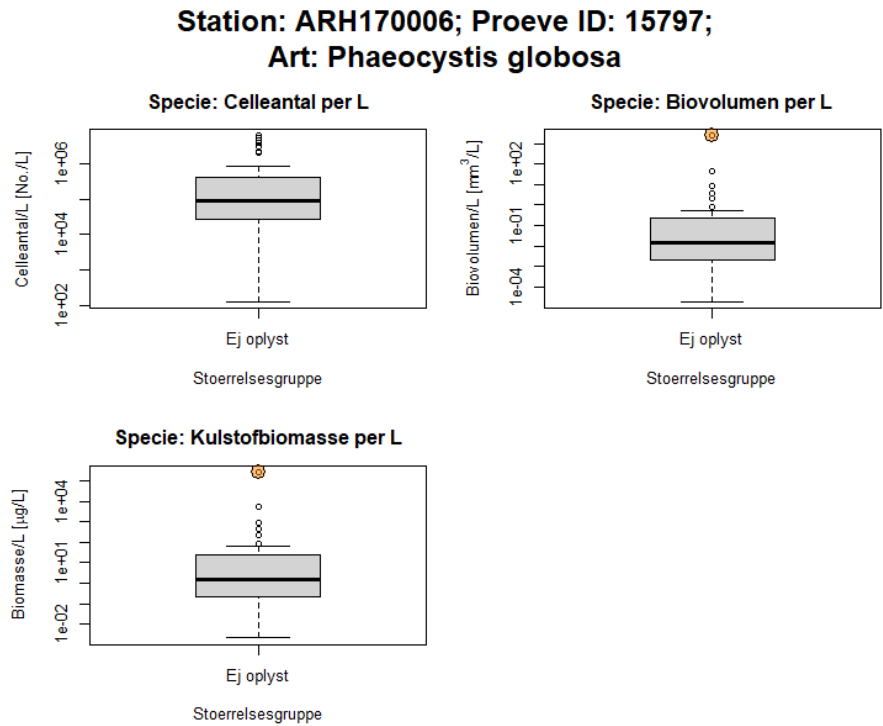
Derefter blev det muligt at undersøge hver fytoplanktonklasse med registreringer i den pågældende prøve. I figur 2.5 vises et eksempel med alle arter i klassen Coccolithophyceae, der udgjorde mindst 5 % af den samlede kulstofbiomasse pr. L eller biovolumen pr. L i en af prøverne fra station ARH170006. I prøven med ID 15.797 blev der kun registreret en enkelt art (*Phaeocystis globosa*), som alene udgjorde 99,9 % af hele prøvens biovolumen pr. L og kulstofbiomasse pr. L. Disse værdier var også de højeste registrerede for denne art på station ARH170006.

Figur 2.6. Eksempel på fordelingen af relative andele for alle arter i fytoplanktonklasse *Coccolithophyceae*, der udgjorde mindst 5 % af den samlede biovolumen eller kulstofbiomasse i en af prøverne fra station ARH170006. Grønne punkter viser værdier for den undersøgte prøve med ID 15.797 fra d. 6. april 1995. Der er kun en enkelt art (*Phaeocystis globosa*), der har betydelig biovolumen og kulstofbiomasse registreret for denne klasse. Grønne punkter markerer potentielle afvigelser, der identificeres i biovolumen og i biomassen.



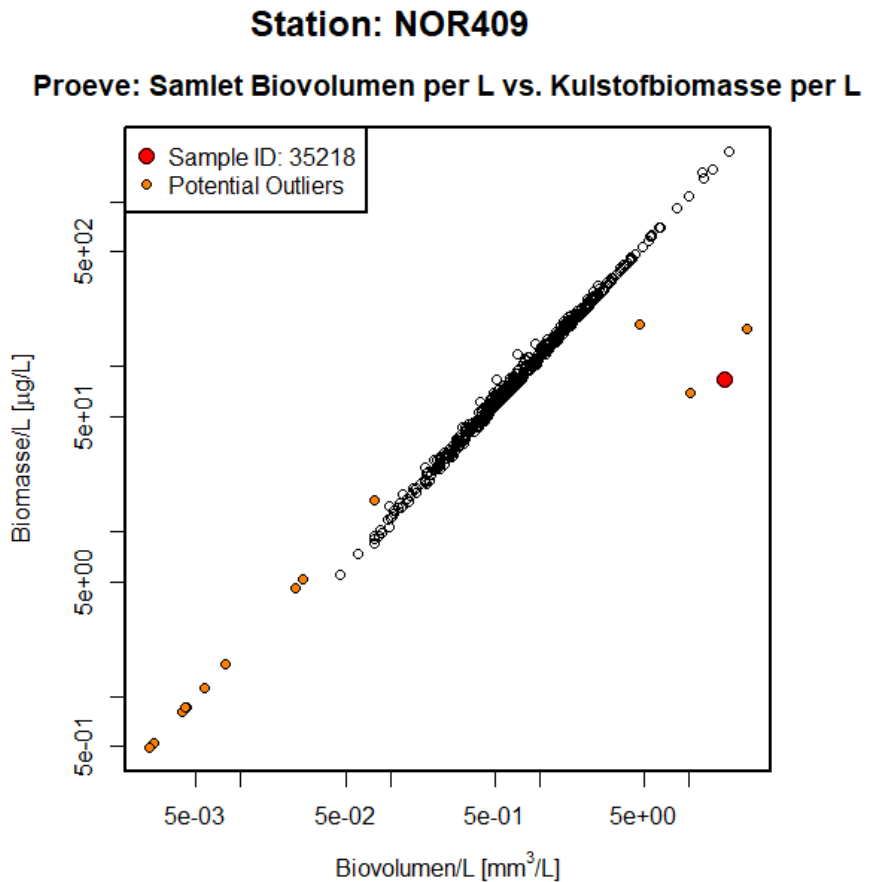
I det sidste trin i denne del af kvalitetssikringsværktøjet blev den bagvedliggende artsregistrering undersøgt, og de underliggende værdier blev sammenlignet med fordelingerne for denne art på tværs af hele datasættet. I figur 2.6 vises et eksempel for *Phaeocystis globosa*. Det bliver nu klart, at denne prøve indeholder de højeste registrerede biovolumen pr. L og kulstofbiomasse pr. L for *Phaeocystis globosa* i hele datasættet. Det må dog bemærkes, at disse meget høje værdier for biovolumen pr. L og kulstofbiomasse pr. L ikke tidligere blev identificeret som potentielle afvigelser i den univariate analyse på artsniveau (se afsnittet tidligere). Denne analyse skulle nu give mulighed for en fagligt funderet beslutning om, hvad der skal gøres med denne prøve. Prøven kunne enten 'accepteres', sættes som 'UK' eller 'afvises' som forkert. I dette tilfælde blev afvigelsen accepteret som et korrekt datapunkt, hvilket betyder, at artsregistreringen blev godkendt som fagligt korrekt.

Figur 2.7. Eksempel på fordelingen af antal celler pr. L, biovolumen pr. L og kulstofbiomasse pr. L for *Phaeocystis globosa* på tværs af hele datasættet i forhold til den undersøgte prøve 15797 fra station ARH170006 (orange). Bemærk: Dette plot kan også indeholde variabelen biovolumen (individ) pr. individ, hvis der ligger tilsvarende data i artsregistreringen for den undersøgte prøve. Ligeledes vil der være flere box-plots ved registrering af >1 størrelsesklasse. Orange punkter markerer potentielle afvigelser, der identificeres i biovolumen og i biomassen.



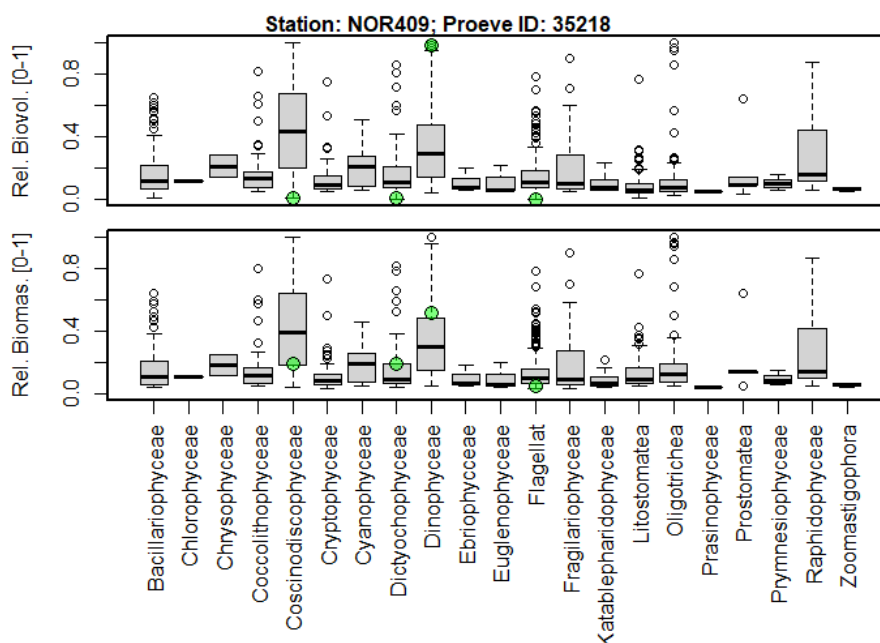
Vi har valgt et andet eksempel fra station NOR409, hvor der var tydelige afvigelser i sammenstillingen af biovolumen pr. L og kulstofbiomasse pr. L for prøverne. Figur 2.8 viser prøven med ID 35.218, som lå langt væk fra hovedfordelingen.

Figur 2.8. Eksempel på en specifik afvigelse, der bliver undersøgt nærmere i de følgende figurer 2.9-2.12. Afvigelsen tilhører prøven med ID 35.218 fra station NOR409 d. 1. december 2008. Orange punkter markerer potentielle afvigelser.



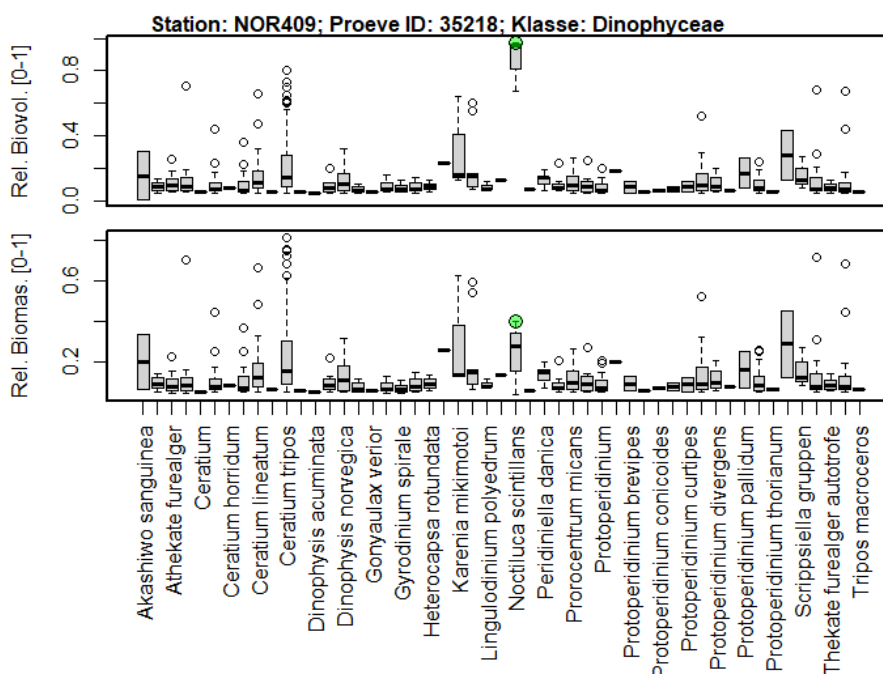
Der blev fundet 4 fytoplanktonklasser med en relativ andel på > 5 % i enten biovolumen pr. L eller kulstofbiomasse pr. L for denne prøve (figur 2.9). Klassen *Dinophyceae* bidrog med 97,9 % af den samlede biovolumen, mens den kun bidrog med 48,6 % til den samlede kulstofbiomasse.

Figur 2.9. Eksempel på fordelingen af relative andele for alle fytoplanktonklasser, der udgjorde mindst 5 % af den samlede biovolumen eller kulstofbiomasse i en af prøverne fra station NOR409. Grønne punkter viser værdier for den undersøgte prøve med ID 35.218 fra d. 1. december 2008. I denne prøve er der 4 klasser med betydelig kulstofbiomasse registreret. Bemærk: For *Dinophyceae* er der stor forskel mellem de relative andele for biovolumen pr. L (97,9 %) og kulstofbiomasse pr. L (48,6 %).



Bagved denne registrering lå hovedsageligt arten *i*, som alene bidrog med 97,5 % af biovolumen og 36,3 % af kulstofmassen til prøvens samlede værdier (figur 2.10).

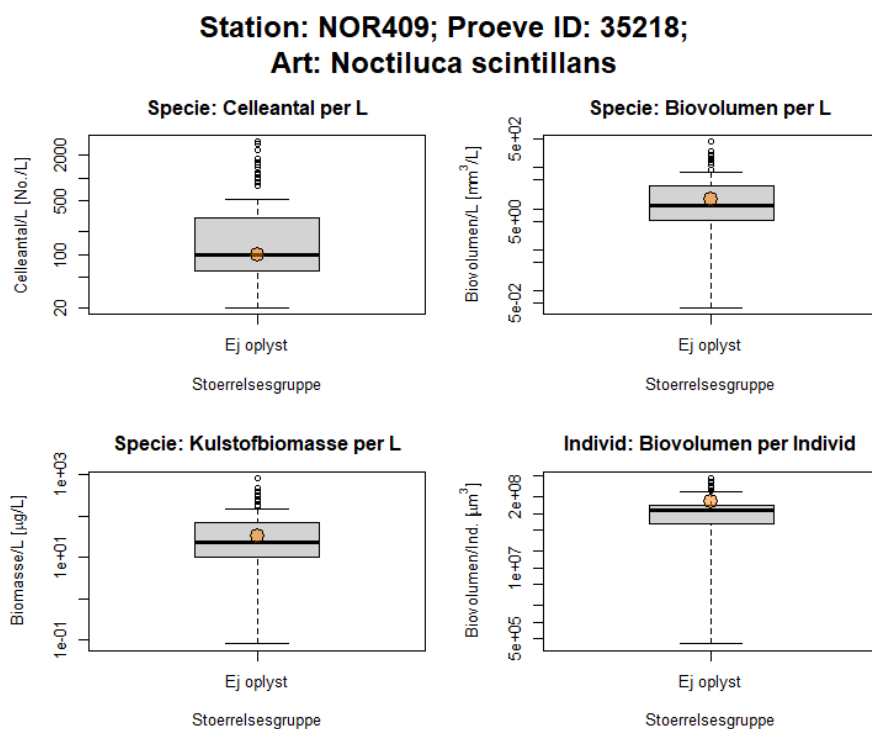
Figur 2.10. Eksempel på fordelingen af relative andele for alle arter i fytoplanktonklasse *Dinophyceae*, der udgjorde mindst 5 % af den samlede biovolumen eller kulstofbiomasse i en af prøverne fra station NOR409. Grønne punkter viser værdier for den undersøgte prøve med ID 35.218 fra d. 1. december 2008. Der er kun en enkelt art (*Noctiluca scintillans*), der har betydelig biovolumen og kulstofbiomasse registreret for denne klasse.



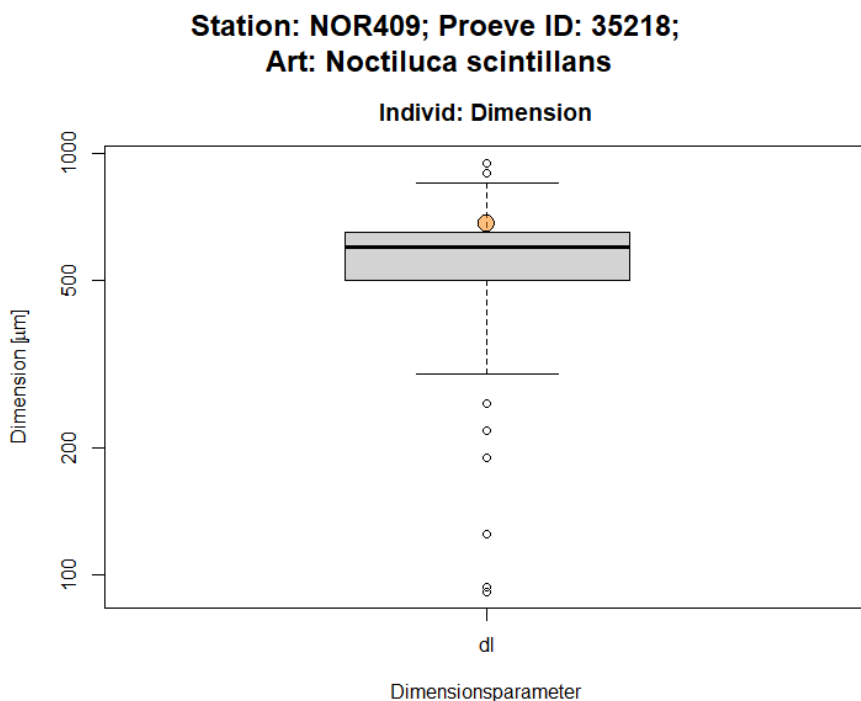
Arts- og individregistreringerne peger ikke på abnormitet i denne registrering for *Noctiluca scintillans* (figur 2.11 og 2.12). Kulfaktoren er årsagen til, at prøven blev udpeget som en afvigelse, fordi omregningen fra biovolumen til kulstofbiomasse er særlig lav for *Noctiluca scintillans*. I tilfældet, hvor netop denne art dominerer, vil den altid falde uden for fordelingen. Dette eksempel

understreger, at der er en logisk, mekanisk (kulfaktoren) forklaring på afvigelsen, og at prøven godkendes som fagligt korrekt.

Figur 2.11. Eksempel på fordelingen på artsniveau af antal pr. L, biovolumen pr. L, kulstofbiomasse pr. L og biovolumen pr. individ for *Noctiluca scintillans* på tværs af hele datasættet i forhold til værdierne for prøve 35218 fra station NOR409 (orange).



Figur 2.12. Eksempel på fordelingen af målte dimensioner for enkeltceller af *Noctiluca scintillans* på tværs af hele datasættet i forhold til værdierne fra den undersøgte prøve 35.218 fra station NOR409 (orange).



2.2 FDC kvalitetssikring

Formålet med FDC KS var at indlæse og compilere data for at vurdere, om enkelte år afveg markant fra de øvrige år. Dette beror i høj grad på et fagligt skøn, men var nødvendigt, da MST KS tidligere havde været uklar (se rapportens

introduktion). Derudover fungerede FDC KS som en metode til at vurdere den overordnede datakvalitet.

Fytoplankton

Formålet med FDC KS var at indlæse og kompilere data for at vurdere, om enkelte år afveg markant fra de øvrige år. Dette beror i høj grad på et fagligt skøn, men var nødvendigt, da MST KS tidligere havde været uklar (se rapportens introduktion). Derudover fungerede FDC KS som en metode til at vurdere den overordnede datakvalitet.

Fytoplankton blev opdelt i forskellige grupper i FDC KS : Heterotroft protistplankton blev inddelt i ciliater, furealger og andre. Der var generel usikkerhed omkring angivelserne af trofi. For nogle arter, herunder kiselalger, blev der angivet heterotrofi, men for de ældre data manglede der i mange tilfælde angivelser af trofi. Heterotrofe arter blev derfor udpeget efter følgende algoritme:

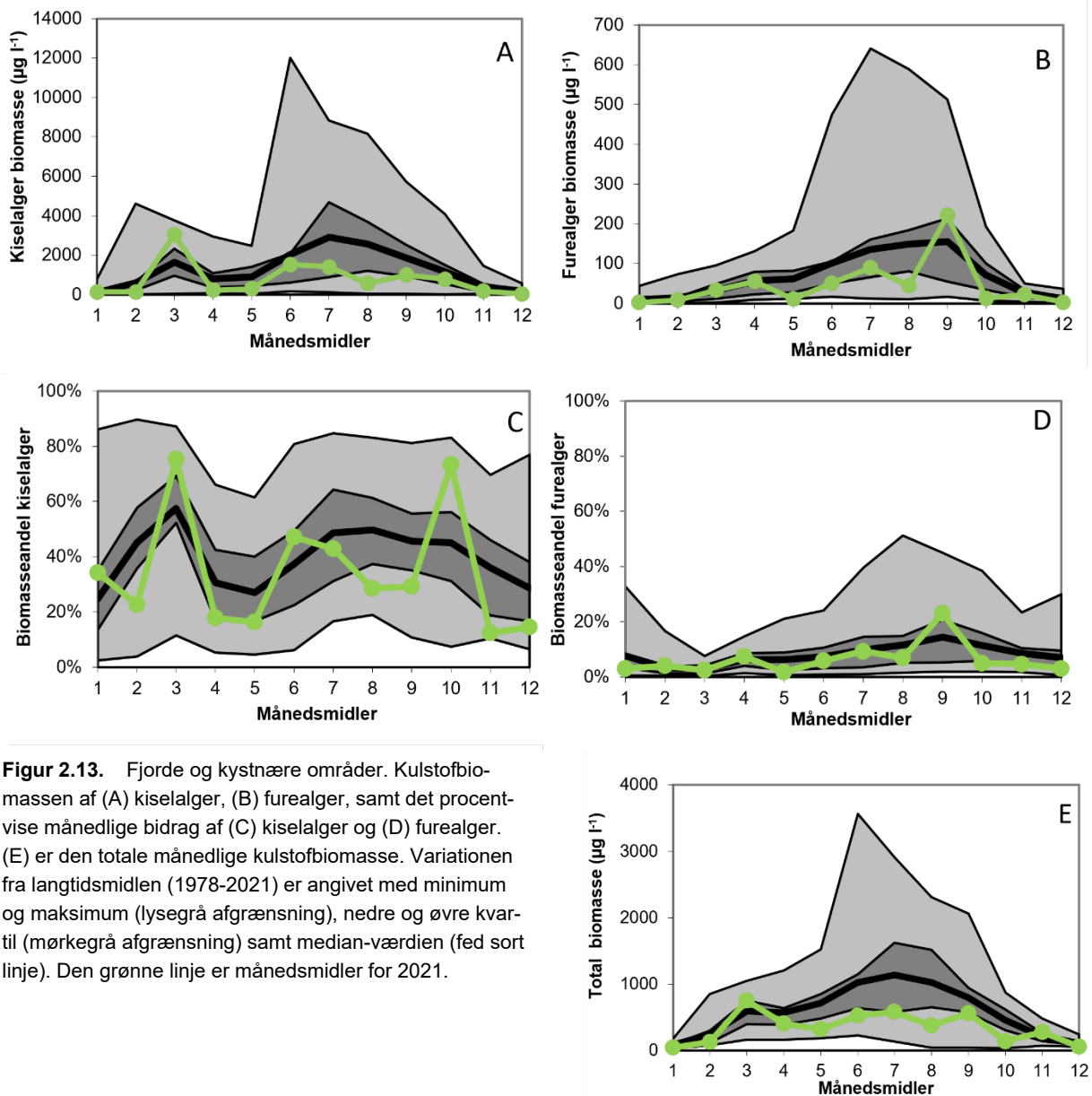
Arter tilhørende gruppen af furealger, såsom *Gymnodinium*, *Gymnodinium/gyrodinium*, *Gyrodinium*, *Dinophyceae*, athekate furealger samt thecate furealger, hvor der var påtegnet heterotrofi, blev medtaget.

Nanoflagellater blev optalt ved hjælp af epi-fluorescens mikroskop, i overensstemmelse med de tekniske anvisninger i TA09, hvor der var angivet heterotrofi.

En liste over i alt 146 heterotrofe arter blev identificeret gennem en gennemgang af alle arter med angivelser af trofi (bilag B2). Arter, der ikke kunne identificeres som heterotrofe, blev forkastet.

Ciliaten *Mesodinium rubrum* blev betragtet som fytoplankton, da denne mixotrofe art blev talt som fytoplankton, da den kun minimalt konsumerede (Kim m.fl., 2017).

Det MST fagligt kvalitetssikrede datasæt indeholdt prøver fra 1978 og frem til 2021. Data blev undersøgt på niveauet for fjorde og kystnære områder samt åbne indre farvande. I beregningen af middelværdier på områdeniveau blev kun værdier fra 1989 og frem medtaget, da dataindsamlingen tidligere var sporadisk med meget få aktive stationer, som ikke nødvendigvis kunne repræsentere hele havområder



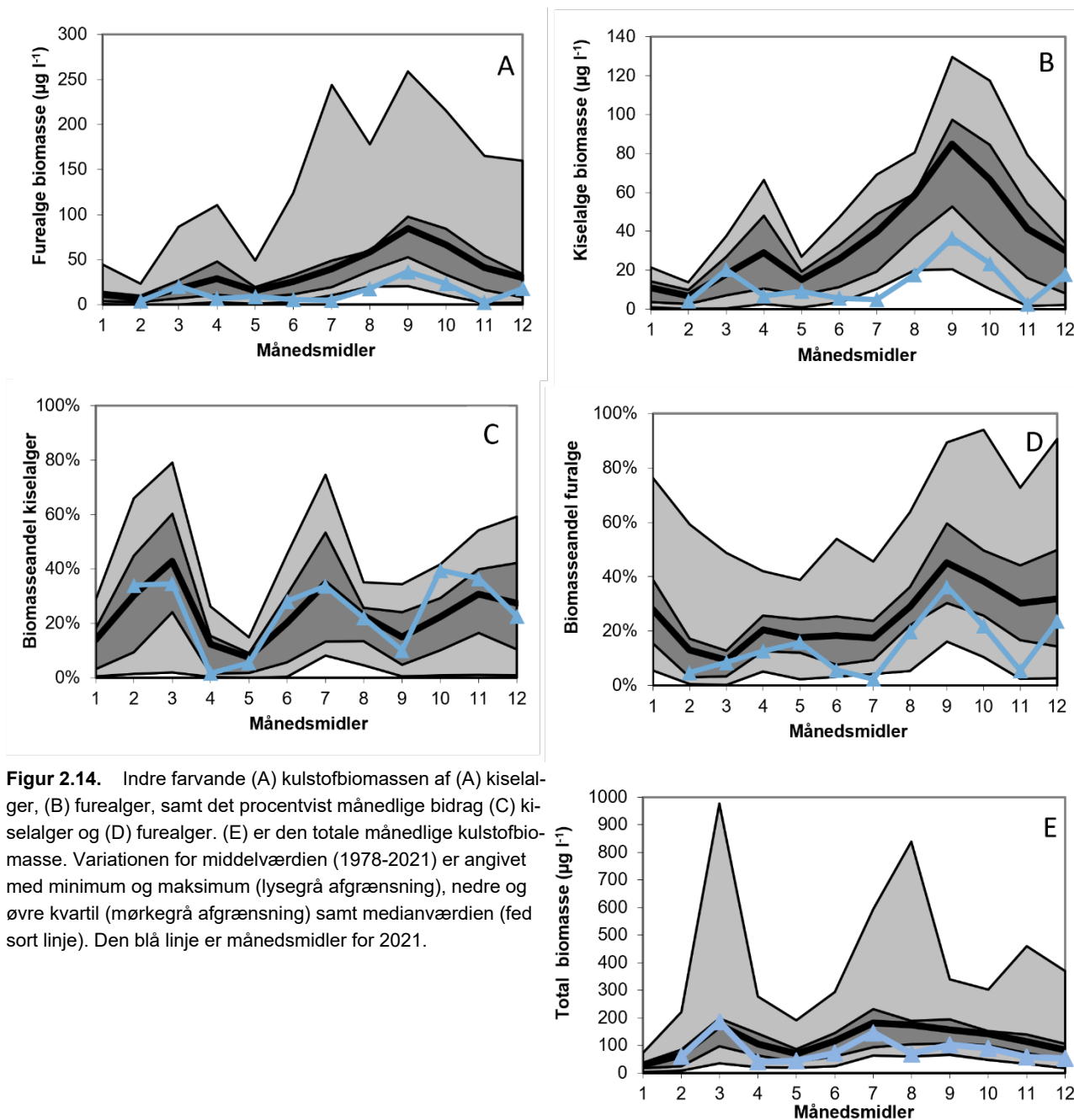
Figur 2.13. Fjorde og kystnære områder. Kulstofbiomassen af (A) kiselalger, (B) furealger, samt det procentvise månedlige bidrag af (C) kiselalger og (D) furealger. (E) er den totale månedlige kulstofbiomasse. Variationen fra langtidsmidlen (1978-2021) er angivet med minimum og maksimum (lysegrå afgrænsning), nedre og øvre kvartil (mørkegrå afgrænsning) samt median-værdien (fed sort linje). Den grønne linje er månedsmidler for 2021.

Den normaliserede fordeling af kiselalger og furealger var mellem 20 % – 60 % for kiselalger og 0,5 % – 17 % af totalbiomassen for furealger (figur 2.13C og D).

Den integrerede kulstofbiomassen for det samlede fytoplanktonsamfund for fjorde og kystnære områder havde en jævnt stigende kulstofbiomasse over året med en lille top i marts, som skyldtes forårsopblomstringen. Derefter steg kulstofbiomassen jævnt, indtil maksimal kulstofbiomasse på omkring $1200 \mu\text{g C L}^{-1}$ blev nået i juli måned, hvorefter kulstofbiomassen faldt jævnt med minimale vinterværdier i december – januar. Det skal bemærkes, at der ikke blev fundet ekstraordinære værdier, og at usikkerheden beregnet på månedsbasis for et enkelt år (grøn linje) var langt større end gennemsnitsværdierne i figur 2.13E.

Gennemsnitskulstofbiomassen (sort linje, figur 2.14A, B, E) for de indre åbne farvande viste forekomst af kiselalger hen over sæsonen, der i forløb mindede om fjorde og kystnære områder. Det betød, at opblomstringen i forårsmåned marts nåede $300 \mu\text{g C L}^{-1}$ for kiselalger, efterfulgt af en sommerperiode med lavere gennemsnitskulstofbiomasse, der blev betegnet som klarvandsfasen, efterfulgt af en stigning i gennemsnitskulstofbiomassen i løbet af efteråret til $140 \mu\text{g C L}^{-1}$ (Figur 2.14A). Furealger viste ikke tegn på forårsopblomstring i de

indre åbne farvande, men havde en jævnt stigende kulstofbiomasse indtil sensommeren i august – september, hvor gennemsnitskulstofbiomassen nåede omkring $90 \mu\text{g C L}^{-1}$. Dette var omkring 25 % mere end fjorde og kystnære områder (figur 2.13B vs. 2.14B).

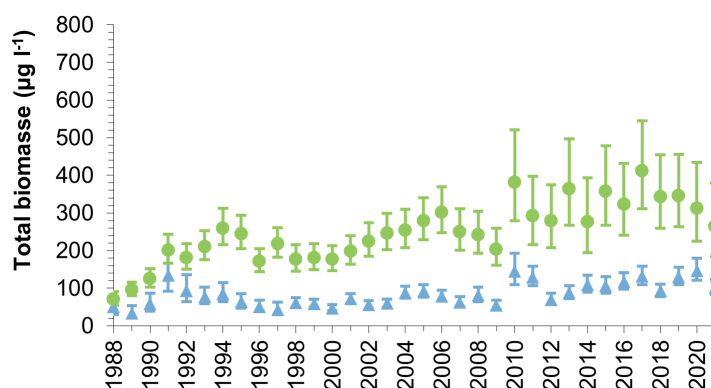


Figur 2.14. Indre farvande (A) kulstofbiomassen af (A) kiselalger, (B) furealger, samt det procentvist månedlige bidrag (C) kiselalger og (D) furealger. (E) er den totale månedlige kulstofbiomasse. Variationen for middelværdien (1978-2021) er angivet med minimum og maksimum (lysegrå afgrænsning), nedre og øvre kvartil (mørkegrå afgrænsning) samt medianværdien (fed sort linje). Den blå linje er månedsmidler for 2021.

Den normaliserede fordeling af kiselalger og furealger var mellem 5 % – 45 % for kiselalger og 8 % – 45 % for furealger af totalbiomassen, med de højeste bidrag fra furealger i sensommeren (figur 2.14C og D). Den integrerede gennemsnitskulstofbiomasse for det samlede fytoplanktonsamfund for de indre farvande viste karakteristisk forårsopblomstring med kulstofbiomasser omkring $180 \mu\text{g C L}^{-1}$ i marts måned, hvorefter gennemsnitskulstofbiomassen faldt jævnt med minimale værdier i maj ($45 \mu\text{g C L}^{-1}$), hvorefter gennemsnitskulstofbiomassen steg til omkring $180 \mu\text{g C L}^{-1}$ (figur 2.14E), primært bestående af kiselalger (figur 2.14C). Herefter faldt den samlede gennemsnitskulstofbiomasse, og

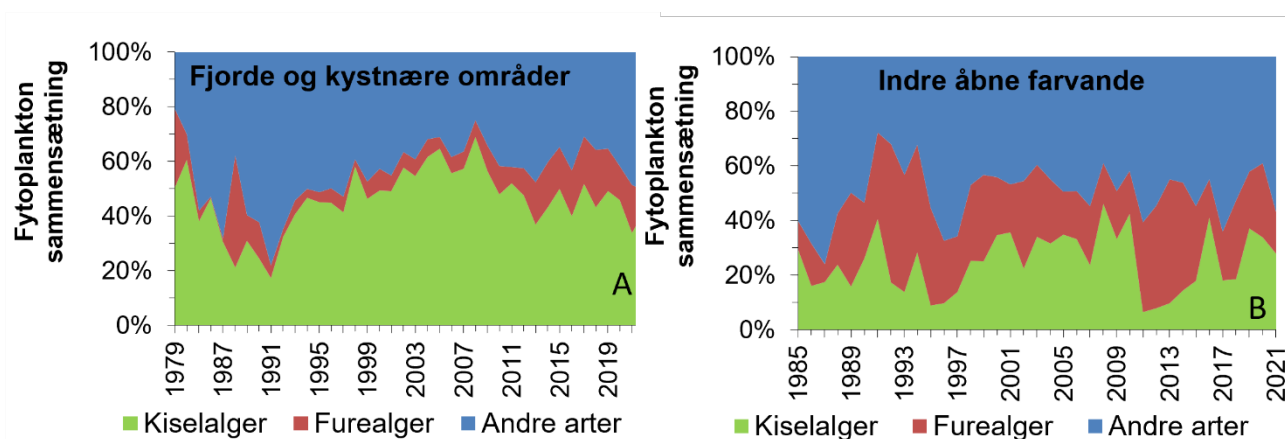
algesamfundet blev erstattet af furealger med maksimal gennemsnitskulstofbiomasse på ca. 150 $\mu\text{g C L}^{-1}$ i september måned (figur 2.14E)

Figur 2.15. Årsmidler (kulstofbiomasse) \pm 95 % konfidensgrænser fjorde og kystnære områder (●) og de indre åbne farvande (▲)



Gennemsnitskulstofbiomassen for fjorde og kystnære områder viste en generel stigning i kulstofbiomassen siden 2008 (figur 2.15). Det samme mønster var ikke helt så tydeligt for de indre åbne danske farvande (figur 2.15).

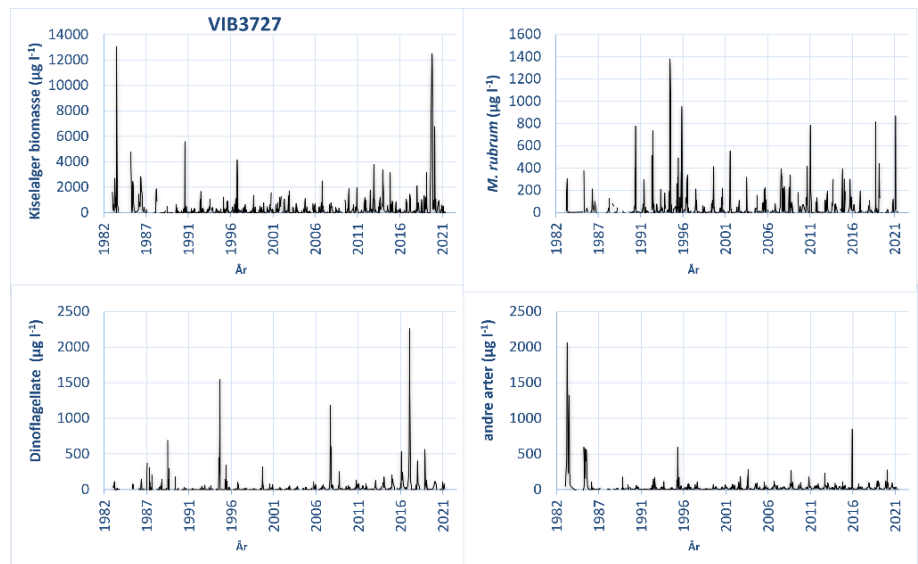
Fordelingen mellem kiselalger, furealger og andre arter blev beregnet ud fra årsmidlerne for fjorde og kystnære områder og pegede på et stigende bidrag af kiselalger i perioden fra 1991 og frem til 2008 (omkring 70%), på bekostning af andre arter. Derefter faldt bidraget fra kiselalger svagt mod 2021, hvor furealger og andre arter fik større betydning. I de indre åbne farvande var der ingen tydelig udvikling i biomassefordelingen, beregnet ud fra årsmidlerne, mellem kiselalger, furalger og andre arter. I dette område var betydningen af kiselalger væsentligt mindre, og både furealger og andre arter bidrog med en større andel af biomassen (figur 2.16A/B).



Figur 2.16. Den årlige fordeling mellem kiselalger, furealger og andre arter for (A) fjorde og kystnære områder og (B) indre åbne farvande.

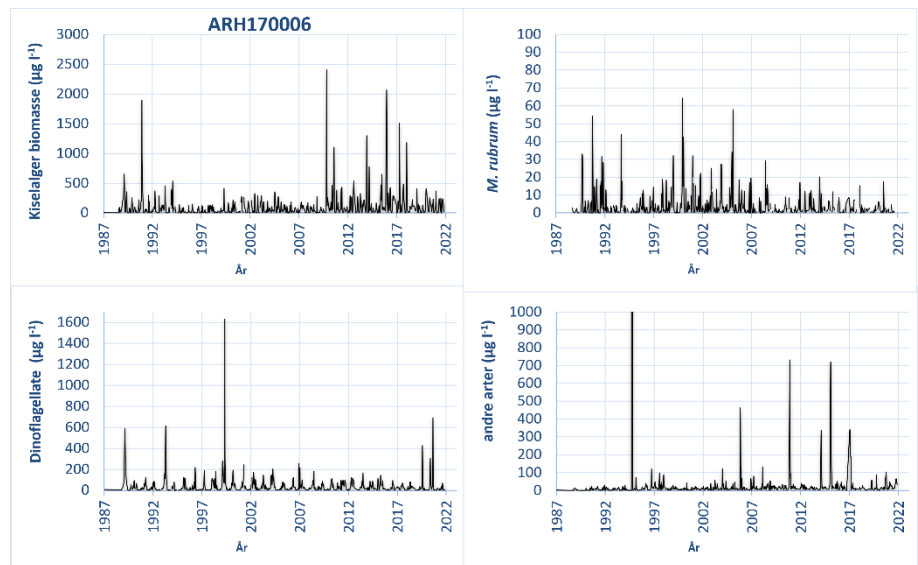
Figur 2.17 og figur 2.18 viste eksempler på udviklingen i kulstofbiomassen på lokale stationer. Figur 2.17 viste stationen VIB3727 (Løgstør Bredning). Der var i perioden fra 2002 og frem til 2010 periodiske høje opblomstringer af kiselalger, som var karakteristiske forårsoptagelser. I perioden efter 2010 faldt frekvensen af algeopblomstringer af kiselalger frem mod 2021. På samme station blev forekomsten af kiselalger erstattet af furealger, som blev den dominerende gruppe i opblomstringen og generelt optrådte sent om sommeren (figur 2.13B).

Figur 2.17. Fytoplanktonsamfundsstrukturen i Løgstør Bredning.



Figur 2.18 var et eksempel på en station, ARH170006 (Århus Bugt), hvor der blev observeret generelt samme udvikling i kiselalger og furealger i perioden fra 1978 til 2021. I midten af perioden (1997 og 2007) optrådte den mixotrofe ciliat med periodiske høje forekomster, som forsvandt i de følgende 15 år. Det skal også bemærkes, at der var periodisk meget høje kulstofbiomasser om foråret i 1996 (6/4/1996).

Figur 2.18. Fytoplanktonsamfundsstrukturen i Århus Bugt.



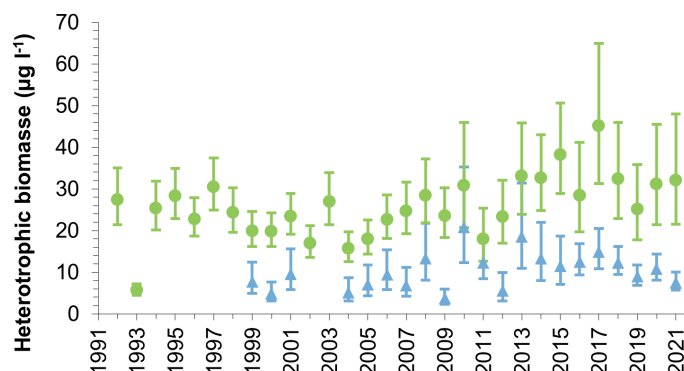
Denne afvigelse blev diskuteret tidligere (Figur 2.2-2.6), og den unormale kulstofbiomassetop blev allerede undersøgt og godkendt i MST's faglige kvalitets sikring. Der var derfor i dette tilfælde god overensstemmelse mellem den faglige og FDC's KS i nærværende eksempel. Der findes flere eksempler på enkeltstation for fytoplankton i bilag B3.

Heterotroft protistplankton

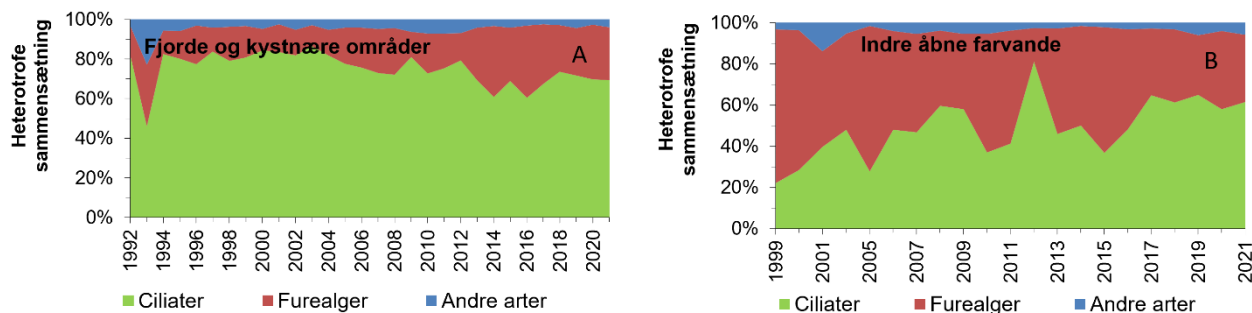
Heterotroft protistplankton var écellede organismer, der var tildelt heterotrofi som ernæringsform. Disse arter blev fundet i bilag B2 og blev indsamlet i henhold til de tekniske anvisninger TA09 (epi-fluorescens) for heterotrofe nanoflagellater og furealger (Fossing og Jakobsen 2015), samt ciliater efter TA10 (Jakobsen og Fossing 2017). Der blev historisk indsamlet prøver for heterotrofe protister fra stationerne ARH170006, BRKBMPK2, DMU444, KBH431, NOR409,

NOR5503, NOR7715, RKB1, ROS60, VIB3708, VIB3727. Bilag B4 viste omfanget af prøvetagningen. I denne analyse opdelte vi det heterotrofe protistplankton i ciliater, heterotrofe furealger samt andre arter (figur 2.19), og stationer og perioder for heterotrofe protister, som blev angivet i tabel 4.1 i bilag B4, blev medtaget. Andre arter bestod af mindre heterotrofe flagellater og nanoflagellater, der typisk spiser bakterier og flagellater < 2 µm. Ciliater spiste typisk partikler mellem 1 og 10 µm, og furealger spiste endnu større partikler, såsom kiselalger.

Figur 2.19. Årsmidler af kulstofbiomasse ± 95 % konfidensgrænser fjorde og kystnære områder (●) og de indre åbne farvande (▲) for det heterotrofe protistplankton i perioden 1992 til 2021.

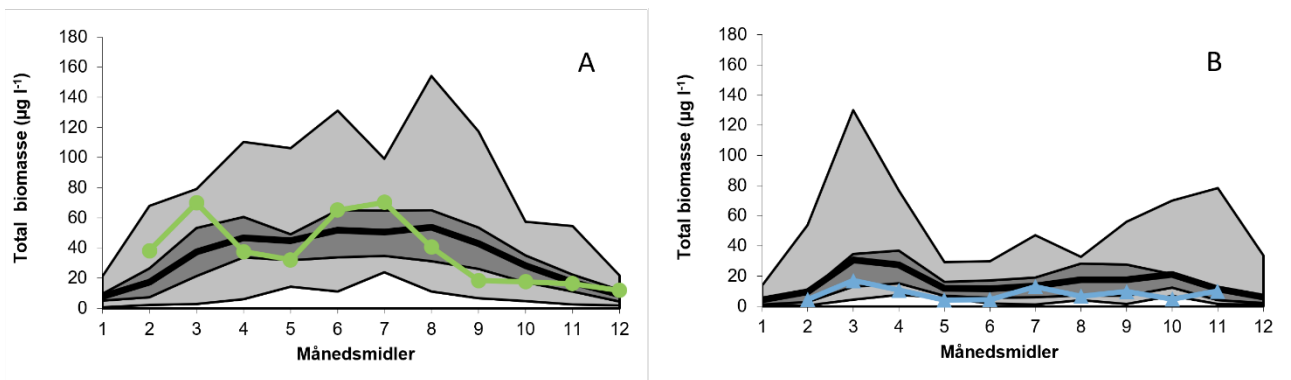


Den integrerede gennemsnitskulstofbiomasse pr. år for det samlede heterotrofe protistplanktonsamfund for fjorde og kystnære områder i perioden fra 1991 viste kulstofbiomasser omkring 25 µg C L⁻¹, indtil omkring 2007, hvorefter de gennemsnitlige kulstofbiomasseværdier kortvarigt faldt til mellem 10 – 20 µg C L⁻¹. Derefter steg de gennemsnitlige kulstofbiomasseværdier til 25 - 45 µg C L⁻¹ og toppede i 2017, hvorefter de begyndte at falde. Udviklingen i det heterotrofe protistsamfund for de indre åbne farvande viste kulstofbiomasseværdier mellem 10 µg C L⁻¹ og 15 µg C L⁻¹, med en faldende tendens mod 2021 (figur 2.19)



Figur 2.20. Kulstofbiomassen som månedsgennemsnit for ciliater og heterotrofe furealger i fjorde og kystnære (1992 – 2021) områder samt indre åbne farvande i perioden 1999 til 2021.

Fordelingen mellem ciliater, heterotrofe furealger og andre arter for fjorde og kystnære områder viste, at ciliater dominerede på tværs af hele perioden og udgjorde typisk > 75 % (Figur 2.20A). Efter 2014 var der et fald i ciliaternes biomasseproportioner til omkring 60 % til fordel for furealger, der steg tilsvarende. I de indre åbne farvande udgjorde ciliaterne sjældent mere end 60 %, og furealger udgjorde typisk 25 %. Derudover var der et stigende bidrag fra ciliater siden 1999 (Figur 2.20B). Heterotrofe mikro- og nanoflagellater udgjorde mindre end 5 % i begge systemer (Figur 2.20 A/B)



Figur 2.21. Månedsmiddelværdier af total kulstofbiomasse af heterotroft protistplankton i fjorde og kystnære områder (1991 - 2021) samt indre åbne farvande (1999 - 2021). Figureerne er angivet med minimum og maksimum (lysegrå afgrænsning), nedre og øvre kvartil (mørkegrå afgrænsning) samt medianværdien (fed sort linje). Den grønne linje er for fjorde og kystnære områder for 2021 og den blå linje er månedsmidler for 2021 indre åbne havområder.

3 Afsluttende bemærkninger og konklusion

Kvalitetssikring blev udført iterativt i flere trin/omgange. I den første kørsel blev der identificeret store afvigelser, og det kunne konstateres, at der manglede bio- og plasmavolumener for en lang række af de ældre data på artsniveau, og at der ikke var overensstemmelse mellem kulfaktorerne og de tekniske anvisninger. Efter erstatning af manglende bio- og plasmavolumener med standardværdier og opdatering af kulfaktorerne fremkom de anvendte algoritmer til afvigelseskontrol med langt færre fejl. Der blev fundet en del afvigelser på individniveau (39.195 ud af 978.070), hvilket typisk skyldtes afvigende dimensionsangivelser i forhold til de anvendte formler. Det var for nærværende ikke muligt at gennemgå alle, og disse individuelle registreringer blev derfor angivet som 'UK' (ukendt). Det var mindre alvorligt for anvendelsen af datasættet, da der i stedet blev anvendt standard bio- og plasmavolumener på artsniveau. En stor del af afvigelserne på individniveau ville kunne rettes efterfølgende, men dette var for omfattende inden for rammen af nærværende projekt.

Vi har identificeret 90.711 individ registreringer, 24.014 artregistreringer samt 568 prøver (ud af 16.377 prøver) og markeret dem UK. Alle typer vil fremadrettet kunne rettes til og hovedparten kunne indgå i det endelige datasæt.

Fytoplanktonkonsolideringsprojektet har kvalitetssikret data fra 1985 og indtil 2021. Det er 37 år, og der tilføres stadig data. Perioden repræsenterer et betydeligt arbejde udført af et veluddannet og erfarent fagligt miljø i de tidligere amter, samt af tidligere konsulenter og nuværende konsulenter. Mange af medarbejderne i amterne overgik til Miljøstyrelsen, da disse blev nedlagt ved kommunalreformen, og de har løbende sikret, at data er systematisk sikret i det nu nedlagte fagsystem STOQ. Dette engagement genkendes i data og er eksemplificeret ved, at kun 586 ud af 16.377 optalte prøver er sat under kontrol efter gennemgang af data.

FDC KS er gennemløbet med flere iterationer, hvor f.eks. systematiske fejl, der ikke fandtes i MST faglige kontrol, blev identificeret. Dette gælder bl.a. udfordringerne med at tilføje manglende cellevolumener. FDC-KS viser sæsonmønstre som forventet.

Der er udviklinger over tid på niveau af fjorde og kystnære og indre, åbne havområder. På stationsniveau peger de præsenterede stationer på lokale dynamikker, hvilket er generelt for mange af de fleste stationer. Der er fundet en del afvigelser i kvalitetskontrollen, der tidligere har været udført og har sat det som UK. Derfor er der intet grundlag for at lægge prøver UK efter FDC-faglig kvalitetskontrol, da afvigelser fundet er rettet og genkendt i FDC faglig kvalitetssikring.

Under FDC's KS blev der beregnet periodegennemsnit for måneder og år for fjorde og kystnære områder samt indre, åbne farvande for hele perioden fra 1978. Det er vigtigt at bemærke, at der for perioden før 1989 var relativt få NOVANA-stationer. Samtidig var NOVANA-programmet under udvikling, og det nuværende metodegrundlag blev endelig fastlagt i 1991 (Olrik 1991). Det betød, at de få aktive stationer før 1989 havde relativt stor vægt i beregningen af periodegennemsnit. Gennemgang af data viste, at biomassen ikke fulgte nogen fast statistisk fordeling (f.eks. normalfordeling eller log-normalfordelinger), og at alle gennemsnitsværdier derfor var underlagt en vis usikkerhed. Denne

skævhed blev delvist imødegået ved at anvende den statistiske model fra Hansen m.fl. (2020).

Implementeringen af bestemmelsen af heterotrofe protister i NOVANA-programmet er sket løbende i halvfemserne, og metoden ser ud til først at være fuldt indarbejdet i overvågningen omkring 1995 på tværs af stationerne i bilag B4. Bestemmelsen af ernæringstype i datasættet pegede på inkonsistente oplysninger om trofi og mangler i angivelsen af ernæringstypen før 2011. Det medførte udviklingen af en liste med 146 kendte heterotrofe arter, som sammen med heterotrofe nanoflagellater, bestemt ved hjælp af fluorescensmikroskopi, udgjorde grundlaget for analysen. Fremadrettet anbefalede AU, at angivelsen af heterotrofi tilføjes til arter, der er opført på listen i bilag B2. Derudover anbefaler AU, at trofi knyttes til artsregisteret. Datagrundlaget for heterotrofi viste troværdige forskelle mellem fjorde og kystnære områder og de indre, åbne områder. Dette peger på systematiske forskelle, som fremadrettet kan implementeres i arbejdet med dele af havstrategien der vedrører biodiversitets- og fødekædedescriptorer.

4 Referencer

Cano EL, Moguerza JH, Corcoba MP (2015) *Quality Control with R*. Springer Verlag, Heidelberg, New York, Dordrecht, London

CEN (2011) *Water quality - Guidance on quantitative and qualitative investigations of marine phytoplankton*. European committee for standardization

CEN (2015) *European Standard 16695:2015 Water quality - Guidance on the estimation of phytoplankton biovolumen*. European committee for standardization

Dansk-Standard (2010) *DS ISO 16269-4: Statistical interpretation of data - Part 4: Detection and treatment of outliers*.

Edler L (1979) *Phytoplankton and chlorophyl*. In: Edler L (ed) *Recommendations on methods for marine biological studies in the Baltic Sea*. Baltic Working Group 9 and The National Swedish Environment Protection Board, Malmö

Fossing H, Hansen JLS, Bruhn A, Krause-Jensen D, Carstensen J, Jakobsen H, Markager S, Hansen JW, Strand J, Larsen MM (2018) *Beskrivelse af metoder som det marine fagdatacenter anvender ved den faglige kvalitetssikring af de enkelte marine dataemner*.

Fossing H, Jakobsen HH (2015) *TA09 Fytoplankton. Tekniske anvisninger*. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi

Hansen JW, Bruhn A, Carstensen J, Dahl K, Galatius A, Göke C, Hansen JLS, Jakobsen HH, Krause-Jensen D, Larsen ML, Markager S, C. M, Petersen IK, Strand J, Teilmann J, Tougaard J (2015) *Marine områder 2015 NOVANA. Videnskabelig rapport fra DCE, Book 208*. Aarhus Universitet DCE - Nationalt Center for Miljø og Energi, *Marine områder 2015 NOVANA*, <https://dce2.au.dk/pub/SR208.pdf>

Hansen JW, Høgslund S, Bruhn A, Carstensen J, Dahl K, Galatius A, Göke C, Hansen JLS, Jakobsen HH, Krause-Jensen D, Kyhn LA, Larsen ML, Markager S, C. M, Petersen IK, Strand J, Stæhr PAU, Sveegaard S, Tairova Z, Teilmann J, Tougaard J (2020) *Marine områder 2020 NOVANA. Videnskabelig rapport fra DCE, Book SR475*. Aarhus Universitet DCE - Nationalt Center for Miljø og Energi, Aarhus <https://dce2.au.dk/pub/SR475.pdf>

Harrison PJ, Zingone A, Mickelson MJ, Lehtinen S, Ramaiah N, Kraberg A, McQuatters A, Sun J, Jakobsen HH (2015) *Cell biovolumes of marine phytoplankton from global data sets*. *Estuarine, Coastal and Shelf Science* 162:130-142.

Høgslund S, Hansen JLS, Bruhn A, Krause-Jensen D, Carstensen J, Jakobsen HH, Markager S, Hansen JW, Dahl K, Galatius A, Tougaard Jb, Sveegaard S, Kyhn L, Feld L, Larsen MM (2019) *Beskrivelse af metoder til KS af dataemner i NOVANA-rapportering for marine områder*. ver. 1.

Jakobsen HH, Carstensen J, Harrison PJ, Zingone A (2015) *Estimating time series phytoplankton carbon biomass: Inter-lab comparison of species*

identification and comparison of volume-to-carbon scaling ratios. *Estuarine Coastal and Shelf Science* 162:143-150. <Go to ISI>://WOS:000360774700015, 10.1016/j.ecss.2015.05.006

Jakobsen HH, Fossing H (2017) TA10 Mikrozooplankton. Tekniske anvisninger, Book 2. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi, Aarhus,

Kim M, Drumm K, Daugbjerg N, Hansen PJ (2017) Dynamics of sequestered cryptophyte nuclei in *Mesodinium rubrum* during starvation and refeeding. *Frontiers in Microbiology* 8<Go to ISI>://WOS:000396907700001, 10.3389/fmicb.2017.00423

Menden-Deuer S, Lessard EJ (2000) Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnology and Oceanography* 45:569-579. <Go to ISI>://WOS:000086754900005, DOI 10.4319/lo.2000.45.3.0569

Olrik K (1991) Planteplanktonmetoder Prøvetagning, bearbejdning og rapportering ved undersøgelser af planteplankton i søer og marine områder. Miljøprojekt Book 187, København, <https://www2.mst.dk/Udgiv/publikationer/1991/87-503-9411-8/pdf/87-503-9411-8.pdf>

Putt M, Stoecker DK (1989) An experimentally determined carbon: volume ratio for marine "oligotrichous" ciliates from euclidean and coastal waters. *Limnology and Oceanography* 34:1097-1103.

Tada K, Pithakpol S, Yano R, Montani S (2000) Carbon and nitrogen content of *Noctiluca scintillans* in the Seto Inland Sea, Japan. *Journal of Plankton Research* 22:1203-1211. <https://doi.org/10.1093/plankt/22.6.1203>, 10.1093/plankt/22.6.1203

Zingone A, Harrison PJ, Kraberg A, Lehtinen S, McQuatters-Gollop A, O'Brien T, Sun J, Jakobsen HH (2015) Increasing the quality, comparability and accessibility of phytoplankton species composition time-series data. *Estuarine Coastal and Shelf Science* 162:151-160. <Go to ISI>://WOS:000360774700016, 10.1016/j.ecss.2015.05.024

5 Bilag

Bilag B1 Tabel over stationer, der er samlet og struktureret i ODA

Station angiver stationsnavn, prioriteret er stationer, der er udpeget i fællesskab med MST, antal prøver er antallet af unikke prøver, der indgår i stationen, og antal år er antallet af årstal, der er registreret på stationen med første og sidste indsamlings år.

Tabel B1.1 Liste over stationer der er undersøgt for dubletter

Station	Position (WG84)	Prioriteret	Antal år	Antal prøver	Periode (første - sidste år)
VIB3708	57.0 N 9.06 Ø	X	38	869	1983-2021
VIB3727	56.6 N 9.08 Ø	X	38	881	1983-2021
ARH170006	56.2 N 10.3 Ø	X	35	777	1978-2021
FYN6300043	55.0 N 10.2 Ø	X	34	641	1985-2021
RKB1	56.1 N 8.22 Ø	X	34	787	1986-2021
ROS60	55.7 N 12.1 Ø	X	34	672	1986-2021
RIB1510007	55.3 N 8.35 Ø	X	32	463	1990-2021
NOR5503	56.7 N 9.97 Ø	X	30	789	1991-2021
VEJ0006870	55.7 N 10.1 Ø	X	25	532	1992-2021
VIB3702	56.6 N 8.41 Ø	X	25	619	1983-2009
NOR409	56.9 N 10.8 Ø	X	24	478	1998-2021
BRKBMPK2	55.3 N 16.0 Ø	X	23	98	1999-2021
FYN6900017	55.5 N 10.5 Ø	X	23	494	1998-2021
VSJ20925	56.1 N 11.2 Ø	X	23	474	1998-2021
KBH431	55.9 N 12.8 Ø	X	22	448	1998-2021
RIB1610002	55.5 N 8.34 Ø		22	471	1993-2016
VEJ0005790	55.9 N 9.91 Ø	X	22	435	1971-2010
FRB65	55.9 N 12.0 Ø		21	240	1986-2006
VEJ0004273	55.7 N 9.67 Ø	X	21	402	1988-2009
ARH190004	56.6 N 10.4 Ø	X	20	413	1989-2009
FRB1877	55.9 N 12.6 Ø		20	205	1987-2006
ROS1727	55.5 N 12.4 Ø		20	265	1987-2006
FRB1939	56.2 N 12.3 Ø		18	190	1989-2006
RIB1510010	55.5 N 8.21 Ø		18	198	1989-2006
RKB23	56.3 N 8.26 Ø	X	18	228	1985-2004
SJY3	55.1 N 8.58 Ø		18	296	1994-2014
VSJ10003	55.9 N 11.8 Ø		16	320	1986-2010
VIB3720	56.5 N 8.67 Ø	X	14	222	1983-2014
FYN6101014	55.6 N 10.1 Ø		13	264	1985-1997
VIB3711	57.0 N 9.63 Ø	X	13	254	1983-1997
RKB21	56.4 N 8.15 Ø	X	12	137	1985-1998
RKB22	56.3 N 8.21 Ø	X	12	284	1998-2010
VIB3723	56.9 N 8.74 Ø	X	12	251	1983-1997
KBH1723	55.6 N 12.4 Ø		11	126	1987-1997
RKB43	56.0 N 8.11 Ø	X	11	120	1989-2000

VSJ10001	55.7 N 11.8 Ø		10	115	1986-1997
VSJ10002	55.8 N 11.7 Ø		10	114	1986-1997
VSJ10006	55.7 N 11.8 Ø		10	110	1986-1997
VSJ44011	55.3 N 11.2 Ø		10	108	1987-1998
VSJ45012	55.3 N 11.2 Ø		10	109	1987-1997
FYN6700053	55.5 N 10.9 Ø	X	9	164	1989-2011
KBH1944	55.6 N 12.7 Ø		9	101	1989-1997
RKB41	56.2 N 8.09 Ø	X	9	82	1988-1997
STO0101015	55.1 N 11.6 Ø		9	169	1988-2011
VSJ42009	55.6 N 11.1 Ø		9	105	1989-1997
VSJ51013	55.3 N 11.3 Ø		9	100	1987-1997
KBH1722	55.6 N 12.5 Ø		8	91	1989-1997
KBK3001	55.6 N 12.5 Ø		8	100	1990-1997
NOR7715	57.6 N 9.92 Ø	X	8	118	2005-2021
ROS1724	55.6 N 12.3 Ø		8	78	1987-1996
SJYKFF5	54.8 N 9.82 Ø		8	176	1998-2005
VSJ20004	56.0 N 11.4 Ø		8	93	1989-1997
VSJ30006	55.8 N 11.2 Ø		8	93	1989-1997
VSJ41008	55.7 N 11.1 Ø		8	92	1989-1997
VSJ50015	55.2 N 11.3 Ø		8	91	1989-1997
VSJ53016	55.2 N 11.5 Ø		8	85	1989-1997
STO802008	55.2 N 11.6 Ø		7	136	1990-2009
DMU444	55.0 N 13.3 Ø	X	6	109	2016-2021
FRB8000	55.9 N 11.8 Ø		6	60	1998-2003
KBK3005	55.7 N 12.7 Ø		6	260	1992-1997
ROS1726	55.4 N 12.3 Ø		6	68	1987-1992
ROS2044	55.7 N 11.8 Ø		6	67	1986-1992
VSJ41007	55.7 N 10.9 Ø		6	74	1989-2007
FRB1935	55.9 N 11.9 Ø		5	56	1988-1992
FRB1993	56.0 N 11.9 Ø		5	144	1993-1997
FYN0018841	55.4 N 10.6 Ø		5	55	1990-1999
FYN6500062	54.9 N 10.5 Ø		5	70	1992-1997
Hjerting Strand#1610	55.5 N 8.35 Ø		5	107	1987-1991
RIB1410003	55.7 N 8.14 Ø		5	59	1996-2006
SJYKFF2	54.8 N 9.49 Ø		5	116	1993-1997
VSJ43010	55.4 N 11.1 Ø		5	54	1987-1992
DMU1026	56.6 N 6.29 Ø	X	4	6	2018-2021
DMU1075	56.1 N 5.20 Ø	X	4	6	2018-2021
DMU1077	56.1 N 3.34 Ø	X	4	6	2018-2021
DMU1130	57.2 N 8.53 Ø	X	4	6	2018-2021
DMU1133	57.5 N 8.41 Ø	X	4	7	2018-2021
FRB75	55.9 N 11.9 Ø		4	17	1986-2003
FRB9000	56.1 N 12.5 Ø		4	9	2001-2004
RIB1410002	55.7 N 8.17 Ø		4	38	1988-2002
ROS2	55.7 N 12.1 Ø		4	46	1989-1992
STO0704010	55.0 N 12.2 Ø		4	39	1988-1992
VSJ30005	55.8 N 11.3 Ø		4	45	1989-1992
VSJ43019	55.5 N 11.1 Ø		4	44	1987-1998

VSJ43021	55.4 N 11.2 Ø		4	45	1987-1998
VSJ51014	55.3 N 11.3 Ø		4	42	1989-1992
station 3#8843	55.7 N 9.13 Ø		4	45	1990-1994
DMU1023	56.6 N 7.77 Ø	X	3	4	2019-2021
DMU1035	56.2 N 7.72 Ø	X	3	4	2018-2021
DMU1042	55.7 N 7.76 Ø	X	3	4	2018-2021
DMU1052	55.3 N 7.97 Ø	X	3	4	2018-2021
FRB2013	56.0 N 12.6 Ø		3	8	1998-2004
FYN0018825	55.5 N 10.6 Ø		3	32	1990-1992
FYN6500051	55.0 N 10.4 Ø		3	66	2005-2011
FYN6910801	55.5 N 10.5 Ø		3	37	1992-1995
KBH5004	55.8 N 12.6 Ø		3	31	1989-1991
Ketting Nor#S057001	55.7 N 9.13 Ø		3	50	2001-2003
RIB1610008	55.5 N 8.42 Ø		3	28	1990-1992
RIB1620014	55.3 N 8.55 Ø		3	22	1990-1992
STO0801006	55.2 N 12.3 Ø		3	21	1988-1990
STO0901016	54.9 N 12.3 Ø		3	75	2001-2003
STO0901032	54.8 N 12.4 Ø		3	73	1998-2000
VIB3709	57.0 N 9.25 Ø	X	3	24	1997-2009
VSJ41018	55.7 N 11.0 Ø		3	37	1991-1993
VSJ43023	55.5 N 11.1 Ø		3	28	1996-1998
st. 52#52	55.7 N 9.13 Ø		3	32	1989-1992
ARH170002	56.2 N 10.4 Ø		2	21	1989-1991
ARH170115	56.1 N 10.5 Ø		2	14	2002-2003
ARH170117	56.0 N 10.4 Ø		2	14	2002-2003
ARH199090	56.6 N 10.7 Ø		2	9	2012-2018
ARH210977	56.3 N 10.8 Ø		2	6	2012-2018
ARH230990	56.6 N 10.3 Ø		2	12	2012-2018
DMU1046	55.7 N 6.30 Ø	X	2	4	2020-2021
FRB1876	56.1 N 12.4 Ø		2	10	1987-2001
FRB62	55.8 N 12.1 Ø		2	21	1986-1991
FYN0018501	54.7 N 10.7 Ø		2	12	1991-2004
FYN0018531	54.8 N 10.7 Ø		2	12	1990-2004
FYN0018843	55.5 N 10.6 Ø		2	9	1991-1992
KBK1728	55.5 N 12.7 Ø		2	23	1987-1988
KBK3003	55.7 N 12.6 Ø		2	22	1990-1991
Limfjorden st. 48#14	55.7 N 9.13 Ø		2	26	1987-1990
NOR403	57.2 N 10.7 Ø		2	3	2006-2007
NOR5504	56.6 N 9.84 Ø		2	5	1997-1999
STO0802008	55.2 N 12.1 Ø		2	39	2010-2011
VEJ0003772	55.7 N 9.59 Ø	X	2	20	1988-1989
VEJ0004669	55.7 N 9.74 Ø	X	2	20	1988-1989
VEJ0005367	55.7 N 9.86 Ø	X	2	18	1988-1989
VSJ10017	55.7 N 11.8 Ø		2	26	1994-1995
VSJ10018	55.7 N 11.8 Ø		2	26	1994-1995
VSJ10021	55.7 N 11.8 Ø		2	26	1994-1995
VSJ20023	56.0 N 11.8 Ø		2	23	1994-1995
ARH160090	56.1 N 10.6 Ø		1	3	2012-2012

ARH250032	56.0 N 10.2 Ø		1	18	2006-2006
FRB1875	56.0 N 11.9 Ø		1	9	1987-1987
FRB1937	56.1 N 12.1 Ø		1	1	2001-2001
FRB8550	55.9 N 12.0 Ø		1	2	2002-2002
FRBILT03-01	56.0 N 12.6 Ø		1	1	2001-2001
FYN0018111	55.4 N 9.81 Ø		1	11	1991-1991
FYN0018121	55.5 N 9.81 Ø		1	11	1991-1991
FYN0018310	55.1 N 10.3 Ø		1	11	1992-1992
FYN0018320	54.9 N 10.5 Ø		1	11	1992-1992
FYN0018361	55.1 N 10.4 Ø		1	11	1989-1989
FYN0018702	55.0 N 10.7 Ø		1	12	1989-1989
FYN0018712	55.1 N 10.7 Ø		1	12	1989-1989
FYN0018842	55.4 N 10.6 Ø		1	1	1991-1991
FYN0018882	55.6 N 10.6 Ø		1	10	1990-1990
FYN6100052	55.7 N 10.3 Ø		1	12	1991-1991
Halkær Bredning st.	55.7 N 9.13 Ø		1	16	1990-1990
KBH1942	55.8 N 12.6 Ø		1	7	1988-1988
KBKSVANE	55.7 N 12.6 Ø		1	7	1988-1988
Limfjorden 132#132	55.7 N 9.13 Ø		1	18	1987-1987
Limfjorden st. 11#11	55.7 N 9.13 Ø		1	24	1987-1987
Limfjorden st. 11#L1	55.7 N 9.13 Ø		1	24	1987-1987
Limfjorden st. 23#12	55.7 N 9.13 Ø		1	18	1987-1987
Limfjorden st. 25#12	55.7 N 9.13 Ø		1	21	1987-1987
Limfjorden st. 42#14	55.7 N 9.13 Ø		1	23	1987-1987
Limfjorden st. 51#15	55.7 N 9.13 Ø		1	5	1987-1987
Limfjorden st. 52#15	55.7 N 9.13 Ø		1	19	1987-1987
Lunkebugten#18691	55.7 N 9.13 Ø		1	12	1991-1991
NOR3310	57.6 N 10.5 Ø		1	21	1997-1997
NOR4410	56.9 N 10.5 Ø		1	1	2006-2006
NOR5542	56.7 N 9.88 Ø		1	2	1999-1999
NOR6602	56.9 N 9.56 Ø		1	16	1990-1990
Nordlige Lillebælt,	55.7 N 9.13 Ø		1	4	2004-2004
RF-benthiske mikroal	55.7 N 9.13 Ø		1	13	2006-2006
RIB1510003	55.4 N 7.92 Ø		1	5	2002-2002
RIB1510006	55.4 N 8.07 Ø		1	1	1990-1990
RIB1510009	55.3 N 8.14 Ø		1	5	2002-2002
RKB4	55.9 N 8.19 Ø		1	14	1992-1992
RKB42	56.2 N 7.97 Ø	X	1	7	1988-1988
RKB45	56.0 N 8.07 Ø		1	8	1993-1993
RKB46	56.0 N 8.03 Ø		1	8	1993-1993
ROS52	55.7 N 12.0 Ø		1	11	1986-1986
ROS55	55.7 N 12.0 Ø		1	11	1986-1986
STO0101050	55.2 N 11.5 Ø		1	1	2011-2011
STO0601076	54.7 N 11.8 Ø		1	11	1992-1992
VEJ0003350	55.5 N 9.53 Ø	X	1	14	2004-2004
VEJ0006489	55.8 N 10.0 Ø	X	1	10	1992-1992
VEJ0007677	55.7 N 10.2 Ø		1	12	1992-1992
VSJ20002	56.0 N 11.8 Ø		1	10	1986-1986

VSJ30020	55.9 N 11.4 Ø	1	11	1991-1991
VSJ42021	55.6 N 11.0 Ø	1	12	1993-1993
VSJ42022	55.6 N 11.1 Ø	1	12	1993-1993
VSJ50019	55.2 N 11.2 Ø	1	12	1990-1990

Bilag B2 Oversigt over de heterotrofe protister

Herunder er en tabel over 146 arter af heterotrofe protister

Sc1067	Phylum	Class
<i>Acanthostomella norvegica</i>	Ciliophora	Oligotrichea
<i>Alexandrium pseudogoniaulax</i>	Dinophyta	Dinophyceae
<i>Amphidinium</i>	Dinophyta	Dinophyceae
<i>Amphidinium crassum</i>	Dinophyta	Dinophyceae
<i>Amphidinium sphenoides</i>	Dinophyta	Dinophyceae
<i>Archaeoperidinium minutum</i>	Dinophyta	Dinophyceae
<i>Askenasia</i>	Ciliophora	Litostomatea
<i>Balanion</i>	Ciliophora	Prostomatea
<i>Calliacantha</i>	Sarcomastigophora	Zoomastigophora
<i>Calliacantha natans</i>	Sarcomastigophora	Zoomastigophora
<i>Ciliophora</i>	Ciliophora	
<i>Craspedophyceae</i>	Sarcomastigophora	Craspedophyceae
<i>Diaphanoeca</i>	Sarcomastigophora	Zoomastigophora
<i>Diaphanoeca grandis</i>	Sarcomastigophora	Zoomastigophora
<i>Diaphanoeca sphaerica</i>	Sarcomastigophora	Zoomastigophora
<i>Didinium</i>	Ciliophora	Litostomatea
<i>Dinophysis</i>	Dinophyta	Dinophyceae
<i>Dinophysis hastata</i>	Dinophyta	Dinophyceae
<i>Diplopelta bomba</i>	Dinophyta	Dinophyceae
<i>Diplopelta pusilla</i>	Dinophyta	Dinophyceae
<i>Diplopsalis</i>	Dinophyta	Dinophyceae
<i>Diplopsalis gruppen</i>	Dinophyta	Dinophyceae
<i>Diplopsalis lenticula</i>	Dinophyta	Dinophyceae
<i>Diplopsalopsis</i>	Dinophyta	Dinophyceae
<i>Diplopsalopsis orbicularis</i>	Dinophyta	Dinophyceae
<i>Ebria</i>		Ebriophyceae
<i>Ebria tripartita</i>		Ebriophyceae
<i>Eutintinnus</i>	Ciliophora	Oligotrichea
<i>Eutintinnus pectinis</i>	Ciliophora	Oligotrichea
<i>Favella</i>	Ciliophora	Oligotrichea
<i>Favella ehrenbergii</i>	Ciliophora	Oligotrichea
<i>Favella serrata</i>	Ciliophora	Oligotrichea
Flagellat	Sarcomastigophora	Flagellat
<i>Gyrodinium britannicum</i>	Dinophyta	Dinophyceae
<i>Gyrodinium dominans</i>	Dinophyta	Dinophyceae
<i>Gyrodinium fusiforme</i>	Dinophyta	Dinophyceae
<i>Gyrodinium lachryma</i>	Dinophyta	Dinophyceae
<i>Gyrodinium spirale</i>	Dinophyta	Dinophyceae
<i>Gyrodinium spirale/britannicum</i>	Dinophyta	Dinophyceae
<i>Helicostomella</i>	Ciliophora	Oligotrichea
<i>Helicostomella subulata</i>	Ciliophora	Oligotrichea
<i>Laboea strobila</i>	Ciliophora	Oligotrichea

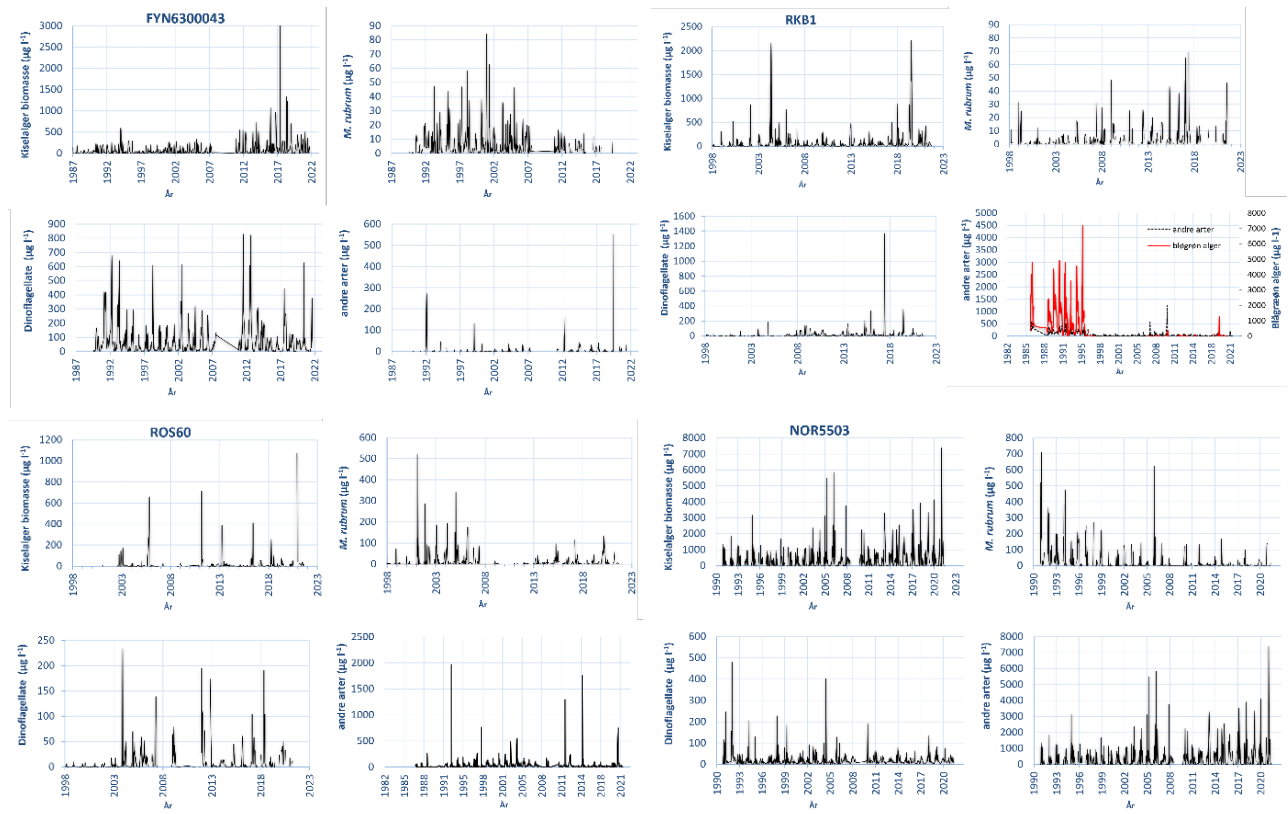
<i>Leegaardiella sol</i>	Ciliophora	Oligotrichea
<i>Leprotintinnus pellucidus</i>	Ciliophora	Oligotrichea
<i>Leucocryptos</i>	Katablepharidophyta	Katablepharidophyceae
<i>Leucocryptos marina</i>	Katablepharidophyta	Katablepharidophyceae
<i>Lingulodinium polyedrum</i>	Dinophyta	Dinophyceae
<i>Lohmanniella</i>	Ciliophora	Oligotrichea
<i>Lohmanniella oviformis</i>	Ciliophora	Oligotrichea
<i>Mesodinium</i>	Ciliophora	Litostomatea
<i>Mesodinium pulex</i>	Ciliophora	Litostomatea
<i>Mesodinium velox</i>	Ciliophora	Litostomatea
<i>Miracanthodinium</i>	Dinophyta	Dinophyceae
<i>Miracanthodinium claytonii</i>	Dinophyta	Dinophyceae
<i>Monosiga</i>	Sarcomastigophora	Zoomastigophora
<i>Nematodinium armatum</i>	Dinophyta	Dinophyceae
<i>Nematopsides vigilans</i>	Dinophyta	Dinophyceae
<i>Oblea</i>	Dinophyta	Dinophyceae
<i>Oblea rotunda</i>	Dinophyta	Dinophyceae
<i>Oligotrichida</i>	Ciliophora	Oligotrichea
<i>Oxyrrhis marina</i>	Dinophyta	Dinophyceae
<i>Pachysphaera</i>	Chlorophyta	Pyramimonadophyceae
<i>Parafavella</i>	Ciliophora	Oligotrichea
<i>Parafavella denticulata</i>	Ciliophora	Oligotrichea
<i>Parvicorbicula socialis</i>	Sarcomastigophora	Zoomastigophora
<i>Pelagostrobilidium spirale</i>	Ciliophora	Oligotrichea
<i>Peridiniella danica</i>	Dinophyta	Dinophyceae
<i>Phalacroma rotundatum</i>	Dinophyta	Dinophyceae
<i>Planctonema</i>	Chlorophyta	Trebouxiophyceae
<i>Polykrikos</i>	Dinophyta	Dinophyceae
<i>Polykrikos kofoidii</i>	Dinophyta	Dinophyceae
<i>Preperidinium meunieri</i>	Dinophyta	Dinophyceae
<i>Procentrum cordatum</i>	Dinophyta	Dinophyceae
<i>Proterospongia</i>	Sarcomastigophora	Zoomastigophora
<i>Protoperidinium</i>	Dinophyta	Dinophyceae
<i>Protoperidinium achromaticum</i>	Dinophyta	Dinophyceae
<i>Protoperidinium bipes</i>	Dinophyta	Dinophyceae
<i>Protoperidinium breve</i>	Dinophyta	Dinophyceae
<i>Protoperidinium brevipes</i>	Dinophyta	Dinophyceae
<i>Protoperidinium cerasus</i>	Dinophyta	Dinophyceae
<i>Protoperidinium claudicans</i>	Dinophyta	Dinophyceae
<i>Protoperidinium conicoides</i>	Dinophyta	Dinophyceae
<i>Protoperidinium conicum</i>	Dinophyta	Dinophyceae
<i>Protoperidinium conicum gruppen</i>	Dinophyta	Dinophyceae
<i>Protoperidinium crassipes</i>	Dinophyta	Dinophyceae
<i>Protoperidinium curtipes</i>	Dinophyta	Dinophyceae
<i>Protoperidinium curvipes</i>	Dinophyta	Dinophyceae

<i>Protooperidinium deficiens</i>	Dinophyta	Dinophyceae
<i>Protooperidinium denticulatum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium depressum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium divergens</i>	Dinophyta	Dinophyceae
<i>Protooperidinium excentricum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium granii</i>	Dinophyta	Dinophyceae
<i>Protooperidinium leonis</i>	Dinophyta	Dinophyceae
<i>Protooperidinium mariebourae</i>	Dinophyta	Dinophyceae
<i>Protooperidinium mite</i>	Dinophyta	Dinophyceae
<i>Protooperidinium oblongum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium ovatum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium pallidum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium paulsenii</i>	Dinophyta	Dinophyceae
<i>Protooperidinium pellucidum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium pentagonum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium punctulatum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium pyriforme</i>	Dinophyta	Dinophyceae
<i>Protooperidinium steinii</i>	Dinophyta	Dinophyceae
<i>Protooperidinium steinii/pyriforme</i>	Dinophyta	Dinophyceae
<i>Protooperidinium subinerme</i>	Dinophyta	Dinophyceae
<i>Protooperidinium thorianum</i>	Dinophyta	Dinophyceae
<i>Protooperidinium thulesense</i>	Dinophyta	Dinophyceae
<i>Ptychocylis urnula</i>	Ciliophora	Oligotrichea
<i>Salpingella</i>	Ciliophora	Oligotrichea
<i>Salpingoeca</i>	Sarcomastigophora	Zoomastigophora
<i>Sclerodinium calyptroglyphe</i>	Dinophyta	Dinophyceae
<i>Scrippsiella trochoidea</i>	Dinophyta	Dinophyceae
<i>Spatulodinium pseudonociluca</i>	Dinophyta	Dinophyceae
<i>Stenosemella</i>	Ciliophora	Oligotrichea
<i>Stenosemella ventricosa</i>	Ciliophora	Oligotrichea
<i>Strobilidium</i>	Ciliophora	Oligotrichea
<i>Strombidinopsis</i>	Ciliophora	Oligotrichea
<i>Strombidium</i>	Ciliophora	Oligotrichea
<i>Strombidium conicum</i>	Ciliophora	Oligotrichea
<i>Strombidium cornucopiae</i>	Ciliophora	Oligotrichea
<i>Strombidium emergens</i>	Ciliophora	Oligotrichea
<i>Strombidium vestitum</i>	Ciliophora	Oligotrichea
<i>Suctorina</i>	Ciliophora	Phyllopharyngea
<i>Telonema</i>		
<i>Telonema subtile</i>		
<i>Tiarina</i>	Ciliophora	Prostomatea
<i>Tiarina fusus</i>	Ciliophora	Prostomatea
<i>Tintinnidae</i>	Ciliophora	Oligotrichea
<i>Tintinnidium</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis</i>	Ciliophora	Oligotrichea

<i>Tintinnopsis beroidea</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis campanula</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis cylindrica</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis fluviatile</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis magna</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis minuta</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis parva</i>	Ciliophora	Oligotrichea
<i>Tintinnopsis rapa</i>	Ciliophora	Oligotrichea
<i>Tontonia</i>	Ciliophora	Oligotrichea
<i>Uronema</i>	Ciliophora	Oligohymenophorea
<i>Uronema/cyclidium</i>	Ciliophora	Oligohymenophorea
<i>Urotrichia</i>	Ciliophora	Prostomatea
<i>Vorticella</i>	Ciliophora	Oligohymenophorea

Bilag B3 Udvalgte stationer som eksempel på variationer

Bilag B3 viser enkelte udvalgte stationer, der er medtaget som eksempel på store lokale variationer mellem de enkelte NOVANA stationer i biomassen.



Bilag B4 NOVANA stationer med heterotrofe protister

Liste over NOVANA stationer med systematisk optælling af heterotrofe protister. Antal år angiver antallet af år hvor der er indsamlet ciliater.

Tabel 4.1. Antal år angiver antallet af år hvor der er indsamlet ciliater, periode angiver tidsrum og mangler angiver år hvor der af ikke kendt grunde ikke er indsamlet data.

Station	Havområde	Longi- tude	Lati- tude	antal år	Periode (første - sidste år)	mangler
ARH170006	fjord og kystnær station	10,32	56,15	33	1989-2021	
BRKBMPK2	indre åben farvand	15,98	55,25	6	2016-2021	
DMU444	indre åben farvand	13,30	55,00	6	2016-2021	
KBH431	indre åben farvand	12,75	55,87	11	2011-2021	
NOR409	indre åben farvand	10,79	56,86	17	1999-2021	2002-2003 og 2012-2015
NOR5503	fjord og kystnær station	9,97	56,66	16	1991-2006	
NOR7715	Nordsø	9,92	57,61	5	2017-2021	
RKB1	fjord og kystnær station	8,22	56,08	27	1989-2015	
ROS60	fjord og kystnær station	12,07	55,71	25	1995-2021	2008 og 2010
VIB3708	fjord og kystnær station	9,06	56,95	37	1985-2021	2007-2008
VIB3727	fjord og kystnær station	9,08	56,62	31	1985-2015	2011

Bilag B5 QA_Dublet_klassifikation

Filen QA_Dublet_Classification_ALL.xlsx giver et overblik over samtlige prøver samt deres dubletstatus (Kolonne G / Active). Dubletterne er inddelt i grupper, hvor den aktive (1) er grøn, og den/de inaktive (0) er gul/gule.

	A	B	C	D	E	F	G	H	I
	Date	Station	Sample_ID	Depth	System	Primary_Sample_ID	Active	Score	Group
93	20021023	VIB3702-00001	28628	358	2	28628	1		16180
94	20021023	VIB3702-00001	28629	335	3	28628	0	99	16180
95	20021024	NOR5503	28320	517	2	28320	1		16181
96	20021024	NOR240008	27952	350	4	28320	0	86	16181
97	20021024	NOR5503	3106	350	11	28320	0	79	16181
98	20021024	NOR5503	28621	517	2	28621	1		16182
99	20021024	NOR240008	27956	502	4	28621	0	91	16182
00	20021024	NOR5503	3110	502	11	28621	0	85	16182
01	20021029	ARH170006	28006	770	4	28006	1		16183
02	20021029	ARH170006	28024	770	4	28024	1		16184
03	20021029	ARH190004	27878	420	4	27878	1		16185
04	20021029	DMU431	28742	520	2	28742	1		16186

Tabel 5.1 Eksempel på dubletter. Den fulde liste findes i filen QA_Dublet_Classification_ALL.xlsx

Sample_ID er prøve-ID, som er unikt. Der er i alt 36.124 unikke prøve-ID'er i datasættet. *Primary_Sample_ID* er den prøve-ID, der er aktiv/bruges i gruppen. Hvis *Sample_ID* og *Primary_Sample_ID* er ens, vil *Active* være 1. Hvis der er prøver uden dubletter, er der derfor kun én linje i gruppen, f.eks. gruppe 16.183 (tabel 5.1). Typisk varierer dybden ofte inden for samme gruppe, hvilket er årsagen til at dybden ikke anvendes i den initiale potentielle dubletgruppering. Tilsvarende er der eksempler på varierende stationsnavne i samme gruppe, og derfor anvendes station ligeledes ikke i den initiale potentielle dubletgruppering.

Bilag B6 QA_Dublet_klassifikation_20021024

Filen QA_Dublet_Classification_20021024.xlsx er et eksempel på håndtering af dubletter på en given dato, nemlig den 24/10/2002, hvor de underliggende data er vist. Eksemplet indeholder følgende 6 prøver fra 3 forskellige systemer (tabel 6.1).

Dubletudredning samt faglig vurdering har identificeret 2 prøver som unikke. De 4 øvrige er dubletter, som deaktiveres (men stadig gemmes). Den faglige vurdering resulterer i, at WSP (system 2) prøverne blev valgt som aktive. Det kunne argumenteres for, at det i virkeligheden kun er en prøve, der er opdelt i 2 sub-arter, men det er uden for dette projekts ramme at vurdere det. Som det også ses, er der forskellige stations- og dybdeangivelser. Dette er generelt for hele datasættet og har nødvendiggjort, at alle kombinationer af dato og artsprofiler er blevet gennemgået for dubletter.

Date	Station	Sample_ID	Depth	System	Primary_Sample_ID	Active
20021024	NOR5503	28320	517	2	28320	1
20021024	NOR5503	3106	350	11	28320	0
20021024	NOR240008	27952	350	4	28320	0
20021024	NOR5503	28621	517	2	28621	1
20021024	NOR5503	3110	502	11	28621	0
20021024	NOR240008	27956	502	4	28621	0

Tabel 6.1 Eksempel på dubletter på en station på datoen 24/10/2002

Tabel 6.1 viser et eksempel på dubletter på en station på datoen 24/10/2002. Eksemplet i tabel 6.1 indeholder 6 prøver, der kan kombineres parvis i 15 mulige kombinationer. Projektet har udviklet en score-algoritme, der implementeres i den videre analyse. Kolonne AN i QA_Dublet_Classification_20021024.xlsx (arket kaldet "Score") indeholder en samlet score, der er udgangspunktet for yderligere kontrol. En score under 50 indikerer en usandsynlig dublet, over 70 er sandsynligvis en dublet, og over 90 er med stor sandsynlighed en dublet. Filen "QA_Dublet_Classification_20021024.xlsx" er vedhæftet som bilag.

	A	B	C	D	E	F	G	H	I	J	K	AN	
1	Date	S1	S2	S1_Station	S2_Station	S1_Depth	S2_Depth	S1_Sample_ID	S2_Sample_ID	S1_Time	S2_Time	Score	
2	20021024		4	11	NOR240008	NOR5503	502	350	27956	3106	1200	903	17
3	20021024		4	11	NOR240008	NOR5503	350	502	27952	3110	1200	903	17
4	20021024		2	4	NOR5503	NOR240008	517	350	28621	27952	903	1200	20
5	20021024		2	4	NOR5503	NOR240008	517	502	28320	27956	903	1200	23
6	20021024		2	11	NOR5503	NOR5503	517	350	28621	3106	903	903	27
7	20021024		4	4	NOR240008	NOR240008	350	502	27952	27956	1200	1200	29
8	20021024		2	11	NOR5503	NOR5503	517	502	28320	3110	903	903	29
9	20021024		11	11	NOR5503	NOR5503	350	502	3106	3110	903	903	30
10	20021024		2	2	NOR5503	NOR5503	517	517	28320	28621	903	903	32
11	20021024		4	11	NOR240008	NOR5503	350	350	27952	3106	1200	903	71
12	20021024		4	11	NOR240008	NOR5503	502	502	27956	3110	1200	903	76
13	20021024		2	11	NOR5503	NOR5503	517	350	28320	3106	903	903	79
14	20021024		2	11	NOR5503	NOR5503	517	502	28621	3110	903	903	85
15	20021024		2	4	NOR5503	NOR240008	517	350	28320	27952	903	1200	86
16	20021024		2	4	NOR5503	NOR240008	517	502	28621	27956	903	1200	91
17													
18													
19													
20													
21	Not Dublet												
22	Dublet												
23	Diff												
24	Identical												

Tabel 6.2 Eksempel på dubletter på scoring af dubletter. Farver illustrerer potentielle dubletprøver.

Tabel 6.2 er nøgletal fra kombinationstabellen med f.eks. artsprofiler, genereret ved at konverterer til de mest opdaterede artskoder, under hensyntagen til forskellige stadier, størrelser, osv.

Messages		
n	S1_Profile_Specie	S2_Profile_Specie
	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...
	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...
	1_1#1152_1#2347_1#2347_2	1_1#1152_1#2347_1#2347_2
	1_1#1152_1#2347_1#2347_2	1152_1#2347_1#2347_2#3865_1
	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...	901_1#1030_1#1036_1#1038_1#1043_1#1049_1#1114_1#...
	1_1#1152_1#2347_1#2347_2	1152_1#2347_1#2347_2#3865_1

Tabel 6.2 Eksempel på artprofil dubletter på scoring af dubletter. Fra filen "QA_Dublet_Classification_20021024.xlsx". Den samlede score ses i ark "20021024")

Kombinationer af prøver i mellemområdet hvor der kan være usikkerhed, tjekkes yderligere og det samlede resultat samt data er sendt videre til faglig vurdering. I den faglige vurdering som ud over personlig faglig og historisk viden benytter tabel data som kan ses i arkene 3106, 3110, 27952, 27956, 28320, 28621 som er de prøveID'er der er dækket dato 20021024.

20021024	Score	28320	3106	27952	28621	3110	27956	2025984	2025988	2114275	2171392
----------	-------	-------	------	-------	-------	------	-------	---------	---------	---------	---------

Tabel 6.3 Eksempel på 6 artprofiler markeret med gult, fra scoring af dubletter. Fra filen "QA_Dublet_Classification_20021024.xlsx". Den samlede score ses i ark "20021024")

Eksemplet i ark 28320 bruges til direkte sammenligning med de øvrige prøver f.eks. 3106 & 27952. Den faglige vurdering resulterer i accept eller afvisning. Arkene 2025984, 2025988, 2114275 & 2171392 indeholder WSP data som kan bruges til at sammenholdes med de data der ligger i ODAM

Station_ID	Station	Number	Longitude	Latitude	Depth	Source_ID				
3884	NOR5503	93610032		9,973667	56,662667	27	80000061			
Activity_ID	Activity_Dato	Activity_Klok	Source_ID							
113654	20021024	903	80005080							
ProeveID	MiddelDybde(cm)	PlanktonSystemKode	ProeveTypeKode	ProeveUdstyrKode	ProeveVolumen(L)	MaskeVidde(um)	Proevetager	Analytiker	LabKo	
28320	517	2		0	38		Uffe Frisenette	Per Andersen		
DybdeID	ProeveID	Dybde(cm)	Source_ID							
1	29409	28320	100	2114275						
2	29410	28320	200	2114275						
3	29411	28320	400	2114275						
4	29412	28320	600	2114275						
5	29413	28320	800	2114275						
6	29414	28320	1000	2114275						
7										
ProeveID	Specie_ID	Sc1067_ID	Sc1067	BestemmelseUsikkerhedKode	StoerrelsesGruppeKode	ErnaeringsBioKode	Specifikation	PlanktonTypeKode	Metod	
3	28320	736851	901 Pseudo-nitzschia delicatissima gruppen	0	0	0				
0	28320	736846	1030 Chaetoceros socialis	0	0	0				
1	28320	736834	1036 Coscinodiscus	2	0	0				
2	28320	736849	1038 Coscinodiscus granii	0	0	0				
3	28320	736840	1043 Cerataulina pelagica	0	0	0				
4	28320	736836	1049 Ditylum brightwellii	0	0	0				
5	28320	736845	1114 Skeletonema costatum	0	0	0				
5	28320	736847	1976 Cryptophyceae	2	-5	0				
7	28320	736837	2016 Mesodinium rubrum	0	6	0				

Tabel 6.4 Ark 20021024.

Sammenhæng kan ses ved at benytte Kilde_ID kan ses i feltet N8 i A_Dublet_Classification_20021024.xlsx

	K	L	M	N
4				
5				
6				
7	LabStartDato	LabSlutDato	Bemaerkning	Source_ID
8	08-12-2002	08-12-2002	Slange_Interval: 1+2+4+6+8+10	2114275
9				
10				

Tabel 6.5 Ark 20021024.

KONSOLIDERING AF PLANKTONDATA I NOVANA

Fytoplanktondata

Denne rapport beskriver arbejdet med at indlæse og kvalitetssikre NOVANA planktondata, indsamlet under de tidligere amter, samt data fra Miljøstyrelsens STOCQ database. Data er blevet gennemløbet adskillige iterationer i kvalitetssikringsprocessen, og der er blevet foretaget en del korrektioner i konverteringskonstanterne. Der er udført to typer kvalitetssikring: 1. multivariat kontrol samt 2. kvalitetssikring efter metoden udviklet af AU. Det endelige godkendte datasæt omfatter 15.809 unikke optalte prøver af fyto- og heterotroft protist plankton