Response to interactive comment of anonymous referee #2

By Hedy M. Aardema in agreement with co-authors.

Reviewer: *This paper analyses spatial and temporal patterns in cruise data with 3 high-resolution monitoring methods: FRRF, Flow-cytometry and Ferrybox. Correlations between the observed variables are also analysed. The large dataset, including many phytoplankton and environmental variables observed together enables the authors to understand the patterns in the various phytoplankton variables. The results could guide the optimal application of such novel monitoring methods in operational monitoring for a.o. MSFD.*

Response: We thank the reviewer for the helpful and critical comments. We rewrote and restructured the manuscript extensively based on these comments. We are happy to hear that the reviewer sees the potential of our applied methods.

General comments

Reviewer: *The paper lacks a clearly stated research question or hypothesis to be tested. Therefore, it is unclear what is the purpose of the various analyses performed and what we can learn from the results. Based on the conclusion that this type of "high-resolution is a very useful supplement to current monitoring", I would expect a hypothesis such as "combined high-resolution monitoring of many phytoplankton variables along with environmental variables allows us to quantify seasonal and meso-scale patterns in phytoplankton biomass, species composition and primary production. The concurrent measurement of different phytoplankton variables allows us to understand the effect of phytoplankton species composition and physiological adaptation processes on the observed patterns in phytoplankton biomass and production". Then the analysis should show how the variables should be combined to provide the most reliable estimates of phytoplankton biomass and primary production.*

Response: Because of the exploratory nature of our research, a hypothesis was not defined. The addition of the suggested sentences does help in making the manuscript easier to follow. We therefore adopted part of the sentences and added of the following sentences to the introduction: "The aim of this study is to test the suitability of these two high-resolution methods to be developed as novel phytoplankton monitoring method. The two high-resolution methods, a flowcytometer and a FRR fluorometer, were deployed concurrently on four 4-day cruises in April, May, June and August to meet a wide range of environmental conditions and phytoplankton community states. These measurements allow for quantification of seasonal and mesoscale spatial patterns in phytoplankton abundance, photophysiology and gross primary production. In this paper we provide an overview of the acquired results, use a spectral cluster analysis to visualize spatial heterogeneity and evaluate the potential of these methods to optimize current monitoring programs."

Reviewer: *There are many observed variables, which are not consistently named in the text, figures and tables. Therefore, it is easy to get lost in the description of patterns for all individual variables. A clear definition of variables that is consistently used throughout the text would help the reader to understand the storyline. Some of the variables observed by the FRRF seem to be very similar. Which of the variables should be used as indicator and which are redundant to answer the research questions?*

Response: We corrected the inconsistent naming. The variables of the FRRf might seem similar under some conditions. However, because these variables vary depending on community composition and environmental conditions, they might deviate when conditions change (Sugget et al., 2009; Kromkamp and Forster, 2003). Therefore, care must be taken into choosing the parameters. For the current study the main interest is on monitoring the phytoplankton community, therefore we are interested in parameters that are informative on physiological adaptation or characteristic for phytoplankton taxons. Additionally, we focus on high resolution measurements, so limit the parameters to the ones attainable at high-resolution. Based on these considerations we decided to

include the current parameters, which give us a broad overview of the photophysiological status of the phytoplankton community.

Reviewer: *In the conclusions section a recommendation on next steps would be much appreciated: what would be required to use the high-resolution methods in scope to provide reliable estimates of phytoplankton biomass, production and species composition for long term monitoring? In the introduction and conclusion the species composition is defined in functional types such as nitrogen fixers, calcifiers or DMS-producers, but these do not correspond to the phytoplankton clusters used in this paper.*

Response: the conclusions were rewritten:
"A good monitoring program monitors the presence of functional types of phytoplankton, including the harmful taxons, the carrying capacity of the ecosystem and changes in biogeochemical cycling. The objective of this study was to evaluate the use of FRR fluorometry and flowcytometry for such monitoring purposes. The four conducted cruises spread over 5 months offered a wide variety of environmental conditions and phytoplankton community states, which the utilized methods were able to visualize. Inclusion of high-resolution methods in monitoring programs allows for analysis of finer scale events. Furthermore, it allows for analysis of living phytoplankton and is thereby able to measure rates and avoid effects of preservation and storage of samples. Another advantage is that high-resolution methods allows for easier comparison between countries, once common protocols have been established. Nevertheless, low resolution methods remain a necessity for more detailed taxonomic analysis, information on vertical heterogeneity, to calibrate and to correct for blanks. Data analysis might be the biggest bottleneck of the implementation of these high-resolution methods. The cluster analysis of flowcytometric data has high potential for improvement to increase the informative value of the method. Especially identification of phytoplankton clusters with a functional quality, such as nitrogen fixers, calcifiers or DMS-producers, would be helpful for interpretation of ecosystem dynamics and biogeochemical fluxes. Regarding the FRRf, the main challenge is converting electron transport rate to gross primary productivity in carbon units. Further research in these topics would benefit implementation of these methods into monitoring protocols. Furthermore, it is important to account for diurnal patterns in monitoring set-up to be able to distinguish between diurnal and spatial variability. Possibly the diurnal variability could be modelled, but more studies with a Langragian based approach would be needed for a better understanding of the impact of diurnal variability in the data. Overall, the in this study presented high-resolution measurement set-up has large potential to improve phytoplankton monitoring in supplement to existing low-resolution monitoring programs."
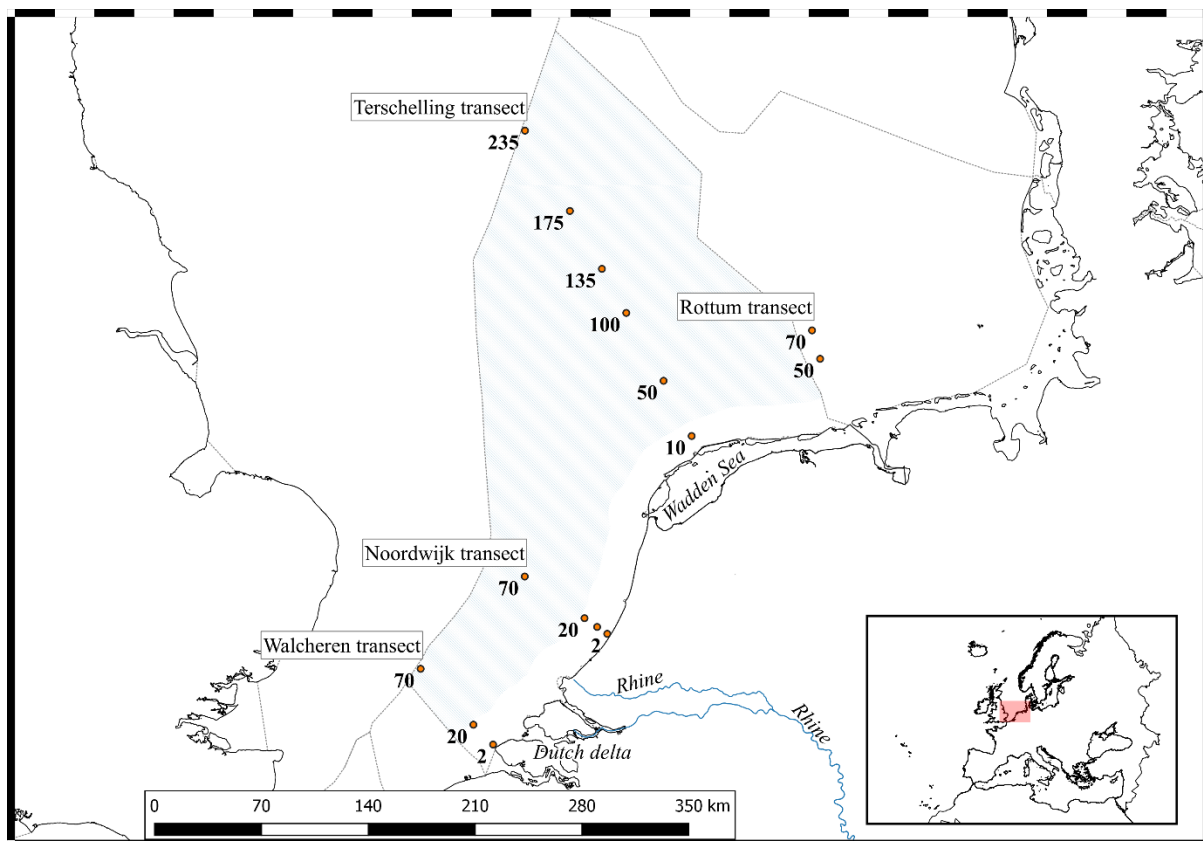
Specific comments

Reviewer: *Sentences are often long: consider breaking up in multiple sentences to improve readability.*

Response: We apologize for the difficulties and hope to have improved the readability in the new manuscript.

Reviewer: *Figure 1: please show only the stations (with names/ abbreviations) used in this study (see table S1) and the areas used in the text (such as Dogger Bank, Wadden, Den Helder, Rhine outflow) so the description of spatial patterns can also be understood by people that are not Dutch.*

Response: We updated the figure to the following:

Reviewer: *Section 2.2: please refer to international protocols/methods rather than internal protocols.*
Response: We added a more detailed description to the method section.

Reviewer: Table 1: it would help to have an additional column stating the interpretation / meaning of this variable, such as total biomass, nutrient stress, maximum growth rate, efficiency of light uptake etc. Then later in the text you can use these 'meaningful' names instead of codes, to facilitate understanding of observed patterns. Also a figure illustrating the meaning of the different variables (alfa,Ek, F', Fm' etc.) could prevent getting lost in all abbreviations.
Response: Unfortunately, the meaning of the different variables is usually not straightforward and dependent on multiple predictors (species, nutrient concentration, light availability etc.; Suggett et al., 2009). Nonetheless, we tried to make the table more information easier to understand.

Reviewer: *Equation 9: why did you use monthly averaged irradiance if you are looking at high-resolution patterns. Why did you not use irradiances measured during the cruise?*
Response: Unfortunately, we were unable to collect reliable irradiance data for all cruises. Clearly, it is preferable to have irradiance (PAR) continuously measured in parallel to the FRRF measurements when aiming to monitor current primary productivity.

Reviewer: *Table 2: Since you use both Length_FWS and O/R ratio as criteria to distinguish the phytoplankton groups, it would be logical to include a column for O/R ratio with the applied criteria.*
Response: Good idea, we added the O/R-ratio to the table.

Reviewer: *It is not entirely clear whether pico-red includes pico-Synecho or not. On page 14, line 30 it says: "Both groups of picophytoplankton (Synechococcus and total)", whereas table 2 and figure 3 suggest the two groups are exclusive.*
Response: Pico-red and Pico-synecho are two different groups, as correctly understood from table 2 and figure 3. We rephrased the sentence to: "Both groups of picophytoplankton (Synechococcus and Pico-red)", and scanned the manuscript for other mixing up.

Reviewer: *Section 2.4: please state with every type of analysis what is the purpose / research questions for that analysis. For example: what are you trying to predict from what and why?*

Response: We added the following the sentences to section 2.4: "To find regions with similar phytoplankton communities, data was spectrally clustered using the uHMM R package (Poisson-caillault and Ternynck, 2016) in the statistical software R (version 3.4.1, R Core Team, 2017)." and "Principal Component Analyses (PCA) were performed to find which variables contributed most to the cluster results."

Reviewer: *Section 3.1: I don't see the value of comparing averages over whole transects (with large spatial variability, which is the subject of this paper), that are not even the same, between months. The only thing you see is seasonal patterns that are well-known from other studies and that can be summarized in section 2.1 in a description of the study area. Most of this section describes the data in table S1. I would replace table 3 with table S1 and remove table S2. N/P ratios address that same question as table S1, but with an indicator that is controversial.*

Response: For the authors the table helped to visualize the seasonal patterns, but we agree on the comment that this table does not add to the already existing knowledge on seasonal patterns. We therefore adopted the suggestion to replace the table with the table S1 from the supplementary material.

Reviewer: *The text in this section (and subsequent sections) is sometimes hard to follow as it is not clearly structured in time and space and variable. We go back and forth in time. Section 3.2 describes first figure 2, then figure 3 and then again figure 2 and then figure 3. I suggest to make one section about phytoplankton biomass (figure 2) and then one section on species composition (figure 3).*

Response: Rewritten

Reviewer: *Page 16, line 14: the southern coastal stations are more strongly affected by the Rhine outflow than the Scheldt outflows (see for example: Lacroix, G., Ruddick, K., Ozer, J., & Lancelot, C. (2004). Modelling the impact of the Scheldt and Rhine/Meuse plumes on the salinity distribution in Belgian waters (southern North Sea). Journal of Sea Research, 52(3), 149-163.).*

Response: We reformulated to Rhine and Scheldt river outflow.

Reviewer: *Figure 4: Please use consistent legends for the same variable between different months, with the same colour scheme and symbols (squares vs. circles) and with blue indicating low values and red indicating high values, so the high values stand out, more than the low values. Also captions in the table per line (red fluorescence, O/R ratio etc.) and per column (april, may etc.) would help to easier understand the figure.*

Response: We remade the figures, see manuscript.

Reviewer: *Section 3.5: I don't see the added value of this analysis. What does it tell us?*

Response: Agreed. We aimed to get a better understanding of the drivers of primary productivity in the Dutch North Sea. However, we realize now that the dataset is not very well suited for this and we therefore removed the analysis.

Reviewer: *Page 24: I suggest to mention in the table all the variables that were included in the analysis and note coefficients or 'ns' for not significant and the p values per explanatory variable. Then readers don't need to reconstruct the overview from the text. Actually, the significance test is likely not valid due to strong spatial autocorrelation in the data.*

Response: The Multiple Linear Regression was removed from the manuscript, because of the lack of information derived from it together with the abundance of literature already addressing the predictors of primary productivity.

Reviewer: *Discussion: Here I would expect to get some advice: How to best estimate phytoplankton biomass from these data? Should we use total red fluorescence (best R2) or F0 (least affected by NPQ)? Is there a way to combine both (with other available variables) to get an even better estimate?*

Response: We added the following paragraph to the discussion:
 "Biomass might be one of the most important parameters to understand phytoplankton dynamics, but its direct measurement is not possible using high-resolution methods. Chlorophyll a concentration is often used as an estimate for biomass, although the Carbon:Chl a ratio is dependent on abiotic conditions and species-specific phenotypic plasticity (Flynn, 1991, 2005; Geider et al., 1997; Alvarez-Fernandez and Riegman, 2014; Halsey and Jones, 2015). Red fluorescence gave a good estimate of chlorophyll a concentration, both using the FRRf (adjusted R2= 0.66) and FCM (adjusted R2=0.90). Both the FRRf and the flowcytometer estimate the chlorophyll a concentration based upon the fluorescence in the red spectrum after excitation in the blue spectrum. There are some slight differences in the optics, the FRRf excites with a 450 nm LED and measures the fluorescence at 682 ± 30 nm, while the FCM excites at 488 nm and filters the red fluorescence over a longpass 650 nm filter towards the red fluorescence detector. The smaller detection range of the FRRf detector is optimized around the maximum emission of PSII and limits contamination by PSI (Franck et al., 2002; Oxborough et al., 2012). The second difference is the fluorescent state of the photosystems, the strong laser of the flowcytometer can only measure the maximum fluorescence (Fm), which is a parameter more prone to quenching than the minimum fluorescence measured by the FRRf. Yet, the biggest difference concerns the method; where the flowcytometer measures the fluorescence per particle, the FRRf does only a bulk measurement. In a bulk measurement other particles in solution scatter the excitation and emission photons, plus the emitted fluorescence of the phytoplankton is subject to reabsorption, especially at higher biomass densities. The latter seems to have the most impact on chlorophyll a concentrations, as the fit of the flowcytometer derived red fluorescence is a better than the FRRf minimum fluorescence. Other studies that use the FCM to estimate chlorophyll a concentrations also showed good relationships, but find better fits using the bulk measurements using a fluorimeter (Thyssen et al., 2015; Marrec et al., 2018). The conversion to biomass may also be done from cell abundances. Some studies use the oversimplified assumption that all cells have a spherical shape and a constant C content per biovolume (Tarran et al., 2006). With the scanning flowcytometer it is also possible to estimate biovolume based on scattering properties of the cell, but this relationship appears to be taxon specific (Rijkeboer, pers. comm.). This relationship will be further explored by comparing the calculated biovolume based on the Image in Flow pictures and the flowcytometric properties of these phytoplankters."

Reviewer: *Can we trust GPP from FRRF as a reliable estimate of primary production or is more work needed to achieve that goal? If so, what needs to be done?*

Response: We added the following paragraph to the discussion:
"The reliability of variable fluorescence as estimate of gross primary productivity is depending on many cell processes from the photon absorbance to carbon assimilation. The variable fluorescence reflects the first step of photosynthesis; the efficiency of which photons are captured and electrons produced and transferred. However, to interpret gross primary productivity in an ecological or biogeochemical meaningful way, the FRR units of electrons per unit time need to be converted to carbon units. Gross photosynthesis correlates well with photosynthetic oxygen evolution (Suggett et al., 2003), and multiple studies have shown good correlation between 14C-derived estimates of primary productivity and FRRf-derived estimates using a constant conversion factor (Melrose et al., 2006; Kromkamp et al., 2008). However, in reality this parameter is not a constant, as along the pathway from electron to carbon atom electrons are consumed by other cell processes (Flameling and Kromkamp, 1998; Halsey and Jones, 2015; Schuback et al., 2016). Therefore, a reliable GPP estimate in carbon units from FRR fluorometry requires more research and estimates provide relative rather than qualitative values. Despite its limitations the fact that the method can measure in situ, with relatively little phytoplankton manipulation before measurement, makes the method promising. Calibration with other methods, such as concurrent C14 of C13 incubations, could help to better understand the processes from electron excitation to carbon fixation. However, it should be recognized that these types of measurements come with their own problems, and measure something

in between net and gross primary productivity depending on the incubation time and growth rate of the phytoplankton (Halsey and Jones, 2015). So it remains a question which method is measuring the 'real' primary productivity. Attempts to calculate primary productivity from flowcytometer data have also been made, which is actually based on the diurnal cycle in cell size caused by cell division (Marrec et al., 2018). Despite the limitations of GPP estimates by variable fluorescence, our results clearly show large spatial variability in gross primary production concurrent to the expected strong variability during the growth season. This spatial heterogeneity is not fully captured by sampling at the standard low-resolution monitoring stations, showing the added value of our approach. Primary productivity was highest in April, and relatively large values were also observed offshore, indicating that a low phytoplankton biomass does not necessarily means that primary production is low. Our GPP rates were based on the same electron requirement for C-fixation (Φe,C). However, this is a likely oversimplification as Φe,C is known to vary with abiotic conditions (Lawrenz et al., 2013) and the changes in nutrient conditions and temperature during the growth season are likely to affect GPP. This will be the topic of a future publication and we expect that the detection of several biogeographic regions will help us in predicting Φe,C."

Reviewer: *It is not really clear whether the diurnal variability in the FRRF variables is a problem that needs to be solved.*

Response: It is not so much a problem that needs to be solved, but it does need to be taken into account when setting up a monitoring program including FRRF variables. It is important to realize that measurements taken at different times of the day, might not be comparable. To be able to include FRRF variables in a long-term monitoring program, the included sampling points should be sampled at the same time of the day.

Reviewer: *Are the clusters in the FCM analysis the relevant ones to provide 'useful' information to science & society? Should we / Can we move on to other clusters that are mentioned in the conclusions?*

We added the following text to the discussion:
"To understand the role of the phytoplankton in biogeochemical cyckes, the FCM clusters would ideally reflect taxonomic or functional groups, as calcifiers, silicifiers, DMS producers (such as Phaeocystis) and nitrogen fixers (le Quéré et al., 2005). The lack of identification of distinct clusters makes this sofar impossible. Other studies manually separate up to 10 phytoplankton groups with the same instrument (Marrec et al., 2018). These groups included Prochlorococcus, which is at the absolute limit of resolving capacity of the FCM because of their small size and low fluorescence. They furthermore distinghuished the Pico-red in three groups based on FLO/FLR-ratio. Nano-cryptophytes group in high and low orange fluorescence and included a micro-eukaryotes group with a size from 10 to 20 uM. But these groups are still made up of many taxonomic genera and, apart from size, won't allow much for further interpretation of their role in the ecosystem or biogeochemical cycles. The same accounts for detection of nuisance phytoplankton; distinct clusters of toxic phytoplankton species are lacking. Although this will remain a challenge because toxicity in phytoplankton can differ within morphotypes and sometimes even differ per strain within a species (Tillman and Rick, 2003). But potentially, further research in flowcytometry can result in suspicious clusters to be flagged and further inspected by a specialist using microscopy. The potential is certainly there, as much of the information retrieved by the FCM is still unexplored; the clustering is performed on totals (area under the peak) instead of the pulse-shape. This in combination with more advanced camera options will need to further distinguish between groups in the future."

Reviewer: *Do the FCM data help to better understand the FRRF data (and vice versa)? For example, do we see diatoms under light limited conditions (high F'/Fm', high alfa, low Ek) and picoplankton under nutrient limited conditions (low F'/Fm')? Other ecological niches that we know from literature? Different conditions promoting Synechococcus compared to other picoplankton?*

Response: We tried to incorporate the link between the methods better, we added the following sentences to the manuscript:

"In this study a large part of the Dutch North Sea shifted from nutrient sufficiency to nutrient limitation between April and May, which was reflected in the low efficiency of PSII (Fv/Fm; Fig. 4). The Fv/Fm recovered between May and June, which suggest that the phytoplankton adapted to nutrient limiting conditions (Kruskopf and Flynn, 2005). However, photophysiological parameters are also varying per taxonomic group; smaller taxa typically have lower Fv/Fm values and higher σPSII values (Kolber et al., 1988; Suggett et al., 2009b). Indeed, by flowcytometry we find that the biggest shift in community composition took place between May and June from a nanophytoplankton dominated community to a picophytoplankton dominated community. These findings demonstrate how flowcytometry and fast repetition rate fluorometry can supplementary improve ecosystem understanding."

Technical comments

Reviewer:
- Collinear should be spelled with 2 ll's throughout the whole text.
- Page 9, line 4 & 5: I guess um means micrometers?
- Page 18, line 4: middle-right, please refer to the label C4 a-x.
- Figure 5: The figure would be easier to read if the colours per group are consistent between the cluster analysis on the right and the map on the left. Labels (A-D for April to August panels) would also help.
- Figure 9: Please add the hours of the day on the x-axis.
- Page 25, line 3: the word influenced is repeated too many times and therefor should get an e in the end.
- Page 28, line 11: estimates are.
- Line 13: parameter without s.
- Page 31, line 8: Jerico-next, without h.

Response: We adopted the suggested technical improvements.

References

Franck, F., Juneau, P., Popovic, R. (2002). Resolution of the Photosystem I and Photosystem II contributions to chlorophyll fluorescence of intact leaves at room temperature. Biochimica et Biophysica Acta - Bioenergetics, 1556(2–3), 239–246. https://doi.org/10.1016/S0005-2728(02)00366-3

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