# High resolution analysis of a North Sea phytoplankton community structure based on *in situ* flow cytometry observations and potential implication for remote sensing.

**M. Thyssen[1*], S. Alvain[1], A. Lefèbvre[2], D. Dessailly[1], M. Rijkeboer[4], N. Guiselin[1], V. Creach[3], L.-F. Artigas[1]**

[1] {*Université Lille Nord de France* – CNRS UMR 8187 Laboratoire d'Océanologie et de Géosciences, Université du Littoral Côte d'Opale, MREN, 32 Avenue Foch, 62930 Wimereux, France}

[*] {now at : *Aix Marseille Université,* CNRS/INSU, IRD, Mediterranean Institute of Oceanography (MIO), UM 110, *13288 Marseille* }

[2] {IFREMER LER, 150 quai Gambetta, 62200, Boulogne sur Mer, France}

[3] {The Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, NR33 0HT Lowestoft, United Kingdom}

[4] {RWS Centre for Water Management. Laboratory for hydrobiological analysis. Zuiderwagenplein 2, 8224 AD Lelystad, The Netherlands}

*Corresponding author: M. Thyssen (melilotus.thyssen@mio.osupytheas.fr)

Key words: Plankton functional type, automated scanning flow cytometry, PHYSAT, North Sea, mapping.

## Abstract

Phytoplankton observation in the ocean can be a challenge in oceanography. Accurate estimations of its biomass and dynamics will help to understand ocean ecosystems and refine global climate models. Relevant datasets of phytoplankton defined at a functional level and on a daily and sub meso scale are thus required. In order to achieve this, an automated, high frequency, dedicated scanning flow cytometer (SFC, Cytobuoy, NL), has been developed to cover the entire size range of phytoplankton cells whilst simultaneously taking pictures of the largest of them. This cytometer was directly connected to the water inlet of a pocket Ferry Box during a cruise in the North Sea, 8-12 May 2011 (DYMAPHY project, INTERREG IV A "2 Seas"), in order to identify the phytoplankton community structure of near surface waters (6 m) with a high resolution spacial basis (2.2 ± 1.8 km). Ten groups of cells, distinguished on the basis of their optical pulse shapes, were described (abundance, size estimate, red fluorescence per unit volume). Abundances varied depending on the hydrological status of the traversed waters, reflecting different stages of the North Sea blooming period. Comparisons between several techniques analyzing chlorophyll a and the scanning flow cytometer, using the integrated red fluorescence emitted by each counted cell, showed significant correlations. For the first, time, the community structure observed from the automated flow cytometry dataset was compared with classical PHYSAT reflectance anomalies over a daily scale. The number of matchups observed between the SFC automated high frequency *in situ* sampling and the remote sensing was found to be two to three times better than when using traditional water sampling strategies. Significant differences in the phytoplankton community structure within the two days for which matchups were available suggest that it is possible to label PHYSAT anomalies using automated flow cytometry to resolve not only dominant groups, but community structure.

## 1. Introduction

Phytoplankton plays a major role in marine ecosystems as the most important primary producer in the ocean (Field *et al.* 1998). Phytoplankton is involved in the long-term trapping of atmospheric carbon and its role in carbon transfer from the upper ocean layers to deep waters highlight its influence on climate (Boyce *et al.* 2010; Marinov *et al.* 2010). Beyond its role in the carbon cycle, phytoplankton also plays a major role in modifying the biogeochemical properties of water masses by converting most of the inorganic matter into available organic matter (nitrogen, phosphate, silicate, sulfur, iron); and determining the structure of the trophic status of marine environments. Given this importance, it is insufficient to use a single proxy, such as chlorophyll *a* measurements, for quantifying and qualifying phytoplankton over large scales when attempting to understand its role in biogeochemical processes (Colin *et al.* 2004). Such proxy does not reflect changes in community structure (Hirata *et al.* 2011) and do not yield robust biomass estimations (Kruskopf and Flynn 2006). Yet this classical proxy is frequently used to study the spatial and temporal variability of phytoplankton from both remotely sensed and *in situ* measurements. LeQuéré (LeQuéré *et al.* 2005) pointed out the importance of taking into account the functionality of phytoplankton species when considering the influence of phytoplankton community structure on biogeochemical processes. This functionality concept (i.e. Phytoplankton Functional Types, PFT) is described as set of species sharing similar properties or responses in relation to the main biogeochemical processes such as the N, P, Si, C and S cycles (diazotrophs for N cycle such as Cyanobacteria, DMSP producers for S cycle such as *Phaeocystis*, silicifiers for Si cycle such as Diatoms, calcifiers for C cycle such as Coccolithophorids, size classes, motility, food web structure mainly used for C cycle).

Representative data sets of phytoplankton functional types, size classes and specific chlorophyll *a* concentrations are the subject of active research using high frequency *in situ* dedicated analysis from automated devices such as spectral fluorometers, particle scattering and absorption spectra recording instruments, or automated and remotely controlled scanning flow cytometry (SFC). Among the high frequency *in situ* techniques used to quantify phytoplankton abundance, community structure and dynamics, SFC is the most advanced instrument, counting and recording cell optical properties at the single cell level. This technology has recently been adapted for the analysis of almost all the phytoplankton size classes and focuses on the resolution of phytoplankton community structure dynamics (Dubelaar *et al.* 1999; Olson *et al.* 2003; Sosik *et al.* 2003; Thyssen *et al.* 2008a; Thyssen *et al.* 2008b). In parallel, remote sensed algorithms have been developed which are dedicated to characterizing phytoplankton groups, PFTs or size classes (Sathyendranath *et al.* 2004; Ciotti *et al.* 2006; Nair *et al.* 2008; Aiken *et al.* 2008; Kostadinov *et al.* 2010; Uitz *et al.* 2010; Moisan *et al.* 2012). One of these algorithms, PHYSAT, has provided a description of the dominant phytoplankton functional types (LeQuéré *et al*., 2005) for open waters on a global scale, leading to various studies concerning the PFT variability (Alvain *et al.* 2005; Alvain *et al.* 2013; Masotti *et al.* 2011; Demarcq *et al.* 2011, Navarro *et al.*, 2014). PHYSAT relies on the identification of water-leaving radiance spectra anomalies, empirically associated with the presence of specific phytoplankton groups in the surface water. The anomalies were labeled thanks to the comparison with high pressure liquid chromatography biomarker pigment match ups. To date, six dominant phytoplankton functional groups in open waters (Diatoms, Nanoeucaryotes, *Prochlorococcus*, *Synecochoccus*, *Phaeocystis*-like, Coccolithophorids) have been found to be significantly related to specific water-leaving radiance anomalies from SeaWifs (Sea-viewing Wide Field-of-view Sensor) sensor measurements at a resolution of 9 km (Alvain *et al.* 2008). These relationships have been verified by theoretical optical models

(Alvain *et al.* 2012). This theoretical study also showed that additional groups or assemblages could be added in the future, once accurate *in situ* observations are available. Describing the community structure on a regional scale will give better quantification and understanding of the phytoplankton responses to environmental change and consequently, support the modification of theoretical considerations regarding energy fluxes across trophic levels. It is critical in understanding community structure interactions and particularly when it is necessary to take into account the meso-scale structure in a specific area (D'Ovidio *et al.* 2010), which is the case in areas under the influence of regional physical forcing such as the English Channel and the North Sea. Long-term changes detected in these regions have been shown to impact local ecosystem functioning by inducing, for instance, a shift in the timing of the spring bloom (Wiltshire and Manly 2004, Sharples *et al.* 2009; Vargas *et al.* 2009; Racault *et al.* 2013) or specific migrations of regional (Gomez and Souissi 2007) or dominant phytoplankton groups (Widdicombe *et al.* 2010). In addition, hydrodynamic conditions have been shown to play a strong role in the phytoplankton distribution on a regional scale (Gailhard *et al.* 2002; Leterme *et al.* 2008). It is therefore crucial to develop specific approaches to characterize the phytoplankton community structure (beyond global-scale dominance) and its high frequency variation in time and space. In order to achieve this, large data sets of *in situ* analyses resolving PFT are essential for specific calibration and validation of regional remote sensing algorithms such as PHYSAT. Flow-through surface water properties analysis for remote sensing calibration optimizes the amount of matchups (Werdell et al., 2010; Chase *et al.*, 2013). For the purpose of collecting high resolution *in situ* phytoplankton community structure, automated SFC technology allows samples to be collected at high frequency, resolving hourly and km scales in a totally automated system. The instrument enables single cell analysis of phytoplankton from 1 to 800 µm and several mm in length for chain forming cells and automated sampling allows large space and time domains to be covered at a high resolution (Sosik *et al.* 2003; Thyssen *et al.* 2008b; Thyssen *et al.* 2009; Ribalet *et al.* 2010).

Based on this approach, a high frequency study of the phytoplankton community structure in the North Sea was conducted. The *in situ* observations from SFC have been used for the first time and as a first trial to label PHYSAT anomalies detected during the sampling period. Thus, the available dataset makes it possible to distinguish between different water-leaving radiance anomaly signatures in which significant distinct phytoplankton community structures can be described, rather than just the dominant communities, as it is the case of previous studies. Our results are an improvement over conventional approaches as they allow the distribution of phytoplankton community structure to be characterized at a high resolution, from both *in situ* and day-to-day water-leaving radiance anomaly maps specific to the study area.

## 2. Materials and Methods

Samples were collected during the PROTOOL/DYMAPHY-project cruise onboard the RV Cefas Endeavour from the 8 to 12 May 2011 in the south-west region of the North Sea (Figure 1). Automated coupled sampling using a Pocket FerryBox (PFB) and a Cytosense scanning flow cytometer (SFC, Cytobuoy, b.v.) started on the 8 May at 9:00 UTC and ended on the 12 May at 4:00 UTC. Water was continuously collected from a depth of 6 m and entered the PFB at a pressure of 1 bar maximum. Sub-surface discrete samples were collected using Niskin bottles on a rosette and analyzed using a second Cytosense SFC (Stations 4, 6 and 13 were used in this paper, Figure 1).

## 2.1. Phytoplankton community structure from automated SFC

Phytoplankton abundance and group description were determined by using two Cytosense SFCs (Cytobuoy, b.v.). These instruments are dedicated to phytoplankton single cell recording, enabling cells from 1 µm to 800 µm and several mm in length to be analysed routinely in 1-10 $cm^3$. For the automated measurements, samples for SFC were automatically collected from a 450 $cm^3$ sampling unit where water from the continuous flow was periodically stabilized. This sampling unit was designed to collect bypass water from the 1 bar PFB inlet. The sampling unit water was replaced within a minute. One of the Cytosenses was directly connected to the sampling unit and two successive analyses with two distinct protocols were scheduled automatically every 10 min. A calibrated peristaltic pump was used to estimate the analysed volumes. Suspended particles were then separated using a laminar flow and subsequently crossed a laser beam (Coherent, 488 nm, 20 mV). The instrument recorded the pulse shapes of forward scatter (FWS) and side ward scatter (SWS) signals as well as red, orange and yellow fluorescence (FLR, FLO, FLY respectively) signals for each chain or single cell. The Cytosense instrument was equipped with two sets of photomultiplier tubes (high sensitivity and low sensitivity modes), resolving a wider range of optical signals from small (~<10 µm) to large particles (~<800 µm). Two trigger levels were applied to discriminate highly concentrated eukaryotic picophytoplankton and cyanobacteria (trigger level: FLR 10 mV; acquisition time: 180 s; sample flow rate: 4.5 $mm^3.s^{-1}$), from lower concentrated nano- and microphytoplankton (trigger level: FLR 25 mV, acquisition time: 400 s; sample flow rate: 9 $mm^3.s^{-1}$). Setting the trigger on red fluorescence was preferred to the commonly FWS or SWS triggering as a tradeoff between representative phytoplankton data sets and non-fluorescing particles/noise recording, but this procedure affected the SWS and FWS pulse shapes to some extent. To ensure good control and calibration of the instrument settings, a set of spherical beads with different diameters was analysed daily. This allowed the definition of estimated-size calibration-curves between Total FWS (in arbitrary units) and actual bead size. This set of beads included 1, 6, 20, 45, 90 µm yellow green fluorescence from Polyscience Fluoresbrite microspheres, 10 µm orange fluorescence Invitrogen polystyrene Fluorosphere, and 3 µm 488 nm Cyto-cal™Alignment standards. To correct for the high refraction index of polystyrene beads that generates an underestimation of cell size, we defined a correcting factor by using 1.5 µm silica beads (Polyscience, Silica microspheres) (Foladori *et al.* 2008). The phytoplankton community was described using several two-dimensional cytograms built with the Cytoclus® software. For each autofluorescing phytoplankton cell analysed, the integrated value of FLR pulse shape (Total red fluorescence TFLR, a.u.) was calculated. For each phytoplankton cluster, the amount of TFLR is reported per unit volume (TFLR.cm$^{-3}$, a.u..cm$^{-3}$). The TFLR.cm$^{-3}$ of each resolved phytoplankton cluster was summed (Total TFLR.cm$^{-3}$) and was used as a proxy for chlorophyll *a* concentration (µg.dm$^{-3}$). The TFLR signal was corrected from high sensitivity PMT saturation signal in the case of highly fluorescing cells (> 4000 mV) by using the low sensitivity PMTs that behaved linearly below this value with the high sensitivity PMT.

Discrete samples were collected during the cruise and analyzed using a second Cytosense SFC equipped with the Image in Flow system. The samples were analysed using settings similar to those of the Cytosense coupled to the PFB and pictures were randomly collected for the largest particles until the predetermined number of pictures was reached.

## 2.2. Temperature and Salinity

The PFB (4H-JENA©) was fixed on the wet laboratory bench, close to the Cytosense, in order to share the same water inlet. This instrument recorded temperature and conductivity (from which salinity was computed) from the clean water supplied by the ships seawater pumping system at a frequency of one sample every minute.

Within the PFB dataset, only data related to automated SFC analyses were selected for plotting temperature – salinity diagrams.

## 2.3. Chlorophyll a

Samples for High Pressure Liquid Chromatography (HPLC) analyses and bench top fluorometry (Turner® fluorometer) were collected randomly within 6 hour periods before or after the supposed MODIS (Moderate Resolution Imaging Spectroradiometer) sensor passage (12:30 pm UTC) to fulfill classical requirements in terms of *in situ* and remotely sensed matchup criteria. Samples were collected from the outlet of the PFB, filtered onto GF/F filters and stored directly in a -80°C freezer. The HPLC analyses were run on an Agilent Technologie, 1200 series. Pigments were extracted using 3 $cm^3$ ethanol containing vitamin E acetate as described by Claustre *et al.* (2004) and adapted by Van Heukelem and Thomas (2001). For bench top fluorometry, the filters were subsequently extracted in 90% acetone. Chlorophyll *a* (chla) concentration was evaluated by fluorometry using a Turner Designs Model 10-AU fluorometer (Yentsch and Menzel 1963). The fluorescence was measured before and after acidification with HCl (Lorenzen 1966). The fluorometer was calibrated using known concentrations of commercially purified chla (Sigma-Aldrich®).

The PFB was equipped with a multiple fixed wavelength spectral fluorometer (AOA fluorometer, bbe©) sampling once every minute to obtain chla values

MODIS chla values were extracted from daily level 2 product determining with a 4 km resolution (L3 Binned data).

## 2.4. Mixed layer depth

Daily water column temperature mapping was obtained from the Forecasting Ocean Assimilation Model 7 km Atlantic Margin model (FOAM AMM7), available at MyOcean data base (http://www.myocean.eu.org/). Model output temperature depths were as follows: 0, 3, 10, 15, 20, 30, 50, 75, 100, 125, 150 m. Average mixed layer depth (MLD) on the 5 sampling days was calculated from daily temperature datasets. MLD was defined as the absolute temperature difference of more than 0.2 °C from one depth to the surface (defined at 10 m, de Boyer Montégut *et al.* 2004).

### 2.5. Matching method between in situ and remote sensed observations for phytoplankton community structure

Remotely sensed observations were selected on the basis of quality criteria that ensured a high degree of confidence in PHYSAT as described in Alvain *et al.* (2005). Thus, pixels were only considered when clear sky conditions were found and when the optical thickness, a proxy of the atmospheric correction steps quality, was lower than 0.15. In addition, as the region of interest included some coastal areas, that are not considered as open waters for remote sensing, we have selected pixels according to their optical properties (Vantrepotte *et al.* 2012). Consequently, this avoided using waters rich in sediment which previously rendered it impossible to use the PHYSAT version. Waters classified as turbid were therefore excluded from the empirical relationship since the PHYSAT method is currently not available for such areas. Waters classified as non-turbid using the same criteria were selected and the PHYSAT algorithm applied. To link coincident *in situ* and remotely sensed observations, a match-up exercise was carried out. Matching points between *in situ* SFC samples (considered as *in situ* data) and 4 km resolution MODIS pixels were selected by comparing their concomitant position day after day. When more than one *in situ* SFC sample was found in a MODIS pixel the averaged value of TFLR (a.u..cm$^{-3}$) for each phytoplankton group was calculated. From the matching points, the PHYSAT method resulted in water leaving irradiance anomalies spectra (Ra) as described in Alvain *et al.* 2008 and 2012.

### 2.6. Statistics

Statistics were run under R software (CRAN, http://cran.r-project.org/). Before running correlation and comparison tests on the different *in situ* sensors (for chla and Total TFLR), the Shapiro normality test was run. When Normality was not applied, a Wilcoxon signed rank test was applied. Correlations between data were defined using Spearman's rank correlation coefficient.

As the PHYSAT approach is based on the link between specific Ra spectra (in terms of shapes and amplitudes) and specific phytoplankton composition, the set of remote sensed data was separated into distinct groups with similar Ra. The PHYSAT Ra found over the studied area and matching the *in situ* SFC samples was differentiated by applying a k-means clustering partitioning method (tested either around means (Everitt and Hothorn 2006) or around menoids (Kaufman and Rousseeuw 1990)). The appropriate number of clusters was decided with a plot of the within groups sum of squares by number of clusters extracted. A hierarchical clustering was computed to illustrate the k-means clustering method. Within each k-mean cluster, SFC-defined phytoplankton community was described and differences between TFLR.cm$^{-3}$ per phytoplankton group were compared within the different PHYSAT spectra clusters using the Wilcoxon signed rank test.

## 3. Results

### 3.1. Temperature, Salinity and Mixed layer depth

The sampling track crossed three North Sea marine zones: Western Humber, Tyne, Dogger, Eastern Humber and Thames (Fig. 1). The PFB measured temperature associated to the SFC samples and varied between 8.83 °C and 12.39°C with an average of 10.67 ± 0.72 °C. Minimal temperatures were found in the western Humber area (53-55°N and -1-1°E) and maximal temperatures were found in the Thames area (54-52°N, 2-4°E) (Fig. 2A). Salinity from the PFB ranged between 34.02 and 35.07 with an average value of 34.6 ± 0.26. Highest salinity values were found in the Dogger area above 55°N and in the limit between the Humber and the Thames areas, 53°N. Lowest salinity values were found in the Tyne area around 55°N, -1 °E and in the Thames area (by the Thames plume; Fig. 2B).

The mixed layer depth calculated from the FOAM AMM7 was used to illustrate the physical environment of the traversed water masses. Different mixed layer depth characterized the sampled area, with deeper MLD in the northern part (15 to 30 m) and a shallower MLD in the southern area (~10 m, Fig.1). A tongue of shallow MLD (~10 m) surrounded by deeper MLD (~20 m) crossed the sampling area at ~55°N and ~3°W.

### 3.2. Phytoplankton community from SFC analysis

A total of 247 SFC validated analysed samples were collected during this experiment. Average distance between samples collected with the automated SFC was of 2.2 ± 1.8 km when the system ran continuously. The sampling rate was 25 ± 45 min. Up to 10 phytoplankton clusters were resolved (Fig. 3) based on their optical fingerprints from SFC analysis. The 10 discriminated clusters were labeled as follows: PicoORG cluster (Fig. 3A), PicoRED cluster (Fig. 3A), NanoSWS cluster (Fig. 3B), NanoRED1 cluster (Fig. 3C), NanoRED2 cluster (Fig. 3B and 3C); Micro1 cluster (Fig. 3C and 3D), MicroLowORG (Fig. 3A), NanoORG and MicroORG clusters (Fig. 3D) and a cluster of large cells, Micro2 cluster (Fig 3D). Pictures were randomly collected (between 20 and 60 pictures per sample within the Micro2 cluster) and were used to illustrate the most frequently encountered class (Fig. 4). Station 4 (Fig. 4A) sampled at 12 m, showed mostly a mixture of dinoflagellate-like cells (25 pictures collected within 47 counted cells). Station 6 (Fig. 4B) sampled at 7 m, showed pictures composed mainly of diatoms (*Thalassiosira* and *Chaetoceros,* 11 images collected among 28 counted cells). Station 13 (Fig. 4C) sampled at 7 m, gave a mixture of diatoms and dinoflagellates (58 pictures shot among the 99 counted cells corresponding to the Micro2 cluster: 5 *Chaetoceros,* 30 *Rhizosolenia,* 10 Dinoflagellates, one flagellate and several unidentified cells).

Cell abundance, average cell size and TLFR.cm$^{-3}$ for each cluster are illustrated on Figures 5, 6 and 7 respectively. Average abundance and sizes of each cluster are addressed in Table 1. PicoRED cells were on average, the most abundant in the studied area (Fig. 5B and Table 1) followed by NanoRED2, PicoORG, NanoRED1 and Micro1 (Fig. 5F, 5A, 5C and 5G respectively, Table 1). The other cluster's abundances were below 1.10$^{2}$ cells.cm$^{-3}$ on average (Fig. 5D, E, H, I, J; Table 1). PicoORG cells were the smallest estimated (Fig. 6A, Table 1), while the largest estimated were MicroORG, MicroLowORG and Micro2 cells (Fig. 6H, 6I and 6J respectively, Table 1).

The western Humber zone (Fig.1) was marked by the highest abundances of PicoRED, PicoORG, MicroORG, MicroLowORG and Micro1 (Fig. 5B, 5A, 5H, 5I and 5G). The eastern part of the Humber zone (Fig.1) was marked by the highest abundances of NanoRED1 and Micro1 (as for the western part) (Fig. 5C, 5G). High values of PicoRED were also observed in this part of the Humber zone. The Tyne zone (Fig.1) had the highest abundance of NanoORG and Micro2 clusters (Fig. 5D, 5J), and the lowest abundance of PicoRED and NanoSWS. High abundance values of MicroORG were also observed (Fig. 5H). The size of the NanoSWS and the NanoRED2 were the greatest in this zone (Fig. 6E, 6F). The Dogger zone (Fig.1) was dominated in terms of abundance by the PicoRED and the PicoORG, where the sizes were the smallest (Fig. 6B and 6A) but did not show the highest abundance values. The cell sizes of Micro1 were the greatest in this zone (Fig. 6G). Observations in the Thames zone

(Fig.1) produced the maximal abundance of NanoSWS and NanoRED2 (Fig. 6E, 6F). Sizes were the greatest for PicoORG, NanoRED1 and NanoSWS (together with the Tyne zone; Fig. 6A, 6C, 6E). TFLR follows similar trends to abundance (Fig. 7).

### 3.3. Comparison between scanning flow cytometry, Total Red Fluorescence and chlorophyll a analysis

Several bench top and *in situ* instruments, i.e. HPLC, Turner fluorometer and the PFB AOA fluorometer, were used to give either exact and/or proxy values of chla. Similarly to temperature and salinity, the PFB AOA fluorometer samples were selected to match SFC samples. Overall values of chla originating from these instruments were superimposed to the Total TFLR.cm$^{-3}$ (by summing up the TFLR.cm$^{-3}$ values of the observed cluster) and the MODIS chla values matching the points on Figure 8. HPLC values varied between 0.21 and 7.58 µg.dm$^{-3}$ with an average of 1.57 ± 2.01 µg.dm$^{-3}$. Turner fluorometer values varied between 0.41 and 2.31 with an average of 1.24 ± 0.7 µg.dm$^{-3}$. AOA fluorometer values varied between 0.73 and 28.53 µg.dm$^{-3}$ with an average of 4.44 ± 5.54 µg.dm$^{-3}$. The Total TFLR.cm$^{-3}$ from SFC, normalized with 3 µm bead red fluorescence varied between 5011 and 399200 a.u..cm$^{-3}$ with an average value of 64394.5 ± 67488.4 a.u..cm$^{-3}$. The Shapiro normality test showed non normality for each of the variables so a Wilcox test was run between techniques involving similar units. HPLC and Turner chla concentrations were significantly not different (n=9, p=0.65) and the correlation was significant (Spearman, r=0.98, Table 2). The absolute values from both techniques were significantly different from the AOA fluorometer values (n=9, p<0.001 for both) but were significantly correlated (Spearman, r=0.86 and r=0.82 for HPLC and Turner fluorometer respectively, Table 2). The SFC Total TFLR (a.u..cm$^{-3}$) summing up the TFLR of all the phytoplankton groups was used for comparison with other chla determinations. Correlations with the AOA fluorometer, the HPLC and the Turner fluorometer results were all significant as shown in Table 2.

### 3.4. PHYSAT anomalies and SFC phytoplankton community composition, extrapolation to the non-turbid classified waters in the North Sea

Considering our database of coincident SFC *in situ* and MODIS remote sensed observations, a total of 56 matching points were identified, from which only 38 points corresponded to non-turbid classified waters. Matching points between *in situ* sampling and remote sensing pixels for the purpose of the PHYSAT empirical calibration were selected in the daytime period 6 - 18 h, the limit of the correlation significance between MODIS chla and the AOA fluorometer chla within the SFC dataset ($r=0.49$, $p=0.06$, $n=15$, Spearman rank test), leaving 15 SFC matching points (Fig. 1 and Fig. 8). The chla values found in the matching points were lower than 0.5 µg.dm$^{-3}$ (Fig. 8).

PHYSAT radiance anomalies (Ra) were calculated based on the 2005 method (Alvain *et al.*, 2005) and the average signal was recalculated to fit the sampling area. The Ra were separated into two distinct anomalies using the within sum of square minimization (Fig. 9A) and illustrated on a dendrogram (Fig. 9B). These two distinct types of anomalies in terms of shape and amplitude are illustrated in Figure 9C and 9D and the anomaly characteristics are summarized on Table 3. The first anomaly set (N1, Table 3) was composed of 5 spectra that had overall higher values than the second anomaly set (N2, Table 3), composed of the other 10 spectra. The corresponding SFC cluster proportion of TFLR.cm$^{-3}$ to the overall Total TFLR.cm$^{-3}$ found within the two anomalies are illustrated in Figures 10 A and B. Similarly, the relative difference of each phytoplankton cluster’s TFLR.cm$^{-3}$ within the two anomalies to its overall TFLR.cm$^{-3}$ median value are illustrated in Figures 10 C and D. Considering our previous analyses, N1 and N2 community structures were dominated by NanoRED2 TFLR.cm$^{-3}$ (Fig. 10A and 10B). Regarding each distinct cluster relative difference to its overall median value, samples corresponding to N1 anomalies had significantly higher NanoRED1 TFLR.cm$^{-3}$, higher NanoORG TFLR.cm$^{-3}$ and higher MicroORG TFLR.cm$^{-3}$; while the samples corresponding to N2 anomalies had only higher PicoRED TFLR.cm$^{-3}$ (Wilcox rank test, N1, n=5; N2, n=10, Fig. 10C and 10D). Temperature, salinity, MODIS chla and SFC Total TFLR.cm$^{-3}$ found in each *in situ* sample corresponding to both sets of anomalies are illustrated in Fig. 11. Samples found in the N1 pixels were significantly warmer (11.3 ± 0.32°C in N1 and 10.94 ± 0.23°C in N2, $p<0.1$, Wilcox rank test, Fig. 11A), not significantly different in terms of salinity, although N1 waters were less salty (Fig. 11B), significantly richer in chla ($p<0.01$, Wilcox rank test, Fig. 11C), but not significantly different in Total TFLR.cm$^{-3}$ values (Fig. 11D).

Considering the specificity of each set of Ra in terms of phytoplankton and environmental conditions, it’s interesting to map their frequency of detection in our area of interest. A pixel is associated to an anomaly when the Ra values at each wavelength fulfilled the criteria of Table 3. The frequencies of occurrence over the sampling period based on a synthesis overlapping the sampling period are illustrated in Fig. 12A and 12B. Pixels corresponding to N1 anomaly were mostly found in the 54-56°N area (Dogger and German, Fig. 1), following the edge between the shallow MLD tong and the deepest MLD zones (Fig. 1), but also near the Northern Scottish coast (Forth, Forties and Cromarty, Fig. 12A), where MLD was shallow (Fig. 1). The N2 anomaly pixels were mostly found in the Forties, Fisher and German area, on much smaller surfaces (Fig. 12B).

## 4. Discussion

The automated SFC used during this study resolves the spacial/temporal issue by its high frequency sampling, reaching sub mesoscale distribution and diel changes in abundances. However, water mass dynamics generates patchiness which modifies phytoplankton community structure and makes it difficult to follow a population over time and at a basin scale. In this context, the daily extrapolation of the community structure using PFT daily remote sensing mapping can help to follow spatial distribution of phytoplankton communities. The improvement of PFT mapping, i.e. from dominant groups to the community structure resolution, is one of the ideas generated in this paper. This paper shows for the first time that SFC datasets can be used for labeling PHYSAT anomalies at the daily scale. The SCF is a powerful automated system aimed to be implemented in several vessels of opportunity and monitoring programs for future PHYSAT anomalies identification at the daily scale and at the community structure level. A recent publication that enables the classification of a large range of anomaly spectra (Ben Mustapha *et al*., 2014) should help to make this easier. Thus, the knowledge and the tools are available, which augurs well for understanding phytoplankton heterogeneity and variability over high frequency spatio-temporal scales.

Indeed, resolving phytoplankton community structure over the sub meso scale and hourly scale is a good way to understand the influence of environmental short scale events (Thyssen et al., 2008a; Lomas *et al.* 2009), seasonal (or not) succession schemes, resilience capacities of the community after environmental changes and impacts on the specific growth rates (Sosik *et al.* 2003, Dugenne et al., 2014). Resolving the community structure and the causes of variations at several temporal and spatial scales has great importance in further understanding the phytoplankton functional role in biogeochemical processes. This scale information is currently lacking for the global integration of phytoplankton in biogeochemical models, mainly due to the lack of adequate technology which are needed to integrate the different levels of complexity linked to phytoplankton community structure.

Phytoplankton community structure from automated SFC is described through clusters of analyzed particles sharing similar optical properties. Thus cluster identification at the species level is speculative and, as any cytometric optical signature, it needs a sorting and genetic or microscopic analysis to be resolved at the taxonomical level. This deep level of phytoplankton diversity resolution requirement is although not needed in biogeochemical processes studies in which functionality is preferred to taxonomy (LeQuéré et al., 2005). In this context, most of the optical clusters could be described at the plankton functional type level because of some singular similarities combining abundance, size, pigments and structure proxies obtained from optical SFC variables (Chisholm *et al.* 1988; Veldhuis and Kraay 2000; Rutten *et al.* 2005; Zubkov and Burkill 2006). The Cytobuoy instrument used in this study was developed to identify phytoplankton cells from picophytoplankton up to large microphytoplankton with complex shapes, even those forming chains. Indeed, the volume analyzed was close to 3 cm$^3$, giving accurate counts of clusters with abundances as low as 30 cells.cm$^{-3}$ (100 cells counted), under which, coefficient of variation exceeds 10% (Thyssen et al., 2008a). Such low abundances were found for some of the clusters identified in this study (NanoORG, MicroORG and Micro2 clusters for which the median abundance value was close to 30 cells.cm$^{-3}$), in agreement with concentrations observed in previous studies for the possible related taxonomical phytoplankton genus, as discussed below, i.e. cryptophytes (Buma *et al.* 1992), diatoms and dinoflagellates (Leterme *et al.* 2006). Previous comparisons between bench top flow cytometry and remote sensing (Zubkov and Quartly, 2003) could technically not include the entire size range of nano-microphytoplankton. The Cytobuoy SFC resolves cells up to 800 µm in theory, but this depends on the counted cells in the volume sampled (which is approximately ten times more than classical flow cytometry). However, the largest part of phytoplankton production in the North Sea is driven by cells < 20 µm (Nielsen *et al.* 1993), and we can consider this size class to be correctly counted with the SFC.

PicoORG cells could be labeled *Synechococcus* (Waterbury *et al.* 1979*;* Li1994) based on their phycoerythrin pigment fluorescence (Fig. 3A), their size estimated between 0.8 and 1.2 µm (Fig. 6A) and their abundances around $10^{2}$ - $10^{4}$ cells.$cm^{-3}$ (Fig. 5A). PicoRED cells could be autotrophic eukaryotic picoplankton, as their cell size varied between 1-3 µm (Fig. 6B) and contained chla as their main pigment. Thus, PicoORG and PicoRED clusters contained the smallest cells found above the so called non-fluorescing/electronical noise background of this instrument (Fig. 3A and 3B). As *Prochloroccocus* is expected to be absent in these waters, we can conclude that the cytometer observed most of the phytoplankton size classes when sufficiently concentrated in the analysed volume. NanoRED1 cells exhibited abundance and sizes close to those of *Phaeocystis* haploid flagellate cells (3-6 µm, Fig. 6C, Rousseau *et al.* 2007 and references therein). Their presence, found mostly in the Humber (Fig. 5C), suggests that this area corresponded to a period between the inter-bloom and the start of the *Phaeocystis* bloom (Rousseau et al., 2007). Similarly, NanoRED2 could be referred to as *Phaeocystis* diploid flagellates or free colonial cells, based on their size and abundance (4-8 µm and 0-50.$10^{3}$ cells.$cm^{3}$ (Fig. 6F and 5F respectively), Rousseau et al., 2007). Their maximal abundance was found in the southern North Sea Thames area. This abundance fell when *Phaeocystis* was blooming (Guiselin 2010).

MicroORG cells, whose abundance and size are close to those of some large cryptophytes cells, were found in the same areas as NanoORG cells (Fig. 5H and 5D respectively), which are related to smaller *Cryptophyceae* cells. MicroLowORG cells with sizes close to that of MicroORG cells and although low in concentration, emitted orange fluorescence and could represent cells with little phycoerythrin content. NanoSWS cluster was composed of high SWS diffusing cells that are consistent with the signature of

*Coccolithophorideae* cells (van Bleijswijk *et al.* 1994; Burkill et al., 2002). The observed abundances did fit with the low *Coccolithophorideae* concentrations observed in the southern North Sea (Houghton, 1991).

The Micro1 cluster could correspond to small nanoplanktonic diatom cells (~10-30 µm, Fig. 6G). Regarding the size range, this cluster could represent several species. They were mainly found within the Humber area. The Micro2 cluster was mostly composed of large diatoms (*Rhizosolenia*, *Chaetoceros*) and dinoflagellates (Fig. 4) within the size range of 40 - 100 µm (Fig. 6J) as observed in the pictures (Fig. 4). The presence of these groups illustrates the boundary between the end of the diatom bloom and the development of a dinoflagellate bloom, from which it could be possible to make a link with the *Dinophysis norvegica* and *Alexandrium* early summer bloom, observed in the Tyne region by Dodge (Dodge 1977). This is in agreement with the stratification observed within the Thames zone (Fig.1).

The data sets from the spatial (km) and the temporal (hourly) scales for phytoplankton community structure based on single cell optical properties are important for validating the methods for describing phytoplankton community structure from space. Ocean algorithms need specific information on water properties and phytoplankton structure and are dependent on validation from *in situ* observations, always complex to collect and limited by sky condition criteria. The PHYSAT method was built on an empirical relationship between dominant phytoplankton functional types from *in situ* HPLC analysis and Ra. The method was thus limited to dominance cases only as HPLC analysis can’t give us more information**.** The remote sensing synoptic extrapolation concerning phytoplankton community structure remains to be established and in spite of a theoretical validation (Alvain *et al.,* 2012), still depends on important *in situ* data point collection in order to build robust empirical relationships. In this study, the combination of phytoplankton high frequency analysis from an automated SFC with the PHYSAT method proved to be an excellent calibration by giving an unprecedented amount of matching points for only two significant sampling days (number of analyzed samples for non-turbid waters matching MODIS pixels: 38, number of used samples between 6 and 18h: 15, corresponding to 39.5 % profitability), compared to the 14% matching points from the GeP&CO dataset (Alvain *et al.*, 2005).

The combination of SFC and PHYSAT has shown that a first set of specific anomalies (N1) can be associated with NanoRED1, NanoORG and MicroORG, which contributed more to the Total TFLR.cm$^{-3}$ (a proxy of chla, Fig. 7, Table 2) than in the second set of anomaly (N2), in which PicoRED cells contributed significantly more to the Total TFLR.cm$^{-3}$, but also, where Micro1 contribution to Total TFLR.cm$^{-3}$ was above its overall median value observed along the matching points (Fig. 10D). Spatial successions between diatoms (as could be found in the NanoRED1 and Micro1 clusters) and cryptophytes (corresponding to the NanoORG and MicroORG specific signatures) revealed differences in stratification, lower salinity and shallower MLD (Moline *et al.* 2004; Mendes *et al.* 2013). Indeed, the N1 anomaly corresponds to areas of low MLD (Fig. 1) following the main North Sea current from the south west to the north east (Holligan *et al.* 1989), surrounding the Dogger bank. This anomaly was also found on the north-western part of the northern North Sea, following the Scottish coastal water current with a shallow MLD (Fig 1 and Fig 11A). The N2 anomaly was observed with the deeper MLD of the Forties, Fisher and German areas (Fig. 1 and 11B). These N2 areas corresponded to a phytoplankton community still blooming while the N1 anomaly areas might be at a stage of late blooming, in which conditions fit cryptophyceae development and grazing (cells of *Myrionecta rubra* were observed when using the Image in Flow, not shown). These organisms were found dominating the areas surrounding the Dogger bank from observations and counts carried out by Nielsen *et al.* (1993) during the same period.

In conclusion, the use of automated SFC Cytosense technology is an area of great interest when coupled with remote sensing algorithms in the study of surface phytoplankton distribution. Further advances in understanding the link between the phytoplankton community composition and distribution, with radiance anomalies are expected from improvements in analyzing larger volumes by automated SFC and by substantially increasing the number of coincidences between remote sensing and *in situ* observations.

## Acknowledgement

This study was funded by the DYMPAHY (Development of a DYnamic observation system for the assessment of MArine water quality, based on PHYtoplankton analysis) INTERREG IVA “2 Mers Seas Zeeën” European cross-border project, co-funded by the European Regional Development Fund (ERDF) and French (ULCO-CNRS-UL1), English (Cefas) and Dutch (RWS) partners. We thank the captain and crew of the RV Cefas “Endeavour”, as well as Anne-Hélène Rêve for chlorophyll a bench top analysis. We also thank Dr. Rodney Forster for his invitation onboard the ship during the EU FP7 PROTOOL (Automated Tools to Measure Primary Productivity in European Seas) cruise. We are also grateful to our funding sources, the CNRS, the CNES-TOSCA/PHYTOCOT project. The authors thank NASA/GSFC/DAAC for providing access to daily L3 MODIS binned products.

## References

Aiken, J., N. J. Hardman-Mountford, R. Barlow, J. Fishwick, T. Hirata and T. Smyth (2008). Functional links between bioenergetics and bio-optical traits of phytoplankton taxonomic groups: an overarching hypothesis with applications for ocean colour remote sensing. Journal of Plankton Research 30(2): 165-181.

Alvain, S., C. Le Quéré, L. Bopp, M. F. Racault, G. Beaugrand, D. Dessailly and E. Buitenhuis (2013). Rapid climatic driven shifts of diatoms at high latitudes. Remote Sensing of Environment 132: 195-201.

Alvain, S., H. Loisel and D. Dessailly (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. Optics Express 20(2): 1070-1083.

Alvain, S., C. Moulin, Y. Dandonneau and F. M. Bréon (2005).Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. Deep Sea Research Part I: Oceanographic Research Papers 52(11): 1989-2004.

Alvain, S., C. Moulin, Y. Dandonneau and H. Loisel (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean : A satellite view. Global Biogeochemical Cycles 22: GB3001.

Ben Mustapha Z., Alvain S.**,** Jamet C, Loisel H. and D. Dessailly. Automatic classification of water leaving radiance anomalies from global SeaWifS imagery : Application to the detection of phytoplankton groups in open ocean waters. Remote Sensing of Environment RSE-08794, 2014.

Boyce, D. G., M. R. Lewis and B. Worm (2010).Global phytoplankton decline over the past century. Nature 466(7306): 591-596.

Buma, A. G. J., W. W. C. Gieskes and H. A. Thomsen (1992).Abundance of cryptophyceae and chlorophyll b-containing organisms in the Weddell-Scotia Confluence area in the spring of 1988. Polar Biology 12(1): 43-52.

Burkill, P.H., Archer, S.D., Robinson, C., Nightingale, P.D., Groom, S.B., Tarran, G.A., and Zubkov, M.V. 2002. Dimethyl sulphide biogeochemistry within a coccolithophore bloom (DISCO): an overview. Deep-Sea Res. II, 49: 2863–2885

Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., Dall’Olmo, G. and T.K. Westberry, (2013). Decomposition of in situ particulate absorption spectra. Methods in Oceanography, 7: 110-124.

Chisholm, S. W., R. J. Olson and C. M. Yentsch (1988). Flow cytometry in oceanography: Status and prospects. Eos, Transactions American Geophysical Union 69(18): 562-572.

Ciotti, A. and A. Bricaud (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by Colored Detrital Matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. Limnol. Oceanogr. Methods 4: 237–253.

Claustre, H., S. B. Hooker, L. Van Heukelem, J.-F.Berthon, R. Barlow, J. Ras, H. Sessions, C. Targa, C. S. Thomas, D. van der Linde and J.-C. Marty (2004). An intercomparison of HPLC phytoplankton pigment methods using in situ samples: application to remote sensing and database activities. Marine Chemistry 85(1-2): 41-61.

Colin, P., I. , C. Le Quéré, E. Buitenhuis, J. House, C. Klaas and W. Knorr (2004). Biosphere dynamics: challenges for Earth system models. The State of the Planet: Frontiers and Challenges,. C.J. Hawkesworth and R.S.J. Sparks (eds), American Geophysical Union

D’Ovidio, F., S. De Monte, S. Alvain, Y. Dandonneau and M. Levy (2010). Fluides dynamical niches of phytoplankton types. PNAS.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar and D. Iudicone (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. Journal of Geophysical Research: Oceans 109(C12): C12003.

Demarcq, H., G. Reygondeau, S. Alvain and V. Vantrepotte (2011).Monitoring marine phytoplankton seasonality from space. Remote Sensing of Environment 117: 211-222.

Dodge, J. D. (1977). The early summer bloom of dinoflagellates in the North Sea, with special reference to 1971. Marine Biology 40: 327-336.

Dubelaar, B. J., P. Gerritzen, A. E. R. Beeker, R. Jonker and K. Tangen (1999). Design and first results of Cytobuoy: a wireless flow cytometer for in situ analysis of marine and fresh waters. Cytometry 37: 247-254.

Dugenne, M., M. Thyssen, D. Nerini, C. Mante, J.-C. Poggiale, N. Garcia, F. Garcia and G. J. Gregori (2014). Consequence of a sudden wind event on the dynamics of a coastal phytoplankton community: an insight into specific population growth rates using a single cell high frequency approach. Frontiers in Microbiology 5: 485. doi: 10.3389/fmicb.2014.00485

Everitt, B. S. and T. Hothorn (2006).A Handbook of Statistical Analyses Using R, Chapman & Hall.

Field, C. B., M. J. Behrenfeld, J. T. Randerson and P. G. Falkowski (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281: 237-240.

Foladori, P., A. Quaranta and G. Ziglio (2008). Use of silica microspheres having refractive index similar to bacteria for conversion of flow cytometric forward light scatter into biovolume. Water Research 42(14): 3757-3766.

Gailhard, I., P. Gros, J. P. Durbec, B. Beliaeff, C. Belin, E. Nézan and P. Lassus (2002). Variability patterns of microphytoplankton communities along the French coasts. Marine Ecology Progress Series 242: 39-50.

Gomez, F. and S. Souissi (2007). Unusual diatoms linked to climatic events in the northeastern English Channel. Journal of Sea Research 58: 283-290.

Guiselin, N. (2010). Etude de la dynamique des communautés phytoplanctoniques par microscopie et cytométrie en flux, en eaux côtières de la Manche orientale. ULCO-MREN. Doctorate (Ph.D.) Thesis in Biological Oceanology, University of Littoral Côte d’Opale (ULCO), 190 pp.

Hirata, T., N. J. Hardman-Mountford, R. J. W. Brewin, J. Aiken, R. Barlow, K. Suzuki, T. Isada, E. Howell, T. Hashioka, M. Noguchi-Aita and Y. Yamanaka (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. Biogeosciences 8(2): 311-327.

Holligan, P. M., T. Aarup and S. B. Groom (1989). The North Sea: Satellite colour atlas. Continental Shelf Research 9(8): 667-765.

Houghton, S. D. (1991). Coccolith sedimentation and transport in the North Sea. Marine Geology 99(1-2): 267-274.

Kaufman, L. and P. J. Rousseeuw (1990). Finding Groups in Data: An Introduction to Cluster Analysis, Wiley-Interscience.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations, J. Geophys. Res., 114, C09015, doi:10.1029/2009JC005303.

Kruskopf, M. and K. J. Flynn (2006). Chlorophyll content and fluorescence responses cannot be used to gauge phytoplankton biomass, nutrient status or growth rate. New Phytologist 169: 525-536.

LeQuéré, C. L., S. P. Harrison, I. Colin Prentice, E. T. Buitenhuis, O. Aumont, L. Bopp, H. Claustre, L. Cotrim Da Cunha, R. Geider, X. Giraud, C. Klaas, K. E. Kohfeld, L. Legendre, M. Manizza, T. Platt, R. B. Rivkin, S. Sathyendranath, J. Uitz, A. J. Watson and D. Wolf-Gladrow (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Global Change Biology 11(11): 2016-2040.

Leterme, S., R. D. Pingree, M. D. Skogen, L. Seuront, P. C. Reid and M. J. Attrill (2008). Decadal fluctuations in North Atlantinc water inflow in the North Sea between 1958-2003: impact on temperature and phytoplankton populations. Oceanologia 50(1): 59-72.

Leterme, S. C., L. Seuront and M. Edwards (2006).Differential contribution of diatoms and dinoflagellates to phytoplankton biomass in the NE Atlantic Ocean and the North Sea. Marine Ecology-progress Series 312: 57-65.

Li, W. K. W. (1994). Primary production of prochlorophytes, cyanobacteria and eukaryotic ultraphyto-plankton: Measurements from flow cytometric sorting. Limnology and Oceanography 39: 169–175.

Lomas, M. W., N. Roberts, F. Lipschultz, J. W. Krause, D. M. Nelson and N. R. Bates (2009). Biogeochemical responses to late-winter storms in the Sargasso Sea. IV. Rapid succession of major phytoplankton groups. Deep Sea Research I 56: 892-909.

Lorenzen, C. J. (1966). A method for the continuous measurement of in vivo chlorophyll concentration. Deep Sea Research I 13: 223-227.

Marinov, I., S. C. Doney and I. D. Lima (2010). Response of ocean phytoplankton community structure to climate change over the 21st century: partitioning the effects of nutrients, temperature and light. Biogeosciences Discuss. 7(3): 4565-4606.

Masotti, I., C. Moulin, S. Alvain, L. Bopp and D. Antoine (2011). Large scale shifts in phytoplankton groups in the Equatorial Pacific during ENSO cycles. Biogeosciences 8: 539-550.

Mendes, C. R. B., V. M. Tavano, M. C. Leal, M. S. Souza, V. Brotas and C. A. E. Garcia (2013). Shifts in the dominance between diatoms and cryptophytes during three late summers in the Bransfield Strait (Antarctic Peninsula). Polar Biology 36(4): 537-547.

Moisan, T. A. H., S. Sathyendranath and H. A. Bouman (2012). Ocean Color Remote Sensing of Phytoplankton Functional Types, ISBN: 978-953-51-0313-4, InTech.

Moline, M. A., H. Claustre, T. K. Frazer, O. Schofield and M. Vernet (2004).Alteration of the food web along the Antarctic Peninsula in response to a regional warming trend. Global Change Biology 10(12): 1973-1980.Navarro, G., S. Alvain, V. Vantrepotte and I. E. Huertas (2014). Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. Remote Sensing of Environment 152(0): 557-575.

Navarro, G., S. Alvain, V. Vantrepotte and I. E. Huertas (2014). Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. Remote Sensing of Environment 152(0): 557-575.

Nair, A., S. Sathyendranath, T. Platt, J. Morales, V. Stuart, M.-H, N. Forget, E. Devred and H. Bouman (2008). Remote sensing of phytoplankton functional types. Remote Sensing of Environment 112(8): 3366-3375.

Nielsen, T. G., B. Lokkegaard, K. Richardson, F. Pedersen and L. Hansen (1993).Structure of plankton communities in the Dogger Bank area (North Sea) during a stratified situation.

Marine Ecology Progress Series 95: 115:131.

Olson, R. J., A. Shalapyonok and H. M. Sosik (2003). An automated flow cytometer for analyzing pico- and nanophytoplankton=FlowCytobot. Deep Sea Research Part I 50: 301-315.

Racault, M. F., C. Le Quéré, E. Buitenhuis, S. Sathyendranath and T. Platt (2013). Phytoplankton phenology in the global ocean. Ecological Indicators 14(1): 152-163.

Ribalet, F., A. Marchetti, K. A. Hubbard, K. Brown, C. A. Durkin, R. Morales, M. Robert, J. E. Swalwell, P. D. Tortell and E. V. Armbrust (2010). Unveiling a phytoplankton hotspot at a narrow boundary between coastal and offshore waters. Proceedings of the National Academy of Sciences 107(38): 16571-16576.

Rousseau, V., M.-J.Chrétiennot-Dinet, A. Jacobsen, P. Verity and S. Whipple (2007). The life cycle of Phaeocystis: state of knowledge and presumptive role in ecology. Biogeochemistry 83(1-3): 29-47.

Rutten, T. P. A., B. Sandee and A. R. T. Hofman (2005). Phytoplankton monitoring by high performance flow cytometry: A successful approach? Cytometry Part A 64A(1): 16-26.

Sharples, J., C. M. Moore, A. E. Hickman, P. M. Holligan, J. F. Tweddle, M. R. Palmer and J. H. Simpson (2009). Internal tidal mixing as a control on continental margin ecosystems. Geophysical Research Letters 36(23): L23603.

Sathyendranath, S., W. Louisa, D. Emmanuel, P. Trevor, C. Carla and M. Heidi (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. Marine Ecology Progress Series 272: 59-68.

Sosik, H. M., R. J. Olson, M. G. Neubert and A. Shalapyonok (2003). Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. Limnology and Oceanography 48(5): 1756-1765.

Thyssen, M., N. Garcia and M. Denis (2009). Sub meso scale phytoplankton distribution in the North East Atlantic surface waters determined with an automated flow cytometer. Biogeosciences 6: 569-583.

Thyssen, M., D. Mathieu, N. Garcia and M. Denis (2008b). Short-term variation of phytoplankton assemblages in Mediterranean coastal waters recorded with an automated submerged flow cytometer. Journal of Plankton Research 30(9): 1027-1040.

Thyssen, M., G. A. Tarran, M. V. Zubkov, R. J. Holland, G. Gregori, P. H. Burkill and M. Denis (2008a). The emergence of automated high-frequency flow cytometry: revealing temporal and spatial phytoplankton variability. Journal of Plankton Research 30(3): 333-343.

Uitz, J., H. Claustre, B. Gentili and D. Stramski (2010). Phytoplankton class-specific primary production in the world's oceans: Seasonal and interannual variability from satellite observations. Global Biogeochemical Cycles 24(3): GB3016.

Van Bleijswijk, J. D. L., R. S. Kempers, M. J. Veldhuis and P. Westbroek (1994). Cell and growth characteristics of types A and B of Emiliania huxleyi (Prymnesiophyceae) as determined by flow cytometry and chemical analyses. . Journal of Phycology 30: 230-241.

Van Heukelem, L. and C. S. Thomas (2001). Computer assisted high performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. Journal of Chromatography A910(1): 31A49.

Vantrepotte, V., H. Loisel, D. Dessailly and X. MÃ©riaux (2012).Optical classification of contrasted coastal waters. Remote Sensing of Environment 123(0): 306-323.

Vargas, M., C. W. Brown and M. R. P. Sapiano (2009).Phenology of marine phytoplankton from satellite ocean color measurements. Geophys. Res. Lett. 36.

Veldhuis, M. J. W. and G. W. Kraay (2000). Application of flow cytometry in marine phytoplankton research: current applications and future perspectives. Scientia Marina 64(2): 121-134.

Waterbury, J. B., S. W. Watson, R. R. L. Guillard and L. E. Brand (1979).Widespread occurrence of a unicellular, marine, planktonic cyanobacterium. Nature 277(293:294).

Werdell, P.J., Proctor, C.W., Boss, E., Leeuw, T. and M. Ouhssain (2013). Underway sampling of marine inherent optical properties on the Tara Oceans expedition as a novel resource for ocean color satellite data product validation. Methods in Oceanography 7: 40-51.

Widdicombe, C. E., D. Eloire, D. Harbour, R. P. Harris and P. J. Somerfield (2010).Long-term phytoplankton community dynamics in the Western English Channel. Journal of Plankton Research 32(5): 643-655.

Wiltshire, K. H. and B. F. J. Manly (2004). The warming trend at Helgoland Roads, North Sea: phytoplankton response. Helgoland marine research 58: 269-273.

Yentsch, C. S. and Menzel, D. W. (1963) A method for the determination of phytoplankton chlorophyll and phaeophytin by fluorescence. Deep Sea Research, 10: 221-231.

Zubkov, M. V. and P. H. Burkill (2006). Syringe pumped high speed flow cytometry of oceanic phytoplankton. Cytometry Part A 69A(9): 1010-1019.

Zubkov, M.V. and G.D. Quartly (2003).Ultraplankton distribution in surface waters of the Mozambique Channel - flow cytometry and satellite imagery. Aquatic Microbial Ecology 33(2): 155-161.

## Figure legends:

Figure 1. Flow cytometry sampling points superimposed on the mixed layer depth (m) calculated with modeled temperature of the water column from the FOAM AMM7 (average values from the 8 to the 12 May 2011). Chosen stations for phytoplankton pictures collection with the flow cytometer are labeled (ST=station, ST4, ST6, ST13). Yellow squares correspond to MODIS matching points for non-turbid waters selected between 6 h and 18 h.

Figure 2. A. Temperature and B. *S*alinity measured with the Pocket Ferry Box. Presented data are selected to match the scanning flow cytometry collected samples. Grey bars delimit the traversed marine areas: H= Humber, T=Tyne, D=Dogger, Th=Thames.

Figure 3. A. TFLO vs TFLR (a.u.) cytogram with a trigger level at 10 mV showing the PicoORG cluster, the PicoRED cluster, the MicroLowORG cluster. B. Maximum SWS (a.u.) vs TFLR (a.u.) cytogram with a trigger level at 10 mV showing the NanoSWS cluster, the NanoRED2 cluster and 3 µm beads. C. TFLR (a.u.) vs TFWS (a.u.) cytogram with a trigger level at 10 mV showing the NanoRED1 cluster, the NanoRED2 cluster, and the Micro1 cluster. D. TFLO vs TLFR (a.u.) cytogram with a trigger level of 25 mV showing the NanoORG1, the MicroORG, the Micro1 and Micro2 clusters and 10 µm beads.

Figure 4: Pictures of cells from the scanning flow cytometer image in flow device collected within the Micro2 cluster. Surface closest stations where Micro2 abundance was the highest (station 4, 6, and 13) are illustrated.

Figure 5. Abundance ($10^3$ cells.cm$^{-3}$) of each phytoplankton cluster resolved with the scanning flow cytometer. Scales are not homogenised for the purpose of distribution evidence. Grey bars separate the traversed marine areas: H= Humber, T=Tyne, D=Dogger, Th=Thames.

Figure 6. Average estimated size for each phytoplankton cluster resolved with the scanning flow cytometer. Scales are not homogenised for the purpose of distribution evidence. Grey bars separate the traversed marine areas: H= Humber, T=Tyne, D=Dogger, Th=Thames.

Figure 7. Scanning flow cytometer Total red fluorescence per unit volume (SFC TFLR.cm$^{-3}$) for each phytoplankton cluster. Superimposed large black squares are the matching points with MODIS pixels in non-turbid waters between 6 h and 18 h. Diamonds correspond to the night SFC samples matching MODIS passage but not taken into account because of the possible differences between day and night community structures. Scales are not homogenised for the purpose of distribution evidence. Grey bars separate the traversed marine areas: H= Humber, T=Tyne, D=Dogger, Th=Thames.

Figure 8. SFC Total TFLR per cm$^{-3}$ compared to chl a analyses using different instruments. Refer to Material and Methods for a detailed description of each method. Blue triangles: AOA fluorometer PFB (chla µg.dm$^{-3}$). Black diamonds: SFC Total TFLR.cm$^{-3}$ (a.u..cm$^{-3}$). Green triangles: Turner fluorometer (chla µg.dm$^{-3}$). Grey triangles: HPLC (chla µg.dm$^{-3}$). Red squares: MODIS chla values corresponding to non-turbid waters (after Vantrepotte et al., 2012) and selected between 6 h and 18 h (chla µg.dm$^{-3}$).

Figure 9: A. Within sum of squares for the optimal number of K-nodes selection corresponding to PHYSAT anomalies. B. Cluster dendrogram defining the two main nodes grouping similar PHYSAT anomalies matchups (N1 and N2). C and D, corresponding Ra (Radiance Anomaly) spectra for N1 and

N2. Red dashed lines correspond to the minima and maxima values of the spectra as described in Table 3.

Figure 10: A and B. Clusters proportional contribution to the Total TFLR.cm$^{-3}$ within each PHYSAT anomaly (N1 and N2). C and D. Within each anomaly, clusters TFLR.cm$^{-3}$ proportional difference to its median value calculated on the entire matching points dataset. Wilcoxon rank test was run for each cluster between the two anomalies. ***p<0,001, **p<0,01, *p<0,1.

Figure 11: Boxplots within each PHYSAT anomaly (N1, N2) of A. Temperature (°C), B. Salinity, C. Chlorophyll *a* (as estimated from MODIS L3 Binned) and D, Total TFLR (a.u..cm$^{-3}$). Wilcoxon rank test was run for each parameter between the two anomalies. ***p<0,001, **p<0,01, *p<0,1.

Figure 12: A and B. Frequency of occurrence of the two distinct anomalies (N1 and N2) over the North Sea during the sampling period (08/05/2011 to the 12/05/2011).

Table 1: Minimal, maximal, average and standard deviation of abundance (cell.cm$^{-3}$) for each defined phytoplankton cluster followed by the size estimated (µm) average ± standard deviation values.

Table 2: Spearman’s rank correlation coefficient between the different methods used for chlorophyll a estimates and with the Total TFLR from the scanning flow cytometer per unit volume. ***p<0,001 ** p<0,01.

Table 3. Minimal and maximal anomaly (Ra) values for each collected MODIS wavelength (nm) that characterizes the edges for the two PHYSAT radiance anomalies spectra (N1 and N2) observed in this study.

Table 1

| Cluster's name | Abundance min-max (cells.cm$^{-3}$) | Average abundance ± SD (cells.cm$^{-3}$) | Average size ± SD (µm) |
|---|---|---|---|
| PicoORG | 25 - 18710 | 1559 ± 2821 | 1.09 ± 0.17 |
| PicoRED | 275 - 26960 | 5674 ± 4647 | 1.83 ± 0.32 |
| NanoRED1 | 97 - 7172 | 888 ± 942 | 2.33 ± 0.33 |
| NanoORG | <10 - 759 | 87 ± 150 | 5.8 ± 2.1 |
| NanoSWS | < 10 - 376 | 99 ± 93 | 10 ± 2.56 |
| NanoRED2 | 200 - 54880 | 4187 ± 7878 | 6.4 ± 1.4 |
| Micro1 | <10 - 4392 | 420 ± 769 | 16.9 ± 5.6 |
| MicroORG | <10 - 306 | 48 ± 60 | 23.5 ± 10 |
| MicroLowORG | <10 - 687 | 69 ± 111 | 23.75 ± 8.6 |
| Micro2 | <10 - 420 | 37 ± 59 | 65.5 ± 21.0 |

Table 2.

| Spearman's correlation coefficient | SFC TFLR.cm$^{-3}$ (a.u.) n=247 | AOA fluorometer (µg.dm$^{-3}$) n=254 | HPLC chla (µg.dm$^{-3}$) n=12 | Turner chla (µg.dm$^{-3}$) n=9 |
|---|---|---|---|---|
| SFC TFLR.cm$^{-3}$ (a.u.) | 1 | 0,93*** | 0,82*** | 0,82*** |
| AOA fluorometer (µg.dm$^{-3}$) | | 1 | 0,86*** | 0,82*** |
| HPLC chla (µg.dm$^{-3}$) | | | 1 | 0,98*** |
| Turner chla (µg.dm$^{-3}$) | | | | 1 |

Table 3.

| Node | Ra (412) nm Min | Ra (412) nm Max | Ra (443) nm Min | Ra (443) nm Max | Ra (488) nm Min | Ra (488) nm Max | Ra (531) nm Min | Ra (531) nm Max |
|---|---|---|---|---|---|---|---|---|
| N1 (n=5) | 1.06 | 1.30 | 0.96 | 1.24 | 0.91 | 1.10 | 0.91 | 1.09 |
| N2 (n=10) | 0.74 | 0.97 | 0.75 | 0.93 | 0.70 | 0.89 | 0.72 | 0.93 |

FIGURE 1

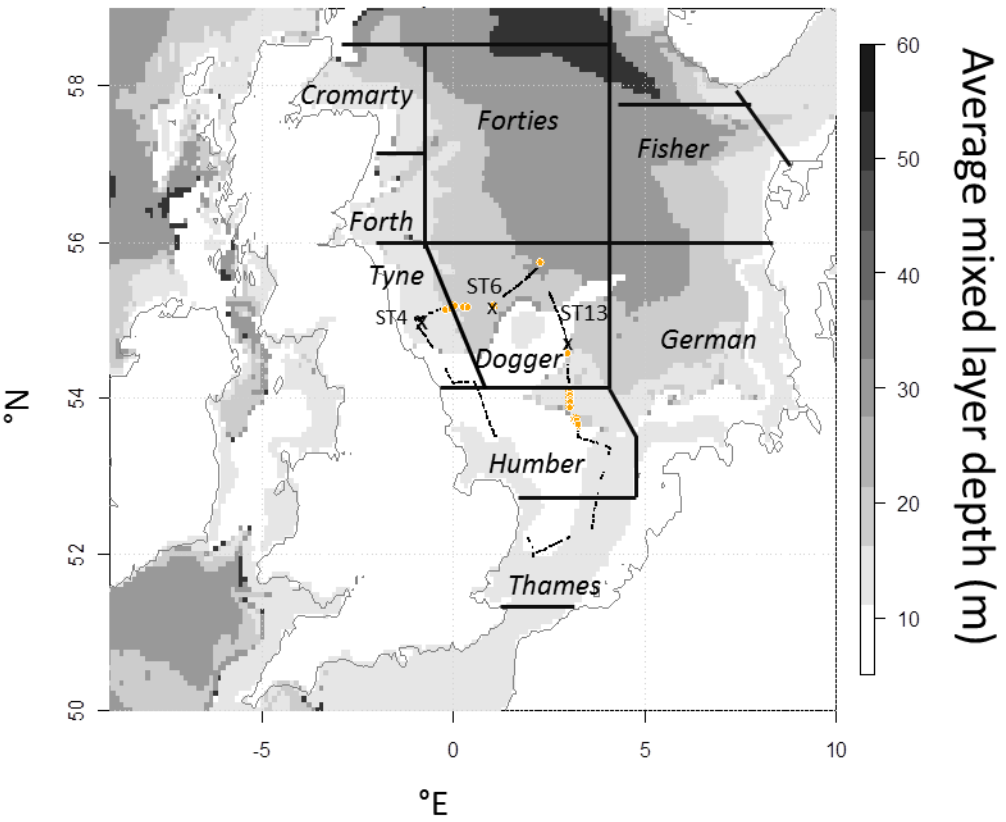
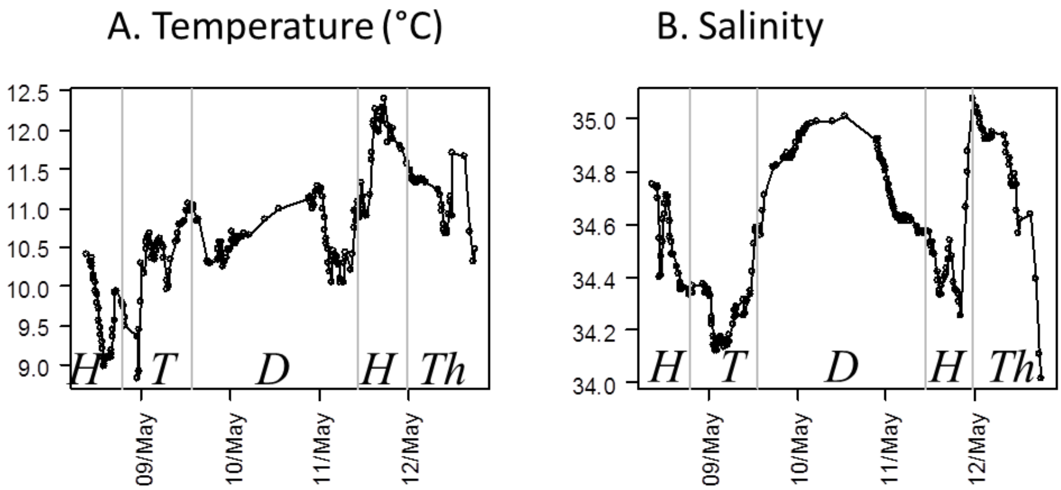

Cromarty
Forties
Fisher
Forth
Tyne
ST6
ST4
ST13
German
Dogger
Humber
Thames
58
56
54
52
50
°N
-5
0
5
10
°E
60
50
40
30
20
10
Average mixed layer depth (m)

FIGURE 2


A. Temperature (°C)

B. Salinity

FIGURE 3

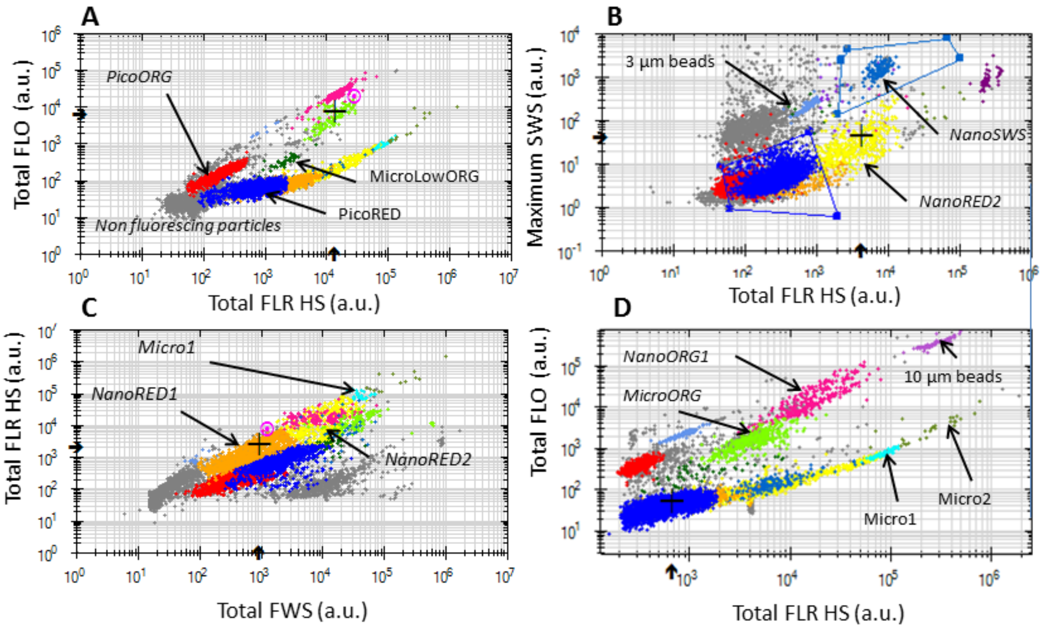
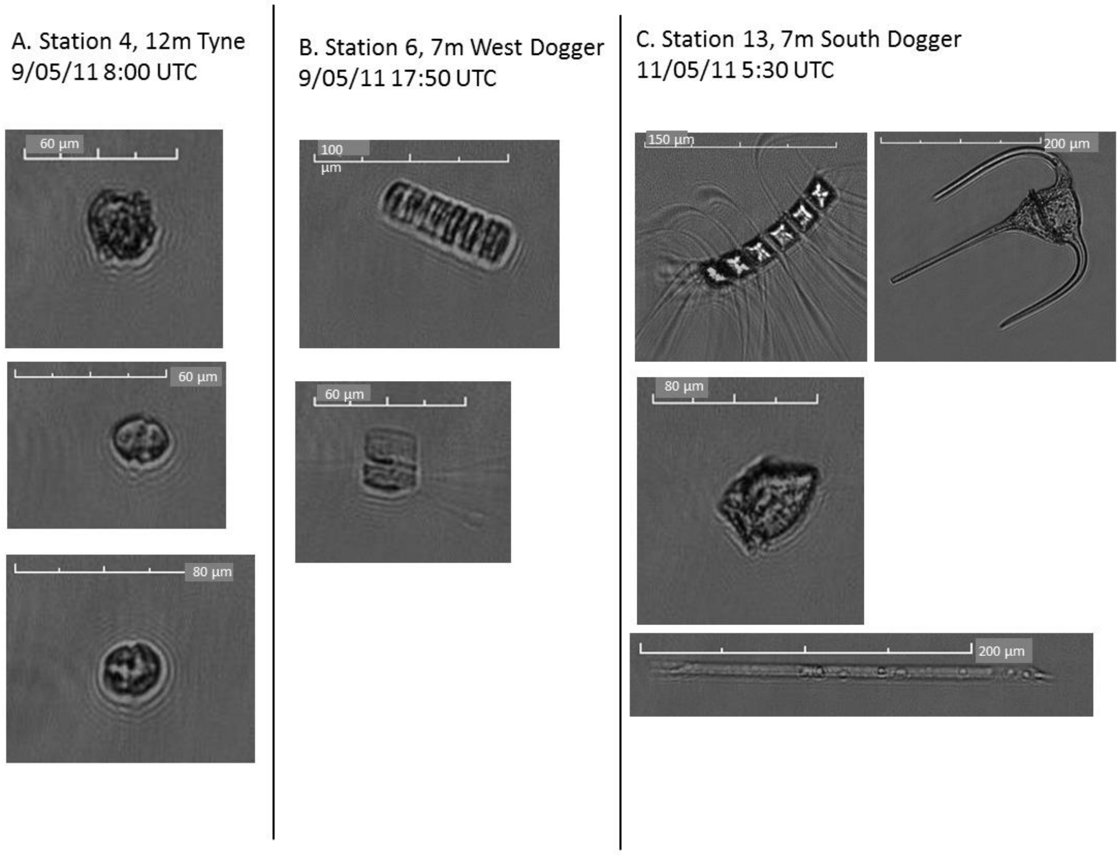
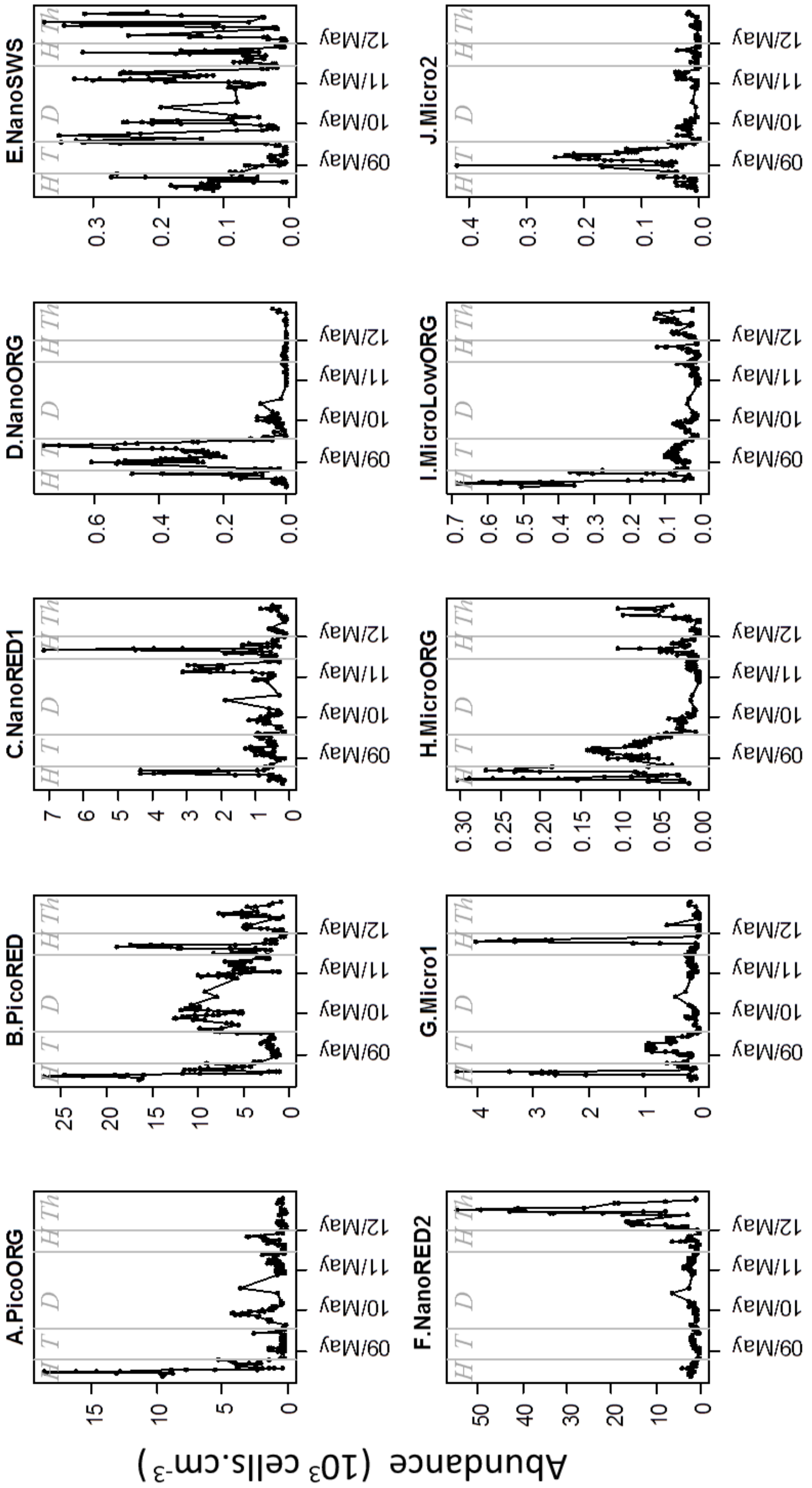

A
Total FLO (a.u.)
Total FLR HS (a.u.)
PicoORG
MicroLowORG
PicoRED
Non fluorescing particles
B
Maximum SWS (a.u.)
Total FLR HS (a.u.)
3 µm beads
NanoSWS
NanoRED2
C
Total FLR HS (a.u.)
Total FWS (a.u.)
Micro1
NanoRED1
NanoRED2
D
Total FLO (a.u.)
Total FLR HS (a.u.)
NanoORG1
MicroORG
10 µm beads
Micro2
Micro1

FIGURE 4


A. Station 4, 12m Tyne
9/05/11 8:00 UTC
60 µm
60 µm
80 µm
B. Station 6, 7m West Dogger
9/05/11 17:50 UTC
100 µm
60 µm
C. Station 13, 7m South Dogger
11/05/11 5:30 UTC
150 µm
200 µm
80 µm
200 µm

966 FIGURE 5

967



968

FIGURE 6

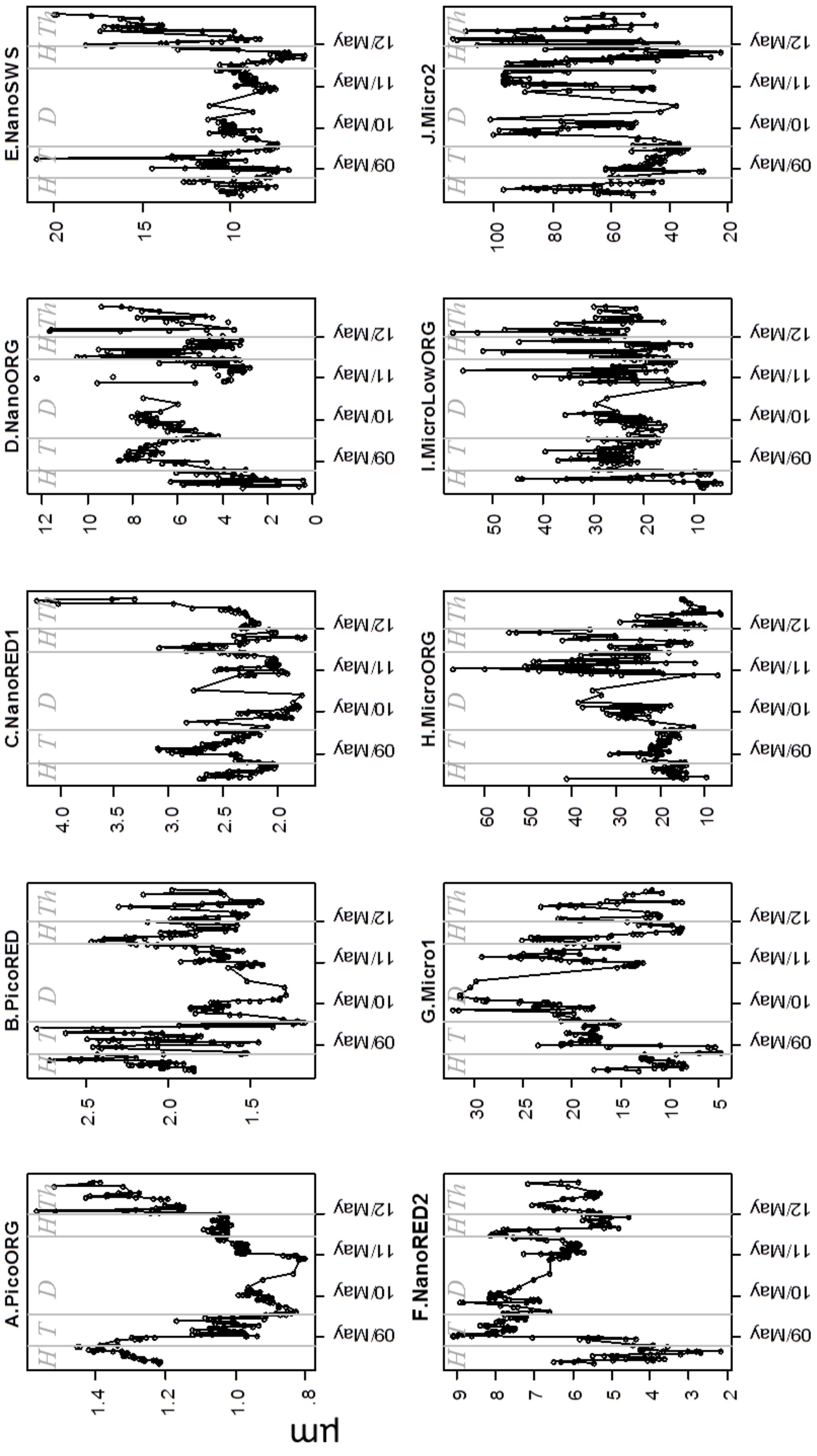

A.PicoORG
B.PicoRED
C.NanoRED1
D.NanoORG
E.NanoSWS
F.NanoRED2
G.Micro1
H.MicroORG
I.MicroLowORG
J.Micro2
µm
H T D H Th
09/May
10/May
11/May
12/May
1.4
1.2
1.0
.8
2.5
2.0
1.5
4.0
3.5
3.0
2.5
2.0
12
10
8
6
4
2
0
20
15
10
9
8
7
6
5
4
3
2
30
25
20
15
10
5
60
50
40
30
20
10
50
40
30
20
10
100
80
60
40
20

FIGURE 7

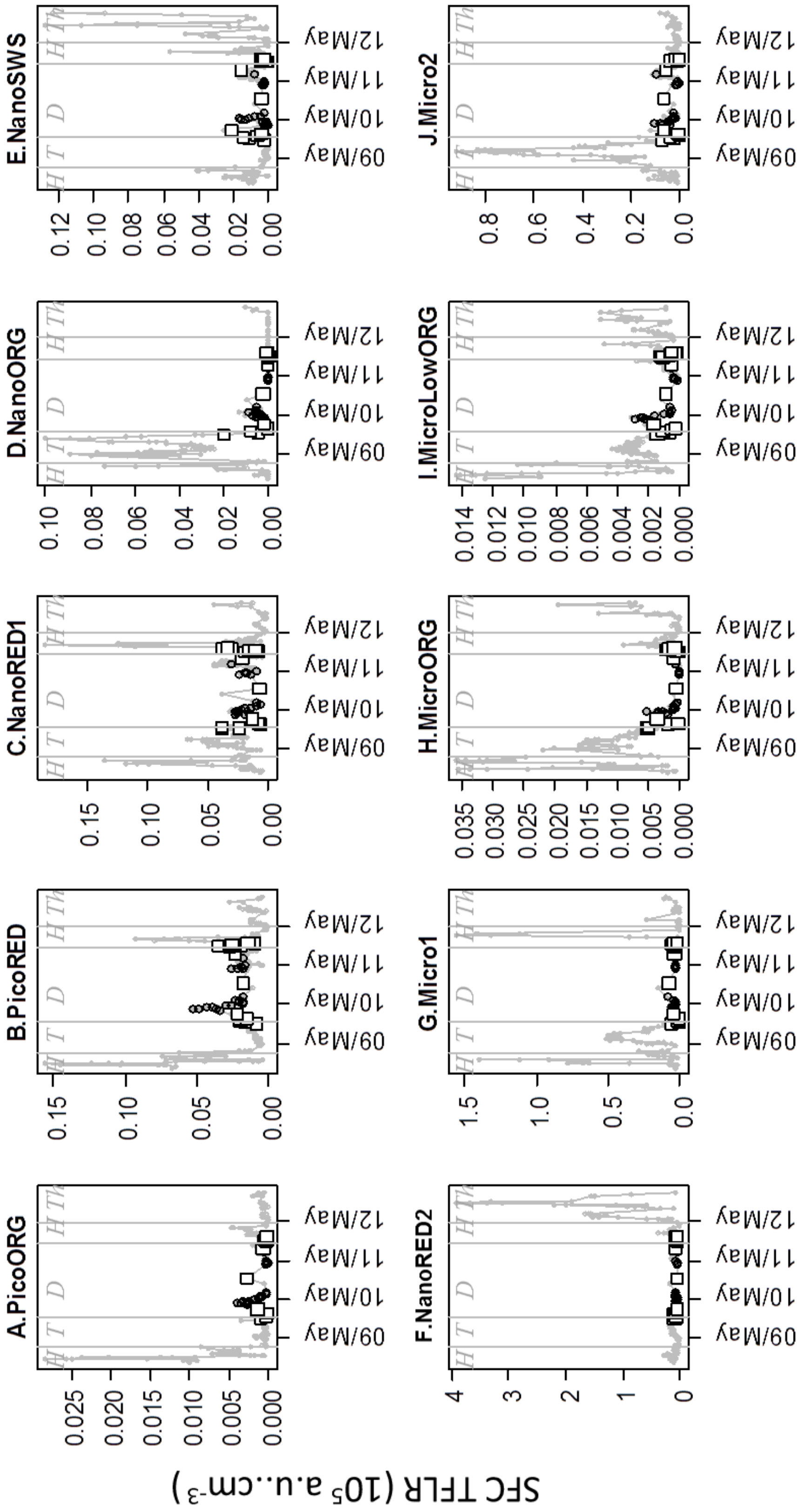

SFC TFLR ($10^5$ a.u..cm$^{-3}$)
A.PicoORG
B.PicoRED
C.NanoRED1
D.NanoORG
E.NanoSWS
F.NanoRED2
G.Micro1
H.MicroORG
I.MicroLowORG
J.Micro2
H
T
D
H
TI
09/May
10/May
11/May
12/May
0.025
0.020
0.015
0.010
0.005
0.000
0.15
0.10
0.05
0.00
0.12
0.08
0.06
0.04
0.02
4
3
2
1
0
1.5
1.0
0.5
0.0
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0.8
0.6
0.4
0.2
0.0

FIGURE 8

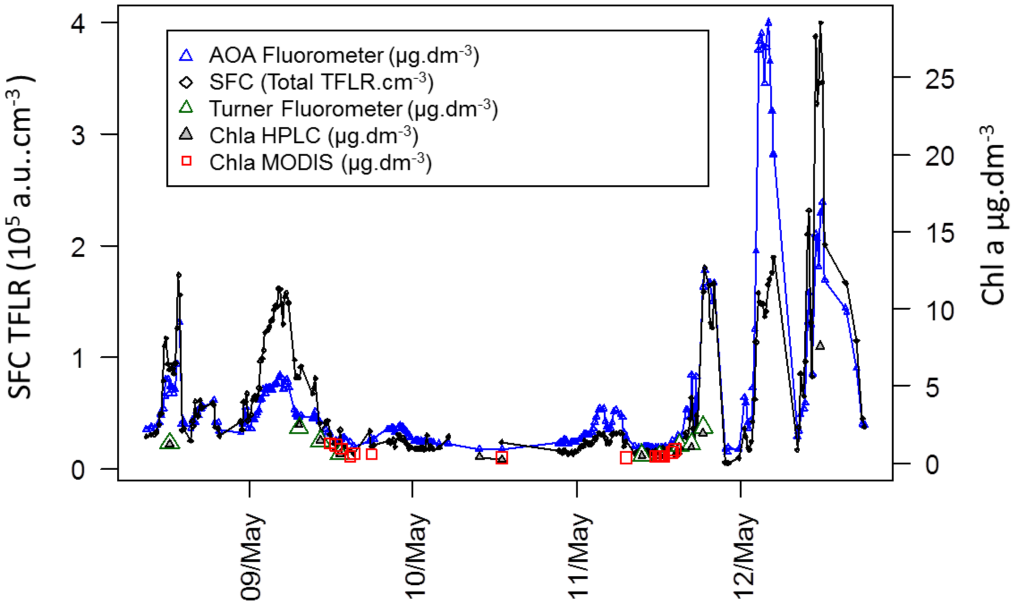

AOA Fluorometer (µg.dm-3)
SFC (Total TFLR.cm-3)
Turner Fluorometer (µg.dm-3)
Chla HPLC (µg.dm-3)
Chla MODIS (µg.dm-3)
SFC TFLR (10^5 a.u..cm-3)
0
1
2
3
4
Chl a µg.dm-3
0
5
10
15
20
25
09/May
10/May
11/May
12/May

FIGURE 9

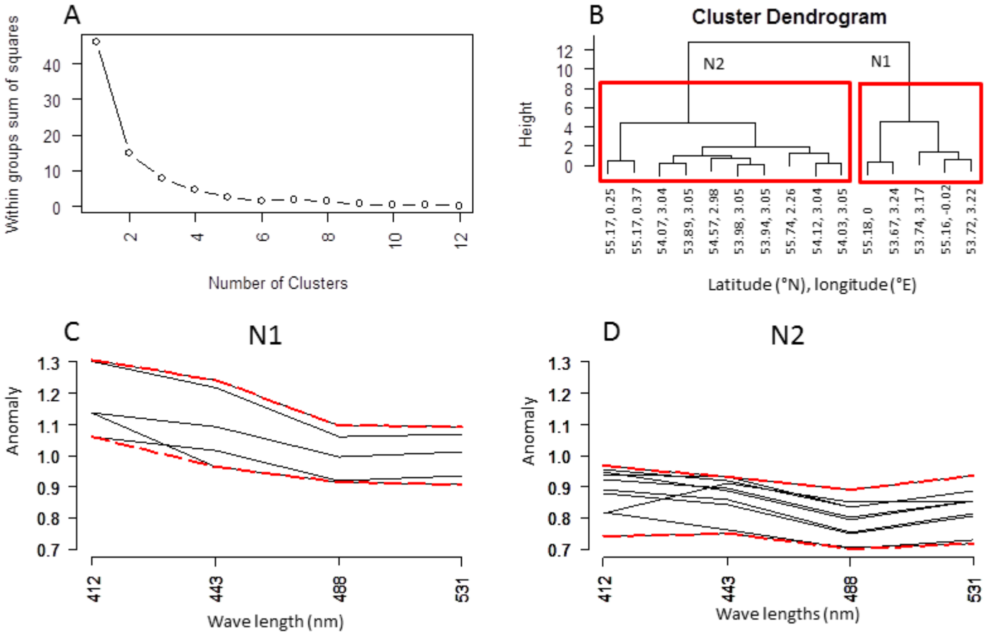

A
Within groups sum of squares
0
10
20
30
40
2
4
6
8
10
12
Number of Clusters
B
Cluster Dendrogram
Height
0
2
4
6
8
10
12
N2
N1
55.17, 0.25
55.17, 0.37
54.07, 3.04
53.89, 3.05
54.57, 2.98
53.98, 3.05
53.94, 3.05
55.74, 2.26
54.12, 3.04
54.03, 3.05
55.18, 0
53.67, 3.24
53.74, 3.17
55.16, -0.02
53.72, 3.22
Latitude (°N), longitude (°E)
C
N1
Anomaly
0.7
0.8
0.9
1.0
1.1
1.2
1.3
412
443
488
531
Wave length (nm)
D
N2
Anomaly
0.7
0.8
0.9
1.0
1.1
1.2
1.3
412
443
488
531
Wave lengths (nm)

FIGURE 10

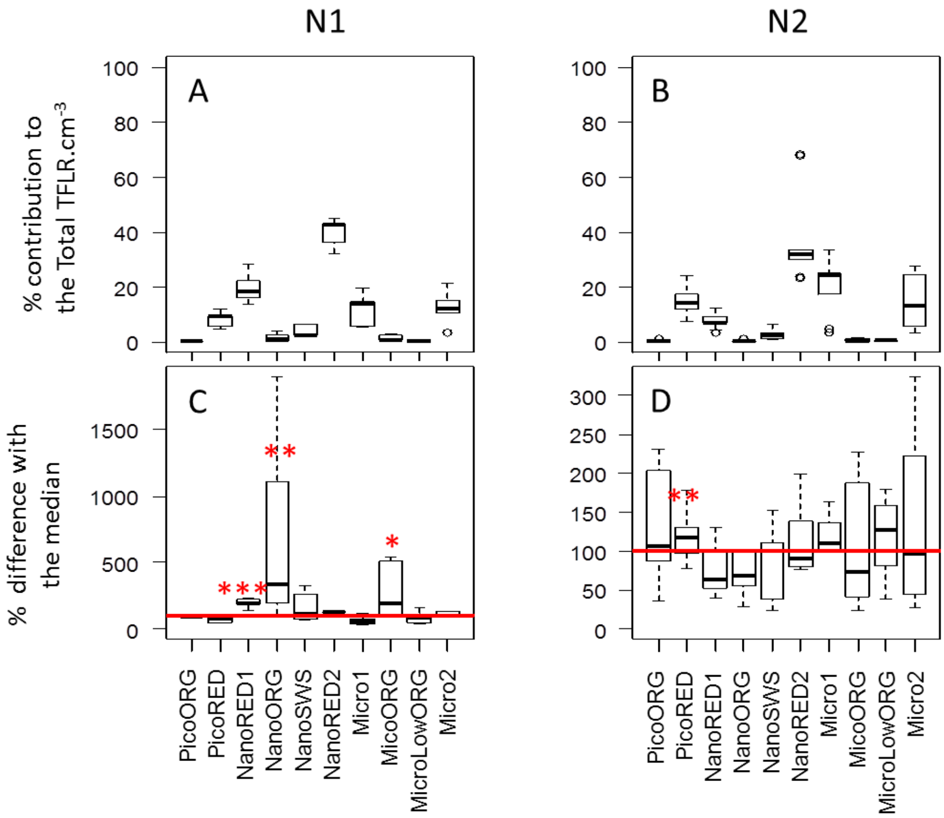
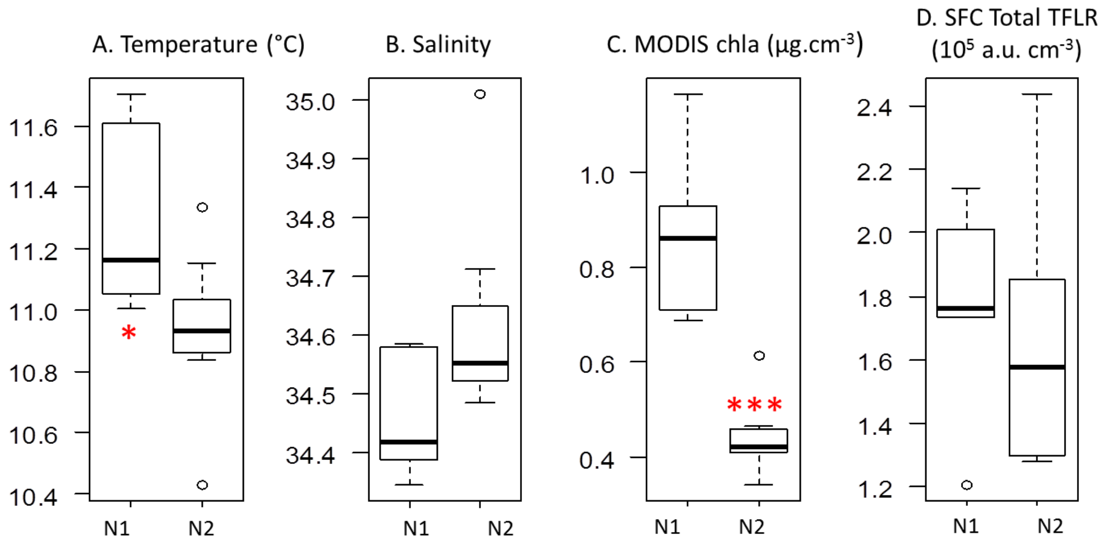
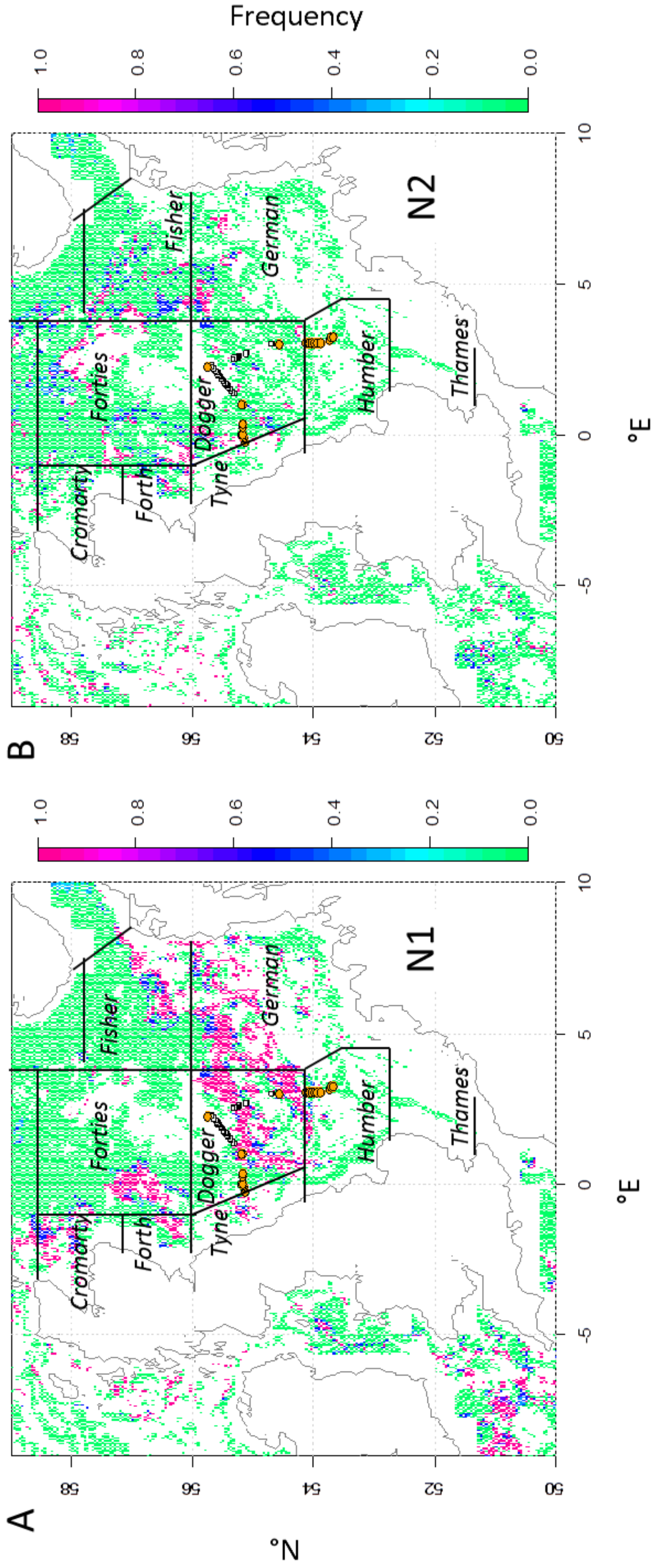

N1
N2
A
B
C
D
% contribution to the Total TFLR.cm$^{-3}$
% difference with the median
100
80
60
40
20
0
1500
1000
500
0
300
250
200
150
100
50
0
**
***
*
**
PicoORG
PicoRED
NanoRED1
NanoORG
NanoSWS
NanoRED2
Micro1
MicoORG
MicroLowORG
Micro2

FIGURE 11


A. Temperature (°C)
11.6
11.4
11.2
11.0
10.8
10.6
10.4
*
N1
N2
B. Salinity
35.0
34.9
34.8
34.7
34.6
34.5
34.4
N1
N2
C. MODIS chla (µg.cm-3)
1.0
0.8
0.6
0.4
***
N1
N2
D. SFC Total TFLR
(105 a.u. cm-3)
2.4
2.2
2.0
1.8
1.6
1.4
1.2
N1
N2

1036 FIGURE 12

1037



1038