IMEKO International Conference on
Metrology for The Sea
Naples, Italy, October 11-13, 2017

Phytoplankton dynamics by autonomous high-frequency flow cytometry from a floating buoy in the Gulf of Naples

Mark.A. van Dijk[1,2], Augusto Passarelli[1], Fabio Conversano[1], Raffaella Casotti[1]

[1] *Stazione Zoologica Anton Dohrn di Napoli, Villa Comunale, Napoli I80121, Italy,* [2] mark.vandijk@szn.it

***Abstract* – A CytoSub (Cytobuoy bv) scanning flow cytometer has been operating for 2 months on a floating buoy at a coastal site in the Gulf of Naples, measuring phytoplankton concentrations every 2 hours. The system was connected to an elastic beacon which provided power and data transmission to the onshore station by wifi. A user-interface was build using R software to automate the data analyses and identify significant cell clusters. Several clusters of autofluorescing phytoplankton were identified, based on different distributions of optical parameters such as scatter and fluorescence. All clusters showed both rapid (in term of days) and gradual (weeks to months) changes in concentrations during the deployment period. Each cluster showed peak of more than twice and up to ten cell numbers within one or few days, rapidly disappearing. This can be due to the internal dynamics of the population, including vertical migrations, and also to advection of water masses passing by the buoy site. These data highlight the importance of high-frequency measurements of cell abundances, due to the intrinsic rapid reactions of phytoplankton to environmental variability and represents a useful tool to detect sudden blooms of species that may represent a hazard for the ecosystem or even human health.**

Introduction

The key role of phytoplankton at the base of the food web in the marine ecosystem and in the global carbon cycle [1] makes it of high importance to monitor their abundances and cell properties and to understand and possibly predict their dynamics in terms of the overall community, and also of functional groups within it. Accessibility to sampling sites, time and resources needed for conventional sampling, as well as the harsh conditions at sea for instrumentation, hamper observations at the appropriate time scales, particularly in coastal areas, where variability often occurs in a narrow boundary between water masses along the coast [2].

In this particular study, a proof of concept is presented for autonomous *in situ* monitoring of the phytoplankton community dynamics at relative high frequency (every 2 h) by means of scanning flow cytometry from a buoy system. The flow cytometer is part of a larger and complex monitoring system operating in the Gulf of Naples and equipped with several probes for meteorological and oceanographic data.

I. METHODS

An automated, submersible high-frequency, scanning flow cytometer (SFCM), CytoSub (Cytobuoy B.V., Woerden, NL), was used from March 29th to May 17th, 2016, placed inside a dedicated buoy, connected by cable to an elastic beacon equipped with solar panels, WIFI connection to land, and environmental parameter sensors (Fig. 1). The system was located offshore in the Gulf of Naples, Italy, within the coastal zone of Naples, an area of interest due to its social-economic relevance in terms of tourism and fisheries and the environmental impact on the ecosystem introduced by the dense urbanization in the region.

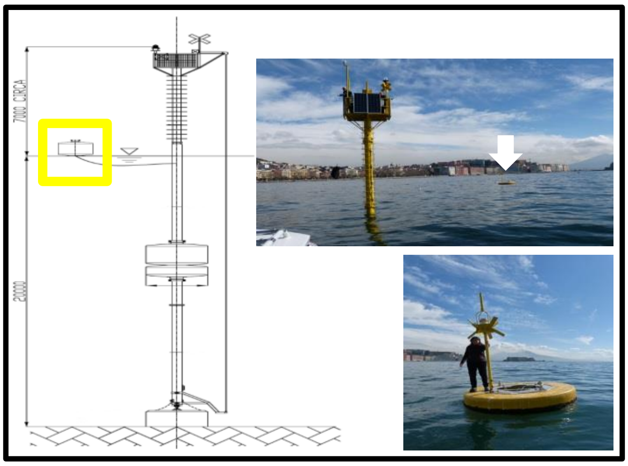

Fig.1 Left, scheme of the floating buoy (yellow square) connected to the main elastic beacon. Right up, pictures of the elastic beacon and the floating buoy (arrow). Right, down, blow up of the floating buoy

hosting the CytoSub.

The CytoSub is equipped with a 488 nm 100 mW blue diode laser and detectors for scatter (forward and sideward), and red, orange and yellow fluorescence. It is designed to handle the majority of the phytoplankton size spectrum [3] and samples directly from the surrounding seawater. Samples were scheduled to be recorded every two hours using overlapping measurement protocols to cover all phytoplankton size classes. Here we show data from the pico (1 to 5 µm ESD, Equivalent Spherical Diameter) and nano (5 to 20 µm ESD) fraction. The size determination of the CytoSub being prone to errors, these sizes are to be considered only indicative. The CytoSub in this configuration could not detect cyanobacteria *Prochlorochoccus* cyanobacteria due to their dim fluorescence and small size at surface.

The data was analyzed manually, using the dedicated clustering software, Cytoclus3, provided with the instrument. Due to the large amount of measurement files (475 for less than two months), gate definitions were drawn based on the position of the clusters in a handful of samples (~10), spread over the time series, followed by a bulk export.

Successively, the data was analyzed with a clustering model incorporated into a user-friendly interface [8] built by us using R and the Shiny package[4,5] to automatically batch-analyze large sets of CytoSub data. The routine used builds upon the work described in [6] which uses the FlowPeaks clustering algorithms[7]. The current version uses a limited amount of parameters from the CytoSub to cluster the data; FL Red Range, FL Orange Range and SWS Length (all log transformed). Signal Range is the signal maximum corrected for the minimum.

First, an artificial mixed sample was created by subsampling *n* particles at random from each individual file in the dataset to obtain a total of 100.000 particles. The representative training sample was then clustered with the FlowPeaks model described in [6], after which the main “noise” populations were removed by the user (noise defined as non-phytoplankton particles and instrument/electronic noise).

In the second clustering stage the cleaned training sample was clustered again to separate and identify the clusters. Large clusters that were cut into smaller ones were grouped together by assigning the same name to the. Once the clusters were identified, a random forest model (1000 iterations) was ran for both stages of clustering. The trained model was then used to process all individual files, followed by concentration calculations. Identified clusters have been grouped successively into larger groups for ease of representation of the concentrations.

The values for multipliers h0 and h of the variance matrices S0 and S, and tolerance values Tol are shown in Table 1.

	Tol	H0	h
Round 1	0.10	0.05	1.5
Round 2	0.05	0.05	1.3

Table 1: parameter settings of the FlowPeaks model

Model results were compared with the manual clustering to ensure their validity. The concentration of *Synechococcus* and the summed concentration of all picoeukaryotes followed the same pattern over time, had an $R^2 > 0.98$, but were on average, respectively, 0.39% lower and 6.26 % higher. The difference for picoeukaryotes is mostly due to the model separating clusters at the boundary between noise and cells while the manual clustering was drawn at a safe distance from the noise.

II. RESULTS

The clusters identified by the clustering routine are shown in Fig. 2. Based on this analysis time dynamics of cyanobacteria *Synechococcus*, belonging to picophytoplankton, total picoeukaryotes, nanophytoplankton and total cryptophytes (pico and nano- sized algae showing orange fluorescence from phycoerythrin pigment) could be monitored in time at the fixed station.

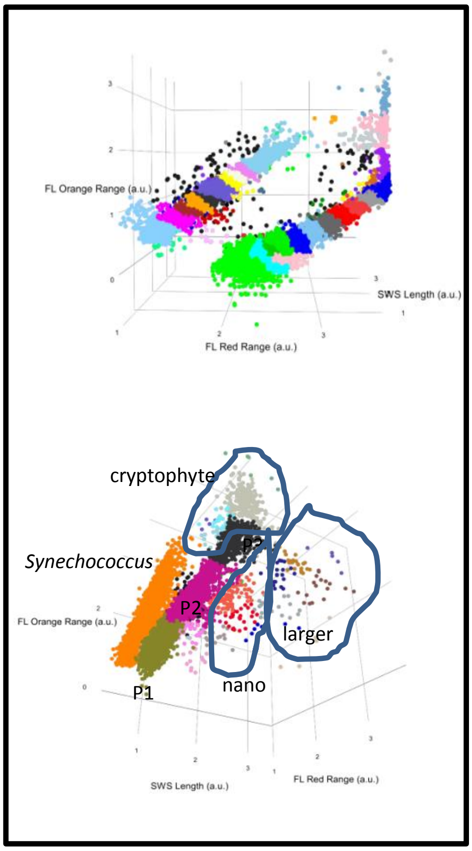

Fig. 2. Upper panel: 3D plot of red vs orange vs scatter showing all the 59 clusters identified after the second round of the Flowpeaks clustering of R. Lower panel: merging of clusters into the different groups described in this paper (*Synechococcus*, picoeukaryotes 1, 2 and 3 (P1, P2, P3), cryptophytes, nanoplankton (nano) and larger phytoplankton (larger).

Data from the pilot 2-months deployment of the CytoSub during late spring 2016 shows a number of peaks in the concentrations of the different clusters (Fig. 3).

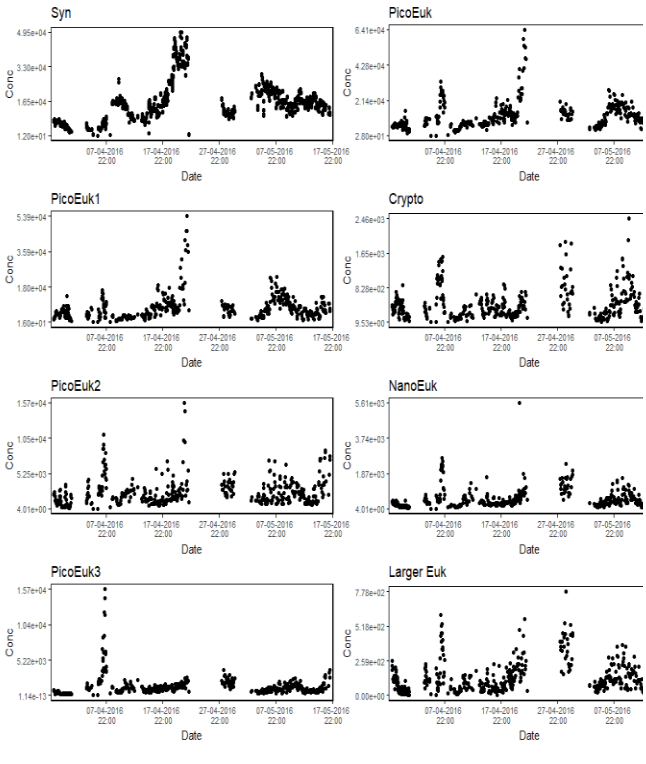

Fig. 3. Cell concentrations (cell ml-) of *Synechococcus* (Syn), total picoeukaryotes (picoeuks), the three subpopulations of picoeukaryotes (PicoEuk1, 2 and 3, cryptophyes (Crypto), nanophytoplankton (nanoEuk) and larger phytoplankton (MicroEuk) during the sampled period.

Around the 7th of April the majority of eukaryotic clusters rapidly increased in concentration by a factor 2 to 10. The increase was not simultaneous, resulting in an overall increase in concentration by a factor 4 for the entire picoeukaryotic community, which took less than a day to establish itself, followed by a decrease over the next 4 days.
A second increase in abundance could be observed from the 15th of April onwards. This increase was observed for all clusters, including *Synechococcus* and the nano- and micro- eukaryotes. The build-up of this peak occurred over a longer period (7 days) and reached higher particle concentrations than the earlier peak.
Towards the end of the recorded period, shortly after the collapse of the main ‘bloom’, concentrations again increased, this time most strongly for picoeukaryotic clusters 2 and 4 and the micro eukaryotic group.

III. DISCUSSION

These preliminary results show the feasibility of high-frequency monitoring of phytoplankton at a coastal fixed station, following the work by [9,10]. It also confirms the usefulness of flow cytometry as a sensor for plankton monitoring to be used in coastal observatories.

The automated routine, while still under development, used for the analysis is very useful when dealing with such large datasets where full manual analysis (file by file) is simply not feasible and needs to be further tested and expanded in order to benefit from its full potential. This includes, but is not limited to, incorporating more relevant parameters recorded by the CytoSub to reduce the number of clusters created by the algorithms (Fig. 2). Here only three parameters were used in combination with a very high sensitivity (model parameter settings for Tol, h and h0) in order to isolate the rare populations, resulting in breaking large clusters up into smaller ones.

Despite this, the results from the automated model compared well to the concentrations of the manual bulk export analysis for the dataset discussed in this paper, which was also the case for a comparison on a much smaller dataset that was manually analyzed file by file [12].

Data show how dynamic the concentration of the overall phytoplankton community is in the Gulf of Naples, as well as that of individual clusters within the community. Despite the gaps in the recording due to technical problems with the instrumentation, the collected information is more detailed than that collected by regular weekly sampling at a nearby station, and shows how small, short lasting peaks in abundance occur frequently, as well as it evidences a longer lasting high abundance period characterized by a more gradual build-up and decline.
The site monitored being very close to the coast, comparison with a more offshore station (LTER-MC)

located two miles offshore and sampled weekly by the Stazione Zoologica Anton Dohrn [11]. will be necessary in order to assess the contribution to the general dynamics of the Gulf and possibly add information thanks to the high frequency of sampling.
The presented data will be extended with analysis of data from dedicated Cytosub measurements of nano- and micro- plankton, which will provide more detailed information on this fraction of the phytoplankton, as well as the recorded environmental data to link the observed dynamics to environmental driving factors.

IV. CONCLUSIONS AND OUTLOOK

The results presented here illustrate clearly how dynamic the concentrations of populations within the phytoplankton community are on the short temporal scale. The peaks in phytoplankton abundance show an increase by multiple factors and can be short-lived as seen here. These variations would not be observed with regular weekly sampling. Also, this data shows the risk of over- or underestimating phytoplankton abundance based on sampling at relatively infrequent intervals.
Having overcome the growing pains with the complete system in terms of technical issues and user experience, in the immediate future the CytoSub will be deployed again in the buoy at Naples as well as a second system in the neighboring bay of Bagnoli to record a more complete dataset for a longer period. In the meantime further improvements will continue as well as further development of automation of the data analysis with the use of clustering algorithms.

REFERENCES

[1] C.B.Field, M.J.Behrenfeld, J.T.Randerson, P.G.Falkowski, "Primary production of the biosphere: integrating terrestrial and oceanic components", Science, 291, 1998, pp. 237-240

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

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

[4] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[5] W.Chang, J.Cheng, J.J.Allaire, Yihui Xie and Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.3. https://CRAN.R-project.org/package=shiny.

[6] M.K.Thomas, S.Fontana, M.Reyes, F.Pomati, "Quantifying cell densities and biovolumes of phytoplankton communities and functional groups using scanning flowcytometry, machine learning and unsupervised clustering", L&O Methods, *in prep.*

[7] Y.Ge et al, FlowPeaks: a fast unsupervised clustering for flow cytometry data via K-means and density peak finding, 2012, Bioinformatics 8(15): pp. 2052-8.

[8] M.A.vanDijk, R.Casotti, "Flow cytometric cluster analysis of Cytosense data using FlowPeaks: An R Shiny user-interface. *In prep.*

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

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

[11] M.Ribera d'Alcalà, F.Conversano, F.Corato, P.Licandro, O.Mangoni, D.Marino, M.G.Mazzocchi, M.Modigh, M.Montresor, M.Nardella, V.Saggiomo, D.Sarno, A.Zingone. "Seasonal patterns in plankton communities in a pluriannual time series at a coastal Mediterranean site (Gulf of Naples): an attempt to discern recurrences and trends. Scientia Marina, 68(S1), 2004, pp. 65-83.

[12] R.Casotti, C.Balestra, M.A. van Dijk, A.Passarelli, J.Sigrist, F.Hammes. "Real-time microbial concentrations by automated on-line flow cytometry for marine coastal monitoring. *In press*, this volume 2017.