Project	AtlantOS – 633211
Deliverable number	7.9
Deliverable title	BIO EOV Report
Description	Biogeochemical EOV assessment and impact of AtlantOS observations
Work Package number	7
Work Package title	Data flow and data integration
Lead beneficiary	PML, UPMC
Lead authors	Giorgio Dall'Olmo and Hervé Claustre
Contributors	N/A
Submission date	2018-April-30
Due date	2018-March-31
Comments	N/A

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement n° 633211.

Last updated: 30 April 2018

Stakeholder engagement relating to this task*

WHO are your most important stakeholders?	☐ Private company If yes, is it an SME ☐ or a large company ☐? ☐ National governmental body ☐ International organization ☐ NGO ☑ others Please give the name(s) of the stakeholder(s): International ocean-science community
WHERE is/are the company(ies) or organization(s) from?	☑ Your own country ☑ Another country in the EU ☑ Another country outside the EU Please name the country(ies): World wide
Is this deliverable a success story? If yes, why? If not, why?	Yes, because it has demonstrated how value can be added to existing AtlantOS datasets by synergistically merging surface and vertically resolved information. Moreover, we have shown that estimates of critical but unmeasured variables can be derived from a set of easily-measureable variables. Our results have important implications for both augmenting existing datasets as well as for better understanding ocean biology and biogeochemistry.
Will this deliverable be used? If yes, who will use it? If not, why will it not be used?	☑ Yes. The international community will use this deliverable to better understand the magnitude and temporal variability of ocean biogeochemical cycles including the biological carbon pump.

NOTE: This information is being collected for the following purposes:

1. To make a list of all companies/organizations with which AtlantOS partners have had contact. This is important to demonstrate the extent of industry and public-sector collaboration in the obs community. Please note that we will only publish one aggregated list of companies and not mention specific partnerships.
2. To better report success stories from the AtlantOS community on how observing delivers concrete value to society.

Executive summary

The overall objective of the work described in this deliverable was to develop techniques to merge satellite and in-situ observations in order to maximise their information content and obtain three-dimensional fields of essential ocean variables (EOVs). To this aim, we have developed different approaches that can be grouped into two broad categories. The first category includes products that have been developed and tested enough to be considered close to operational. Products that are still at the research stage are instead included in the second category.

Three "quasi-operational" products were developed based on artificial neural networks trained using in-situ data from cruises and Biogeochemical-Argo (BGC-Argo) floats. Two of these products can be used to merge satellite estimates of surface chlorophyll concentration (*chl-a*) and optical backscattering (b_{bp}, a proxy for suspended particles in the ocean) with T/S profiles collected by Argo floats to derive 3-dimensional fields of *chl-a* and b_{bp}. These products are important because they allow one to extend surface properties to depth and can be used to improve estimates of net primary production (chl-a) and of stocks of particulate organic carbon. The third product can be used to estimate the concentrations of nitrate, phosphate and silicate as well as the parameters of the carbonate system from profiles of temperature, salinity, hydrostatic pressure and oxygen. This latter set of products is important because not only it allows one to augment the parameter space of existing datasets, but also it can be used as an automatic quality-control procedure for new datasets including from BGC-Argo floats.

Research products focused on expanding our understanding of the biological carbon pump, i.e., the suite of processes responsible for transferring organic carbon from the surface to the deep ocean. We first investigated the role of the seasonal mixed-layer pump at the global scale by merging satellite estimates of particulate organic carbon with mixed-layer depths estimated from core-Argo floats. Results showed that, in high-latitude regions where mixed layers are usually deep, this pump contributes a flux of organic carbon that is on average about 20%, but can be as high as 100%, of the expected gravitational carbon flux. Our main conclusion is that this pump must supply a substantial energy flux to the mesopelagic ecosystem in these regions. Second, we analysed data from a subset of BGC-Argo floats deployed in the North Atlantic to better understand the mechanisms that regulate the export of large fresh aggregates that is typically recorded in correspondence of the spring bloom. We found that 1) export events occur between within two weeks of the seasonal peak of surface chl-a and 2) that this time increases with the magnitude of the surface bloom, 3) but is not related to the estimated aggregate sinking rates, and 4) that the magnitude of the export event is positively related to the magnitude of the surface bloom. These results are important because for the first time they exploit vertically-resolved high-resolution data of optical properties to provide a new view of the biological carbon pump.

Introduction

The development of new products and applications based on the synergistic use of ocean-colour satellite data and physical or bio-optical data acquired by core-Argo and Biogeochemical-Argo (BGC-Argo) profiling floats has been identified as a promising area of research (IOCCG, 2011). This field of research is expected to progress rapidly as the number of deployed BGC-Argo floats will increase in the near future. Indeed, the outlines of the future BGC-Argo network are already defined and its implementation, already initiated on a regional scale (pilot projects including a strong focus in the North Atlantic), is increasing year by year.
This document summarises the results of a series of AtlantOS activities aimed at augmenting the information contained in BGC-Argo data by synergistically merging them with other observations. The document is divided in two sections: one dedicated to quasi-operational products and one to products that are still at the research stage.

"Quasi-operational" products

Since satellites only sense approximately 1/5 of the sunlit ocean layer, in-situ measurements represent the natural complement to extend the surface information over the vertical dimension. Benefiting from the new BGC-Argo database, it should be possible therefore to develop 3D parameterizations of certain bio-optical variables.

We began our work by developing algorithms based on neural networks to estimate the 3D distribution of the backscattering coefficient (b_{bp}) from satellite observations of ocean colour (chlorophyll-a concentration and b_{bp}) and concomitant T/S Argo profiles. The method uses a database consisting of vertical profiles of temperature, salinity and $b_{bp}(700)$ acquired by BGC-Argo floats to train a neural network (Figure 1).

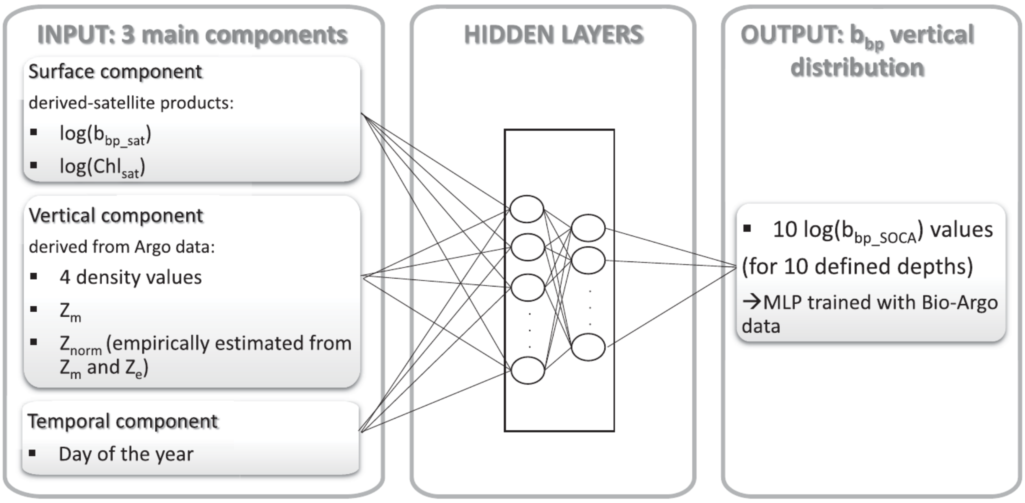

Figure 1. Schematic overview of the SOCA-BBP MLP-based algorithm that retrieves the vertical distribution of b_{bp} from merged ocean colour satellite and Argo data associated with the day of the year of the considered satellite-to-Argo match up.

With this method, all the temperature and salinity measurements from the Argo network in a given area could potentially be exploited simultaneously with satellite observations to predict the vertical distribution of $b_{bp}(700)$ (Figure 2). This method, which takes advantage of the higher density of the core-Argo T/S measurements, makes it possible to considerably increase the density of b_{bp} profiles. The first validations seem satisfactory (Figure 3). Moreover, by construction, the estimated profiles are intrinsically homogeneous with the satellite observations (Sauzède et al., 2016).

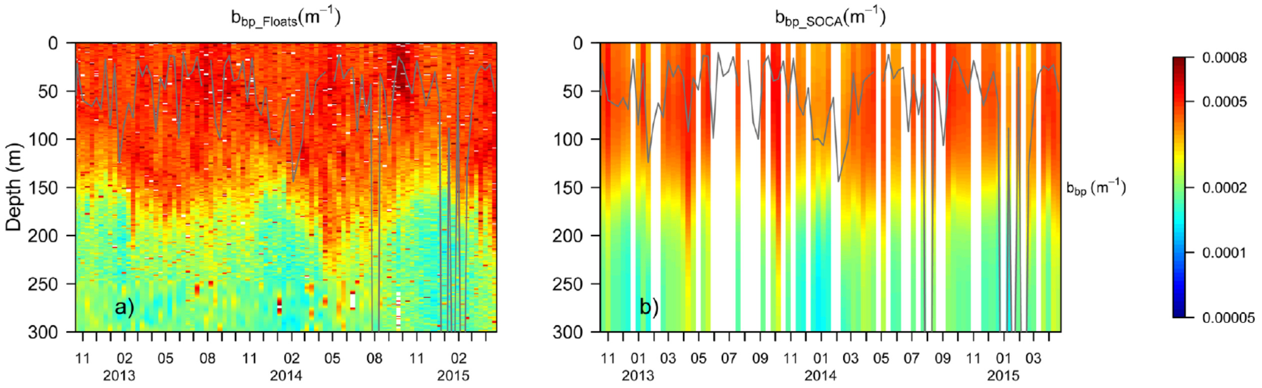

Figure 2. Comparison of the reference b_{bp} measurements acquired by Bio-Argo floats, bbp_Floats (a), with the values predicted by SOCA-BBP (b). Timeseries for a BGC-Argo float deployed in the North Atlantic Subtropical Gyre (WMO 6901472). The grey line in each plot indicates the depth of the mixed layer.

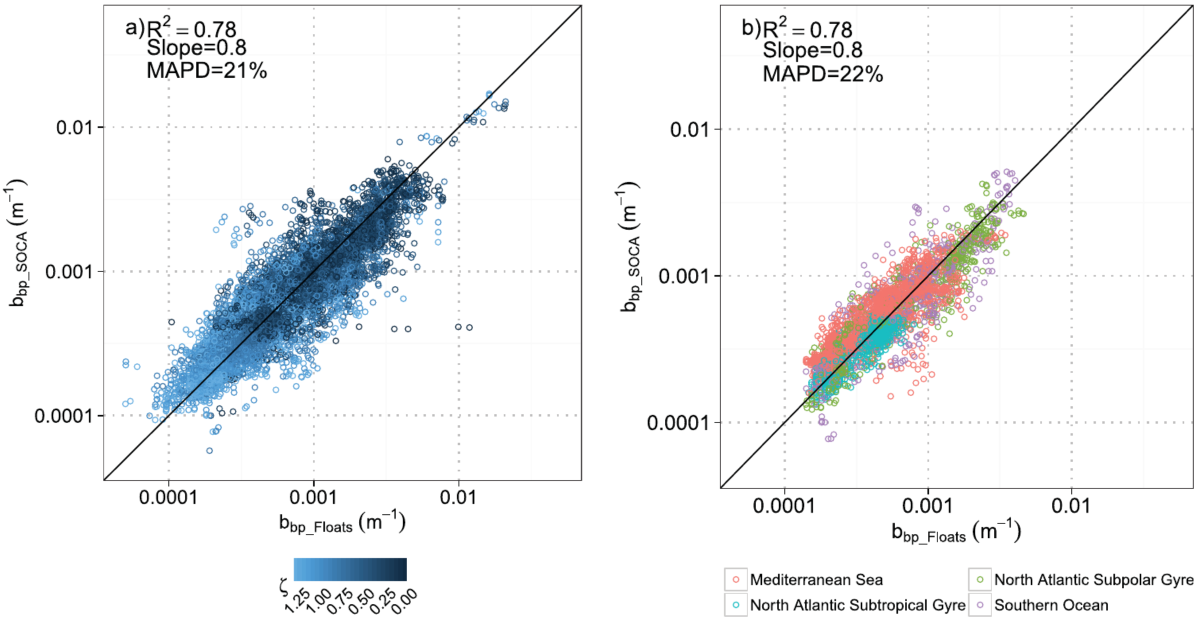

Figure 3. Comparison of the bbp values retrieved by SOCA-BBP (bbp_SOCA) to the reference b_{bp} measurements acquired by the BGC-Argo floats (bbp_Floats) using two different data sets: (a) the validation database (i.e., 20% of the entire database chosen randomly) with data ordered according to the dimensionless depth f; (b) the independent data acquired by four BGC-Argo floats not integrated in the training and validation databases with the colour code indicating the oceanic basins in which the BGC-Argo floats were deployed. The 1:1 line is shown in each plot.

Additional neural networks were developed (Fig. 4, Sauzède et al., 2017) to obtain a seasonal view of phytoplankton biomass together with macro-nutrients as well as phytoplankton primary production and growth rate. Results from these networks demonstrate that a basic set of vertically-resolved variables can be used to estimate with relatively good accuracy a series of variables that are critical to understand ocean biology and biogeochemistry (Table 1).

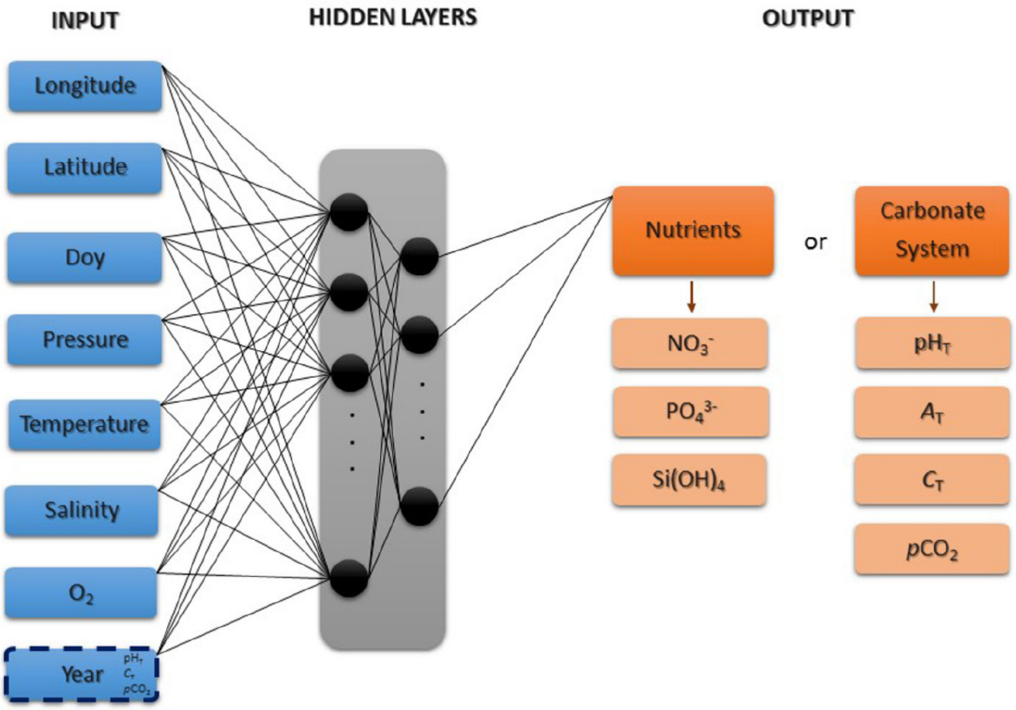

Figure 4. Schematic representation of the CANYON MLP-based neural-network algorithm that retrieves the concentrations of nutrients [NO_3^-, PO_4^{3-}, and $Si(OH)_4$] and the parameters of the carbonate system in seawater (pH_T, AT, C_T, and pCO_2). The input variables are measured temperature, salinity, O_2, and hydrostatic pressure (i.e., depth) together with the geolocation and time of sampling. Doy, day of year.

Table 1. CANYON retrieval accuracy (RMSE) for each variable in each of the eight independent zones.

	NO_3^- (μmol kg^{-1})	PO_4^{3-} (μmol kg^{-1})	$Si(OH)_4$ (μmol kg^{-1})	pH_T	A_T (μmol kg^{-1})	C_T (μmol kg^{-1})	pCO_2 (%)
Sub-Equatorial Pacific	0.99 (284)	0.073 (284)	4.4 (284)				5.7 (274)
Sub-Equatorial Indian	0.37 (886)	0.036 (903)	2.8 (919)		5 (408)	4 (418)	3.4 (367)
North Atlantic Subtropical Gyre	0.59 (4,526)	0.052 (3,848)	1.1 (4,660)	0.014 (1,468)	5 (2,031)	5 (2,565)	3.8 (1,884)
North Atlantic Subpolar Gyre	0.60 (2,889)	0.045 (1,293)	1.7 (3,168)	0.014 (1,960)	7 (1,428)	6 (1,469)	3.0 (1,333)
North Pacific	0.94 (1,005)	0.093 (1,017)	4.5 (1,005)	0.024 (341)	6 (343)	8 (331)	6.3 (330)
South Atlantic	0.85 (1,395)	0.065 (1,323)	2.5 (1,396)	0.016 (525)	6 (516)	7 (571)	4.5 (509)
South Indian	0.69 (1,511)	0.039 (1,490)	1.8 (1,512)	0.015 (833)	6 (841)	5 (900)	4.3 (823)
South Pacific	0.54 (1,406)	0.037 (1,406)	1.7 (1,406)	0.016 (754)	3 (760)	5 (762)	3.0 (754)
Validation dataset (20 %)	0.93 (137,219)	0.066 (128,264)	3.0 (137,963)	0.019 (54,161)	7 (53,072)	10 (62,678)	5.1 (48,592)

In the last row, comparable information is provided for the validation dataset (Section Overall CANYON Performance), as reference. In brackets figures the number of observations used to compute each RMSE.

Research products

The second part of our AtlantOS work focused on investigating the extent to which new datasets can allow us to better understand the biological carbon pump.

First the synergy of satellite and in-situ data from the core-Argo and BGC-Argo networks was to investigate the global magnitude of the seasonal mixed-layer carbon pump (Dall'Olmo et al., 2016). This process refers to the organic carbon exported to the mesopelagic zone by the seasonal entrainment of surface waters in deeper layers (Figure 5). Our results demonstrated that the mixed-layer pump supplies an important seasonal flux of organic carbon to the mesopelagic zone. We estimated that this process is responsible for a global flux of 0.1–0.5 Pg C yr^{-1}. In high-latitude regions where the mixed layer is usually deep this flux amounts on average to 23%, but it can be greater than 100% of the carbon supplied by fast sinking particles (Figure 6). We conclude that the seasonal mixed-layer pump is an important source of organic carbon for the mesopelagic zone.

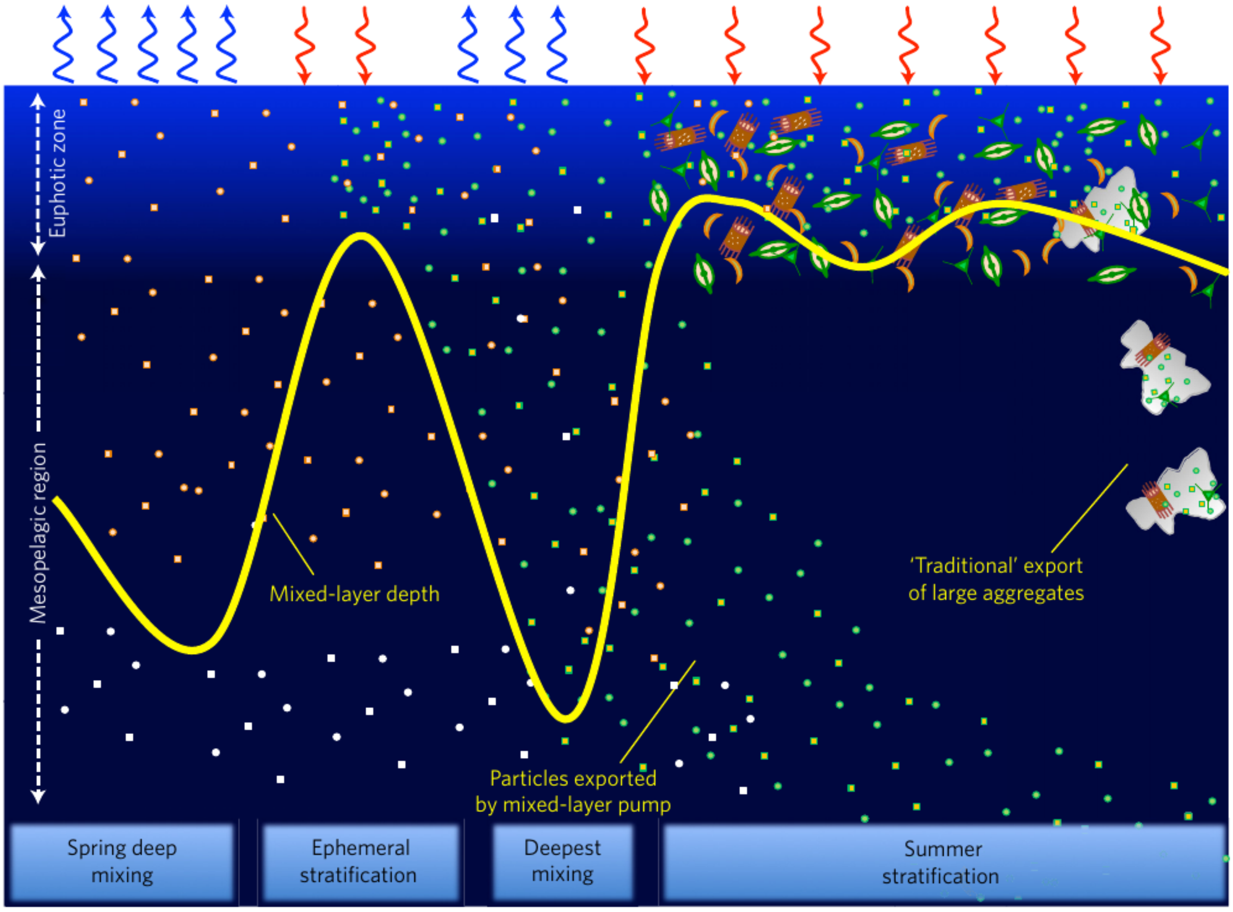

Figure 5. Schematic representation of the seasonal mixed-layer pump. In the spring the depth of the mixed layer (yellow line) reaches its maximum annual value. Before and during this event, ephemeral stratification events can occur due to, for example, intermittent changes in the heat flux from negative (out of the ocean, blue arrows) to positive (into the ocean, red arrows). These stratification events can result in new accumulation of organic matter, which is then redistributed over the water column by subsequent deep mixing. Eventually, when the summer stratification is established, the deeply mixed organic matter remains isolated below the sunlit layer, resulting in an export of carbon. Orange, white and green squares and circles represent small particles accumulated within and below the surface mixed layer during the previous summer, and produced due to the ephemeral stratification events, respectively.

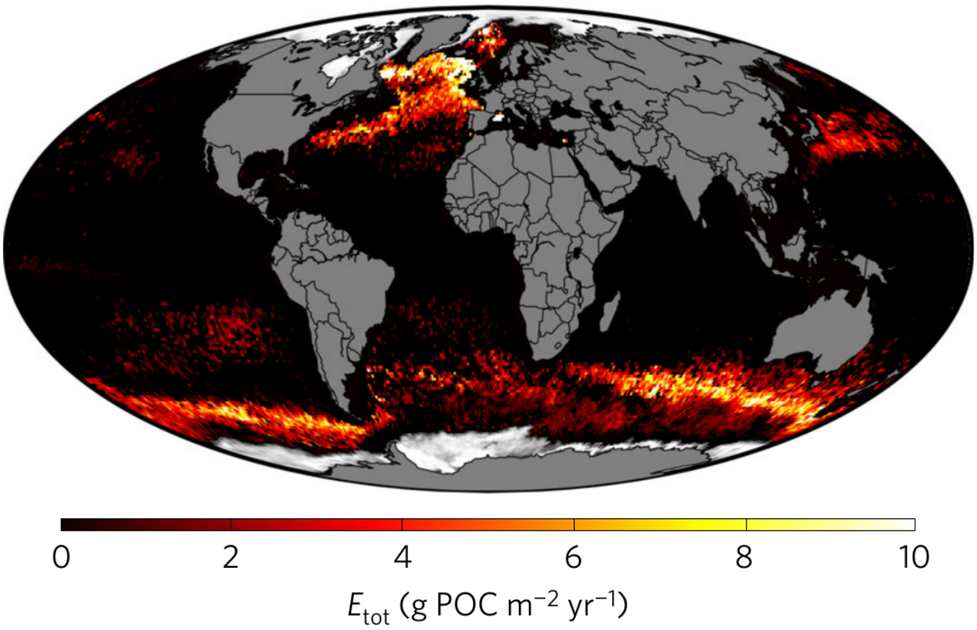

Figure 6. Global estimates of particulate carbon export by the mixed-layer pump (E_{tot}). Black colours refer to regions with winter mixed layers shallower than 100 m not considered in this study.

Second, in order to gain a more mechanistic understanding of the processes regulating the gravitational export of particles, we have focused on the most important seasonal biological event that injects particles in the mesopelagic region: the spring phytoplankton bloom. Our study area is the North Atlantic.

Data from BGC-Argo floats were analysed to detect and extract spikes in optical data (a proxy for large aggregates, Briggs et al., 2011) and to relate these spikes to the surface expression of the bloom (i.e., surface chlorophyll concentration). To this aim, we developed an automatic method to detect plumes of chlorophyll spikes at the time of the spring bloom (Figure 7).

Results show that export of large aggregates containing fresh algal material typically occur when the surface chlorophyll concentration begins to decline after its maximum surface concentration in the spring (Figure 8).

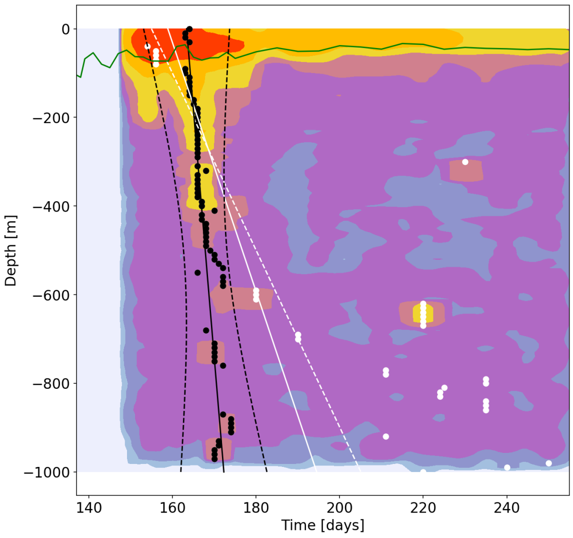

Figure 7. Example of depth vs. time contours of chlorophyll spikes extracted from a BGC-Argo float (WMO 6901523). Black points represent chlorophyll spikes that were considered part of the export plume. Black lines are the fit of the depth of the spikes vs. time (continuous) and related 95% confidence intervals (dashed): the slope of this fit is used as an estimate of the sinking rate of the large aggregates (Briggs et al., 2011). White points and lines are the automatically discarded outliers and fits that included the outliers. The green line is the depth of the euphotic zone and its intersection with the black continuous line is used as an estimate of the start of the export event.

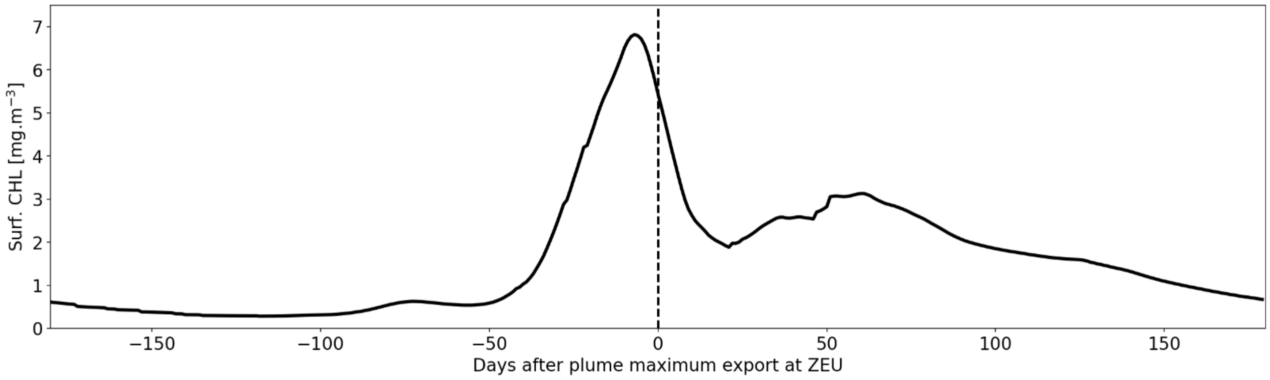

Figure 8. Plumes of fresh large aggregates occur after the chlorophyll reaches its peak in the spring in the North Atlantic. The black line is the average surface concentration from BGC-Argo floats that recorded measurable plumes of spikes at the time of the spring bloom. The x-axis is the time from the start of the export plume in days. The vertical dashed line is the time of the export plume.

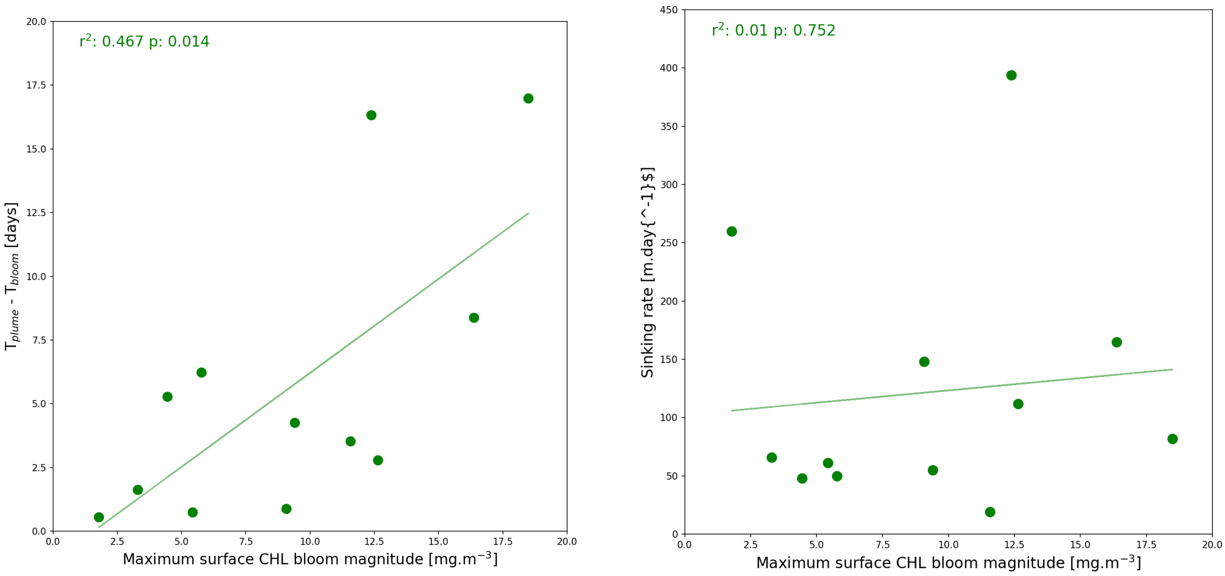

Figure 9. (Left) Delay between export event and time of chlorophyll maximum and (right) sinking rate vs. maximum surface chlorophyll concentration.

Specifically, plumes of fresh algal material are typically exported between 1 and 17 days after the chlorophyll reaches its maximum surface concentration (Figure 9 left). Moreover, our dataset suggests that the time between the chlorophyll maximum and the export plume is positively related to the magnitude of the chlorophyll maximum (Figure 9 left). On the other hand, no clear relationship was found between the maximum concentration of chlorophyll at the surface and the estimated sinking rates (Figure 9 right).

The last step of our analysis was to investigate the extent to which the intensity of the export flux can be predicted from surface measurements. Following Briggs et al. (2011), export rate was defined as the product of the sinking rate and spike intensity at the start of the export event. The depth at which the export flux is computed is the bottom of the euphotic zone. Among the different surface indicators analysed, the integral of the surface chlorophyll concentration from before the bloom to the time of the peak of the bloom showed the strongest relationship with the export flux (Figure 10): export rate was positively correlated to the integrated chlorophyll concentration.

In summary, this investigation demonstrated that, using surface data detectable from space, we can predict 1) when an export event of fresh algal material occur; and 2) the approximate strength of the export flux.

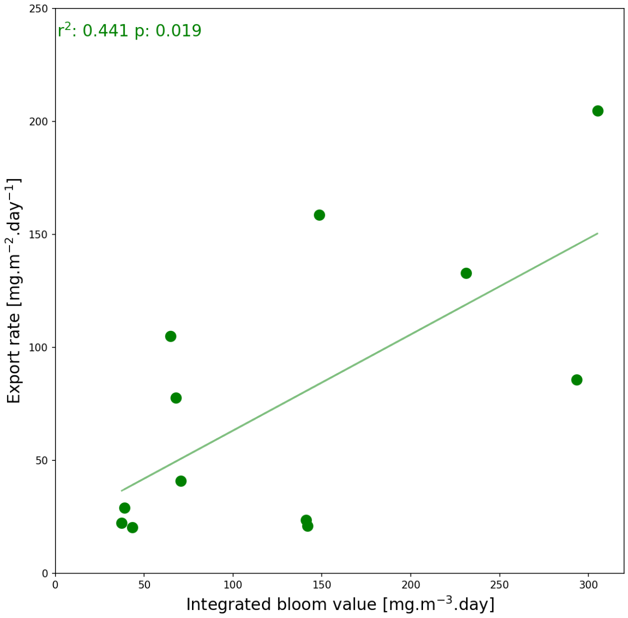

Figure 10. Rate of export of fresh aggregates as a function of the maximum surface chlorophyll concentration.

Concluding remarks

- Artificial neural networks are an important tool to exploit the large amounts of data collected by satellite and BGC-Argo floats in order to predict oceanic variables at depth from surface measurements.
- Artificial neural networks can also be exploited to estimate unmeasured sub-surface variables using data from BGC-Argo floats or other autonomous platforms.
- The potential of the synergy between satellite and BGC-Argo data was demonstrated by investigating the seasonal mixed-layer pump, a previously overlooked process that is part of the global biological carbon pump.
- High-resolution measurements from BGC-Argo floats were exploited to improve our understanding of the flux of fast sinking aggregates containing fresh algal material.

References acknowledging AtlantOS

Dall’Olmo, G.; Dingle, J.; Polimene, L.; Brewin, R. J. W. & Claustre, H. (2016) Substantial energy input to the mesopelagic ecosystem from the seasonal mixed-layer pump. *Nature Geoscience*, 9 (11), 820-823. doi:10.1038/ngeo2818.

Sauzède, R., H. Claustre, C. Jamet, J. Uitz, J. Ras, A. Mignot, and F. D'Ortenzio (2015). Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: A method based on a neural network with potential for global-scale applications. Journal of Geophysical Research-Oceans, 120(1), 451-470. doi:10.1002/2014JC010355

Sauzède, R., H. Claustre, J. Uitz, C. Jamet, G. Dall’Olmo, F. D’Ortenzio, B. Gentili, A. Poteau, and C. Schmechtig (2016), A neural network-based method for merging ocean color and Argo data to extend surface bio-optical properties to depth: Retrieval of the particulate backscattering coefficient, J. Geophys. Res. Oceans, 121, doi:10.1002/2015JC011408.

Sauzède, R., H.C. Bittig, H. Claustre, O. Pasqueron de Fommervault, J.P. Gattuso, L. Legendre, and K.S. Johnson (2017). Estimates of Water-Column Nutrient Concentrations and Carbonate System Parameters in the Global Ocean: A Novel Approach Based on Neural Networks. Frontiers in Marine Science, 4(128). doi:10.3389/fmars.2017.00128.

Other cited references

IOCCG (2011). Bio-Optical Sensors on Argo Floats. Claustre, H. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 11, IOCCG, Dartmouth, Canada

Briggs, Perry, Cetinić, Lee, D'Asaro, Grayb, Rehm (2011), High-resolution observations of aggregate flux during a sub-polar North Atlantic spring bloom. Deep Sea Research, 58(10), 1031-1039.