Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll a Concentrations

A. Mignot, F. d'Ortenzio, V. Taillandier, G. Cossarini, S. Salon

SPi	Journal Code				Article ID					Dispatch: 26.03.19	CE: RJS
	G	R	L		5	8	8	0	7	No. of Pages: 8	ME:

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL080541

Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll *a* Concentrations

A. Mignot[1,2], F. D'Ortenzio[1], V. Taillandier[1], G. Cossarini[3], and S. Salon[3]

[1]Sorbonne Universités, UPMC Univ. Paris 06, CNRS, UMR 7093, Laboratoire d'Océanographie de Villefranche (LOV), Villefranche-sur-mer, France, [2]Now at Mercator Ocean International, Ramonville-Saint-Agne, France, [3]Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Sgonico TS, Italy

Key Points:
- The triple collocation analysis is used to estimate biases and root-mean-square errors in BGC-Argo float data
- Adjusted data sets of oxygen, nitrate, and chlorophyll *a* concentrations are evaluated
- The errors and biases in adjusted data sets are less than or equal to 10%

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Mignot,
mignot@mercator-ocean.fr

Citation:
Mignot, A., D'Ortenzio, F., Taillandier, V., Cossarini, G., & Salon, S. (2019). Quantifying observational errors in Biogeochemical-Argo oxygen, nitrate, and chlorophyll *a* concentrations. *Geophysical Research Letters*, *46*. https://doi.org/10.1029/2018GL080541

Received 19 SEP 2018
Accepted 20 MAR 2019
Accepted article online 25 MAR 2019

Abstract Biogeochemical (BGC)-Argo floats observations are becoming a major data source for assimilation into and constraining of ocean biogeochemical models. An important prerequisite for a successful synthesis between models and observations is the characterization of the observational errors in BGC-Argo float data. The root-mean-square error and multiplicative and additive biases in quality-controlled data sets of oxygen, nitrate, and chlorophyll *a* concentrations collected with 17 BGC-Argo floats in the Mediterranean Sea between 2013 and 2017 are assessed using the triple collocation analysis. The analysis suggests that BGC-Argo float oxygen, nitrate and chlorophyll *a* data suffer from an additive bias of 2.9 ± 5.5 µmol/kg, 0.46 ± 0.07 µmol/kg, and -0.06 ± 0.02 mg/m^3, respectively. The root-mean-square error is evaluated at 5.1 ± 0.8 µmol/kg, 0.25 ± 0.07 µmol/kg, and 0.03 ± 0.01 mg/m^3. Additional studies should determine whether these values are applicable to the global ocean.

Plain Language Summary The Biogeochemical-Argo program is a network of ocean robots whose sensors monitor oxygen, nitrate, and chlorophyll *a* concentration information that is needed to detect decadal changes in biological carbon production, ocean acidification, ocean carbon uptake, and hypoxia in the world ocean. One of the goals of the Biogeochemical-Argo program is to incorporate these observations into ocean models to understand and forecast the changing state of the carbon cycle. The successful integration of the float data into numerical models, however, requires the specification of the observational errors. This study provides, for the first time, the biases and errors of the three cores variables of the Biogeochemical-Argo floats network: oxygen, nitrate, and chlorophyll *a* concentrations.

1. Introduction

Biogeochemical (BGC)-Argo floats are autonomous profiling platforms equipped with physical and biogeochemical sensors that can acquire time series of the vertical distribution of key physical, biological, and chemical variables at all sea conditions and over complete annual cycles (D'Ortenzio et al., 2014; Mignot et al., 2014, 2016, 2018). Over the last decade, the number of biogeochemical profiles acquired by these platforms has steadily increased. For instance, by 2011, the total number of BGC-Argo profiles (all parameters combined) collected in the global ocean approached ~45,000, and by 2017, there were almost 390,000 profiles (Argo, 2018).

Consequently, the BGC-Argo floats network is becoming a crucial ocean observing system to monitor, understand, and forecast the changing state of the ocean ecosystem (Biogeochemical-Argo Planning Group, 2016).

One of the goals of the BGC-Argo float network is to improve our knowledge about the biogeochemical state of the ocean by assimilating the float data into ocean biogeochemical models (Biogeochemical-Argo Planning Group, 2016). The assimilation of oxygen, nitrate, and chlorophyll *a* (O_2, NO_3^-, and Chl*a*) concentrations is particularly promising as O_2, NO_3^-, and Chl*a* are core state variables of most ocean biogeochemical models and they are the three parameters sampled most frequently by the floats. Such synthesis between float observations and model has the potential to improve our understanding of the ocean carbon cycle as well as its changing state.

A proper specification of observational errors is essential for an effective data assimilation because it controls the weight that the assimilated observations have on the model solution. However, little is known about the BGC-Argo O_2, NO_3^-, and Chl*a* error statistics in part because these observations are relatively recent and the errors provided by the sensors manufacturers were generally considered sufficient. With the dramatic augmentation of BGC-Argo observations and the growing interest in assimilating these data into numerical models, more rigorous approaches are needed.

The measurements of O_2, NO_3^-, and Chl*a* collected with the BGC-Argo floats, as any other measurement, are affected by systematic and random errors (Taylor, 1997). Systematic errors are a type of error that make measurements deviate from their true value by the same amount (hereinafter additive bias or offset) or the same fraction (hereinafter denoted multiplicative bias or gain) all the time. Systematic errors are in principle predictable and are introduced by imperfect initial calibration, sensor drift, or systematic changes in the environment during the measurement process. Random errors, on the other hand, are a type of error that shift measurements from their true value by a random amount. Random errors are unpredictable and are caused by unknown random changes in the measurement system or in the environment. Random errors are, in general, characterized by their standard deviation as they tend to cancel each other out, and their mean approaches 0.

A lot of effort has been made to correct the float data from systematic errors and improve the overall accuracy of the data collected with the BGC-Argo floats network (Johnson et al., 2017). This work has led to the emergence of quality control (QC) procedures that correct the systematic errors in the raw data acquired by the floats and produce adjusted data sets. On the contrary, random errors in float observations cannot be eliminated because of their randomness and unpredictability. However, with the objective and the need to make the full use of these observations, especially for data assimilation in ocean biogeochemical models, the accurate representation of random errors needs to be specified.

Random errors cannot be characterized by the methods that are traditionally used to evaluate systematic errors. Systematic errors are usually assessed by dual comparison with discrete sea water samples collected from a ship during the deployment of the floats. In these comparisons, the ship samples are assumed to be perfectly calibrated and to represent the “truth” and all discrepancies between the floats and the ship data are attributed to the float data. In reality, however, even though the ship measurements may be considered perfectly calibrated, they do include their own random errors. Consequently, the variability that can be observed between the ship and the float data does not necessarily result from the floats data alone. In this context, dual comparison with water samples collected from a ship is not adequate to quantify the random errors in BGC-Argo data sets and more elaborate tools are required.

The triple collocation (TC) analysis is a method for quantifying the random error standard deviation (hereinafter denoted root-mean-square error, RMSE) of three data sets of the same geophysical variable (Stoffelen, 1998) by combining the covariances between the data sets. The TC analysis requires three spatially and temporally collocated data sets of the same target variable with uncorrelated random errors, but it does not require a high precision data set. If one data set is assumed to be perfectly calibrated, the TC analysis can also provide the additive and multiplicative biases of the two other data sets (Stoffelen, 1998; Yilmaz & Crow, 2013). The TC analysis is a widespread method to estimate the RMSE in measurements and model predictions of oceanic (Caires, 2003; Janssen et al., 2007; O'Carroll et al., 2008; Ratheesh et al., 2013; Scott et al., 2014; Stoffelen, 1998), atmospheric (Alemohammad et al., 2015; Gao et al., 2012; Roebeling et al., 2012), and terrestrial variables (D'Odorico et al., 2014; Fang et al., 2013; Scipal et al., 2008).

The main limitation for the application of the TC analysis on measurements of O_2, NO_3^-, and Chl*a* collected with the BGC-Argo floats is to find two independent data sets collocated in space and in time. Obviously, the ship measurements of O_2, NO_3^-, and Chl*a* performed at the time of the float deployment can be used as a data set since the three parameters are estimated with different measurement systems from those performed on the floats. It should be noted at this point that the three data sets used in the TC analysis do not have to come from observations alone but can also be derived from models (Caires, 2003; Gao et al., 2012; Janssen et al., 2007; Ratheesh et al., 2013; Scipal et al., 2008). Therefore, vertical profiles of O_2, NO_3^-, and Chl*a* that are forecasted at high spatial and temporal resolution in some regions of the global ocean by operational biogeochemical model systems can be used as a third data set.

The Mediterranean Sea is an ideal region for the application of the TC analysis on BGC-Argo floats data because (1) there is one of the largest number of matchups between ship and floats thanks to the Bio-

Argo Med program (Taillandier et al., 2018) and (2) of the availability of an operational biogeochemical model system (MedBFM, Teruzzi et al., 2014).

In this study, we use the TC analysis to evaluate the RMSE in measurements of O_2, NO_3^-, and Chl*a* collected with the BGC-Argo floats deployed in the Mediterranean Sea between 2013 and 2017. We also evaluate the additive and multiplicative biases in the float data sets assuming that the ship data sets are perfectly calibrated.

The paper is organized as follows. Section 2 presents the data sets used in the study, and section 3 presents their TC. In section 4, the TC analysis is applied to estimate the RMSE and biases. Finally, section 5 summarizes the results and conclude the study. For the sake of brevity, the TC analysis is not reviewed in the main text but in the supplementary information.

2. Data

The TC analysis was performed on vertical profiles of O_2, NO_3^-, and Chl*a* acquired with 17 BGC-Argo floats from 2013 to 2017 in the Mediterranean Sea. We selected ascending profiles that were collocated together with a conductivity-temperature-depth (CTD)-rosette cast where discrete water samples were collected. The oceanographic stations typically occurred during float deployment and occasionally during float recovery. The ship-based estimates constitute the reference data set in our analysis. The third data set includes high spatial and temporal resolution outputs from an operational biogeochemical model system of the Mediterranean Sea.

2.1. Float Data Set

All floats used in this study were PROVOR CTS-4 profilers equipped with a Seabird SBE41 CPCTD sensor, a Satlantic OC4 radiometer measuring downwelling irradiance at 380, 412, and 490 nm and photosynthetically available radiation integrated between 400 and 700 nm, a WET Labs ECO-triplet comprising a Chl*a* fluorometer, a color dissolved organic matter fluorometer and a backscattering sensor at 700 nm. The backscattering, color dissolved organic matter, photosynthetically available radiation, and irradiance data were not used in this study. Eleven floats also included an AANDERAA optode oxygen sensor. Among these 11 floats, 9 floats were equipped with a Satlantic SUNA nitrate sensor.

The floats' nominal mission included profile measurements from a depth of 1,000 m up to the surface. The CTD and oxygen sensors sampling resolution was 1 m from the surface to 250 m, and 10 m from 250 to 1,000 m. The ECO-triplet sensor sampling resolution was 0.2 m from the surface to 10 m, 1 m from 10 to 250 m, and 10 m from 250 to 1,000 m. The nitrate sensor sampling resolution was 5 m from the surface to 30 m, 10 m from 30 to 250 m, and 25 m from 250 to 1,000 m. The floats typically emerge from the sea around local noon. In this study, only the first ascending profiles with a sampling depth greater than 800 m were used. The 800-m criterion was required to apply the QC procedures.

The float data were downloaded from the Argo Global Data Assembly Centre in France (Carval et al., 2017). The CTD and trajectory data were quality controlled using the standard Argo protocol (Wong et al., 2015). Following the BGC-Argo protocols (Johnson et al., 2016; Schmechtig et al., 2015; Thierry et al., 2016), the raw signals were transformed into the three states variables of interest: Chl*a*, O_2, and NO_3^-. All vertical profiles used in this study were visually inspected and we did not detect profiles obviously impacted by sensor failure. Finally, QC procedures were applied to adjust the raw data and improve the accuracy of the concentration estimates. The QC procedures applied on each variable are detailed in the supporting information.

2.2. Ship Data Set

The floats' vertical profiles used in this study were concomitant with a CTD-rosette cast where water samples were also collected to measure Chl*a*, O_2, and NO_3^-. We verified that the water masses sampled by the CTD-rosette casts corresponded those observed by the floats. The time lag between the CTD-rosette casts and the first ascending profiles was on average 23 hr. This lag mainly results from the fact that the floats were deployed at the end of every sampling station to prevent the floats from profiling under the ship (Taillandier et al., 2018). Furthermore, most of the floats performed one or two shallow dives to make sure that the float and the sensors operated properly. This lag of 23 hr is of the same order of magnitude than other studies that have made comparisons between ship-based and float measurements (Johnson et al., 2015). For the CTD-rosette casts that occur just after the float recovery, the lag was on average 1.6 hr.

A complete description of how the ship-based samples have been collected, measured, and processed can be found in Taillandier et al. (2018). Briefly, oxygen concentrations were measured by Winkler titration (Winkler, 1888) with potentiometric endpoint detection (Oudot et al., 1988), nitrate concentrations by a standard automated colorimetric system set up following Aminot and Kérouel (2007) using a Seal Analytical continuous flow AutoAnalyzer III. Finally, Chl*a* concentrations were estimated by high-performance liquid chromatography pigment analysis (Ras et al., 2008). As detailed in Taillandier et al. (2018), the data were acquired following best practice protocols and they were subjected to rigorous QC procedures so that they can be used as a reference data set to assess the accuracy of the BGC-Argo floats. Therefore, we can reasonably assume that the ship data set do not suffer from major systematic errors and can be considered as perfectly calibrated.

Occasionally, two rosette casts were performed within the same day and for the same float vertical profile. In that case, we first made sure that the two casts sampled the same water mass. Then, assuming that the biogeochemical temporal variability was longer than the time difference between the two casts (<1 day), the two casts were averaged into one.

2.3. Model Data Set

The third data set required to perform TC analysis consists of outputs from an operational biogeochemical model system of the Mediterranean Sea (MedBFM). The main characteristics of the MedBFM are given in the supporting information. The three-dimensional fields of the daily-mean Chl*a*, O_2, and NO_3^- in the Mediterranean Sea were downloaded from the Copernicus Marine Environmental Monitoring Service (Bolzon et al., 2017). We extracted the daily model outputs that were collocated in time and the closest to the CTD-rosette casts positions.

3. TC

The triplet matchups were generated by interpolating the float and model data to the sampling depth of the ship data. Following this interpolation, the O_2, NO_3^-, and Chl*a* data sets comprised 129, 108, and 117 triplets, respectively.

The covariances, and therefore the TC analysis, are very sensitive to outliers. Consequently, as proposed by Tukey (1977), values 1.5 times larger than the interquartile range of a given data set were discarded. This procedure removed 10 Chl*a* triplets.

The TC analysis requires a large number of triplets (~500; Zwieback et al., 2012) to be able to provide a robust estimate of the covariance, which was not the case in our analysis. Measurements derived from water samples collected from ships are time and cost consuming, and any statistical analysis using these measurements is generally realized with small samples. To assess the impact of having only a reduced number of observations on the covariance estimates, we followed the approach of Alemohammad et al. (2015). For a given state variable, 1,000 RMSEs, gains, and offsets were generated from 1,000 bootstrap simulations by sampling with replacement the available number of triplets (i.e., 129,108, or 107). The mean and the standard deviation of the bootstrap estimates were then computed. If the number of triplets is not too small, then the standard deviation will be much smaller than the mean. The standard deviations are also useful for comparing different estimators between data sets as it reflects their range of variability.

4. Results

In this section, we apply the TC analysis to estimate the RMSE and biases in O_2, NO_3^-, and Chl*a*.

4.1. Oxygen

Figures 1a and 1b show the comparison between ship, float, and model O_2. The additive and multiplicative biases resulting from the application of the TC analysis are indicated as regression lines. The mean and the standard deviation of the 1,000 bootstrap estimates of the biases and RMSEs are indicated in Table 1.

10.1029/2018GL080541

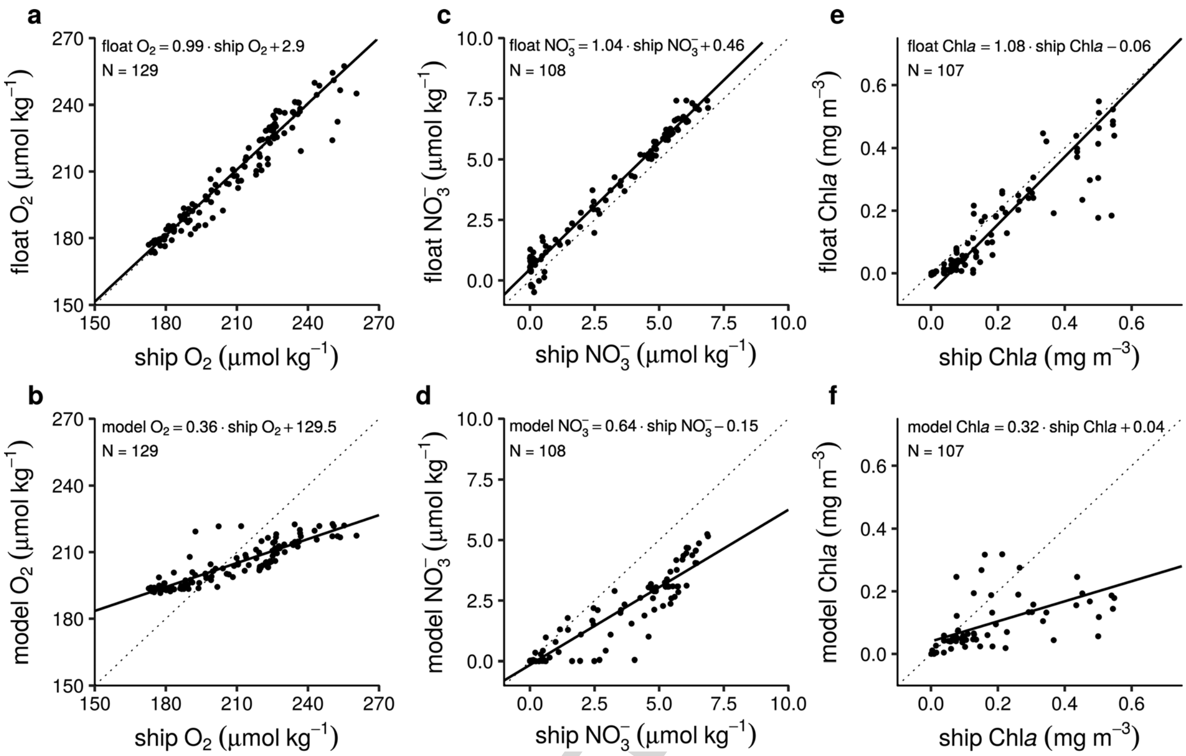

Figure 1. Scatter plots of float and model O_2, NO_3^-, and chlorophyll *a* (Chl*a*) as a function of ship data. (a) Float O_2 versus ship O_2. (b) Model O_2 versus ship O_2. (c) Float NO_3^- versus ship NO_3^-. (d) Model NO_3^- versus ship NO_3^-. (e) Float Chl*a* versus ship Chl*a*. (f) Model Chl*a* versus ship Chl*a*. The dashed lines represent the line 1:1. The plain lines are the linear regression lines between the two data sets considered in each panel. The coefficients of the linear regression lines (gain and offset) were computed using equations (6) and (7) in the supporting information.

Table 1
TC Estimates of Gain, Offset, RMSE, and Rescaled RMSE for the Ship, Float, Model O_2, NO_3^-, and Chla Data Sets

Variable	TC parameters	Data set		
		ship	float	model
O_2	gain	1	0.99 (0.03)	0.36 (0.01)
	offset (μmol/kg)	0	2.9 (5.5)	129.5 (2.5)
	RMSE (μmol/kg)	2.7 (1.0)	5.1 (0.8)	4.2 (0.6)
	rescaled RMSE (μmol/kg)		5.2 (0.9)	11.8 (1.7)
NO_3^-	gain	1	1.04 (0.02)	0.64 (0.02)
	offset (μmol/kg)	0	0.46 (0.07)	−0.15 (0.06)
	RMSE (μmol/kg)	0.25 (0.07)	0.25 (0.07)	0.55 (0.06)
	rescaled RMSE (μmol/kg)		0.24 (0.06)	0.86 (0.09)
Chl*a*	gain	1	1.08 (0.09)	0.32 (0.04)
	offset (mg/m^3)	0	−0.06 (0.02)	0.04 (0.01)
	RMSE (mg/m^3)	0.07 (0.01)	0.03 (0.01)	0.07 (0.01)
	rescaled RMSE (mg/m^3)		0.03 (0.02)	0.21 (0.03)

Note. The mean and standard deviation (given in parentheses) of the 1,000 bootstrap estimates are shown for each estimate, except for the biases and rescaled root-mean-square error (RMSE) for the ship data sets. The gains and offsets were computed using equations (6) and (7) in the supporting information. The RMSEs and rescaled RMSEs were computed using equations (5) and (8) in the supporting information. TC = triple collocation.

Overall, the ship and float O_2 data are in very good agreement with a gain not different from 1 and a positive offset of 2.9 μmol/kg. The model performs worse than the floats in estimating O_2. The model tends to overestimate O_2 at low concentrations and to underestimate O_2 at high concentrations. The model data set shows a gain of 0.36 and a positive offset of 129.5 μmol/kg.

The RMSEs estimated for the ship, float, and model are 2.7, 5.1, and 4.23 μmol/kg. In order to compare the three data sets with each other, the RMSE estimates must be rescaled within the ship data space using equation (8) in the supporting information. This rescaling yields to float and model RMSEs of 5.2 and 11.8 μmol/kg, respectively, suggesting that oxygen concentrations are more precisely determined by Winkler titration than the oxygen sensors mounted on the floats or model predictions.

The standard deviations of the bootstrap estimates are at least a factor of 3 smaller than the mean for all estimates except for the offset in float O_2. In this case, the standard deviation is 3 times larger than the mean. We can then conclude that the TC estimates of RMSEs and biases for the oxygen concentrations are overall robust. However, more triplets are necessary to precisely determine the additive bias in the float O_2 data set.

4.2. Nitrate

The TC estimates for NO_3^- are presented in Table 1, and scatter plots are in Figures 1 c and 1d.

The ship and float NO_3^- data are in good agreement with a gain of 1.04 and a positive offset of 0.46 μmol/kg. On the other hand, the model always underestimates NO_3^- when compared to the ship estimates. The model data set shows a gain of 0.64 and a negative offset of 0.15 μmol/kg.

The TC estimates of RMSE for the ship, float, and model are 0.25, 0.25, and 0.55 μmol/kg, respectively. The float rescaled RMSE is similar to the ship RMSE, suggesting that the float and the ship NO_3^- data sets are equally precise.

The standard deviations of the bootstrap estimates are at least 2 to 3 times smaller than the mean for all estimates, indicating that the TC estimates for the ship, float, and model NO_3^- data sets are independent of the choices of the triplet.

4.3. Chl*a*

The TC estimates for Chl*a* are presented in Table 1, and scatter plots are in Figures 1e and 1f.

The ship and float data sets are in favorable agreement with a gain of 1.08 and a negative offset of 0.06 mg/m^3. The model predictions, on the other hand, strongly underestimate the Chl*a* concentrations when compared to the ship measurements. The model data set shows a gain of 0.32 and a negative offset of 0.04 mg/m^3.

The RMSEs in the ship, float, and model data sets are 0.07, 0.03, and 0.07 mg/m^3, respectively. Rescaling the float and model RMSEs reveals that Chl*a* concentrations are more precisely determined by fluorometers mounted on the floats than by ship-based measurements or predictions from the model.

The standard deviations of the bootstrap estimates are at least a factor of 3 smaller than the means, suggesting that the sampling size error is not important.

5. Conclusion

In this study, we used the TC analysis to estimate the RMSE and multiplicative and additive biases in data sets of oxygen, nitrate, and Chl*a* concentrations collected with 17 BGC-Argo floats in the Mediterranean Sea between 2013 and 2017. The accuracy of the float data was adjusted using dedicated QC procedures before applying the TC analysis. Collocations were made using ship-based measurements performed at the time of the float deployment as the reference data set and outputs from an operational biogeochemical model of the Mediterranean Sea as the third data set.

The TC analysis reveals that the performance of the quality-controlled data sets is fairly good. The float O_2, NO_3^-, and Chl*a* data sets have an additive bias of 2.9 μmol/kg, 0.46 μmol/kg and −0.06 mg/m^3, respectively, and all three data sets have a multiplicative bias close to 1. The RMSE in the O_2, NO_3^-, and Chl*a* data sets were evaluated at 5.1 μmol/kg, 0.25 μmol/kg, and 0.03 mg/m^3. Normalized to the range of each data set (supporting information Table 1), these estimates correspond to relative biases of 3.4%, 5.8%, and −10.7% and a relative error of 6.1%, 3.2%, and 5.4%.

The analysis presented in this study proposes a theoretical framework to appreciate and evaluate the quality of the BGC-Argo QC procedures. These methods are continuously evolving by introducing correction factors or improved data processing. The analysis presented here could be particularly efficient to highlight and quantify the amelioration introduced by a new QC protocol, by estimating the modification on the biases and the RMSE.

Moreover, a proper estimation of the observational errors is crucial for data assimilation as it determines, together with the model error, how much the model solution is modified at each assimilation step. Until now, little was known about these errors. This limitation has certainly dampened the utilization of these data by the BGC modeling community. For example, in a recent BGC-Argo data assimilation experiment, Cossarini et al. (2018) has proposed an objective procedure for tuning the BGC-Argo Chl*a* random errors in the Mediterranean Sea. They found a similar value than our estimate, but their method required substantial computational efforts, which might be not an optimal solution for operational activities. Therefore, this study should strengthen the exploitation of the BGC-Argo data by the modeling community.

The main limitation in our study is the sample size, due to the scarcity of matchups between float and ship. The impact of having only a limited number of observations was assessed with a bootstrap analysis, and we found that the sampling errors were overall small and only impact the retrieval of the float O_2 data set offset.

The overall number of match ups will increase with time as the number of floats deployed grows and ship samples are collected at the deployment more and more frequently. Therefore, we can expect that the sampling error will be reduced in future studies.

Our biases estimates for O_2 and NO_3^- are close to the values found in the Southern Ocean by Johnson et al. (2017; i.e., 3.2 and −0.5 μmol/kg, respectively). However, the RMSE and biases estimated in this study were evaluated in the Mediterranean Sea. The range of oxygen, nitrate, and chlorophyll concentrations being smaller than in the global ocean, the utilization of these values in other regions should require further verification. Q7

Acknowledgments
This work was carried out as part of the Copernicus Marine Environment Monitoring Service MASSIMILI project. This work is also a contribution to the French "Equipement d'avenir" NAOS project (ANR J11R107-F). The BGC-Argo float data can be downloaded from http://doi.org/10.17882/42182 and were collected and made freely available by the International Argo Program and the national programs that contribute to it (http://www.argo.ucsd.edu, http://argo.jcommops.org). The Argo program is part of the Global Ocean observing System. The model data can be downloaded from the Copernicus Marine Environmental Monitoring Service (ftp://cmems-med-mfc.eu:2121/Core/V2/MEDSEA_ANALYSIS_FORECAST_BIO_006_006/, ftp://cmems-med-mfc.eu:2121/Core/V2/MEDSEA_ANALYSIS_FORECAST_PHYS_006_001/). The ship data are available at http://www.seanoe.org/data/00405/51678/. We want to thank the crews of the French R/V Le Suroit and R/V Tethys-II for their invaluable support during the field works. The French MOOSE and MERMEX projects are also gratefully thanked. We express our gratitude to Antoine Poteau and Catherine Schmechtig for their help in the processing of the BGC-Argo float data. The authors also acknowledge the valuable comments and suggestions made by three anonymous reviewers.

References Q8

Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., & Stoffelen, A. (2015). Characterization of precipitation product errors across the United States using multiplicative triple collocation. *Hydrology and Earth System Sciences*, *19*(8), 3489–3503. https://doi.org/10.5194/hess-19-3489-2015

Aminot, A., & Kérouel, R. (2007). *Dosage automatique des nutriments dans les eaux marines: méthodes en flux continu*. France: Editions Ifremer.

Argo. (2018). Argo float data and metadata from global data assembly Centre (Argo GDAC). SEANOE. https://doi.org/10.17882/42182

Biogeochemical-Argo Planning Group. (2016). The scientific rationale, design and implementation plan for a Biogeochemical-Argo float array (p.). https://doi.org/10.13155/46601

Bolzon, G., Cossarini, G., Lazzari, P., Salon, S., Teruzzi, A., Crise, A., & Solidoro, C. (2017). Mediterranean Sea biogeochemical analysis and forecast (CMEMS MED AF-Biogeochemistry 2013-2017)". [Data set]. Copernicus Monitoring Environment Marine Service. https://doi.org/10.25423/MEDSEA_ANALYSIS_FORECAST_BIO_006_006

Caires, S. (2003). Validation of ocean wind and wave data using triple collocation. *Journal of Geophysical Research*, *108*(C3), 3098. https://doi.org/10.1029/2002JC001491

Carval, K., Keeley, R, Takatsuki, Y., Yoshida, T., Schmid, C., Goldsmith, R., et al. (2017). Argo user's manual V3.2. https://doi.org/10.13155/29825

Cossarini, G., Lazzari, P., & Solidoro, C. (2015). Spatiotemporal variability of alkalinity in the Mediterranean Sea. *Biogeosciences*, *12*(6), 1647–1658. https://doi.org/10.5194/bg-12-1647-2015 Q9

Cossarini, G., Mariotti, L., Feudale, L., Mignot, A., Salon, S., Taillandier, V., et al. (2018). Towards operational 3D-Var assimilation of chlorophyll Biogeochemical-Argo float data into a biogeochemical model of the Mediterranean Sea. *Ocean Modelling*. https://doi.org/10.1016/j.ocemod.2018.11.005

Dobricic, S., & Pinardi, N. (2008). An oceanographic three-dimensional variational data assimilation scheme. *Ocean Modelling*, *22*(3–4), 89–105. https://doi.org/10.1016/j.ocemod.2008.01.004 Q10

D'Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E., & Schaepman, M. E. (2014). Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. *Remote Sensing of Environment*, *142*, 141–154. https://doi.org/10.1016/j.rse.2013.12.005

D'Ortenzio, F., Lavigne, H., Besson, F., Claustre, H., Coppola, L., Garcia, N., et al. (2014). Observing mixed layer depth, nitrate and chlorophyll concentrations in the northwestern Mediterranean: A combined satellite and NO_3 profiling floats experiment. *Geophysical Research Letters*, *41*, 6443–6451. https://doi.org/10.1002/2014GL061020

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J. M., et al. (2013). Characterization and intercomparison of global moderate resolution leaf area index (LAI) products: Analysis of climatologies and theoretical uncertainties: INTERCOMPARISON OF GLOBAL LAI PRODUCTS. *Journal of Geophysical Research: Biogeosciences*, *118*, 529–548. https://doi.org/10.1002/jgrg.20051

Foujols, M.-A., Lévy, M., Aumont, O., & Madec, G. (2000). OPA 8.1 Tracer Model reference manual [Publication - Report]. Retrieved December 28, 2017, from http://nora.nerc.ac.uk/id/eprint/164848/ Q11

Gao, F., Zhang, X., Jacobs, N. A., Huang, X.-Y., Zhang, X., & Childs, P. P. (2012). Estimation of TAMDAR observational error and assimilation experiments. *Weather and Forecasting*, *27*(4), 856–877. https://doi.org/10.1175/WAF-D-11-00120.1

Garcia, H. E., Boyer, T. P., Locarnini, R. A., Antonov, J. I., Mishonov, A. V., Baranova, O. K., et al. (2014). *World Ocean atlas 2013. Volume 3, dissolved oxygen, apparent oxygen utilization, and oxygen saturation*. Q12 Q13

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series*, *148*(1–3), 187–200. https://doi.org/10.3354/meps148187 Q14

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., & Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation*, *45*, 200–211. https://doi.org/10.1016/j.jag.2015.09.002 Q15

Janssen, P. A. E. M., Abdalla, S., Hersbach, H., & Bidlot, J.-R. (2007). Error estimation of buoy, satellite, and model wave height data. *Journal of Atmospheric and Oceanic Technology*, *24*(9), 1665–1677. https://doi.org/10.1175/JTECH2069.1

Johnson, K, Pasqueron De Fommervault, O., Serra, R., D'Ortenzio, F., Schmechtig, C., Claustre, H., & Poteau, A. (2016). Processing Bio-Argo nitrate concentration at the DAC Level (p.). https://doi.org/10.13155/46121

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., et al. (2017). Biogeochemical sensor performance in the SOCCOM profiling float array: SOCCOM BIOGEOCHEMICAL SENSOR PERFORMANCE. *Journal of Geophysical Research: Oceans*, *122*, 6416–6436. https://doi.org/10.1002/2017JC012838

Johnson, K. S., Plant, J. N., Riser, S. C., & Gilbert, D. (2015). Air oxygen calibration of oxygen optodes on a profiling float Array. *Journal of Atmospheric and Oceanic Technology*, *32*(11), 2160–2172. https://doi.org/10.1175/JTECH-D-15-0101.1

Körtzinger, A., Schimanski, J., & Send, U. (2005). High quality oxygen measurements from profiling floats: A promising new technique. *Journal of Atmospheric and Oceanic Technology*, *22*(3), 302–308. https://doi.org/10.1175/JTECH1701.1 Q16

Lazzari, P., Solidoro, C., Ibello, V., Salon, S., Teruzzi, A., Béranger, K., et al. (2012). Seasonal and inter-annual variability of plankton chlorophyll and primary production in the Mediterranean Sea: A modelling approach. *Biogeosciences*, *9*(1), 217–233. https://doi.org/10.5194/bg-9-217-2012 Q17

Lazzari, P., Solidoro, C., Salon, S., & Bolzon, G. (2016). Spatial variability of phosphate and nitrate in the Mediterranean Sea: A modeling approach. *Deep Sea Research Part I: Oceanographic Research Papers*, *108*, 39–52. https://doi.org/10.1016/j.dsr.2015.12.006 Q18

Mignot, A., Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., & Xing, X. (2014). Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: A Bio-Argo float investigation. *Global Biogeochemical Cycles*, *28*, 856–876. https://doi.org/10.1002/2013GB004781

Mignot, A., Ferrari, R., & Claustre, H. (2018). Floats with bio-optical sensors reveal what processes trigger the North Atlantic bloom. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02143-6

Mignot, A., Ferrari, R., & Mork, K. A. (2016). Spring bloom onset in the Nordic Seas. *Biogeosciences*, *13*(11), 3485–3502. https://doi.org/10.5194/bg-13-3485-2016

O'Carroll, A. G., Eyre, J. R., & Saunders, R. W. (2008). Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations. *Journal of Atmospheric and Oceanic Technology*, *25*(7), 1197–1207. https://doi.org/10.1175/2007JTECHO542.1

Oddo, P., Adani, M., Pinardi, N., Fratianni, C., Tonani, M., & Pettenuzzo, D. (2009). A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting. *Ocean Science*, *5*(4), 461–473. https://doi.org/10.5194/os-5-461-2009 Q19

Oudot, C., Gerard, R., Morin, P., & Gningue, I. (1988). Precise shipboard determination of dissolved oxygen (Winkler procedure) for productivity studies with a commercial system1. *Limnology and Oceanography*, *33*(1), 146–150. https://doi.org/10.4319/lo.1988.33.1.0146

Ras, J., Claustre, H., & Uitz, J. (2008). Spatial variability of phytoplankton pigment distributions in the subtropical South Pacific Ocean: Comparison between in situ and predicted data. *Biogeosciences*, *5*(2), 353–369. https://doi.org/10.5194/bg-5-353-2008

Ratheesh, S., Mankad, B., Basu, S., Kumar, R., & Sharma, R. (2013). Assessment of satellite-derived sea surface salinity in the Indian Ocean. *IEEE Geoscience and Remote Sensing Letters*, *10*(3), 428–431. https://doi.org/10.1109/LGRS.2012.2207943

Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., & Leijnse, H. (2012). Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data. *Journal of Hydrometeorology*, *13*(5), 1552–1566. https://doi.org/10.1175/JHM-D-11-089.1

Roesler, C., Uitz, J., Claustre, H., Boss, E., Xing, X., Organelli, E., et al. (2017). Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. *Limnology and Oceanography: Methods*, *15*(6), 572–585. https://doi.org/10.1002/lom3.10185 Q20

Sauzède, R., Bittig, H. C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J.-P., Legendre, L., & Johnson, K. S. (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*. https://doi.org/10.3389/fmars.2017.00128 Q21

Schmechtig, C., Poteau, A., Claustre, H., D'Ortenzio, F., & Boss, E. (2015). *Processing bio-Argo chlorophyll-A concentration at the DAC level*. Ifremer. https://doi.org/10.13155/39468 Q22

Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., & Wagner, W. (2008). A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophysical Research Letters*, *35*, L24403. https://doi.org/10.1029/2008GL035599

Scott, K. A., Buehner, M., & Carrieres, T. (2014). An assessment of sea-ice thickness along the Labrador coast from AMSR-E and MODIS data for operational data assimilation. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(5), 2726–2737. https://doi.org/10.1109/TGRS.2013.2265091

Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, *103*(C4), 7755–7766. https://doi.org/10.1029/97JC03180

Storto, A., Masina, S., & Dobricic, S. (2014). Estimation and impact of nonuniform horizontal correlation length scales for global ocean physical analyses. *Journal of Atmospheric and Oceanic Technology*, *31*(10), 2330–2349. https://doi.org/10.1175/JTECH-D-14-00042.1 Q23

Taillandier, V., Wagener, T., D'Ortenzio, F., Mayot, N., Legoff, H., Ras, J., et al. (2018). Hydrography and biogeochemistry dedicated to the Mediterranean BGC-Argo network during a cruise with RV *Tethys 2* in May 2015. *Earth System Science Data*, *10*(1), 627–641. https://doi.org/10.5194/essd-10-627-2018

Takeshita, Y., Martz, T. R., Johnson, K. S., Plant, J. N., Gilbert, D., Riser, S. C., et al. (2013). A climatology-based quality control procedure for profiling float oxygen data: QC procedure for profiling float oxygen. *Journal of Geophysical Research: Oceans*, *118*, 5640–5650. https://doi.org/10.1002/jgrc.20399 Q24

Taylor, J. (1997). *Introduction to error analysis, the study of uncertainties in physical measurements*, (2nd ed.). 648 Broadway, Suite 902, New York, NY 10012. Retrieved from: University Science Books. http://adsabs.harvard.edu/abs/1997ieas.book.....T

Teruzzi, A., Dobricic, S., Solidoro, C., & Cossarini, G. (2014). A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties: 3D-VAR IN BIOGEOCHEMICAL MODELS. *Journal of Geophysical Research: Oceans*, *119*, 200–217. https://doi.org/10.1002/2013JC009277

Thierry, V., Bittig, H., Gilbert, D., Kobayashi, T., Kanako, S., & Schmid, C. (2016). Processing Argo oxygen data at the DAC level cookbook (p.). https://doi.org/10.13155/39795

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley Publishing Company.

Winkler, L. W. (1888). Die Bestimmung des im Wasser gelösten Sauerstoffes. *Berichte der Deutschen Chemischen Gesellschaft*, *21*(2), 2843–2854. https://doi.org/10.1002/cber.188802102122

Wong, A., Keeley, Robert, Carval, Thierry, & Argo Data Management Team (2015). Argo quality control manual for CTD and trajectory data. https://doi.org/10.13155/33951

Xing, X., Claustre, H., Blain, S., D'Ortenzio, F., Antoine, D., Ras, J., & Guinet, C. (2012). Quenching correction for in vivo chlorophyll fluorescence acquired by autonomous platforms: A case study with instrumented elephant seals in the Kerguelen region (Southern Ocean): Quenching correction for chlorophyll fluorescence. *Limnology and Oceanography: Methods*, *10*(7), 483–495. https://doi.org/10.4319/lom.2012.10.483 Q25

Yilmaz, M. T., & Crow, W. T. (2013). The optimality of potential rescaling approaches in land data assimilation. *Journal of Hydrometeorology*, *14*(2), 650–660. https://doi.org/10.1175/JHM-D-12-052.1

Zwieback, S., Scipal, K., Dorigo, W., & Wagner, W. (2012). Structural and statistical properties of the collocation technique for error characterization. *Nonlinear Processes in Geophysics*, *19*(1), 69–80. https://doi.org/10.5194/npg-19-69-2012