AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2016JC011833

Special Section:
Dense Water Formations in the North Western Mediterranean: From the Physical Forcings to the Biogeochemical Consequences

Long-term monitoring of ocean deep convection using multisensors altimetry and ocean color satellite data

Marine Herrmann[1], Pierre-Amael Auger[2], Caroline Ulses[3], and Claude Estournel[3]

[1]LEGOS, Université de Toulouse, IRD, CNES, CNRS, UPS, Toulouse, France, [2]Instituto Milenio de Oceanogriafa and Escuela de Ciencias del Mar, Pontificia Universidad Catolica de Valparaiso, Valparaiso, Chile, [3]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse, France

Abstract Deep convection occurs in oceanic regions submitted to strong atmospheric buoyancy losses and results in the formation of deep water masses (DWF) of the ocean circulation. It shows a strong interannual variability, and could drastically weaken under the influence of climate change. In this study, a method is proposed to monitor quantitatively deep convection using multisensors altimetry and ocean color satellite data. It is applied and evaluated for the well-observed Northwestern Mediterranean Sea (NWMS) case study. For that, a coupled hydrodynamical-biogeochemical numerical simulation is used to examine the signature of DWF on sea level anomaly (SLA) and surface chlorophyll concentration. Statistically significant correlations between DWF annual indicators and the areas of low surface chlorophyll concentration and low SLA in winter are obtained, and linear relationships between those indicators and areas are established. These relationships are applied to areas of low SLA and low chlorophyll concentration computed, respectively, from a 27 year altimetry data set and a 19 year ocean color data set. The first long time series (covering the last 2 decades) of DWF indicators obtained for the NWMS from satellite observations are produced. Model biases and smoothing effect induced by the low resolution of gridded altimetry data are partly taken into account by using corrective methods. Comparison with winter atmospheric heat flux and previous modeled and observed estimates of DWF indicators suggests that those DWF indicators time series capture realistically DWF interannual variability in the NWMS. The advantages as well as the weaknesses and uncertainties of the method are finally discussed.

Key Points:
- Quantitative monitoring of deep convection using multisensors altimetry and ocean color satellite data
- Hydrodynamical-biogeochemical model to establish linear relations relationships between sea level, surface chlorophyll, and deep convection
- Our long time series captures correctly the interannual variability of deep convection in the Mediterranean Sea

Correspondence to:
M. Herrmann,
marine.herrmann@ird.fr

Citation:
Herrmann, M., P.-A. Auger, C. Ulses, and C. Estournel (2017), Long-term monitoring of ocean deep convection using multisensors altimetry and ocean color satellite data, *J. Geophys. Res. Oceans*, *122*, 1457–1475, doi:10.1002/2016JC011833.

Received 24 MAR 2016
Accepted 25 JAN 2017
Accepted article online 31 JAN 2017
Published online 25 FEB 2017

1. Introduction

Open-ocean deep convection occurs in a few regions of the world ocean submitted to strong surface buoyancy losses that induce an increase of sea surface water density, resulting in the vertical mixing of the water column. It is at the origin of the formation of deep water masses of the ocean circulation [*Marshall and Schott*, 1999]. It shows a strong interannual variability [*Yashayaev*, 2007; *Herrmann et al*., 2010]: yearly maximum of mixed layer depth (MLD) varies from very shallow to intermediate and even sea bottom values. Both observational and modeling studies suggested that deep convection could drastically weaken under the influence of climate change [*Somot et al*., 2006; *de Lavergne et al*., 2014].

Northwestern Mediterranean Sea (NWMS) is a region of dense water formation (DWF) that can be considered as a golden case study, due to its easier access compared to other convection regions (e.g., Labrador and Greenland seas). DWF occurs there in winter under the influence of cold northerly winds and results in the formation of Western Mediterranean Deep Water, one of the main water masses of the Mediterranean thermohaline circulation. DWF in the NWMS does not only play an important role in the hydrodynamical functioning of the Mediterranean sea, it also influences the ecosystems: the associated winter vertical mixing is responsible for the nutrients enrichment of the surface layer, and therefore contributes to the following spring bloom [*Herrmann et al*., 2013, 2014]. As a result, NWMS is one of the biologically most productive areas of the Mediterranean sea [*Bosc et al*., 2004].

Understanding and monitoring the interannual variability and long-term evolution of DWF is essential for studies of ocean circulation and ecosystems. Long-term in situ monitoring of deep convection is costly, we therefore explore the possibility to monitor DWF using satellite data. First, as can be seen in altimetry data

during the strong NWMS convection winter of 2005 (Figure 1), DWF is associated with a lowering of sea surface due both to an increase of the water density (steric effect) and to an activation of the cyclonic circulation (dynamic effect) [*Herrmann et al*., 2008]. This influence of DWF on sea surface level was the starting point of studies that proposed methods to monitor DWF using altimetry satellite data [*Herrmann et al*., 2009; *Gelderloos et al*., 2013]. For the NWMS, *Herrmann et al*. [2009] used a numerical oceanic simulation performed over the Mediterranean Sea to establish a relationship between sea level and DWF, and applied this relationship to along-track altimetry data. Second, DWF is also associated with strong vertical displacements that induce the decrease of surface chlorophyll concentration [*Herrmann et al*., 2013]. Since phytoplankton cannot stay stably in the surface euphotic layer where photosynthesis can occur, primary production indeed stops (light limitation effect). Moreover, the chlorophyll stock initially present in the surface layer is vertically diluted throughout the whole mixed column (dilution effect) [*Auger et al*., 2014]. DWF consequently has a signature on surface chlorophyll concentration that can be observed on ocean color satellite data (see for example the strong convection winter of 2005, Figure 1), suggesting that those data could be used to detect and monitor DWF.

Several authors attempted to use the chlorophyll depleted area estimated from satellite data as an indicator of deep convection intensity. These studies focused on individual cases or short time series of DWF. *Herrmann et al*. [2010] used it qualitatively to estimate the ability of their model to represent correctly the spatial extension of convection for winter 2005. *Somot et al*. [2016] used this area as an indicator of DWF intensity for winters 2007–2013. Some authors multiplied empirically this area by the bottom depth (∼2200 m) to provide estimates of the volume of dense water formed during respectively winter 2012 [*Durrieu de Madron et al*., 2013] and winters 2007–2013 [*Houpert et al*., 2016]. These latter studies therefore assumed that when convection occurs, it reaches the bottom. Their method is therefore only suitable for cases of either null or bottom convection, but not of intermediate convection.

Here based on the results of a coupled hydrodynamical-biogeochemical ocean simulation, we propose a method to monitor annual DWF intensity on the long term using both altimetry and ocean color satellite observations. We apply and test this method for the NWMS. The numerical tool and satellite data sets as well as the existing estimations of DWF rates are presented in section 2. We use the model to establish significant linear relationships between the DWF intensity on one side and the sea level anomaly (SLA) and surface chlorophyll concentration on the other side (section 3). We apply those linear equations to SLA and surface chlorophyll concentration obtained from real satellite data. This allows us to produce long-term time series of annual DWF intensity in terms of volume of mixed water, volume of newly formed dense water and mixed layer depth (section 4). Advantages and weaknesses of our method and uncertainties associated with those time series are discussed in section 5. Concluding remarks are given in section 6.

2. Methods and Tools

2.1. The Numerical Simulations

A 38 year hydrodynamical simulation was performed at 2.5 km resolution over the western Mediterranean (0°40′W–11°40′E; 36°25′N–44°25′N, see Figure 2) for the period 1975–2013 with the 3-D primitive equations, sigma-coordinate (40 levels), free surface ocean model SYMPHONIE [*Marsaleix et al*., 2009]. The explicit nonlinear free surface scheme described in *Marsaleix et al*. [2008] is used. The 2.5 km resolution enables to reproduce realistically NWMS deep convection and associated mesoscale structures [*Herrmann et al*., 2008]. The model was initialized and forced at the lateral boundaries by the results of a Mediterranean basin scale simulation performed with the NEMOMED8 model [*Herrmann et al*., 2010] and at the surface by the atmospheric fluxes of the ARPERA data set [*Herrmann and Somot*, 2008].

A twin tridimensional biogeochemical 38 year simulation was performed for the same period with the biogeochemical model Eco3M-S forced by the results of the hydrodynamical SYMPHONIE simulation. This biogeochemical simulation is described in details in *Auger et al*. [2014] and the hydrodynamical and biogeochemical simulations were examined and validated by *Auger et al*. [2014] and *Ulses et al*. [2016]. This coupled hydrodynamical-biogeochemical tool was also used to study the impact of interannual variability and long-term evolution of atmospheric and oceanic conditions, in particular deep convection, on the NWMS pelagic planktonic ecosystem and associated carbon cycle [*Herrmann et al*., 2013, 2014; *Ulses et al*.,

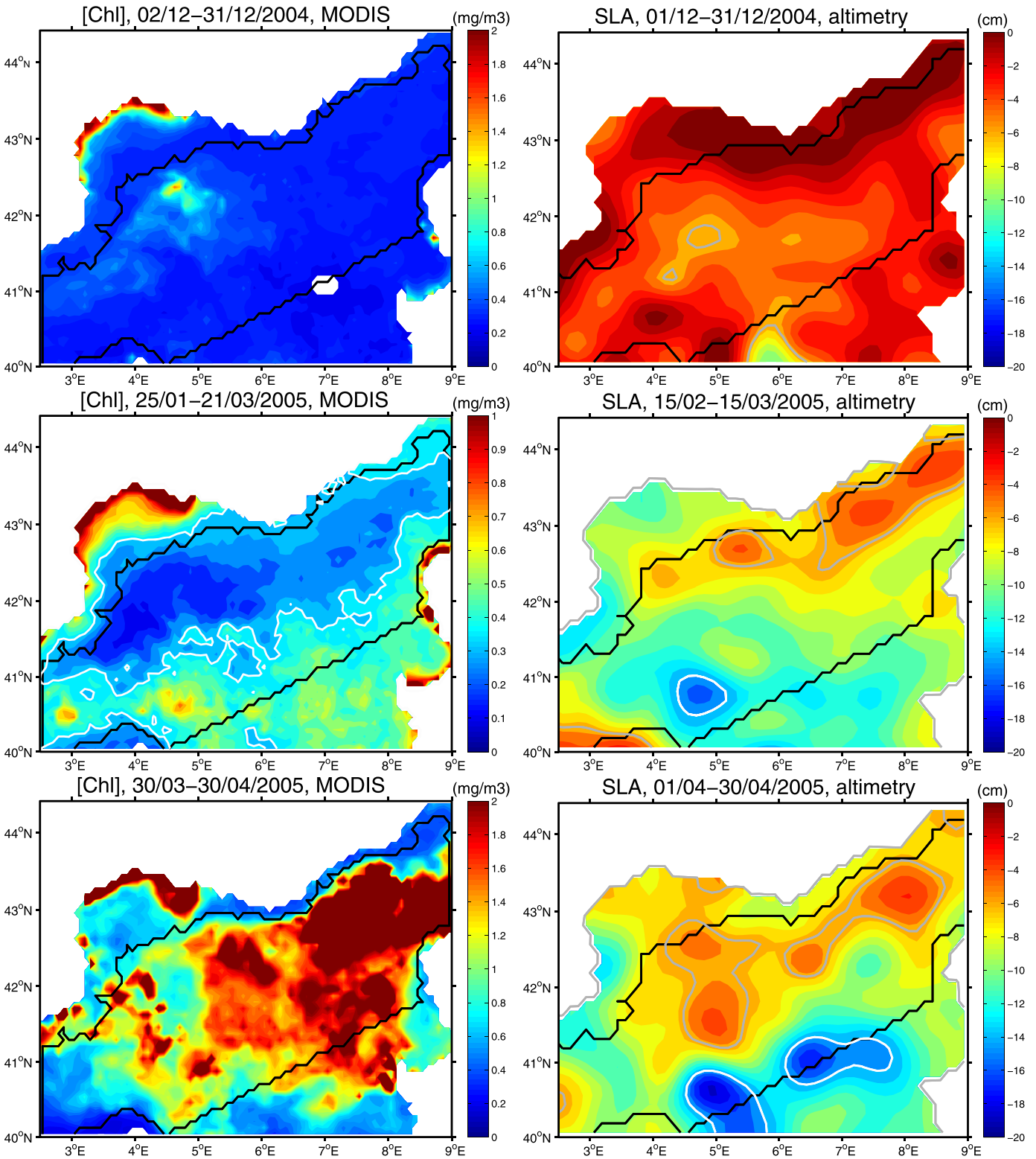

Figure 1. (left) Sea surface chlorophyll concentration (mgChl m^{-3}) and (right) sea level anomaly (cm) for winter 2005 in satellite ocean color data and altimetry. From top to bottom: averages for December 2004; averages between 25 January and 21 March 2005 for chlorophyll concentration and between 15 February and 15 March 2005 for SLA; averages for April 2005. White line corresponds to the 0.35 mgChl m^{-3} isoline for surface chlorophyll concentration. White and gray lines correspond, respectively, to the −14.0 and −5.5 cm isolines for SLA. Black line corresponds to the limits of the RDC region where A_{lowChl} (defined in equation (5)) and A_{lowSLA} (defined in equation (6)) are computed.

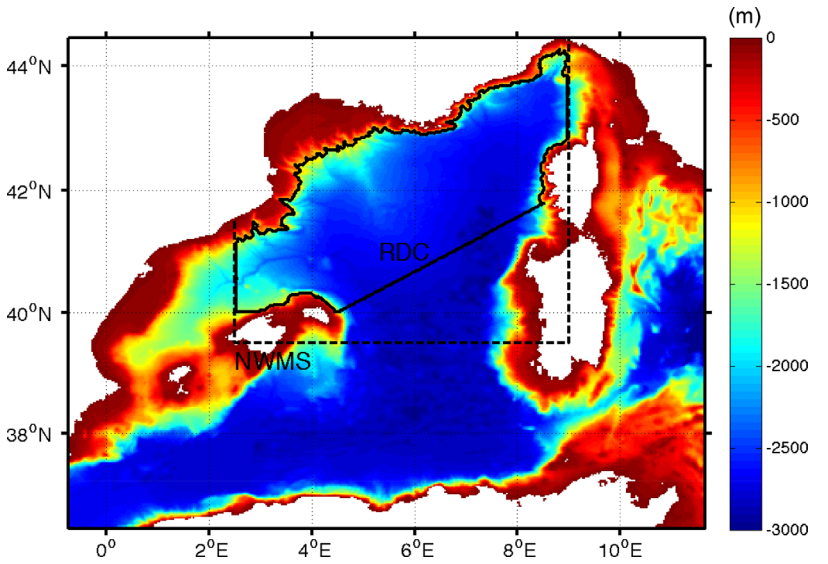

Figure 2. Bathymetry of the modeled domain (m). Black dotted line corresponds to the limits of the NWMS region, and black full line corresponds to the limits of the RDC region.

2016]. Those studies showed that our coupled model represents realistically NWMS ocean dynamics, in particular deep convection, as well as the interactions between dynamics and biogeochemistry.

Due to the Boussinesq approximation, SYMPHONIE is not able to reproduce the temporal variability of sea level associated with the steric effect. *Greatbatch* [1994] showed that sea level calculated by models making the Boussinesq approximation can be corrected for this problem by adding to the modeled sea level field a spatially uniform but time-dependent constant that accounts for any net expansion/contraction of the global ocean. To compute this temporally varying constant, we use the same method as *Lombard et al*. [2005] and *Bouffard et al*. [2008], using the monthly temperature and salinity fields from the NEMOMED8 simulation over the region between 2.5°E and 9°E, north of 39.5°E. We finally remove the long-term linear trend of the modeled sea level over the 1975–2013 period to remove the large scale sea level trend signal.

We define three indicators to quantify the intensity of DWF in the NWMS. First, the maximum depth reached by the mixed layer during winter is an indicator that has been widely used in previous modeling and observations studies. We define the annual maximum of the MLD averaged over the convection area, MLD_{mean}, as:

$$MLD_{mean} = max_{t\in DJFM}\left(\frac{\iint_{(x,y)\in NWMS/MLD(x,y,t)>500} MLD(x,y,t)\,dx\,dy}{\iint_{(x,y)\in NWMS/MLD(x,y,t)>500} dx\,dy}\right) \quad (1)$$

where DJFM stands for the December to March winter period and NWMS is defined as the region between 2.5°E and 9°E and north of 39.5°N (see Figure 2). We consider values of MLD larger than 500 m to ensure that we are in the convection area. In the model, the MLD is defined using a threshold value of 4 $cm^2\ s^{-1}$ for the vertical diffusion coefficient [*Herrmann et al*., 2008].

Second, the volume of water affected each year by DWF is also an indicator of its intensity. It is a key value for studying the formation and fate of water masses involved in the thermohaline circulation. Following previous studies (see section 2.3), we consider two kinds of yearly volumic DWF indicators, as defined in *Herrmann et al*. [2008]: the maximum volume of mixed water, V_{MLD}, and the rate of dense water formed annually, $\tau_{29.11}$. V_{MLD} is the winter maximum of the spatial integral of the MLD over the convection area:

$$V_{MLD} = max_{t \in DJFM} \left(\iint_{(x,y) \in NWMS / MLD(x,y,t) > 500} MLD(x,y,t)\, dx\, dy \right) \quad (2)$$

In our 38 year simulation, the densest water masses formed in winter have densities exceeding 29.11 kg m^{-3}, in agreement with values from previous modeling studies and observations (reported for example in *Herrmann et al*. [2010]) and with values observed for 2012–2013 (see section 2.3). This value is therefore taken as the criterion to define the volume of dense water formed $V_{29.11}$, computed as the volume of water of density higher than 29.11 kg m^{-3}:

$$V_{29.11}(t) = \iiint_{(x,y,z) \in NWMS / \rho(x,y,z,t) \geq 29.11} dx\, dy\, dz \quad (3)$$

$\tau_{29.11}$ is then defined as the annual rate of dense water formed. It is computed as the difference between the winter maximum and minimum of $V_{29.11}$:

$$\tau_{29.11} = max_{t \in DJFM}(V_{29.11}) - min_{t \in DJFM}(V_{29.11}) \quad (4)$$

Both volumic DWF indicators V_{MLD} and $\tau_{29.11}$ are quantified in Sv by dividing the cubic meters volumes by the number of seconds in one year. In the following, we therefore focus on three DWF indicators over the NWMS: MLD_{mean} (m), V_{MLD} (Sv), and $\tau_{29.11}$ (Sv). The time series of these three annual indicators computed from the model results are presented in Figure 3 (gray curves).

2.2. Satellite Data

2.2.1. Altimetry Data

We use the L4 daily multimissions satellite SLA data set generated at 1/4° resolution by the SSALTO/DUACS Delayed Time (DT) processing system for the period 1993–2015 and by the Near Real Time (NRT) processing system for 2016. The multisatellite component of SSALTO/DUACS system is responsible for the processing of HY-2A, Saral/AltiKa, Cryosat-2, Jason-1, Jason-2, Topex/Poseidon, Envisat, GFO, ERS1/2, and Geosat data in order to provide a homogeneous, intercalibrated and highly accurate long time series of gridded SLA altimeter data. The processing system uses a mapping procedure based on optimal interpolation with realistic correlation functions to produce gridded SLA at a given date. These altimeter products cover the period 1993–2016 and are now distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS, http://marine.copernicu.eu). As done for the model, we remove from those data their linear long-term trend. We also remove from both modeled and observed SLA their temporal average over the period 1993–2016 in order to use the same reference for model results and satellite data.

2.2.2. Ocean Color Data

We use the standard L3 near-surface chlorophyll-a concentrations data at 9 km resolution computed daily for the global ocean from SeaWIFS (1998–2010) and MODIS (2003–2016) satellite observations using the OC algorithm [*O'Reilly and et al*., 2000]. Those data are available on http://oceandata.sci.gsfc.nasa.gov.

2.3. In Situ Data and Existing Estimations of Deep Water Formation Rate

Between summer 2012 and summer 2013, four oceanographic cruises were conducted in the NWMS in the framework of the MERMEX (Marine Ecosystems Response to climatic and anthropogenic forcings in the Mediterranean), HYMEX (Hydrological Cycle of the Mediterranean Experiment), DEWEX (Impacts of Deep water formation on the Mediterranean pelagic ecosystems), and MOOSE (Mediterranean Ocean Observing System Experiment) programs: July–August 2012, February 2013, April 2013, and June 2013. They are presented in details in *Testor* [2013] and *Conan* [2013]. A large set of CTD profiles (in average 70 per cruise) was collected during these cruises. One of the main goals of those cruises was to estimate the seasonal and annual variations of dense water volume in the NWMS. Several estimations of the DWF rate based on those in situ CTD profiles were computed for winter 2012–2013. Performing optimal interpolation of those profiles and using a numerical model to assess the uncertainty associated with their DWF rate estimates, *Waldman et al*. [2016] obtained a DWF rate of 2.3 ± 0.5 Sv for $\tau_{29.11}$. *Bosse* [2015] performed an objective analysis of in situ data provided by those profiles, gliders surveys, ARGO floats and moorings, obtaining a 2 Sv DFW rate for $\tau_{29.11}$. Several studies attempted to estimate the DWF rate for winter 2012–2013 using other sources, i.e., numerical model results or satellite data. *Estournel et al*. [2016] used the SYMPHONIE model at 1 km

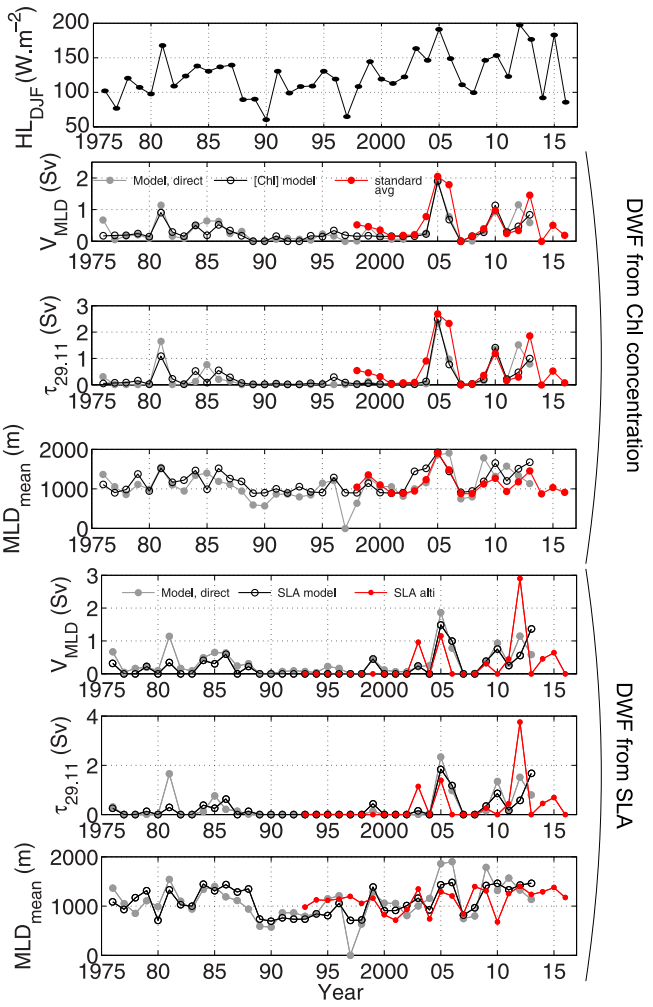

Figure 3. Annual time series of atmospheric and DWF indicators between 1975 and 2016. (top) Winter heat loss over the NWMS, HL_{DJF}, computed from NCEP reanalysis [*Kalnay et al.*, 1996]. (middle) DWF indicators V_{MLD}, $\tau_{29.11}$, and MLD_{mean} computed directly in the model (gray), and predicted by applying the relationships established in section 3.1 to A_{lowCHL} computed in the model (black) and to A_{lowCHL} obtained from merged SeaWIFS and MODIS satellite data (red). (bottom) DWF indicators computed directly in the model (gray) and predicted by applying the relationships established in section 3.2 to A_{lowSLA} computed in the model (black) and to A_{lowSLA} obtained from SSALTO/DUACS DT satellite data (red). When applying the equations to real satellite data, adjustment methods explained in sections 4.1 and 4.2 were used.

resolution to perform a realistic numerical simulation that closely reproduces the observed characteristics of the water column during 2012–13. They obtained a 1.6 Sv rate for $\tau_{29.1125}$. Using the oceanic model NEMO at 1/36° resolution, *Léger et al.* [2016] ran three sensitivity experiments with different initial ocean states. They obtained DWF rates varying between 0.6 Sv and 2.6 Sv for $\tau_{29.11}$.

Previous observations and modeling studies identified other winters during which DWF was stronger than the average: 1999 [*Béthoux et al.*, 2002, from in situ observations], 2005 [*Herrmann et al.*, 2010, gave an estimate of 1.2 Sv from model results and suggested that winter 2005 was the most convective over the period 1960–2006 due to considerable atmospheric heat loss], 2006 [*Schroeder et al.*, 2008, gave an estimate of 2.4 Sv for winters 2005 and 2006 from in situ observations], 2012 [*Durrieu de Madron et al.*, 2013, gave an estimate of 1.1 Sv from satellite color data using the same method as *Houpert et al.* [2016] with a threshold criteria of 0.1 mgChl.m^{-3}]. *Houpert et al.* [2016] used L3 MODIS Aqua surface chlorophyll concentration satellite data to estimate DWF rates for the period 2007–2013: they took the maximum extension of the low concentration area (defined with a threshold criteria of 0.15 mgChl m^{-3}) and assumed that the mean MLD below this area was 2200 m. They obtained the following yearly values for the period 2007–2013: 0 Sv, 0 Sv, 1.14 Sv, 0.91 Sv, 1.10 Sv, 1.25 Sv, and 1.65 Sv. *Somot et al.* [2016] performed a simulation over the Mediterranean Sea for the 1980–2013 period using a coupled ocean-atmosphere model (ALADIN-Climate—NEMOMED8) to investigate the factors responsible for the interannual variability of deep convection in the NWMS. In their paper, they produced a time series of DWF rate (their Figure 6), identifying winter 2005 as the most convective of the period.

3. Signature of Deep Convection on Sea Surface Chlorophyll Concentration and Sea Level Anomaly in the Coupled Simulation

To assess DWF annual intensity using ocean color and altimetry satellite data, we first need to establish relationships between the sea surface characteristics and the intensity of deep convection. We use the 38 year hydrodynamical-biogeochemical simulation to establish relationships between the surface chlorophyll concentration and SLA on one side and the DWF indicators on the other side.

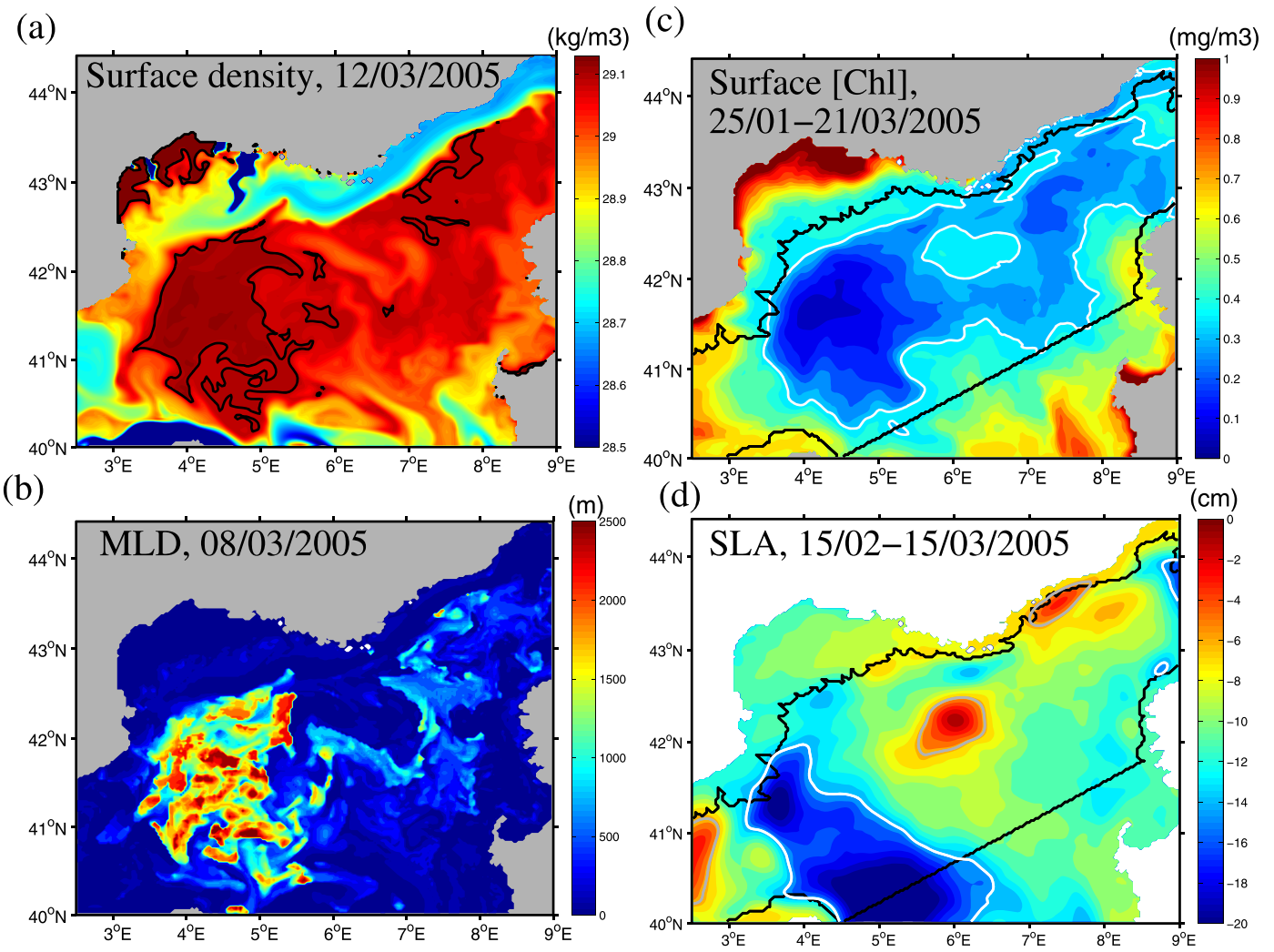

Figure 4. Sea surface characteristics for winter 2005 in the model. (a) Surface density (kg m^{-3}) on the day when $V_{29.11}$ (defined in equation (4)) is maximum (12 March 2005). Dark grey line corresponds to the 29.11 kg m^{-3} isoline. (b) MLD (m) on the day when V_{MLD} (defined in equation (2)) is maximum (8 March 2005). (c) Surface chlorophyll concentration (mgChl m^{-3}) averaged between 25 January and 21 March. White line corresponds to the 0.35 mgChl m^{-3} isoline. (d) SLA (cm) averaged between 15 February and 15 March. White and gray lines correspond, respectively, to the −14.0 and −5.5 isolines. Black line corresponds to the limits of the region RDC where A_{lowChl} (defined in equation (5)) is computed.

3.1. From Sea Surface Chlorophyll Concentration to Dense Water Formation

As explained above, very low chlorophyll concentration values are observed during the period and over the surface of DWF (see the example of winter 2005 in Figures 1 and 4, the most convective year over the simulated period, Figure 3). We therefore look for relationships between the size of the chlorophyll depleted area, A_{lowCHL}, and the three indicators of DWF intensity defined in section 2.1. For a given period T and a given threshold Chl_{crit}, A_{lowCHL} is defined as the surface where the chlorophyll concentration averaged over the period T is lower than Chl_{crit}:

$$A_{lowCHL} = \iint_{(x,y)\in RDC/\left(\int_{t\in T} C\, hl(x,y,t)dt\right) \le Chl_{crit}} dx\, dy. \quad (5)$$

where RDC (Region of Deep Convection) is the region of the NWMS where DWF occurs in the model and where low chlorophyll concentrations are observed in the model and in satellite data. It corresponds to the region between 2.5°E and 9°E in the east-west direction, and north of 40°N and of a line going from [40°N;4.5°E] to [42°N;9°E] in the south-north direction. Since we study open ocean convection and not shelf

dense water formation, we consider only the region where the bathymetry exceeds 1000 m. The selected RDC (see black line in Figures 1, 2, and 4) is consistent with the "blooming" bioregion defined by *D'Ortenzio and Ribera d'Alcalà* [2009], who applied a K-means cluster analysis on time series of chlorophyll concentration computed from SeaWIFS satellite data to characterize the biogeography of the Mediterranean Sea.

To compute annual time series of A_{lowCHL}, we then need to define the period of averaging, T, and the surface chlorophyll concentration criteria, Chl_{crit}. For that, we use a simple optimization procedure. We vary T inside the January-March period (during which DWF occurs) and vary the value of Chl_{crit} in the range [0.00–1.00] mgChl m^{-3}. We compute the corresponding correlations between the annual time series of each DWF intensity indicators and A_{lowCHL}. For each indicator, we finally select the period and chlorophyll concentration criteria for which we obtain the highest correlation. For V_{MLD} and $\tau_{29.11}$, we obtain the 56 day period 25 January –to 21 March and $Chl_{crit} = 0.35$ mgChl m^{-3}, with statistically significant correlations (significant level SL > 0.9999) of, respectively, 0.88 and 0.89 (Figure 5). For MLD_{mean}, we obtain the 48 day period 25 January to 31 March and $Chl_{crit} = 0.50$ mgChl m^{-3}, with a statistically significant correlation (SL > 0.9999) of 0.64 (Figure 5). Given these high correlation levels, we then perform linear regression analysis under the form $y = ax + b$ where y is the value of the DWF indicator and x is the value of A_{lowCHL}: the values of a and b are given for each DWF indicator in Figure 5. When quantifying the DWF indicators in m^3 instead of Sv, the values of a are equal to 751 and 1057 m for V_{MLD} and $\tau_{29.11}$, respectively, much smaller than the 2200 m value used by *Durrieu de Madron et al*. [2013] and *Houpert et al*. [2016]. We apply those linear relationships to the values of A_{lowCHL} given by the model, obtaining time series of V_{MLD}, $\tau_{29.11}$, and MLD_{mean} predicted from those relationships (Figure 3, black curves). The normalized root-mean-square error (NRMSE) between the DWF indicators time series computed directly from the hydrodynamical simulation and the predicted time series are of 10.4%, 11.0%, and 14.9% for V_{MLD}, $\tau_{29.11}$ and MLD_{mean}, respectively.

The strong correlation obtained in the model between the volume of water affected by convection and the chlorophyll depleted area can be mainly explained by the vertical dilution effect of convection. Before winter, the water column is still stratified, chlorophyll is mainly present in the surface euphotic layer and the total chlorophyll content of the water column is very similar for every year (Figure 6, left). When convection occurs, this initial chlorophyll content is diluted over the mixed column. The resulting surface chlorophyll concentration should be approximately inversely proportional to the MLD. This inversely proportional pattern can be observed in the model on the scatterplot of daily values of MLD versus surface chlorophyll concentration at the center of the convection area during the January to February period (Figure 6, right). This suggests that the dilution effect is indeed the main factor responsible for the chlorophyll depletion of the surface water. Moreover, most of MLD values greater than ~500 m are associated with chlorophyll surface concentration values lower than ~0.35 mgChl m^{-3} (Figure 6, right). The 0.35 mgChl m^{-3} threshold obtained for V_{MLD} and $\tau_{29.11}$ therefore approximately corresponds to the value below which the surface chlorophyll concentration falls when the mixed layer reaches significant depths (MLD > ~500 m). During winter 2005, for example, the area of deep convection (where MLD > 500 m) is indeed approximately the same as the area inside the 0.35 mgChl m^{-3} isoline for the surface chlorophyll concentration (Figure 4). The dilution effect therefore enables us to physically justify the strong correlations found above but also the 0.35 mgChl m^{-3} threshold.

3.2. From Sea Surface Height to Dense Water Formation

Herrmann et al. [2009] used a 9 year simulation at 1/16° (~5 km resolution) together with along track altimetry data. Our goal is to strengthen the robustness of the relationships that they obtained between sea level and deep convection. For that we use a longer (38 year) simulation with a higher resolution (2.5 km), hence a simulation who represents more realistically ocean dynamics in the deep convection region [*Herrmann et al*., 2008]. Moreover, we use gridded altimetry data, where submesoscale structures, which are highly active during convection [*Herrmann et al*., 2008], should be filtered out compared to alongtrack data. To establish relationships between SLA and DWF indicators, we proceed the same way as for the surface chlorophyll: we look for relationships between those indicators and the surface of low SLA, A_{lowSLA}. For a given period T and a given threshold SLA_{crit}, A_{lowSLA} is defined as the surface where the SLA averaged over the period T is lower than SLA_{crit}:

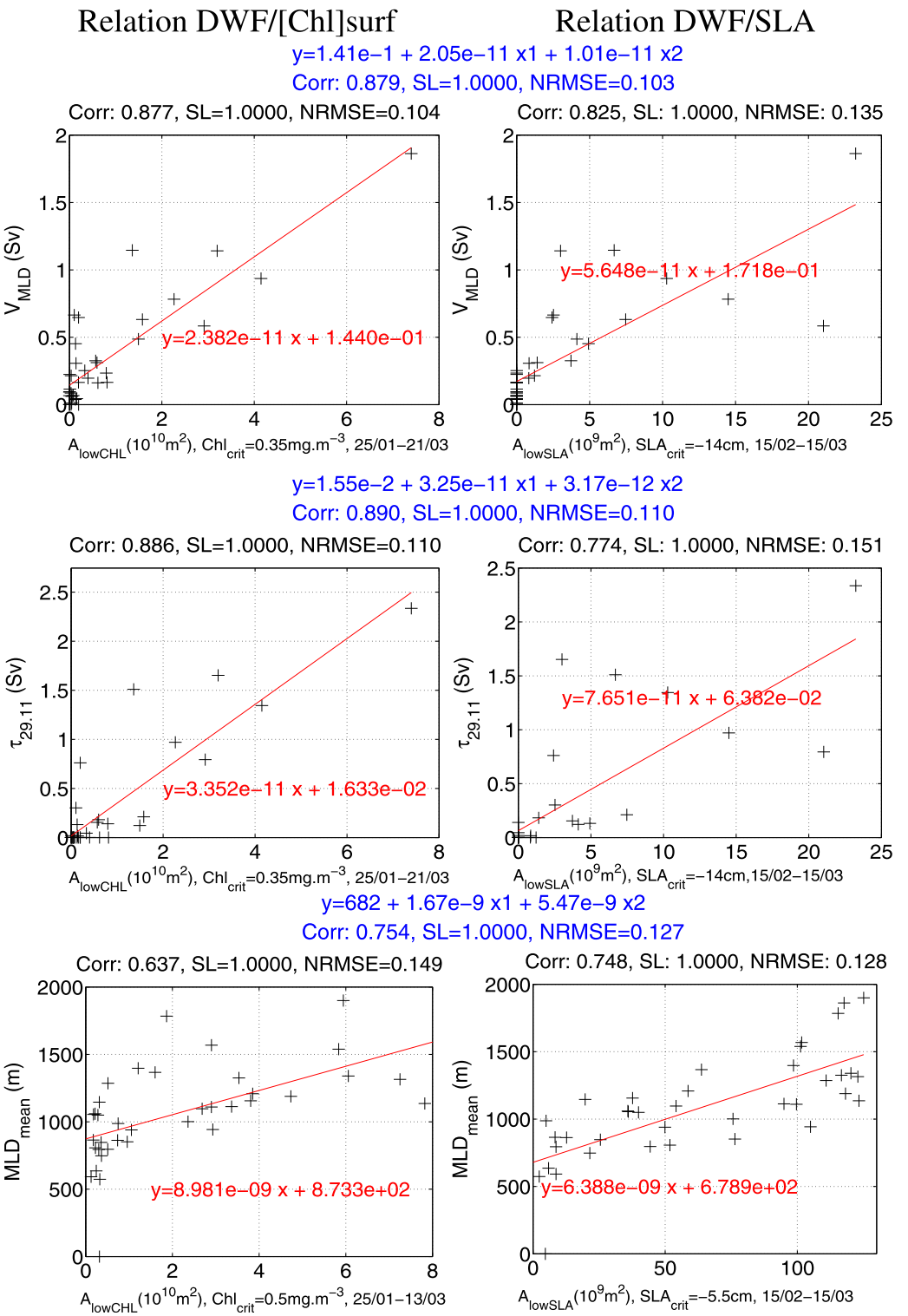

Figure 5. Relationships established in sections 3.1 and 3.2 between the modeled annual DWF intensity indicators (top) V_{MLD}, (middle) $\tau_{29.11}$, (bottom) MLD_{mean}, and the (left) modeled low surface chlorophyll concentration area A_{lowCHL} and (right) low SLA area A_{lowSLA}. For each DWF indicator, bilinear regression analysis coefficients established in section 3.3 and correlation coefficients and NRMSE between corresponding predicted time series and direct modeled time series are indicated in blue.

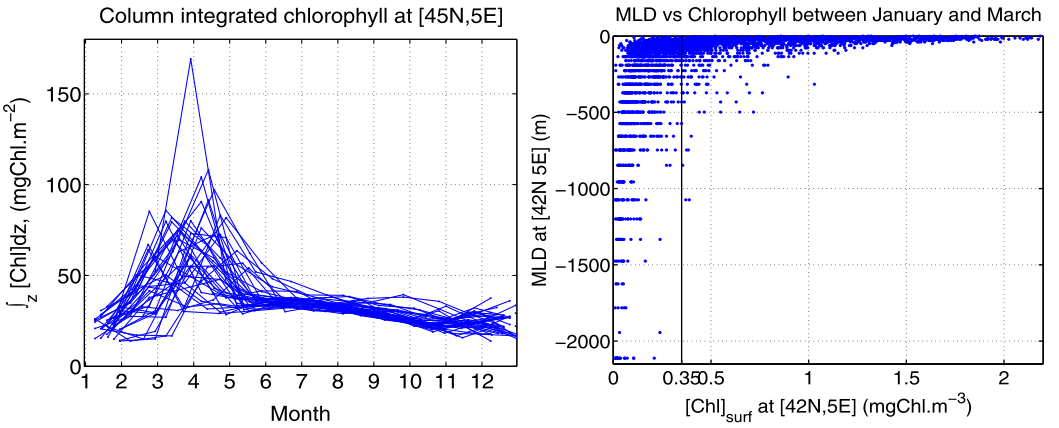

Figure 6. Relationship between MLD and chlorophyll concentration in the model. (left) Annual evolution of the column-integrated chlorophyll content over the water column at [42°N 5°E] for the 38 years of the simulation. (right) Scatterplot of the daily values of MLD versus surface chlorophyll concentration at the center of the convection area, [42°N 5°E], between January and March for the 38 years of the simulation.

$$A_{lowSLA} = \iint_{(x,y)\in RDC/\left(\int_{t\in T\,S} LA(x,y,t)dt) \le SLA_{crit}\right)} dx\,dy. \tag{6}$$

We consider the same RDC region as for the surface chlorophyll (black line in Figure 4). We then vary the period T and the criteria SLA_{crit}, and choose the values that maximize the correlations between the DWF indicators and A_{lowSLA}. The optimal averaging period is 15 February to 15 March, with SLA_{crit} of −14.0, −14.0, and −5.5 cm for V_{MLD}, $\tau_{29.11}$, and MLD_{mean}, respectively. We obtain statistically significant correlations (SL > 0.9999) between these indicators and A_{lowSLA} of 0.83, 0.77, and 0.75, respectively (Figure 5). We perform linear regression analysis under the form $y=ax+b$ where y is the value of the DWF indicator and x is the value of A_{lowSLA}. Values of a and b obtained for each DWF indicator are given in Figure 5. We apply these relationships to the modeled A_{lowSLA} to obtain predicted DWF indicators time series. The NRMSE between the DWF indicators time series computed directly from the simulation and the predicted time series are 13.5%, 15.1%, and 12.8% for V_{MLD}, $\tau_{29.11}$, and MLD_{mean}, respectively (Figure 5).

3.3. Using Together Sea Surface Height and Chlorophyll Concentration

To combine the information provided both by altimetry and ocean color data, we also establish bilinear relationships under the form $y=a_1x_1+a_2x_2+b$ where y is the value of the DWF indicator, x_1 is the value of A_{lowCHL}, and x_2 is the value of A_{lowSLA}. The values of a_1, a_2, and b are indicated in blue in Figure 5, as well as the correlation coefficients and NRMSE between the time series of DWF indicators given directly by the model and the time series predicted by applying the bilinear relationships to the modeled A_{lowCHL} and A_{lowSLA}. For V_{MLD}, $\tau_{29.11}$, and MLD_{mean}, those correlation coefficients are equal to 0.879, 0.890, 0.754 (SL > 0.999) and the NRMSE to 10.3%, 11.0%, and 12.7%, respectively, only marginally higher than the best coefficients obtained for each indicator: 0.877 (NRMSE 10.4%), 0.886 (NRMSE 11.0%), and 0.748 (NRMSE 12.8%) obtained with A_{lowCHL}, A_{lowCHL}, and A_{lowSLA}, respectively. This is due to the fact that performing this multivariate regression analysis is actually equivalent to applying a weighted average to the linear relationships $y=ax+b$ established in sections 3.1 and 3.2. The strongest weight is given to the relationship associated with the highest predicted versus direct modeled time series correlation. This can be seen when comparing the $y=a_1x_1+a_2x_2+b$ equation with both $y=ax+b$ equations (Figure 5). In the model, for a given DWF indicator, multivariate DWF indicator time series, showed in black in Figure 7, is therefore close from the univariate time series associated with the strongest weight (Figure 3).

4. From Satellite Data to DWF Intensity

In this section, we apply the $y=ax+b$ and $y=a_1x_1+a_2x_2+b$ relationships established for the model in section 3 to the areas of low SLA and low surface chlorophyll concentration computed from real satellite

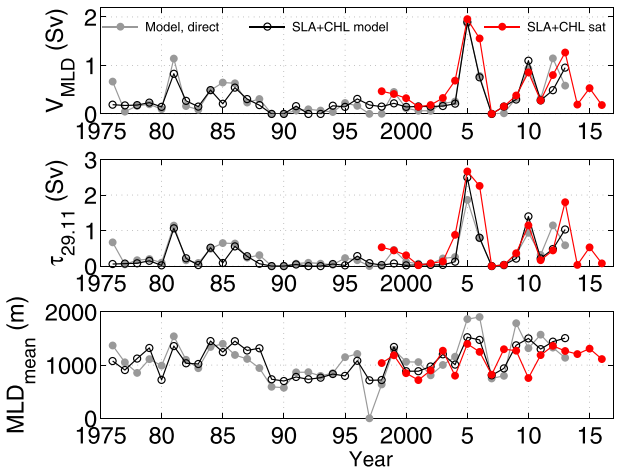

Figure 7. Annual time series of DWF indicators between 1975 and 2016: V_{MLD}, $\tau_{29.11}$, and MLD_{mean} indicators computed directly in the model (gray), and predicted by applying the multivariate relationships established in section 3.3 to A_{lowCHL} and A_{lowSLA} computed in the model (black) and to A_{lowCHL} and A_{lowSLA} obtained from merged SeaWIFS and MODIS satellite data and from SSALTO/DUACS DT satellite data (red). When applying the equations to real satellite data, adjustment methods explained in sections 4.1 and 4.2 were used.

observations data sets, in order to produce DWF indicators time series from those data sets. For that we first compare the modeled and observed values of SLA and surface chlorophyll concentration to determine how these relationships can be applied to real data sets.

4.1. Adjustment of Chl_{crit} for Ocean Color Data

Time series of the mean surface chlorophyll concentration and chlorophyll depleted area over RDC computed from the model and from ocean color data are presented in Figure 8. The length of the common period between ocean color data sets and the numerical simulation is 12 years for SeaWIFS (no data in 2008) and 11 years for MODIS. The correlation factors between time series of the mean surface concentration computed in the model and in SeaWIFS and MODIS data sets are, respectively, equal to 0.59 (SL = 0.96) and 0.76 (SL = 0.99) for the period 25 January to 21 March, and to 0.75 (SL = 0.99) and 0.83 (SL > 0.999) for the period 25 January to 13 March . The model overestimates this mean surface chlorophyll concentration: +0.17 mgChl m^{-3} compared with SeaWIFS for the 1998–2010 period; +0.06 mgChl m^{-3} compared with MODIS for the 2003–2013 period.

Both satellite data sets are extremely close, showing the same variability and very similar values. We therefore merge those data sets to produce a 19 year time series for the period 1998–2016. We will apply the relationships established from the model in section 3 to this merged data set. To account for the model overestimation, we adjust the chlorophyll concentration criteria Chl_{crit} used to compute A_{lowCHL}. For the model, Chl_{crit} is, respectively, equal to 0.35 and 0.50 mgChl m^{-3} for the periods 25 January to 21 March and 25 January to 13 March (established in section 3.1). For the merged satellite data set, we vary Chl_{crit} and use the value that maximizes the temporal correlation and minimizes the NRMSE between A_{lowCHL} computed from the merged data set and A_{lowCHL} computed from the model. This results in criteria Chl_{crit} equal to 0.35 mgChl m^{-3}, both for the periods 25 January to 21 March and 25 January to 13 March. Correlations between time series of A_{lowCHL} obtained from the model and from the merged data set are, respectively, equal to 0.85 and 0.79 for the periods 25 January –to 21 March and 25 January –to 13 March, with level of significance SL > 0.999 and RMSE ∼ 20% for both periods (Figure 8).

4.2. Adjustment for Altimetry Data

Time series of the mean SLA over RDC and over the period 15 February to 15 March computed from the model and from altimetry are presented in Figure 9. Over the 1993–2013 period, the correlation between observed and modeled time series is equal to 0.41 (SL > 0.94) with a mean bias of −0.4 cm and an NRMSE of 28.0% (Figure 9). The model reproduces correctly the range of observed sea level values, but its representation of their interannual chronology is not very good. There are two main reasons for that. First, the model only represents the monthly variations of the steric effect (see section 2.1) and can therefore miss its high frequency variations. Second, the spatial and temporal resolution of the SLA tracks used to produce the SLA gridded data set is not very high (∼10 days, O(100 km)), much coarser than in the model, also preventing altimetry to capture correctly the high frequency of the SLA spatial and temporal variability. As a result, SLA patterns are strongly smoothed in altimetry data, with spatial extrema less peaked than in reality than in the model, as can be seen in Figures 1 and 6: regions of extrema are of larger extension but of smaller extrema values than in the model. This smoothing effect therefore results in an overestimation of A_{lowSLA}. We need to account for this spatial smoothing effect when applying to the data the linear relationship established in the model. For that, we apply a corrective multiplying factor C on A_{lowSLA}, computed by taking the

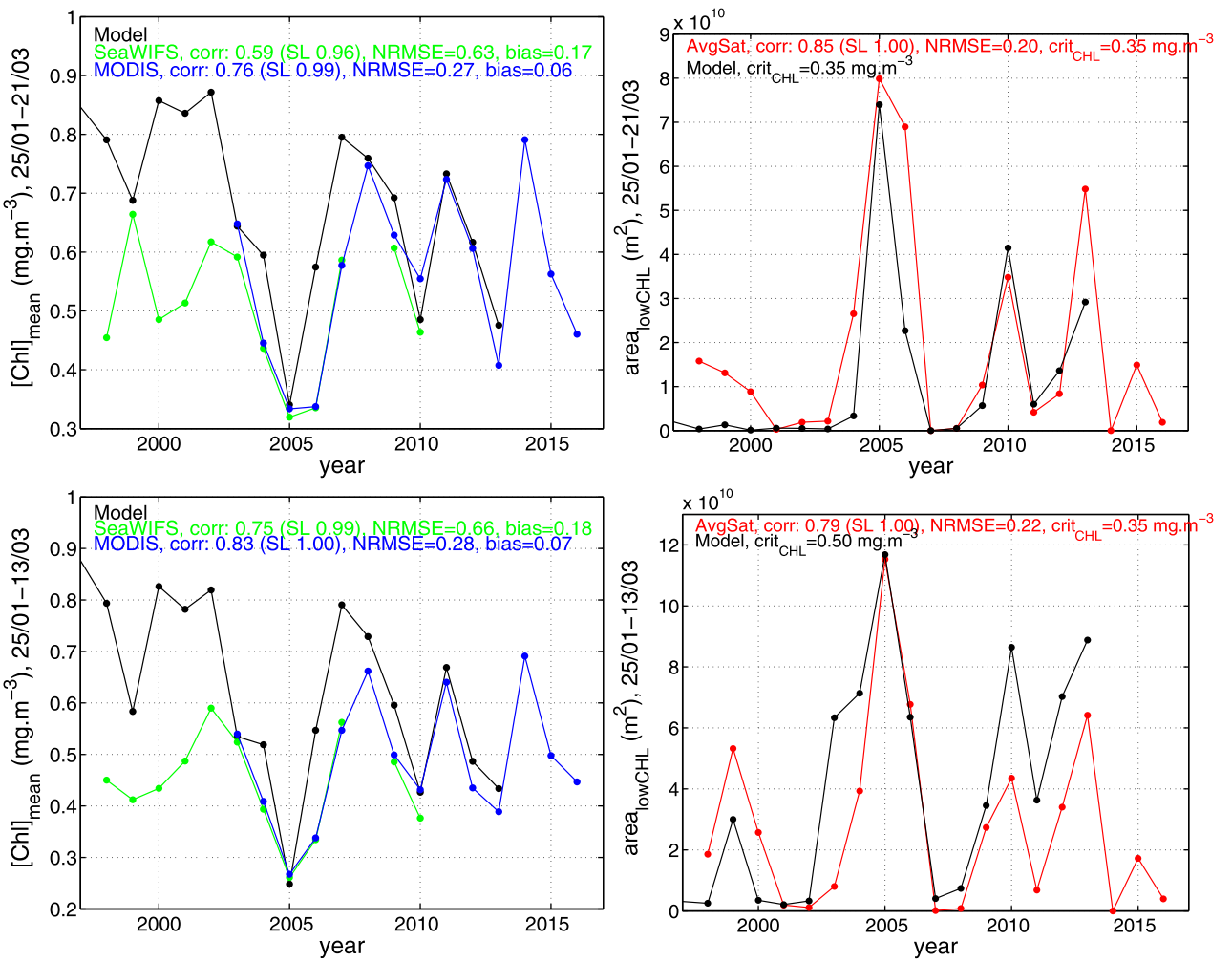

Figure 8. Comparison of surface chlorophyll concentration in the model and in ocean color data. Annual time series (left) of the surface chlorophyll concentration averaged over the region RDC (mgChl m^{-3}) and (right) of the chlorophyll depleted area A_{lowCHL} (m^2) for the periods 25 January to 21 March (top) and 25 January to 14 March (bottom), computed in the model (black), obtained from SeaWIFS (green), and MODIS (blue) satellite data and from merged SeaWIFS and MODIS data (red). When computing A_{lowChl} from real satellite data, criteria adjustment explained in section 4.1 is used.

ratio between the modeled A_{lowSLA} and the observed A_{lowSLA}, both averaged over the common period 1993–2013:

$$A_{lowSLA}^{altimetry,corrected}(y) = C \times A_{lowSLA}^{altimetry}(y) \quad \text{with} \quad C = \frac{\overline{A_{lowSLA}^{model}}}{\overline{A_{lowSLA}^{altimetry}}} \tag{7}$$

where A_{lowSLA}^{model} and $A_{lowSLA}^{altimetry}$ are the values of the area computed, respectively, from the model and the altimetry, $A_{lowSLA}^{altimetry,corrected}$ is the value obtained from the altimetry after applying the corrective factors, and the overbar denotes the average over the 1993–2013 period. For values of SLA_{crit} of, respectively, −14.0 and −5.5 cm established in section 3.2, we obtain corrective factors C of 5.629 and 0.888. The resulting time series of $A_{lowSLA}^{altimetry,corrected}$ are shown in Figure 9 (red curve).

4.3. Time Series of DWF Indicators Produced From Real Ocean Color and Altimetry Data Sets

We finally apply the linear relationships established from the model in sections 3.1 and 3.2 to the time series of A_{lowCHL} and A_{lowSLA} computed from real data sets. Time series obtained for V_{MLD}, $\tau_{29.11}$, and MLD_{mean} are

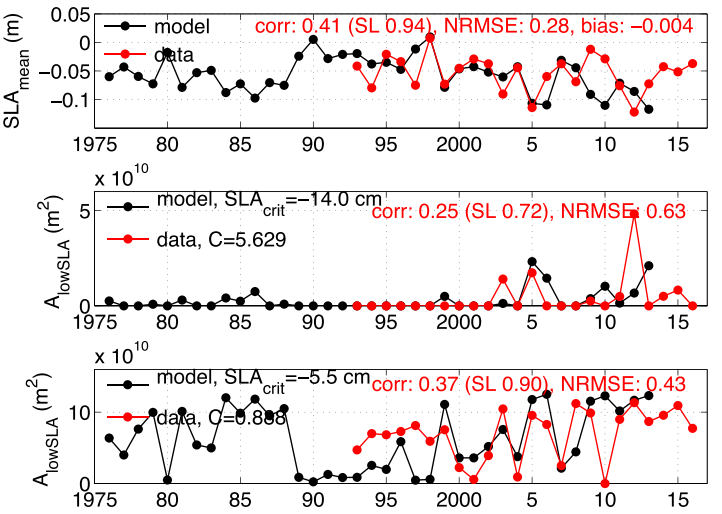

Figure 9. Comparison of SLA in the model and in altimetry data. Annual time series of the average over 15 February to 15 March of the mean SLA over RDC (top) and of the low SLA area A_{lowSLA} computed in the model (black) and in the altimetry data (red) for SLA_{crit} = −14.0 cm (middle) and −5.5 cm (bottom). When computing A_{lowSLA} from altimetry, adjustment method explained in section 4.2 is used.

shown in Figure 3 and given in Table 1. DWF stronger than the average is observed by ocean color data during winters 2004, 2005, 2006, 2010, 2013, 2015, and by altimetry during winters 2003, 2005, 2012 and to a lesser extent 2009, 2011, 2014, 2015. Previous observations and modeling studies (section 2.3) provide a list of known strong DWF winters: 1999, 2005, 2006, 2012, and 2013. Both types of data miss several strong DWF winters (2012 for color data, 2006, 2010, and 2013 for altimetry). This is discussed in the following section 5. Ocean color data give 2005 as the most convective winter, as suggested in the literature as well [*Schroeder et al*., 2008; *Herrmann et al*., 2010; *Somot et al*., 2016]. Altimetry ranks it as the second most convective winter, ranking 2012 as the most convective winter. To our knowledge, winter 2012–2013 was the only winter for which the in situ measurements coverage allowed to produce a robust rate of DWF (respectively, 1.8 and 2.3 ± 0.5 Sv for $\tau_{29.11}$ in *Bosse* [2015] and *Waldman et al*. [2016]). Our color data time series falls in this range, whereas altimetry data misses this convective winter (Figure 3).

Herrmann et al. [2010] showed that the intensity of DWF was significantly correlated to atmospheric conditions during the DWF period, in particular to the average winter (December to February) heat loss over the NWMS, HL_{DJF}. Examining jointly HL_{DJF} time series with the DWF indicators time series obtained here therefore provides an indication about the DWF interannual variability and about the ability of the satellite data to reproduce correctly this variability. We compute HL_{DJF} for the period 1976–2016 using NCEP reanalysis outputs [*Kalnay et al*., 1996, Figure 3]. We then compute the correlation between HL_{DJF} time series and DWF indicators time series obtained from various methods (direct model, model SLA, satellite altimetry, model chlorophyll concentration, satellite ocean color, combined altimetry, and ocean color, Table 2). The correlation between DWF indicators obtained directly from the 38 year simulation and HL_{DJF} is equal to, respectively, 0.68, 0.75, and 0.71 (SL > 0.999) for V_{MLD}, $\tau_{29.11}$, and MLD_{mean} over the 1976–2013 period. This confirms that the model reproduces realistically the interannual variability of DWF in the NWMS. The correlation between DWF volumic indicators V_{MLD} and $\tau_{29.11}$ predicted from altimetry data and HL_{DJF} is statistically significant, equal to 0.60 (SL > 0.99) over the 1993–2016 24 year period. However, the correlation is weaker and less significant for MLD_{mean} (0.24, SL = 0.74). This will be discussed in section 5. The time series produced from ocean color merged data set also shows a statistically significant correlation with HL_{DJF} (0.67 (SL > 0.99), 0.58 (SL > 0.99), 0.57 (SL > 0.98) for V_{MLD}, $\tau_{29.11}$, and MLD_{mean} over the 1998–2016 19 year period). These results suggest that the method built here, using gridded altimetry data and standard ocean color data, allows to produce time series that correctly capture the interannual variability of DWF in the NWMS. *Somot et al*. [2016] confirmed that winter heat loss plays a key role in DWF, but also that the initial stratification of the water column influences the convection, explaining that some years with strong buoyancy loss show convection weaker than expected. For example in their simulation $\tau_{29.11}$ was 1/2 smaller for 2012 than for 2013 although the winter buoyancy loss was stronger. Though it provides an indication about the ability of satellite data to reproduce realistically DWF indicators time series, it should therefore be underlined that the correlation between our DWF time series and NCEP HL_{DJF} time series must only be considered as an indication of the realism of DWF time series in terms of interannual variability, and not as an exact metric of this realism.

DWF time series obtained by applying the bilinear relationships established in section 3.3 to altimetry and ocean color data sets are shown in Figure 7 and given in Table 1. The correlation coefficients between those time series and HL_{DJF} are given in Table 2. As explained in section 3.3, those bilinear relationships give the

Table 1. Annual Time Series of DWF Indicators Between 1993 and 2016: MLD_{mean} (m), V_{MLD} (Sv), and $\tau_{29.11}$ (Sv) Indicators Predicted by Applying the Relationships Established in Sections 3.1–3.3 to A_{lowCHL} and A_{lowSLA} Computed From Merged Satellite Ocean Color Data From SeaWIFS and MODIS and From the SSALTO/DUACS DT Altimetry Satellite Data[a]

	Altimetry			Ocean Color			Combined Altimetry + Ocean Color		
Year	MLD_{mean}	V_{MLD}	$\tau_{29.11}$	MLD_{mean}	V_{MLD}	$\tau_{29.11}$	MLD_{mean}	V_{MLD}	$\tau_{29.11}$
1993	980	0,00	0,00						
1994	1126	0,00	0,00						
1995	1116	0,00	0,00						
1996	1145	0,00	0,00						
1997	1198	0,00	0,00						
1998	157	0,00	0,00	1040	0,52	0,55	1037	0,47	0,53
1999	1160	0,00	0,00	1351	0,46	0,46	1183	0,41	0,44
2000	822	0,00	0,00	1104	0,35	0,31	847	0,32	0,30
2001	715	0,00	0,00	891	0,15	0,02	716	0,15	0,02
2002	929	0,00	0,00	883	0,19	0,08	898	0,18	0,08
2003	1346	0,96	1,14	945	0,19	0,09	1267	0,33	0,13
2004	738	0,00	0,00	1227	0,78	0,91	798	0,69	0,88
2005	1289	1,15	1,39	1908	2,05	2,69	1397	1,96	2,67
2006	1207	0,00	0,00	1481	1,79	2,33	1248	1,56	2,26
2007	839	0,00	0,00	874	0,00	0,00	819	0,00	0,00
2008	1395	0,00	0,00	880	0,16	0,03	1297	0,15	0,03
2009	1311	0,31	0,25	1119	0,39	0,36	1269	0,38	0,36
2010	679	0,00	0,00	1264	0,97	1,18	754	0,86	1,15
2011	1254	0,45	0,44	934	0,24	0,15	1186	0,28	0,17
2012	1402	2,90	3,76	1179	0,34	0,30	1359	0,80	0,44
2013	1234	0,00	0,00	1450	1,45	1,86	1264	1,27	1,80
2014	1290	0,45	0,45	873	0,00	0,00	1205	0,19	0,03
2015	1377	0,64	0,69	1028	0,50	0,52	1309	0,53	0,53
2016	1173	0,00	0,00	909	0,19	0,08	1111	0,18	0,08

[a]When applying the equations to real satellite data, adjustment methods explained in sections 4.1 and 4.2 were used.

strongest weight to the area A_{lowCHL} or A_{lowSLA} for which the correlation between predicted and direct modeled DWF indicators time series is the strongest. As a result, for each indicator, the time series computed from combined altimetry and ocean color data is very similar to the time series obtained from the data set associated with this strongest weight (ocean color for V_{MLD} and $\tau_{29.11}$, altimetry for MLD_{mean}) and the correlation with HL_{DJF} is weakly improved. For MLD_{mean}, the difference of coefficient for combined data set (0.43) versus altimetry (0.24) is only due to the fact that the period considered differs. For V_{MLD}, the correlation improvement is mainly due to the fact that altimetry overestimates DWF for 2012, correcting the fact that ocean color misses it.

5. Discussion

The strength of the method developed here compared to empirically deduced relationships is the fact that it is based on the physical links, reproduced by the numerical model, between DWF, SLA, and chlorophyll

Table 2. Correlation Factors (With Significant Levels SL) Between Time Series of HL_{DJF} Computed From NCEP Reanalysis [*Kalnay et al.*, 1996] and Time Series of DWF Indicators Given Directly by the Model, Predicted by Model Chlorophyll Concentration and Ocean Color Data Applying the Linear Equations Established in Section 3.1, Predicted by Model SLA and Altimetry Data Applying the Linear Equations Established in Section 3.2, and Predicted by Combined Model SLA and Chlorophyll Concentration and Combined Altimetry and Ocean Color Data Applying the Bilinear Equations Established in Section 3.3[a]

Predicted From	MLD_{mean}	V_{MLD}	$\tau_{29.11}$	Period	Length (years)
Direct model results	0.681 (>0.999)	0.746 (>0.999)	0.707 (>0.999)	1976–2013	38
Modeled (chlorophyll)	0.777 (>0.999)	0.669 (>0.999)	0.644 (>0.999)	1976–2013	38
Ocean color data	0.667 (0.998)	0.583 (0.991)	0.568 (0.989)	1998–2016	19
Modeled SLA	0.683 (>0.999)	0.685 (>0.999)	0.667 (>0.999)	1976–2013	38
SLA from altimetry	0.241 (0.744)	0.602 (0.998)	0.596 (0.998)	1993–2016	24
Modeled SLA + CHL	0.725 (>0.999)	0.682 (>0.999)	0.647 (>0.999)	1976–2013	38
Altimetry + ocean color	0.434 (0.937)	0.687 (>0.999)	0.594(0.993)	1998–2016	19

[a]When applying the equations to real satellite data, adjustment methods explained in sections 4.1 and 4.2 were used.

concentration. *Durrieu de Madron et al*. [2013] and *Houpert et al*. [2016] used chlorophyll satellite data to propose DWF rate estimates (see section 2.3). Their method assumes that when convection occurs it reaches the bottom hence overestimates intermediate convection cases (see section 1). Indeed, for years 2009–2013, the range of DWF rate proposed by *Somot et al*. [2016] (0.2–1.7 Sv for $\tau_{29.11}$) was larger than the one proposed by *Houpert et al*. [2016] (0.9–1.7 Sv), with same maximum value but lower weak and intermediate values. Moreover, the day of maximum extension used by *Durrieu de Madron et al*. [2013] and *Houpert et al*. [2016] can be missed since the temporal coverage of the daily data is not perfect (we computed a temporal coverage of 20–35% for the daily data set proposed by SeaWIFS and MODIS). Our method allows to reduce the temporal coverage issue since we compute the depleted area using values averaged over ~2 months (see section 3.1). Using those averaged values instead of instantaneous values also enables to consider the full range of convection cases. Averaging the surface chlorophyll concentration indeed allows to integrate both spatial and temporal information about the intensity of the convection: deeper convection cases are associated with longer durations and to chlorophyll depleted areas of larger extension (and inversely). Finally, our method also allows to cope with the intermediate convection issue because it is based on a linear relationship determined from all the convection cases obtained in the model over a 38 year period. We consequently determined values of the relationship coefficient by which we multiply the chlorophyll depleted area of 751 and 1057 m for, respectively, V_{MLD} and $\tau_{29.11}$ (see section 3.1), i.e., not implying that convection obligatorily reaches the bottom under the chlorophyll depleted area.

The method proposed in this studies has, however, some weaknesses.

First, as seen above, comparison with atmospheric heat loss and previous observations and modeling studies suggests that time series of DWF indicators obtained from satellite data underestimate or overestimate DWF in some cases.

Time series from satellite chlorophyll concentration seems to overestimate DWF for 2006 and 2013 and to underestimate it for 2012 (Figure 3). This is due to the particular chronology of those convection events. Figure 10 shows the daily evolution of the mixed volume and average SLA and SST over the NWMS for winters during which HL_{DJF} is stronger than the average (2005, 2006, 2010, 2012, 2013). In 2012, the second most convective winter in the simulation in terms of volumes (Figure 3), the convection event is very short compared to other years, leading to an underestimation of A_{lowCHL} hence of DWF, both in model and in data. 2006 and 2013 are not the most convective years in terms of volumes, however in 2006, the convection begins very early, and for both years it lasts throughout the winter with several peaks. The surface is therefore depleted in chlorophyll during a long time, and the resulting A_{lowCHL} is high for satellite observations. The model does not reproduce this. This is due to an overestimation of the modeled chlorophyll concentration during the periods and in the regions of weak convection, that results in an underestimation of A_{lowCHL}. Using in situ and satellite observations, *Auger et al*. [2014] indeed showed that the model tends to overestimate the winter chlorophyll concentration (February to mid-March, their Figures 3 and 8). This overestimation of chlorophyll concentration in the model, also discussed in section 4.1, can be seen on time series and maps of winter average chlorophyll concentration: the mean chlorophyll concentration during the period 25 January to 21 March is overestimated by, respectively, 70% and 20% for 2006 and 2013, and is generally overestimated for years of weak convection (Figure 8); the regions outside the depleted area show a positive chlorophyll concentration bias in the model (Figures 1 and 4).

Time series from altimetry SLA seems to underestimate DWF for 2006, 2010, and 2013 and to overestimate it for 2012 (Figure 3). This is not the case for time series obtained from modeled SLA. Discrepancies between the modeled and satellite SLA are due partly to the representation of the steric effect in the model, which is computed from monthly fields (section 2.1). High frequency variations of the steric effect that can, for example, be induced by a sudden surface cooling/warming are therefore not included in the model. In average, since we consider the average SLA over the period 15 February to 15 March, the impact is limited, but for some particular cases it can impact the computation of A_{lowSLA}. Though 2012 is not the strongest convective year, the surface is very cold between 15 February and 15 March (see Figure 10). This cooling seems to be captured by the altimetry data that contains the real steric effect, contrary to the model. In 2006, 2010, and 2013, on the contrary, there are several warm events between 15 February and 15 March. As a result, the altimetry SLA is not very low and A_{lowSLA} is not very high, resulting in an underestimation of the DWF indicators by altimetry that is not observed in the model.

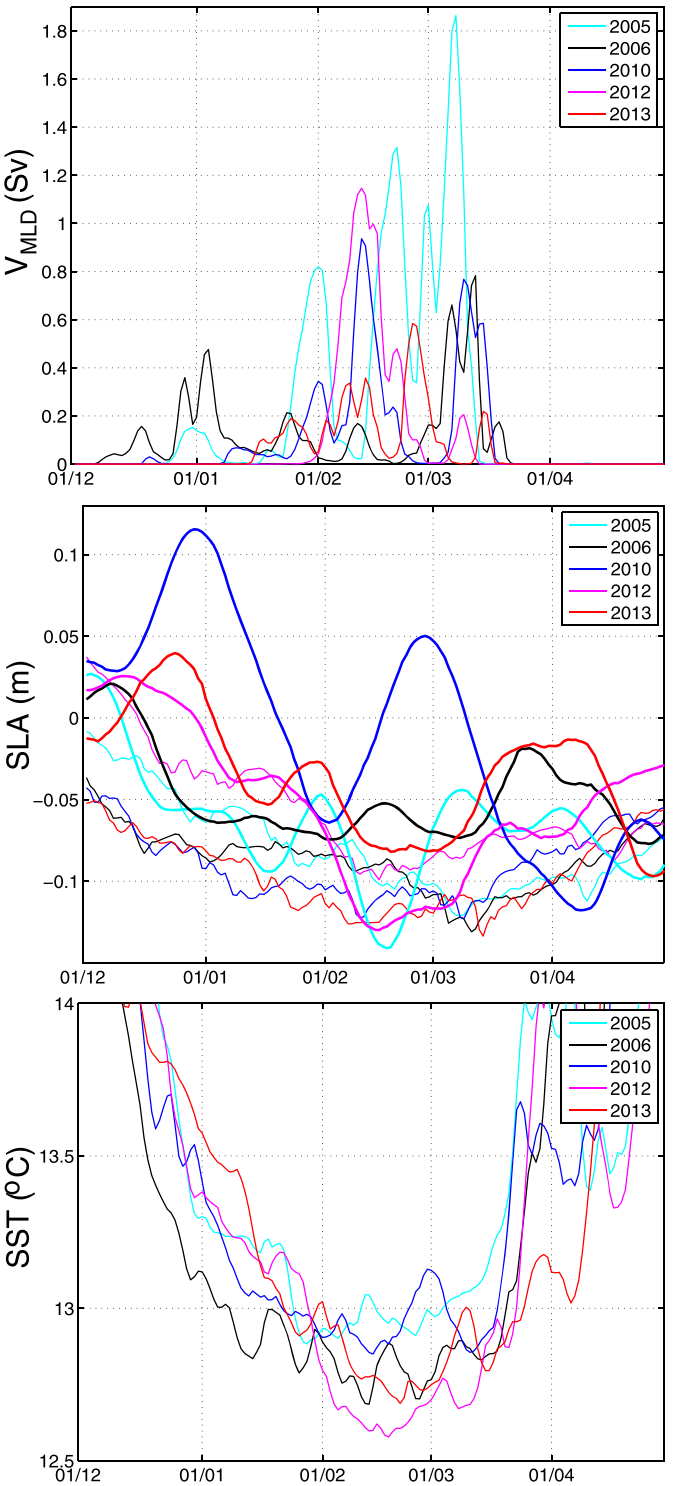

Figure 10. Evolution between 1 December and 30 April of the modeled daily mean of the (top) mixed volume, (middle) SLA, and (bottom) SST averaged over the NWMS for winters 2004–2005, 2005–2006, 2009–2010, 2011–2012, and 2012–2013.

Second, the relationships between DWF, SLA, and chlorophyll concentration were established from the model results. As seen above, this model shows some weaknesses in the representation of physical and biogeochemical processes and of their interactions. The established relationships are therefore not completely adapted for real altimetry and chlorophyll concentration data. Using corrective methods detailed in section 4.1 and 4.2 when applying the linear relationships found in the model to the real data set partly corrected those weaknesses. However, increasing the realism of the coupled model is essential to increase the ability of the linear relationships established in the model to represent the real physical interactions between DWF, SLA, and chlorophyll concentration. The performance of the coupled model should be improved in particular by recalibrating biogeochemical model parameters on the 2012–2013 well documented period, and by providing daily variable lateral boundary conditions to the physical model in order to better represent high frequency of steric effect.

Third, results from the numerical model have by definition a complete and high resolution spatial and temporal coverage. On the contrary, the coverage of satellite data is not perfect, due to the spatial and temporal resolution of the measurements and to external factors which hinder the measurements. Altimetry tracks are indeed spaced by several days and hundreds of kilometers in the NWMS. Ocean color data, though made daily at a high spatial resolution, are strongly impacted by cloud cover, showing an average coverage of 20–35%. These weaknesses of satellite data coverage can impact their ability to capture correctly the high frequency variations of SLA and chlorophyll concentration, which can be a reason for the misrepresentation of some DWF cases. The precision and accuracy of the satellite measurements and of the algorithms used to produce the data are an additional source of uncertainty in our time series.

Fourth, this study is based on the hypothesis that there is a strong linear relationship between DWF, SLA, and chlorophyll concentration. The reality is of course more complex, and SLA and chlorophyll concentration are impacted by other factors. The source of uncertainty linked to the linear regression analysis was estimated by giving the values of NRMSE and correlation between time series computed directly from the model results, and computed by applying the relationships to the model SLA and chlorophyll outputs: NRMSE varies between 10% and 15%, and correlation factors between 0.64 and 0.89 (SL $>$ 0.99). In particular MLD_{mean} is by construction a less integrated indicator than $\tau_{29.11}$ and V_{MLD}, which take into account both the depth and the area impacted by deep convection. The physical link between MLD_{mean} and the SLA and chlorophyll concentration is therefore less direct than for the volumic indicators. This explains that the correlation between MLD_{mean} and A_{lowCHL} and A_{lowSLA} (respectively, 0.64 and 0.75 versus 0.88 and 0.83 for V_{MLD} and 0.89 and 0.77 for $\tau_{29.11}$, Figure 5), though still significant at more than 0.999, is weaker than for the other indicators. This explains the weaker correlation between V_{MLD} time series obtained from altimetry data and HL_{DJF} (0.24, SL = 0.74 versus 0.60, SL $>$ 0.99 for V_{MLD} and $\tau_{29.11}$, Table 2 and Figure 3). For the time series obtained from ocean color data on the contrary, MLD_{mean} shows the highest correlation with HL_{DJF} (0.67 versus 0.58 and 0.57 for V_{MLD} and $\tau_{29.11}$, Table 2), though the correlation between A_{lowCHL} and MLD_{mean} is the weakest of all (0.64 versus 0.88 and 0.89 for V_{MLD} and $\tau_{29.11}$, Figure 5). The link between the atmospheric heat flux and the primary production is actually not only due to the effect of vertical mixing induced by cold atmospheric events on chlorophyll concentration, but also to the influence of the surface layer temperature, that largely depends on atmospheric heat flux, on primary production [*Herrmann et al*., 2014]. This high correlation between the time series of MLD_{mean} obtained from ocean color data and atmospheric heat loss therefore rather reflects this link between surface temperature and chlorophyll concentration than the ability of ocean color data to capture MLD_{mean} interannual variability.

In this method, we use altimetry and ocean color data as proxies of DWF. Since DWF results from surface buoyancy loss, that is mostly associated with cooling in the NWMS [*Herrmann et al*., 2010], one could consider using sea surface temperature (SST) as a proxy of DWF, similarly as what we did for SLA and chlorophyll concentration. However, the correlation between the winter SST and DWF indicators is much lower ($<$0.50) and less significant in our simulation than the correlations between winter SLA and chlorophyll concentration and DWF indicators. This is due to several reasons. First, the temperature does not decrease regularly with depth in the NWMS, due to the presence of the warm Levantine Intermediate Water (LIW) between the colder surface Modified Atlantic Water (MAW) and Western Mediterranean Deep Water. When the mixed layer deepens, the temperature consequently first increases when it reaches the LIW, then decreases. Second, atmospheric events but also advection of cold and fresh (due to the Rhone river input) thus light water produced on the Gulf of Lions shelf toward the open sea can induce strong but brief cooling events of surface water in the convection area, not necessarily associated with DWF. Those elements suggest that determining an SST criteria and building a DWF indicator from SST satellite data would not be possible.

6. Conclusion

Estimating the volume of dense water produced by deep convection and its interannual and long-term variability is of primary importance for the study of ocean circulation and ecosystems. Our objective in this paper was to propose a method allowing to assess the interannual variability of DWF using multisensors gridded altimetry and ocean color satellite data. For that we took the case of the NWMS which can be considered as a golden case study of DWF. We used the results of a 38 year simulation performed over the NWMS with the hydrodynamical-biogeochemical coupled high resolution model SYMPHONIE-Eco3M-S. Statistically significant correlations were computed in the model between the areas of low SLA and low surface chlorophyll concentration in winter on one side, and the DWF intensity estimated in terms of depth and volumes of affected water (mean MLD, mixed volume, and volume of newly formed dense water) on the other side. This allowed us to establish linear relationships between sea surface height and chlorophyll concentration and DWF. These relationships are not empirical but are obtained from a model that reproduces realistically the physical links between the ocean dynamics and the biogeochemistry. Using a 4 times longer simulation at a twice higher resolution than *Herrmann et al*. [2009], including in particular the period 2005–2013 with several convective years (see Figure 3), we increased the robustness of those relationships. We then applied those relationships to time series of areas of low SLA and low surface chlorophyll

concentration computed, respectively, from SSALTO-DUACS altimetry 24 year data set and SeaWIFS and MODIS ocean color data, merged to produce a 19 year data set. The chlorophyll concentration bias between the model results and the data was taken into account by adapting the threshold used to compute the depleted area. The smoothing effect of the gridded altimetry data was taken into account by applying a multiplying corrective factor to the low SLA area observed from altimetry. This allowed us to produce for the NWMS the first long time series of DWF indicators obtained from observations covering the last 2 decades. By comparison with existing estimations of DWF indicators and with the atmospheric heat loss over the region, we showed that the DWF indicators time series obtained from SSALTO-DUACS and merged standard ocean color data sets reproduce well the interannual variability and range of DWF intensity. We discussed the advantages and strengths but also the weaknesses and uncertainties of our method (misrepresentation of several DWF cases; realism of the linear relationships; ability of the numerical model to represent realistically the physical and biogeochemical processes, their interactions and their variability; spatiotemporal coverage, accuracy and precision of satellite data sets).

Using a combination of altimetry and ocean color data through bilinear relationships does not significantly modify the resulting DWF indicators time series. This is due to the fact that in the model correlations between the predicted and direct time series for a given DWF indicator are significantly different for A_{lowSLA} and A_{lowCHL}. Using bilinear relationships could be more efficient if those correlations were higher and more similar, which could be obtained when correcting the model weaknesses in the representation of those variables.

Smoothing effect of gridded altimetry data should disappear in the SWOT wide-swath satellite altimetry mission (Surface Water and Ocean Topography, from 2020), that will provide sea level data with a complete spatial coverage at a much higher resolution. This should allow to increase significantly the quality of the DWF indicators time series obtained from altimetry data set.

Our method is not proposed as a replacement of in situ measurement. In situ measurements methods dedicated to the observation of water masses in DWF sites are necessary to estimate the volume of those water masses and to evaluate and correct the models used to study and forecast DWF, hence to build our method. However, satellite data are highly complementary to in situ measurements given the length and quality of their spatial and temporal coverage. Such a coverage would be too expensive to be reached through in situ measurements and allows to monitor the interannual and long-term evolution of processes implied in ocean circulation like DWF.

The feasibility of our method was examined for the NWMS convection region, but deep convection occurs in other regions of the world ocean, in particular the Greenland and Labrador seas. Deep water masses formed in those regions play a key role in the Atlantic and global ocean circulation. Numerical simulations suggested a weakening of global overturning circulation due to a decrease of dense water formation under the influence of climate change. Long-term quantitative monitoring of DWF in deep convection regions of the world ocean and of its long-term trend is therefore of great importance. Our method applied to those regions could contribute to this monitoring.

Acknowledgments
This work is a contribution to the MISTRALS/HyMeX and MISTRALS/Mermex programmes. It has received funding from the French National Research Agency (ANR) projects ASICS-MED (contract ANR-12-BS06-0003) and additional support during the writing phase was provided by the Instituto Milenio de Oceanografia (IMO-Chile), funded by the Iniciativa Cientifica Milenio (ICM-Chile). We warmly thank the researchers that provided their estimations of DWF rates for the discussions that we had together. Standard L3 near-surface chlorophyll-a concentrations daily SeaWIFS and MODIS are available on http://oceandata.sci.gsfc.nasa.gov. SSALTO/DUACS DT and NRT data are distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS, http://marine.copernicu.eu). Results from numerical simulations are available on request (marine.herrmann@ird.fr).

References

Auger, P.-A., C. Estournel, C. Ulses, L. Stemmann, S. Somot, and F. Diaz (2014), Interannual control of plankton ecosystem in a deep convection area as inferred from a 30 year 3D modeling study: Winter mixing and prey/predator interactions in the NW Mediterranean, *Prog. Oceanogr.*, *124*, 12–27, doi:10.1016/j.pocean.2014.04.004.

Béthoux, J.-P., X. Durrieu de Madron, F. Nyffeler, and D. Tailliez (2002), Deep water in the western Mediterranean: Peculiar 1999 and 2000 characteristics, shelf formation hypothesis, variability since 1970 and geochemical inferences, *J. Mar. Syst.*, *33–34*, 117–131.

Bosc, E., A. Bricaud, and D. Antoine (2004), Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWifs observations, *Global Biogeochem. Cycles*, *18*, GB1005, doi:10.1029/2003GB002034.

Bosse, A. (2015), Circulation générale et couplage physique-biogéochimie à (sous-) mesoéchelle en Méditerranée Nord-Occidentale à partir de données in situ, PhD thesis, Univ. Paris 6, Pierre et Marie Curie, Paris, France.

Bouffard, J., S. Vignudelli, M. Herrmann, F. Lyard, P. Marsaleix, Y. Ménard, and P. Cipollini (2008), Comparison of ocean dynamics with a regional circulation model and improved altimetry in the northwestern Mediterranean, *Terr. Atmos. Oceans Sci.*, *19*(1–2), 117–133, doi:10.3319/TAO.2008.19.1-2.117(SA).

Conan, P. (2013), DEWEX-MERMEX 2013 LEG2 cruise, *RV Le Suroit*, doi:10.17600/13020030.

de Lavergne, C., J. Palter, E. Galbraith, R. Bernardello, and I. Marinov (2014), Cessation of deep convection in the open Southern Ocean under anthropogenic climate change, *Nat. Clim. Change*, *4*, 278–282, doi:10.1038/nclimate2132.

D'Ortenzio, F., and M. Ribera d'Alcalà (2009), On the trophic regimes of the Mediterranean Sea: A satellite analysis, *Biogeosciences*, *6*, 139–148.

Durrieu de Madron, X., et al. (2013), Interaction of dense shelf water cascading and open-sea convection in the northwestern Mediterranean during winter 2012, *Geophys. Res. Lett.*, *40*, 1379–1385, doi:10.1002/grl.50331.

Estournel, C., et al. (2016), High resolution modelling of dense water formation in the north-western Mediterranean: Benefits from an improved initial state in summer, *J. Geophys. Res. Oceans*, *121*, 5367–5392, doi:10.1002/2016JC011935.

Gelderloos, R., C. A. Katsman, and K. Vage (2013), Detecting Labrador sea water formation from space, *J. Geophys. Res. Oceans*, *118*, 2074–2086, doi:10.1002/jgrc.20176.

Greatbatch, R. J. (1994), A note on the representation of steric sea level in models that conserve volume rather than mass, *J. Geophys. Res.*, *99*, 12,767–12,771.

Herrmann, M., and S. Somot (2008), Relevance of ERA40 dynamical downscaling for modeling deep convection in the Mediterranean Sea, *Geophys. Res. Lett.*, *35*, L04607, doi:10.1029/2007GL032442.

Herrmann, M., S. Somot, F. Sevault, C. Estournel, and M. Déqué (2008), Modeling the deep convection in the Northwestern Mediterranean sea using an eddy-permitting and an eddy-resolving model: Case study of winter 1986–87, *J. Geophys. Res.*, *113*, C04011, doi:10.1029/2006JC003991.

Herrmann, M., J. Bouffard, and K. Béranger (2009), Monitoring open-ocean deep convection from space, *Geophys. Res. Lett.*, *36*, L03606, doi:10.1029/2008GL036422.

Herrmann, M., F. Sevault, J. Beuvier, and S. Somot (2010), What induced the exceptional 2005 convection event in the northwestern Mediterranean basin? Answers from a modeling study, *J. Geophys. Res.*, *115*, C08029, doi:10.1029/2009JC005749.

Herrmann, M., F. Diaz, C. Estournel, P. Marsaleix, and C. Ulses (2013), Impact of atmospheric and oceanic interannual variability on the Northwestern Mediterranean Sea pelagic planktonic ecosystem and associated carbon cycle, *J. Geophys. Res. Oceans*, *118*, 5792–5813, doi:10.1002/jgrc.20405.

Herrmann, M., C. Estournel, F. Diaz, and F. Adloff (2014), Impact of climate change on the Northwestern Mediterranean Sea pelagic planktonic ecosystem and associated carbon cycle, *J. Geophys. Res.*, *119*, 5815–5836, doi:10.1002/2014JC010016.

Houpert, L., et al. (2016), Observations of open-ocean deep convection in the Northwestern Mediterranean Sea: Seasonal and interannual variability of mixing and deep water masses for the 2007–2013 period, *J. Geophys. Res. Oceans*, *121*, 8139–8171, doi:10.1002/2016JC011857.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Léger, F., C. Lebeaupin Brossier, H. Giordani, T. Arsouze, J. Beuvier, M.-N. Bouin, E. Bresson, V. Ducrocq, N. Fourrié, and M. Nuret (2016), Dense Water Formation in the North-Western Mediterranean area during HyMeX-SOP2 in 1/36 ocean simulations: Sensitivity to initial conditions, *J. Geophys. Res. Oceans*, *121*, 5549–5569, doi:10.1002/2015JC011542.

Lombard, A., A. Cazenave, P. Y. L. Traon, and M. Ishii (2005), Contribution of thermal expansion to present-day sea-level change revisited, *Global Planet. Change*, *47*, 1–16.

Marsaleix, P., F. Auclair, J. W. Floor, M. J. Herrmann, C. Estournel, I. Pairaud, and C. Ulses (2008), Energy conservation issues in sigma-coordinate free-surface ocean models, *Ocean Modell.*, *20*, 61–89, doi:10.1016/j.ocemod.2007.07.005.

Marsaleix, P., C. Ulses, I. Pairaud, M. J. Herrmann, J. W. Floor, C. Estournel, and F. Auclair (2009), Open boundary conditions for internal gravity wave modelling using polarization relations, *Ocean Modell.*, *29*, 27–42, doi:10.1016/j.ocemod.2009.02.010.

Marshall, J., and F. Schott (1999), Open-ocean convection: Observations, theory, and models, *Rev. Geophys.*, *37*(1), 1–64.

O'Reilly, J., et al. (2000), SeaWiFS postlaunch calibration and validation analyses, Part 3, *Tech. Rep. 2000-206892,11*, 49 pp., NASA Goddard Space Flight Cent., Greenbelt, Md.

Schroeder, K., A. Ribotti, M. Borghini, R. Sorgente, A. Perilli, and G. P. Gasparini (2008), An extensive western Mediterranean deep water renewal between 2004 and 2006, *Geophys. Res. Lett.*, *35*, L18605, doi:10.1029/2008GL035146.

Somot, S., F. Sevault, and M. Déqué (2006), Transient climate change scenario simulation of the Mediterranean Sea for the 21st century using a high resolution ocean circulation model, *Clim. Dyn.*, *27*(7), 851–879, doi:10.1007/s00382-006-0167-z.

Somot, S., et al. (2016), Characterizing, modelling and understanding the climate variability of the deep water formation in the North-Western Mediterranean Sea, *Clim. Dyn.*, 1–32, doi:10.1007/s00382-016-3295-0.

Testor, P. (2013), DEWEX-MERMEX 2013 LEG1 cruise, *RV Le Suroit*, doi:10.17600/13020010.

Ulses, C., P.-A. Auger, K. Soetaert, P. Marsaleix, F. Diaz, L. Coppola, M. Herrmann, F. Kessouri, and C. Estournel (2016), Budget of organic carbon in the North-Western Mediterranean Open Sea over the period 2004–2008 using 3D coupled physical-biogeochemical modeling, *J. Geophys. Res. Oceans*, *121*, 7026–7055, doi:10.1002/2016JC011818.

Waldman, R., et al. (2016), Estimating dense water formation rates with an Observing System Simulation Experiment (OSSE): Case study in the Northwestern Mediterranean sea, *J. Geophys. Res.*, doi:10.1002/2016JC011694.

Yashayaev, I. (2007), Hydrographic changes in the Labrador Sea, 1960–2005, *Prog. Oceanogr.*, *73*, 242–276, doi:10.1016/j.pocean.2007.04.015.