VALIDATION OF A NUMERICAL SIMULATION OF THE NORTH ATLANTIC OCEAN WITH ARGO DATA

Degree in Marine Sciences

Laura Gómez Navarro

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

FACULTAD DE CIENCIAS DEL MAR
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Tutor: Alonso Hernández Guerra

Co-tutor: Joaquim Ballabrera Poy

June 2014

Abstract

The outputs of a numerical simulation have been validated with Argo data; with the idea of being able to use these outputs to assimilate SSS information retrieved by the satellite SMOS. In addition to Argo data, the model is also compared against the WOA09 climatology. There are a considerable amount of valid profiles available during year 2012, and the coverage does not have temporal biases, as the amount of data is homogeneous throughout the year. This demonstrates that choosing Argo to validate the simulation was a correct choice. The results show that there exists a reasonable agreement between model and Argo data in the open Ocean. However, it has been observed that the simulation still contains spatial and temporal errors probably due to a bad representation of certain oceanic processes. Nevertheless, as the errors are easily identified and do not cover the open Ocean in the tropical and subtropical regions, the simulation could still be considered as an appropriate tool to improve the remote sensing of SSS data.

Table of Contents

List of figures

List of tables

1. Introduction

The Atlantic Ocean has been thoroughly studied for more than a century, and it is still a subject of great focus for various reasons. It has a significant role in regulating the climate, of the northern hemisphere (for example, the heat transported by the water of the Gulf Stream contributes to warm the European Subcontinent). In addition, deep convection occurs in the North Atlantic Ocean. Finally, it has a human importance in maritime security, economic, social and military activities.

All of these roles of the Atlantic Ocean are due by the wind and heat forcing, but regulated by its water masses and circulation which can be seen in Figure 1.

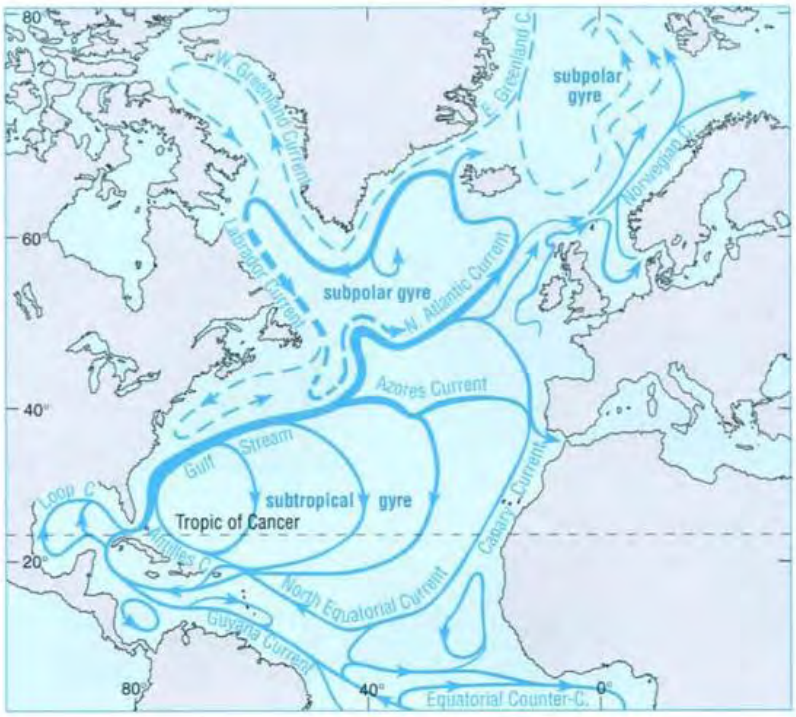


Figure 1. Schematic of the surface circulation of the Atlantic Ocean. Continuous lines are warm and discontinuous cold currents (Brown *et al*., 1988).

The Ocean circulation is modulated by local density, which is a function of temperature and salinity. According to Emery and Thomson (2004), in oceanography, *water masses* are defined in diverse ways; "real, objective physical entities, building blocks from which the oceanic stratification (vertical structure) is constructed", "mainly descriptive words, summary shorthand for pointing to prominent features in property distributions" or "a single point on a characteristic diagram such as a Temperature-Salinity (T-S) curve".

Thus, the analysis of the characteristics of the water masses present in a given region is important as it provides elements helping to describe the Oceans' circulation. In the Atlantic Ocean different water masses are found:

- North Atlantic Central Water (NACW)
    - North Atlantic East Central Water (NAECW)
    - North Atlantic West Central Water (NAWCW)
- South Atlantic Central Water (SACW)
- Antarctic Intermediate Water (AIW)
- Arctic Intermediate Water (AAIW)
- Mediterranean Intermediate Water (MIW)
- North Atlantic Deep Water (NADW)
    - North Atlantic East Deep Water (NAEDW)
    - North Atlantic West Deep Water (NAWDW)
- Antarctic Bottom Water (ABW)

Each one can be identified by their particular temperature, salinity and depth. For example the MIW, is identified easily by its high salinity and temperature. In addition, different regions have different prominent water masses (Figure 2).

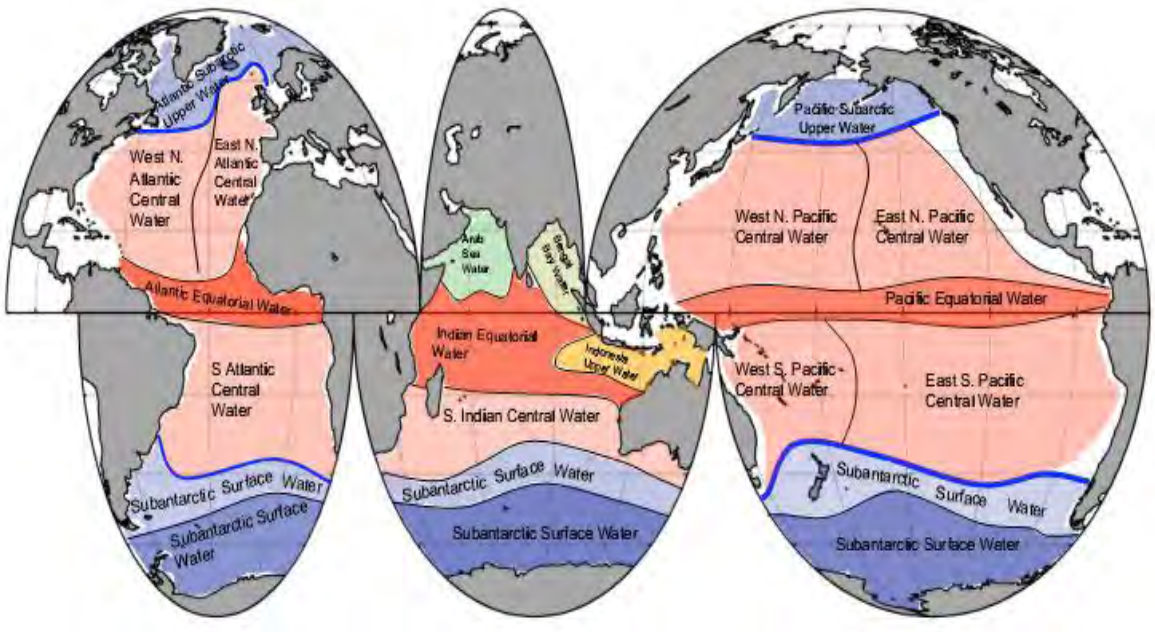


Figure 2. World distribution of the upper Oceans' water masses (Chen, 2009).

In this study, the outputs of a numerical simulation of the North Atlantic Ocean are validated. This work represents a first step towards the goal of producing a better Seas Surface Salinity (SSS) remotely-sensed product in the North Atlantic Ocean. Indeed, since the launch of the Soil Moisture/Ocean Salinity (SMOS) satellite, members of the Institut de Ciències del Mar (CSIC) are working to produce worldwide maps of SSS from the information gathered by the satellite. However, the SSS retrievals are particularly noisy in the Atlantic Ocean as the satellite measures (in the L-band, 1.4GHz) are contaminated by many radio frequency interferences in this region. The numerical model used here is going to be used to help reduce the noise to signal ratio of the SMOS SSS in the North Atlantic using various data assimilation techniques that will use the model as a dynamical interpolator.

2. Data

2.1. Model and simulation description

The numerical model solves the so called Primitive Equations with a free surface formulation (Brodeau *et al*., 2010). The regional, eddy-permitting configuration corresponds to the North Atlantic Ocean. The southern and northern open boundaries are relaxed towards the World Ocean Atlas 09 (WOA09) climatology data, to represent more realistically the Meridional Overturning Circulation (Hoareau *et al*., 2014). This configuration couples the Nucleus for European Modelling of the Ocean (NEMO) oceanic model and the Louvain-la-Neuve Ice Model (LIM2.0), which is inside the ORCA4 ¼º grid global model. The horizontal domain of this grid includes the Atlantic Ocean's basin from 80ºN to 20ºS, the Nordic Seas, the Denmark Strait and part of the Mediterranean Sea, up to 23ºE (Minvielle *et al*., 2011). The spatial resolution is 0.25 degrees. Its vertical domain is of 45 geopotential levels, and the grid spacing is 6 m at the surface and 250 m at the bottom (Hoareau *et al*., 2014). River runoff is simulated by adding a flux of water with a salinity of 0 psu and with the same temperature as the one at its mouth. The solutions are advanced in time using a Leap-Frog method, combined with a Robert-Asselin filter to ensure the stability of the scheme. The time step of the simulation is 40 minutes.

A 2001-2012 simulation using realistic forcing fields has been carried out by Ms. Nina Hoareau. The simulation requires about 6 hours for each simulated year when 60 Central Processing Units (CPUs) are used. Each output file of the simulation was saved in NetCDF format, occupied 106 Mb, and includes the five-day average of the stored parameters.

2.2. World Ocean Atlas

Climatology is, as defined by the National Weather Service of the National Oceanic and Atmospheric Administration (NOAA), "a quantitative description of climate showing the characteristic values of climate variables over a region" (www.nws.noaa.gov/climate/help/glossary.php). Climatologies are static, averaged values of a given parameter, calculated during a certain period of time. Therefore, it does not provide, in general, an observable value, but it should give an idea of the most probable value to be measured.

In particular, temperature and salinity climatologies in a given area are usually estimated by the average values of all observations ever measured at that area, and thus, they would define the mean state of the Ocean if the sampling were appropriate. The T and S observations are collected from measured surface data (as thermosalinographers) and/or profiles. However, as measurements are neither constant in time nor space, they require time-spatial averaging and smoothing (Higgison *et al*., 2009).

The World Ocean Atlas, WOA (Antonov *et al*., 2010; Locarnini *et al*., 2010) is the most widely used climatology of oceanic temperature and salinity. This World Ocean climatology was first published by Levitus in 1982, and throughout the years it has been improved by including more data and using new methods to calculate climatology, and therefore releasing new atlases every few years (Chatterjee *et al*., 2012).

The uses of these climatology atlases are diverse: boundary and/or initial conditions in Ocean circulation numerical models and Atmosphere-Ocean models, verification of numerical simulations of the ocean, as a form of "sea truth" for satellite

retrievals, for computation transports, and for planning oceanographic expeditions (Locarnini *et al*., 2010).

The WOA is distributed as a set of objectively analyzed climatological fields with a grid resolution of 1 degree, of *in situ* temperature, salinity, dissolved oxygen, Apparent Oxygen Utilization (AOU), percent oxygen saturation, phosphate, silicate, and nitrate for the World Ocean. These parameters are given at standard depth levels from 0 to 1500m, both annually, seasonally and monthly, and the data is obtained by the NOAA at the National Oceanographic Data Centre (NODC).

2.3. Argo

Argo is an upper-Ocean (1500 m) observing network including approximately 3000 profiling floats that drift freely through the global Oceans (Figure 3).

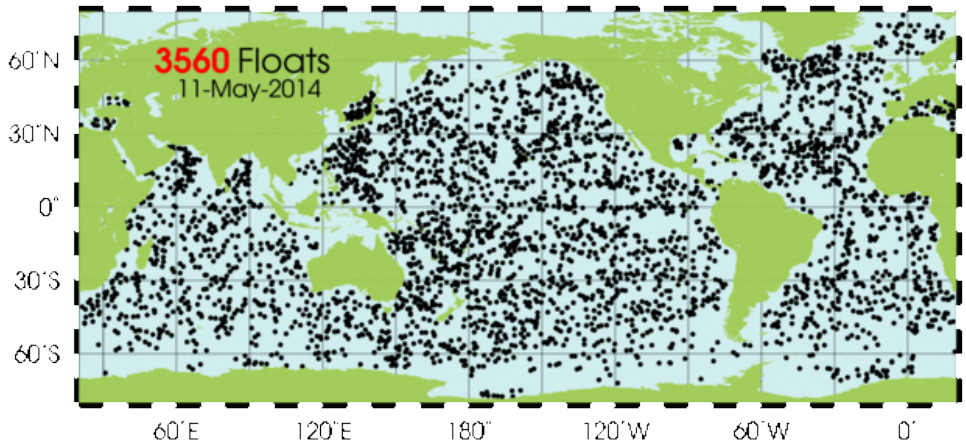


Figure 3. Argo floats distribution on 11th of May 2014 (http://www.argo.ucsd.edu/).

These floats measure pressure in decibars, *in situ* sea water temperature in degrees Celsius and conductivity from which practical salinity (dissolved salt content) is retrieved. The depths of the measurements range from the Ocean's surface to a maximum depth of 2000 m. Information about surface currents and deep oceanic circulation can be obtained from the analysis of changes in the profiler's position.

Before the year 2000 the main source of information about the Oceans' vertical structure were the Expendable Bathythermograph (XBT) probes, usually launched from commercial vessels along the main shipping lanes. The major drawbacks of XBT probes

are that they only measure temperature, and that the measures refer to the upper 500 m of the Ocean. The Argo programme was born to overcome these three limitations: The profilers are, after deployment, autonomous probes that periodically leave their parking depth to rise to the surface while taking measurements; they usually provide information about both temperature and salinity; and data is available to depths up to 2000 m (http://www.metoffice.gov.uk/weather/marine/observations/gathering_data/argo.html).

The Argo floats are battery-powered and weigh approximately 25 kg. In general, they last up to 4-5 years and make about 150 cycles (http://www.argo.ucsd.edu/). Once the floats are deployed from a vessel they follow a 10 day cycle (Figure 4). This cycle consists in a descent from surface to the parking or cruising depth, where they become neutrally buoyant. They are left to drift at this depth, and after approximately 9 days, they rise to the surface by pumping fluid into their bladder, and measure pressure, temperature and salinity whilst doing so.

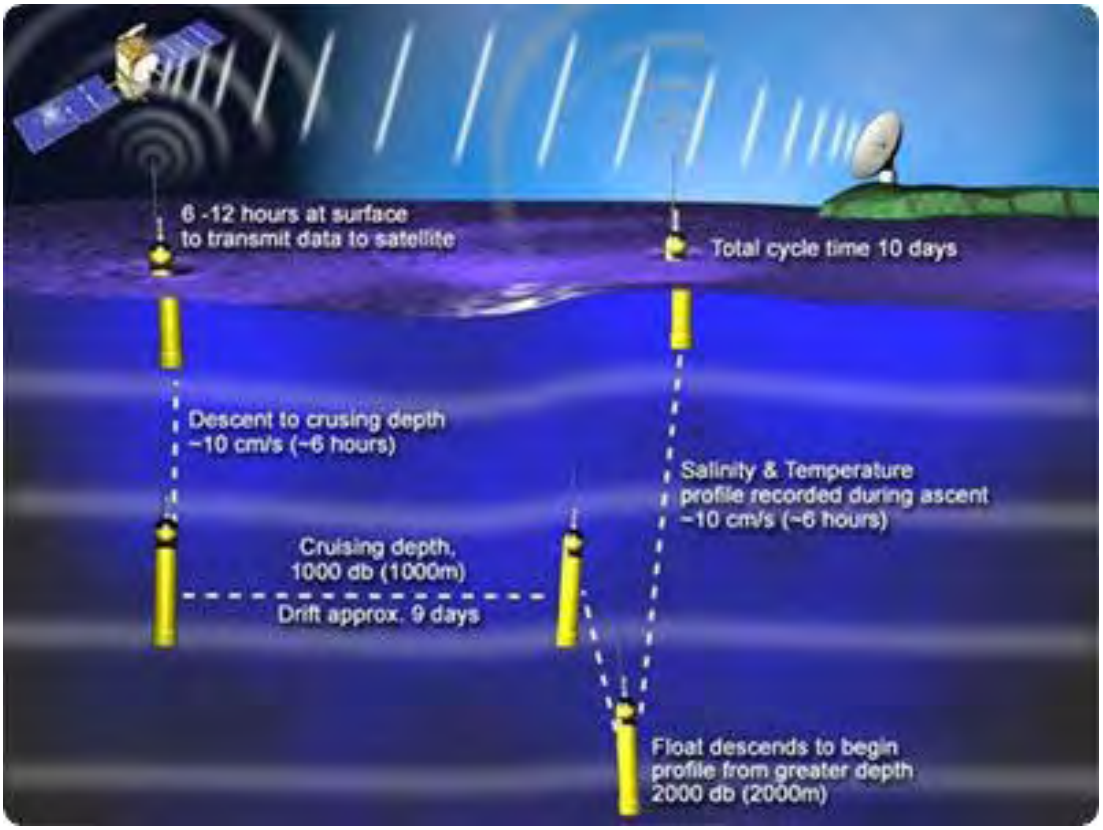


Figure 4. Description of Argo floats' cycle (http://www.argo.ucsd.edu/).

The location of the bladder and of the different sensors can be observed in Figure 5. The measurements are stored on board the float, and once at the surface, they send their information to a satellite. The majority of the floats transmit the data via the Argos location and data transmission system, which requires the float to be at the surface between 6 and 12 hours. After transmission, the profiler sinks again by deflating its bladder. Some of the new floats transmit their data using the Iridium satellites (http://www.argo.ucsd.edu/) that reduces the transmission time to 40 minutes, and allows for two way communication.

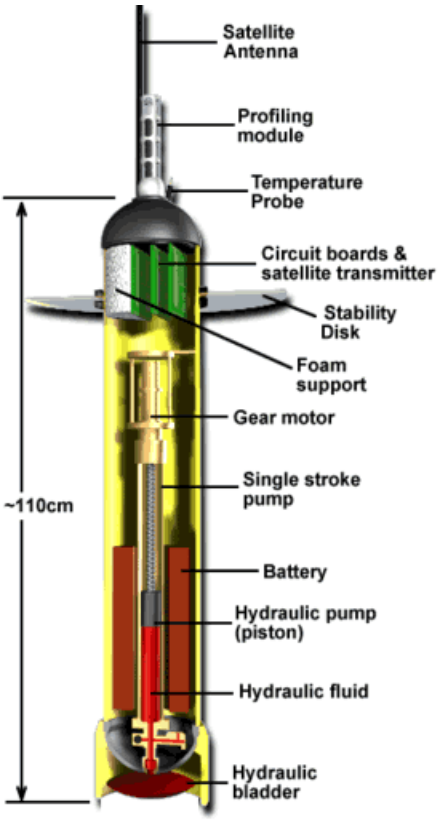


Figure 5. Different parts of one of the Argo floats models (http://www.argo.ucsd.edu/).

As a result, by now there are about five times more observations of the World Oceans than all the observations gathered before Argo. Some salinity measurements come from regions never sampled before. In addition, Argo measurements have increased. However, limits still continue to exist with Argo; they have to be deployed at ice-free Ocean so some regions are not well represented during some times of the year or at all. Table 1 shows the advantages Argo floats have provided in comparison to measurements made from ships and XBTs.

Table 1. Summarizes the advantages of the Argo dataset in comparison to other sources of data (http://www.argo.ucsd.edu/).

Observation type	T/S	Number per year	Max Depth	Geographical restriction
Ship-based temperature and salinity	T + S	5000 (to 1000m)	Full water depth	• Limited by ship endurance (100 per month) • Few at high latitude in winter • Typically along lines
Expendable XBT from merchant ships	T	25,000	750m	• Along shipping routes • Avoid high latitude in winter • Many areas unsampled
Argo	T + S	42,000 (May 2004) 100,000+ (2013)	2000m	• Ice free areas deeper than 2000m

Other advantages are that there is little or no seasonal bias in the sampling of the Ocean, as the floats can operate throughout all the year; and the fact that the data is freely available through the two Argo Global Data Acquisition Centres: the United States Global Ocean Data Assimilation Experiment (USGODAE) server and Coriolis (in France). The data is made available in two different modes: real-time and delayed mode. The real-time data is available within 24 hours after the Argo has transmitted it. The delayed-mode data has gone through a quality control and the data is available within 1 year. This strategy allows the data to be monitored continuously while, after going through a first set of filters in the different data centres, being made available quickly to the public.

The quality controls mentioned above involves different criteria and an inflicted number depending on the quality control applied to the data, for example if it passed a statistical test or if the data was interpolated (Table 2). In the Argo data files, each variable has a quality control number, for example Z_QC, T_QC, S_QC, for pressure, temperature and salinity respectively.

Table 2. Argo quality control flag scale (Carval et al., 2012).

n	Meaning	Real-time comment	Delayed-mode comment
0	No QC was performed	No QC was performed.	No QC was performed.
1	Good data	All Argo real-time QC tests passed.	The adjusted value is statistically consistent and a statistical error estimate is supplied.
2	Probably good data	Not used in real-time.	Probably good data.
3	Bad data that are potentially correctable	Test 15 or Test 16 or Test 17 failed and all other real-time QC tests passed. These data are not to be used without scientific correction. A flag '3' may be assigned by an operator during additional visual QC for bad data that may be corrected in delayed mode.	An adjustment has been applied, but the value may still be bad.
4	Bad data	Data have failed one or more of the real-time QC tests, excluding Test 16. A flag '4' may be assigned by an operator during additional visual QC for bad data that are not correctable.	Bad data. Not adjustable.
5	Value changed	Value changed	Value changed
6	Not used	Not used	Not used
7	Not used	Not used	Not used
8	Interpolated value	Interpolated value	Interpolated value
9	Missing value	Missing value	Missing value

3. Methodology and results

3.1. Preliminary data exploration.

Four different variables of the numerical simulation were analyzed: potential temperature (Tm), salinity (Sm), Sea Surface Height (SSH) and the Mixed Layer Depth (MLD). To read these variables from the output files and to carry out a first analysis of them, a Fortran 90 program was created using the NetCDF library.

The simulation reproduces important oceanic processes and features observed in the North Atlantic Ocean. Figure 6 depicts the raw output of one of the SSH: the Gulf Stream, the North Atlantic Subtropical gyre, the equatorial upwelling, and the extent of Arctic waters can be detected.

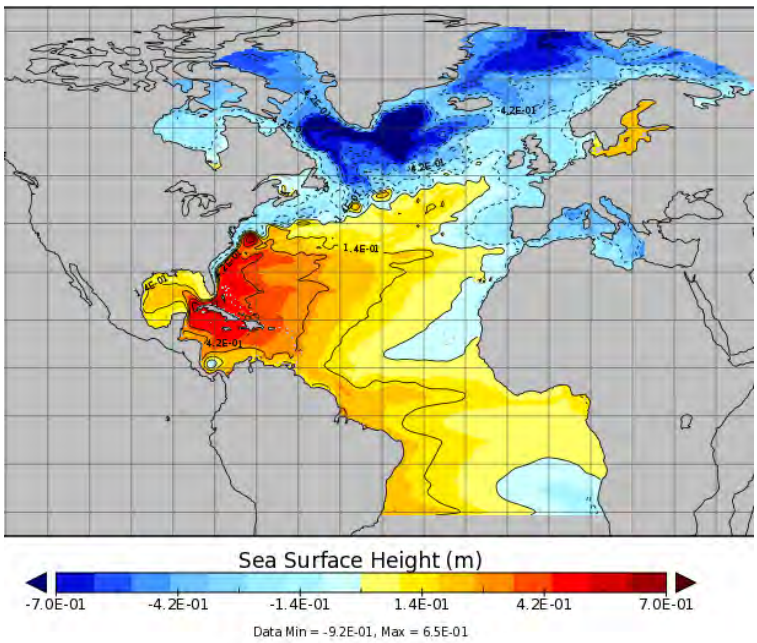


Figure 6. Sea Surface Height (m) in 2008 (obtained from Ms. N. Hoareau).

In a first preliminary step, only one output file of the simulation was processed. The data corresponded to the time period from the 22/01/2001 to the 26/01/2001. In this analysis, two types of graphs were made, using the Matlab: histograms and contour maps. In the case of Tm and Sm, each depth of the model was processed separately. In addition, a set of statistical parameters were calculated (Table 3).

Table 3. Description of the statistical parameters calculated.

Statistical parameter	Abbreviation	How it was calculated
Maximum	Max	Maxval
Minimum	Min	Minval
Range	R	Max-Min
Robust maximum	p99	*
Robust minimum	p01	*
Robust range	Rr	p99 – p01
Upper quartile or percentile 75	UQ	*
Lower quartile or percentile 25	LQ	*
Interquartile range	IQR	p75 – p25
Median or percentile 50	Med	*
Mean	Av	**
Variance	Var	**
Standard deviation	Sd	Using the intrinsic function of Fortran 90 to obtain the square root, sqrt(var).

** To see how it was calculated go to Appendix 2.*

*** To see how it was calculated go to Appendix 3.*

The maximum and minimum parameters, when compared to the robust maximum and minimum, helped to identify anomalous results, very particular of certain regions and/or of moments in time. On the other hand, p99 and p01 are more representative of the data set as they eliminate the one percent most-extreme cases.

What could be observed in this first analysis, specifically in the contour maps, was that the land values had a value of 0 for Tm and of 0.1 for Sm. This had to be taken into account and thus the programme created in the first place had to be modified.

The analysis of the histogram revealed the prominence of Arctic water and, in a lesser extent, of tropical water. Because any statistical analysis of the basin's statistics would be strongly influenced by these two water masses, it was decided to reduce their impact by limiting the latitude range of the data to 60ºN and 5ºS. In addition, the dynamics of the model near the north and south open boundaries is perturbed by the relaxation of the model solutions to the climatological values. Near the open boundaries, some spurious values can be sometimes observed.

Figures 7 and 8 illustrate the effect that limiting the latitudinal range has in the potential temperature. Figure 7 shows the histogram of values of the first layer (thickness is 6 m, and the centre of the layer is located at 3 m below the surface). By limiting the latitudinal extent of the data to be analysed, the amount of water below 10 degrees C is reduced, in particular, the amount of water near, or below 0. Figure 8 shows the mean and the median of the potential temperature as a function of depth. While the limitation of the meridional extent leads to an increase of the average

temperature, it is interesting to notice the reduction of the differences between the mean and the median. The larger the differences between the mean and the median, the larger is the impact of the tails of the distribution, and the less significant are the values of the statistical parameters listed in Table 3. On the other hand, the results shown in Table 4 indicate that the statistical description of SSH is little affected by the latitude limits, but the MLD is.

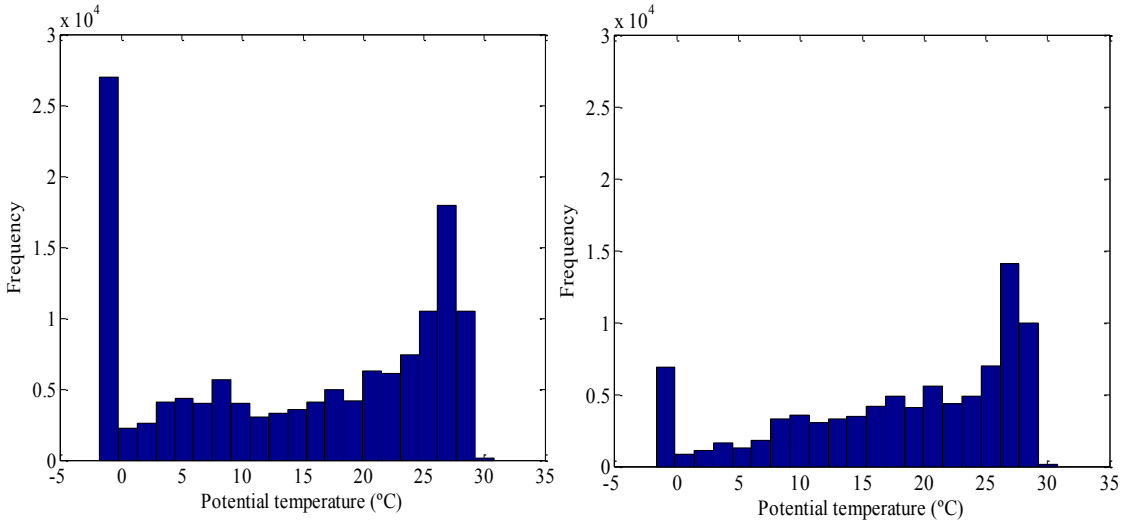


Figure 7. Histograms of potential temperature data at layer 1 (3m). Left: Original data. Right: With latitude limits and land values changed.

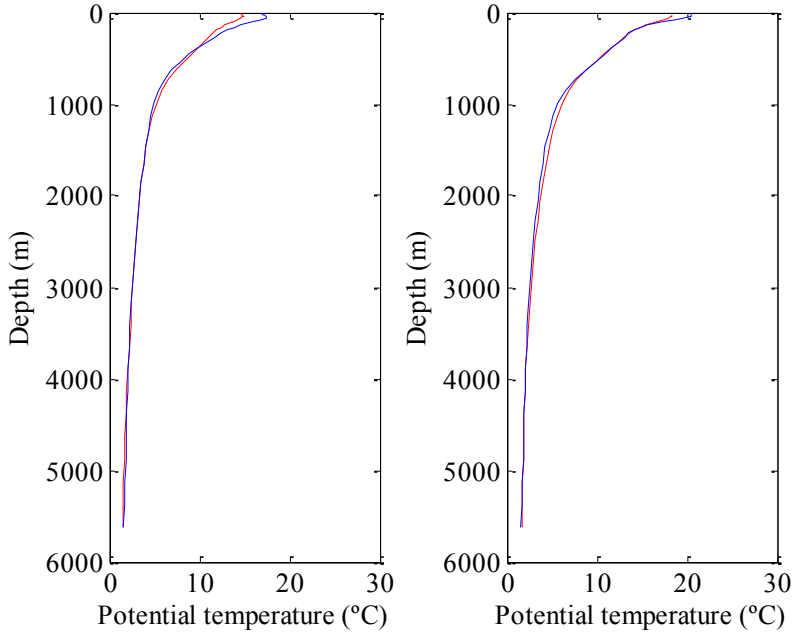


Figure 8. Change of mean (red line) and median (blue line) values with depth for the original data (left) and with latitude limits and land values changed (right).

Table 4. Statistical parameters of SSH and MLD for the original data and columns marked with* are for data between 60ºN and -5ºN.

Parameters	SSH (m)	SSH* (m)	MLD (m)	MLD* (m)
Maximum	0,79	0,79	393,84	249,92
Minimum	-0,94	-0,94	12,84	12,84
Range	1,74	1,74	381,00	237,08
Percentile 99	0,53	0,55	189,89	139,91
Percentile 1	-0,87	-0,89	12,84	12,84
Robust range	1,40	1,44	177,05	127,08
Percentile 75 (Upper quartile)	0,13	0,17	74,89	77,93
Percentile 25 (Lower quartile)	-0,38	-0,24	27,09	28,89
Interquartile range	0,51	0,41	47,79	49,04
Median	-0,12	0,03	45,82	59,89
Mean	-0,12	-0,04	56,28	57,62
Variance	0,11	0,10	1530,65	1010,65
Standard deviation	0,33	0,32	39,12	31,79

Even though the presence of extreme values has been reduced, some histograms remain bimodal (see Figure 7). With the help of salinity maps (like the one in Figure 9), some regions could be identified, where there was still anomalous data. These corresponded to peripheral regions, specifically the Mediterranean Sea, the North and Baltic Seas and the Hudson Bay. Figure 9 shows one of the contour maps used to establish the peripheral regions, particularly in this case the Baltic Sea and Hudson Bay due to their low salinity.

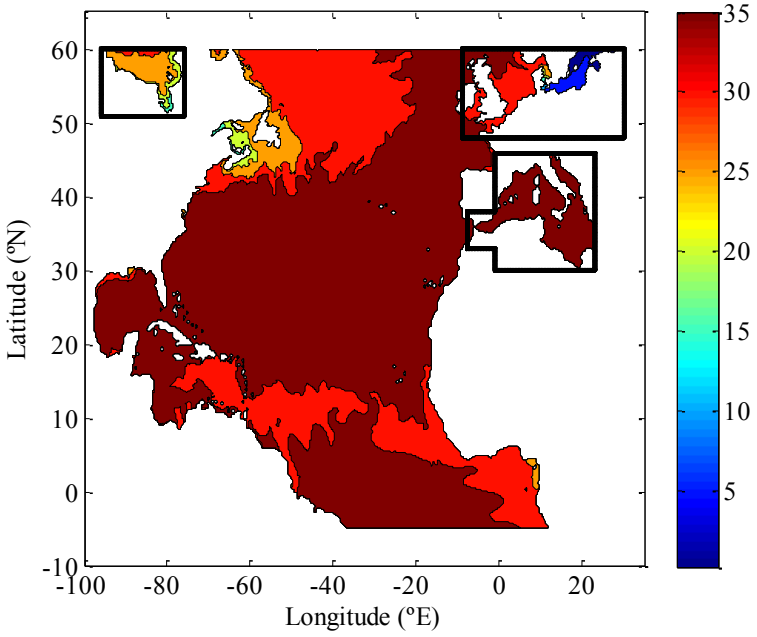


Figure 9. Contour map of salinity (psu) at layer 1 (3m) with the peripheral regions marked.

Figure 10 shows a T-S diagram for the whole region (including peripheral regions). Some characteristic water masses stand out from the rest of the system. The Mediterranean water (12-20ºC S>35), the Baltic and North seas' water (T<10, wide range of S), and the Hudson Bay water (T<3ºC, 15<S<33), all appear concentrated around localized regions of the TS plot. Moreover, the circles circled in purple correspond to the Chesapeake Bay (15<T<18ºC, 25<S<35) that also stands out from the rest of the system.

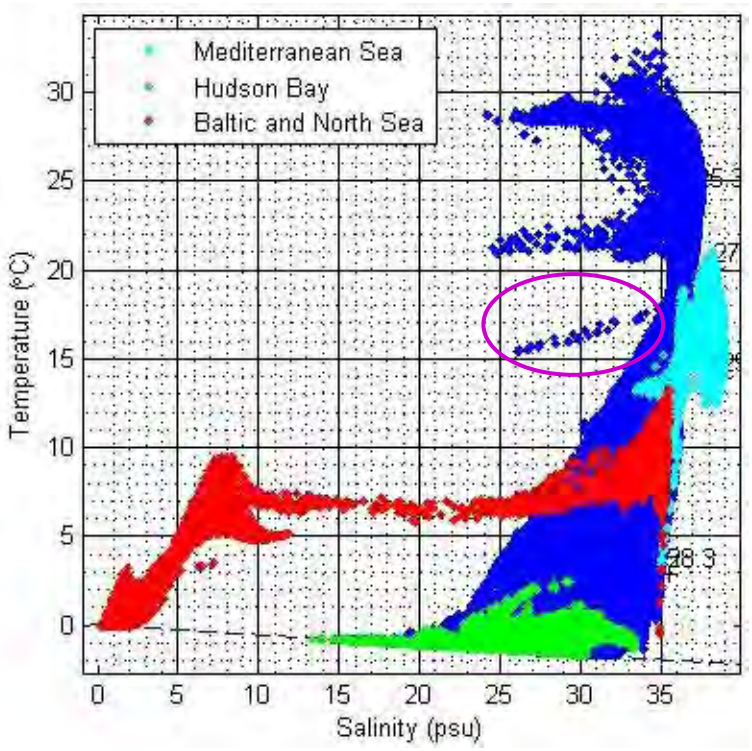


Figure 10. T-S diagram of all the simulation data.

The T-S diagram also allows to identify water from particular regions: East coast of Canada (from the Gulf of Saint Lawrence up to Ungava Bay), Mississippi-Alabama Shelf and Marshland Island, Gulf of Paria (Trinidad Island), Amazon River delta, Gulf of Guinea, particularly at the Bight of Bonny and Chesapeake Bay already identified as a region of anomalous salinity (Figure 11).

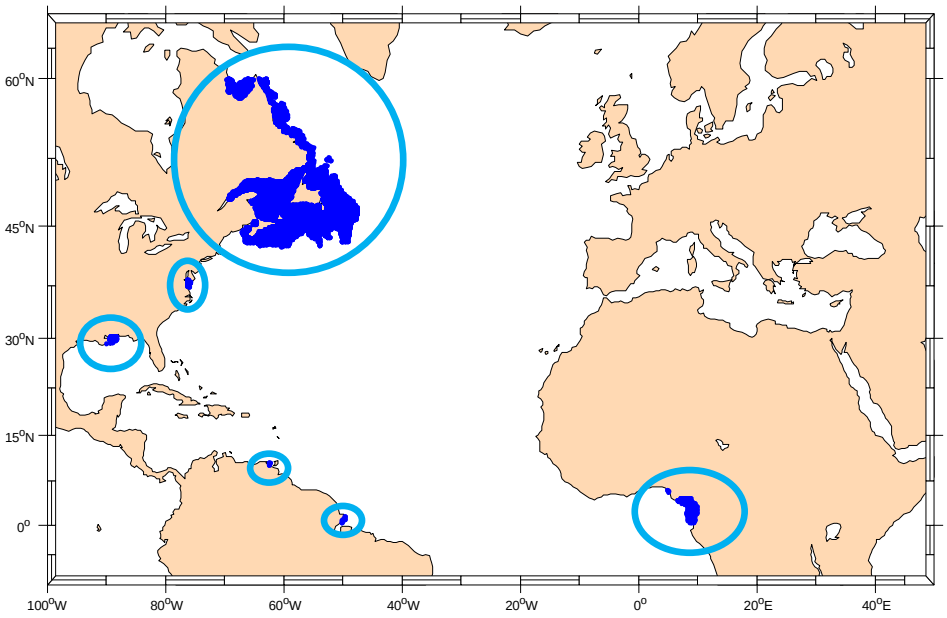


Figure 11. Location of Sm values lower than 30 psu.

3.2. The 2012 North Atlantic numerical simulation

After the first set of preliminary experiments was done, all the files corresponding to the year 2012 were processed. This corresponds to a total of 73 files in NetCDF format. The procedure carried out was the same as the previous one with the exception that the statistical parameters were calculated both in space (2D) and time (1D). For example, at each grid-point, the maximum value reached during the whole 2012 can be estimated, and a (2D) map of the maximum values can be constructed. Alternatively, the maximum value at each snapshot can be calculated over the region of interest, and a (1D) time series of regional maximum values can be created.

Figure 12 displays the map showing the maximum 2012 values of the sea level and the mixed layer depth. Although the graph does not include information about the time of occurrence of the maximum value, it clearly is a tool to delineate the different dynamical provinces of the North Atlantic Ocean. It also helps to illustrate the main mechanisms at work in the region as the 1 m sea level near the western boundary due to the wind driven oceanic circulation (combined with thermal effects), as well as the relative size of the resulting meridional and zonal gradients at the region. On the other hand, the largest values of the mixed layer depth are found at the western Atlantic Ocean, near the path of the Gulf Stream. The lowest values are concentrated at the

north-eastern coast of Canada and at the Gulf of Guinea, two regions strongly influenced by the presence of surface fresh water fluxes.

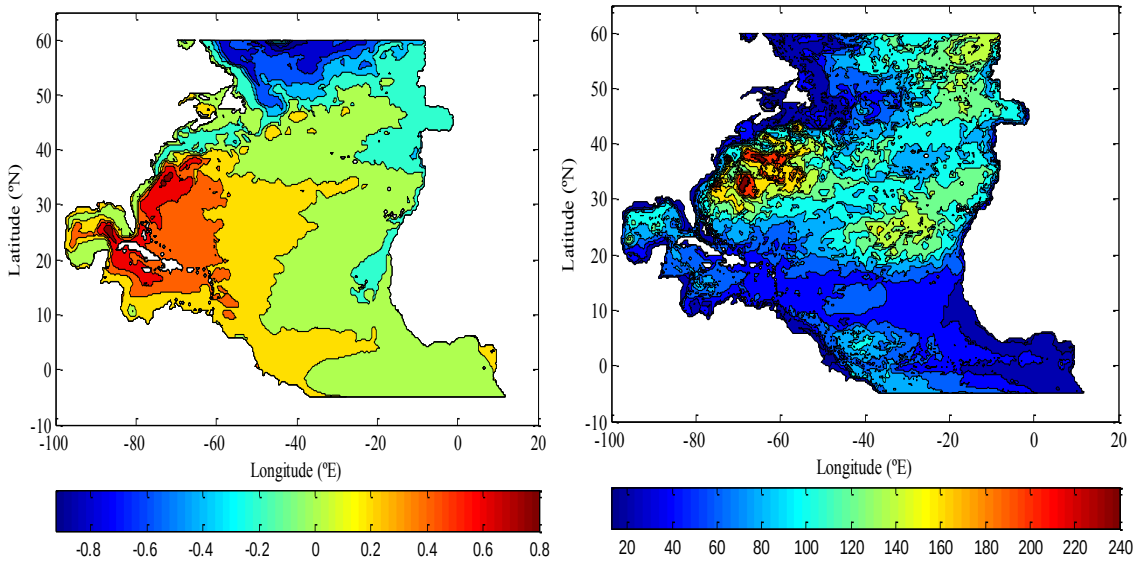


Figure 12. Maximum values' distribution in metres for SSH (left) and MLD (right).

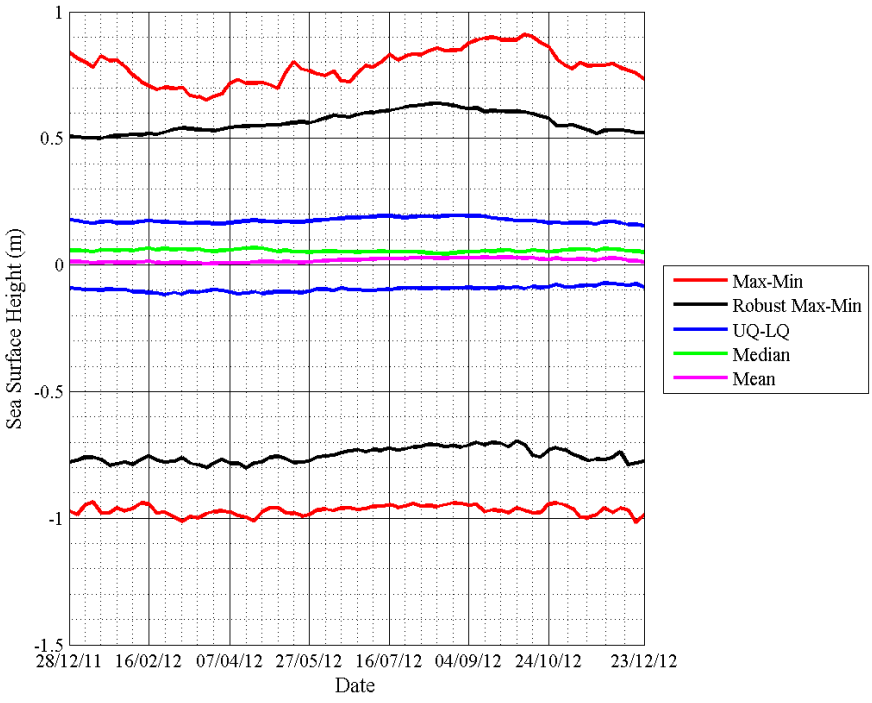


Figure 13. SSH's statistical parameters variation with time.

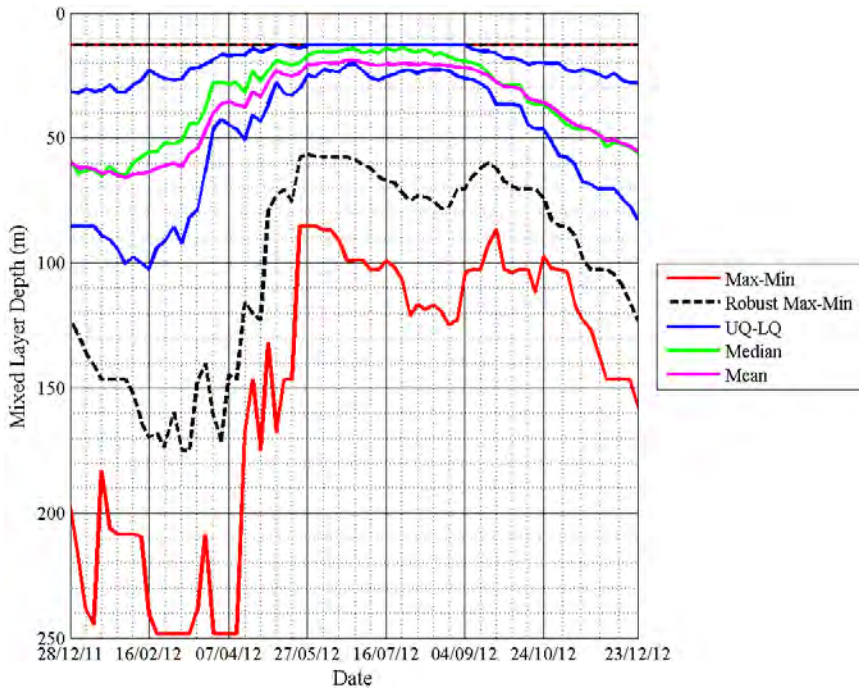


Figure 14. MLD's statistical parameters variation with time.

An example of time series of the statistical properties of the system is given in Figure 13, where the time evolution of the minimal, central, and largest sea level values is shown. The mean value remains near to zero as the model uses a volume conservation constrain in the parameterization of the open boundaries. The departures of zero are due to the non-accounted contribution of the peripheral seas. The time evolution of the robust maximum (the 99 percentile) reaches its largest value during the month of August, indicating the seasonal warming of the upper Ocean. Contrarily, the time evolution of the absolute maximum has two maximum values (at the beginning of the year and towards the end of the month of October). The more reasonable evolution of the robust maximum illustrates the need of removing extreme values before the statistical exploration of the data.

Figure 14 helps to visualize the strong seasonal variation of the MLD. As it would be expected, the maximum values occur during winter due to the combined effect of low stratification and strong wind stirring, and the lowest values occur during the summer season due to the stabilizing effect of the surface seasonal warming.

Figure 15 shows the corresponding time series for the salinity at layer 15 (centred on 228 m depth). The most striking feature is the anomalous behaviour of the minimum salinity towards the end of the year. Further analysis have shown that such

sudden reduction of salinity happens in some isolated grid-points, and that are not linked to any physical process. The analysis was repeated to the year 2011, and it was found that some grid-points, located at different depths, had similar errors. All points affected by this behaviour are located at the bathymetry, indicating a problem in the parameterization of the tracers in the bottom mixed layer.

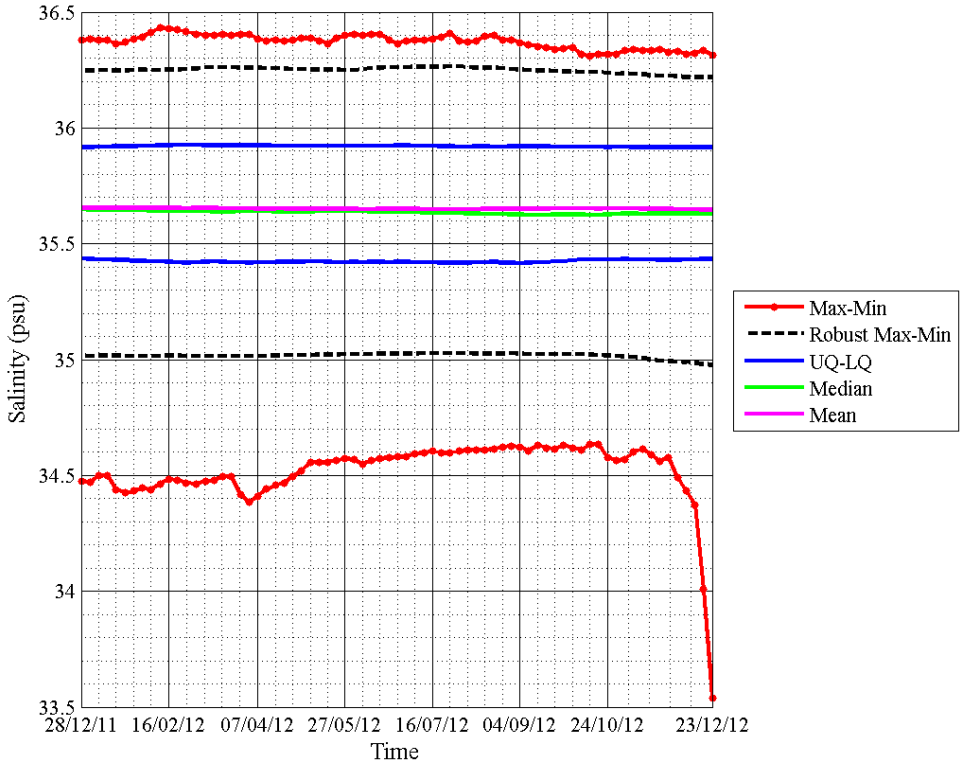


Figure 15. Salinity 1D plot at layer 15 (228m) for year 2012.

The spatial distribution of the minimum salinity during the year 2012 in the Gulf of Saint Lawrence are shown in Figures 16 (32 m depth) and 17 (191 m depth). The model values are compared with the WOA09 salinity data. The model surface salinity can be much lower than the climatological salinity (more than 10 psu fresher). At depth, the differences are of about 1 psu. Such a 10 psu surface salinity differences between the model and climatology indicates deficiencies of the model in obtaining realistic fresh-water balances in such region dominated by local water runoff and shallow bathymetry.

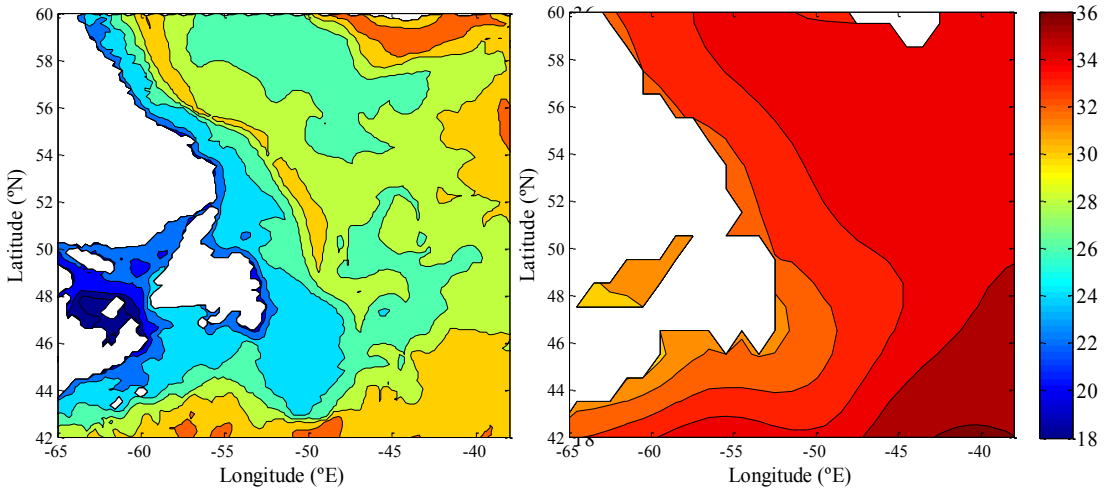


Figure 16. Left: Minimum salinity values' distribution for 2012 layer 5 (32m). Right: WOA09 Salinity values for December 2012 layer 4 (30m).

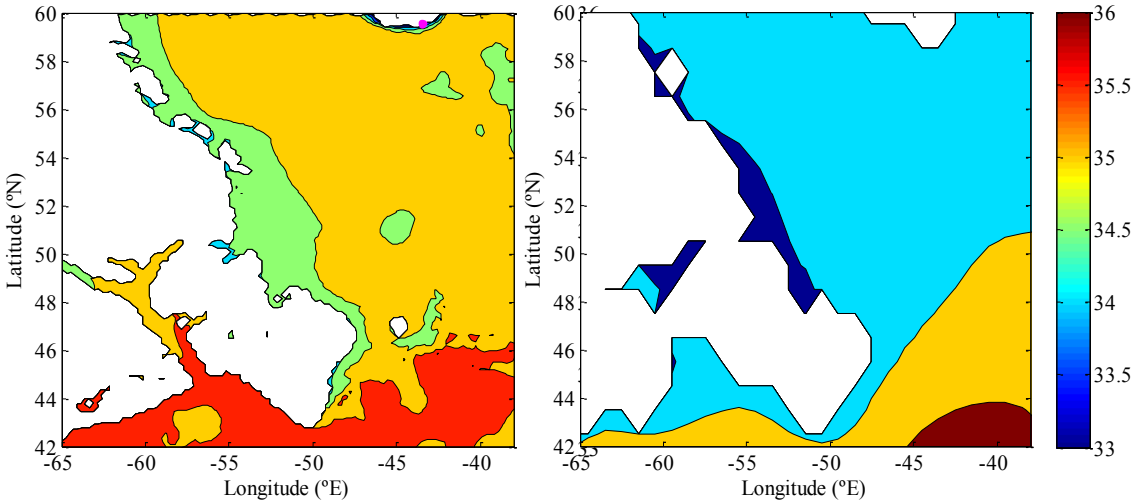


Figure 17. Left: Minimum salinity values' distribution for 2012 layer 14 (191m). The pink point is where the lowest value is located. Right: WOA09 Salinity values for December 2012 layer 10 (200m).

3.3. The 2012 North Atlantic Argo data

The Argo data corresponding in time and space to the simulation was obtained from the Barcelona Expert Centre (BEC) database which is a mirror of the Coriolis database. Only the profiles which were found in the Atlantic Ocean, between 60ºN and 5ºS and of the year 2012, were retrieved from the database.

Only Delayed mode data and profiles outside the peripheral regions have been considered in this part of the work. The hydrographic profiles where then processed following these steps:

a. Adjusted pressure, temperature and salinity variables were read.
b. If the number of measurement levels is less than 20, the whole profile is dismissed
c. If Z_QC at a point was not equal to 1, the corresponding pressure, temperature and salinity values were set to 99999 (missing value).
d. If T_QC of a point was not equal to 1, the corresponding temperature and salinity values were set to missing.
e. If S_QC of a point was not equal to 1, the corresponding salinity value was set to missing. This is done as the salinity values depend on both temperature and pressure, while temperature value depends only on pressure.
f. For each retained Argo profile, a profile is created in such a way that only valid salinity data are stored.
g. The number of data in each profile for these new variables was counted. If there were not at least 20 values or if more than 40% of the data in the original profile was discarded, the whole profile was discarded. Also, for this data filter, all the profiles which their first valid salinity measure was at depth lower than 100m were discarded too.
h. As the temperature measured by the buoys is *in situ* the potential temperature was calculated. See Appendix 1 to see the programme used for its calculation.
i. The potential temperature values were vertically interpolated.
j. The data was interpolated using three interpolation methods:
    - *Akima splines*
    - *Cubic splines*
    - *Third-order polynomial fitting*
k. When the Argo data was interpolated to the simulation's vertical levels, any level of the simulation deeper than the last depth of the profile and shallower than the first valid depth of the profile was changed to missing.
l. After the interpolation the mean was calculated.
m. If the absolute difference between each individual interpolation and the mean was greater than 5% of the mean, the corresponding potential temperature value was changed to missing.
n. Steps g) to k) were repeated for salinity.

The final results of potential temperature and salinity of the Argo data hereafter referred to as Ta and Sa respectively, were saved in a NetCDF file. In this file, information about each profile was saved too: its longitude, latitude, date and the ID of the profile (Did) in the BEC database. In addition, the additional parameters were calculated and included in this file. The indices ipos and jpos are the values of the x and y dimension respectively, of the point in the model's grid which is closest to the Argo profile's position. The tpos variable points to the model snapshot that corresponds to the Argo measurement.

In addition, the number of valid Ta and Sa values for 3 different surface depths; 10 m, 25 m and 50 m, and in different points of the data processing (after having applied different filters). This was done to know how much data was really available and valid, as particularly in the surface, there are many Argo data problems, but on the other hand, it's the surface data which is of greater interest to validate, for its use for SMOS SSS.

From a total of 4121 Argo profiles, only 3117 profiles have passed the Quality Control described above (Figure 18). Some regions, as the North Atlantic subtropical gyre, show a clear lack of Argo data, due to the free advective nature of the buoys.

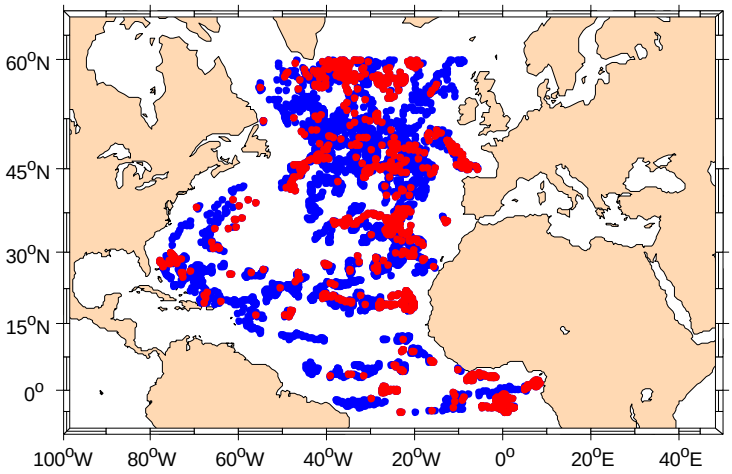


Figure 18. Argo profiles for 2012. The valid profiles are marked in blue and the invalid in red.

The spatial distribution of the Ta (Figure 19) and Sa (Figure 20) is shown for layers 1 (3 m) and 2 (9 m). Notice the lack of Argo data for the first layer of the model.

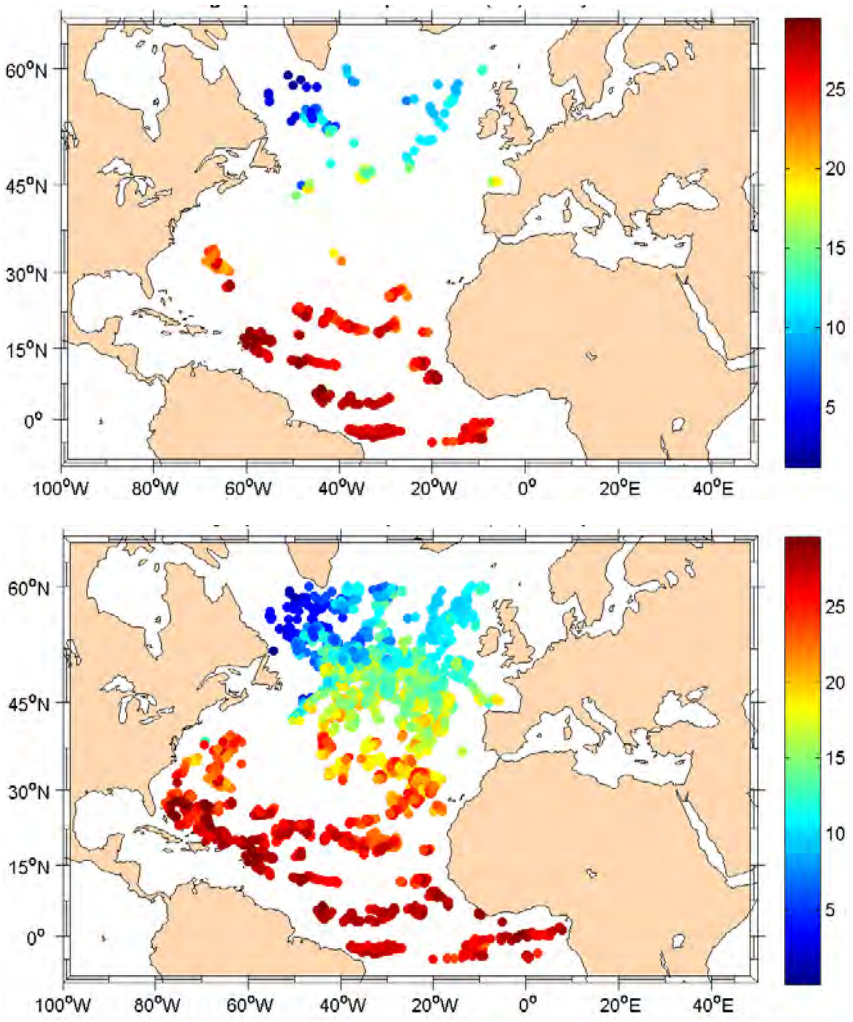
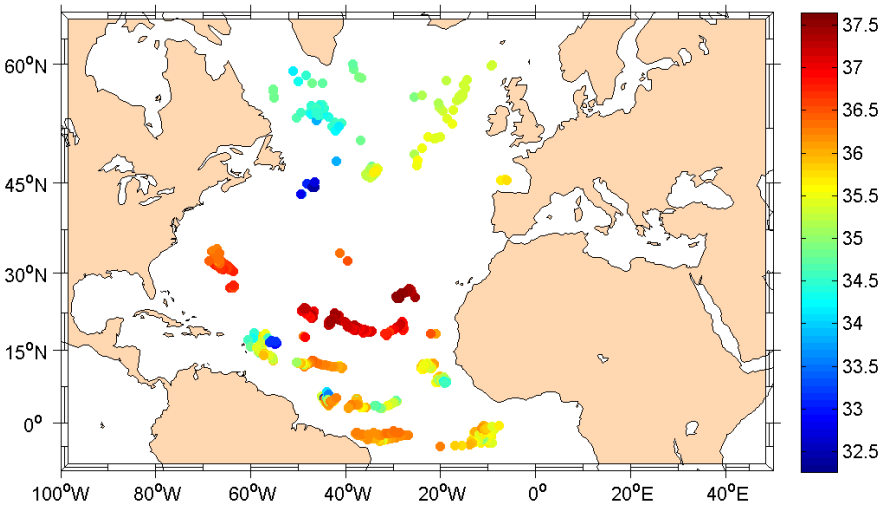


Figure 19. Potential temperature (ºC) of the Argo profiles for layer 1 (top) and layer 2 (bottom).



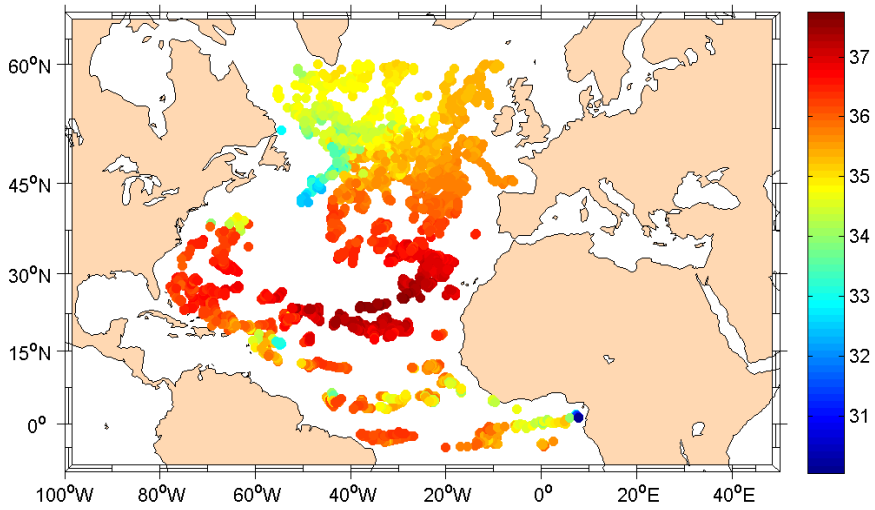


Figure 20. Salinity (psu) of the Argo profiles for layer 1 (top) and layer 2 (bottom).

The time series of the statistical parameters are shown in Figure 21. The seasonal variation can be observed for both Ta and Sa at the surface layers. Comparing Figure 21 with Figures 13, 14 and 15, it can be seen that the Argo statistics display less variability than the model. This might be due to the fact that the Argo data correspond to the pelagic Ocean, and does neither sample at all the continental shelf, the coastal, nor the inland waters.

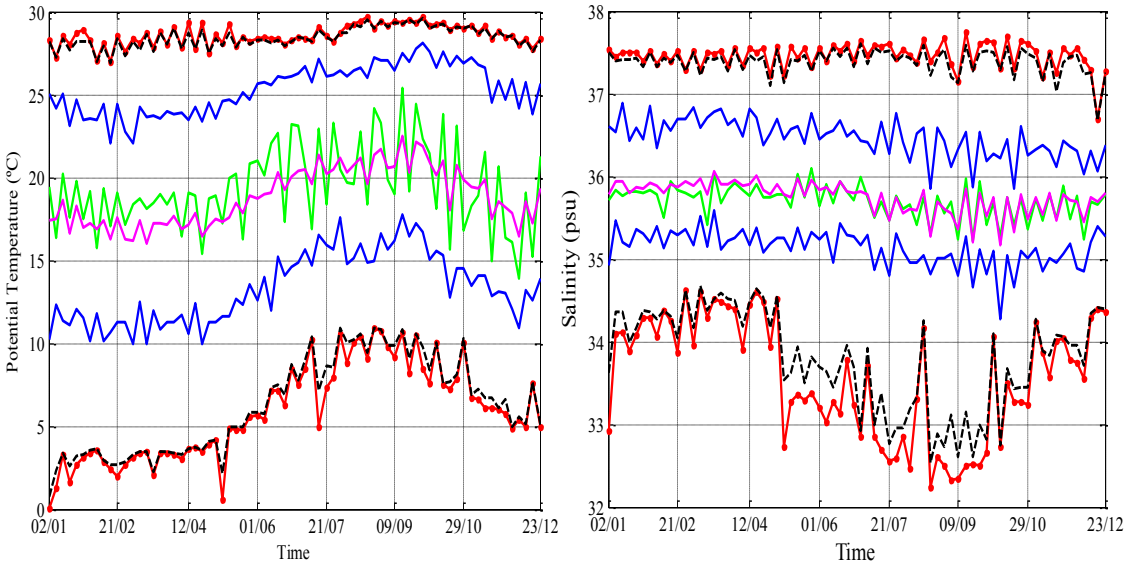


Figure 21. Ta (left) and Sa (right) statiscal parameters variation with time for layer 3 (16m). The legend is the same as in figure 15.

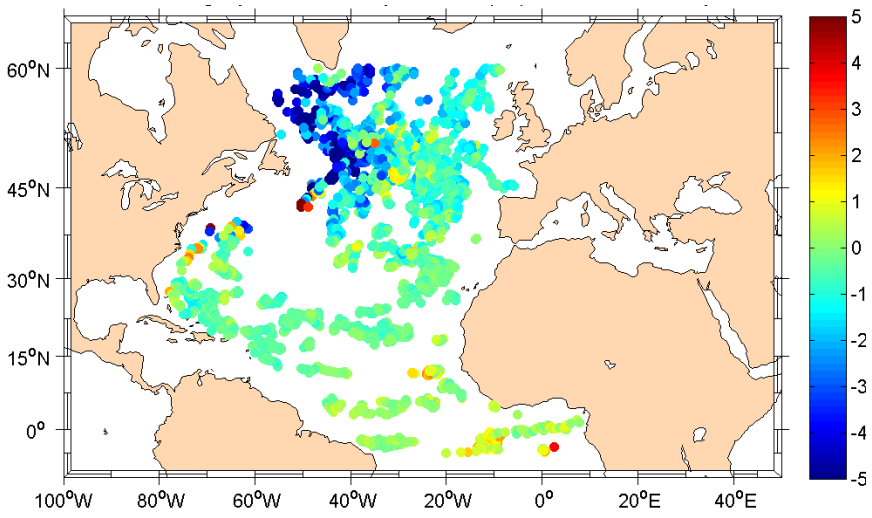
3.4. Simulation-Argo differences.

Finally, the difference between the model and Argo is calculated by taking the model grid-point closest to the Argo position:

$$\Delta T = Tm\ (ipos, jpos, layer, tpos) - Ta\ (layer, profile)$$

$$\Delta S = Sm\ (ipos, jpos, layer, tpos) - Sa\ (layer, profile)$$

In the scatter maps for ΔT, it can be observed that the model is colder than Argo data at the surface layers, warmer than Argo at intermediate layers, and very similar at deep layers (Figure 22). At the surface, the model is colder (up to 5ºC at high latitudes), but slightly warmer (1ºC) at the equator. The spatial coherence of these differences indicates that, at firs order, the errors of the model are not random, but arising from errors of the physical processes that the model fails to properly represent. Similarly, the error at layers 14 and 29 also displays a clear spatial coherence.



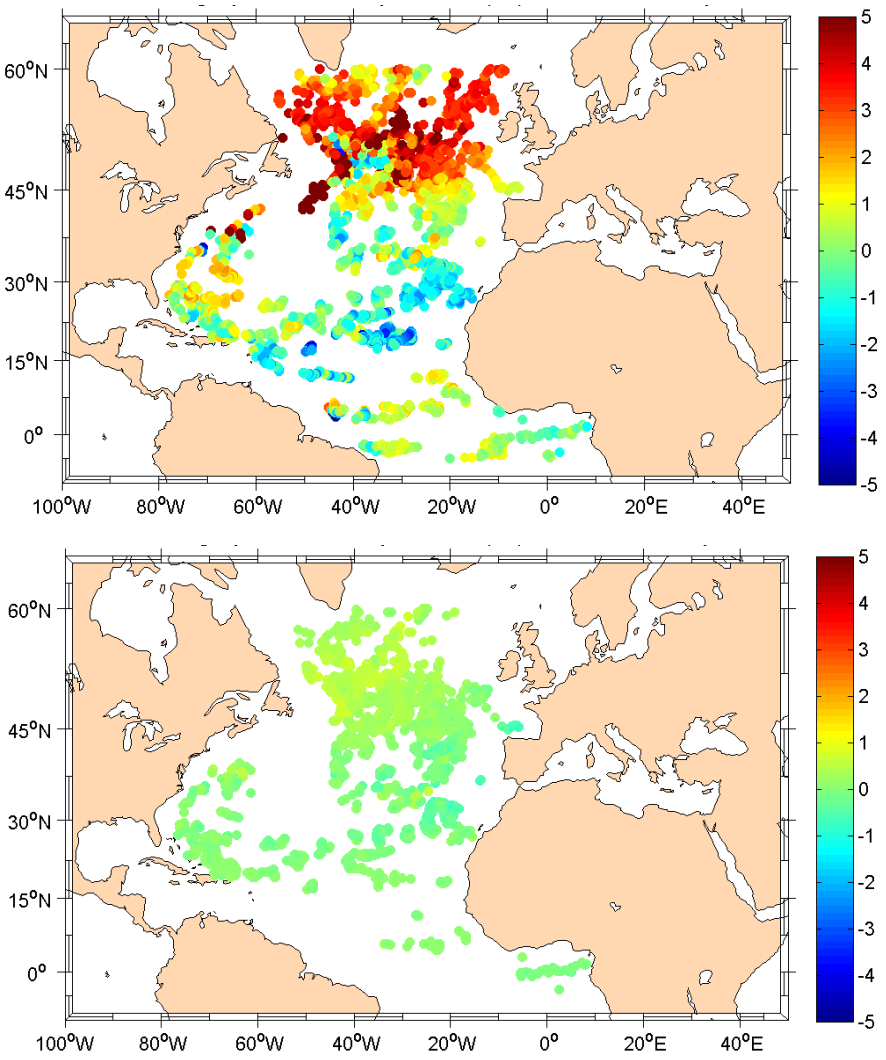


Figure 22. Scatter map of the ΔT (ºC) for layer 2, 9 m (top), layer 14, 191 m (middle) and layer 29, 1850 m (bottom).

For the ΔS, it can be observed that the model surface salinity is much fresher than Argo (Figure 23). Similar excessive freshness has been found up to layer 12 (134 m), especially at the northwest of the Atlantic Ocean. Deeper than layer 12 the difference between sets of data is much smaller. The size of the differences is further reduced as depth increases.

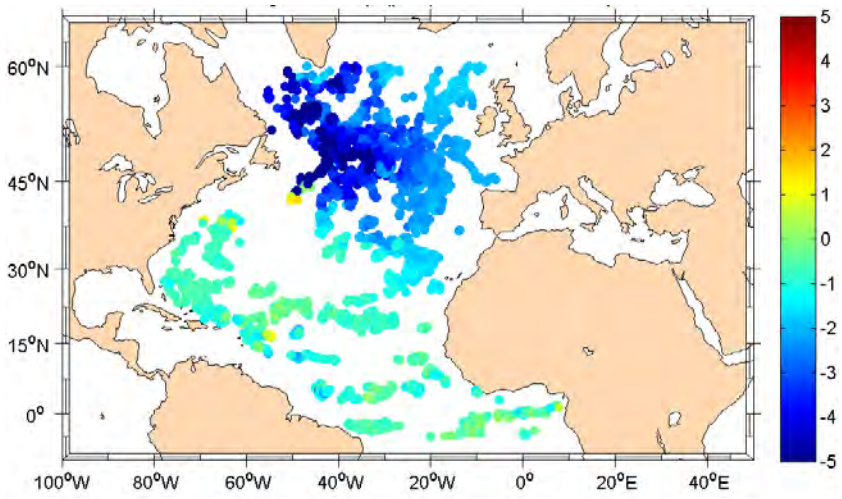


Figure 23. Scatter map of ΔS (psu) for layer 2 (9m).

Figure 24 shows the difference of temperature and salinity as a function of the buoy number. All the profile is shown. The salinity has a clear negative bias. In contrast, the ΔT values have a positive bias, although smaller than the salinity one. Figure 25 shows the location of all the points with a model minus Argo difference larger than 10ºC (blue circles), and salinity differences larger than 6 psu (red circles).

While all the large salinity differences appear located off the Canadian Coast, the temperature differences are located both at high and low latitudes. Although the differences in the Gulf Stream region may be just due to misplacement of the temperature front, the differences at low and high latitudes might be due to systematic model errors.

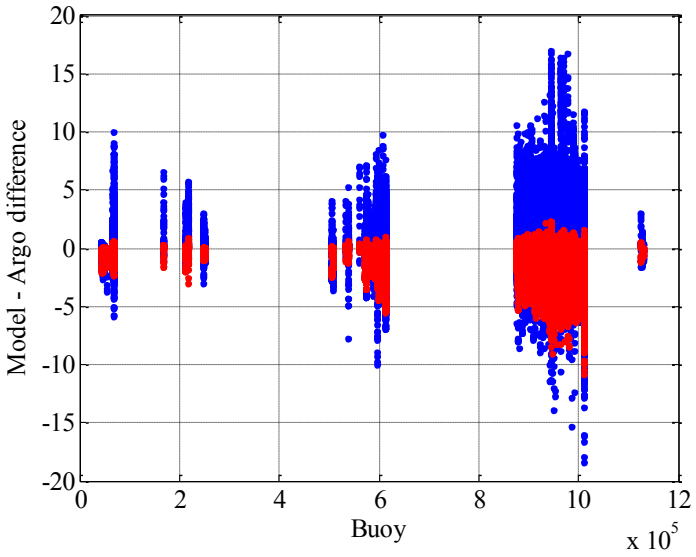


Figure 24. ΔT (blue) and ΔS (red) values for every buoy.

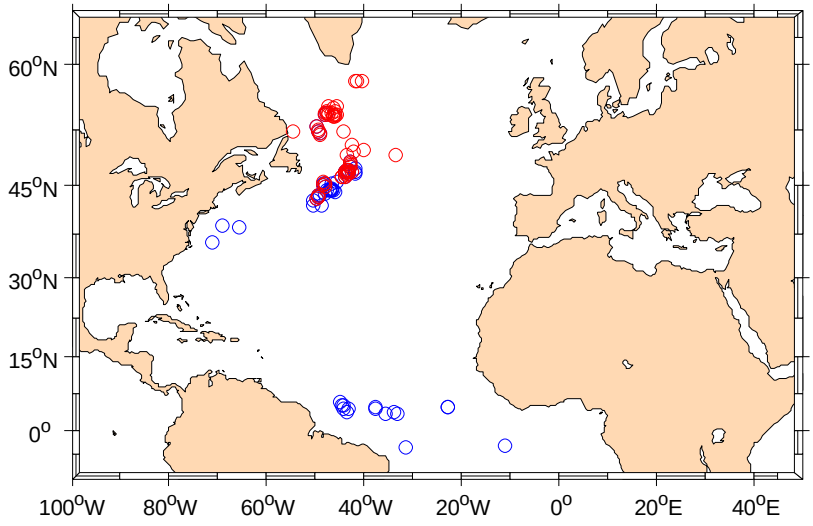


Figure 25. Location of the Argo buoys with an absolute value of ΔT greater than 10ºC (blue circles) and with an absolute value of ΔS greater than 6 psu (red circles).

Moreover, the differences between model and Argo seem to have a seasonal dependency at the surface layers, especially for the large temperature and salinity differences. The difference is larger during the end of summer beginning of autumn (Figure 26).

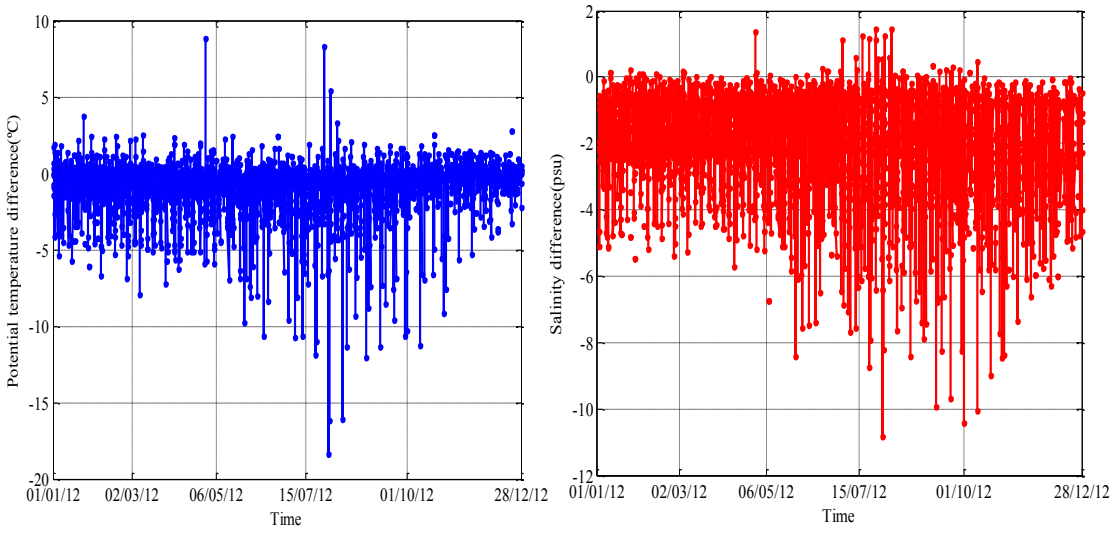


Figure 26. Variation of ΔT (left) and ΔS (right) with time for layer 2 (9m).

Finally, as the T-S diagram shown in Figure 10 did not really help in identifying the different water masses present in the region, a set of 12 T-S diagrams have been created for three randomly chosen Argo profiles in four different regions (Figure 27).

These 4 regions are the ones shown in Figure 2 (Chen, 2009). The dates of the 12 randomly selected profiles are shown in Table 5.

Table 5. Dates of the 12 profiles selected for the T-S diagrams.

Buoy number	Day	Month	Year
1	12	10	2012
3	6	1	2012
2	11	10	2012
4	6	2	2012
5	1	2	2012
6	3	7	2012
7	23	11	2012
8	25	3	2012
9	4	4	2012
10	23	7	2012
11	20	2	2012
12	8	10	2012

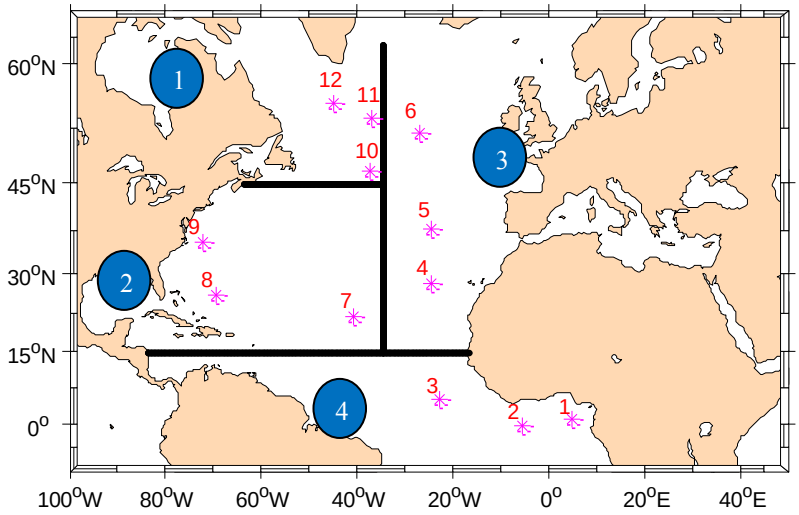


Figure 27. Positions of the selected buoys for the T-S diagrams. The blue circles indicate the region.

The T-S diagrams in Figure 28 show that, for region 4 (bottom), both T-S diagrams are similar. Similarly, in region 2 (second row) the model and Argo agree with the exception of the surface layers. On the other, for regions 1 (top) and 3 (third row) the T-S diagrams for some of the profiles are quite different, especially near the surface. Again, these T-S diagrams provide further evidence that the numerical simulation is less accurate at high latitudes, as the ones with the greatest differences are those of regions 1 and 3, which are the regions with buoys selected at higher latitudes.

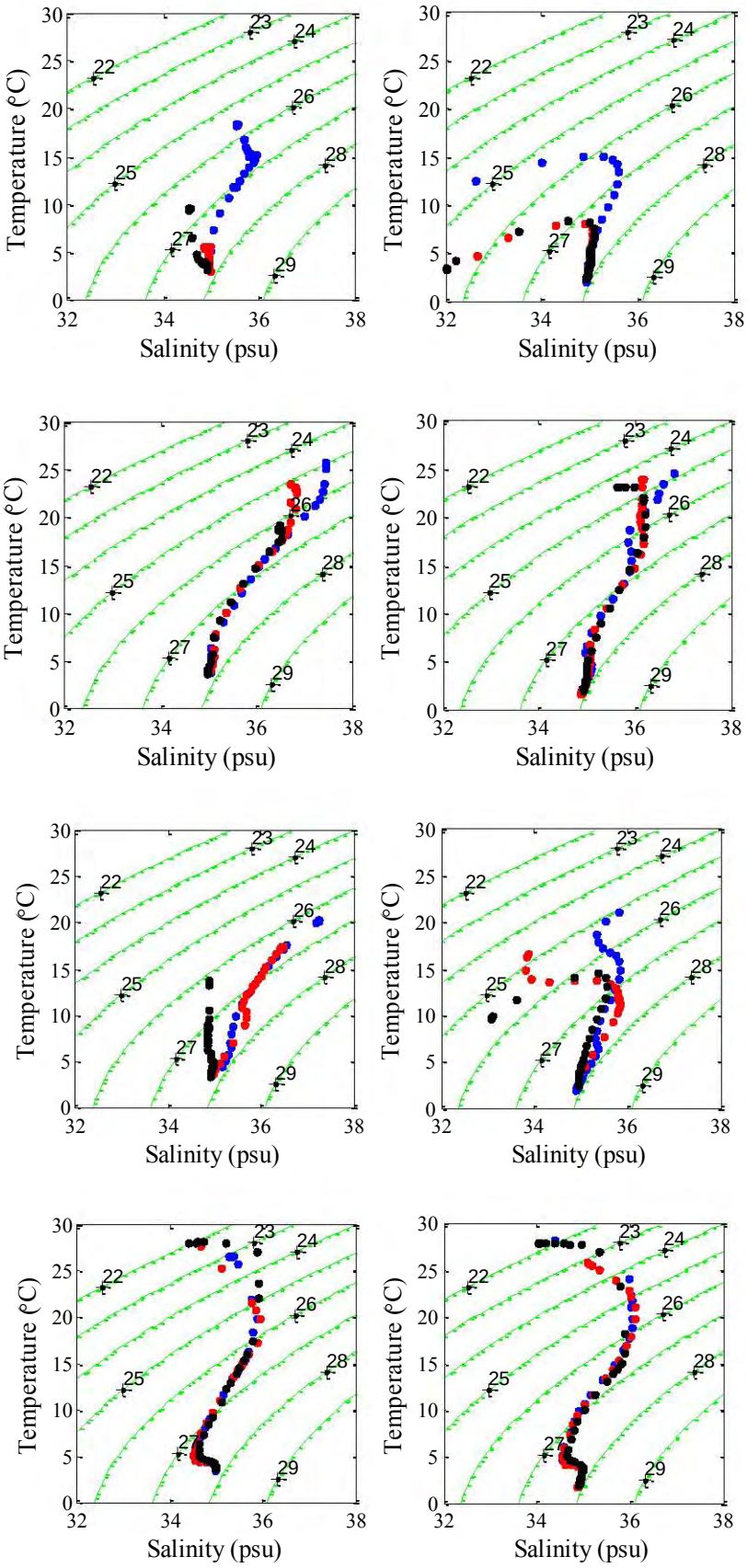


Figure 28. T-S diagrams on the left correspond to the Argo data, and on the right to the simulation's outputs. From bottom to top they correspond to points from region 4 to 1, and following the colour order blue, red, black, they correspond to the points 1 to 12, as labelled on Figure 27. The green, dashed lines are the isopycnals.

4. Conclusions

The analysis of the outputs of the numerical simulation has indicated that north and south regions, close to the open boundaries of the model, together with the peripheral regions are identified as problematic. As described in Section 2.1, the open boundaries of the simulation are relaxed towards climatology data. Therefore, its variability is not reflected accurately. In addition, at high latitudes, an important process is not represented at all by the simulation: the interannual variation of the cold, freshwater input from the Arctic Ocean.

In the peripheral regions, the characteristics of the water masses are very different than the characteristics of the pelagic Ocean that is the main objective of the simulation. The characteristics of the peripheral regions arise from the complex interaction between large river runoff, semi-closed domain, shallow bathymetry and atmospheric forcing. Errors in any of these factors (for example, the river runoff) may strongly affect the realism of the simulation. As there are very few, or none, Argo profile sampling these regions, they are eliminated from our analysis.

The eastern Canadian coast could have also been considered a peripheral region as it clearly follows a different data distribution, than the rest of the system. The simulation at the Gulf of Saint Lawrence has been found excessively fresh by respect to the climatology data and by respect to published data from the region (Petrie *et al*., 2006). The minimum salinity observed in the Atlas at a depth of 30 m during the winter is of 29 psu (Figure 29), whilst the corresponding value of the simulation is 18 psu. This is a very important difference in salinity values, which could not be detected with the Argo data as this is one of the regions which are poorly represented by the Argo network (Figure 17).

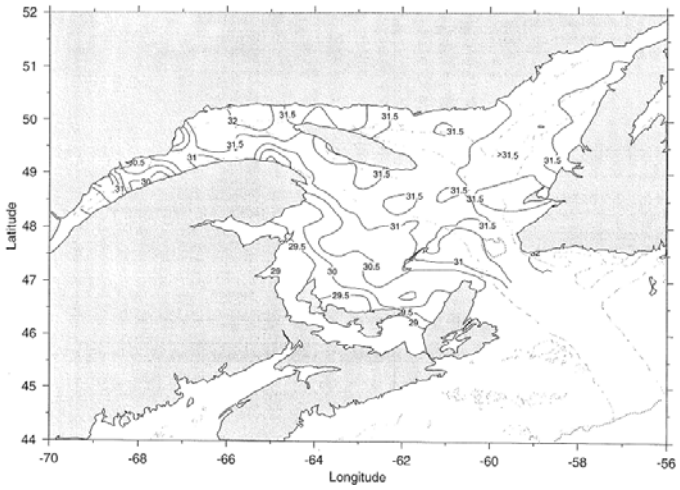


Figure 29. Salinity values at the Gulf of Saint Lawrence, for the 15th of November at 30m (Petrie *et al*., 2010).

The region of the Gulf of Saint Lawrence is a particular case, as freshwater seems to accumulate there. This could be due to the combined effect of the river runoff, the presence of ice during certain times of the year and due to bathymetry errors.

As stated in Buongiorno (2012), temperature and salinity values, especially the surface ones, are sensible to external forces, like freshwater inputs from rivers or precipitation, and to the Ocean's internal dynamics, like convection, advection and mixing processes. Consequently, if these inflows of freshwater and internal processes are not accurately represented by the model, its effect can be clearly observed in the surface layers. As is the case of figure 10 where low values of salinity can be detected at the mouth of great rivers, as the model attributes a salinity of 0 psu to them, which is quite an underestimate.

The anomalous results obtained from the simulation, as already discussed above, could have various reasons and explanations. In Thorpe *et al*., 2005 it is suggested that possible sources of error are low vertical resolution, poor representation of the bathymetry, and issues with the vertical mixing schemes (as described by Buongiorno, 2012), which lead to an unrealistic oceanic circulation, which can cause certain water masses to not be detected in simulations. The importance of the bathymetry used in simulations, is also highlighted in Minvielle *et al*. (2011) and Girton and Sanford (2003). They state that its importance is due to that it affects bottom layer processes which control the deep oceanic circulation.

The T-S diagrams (Figure 28) and the difference between the model and Argo data (Figures 22, 23 and 24), show how the temperature of the model has a positive bias and the salinity has a negative one. The negative salinity tendency is much more marked than the positive one seen for the temperature values. This means that the model is too 'fresh', and could be due to all the reasons exposed above. They also show that the simulation is more inaccurate near the surface, but also it should be kept in mind that here is where the oceanic parameters show a greater variability.

The shallowest depth at which the simulation could be validated was 9 m, due to the lack of valid Argo data near the surface, and also there were regions like the North Atlantic Subtropical gyre and the Gulf of St. Lawrence between others were there was a lack of data too. In the latter region, another source of data had to be used to validate the outputs of the simulation there. This lack of valid Argo data at the surface was expected, as in many occasions when the buoys are near the surface the hydraulic pump is stopped and so the salinity sensor stops measuring.

More Argo data at the surface data could have been obtained if not only the profiles flagged with a QC = 1 have been considered, but also the ones flagged with a QC of 2, 5 or 8 (Table 2). Due to the limited amount of time for this study, only the data with QC=1 were selected. However, it has been noticed that the Argo observing system is able to represent seasonal variability of the Ocean (Figure 21). There is a considerable amount of valid profiles (3117) for year 2012, with the exception of some points and buoys the data they measure is very reasonable (Figure 20) and there are no temporal biases as they provide data throughout the year. This demonstrates that choosing Argo to validate the simulation was a correct choice.

To conclude, in this work I have contributed to the first efforts of the Department of Physical and Technological Oceanography of the Institut de Ciències del Mar (CSIC) to the validation of their North Atlantic Ocean simulation. This simulation couples two of the state-of-art oceanic models: NEMO (oceanic dynamics) and LIM-2 (sea-ice dynamics). By comparing the outputs of the model against the Argo hydrographic profiles, I have found that the pelagic circulation of the model has some reasonable similarities with the observations. Although important deviations between the model and data exist near the meridional borders and off the Canadian Coast, our results indicate that the pelagic circulation of the model has some reasonable similarities

with observations. This is important because the ability of retrieving SSS from remote sensing is extremely low in cold waters and near the coast (due to land sea abrupt emissivity transition). That is, the model provides the largest compatibility with data right in the region where it would be more useful to provide a coherent map of salinity in the North Atlantic Ocean.

5. References

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Garcia, H. E., Baranova, O. K., Zweng, M. M. and Johnson, D. R. (2010): World Ocean Atlas 2009 Volume 2: Salinity. S. Levitus Ed. NOAA Atlas NESDIS 69, U.S. Gov. Printing Office, Washington, D.C., 184 pp.

Brodeau, L., Barnier, B., Treguier, A. M., Penduff, T. and Gulev, S. (2010): An ERA40-based atmospheric forcing for global ocean circulation models. Ocean Modelling 31, 88–104.

Brown, E., Colling, A., Park, D., Phillips, J. and Rothery, D. (1988): Ocean Circulation. The Open University, 287 pp.

Buongiorno Nardelli, B. (2012): A Novel Approach for the High-Resolution Interpolation of In Situ Sea Surface Salinity. J. Atmos. Oceanic Technol., 29, 867–879.

Carval, T., Keeley, B., Takatsuki, Y., Yoshida, T., Loch, S., Schmid, C., Goldsmith, R., Wong, A., McCreadie, R., Thresher, A. and Tran, A. (2012): ARGO USER'S MANUALVersion 2.4, 85 pp.

Chatterjee, A., Shankar1, D., Shenoi, S. S. C., Reddy, G. V., Michael, G. S., Ravichandran, M., Gopalkrishna, V. V., Rama Rao, E. P., Udaya Bhaskar, T. V. S., and Sanjeevan, V. N. (2012): A new atlas of temperature and salinity for the North Indian Ocean. J. Earth Syst. Sci.121, No. 3, 559–593.

Chen, C. (2009): Ocean Water Masses-T-S Diagrams and Upper Ocean Waters. General Physical Oceanography, MAR 555, School for Marine Sciences and Technology, Umass-Dartmouth, 18 pp.

Emery, W. J., and Thomson, R. E. (2004): Data analysis methods in physical oceanography 2nd edition. Elsevier Ed., 654 pp.

Girton, J.B. and Sanford, T. B. (2003): Descent and modification of the overflow plume in the Denmark Strait. J. Phys. Oceanogr., 33, 1351–1364.

Locarnini, R. A., Mishonov, A. V., Antonov, J.I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M. and Johnson, D. R. (2010): World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed.,NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.

Higgison, S., Thompson, K. R. and Liu, Y. (2009): Estimating ocean climatologies for short periods: A simple technique for removing the effect of eddies from temperature and salinity profiles. GEOPHYSICAL RESEARCH LETTERS, 36, 4 pp.

Hoareau, N., Portabella, M., Garcia-Ladona, E., Turiel, A. and Ballabrera-Poy, J. Meridional variability of North-Atlantic Sea Surface Salinity: Wavenumber spectra derived from climatology, ocean model and satellite observations. Submitted to J. Geophys. Res.

MetOffice, United Kingdom:

http://www.metoffice.gov.uk/weather/ marine/observations/gathering_data/argo.html

Minvielle, M., Cassou, C., Bourdallé-Badie, R., Terray, L. and Najac, J. (2011): A statistical–dynamical scheme for reconstructing ocean forcing in the Atlantic. Part II: methodology, validation and application to high-resolution ocean models. J. Climate Dynamics, 36, 401-417.

National Weather Service, NOAA: www.nws.noaa.gov/climate/help/glossary.php

Petrie, B., Drinkwater, K., Sandström, A., Pettipas, R., Gregory, D., Gilbert, D. and Sekhon, P. (1996): Temperature, salinity and sigma-t Atlas for the Gulf of St. Lawrence. Canadian Technical Report of Hydrography and Ocean Sciences 178, 263 pp.

Thorpe, S. E., Stevens, D. P. and Heywood, K. J. (2005): Comparison of two time-variant forced eddy-permitting global ocean circulation models with hydrography of the Scotia Sea. Ocean modeling, 9, 105-132.

University of California, San Diego: http://www.argo.ucsd.edu/

6. Appendices

6.1. Appendix 1

```
SUBROUTINE sw_ptmp(n,S,T,P,PR,TP)
IMPLICIT NONE
INTEGER, INTENT(in)                  :: n
real(kind=8), dimension(n), intent(in)   :: T,P
real(kind=8),               intent(in)   :: PR
real(kind=8), dimension(n), intent(in)  :: S
real(kind=8), dimension(n), intent(out) :: TP
! ... Local variables
real(kind=8), dimension(n)
  :: ADGT,del_P,del_th,q,th,PP
! ... theta1
  del_P = PR - P
  CALL sw_adtg(n,S,T,P,ADGT)
  del_th = del_P * ADGT
  th = T + 0.5D0 * del_th
  q = del_th
! ... theta2
  PP = P+0.5D0*del_P
  CALL sw_adtg(n,S,th,PP,ADGT)
  del_th = del_P * ADGT
  th = th + (1.0D0 -1.0D0/sqrt(2.0D0))*(del_th -q)
  q = (2.0D0-sqrt(2.0D0)) * del_th +
(-2.0D0+3.0D0/sqrt(2.0D0))* q
! ... theta3
  PP = P+0.5D0*del_P
  CALL sw_adtg(n,S,th,PP,ADGT)
  del_th = del_P * ADGT
  th = th + (1.0D0 +1.0D0/sqrt(2.0D0))*(del_th -q)
  q = (2.0D0+sqrt(2.0D0)) * del_th
 + (-2.0D0-3.0D0/sqrt(2.0D0))* q
! ... theta4
  PP = P+del_P
  CALL sw_adtg(n,S,th,PP,ADGT)
  del_th = del_P * ADGT
  TP = th + (del_th -2.0D0*q)/6.0D0
END SUBROUTINE sw_ptmp
!%%%%%%%%%%%%%%%%%%%
! Adiabatic temperature gradient
SUBROUTINE sw_adtg(n,S,T,P,ADGT)
IMPLICIT NONE
integer, intent(in)                  :: n
real(kind=8), dimension(n), intent(in)   :: S,T,P
real(kind=8), dimension(n), intent(out) :: ADGT
! ... Parameters:
real(kind=8), parameter  :: a0 = 3.5803D-5
real(kind=8), parameter  :: a1 = 8.5258D-6
real(kind=8), parameter  :: a2 = -6.836D-8
real(kind=8), parameter  :: a3 = 6.6228D-10
real(kind=8), parameter  :: b0 = 1.8932D-6
real(kind=8), parameter  :: b1 = -4.2393D-8
real(kind=8), parameter  :: c0 = 1.8741D-8
real(kind=8), parameter  :: c1 = -6.7795D-10
real(kind=8), parameter  :: c2 = 8.733D-12
real(kind=8), parameter  :: c3 = -5.4481D-14
real(kind=8), parameter  :: d0 = -1.1351D-10
real(kind=8), parameter  :: d1 = 2.7759D-12
real(kind=8), parameter  :: e0 = -4.6206D-13
real(kind=8), parameter  :: e1 = 1.8676D-14
```

```
real(kind=8), parameter  :: e2 = -2.1687D-16

 ADGT = a0 + (a1 + (a2 + a3 * T)*T)*T + (b0 + b1
*T)

* (S -35.0D0) + ((c0 + (c1 + (c2 + c3 *T) *T)*T) +
(d0 + d1 *T)*(S-35.0D0))*P + ( e0 + (e1 + e2 *T)*T)

*P *P

END SUBROUTINE
```

6.2. Appendix 2

```
SUBROUTINE percentiles (n,x,np,p,xp)

IMPLICIT NONE

integer, intent(in)                  :: n,np

real(kind=8), dimension(np), intent(in)  :: p

real(kind=8), dimension(n), intent(in)   :: x

real(kind=8), dimension(np), intent(out) :: xp

integer i,kk

real(kind=8) rr,dd

integer indx(n)

CALL indexx(n,x,indx)

DO i=1,np

  rr = p(i)*(n-1.0D0)/100.0D0 + 1.0D0

  kk = FLOOR(rr)

  dd = rr - kk

  IF (kk.EQ.0) THEN

    xp(i) = x(indx(1))

  ELSE IF (kk.EQ.n) THEN

    xp(i) = x(indx(n))

  ELSE

    xp(i) = x(indx(kk)) + dd*(x(indx(kk+1))-x(indx(kk)))

  ENDIF

ENDDO

RETURN

END
```

6.3. Appendix 3

```
SUBROUTINE avevar (data,n,ave,var)

IMPLICIT NONE

integer, intent(in)                  :: n

real(kind=8), dimension(n), intent(in)   :: data

real(kind=8), intent(out)              :: ave,var

! ... Local variables:

integer j

real(kind=8) s,ep

ave = SUM(data(1:n))/n

var = 0.0d0

ep  = 0.0d0

do j=1,n

  s   = data(j) - ave

  ep  = ep + s

  var = var + s*s

ENDDO

var = (var-ep**2/n)/(n-1)

RETURN

END

!  (C) Copr. 1986-92 Numerical Recipes Software *5sV1.

! ...

! ==================================

! ...

SUBROUTINE avevar4 (data,n,ave,var)

IMPLICIT NONE

integer, intent(in)                  :: n

real(kind=4), dimension(n), intent(in)   :: data

real(kind=4), intent(out)              :: ave,var

! ... Local variables:

integer j

real(kind=8) s,ep,v8

ave = SUM(data(1:n))/n

v8 = 0.0d0

ep = 0.0d0

do j=1,n

  s  = DBLE(data(j) - ave)

  ep = ep + s

  v8 = v8 + s*s

ENDDO

v8 = (v8-ep**2/n)/(n-1)

var = SNGL(v8)

RETURN

END

!  (C) Copr. 1986-92 Numerical Recipes Software *5sV1.
```

7. Additional information about the development of this study

7.1. Detailed description of the activities carried.

This study was divided in 3 parts:

7.1.1. Statistical characterization of the simulation's outputs.

The first part consisted on analyzing 4 variables of the numerical simulation. The variables I analyzed were potential temperature, salinity, sea surface height and mixed layer depth, hereafter referred to as Tm, Sm, SSH and MLD respectively. To obtain and carry out a preliminary analysis of these values, specific programmes were created to read the NetCDF files.

For this first analysis I only processed the data of one output file of the simulation, which in this case, corresponded to data between the 22/01/2001 and the 26/01/2001. With this set of data, I did two types of figures: histograms and contour maps. For Tm and Sm, I created a figure for each layer of the simulation. In addition, I calculated a set of statistical parameters throughout programs I created in Fortran.

In this preliminary analysis I could detect that the land values for Tm were 0, while for Sm they were 0.1. I had to take this into account when processing the data in my Fortran programmes. In a second version of the programmes, I also modified the merdional limits to 60ºN and 5ºS.

I obtained again the histograms, contour maps and statistical parameters of the 4 variables and compared them. In addition, this time I plotted the variation of 4 of the statistical parameters of Tm and Sm with depth in Matlab, and compared them. In the same figure two subplots were made: depth against the mean and the median, and the depth against the interquartile range and the standard deviation.

However, anomalous results and bimodal distributions were still present in the histograms. Using the contour maps of the minimum and maximum values of the parameters, I could identify the regions which modified the mean values of these

parameters, and I established certain peripheral regions, which from this moment on I treated separately.

Lastly, I used Matlab to make a T-S diagram of all the regions, for each peripheral region and for the North Atlantic Ocean without the peripheral regions. I analysed these diagrams and I could detect other regions of anomalous Tm and Sm values, and take these regions into account for the validation results obtained later.

After this part of the study, I analysed the output files of the simulation which corresponded to January 2012 (a total of 7 NetCDFs). The procedure followed was the same as the one previously carried out with the exception that the statistical parameters were calculated in space and time.

Once I had successfully created the necessary programmes in Fortran and Matlab for the analysis of the simulation's outputs, I analysed all the outputs of the year 2012. As a greater quantity of results were now obtained, instead of saving the new obtained data in a data file (as I had been doing previously), I saved it in the same format as the original files (in a NetCDF). In addition to being able to save a greater amount of data, saving the new files in NetCDF format allowed me to use the application ncview, from which I could quickly visualize the data obtained.

Having obtained all the NetCDF files and exported them to Matlab, I created the figures of the statistical parameters in time, in space and their variation with depth. Due to anomalous results found in the statistical parameters in time, the whole of the procedure previously described was repeated for data corresponding to December 2011, to be able to compare the results. Anomalous results were found too in the statistical parameters in space, and to compare these results, this time the climatology data of WOA09 was downloaded as a NetCDF file.

7.1.2. Recovery and analysis of the data measured by Argo floats.

In this next part of the study, I created programmes in Fortran to identify the Argo profiles which corresponded to my region of study. To do so I used an application developed at BEC called locateArgo, which after specifying the desired characteristics of the Argo profile; it returns the identity of the profile and its location in BEC's database. I carried out a second Argo profile selection creating another program in

Fortran in which I saved the data of the desired Argo profiles in a text file. After doing so I followed steps a) to n) as described in section 3.3.

The final potential temperature (Ta) and salinity (Sa) results of the Argo data were saved by profile and depth in a NetCDF file. In this file, I also saved the latitude, longitude, date, profile ID in BEC's database, ipos, jpos, tpos and the valid measurements at the different depths for each profile.

I exported this file to Matlab and firstly, I plotted a map of the positions of the selected Argo profiles. Thanks to this figure I could identify and correct an error in one of my Fortran programmes. Then I created graphs of temperature-depth and salinity-depth of all the profiles to visually and quickly analyze, and try to detect any anomalous profiles. I also obtained scatter maps of Ta and Sa for every depth. Lastly, I created a program in Fortran to obtain the time series of the statistical parameters of the Argo data.

7.1.3. Global comparison of the numerical model and Argo data.

Like in the previous parts of the study, I began creating a program in Fortran to obtain a NetCDF of the data. This NetCDF included the difference between Ta or Sa (Argo) and the corresponding value of Tm or Sm (model). The results were represented in scatter maps, and plots of the differences against time and against the buoy number. From this last plot I considered it was necessary to find the points with a model minus Argo difference larger than 10ºC, and salinity differences larger than 6 psu. To do this a used both Fortran and Matlab.

Once these points were identified, I plotted their temperature and salinity profiles using the simulation data, the original Argo data and the interpolated Argo data, so they I could try to discover is the differences were due to one source of data or another.

Finally, I decided to create new T-S diagrams of particular points of the North Atlantic Ocean to be able to better distinguish the different water masses. I randomly selected 12 points of an Argo data salinity map, and using Matlab found the positions of these points. I then exported this information to Fortran where I created a program to obtain the corresponding simulation and Argo data. Afterwards I went back to Matlab

to plot the T-S diagrams as well as the vertical temperature and salinity profiles for these points, for both sources of data.

7.2. Training received

During this study I have acquired both theoretical and practical knowledge. As theoretical knowledge I have learned about mathematical models used in oceanography, the dynamics of the Atlantic Ocean and the functioning of Argo buoys.

As practical knowledge, I have learnt more about programming, not only with Matlab, but with Fortran 90 too, and I have become familiar with the use of ncview. This has made me be quicker and feel more comfortable whilst using this software. I have also learnt to work with big data sets of different sources, and to organize them in an efficient way. Lastly, I have also become familiar with the Linux operating system.

Every one or two weeks seminars are organized at the Institute about the different investigations taking place there, which also helped me learn about different aspects of marine sciences in general and about physical oceanography.

7.3. Level of integration and involvement in the department and relationship with personnel.

To be able to better understand the results and errors obtained, I sometimes needed a detailed explanation of some general aspects about mathematical models and of the particular model I was working with. Both my placement tutor as well as his PhD student Nina Hoareau found the patience and time to explain to me these concepts. In addition, with any technical problem I encountered (as I sometimes had to us the Institute's internal system, Gaia), there was always someone to kindly help me, and if they could not, they would tell me who could help me. The scientist with whom I shared the office where I worked was very kind and helped me both with any theoretical doubt, or related with the Institute. He also informed me of when the next seminar was going to take place.

7.4. Positive and negative aspects related to the development of this study.

The positive aspects of this study are that it has provided me with a greater vision about the different applications of physical oceanography, new knowledge about mathematical models and their use, an understanding of climatology and knowledge about the different systems used to measure salinity (remote sensing, mathematical models and Argo buoys). At a personal level, it has helped me gain confidence when deciding how to best solve a problem that arises during an investigation.

On the other hand, I have found the amount of time available for this study a bit limiting due to the great amount of data I had to process and analyze. I would have liked to analyze further some of the results I obtained. Also, programming took me more time than I expected, and not as much the theoretical part behind this study.

7.5. Personal appraisal of the learning achieved during this study.

Thanks to carrying out this study, I have been able to learn new programming languages, validation procedures and how to process and analyze big data sets from different sources. It has also helped me to consolidate the knowledge and techniques learnt during my degree.

I think that the most important thing I have achieved in addition to new knowledge is confidence when deciding how to best solve a problem. In other words, autonomy whilst working, as I have been able to realize that I can put into practice all that I have learnt during the past years.