Progress in Oceanography xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Progress in Oceanography

journal homepage: www.elsevier.com/locate/pocean

Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean

Guillaume Maze [a,*], Herlé Mercier [b], Ronan Fablet [c], Pierre Tandeo [c], Manuel Lopez Radcenco [c], Philippe Lenca [c], Charlène Feucher [a], Clément Le Goff [a]

[a] Ifremer, Univ. Brest, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, 29280, Plouzané, France
[b] Univ. Brest, CNRS, IRD, Ifremer, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, 29280, Plouzané, France
[c] Institut Mines-Telecom, Télécom Bretagne, UMR CNRS 6285 Lab-STICC, Brest, France

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form 9 December 2016
Accepted 30 December 2016
Available online xxxx

Keywords:
Heat content
Classification
North Atlantic
Stratification
Water mass
Thermocline
Argo
Pattern

ABSTRACT

A quantitative understanding of the integrated ocean heat content depends on our ability to determine how heat is distributed in the ocean and identify the associated coherent patterns. This study demonstrates how this can be achieved using unsupervised classification of Argo temperature profiles. The classification method used is a Gaussian Mixture Model (GMM) that decomposes the Probability Density Function of a dataset into a weighted sum of Gaussian modes.

It is determined that the North Atlantic Argo dataset of temperature profiles contains 8 groups of vertically coherent heat patterns, or classes. Each of the temperature profile classes reveals unique and physically coherent heat distributions along the vertical axis. A key result of this study is that, when mapped in space, each of the 8 classes is found to define an oceanic region, even if no spatial information was used in the model determination. The classification result is independent of the location and time of the ARGO profiles.

Two classes show cold anomalies throughout the water column with amplitude decreasing with depth. They are found to be localized in the subpolar gyre and along the poleward flank of the Gulf Stream and North Atlantic Current (NAC). One class has nearly zero anomalies and a large spread throughout the water column. It is found mostly along the NAC. One class has warm anomalies near the surface (50 m) and cold ones below 200 m. It is found in the tropical/equatorial region. The remaining four classes have warm anomalies throughout the water column, one without depth dependance (in the southeastern part of the subtropical gyre), the other three with clear maximums at different depths (100 m, 400 m and 1000 m). These are found along the southern flank of the North Equatorial Current, the western part of the subtropical gyre and over the West European Basin. These results are robust to both the seasonal variability and to method parameters such as the size of the analyzed domain.

1. Introduction

As revealed by in situ and satellite observations, the ocean has undergone significant changes in recent decades. In particular, since the early 1970s, the ocean has stored 93% of the excess of heat added to the Earth's climatic system by the anthropogenically modified radiative balance at the top of the atmosphere (Stocker et al., 2013). The ocean has also been found to be more stratified (Levitus et al., 2012) and Western Boundary Currents have probably shifted poleward and intensified (Wu et al., 2012; Yang et al., 2016). To understand the drivers of these changes requires a quantitative understanding of the integrated ocean heat content.

The ocean temperature structure is very complex but a simple first-order description is possible. Near the surface, ocean temperature is primarily driven by air-sea heat fluxes and modulated by horizontal heat mean and eddy transports, especially in Western Boundary Current systems (Kwon et al., 2010). This causes the ocean surface temperature to decrease poleward. However, ocean currents are three-dimensional and redistribute heat at different depths. At mid-latitudes, a negative wind-stress curl forces a downward doming of isopycnal and isothermal surfaces (Vallis, 2006), which results in the temperature at depth, e.g. 500 m, to be higher in subtropical gyres than at the equator (Talley et al., 2011). This 3-dimensional redistribution of heat in the ocean makes our ability to identify remarkable heat patterns in the horizontal and vertical plans crucial to the understanding of the integrated ocean heat content.

* Corresponding author.
E-mail address: gmaze@ifremer.fr (G. Maze).

http://dx.doi.org/10.1016/j.pocean.2016.12.008
0079-6611/

Please cite this article in press as: Maze, G., et al. Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean. Prog. Oceanogr. (2017), http://dx.doi.org/10.1016/j.pocean.2016.12.008

Along the vertical axis, remarkable patterns may be defined a priori using known water masses such as shallow, intermediate and deep waters or layers such as the mixed and Ekman layers and the permanent thermocline. These patterns can be used to partition horizontal or vertical heat transports (Talley, 2003; McCarthy et al., 2012; Buckley et al., 2014). However, they are not used to partition heat content variability despite efforts to formalize the use of reference surfaces in vertically integrating heat content (Palmer and Haines, 2009). The problematic is that the general lack of clear objective definition for vertical patterns, despite a recent effort with regard to the permanent pycnocline (Feucher et al., 2016), impedes their description, especially over long timescales during which their defining properties can change (e.g. Yang and Wang, 2009; Fiedler, 2010).

In the horizontal plan, remarkable large scale patterns are not defined per se. Simple geographical boxes of fixed size and shape are preferred. One is left with the difficult task of looking for relevant boxes to explain the large scale structure and variability of the heat content. Many studies define the subtropical and/or subpolar gyres as rectangular boxes from which box-averaged statistics are computed (e.g. Lozier et al., 2010; Bryden et al., 2014; Häkkinen et al., 2015; Grist et al., 2015). Due to limited availability of historical measurements, one can even find signals of entire regions to be approximated by a single location dataset (e.g. Curry and McCartney, 2001). Obviously, a serial issue with a rectangular box is that it does not take into account the complex structure of the ocean, which is not aligned along latitudes and longitudes. The problematic is that, although it is always possible to use more complex polygons than a rectangle to describe a region (e.g. Barrier et al., 2015), this will be difficult, if not possible , if a region has to be bounded by a dynamical structure such as a Western Boundary Current.

To identify remarkable heat patterns in the horizontal and vertical plans, their variability and climatology thus remains a challenge. In this study, we propose to tackle this problem with a method that belongs to the class of unsupervised classification methods. Classification, or clustering, is a statistical method that groups data into classes, or clusters, according to a given similarity metric.

Profile classification has already been used in oceanographic applications but for other purposes. Hjelmervik and Hjelmervik (2013) used a classification method on in situ profiles to predict the local vertical structure of temperature and salinity at a given location, without surface information. Indeed, it is rather common to predict the interior structure of the ocean based on surface data, such as sea surface height, and a model either based on physical principals (Ponte and Klein, 2013) or on historical local regressions (Guinehut et al., 2012). To do so without a surface information is much more complicated though. Hjelmervik and Hjelmervik (2013) grouped profiles according to their Euclidean distance in a reduced dimensional space for latitude/longitude/temperature/sali nity and derived a prediction model of climatological profiles at a given latitude/longitude location. For the method to perform better than a classic box averaging method, they determined that 26 groups of profiles were necessary for the North Atlantic Ocean. They later adapted the method to real-time profile prediction using partial observations, and decreased the number of groups to 18 for the method to perform well (Hjelmervik and Hjelmervik, 2014). A classification based prediction method to fill in gaps in two-dimensional data has also been used for satellite measurements with clear success (Aretxabaleta and Smith, 2013). The work by Hjelmervik and Hjelmervik (2013, 2014) extents this idea to vertical profiles with strong promises.

Classification based prediction methods strive in dealing with non-Gaussian statistics, such as observed in frontal regions (Sura, 2010). However, for our goal, which is to characterize remarkable heat patterns, they suffer from two limitations: (i) they take data latitude and longitude as parameters and (ii) they require a rather large number of classes to perform well. We understand that these requirements are imposed to ensure a satisfactory prediction performance, but here, we are interested in identifying remarkable patterns in vertical temperature profiles, their corresponding regional distributions (if any) and their climatology. Therefore, on the one hand there is no reason to impose data coordinates in the classification. Indeed, we should let the classification reveal the spatial coherence, or lack thereof, of temperature profiles rather than impose it. Tandeo et al. (2014) demonstrated how unsupervised mixture modeling can be used to classify sea surface temperature and height anomalies into small scale dynamical modes without using the latitude and longitude of the data. We shall demonstrate in this study that profile based classes are indeed coherent in space in the North Atlantic. On the other hand, we aim to understand, and therefore reduce, the information contained into a large collection of temperature profiles. Thus, it is crucial for the information to be contained into a limited number of classes. This however will depend on how many *remarkable patterns* are relevant for a given usage.

The paper is organized as follows: in Section 2, the dataset is presented; in Section 3, the classification method is introduced as well as the method we employed to apply it to the Argo temperature dataset; in Section 4, we apply classification to analyze the vertical integral of the heat content and in Section 5, the classification of temperature profiles is performed to reveal the ocean internal heat content structure in the North Atlantic. Last, discussion/conclusion are drawn in Section 6 while the appendices provide technical details about optimization (Appendix A) and sensitivity experiments (Appendix B).

2. Data

In this study, we used data from the Argo array. Argo is a real-time global ocean observation network. It consists of about 3000 autonomous profilers randomly distributed in all oceans to observe the large scale open ocean out of the high latitudes and marginal seas. Most of the profilers drift freely at a parking depth around 1000 m and every 10 days descend down to 2000 m to then rise up to the surface, measuring pressure, temperature and salinity. Once at the surface, profile data are transmitted to data assembly centers by satellite, after which profilers descend back to their parking depth and start another 10 day cycle. Argo data are now used routinely in physical oceanography and are key to the observation of the ocean climate (Riser et al., 2016).

The Argo database is a collection of more than 1.5 million temperature and salinity profiles going from the surface to 2000 m, evenly distributed throughout the seasonal cycle and with approximately 1 profile per month per $3^\circ \times 3^\circ$ cell between 2000 and 2014. We extracted the database in December 2014 from the Coriolis GDAC (Argo, 2014). We selected profiles located in the North Atlantic between the equator and 70N and between 90W and 0E. The collection was reduced to profiles and measurements with correct quality control flags (1, 2, 5 or 8, following the Argo reference Table 2 of the user manual, Carval et al., 2015) between the surface and 1400 m. The depth level of 1400 m was chosen as a compromise between the total number of profiles (the shallower the larger) in the analyzed dataset and the vertical extent of the analysis. We finally interpolated the data on a regular vertical grid with a 5 m resolution (the original resolution ranges from less than 10 m at the surface to 200 m at the bottom of the profile).

Fig. 1A shows the spatial density of the final collection of 100,684 profiles. The North Atlantic is a well observed basin and the spatial density is such that there are around 30 profiles per

Please cite this article in press as: Maze, G., et al. Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean. Prog. Oceanogr. (2017), http://dx.doi.org/10.1016/j.pocean.2016.12.008