APRIL 2017

#55

SPECIAL ISSUE WITH

Coriolis

MERCATOR OCEAN JOURNAL

SPECIAL ISSUE JOINTLY COORDINATED
BY MERCATOR OCEAN AND CORIOLIS FOCUSING ON THE LATEST OUTCOMES
FROM THE IN SITU OCEAN OBSERVING NETWORK

#55

EDITORIAL

Greetings all,

Once a year, the Mercator Ocean Forecasting Center in Toulouse and the Coriolis Infrastructure in Brest publish a common journal with articles dedicated to observations only.

Articles are dedicated to the following thematics:

Cancouët et al. present the Euro-Argo MOCCA project (Monitoring the Ocean and Climate Change with Argo) scheduled for a 5-year period (2015-2020). Its goal is to procure and deploy 150 European Argo floats, along with corresponding data collection, analysis, management, processing and dissemination in the years up to 2020 over the operating lifetime of the floats.These floats will complement the Euro-Argo countries contribution to the international network.

Pouliquen et al. next present the AtlantOS (Atlantic Ocean Observing System) project scheduled for a 4 years period which objective is to obtain an international, more sustainable, efficient, integrated, and fit-for-purpose observing system in the Atlantic Ocean and in particular the activities related to data management and interoperability.

Next paper by **Thierry et al.** presents the work done to share dissolved oxygen data processing methods and converge towards common methods. Dissolved oxygen concentration is an important parameter in the context of climate change and the monitoring of marine ecosystems.

Quentin et al. then describe the progress in France towards the development of high frequency ocean radars in order to observe coastal areas. High Frequency Radars are shore-based remote-sensing instruments to accurately monitor the ocean surface current field and transport, over large areas (typically 60 X 60 km), in real time and at high spatial (3-5 km) and temporal resolutions (10 min-1 hour).

The following paper by **Gourrion et al.** presents a new statistical approach for in situ temperature and salinity observation quality control developed within the In Situ Thematic Assembly Center of the Copernicus Marine and Environmental Monitoring Service. This new approach based on direct inference of local extreme values improves product quality assessment and reduce subjective decisions by human operators..

Maze et al. present a data mining statistical method called «Profile Classification Model». When applied to in situ data, it allows sorting out all in situ profiles into a small number of classes, each capturing the diversity of all possible vertical structures. Such method is for example useful for observation data centers to better select reference data for quality control procedures. It is also useful for scientists to better comprehend ocean processes and assess the realism of a numerical model output. **V. Thierry et al.** and **L.Coppola et al.** tell us about the implementation of an Argo-O2 array in respectively the North-Atlantic Ocean and the Mediterranean Sea.

We wish you a pleasant reading,

Laurence CROSNIER and Sylvie POULIQUEN,
Editors

#55

TABLE OF CONTENTS

#55

TABLE OF CONTENTS

THE EURO-ARGO MOCCA PROJECT

BY
R. CANCOUËT[1], G. OBOLENSKY[1], S. POULIQUEN[1] AND EURO-ARGO ERIC PARTNERS[2]

ABSTRACT

Officially set up in 2014, the Euro-Argo ERIC is a European legal entity whose main objective is to organize a long-term European contribution to the international Argo array of oceanographic profiling floats. In 2015 the ERIC submitted a proposal to DG-MARE through EASME that was accepted. In the framework of the action "Monitoring the Oceans" of EMFF Work Programme 2015, the MOCCA project (Monitoring the Ocean and Climate Change with Argo) started in June 2015 and is scheduled for a 5-year period. With a EU contribution of 4M€, the ERIC with its members added an additional 20% that generated a total of 5M€, allowing three actions: procurement of 150 T/S Argo floats during 2015-2016, arrangement for their deployment in 2016-2017, and data processing in real-time and delayed-mode during the period 2016-2020. Within MOCCA, the Euro-Argo ERIC is progressing towards the establishment of a sustainable European contribution to the Argo programme and is demonstrating its operational capabilities.

[1] **Euro-Argo ERIC**, Plouzané, France.

[2] http://www.euro-argo.eu/About-us/The-Partners

CONTEXT

The international Argo[1] programme was initiated in 1999 as a pilot project endorsed by the Climate Research Programme of the World Meteorological Organisation, GOOS, and the Intergovernmental Oceanographic Commission. The Argo network is a global array of autonomous instruments, deployed over the world ocean, reporting subsurface ocean properties to a wide range of users via satellite transmission links to data centres.

Euro-Argo[2] is a European Research Infrastructure Consortium (ERIC) that involves 25 organisations from 12 countries. All partners wish to optimise their collective contribution to relevant EU policies, programmes and projects. The overall objectives of Euro-Argo in monitoring the oceans are:

- to maintain an array of around 800 floats at any time (a European contribution of ¼ of the global array),
- to supply enhanced coverage in European regional seas,
- to provide quality controlled data and access to the data sets and data products to the research (ocean and climate) and operational oceanography (e.g. Copernicus Marine Service) communities,
- to prepare and contribute to the extensions of Argo (e.g. marginal seas, biogeochemistry, deep ocean, polar regions).

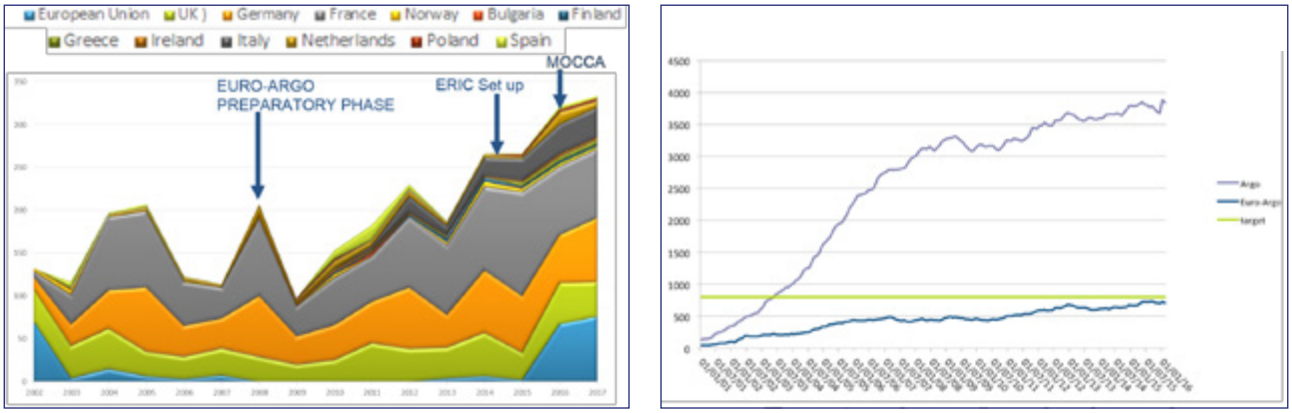


FIGURE 1

Evolution of the number of floats deployed per European countries in the past 15 years (left) and Euro-Argo distinct floats distributing data at GDACs vs Argo (monthly values, right).

Even though the number of floats deployed per European Countries increased in the past 15 years (see Figure 1), the European Contribution to Argo, based on national funds only, has reached a plateau. Since the start, Euro-Argo partners have been discussing closely with the European Commission (DG-MARE, DG-RESEARCH and DG-GROWTH) on supporting the Euro-Argo Infrastructure to develop a durable European contribution to Argo.

In April 2015, a proposal has been submitted to DG-MARE through EASME[3] and definitely signed end of June 2015. In the framework of the action "Monitoring the Oceans" of the EMFF (European Maritime and Fisheries Fund) Work Programme 2015, the MOCCA[4] project (Monitoring the Ocean Climate Change with Argo) started in June 2015 and is scheduled for a 5-year period (2015-2020).

The goal of the MOCCA project is to progress towards the Euro-Argo objectives in monitoring the oceans and increase the European effort under the international Argo programme by procuring and deploying 150 Argo floats, as well as the data collection, analysis, management, processing and dissemination in the years up to 2020 over the operating lifetime of the floats.

MOCCA PROJECT ORGANISATION

The MOCCA project is structured along 4 workpackages (see Figure 2) and will last 60 months. The Euro-Argo ERIC organizes the splitting of activities between the ERIC and its members, and reports to EASME (WP1).

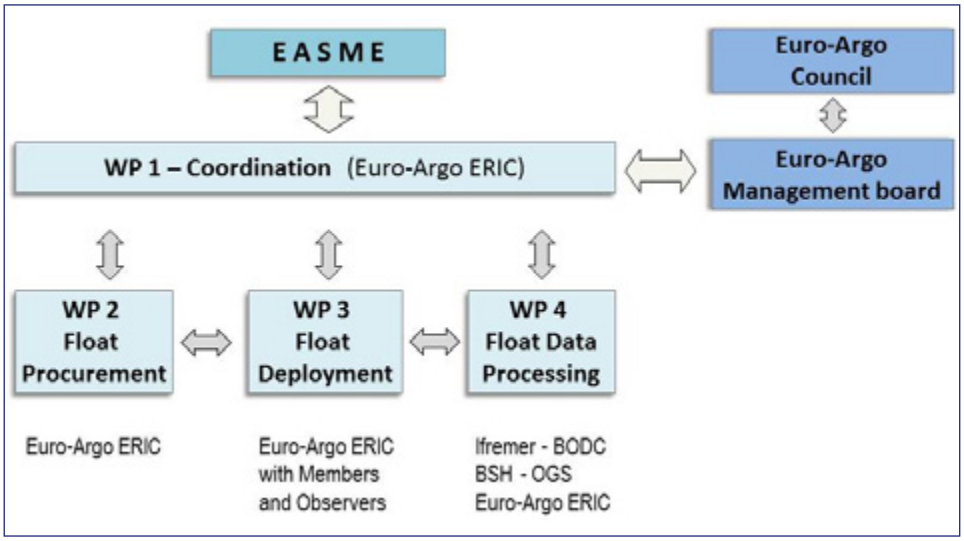


FIGURE 2

MOCCA project logic diagram

The ERIC itself procures all the floats as part of the EASME grant (WP2). The co-funding (20%) for floats has been organized through the transfer of funds from several ERIC members/observers towards the ERIC. The procurement has been done through a EU call for tenders organised by the ERIC and issued in July 2015. NKE Instrumentation (France) won the tender and the following Argo floats were purchased:

- 130 Arvor with Sea-Bird SBE41 CTD (Temperature and Salinity), iridium satellite transmissions,
- 20 Arvor with Sea-Bird SBE41 CTD (Temperature and Salinity), Argos satellite transmissions.

To distinguish between co-financed floats and truly EASME MOCCA floats in the Argo Identification Center (AIC) of JCOMMOPS[5], specific programmes have been defined in the database: MOCCA-GER for German floats (10), MOCCA-IT for Italian floats (2), MOCCA-POL for Polish floats (2), MOCCA-NETH for Dutch floats (14) and MOCCA-EU for the 120 DG-MARE EASME financed floats.

The 150 floats have been received in ERIC premises in 4 batches between April and September 2016. As part of MOCCA project activities, preparing and testing the floats before their shipment to deployment ships is a core action of the ERIC (WP3). These acceptance tests have been performed in the Ifremer tank by the Euro-Argo technical team and consisted in checking all the float components: mission parameters, hydraulic pump and solenoid valve actions, T&S sensors with inter-comparison between floats measurements (see Figure 3), and satellite communications (2-way).

FIGURE 3

Euro-Argo ERIC warehouse, with MOCCA floats waiting for acceptance tests and ship deliveries (left). Floats performing their ascent to surface during acceptance tests in Ifremer pool (right).

Acceptance experiments were made possible by the availaibility of Ifremer test facilities (20m-deep pool, pressure test bank etc.) and bring great added-value to Euro-Argo partners, improving the reliability of floats. For instance a problem with some pressure sensors was detected during the MOCCA tests, with a pressure shift up to 30%, causing errors in CTD plots and computed salinity values. Sea-Bird released an official statement to the Argo community in order to describe the problem, and replaced the defect sensors on the 4 MOCCA floats.

The ERIC with support from its members organizes the deployment of floats and at-sea monitoring (WP3). Ifremer, NERC-BODC, OGS, BSH and Met Office are in charge of the real-time and delayed-mode data processing (WP4) and are identified as partners in the MOCCA project[6]. Finally, the Euro-Argo ERIC with the Management Board analyse the float procurement, deployment, monitoring at sea, processing and dissemination and will provide recommendations for managing the European contribution to the Argo international programme.

COORDINATION OF FLOAT DEPLOYMENT

The Euro-Argo ERIC coordinates operations at sea and associated logistics. The succession of operations is as follows: elaboration of deployment plan, preparation of the floats, shipment, training of the team in charge of the deployment and preparation of metadata for real-time data float processing.

MOCCA deployment plan is elaborated by the ERIC and Euro-Argo members, taking into account the following elements:

- "Strategy for evolution of Argo in Europe" document[7],
- National plans,
- Argo density/age maps (from JCOMMOPS),
- Cruises opportunities from partners and others,
- Recommendations from STAG (Euro-Argo Scientific and Technical Advisory Group).

This leads to the following target deployment areas for the MOCCA floats:

- Southern Ocean (ice-free): poor density in Argo network + recommendation from STAG;
- Deployment of iridium floats in equatorial regions was also deemed important (in terms of extreme events);
- Nordic Seas (ice free): based on actual national plans, gap in the area from target identified in Strategy Document;
- Marginal Seas enhancement: Black Sea, Mediterranean Sea (Aegean, Levantine), Baltic Sea;
- Global Ocean.

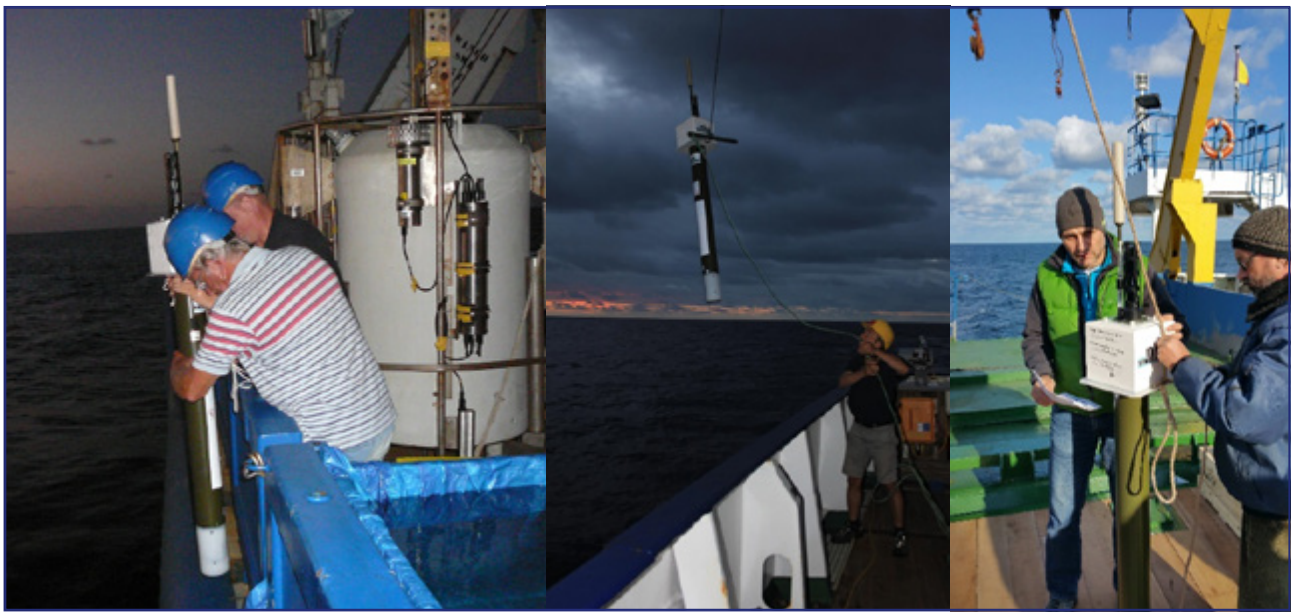


FIGURE 4

MOCCA floats launches in the Caribbean Sea (R/V Pelagia, left), South Tropical Atlantic Ocean (R/V Meteor, middle) and Black Sea (R/V Mare Nigrum, right).

Floats are deployed during research cruises or from ships of opportunity. In February 2017, 66 floats were already deployed and operationnal (see figures 5), and 16 were scheduled for the beginning of 2017. The remaining 68 floats will be deployed in 2017 and early 2018 taking into account international recommendations, research projects, specific European interests and requirements from the Copernicus Marine Service and DG-MARE/EMODNET.

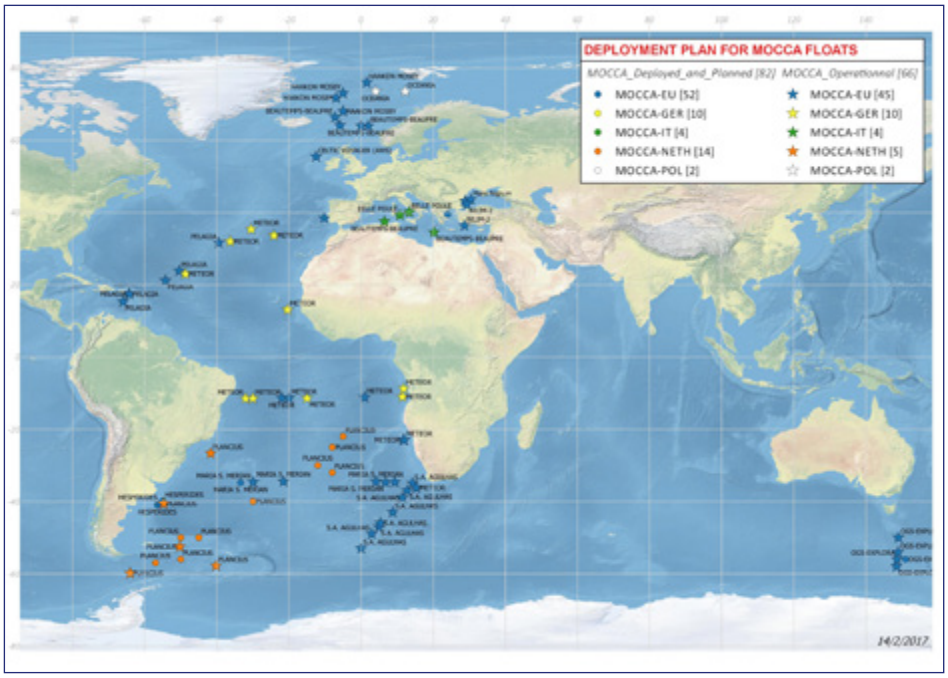


FIGURE 5

MOCCA floats deployment plan. Stars show floats already operational and dots the current plans (including floats already deployed) for the end of 2016 and 2017..

MONITORING THE EUROPEAN FLEET

Monitoring of the behavior of the floats at sea will be coordinated by the Euro-Argo ERIC with the help of national technology specialists (NERC-BODC, Ifremer, BSH and OGS). This monitoring will be targeted to rapidly detect failures that would need to stop deployments and also to provide feedback to manufacturers on anomalies or improvements that need to be studied (see Figure).

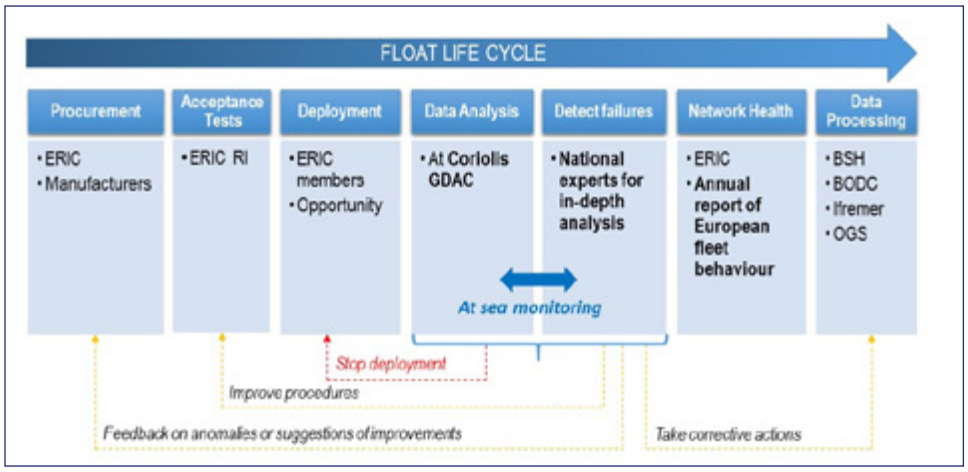


FIGURE 6

At-sea monitoring of MOCCA Argo floats.

The tools developed by Ifremer[8] will be extended to monitor the European fleet. All the technical information provided by Ifremer and BODC DACs[9] will be loaded in the Coriolis database. A technical person from the Euro-Argo ERIC Office will periodically analyse the float behaviour (in particular the failures) and provide a summary report that will be distributed to Euro-Argo partners.

The MOCCA fleet data processing (WP4) is organized through Euro-Argo data centres and delegated to its members. Coriolis (France) and BODC (UK) share the Real-Time processing of the MOCCA floats, according to Argo standard procedures. All MOCCA data are available at http://www.euro-argo.eu/EU-Projects-Contribution/MOCCA/Access-to-MOCCA-Data. Delayed-Mode Quality Control (DMQC, elaborated procedures based on statistical methods and scientific expertise from principal investigators, in order to detect and correct subtle errors in the datasets) will be distributed among the 4 institutes involved in DMQC, based on the areas of deployment: OGS for Mediterranean and Black Sea, BSH for Nordic Seas and part of the Atlantic Ocean, BODC for Southern Ocean and part of the Atlantic Ocean, and Ifremer for part of the Atlantic Ocean. This activity will start in 2018, as a minimum of one year of data is needed before the delayed-mode processing can be performed.

CONCLUSIONS

The goal of the EU EASME MOCCA project is to progress towards the Euro-Argo objectives in monitoring the oceans and the European effort under the international Argo programme by procuring and deploying 150 Argo floats, as well as the data collection, analysis, management, processing and dissemination in the years up to 2020 over the operating lifetime of the floats.

With the MOCCA project Euro-Argo ERIC demonstrates its operational capabilities.

These data are being made freely available as a European contribution to the international Argo programme. The Euro-Argo ERIC is organizing interfaces with operational users and, in particular, with the Copernicus Marine Service and with the ocean and climate change research communities. This will ensure that this action will contribute to an improved monitoring, understanding and prediction of the oceans and climate change, as well as improved ocean services for the development of the blue economy in Europe.

With the MOCCA project Euro-Argo ERIC demonstrates its operational capabilities. This would assist the ERIC in continuing working with the European Commission to sustain such funding for Euro-Argo, to be able to pursue its contribution to Argo and its extensions.

NOTES

1 International Argo Programme, http://www.argo.ucsd.edu

2 Euro-Argo European Research Infrastructure Consortium (ERIC), http://www.euro-argo.eu

3 Executive Agency for Small and Medium-sized Enterprises (EASME) has been set-up by the European Commission to manage on its behalf several EU programmes. See more on https://ec.europa.eu/easme

4 MOCCA Grant Agreement N° EASME/EMFF/2015/1.2.1.1/SI2.709624, Monitoring the Oceans («Monitoring the Oceans and Climate Change with Argo (MOCCA)»). Project info available on http://www.euro-argo.eu/EU-Projects-Contribution/MOCCA

5 WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology in-situ Observing Programmes Support Centre, http://www.jcommops.org

6 Affiliated Euro-Argo partners involved in MOCCA:

Ifremer (French Research Institute for Exploitation of the Sea), France

BSH (Federal Maritime and Hydrographic Agency), Germany

OGS (National Institute of Oceanography and Experimental Geophysics), Italy

BODC (British Oceanographic Data Centre), UK

Met Office, UK

7 Euro-Argo ERIC (2016). Strategy for evolution of Argo in Europe. http://archimer.ifremer.fr/doc/00374/48526/

https://www.euro-argo.eu/content/download/104770/1502786/version/1/file/EA-2016-ERIC-STRAT-V3.pdf

8 http://www.coriolis.eu.org/Data-Products/At-sea-monitoring

9 The international Argo data system is based on two Global Data Assembly Centres (GDACs), a series of 11 national Data Assembly Centres (DACs) and several Delayed-Mode (DM) operators. Their functions are summarized here: http://www.euro-argo.eu/Activities/Data-Processing/Argo-Data-System

PROGRESS ON DATA INTEGRATION WITHIN ATLANTOS

BY
S. POULIQUEN[1], V. HARSCOAT[1], AND WP7 ATLANTOS PARTNERS[2]

THE ATLANTOS PROJECT: CONTEXT, OBJECTIVES AND STRUCTURE

Atlantic Ocean observation is currently undertaken through loosely-coordinated, in-situ observing networks, satellite observations and data management arrangements of heterogeneous international, national and regional design to support science and a wide range of information products. Thus there is tremendous opportunity to develop the systems towards a fully integrated Atlantic Ocean Observing System (AtlantOS) consistent with the recently developed 'Framework of Ocean Observing' (FOO).

The vision of AtlantOS is to improve and innovate Atlantic observing by using the Framework of Ocean Observing to obtain an international, more sustainable, more efficient, more integrated, and fit-for-purpose system. Hence, the AtlantOS initiative will have a long-lasting and sustainable contribution to the societal, economic and scientific benefit arising from this integrated approach. This will be archived delivered by improving the value for money, extent, completeness, quality and ease of access to Atlantic Ocean data required by industries, product supplying agencies, scientist and citizens.

The overarching target of the AtlantOS initiative is to deliver an advanced framework for the development of an integrated Atlantic Ocean Observing System that goes beyond the state-of–the-art, and leaves a legacy of sustainability after the life of the project.

[1] **IFREMER**, Plouzané, France.

[2] https://www.atlantos-h2020.eu

The legacy will derive from the AtlantOS aims to:

- improve international collaboration in the design, implementation and benefit sharing of ocean observing,
- promote engagement and innovation in all aspects of ocean observing,
- facilitate free and open access to ocean data and information,
- enable and disseminate methods of achieving quality and authority of ocean information,
- strengthen the Global Ocean Observing System (GOOS) and to sustain observing systems that are critical for the Copernicus Marine Environment Monitoring Service and its applications and
- contribute to the aims of the Galway Statement on Atlantic Ocean Cooperation

The EU Horizon 2020 AtlantOS project pools the efforts of 57 European and 5 non-European partners (research institutes, universities, marine service providers, multi-institutional organizations, and the private sector) from 18 countries to collaborate on optimizing and enhancing Atlantic Ocean observing. The project has a budget of € 21M for 4 years (April 2015 – June 2019) and is coordinated by GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany (Prof. Dr. Martin Visbeck).

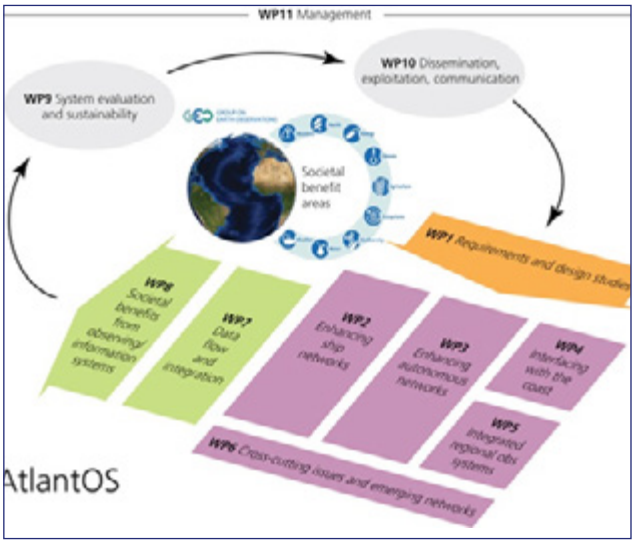


FIGURE 1

AtlantOS work packages schema.

The project is organized along 10 high-level work packages on:

- Observing system requirements and design studies,
- Enhancement of ship-based and autonomous observing networks,
- Interfaces with coastal ocean observing systems,
- Integration of regional observing systems,
- Cross-cutting issues and emerging networks,
- **Data flow and data integration,**
- Societal benefits from observing /information systems,
- System evaluation and resource sustainability.

Engagement with wider stakeholders including end-users of Atlantic Ocean observation products and services will also be key throughout the project.

The AtlantOS initiative contributes to achieving the aims of the Galway Statement on Atlantic Ocean Cooperation that was signed in 2013 by the EU, Canada and the US, launching a Transatlantic Ocean Research Alliance to enhance collaboration to better understand the Atlantic Ocean and sustainably manage and use its resources.

TOWARDS AN INTEGRATED EU DATA SYSTEM

One goal of AtlantOS is to ensure that data from different and diverse in-situ observing networks are readily accessible and useable to the wider community, including international ocean science community and other stakeholders in this field.

To achieve that, the strategy is to move towards an integrated data system within AtlantOS that harmonizes work flows, data processing and distribution across the in-situ observing

network systems, and integrates in-situ observations in existing European and international data infrastructures and Portals. The targeted integrated system shall deal with data management challenges for efficient and reliable data service to users: (1) quality control commons for heterogeneous and nearly real time data, (2) standardization of mandatory metadata for efficient data exchange and (3) interoperability of network and integrator data management systems.

The actors: Networks and Integrators

The **Networks** involved in Data integration for AtlantOS are:

- Ship-based observation Networks: GO-SHIP (Global Ocean Ship-based Hydrographic Investigations Program), VOS (Voluntary Observing Ship)/SOOP (Ship of Opportunity Program), CPR (Continuous Plankton Recorder), fish and plankton surveys, seafloor mapping
- Autonomous observing Networks: Argo, Gliders, Drifters, OceanSITES, EATN (European Animal Tracking Network)
- Coastal observing systems: Ferrybox, FOS (Fishery Observing System), coastal profilers, fixed moorings

Some Networks are organized with DACs and GDACs components. A DAC is a Data Assembly Centre typically operating at either the national or regional scale. A DAC manages data and metadata for its area with a direct link to scientists and operators. The DAC pushes observations to the network GDAC. A GDAC is a Global Data Assembly Centre. It is designed for a global observation network such as Argo, OceanSITES, EGO for Gliders, etc. The GDAC aggregates data and metadata provided by Network DACs, in RT (Real Time) and DM (Delayed Mode).Therefore it's a portal to access, at any time, the "best version" of the Network data.

The European infrastructures or global assembly centres involved as **Integrators** in AtlantOS are:

- For marine environmental data: SeaDataNet for validated and archived data; and the In-Situ Thematic Assembling Centre (INS TAC) component of Copernicus Marine Environment Monitoring Service (CMEMS) for NRT data and for the past 60 years of historical data assembled for reanalysis needs
- for marine biodiversity data: the ICES system, and EurOBIS

The Portals involved as **Integrators** in AtlantOS are:

- EMODnet lots (physics, chemistry, bathymetry, biology) fed by Copernicus INS TAC, SeaDataNet and EurOBIS
- GEOSS

The roadmap: starting point and target

To summarize the situation at the beginning of AtlantOS project, the data acquired by the different in situ observing networks contributing to the AtlantOS project were processed and distributed using different methodologies and means. Depending on the network, the data were either processed following recommendations elaborated by the network and made accessible through a unique portal (FTP or Web), or were processed by individual scientific researchers and made available through National Data Centres or at the Institution level. Some datasets were available through Integrators by ad-hoc links that were developed in past years within projects such as Copernicus, EMODNet, SeaDataNet, etc.

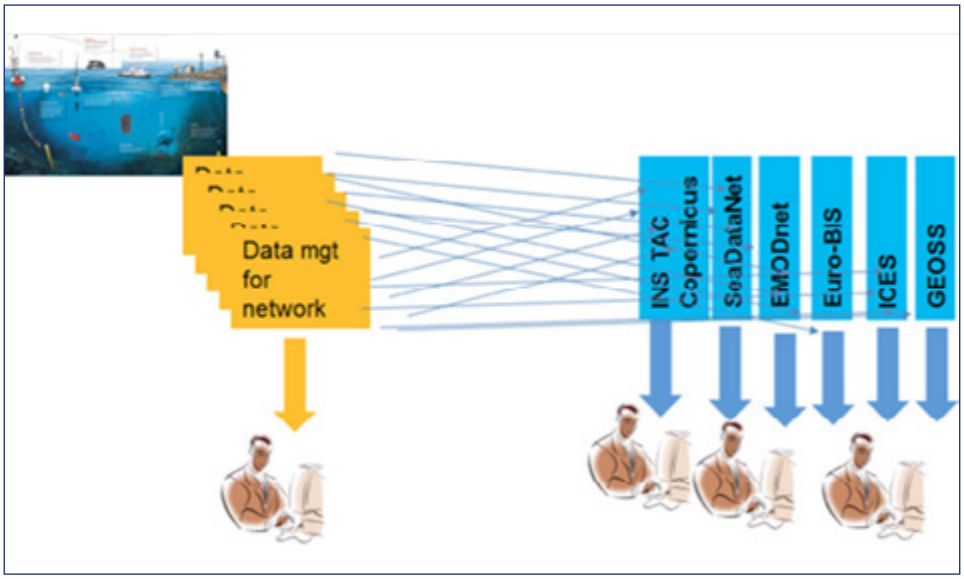


FIGURE 2

Map of the data management situation at the beginning of AtlantOS project

The Networks and Integrators in the AtlantOS project are overall mature systems with long-term experience in the collection, handling, curation and dissemination of data and meta-data. In this regard, the Networks overall have established work-flows and policies for their data-management. Consequently, trying to implement a sovereign and rigid set of for all the Networks and Integrators in AtlantOS to comply with, would be highly challenging and not in the best interest of AtlantOS. The scope of "Data flow and integration" work package was therefore first to explore the data landscape and hereby identify needs for improvements to facilitate the access to the broad array of Atlantic observations and avoid "mixing apples and oranges", which will be to the benefit of all Networks and the users. In this regards, significant progress has been made on the following harmonization and standardization tasks relying on existing international standards and protocols, involving data providers, both Networks and Integrators:

- Identifying the data landscape and prioritizing a list of AtlantOS Essential Variable across the Networks involved in Data flow and integration work package.
- Identifying a minimum set of metadata common vocabularies to be used by all Networks.
- Giving recommendations for providing a minimum level of Near Real Time Quality Control Procedures (NRT QC) for selected Essential Variables (T, S, Current, O2, Chl, Nitrate, Sea Level, Carbon).
- Identifying and improving gaps or impediments in basic services (discovery, viewing and downloading) to distribute the data.

Then a data exchange backbone has been defined to facilitate discovery, viewing and downloading by the users. At the Network level, tools can be set up to: (1) plug the data on the backbone, and (2) to facilitate integration into the Integrators. And finally services to the users shall be enhanced to ease access to existing observations.

The roadmap towards such an integrated EU data system is:

- for the Networks to (1) implement the AtlantOS recommendations for standardization across Networks, (2) plan NRT QC procedures enhancement if needed and (3) facilitate access to Network data,

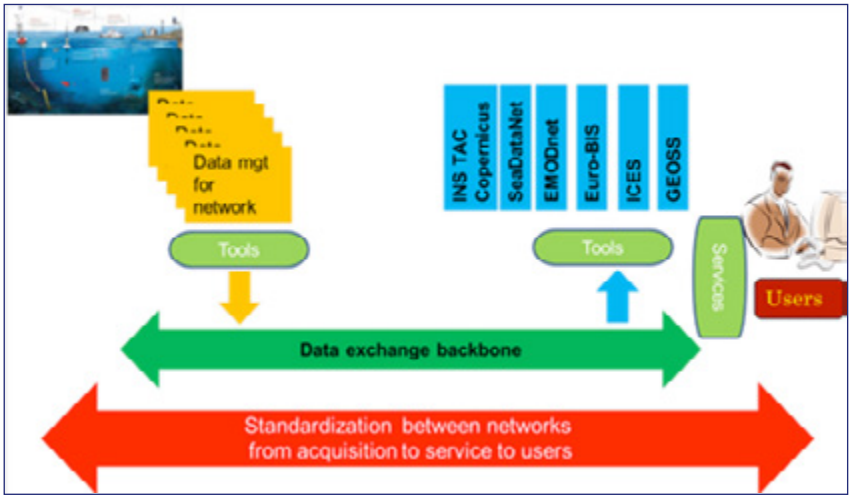


FIGURE 3

Map of the integrated system to be developed within the AtlantOS project

- for the Integrators to enhance (1) their tools for data integration (update their ingestion procedure to integrate new Network data, enhance viewing and downloading services on Network data, perform cross Network assessments and provide feedback to Networks, develop traceability and monitoring facilities for providers and users) and (2) the services to users (facilitate discovery through Network and product catalogues based on ISO standards, provide OGC services (WMS,WFS) to facilitate development of customized user interfaces (either through Integrators or directly from Networks), provide enhanced download facilities (either through Integrators or directly from Networks), facilitate visibility of existing data and provide gap identification).

HARMONIZATION AMONG ATLANTOS NETWORKS

AtlantOS set of Essential Variables and controlled vocabularies

The Networks in AtlantOS offer a wide variety of marine data related to many different scientific disciplines. The data ranges from standard parameters such as common physical ocean measurements, such as conductivity, temperature and density to specialized variables such as isotopes of O2, N2, and Fish and Plankton surveys (ICES). A vast pool of very heterogeneous data is collected on different spatial and temporal scales and with different instrumentations.

The recent decade's rapid development of marine technology allows for the deployment of more and more autonomously operated observation systems. This opportunity to collect almost unlimited amounts of data has also accommodated a significant need for a prioritization of parameters measured in the global oceans as well as in the Atlantic. Various groups and organizations have in recent years debated the identification Essential Variables, such as EOVs (Essential Ocean Variables) or ECVs (Essential Climate Variables), for physics, biogeochemistry and biology/ecosystems variables as part of the Framework on Ocean Observing (GOOS). The prioritization of Essential Variables for AtlantOS was incorporated into the data management plan and is synthetized below.

		WP2 - ship based observing networks					WP3 - autonomous observing networks					WP4 - coastal observing systems				
		GOSHIP	SOOP	CPR	Fish + plankton survey	Seafloor mapping	Argo	Glider	Drifter	OceanSITES	EATN	Tide Gauges	Ferrybox	FOS	Coastal profilers	Fixed moorings
EOV	Physics															
	Temperature	M	A	A	A	A	M	100%	100%	M	A	A	M	M	M	M
	Salinity	M	A	A	A	A	M	100%	4%	M			M	M	M	M
	Current	M					A	100%	100%	A						A
	Sea Level											M				
	bottom depth					M	A				A					
	Air temperature									M		A	A			A
	Air humidity									M			A			A
	Atmospheric pressure								100%	M		A	A			A
	Wind speed									M		A	A			A
	Wind direction											A	A			A
	Rainfall									A						A
	Waves									A						
	Radiative fluxes									A			A			
	Biogeochemistry															
	Oxygen	M			A		M	80%		M	A		M		A	M
	Chla/Fluo	M		A			M	50%		M			M		A	M
	Nutrients (nitrate NO3,...)	M					M	2%		M			A			
	Carbonate system (inorganic carbon)	M	M				not yet			A			M			
	Dissolved Organic Matter	A								A			A			
	Transient Tracers	M						25%								
	Nitrous Oxide	A														
	Turbidity							50%					M	A	A	M
	Biology/ ecosystems															
	Zooplankton			M	M											
	Phytoplankton			M												
	Species			M	M						M					
	Eggs and larvae			M	M											

FIGURE 4

AtlantOS Essential Variables landscape (M stands for Major and A for Additional)

The AtlantOS Essential Variables list of terms (aggregated level), related to ECV –EOV or other, has been defined and was published in June 2016 on the NERC/BODC Vocabulary Server as A05 vocabulary[1], This new vocabulary is mapped to the standards recommended for AtlantOS parameter metadata: P01 (parameter), P07 (CF variable), P06 (units) from SeaDataNet controlled vocabularies managed by NERC/BODC and the internationally assured AphiaID from the WOrld Register of Marine Species (WoRMS)[2], The A05 vocabulary and associated mappings is updated and adjusted when EOVs (especially biological EOVs) and network measurements are confirmed.

Each Network has to define the mapping between the metadata for the parameters in their data and the standards recommended. By doing this, a Network allows mapping on the fly without having to change its dataset.

1 https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/A05/

2 http://www.marinespecies.org/aphia.php?p=webservice

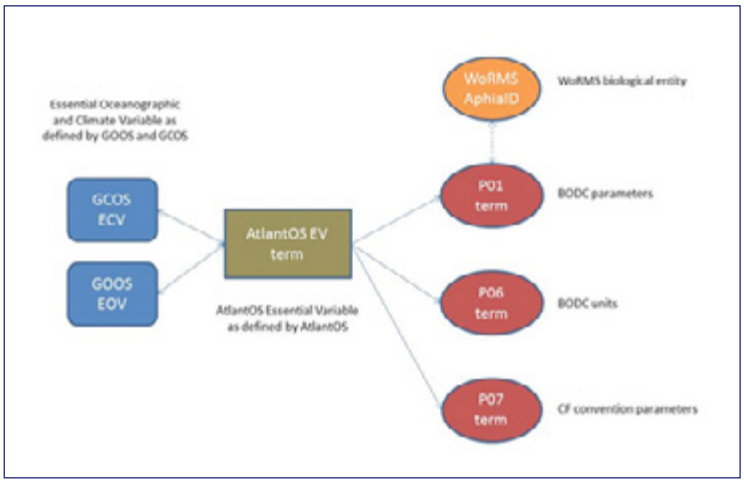


FIGURE 5

Map of AtlantOS recommended vocabularies

NRT QC for selected EOVs

A core of seven EOVs are selected for implementation of common QC procedures because they are acquired and controlled automatically in NRT (24h to several days) by more than one Network among the Networks involved in AtlantOS integration activity

The selected EOVs are:

- Physics : temperature (T), Salinity (T), Current for surface and subsurface and Sea level
- Biogeochemistry : Oxygen (O2), Chlorophyll-A, Nitrate (NO3) and Carbon (pCO2) for surface and subsurface

The recommendations have been compiled by experts on those EOVs and validated by the Networks acquiring those EOVs and performing NRT QC. Also the harmonization recommendations across Networks include QC information to be attached to the data. These include both Quality flags that can be mapped to the SeaDataNet flag scale and when known processing level information ("qualified in NRT using automated procedures" or "processed in DM by Scientist").

Unique identification for platform/station and data providers

The first objective is to identify, without ambiguity, the platform and/or station that has acquired data by including the unique ID of each platform and station in the dataset metadata. These unique IDs will help in the traceability of datasets, identifying which platform carried out the measurements and/or at which station. This will enable better comparisons of/combining of NRT and DM validated data from the same platform/station that are supplied by different routes. For Networks involved in AtlantOS, the two catalogues agreed for unique IDs management are: (1) C17 controlled vocabulary of SeaDataNet listing the codes for all platforms except stations, and (2) ICES station directory for stations. The second catalogue has a geospatial component not present in the C17 SeaDataNet catalogue and thus is more suitable as a station can be relocated and then spatial metadata are needed.

The second objective is to give visibility to the Institutions that provide data. Thus, the mandatory and minimal information for data providers to put in a data file is the Institution code from EDMO. EDMO is the European Directory of Marine Organizations developed under SeaDataNet, and it can be used to register any marine organization involved in the collection of datasets (operators, funders, data holders etc.). It delivers a code for the organization to be included in the data or metadata to harmonize the information (compared to free text) and optimize the discovery of datasets and allows feedback to institution on traceability of use.

DATA EXCHANGE BACKBONE

The aim of this element of the AtlantOS integrated system is to ease discovery, viewing and downloading by users. It is based on the minimum set of agreed standards.

Platform catalogue at the GDAC level

This catalogue, located at the root on an FTP portal, aims to describe the available datasets and platforms of the Network. This facility enables (1) the users to discover more easily and rapidly the data from a Network, and (2) set up monitoring services. Such catalogues are populated (built and updated) from the metadata in the data files on Network FTP sites (minimum data access mean recommended).

It is based on a simple catalogue technique that consists of populating, continuously (creation and update) on file arrival/update, two types of indexes as simple ASCII files besides the data files made available on FTP:

- An index of data files (one line per file described), that contains all the relevant metadata to describe each individual data file, in particular the "provider" with at least the unique institution code defined for AtlantOS (and not an alphanumeric string)
- An index of platforms (one line per platform described) aggregated from the metadata in the data file, that contains all the relevant metadata to describe each platform

This kind of catalogue exists in the Integrator Copernicus INS TAC (content at the end of 2015: 100 000 data files and 30000 platforms). Such index files are useful for setting up synchronization between the GDAC and the user space. They can also be used to create KPIs (Key Performance Indicators) for monitoring purposes on the networks availability, statistics on institutions or countries providing data, maps of the latest data available parameters provided, delays, etc.

Detailed network and platform metadata

A harmonized way of describing platforms helps to trace the provenance of the data and how it was acquired. A trigger to go to the SWE (Sensor Web Enablement) standard from OGC occurs when a Network needs to register a significant amount of metadata to describe a platform or a deployment. In this context, SensorML, one of several implementation standards produced under OGC's SWE activity, is flexible and hierarchical allowing to describe the sensor, instance of sensor on a platform or deployment of a platform. Networks that want to exchange such data should develop the capability to describe their metadata using a SensorML profile.

Concerning these metadata for platform and sensors, it was agreed that it was an issue to be solved at Network level and that harmonization across networks was not seen as a priority. Nevertheless, a recommendation to implement a SensorML marine profile whenever possible will be issued in partnership with other projects such as FIXO3, ODIP2, ENVRI+, SeaDataCloud.

AtlantOS network and product catalogue

As a front window for the WP7 AtlantOS efforts to aggregate and federate observations, it is proposed to build a catalogue of data products and present it in searchable web pages. The catalogue is implemented with the GeoNetwork component of the Sextant Spatial Data Infrastructure. This catalogue will also feed GEOSS common infrastructure.

A template for the AtlantOS data product descriptions to be filed in by the Network representatives and the Integrators in AtlantOS is present has been defined and product sheets will soon be published for Networks and Integrators.

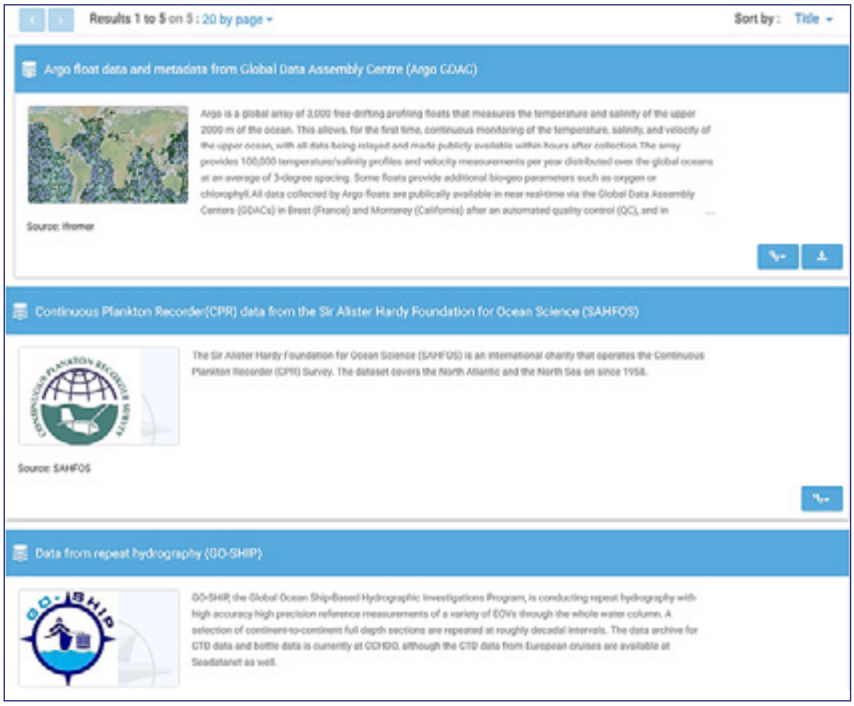


FIGURE 6

Snapshot of Networks in AtlantOS catalogue

Implementation of data citation

To be able to operate observing systems on a long-term basis, operators are often asked to provide evidence that their platform data are essential not only for their study but also for multiple uses. Sharing data with other communities contributes to foster multiple uses of observations but makes it more difficult to trace its effective use.

A DOI (Digital Object Identifier) is a unique identifier for an electronic document or a dataset. Networks can assign DOIs to documents and datasets for two main objectives:

- **Citation** (in a publication the DOI is efficiently tracked by bibliographic surveys)
- **Traceability** (the DOI is a direct and permanent link to the document or data set used in a publication).

In the past years, a lot of progress has been made on data citation, and it is now possible to assign a DOI to the network and link to DOIs assigned to frozen fragments that are archived forever. The proposed way to cite dynamic data (continuously updated in time), is to have a unique DOI plus an additional date stamp (#date) for frozen fragments (like http://dx.doi.org/10.12770/128...cfc9#date).

That is the way DOIs are implemented for Argo Network (http://www.argodatamgt.org/Access-to-data/Argo-DOI-Digital-Object-Identifier). Assigning DOI is also underway for French cruises linked to a Cruise Summary Report (suggestion to add the EXPOCODE = ICES code + date of the Cruise).

A "master" DOI is assigned to each French oceanographic cruise (past and future). Each data set from the cruise is assigned a DOI, linked to the cruise DOI. The cruise landing page is an efficient and dynamic support to give access to all data sets (with their own DOIs) produced by the cruise and the bibliography of publications using these DOIs.

A reference guide name "DOIs for ocean data, general principles and selected examples (Argo, French cruises)" has been issued and is available at http://dx.doi.org/10.13155/44515

DATA INTEGRATION AND SERVICES TO USERS

To facilitate access to AtlantOS data it has been decided to work at two levels:

- **At Networks level** to provide integrated access to all available data. The importance of enhancing services at the Network level is that data managers are close to platform operators and can design the system to fit the platform specificities.
- **At Integrators level** that build thematic services for additional targeted users and will be able to enhance their services with the help of the Networks (integration, update process, archive, etc.).

			WP2 - ship based observing networks					WP3 - autonomous observing networks					WP4 - coastal observing systems			
			GO-SHIP	VOS/SOOP	CPR	Fish + plankton survey	Seafloor mapping	Argo	Glider	Drifter	OceanSITES	EATN	Ferrybox	FOS	Coastal profilers	Fixed moorings
Integrator	Portal	GEOSS	X	Partially	Partially	Partially	X	Partially	Partially	X	Partially		Partially			
		Emodnet-physics	Partially	X				X	X	X	X		X	Partially	Partially	Partially
		Emodnet-chemistry	Partially	Partially				not yet	not yet		not yet		Partially			
		Emodnet-biology			X	X			not yet			target				
		Emodnet-bathymetry					X									
	Infrastructure	Copernicus INS TAC	Partially	Partially				X	X	X	X		Partially	Partially	Partially	Partially
		Seadatanet	Partially	Partially		Underway	Allmost all	Partially	not yet		Partially		Partially			
		Eur-OBIS			X	X						target				
	Data center	Global assembling	CCHDO, CDIAC	Sea surfaces: CDIAC/USA ; PANGAEA for atlas Subsurfaces: NOAA	SAHFOS database	ICES		GDAC	GDAC	GDAC	GDAC	to be set up in Task3.7	GDAC			
		national level	NODC		BODC, OBIS	IMR, IFREMER, DTU,...	NGDC	DAC	DAC	DAC	DAC		DAC	DAC	DAC	DAC

FIGURE 7

Summary table of data integration at the beginning of AtlantOS project

Facilitate access to Network data

The way to facilitate access to Network data for users is to set up a central point from where the data can be uploaded, or rely on existing Integrators to distribute more widely their data. This central point can be either a GDAC for the Network, or a portal with files on FTP and/or web services, allowing machine-to-machine downloading and sub-setting services.

Enhance integration in Integrators and services to users

The started enhancements are:

- All Integrators are planning to connect to new Network GDACs that are setting up to achieve:
  - A more complete data coverage in time and space
  - Better quality of the integrated data as update processes will be easier
  - Extension to more biogeochemistry data essential for Ecosystem modelling
  - Facilitate also links between Integrators (Copernicus INS TAC <-> SeaDataNet, Copernicus INS TAC <-> EMODnet)
- All Integrators are updating their data system to implement the AtlantOS recommendations on metadata and vocabularies for parameters
- Surveys were performed to identify the AtlantOS data that were not integrated yet, and activities are going on with Networks to improve the situation for Copernicus INS TAC, SeaDataNet and consequently EMODnet
- Implement traceability of AtlantOS observations and use methods, and develop monitoring tools and dashboards
- AtlantOS is contributing to GEOSS through different channels including teaming up with ODIP and in addition is promoting its use as the central hub to discover environmental data and information. However, as the GEOSS Common Infrastructure is going through a transitional phase, AtlantOS will explore the best strategy for taking new initiatives like the GEOSS European Data Hub into account.

CONCLUSIONS

The benefit for the Atlantic community will be at different levels. For the Network operators, it allows targeting new users through wider data availability. They also take advantages of new tools and methods to improve traceability of use and monitoring of the network data availability. Furthermore they will be able to implement internationally agreed recommendations for data citation strategy and mapping between network parameters and AtlantOS Essential Variables.

They also take advantages of new tools and methods to improve traceability of use and monitoring of the network data availability.

Operational users will have access to enhanced products with extended time and space coverage for present parameters (T&S Current Sea Level Wave O2 Chl) both for forecast and reanalysis, but also access to enhanced products for Ecosystem model validation. The benefit will probably be more visible in European Seas, but collaboration started also for international partnership and integration of new platforms.

In addition the research community will also benefit of enhanced quality of the historical products in partnership with the Networks and the Integrators.

Finally new monitoring tools developed with WP9 will allow the AtlantOS coordination to have more visibility on what data is freely available for users and provide inputs for the elaboration of the AtlantOS Blueprint, which aims at providing an integrated vision and plan for Atlantic Ocean observations.

A NATIONAL WORKING GROUP ON O2 DATA ACQUISITION AND PROCESSING

BY
V. THIERRY[1], L. COPPOLA[2], C. LAGADEC[1], G. REVERDIN[3], H. BITTIG[2]

ABSTRACT

Dissolved oxygen concentration (referred to as O2 in the following) is a classical parameter in oceanography. O2 can be measured with different methods and from various platforms. A great effort is currently undertaken within the scientific community at national and international levels for the acquisition and processing of O2 data. In this context, there was a strong need to make a state of the art of the various methods implemented in the French laboratories to process O2 data acquired by fixed or mobile platform, to share expertise and eventually to converge toward consistent and common processing methods. To achieve this objective, two meetings in Paris in September 2015 and October 2016 were organized. Highlights of the 2016 meeting discussions are summarized in this paper.

[1] **IFREMER Laboratoire d'Océanographie Physique et Spatiale, UMR 6523 CNRS-IFREMER-IRD-UBO**, Plouzané, France.

[2] **Sorbonne Universités (UPMC Univ. Paris 06), UMR 7093, Laboratoire d'Océanographie de Villefranche, Observatoire Océanologique**, Villefranche/mer, France

[3] **Sorbonne Universités (UPMC, Univ Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory**, Paris, France

CONTEXT

Dissolved oxygen concentration (referred to as O2 in the following) is a classical parameter in oceanography that is used to get insights into oceanic biogeochemical (eg. Riser et al, 2008) and physical processes (Piron et al, 2016). O2 data are also important for operational oceanography that aims now at monitoring and predicting the biogeochemical state of the ocean and marine ecosystems (Brasseur et al, 2009). Deoxygenation of the ocean is also a major concern in the context of climate change (Keeling et al., 2010). Finally there are ongoing work to exploit accurate O2 measurements to estimate, through transfer function derived from neural network techniques, concentration of nutrients (nitrate, phosphate and silicate) and carbonate system parameters (total alkalinity, dissolved inorganic carbon, pH and partial pressure of CO2) (Sauzède et al., in revision). There is thus a strong need in the scientific community to set up and maintain a long term observing systems that will produce a homogeneous and validated O2 dataset.

O2 can be measured with different methods (Winkler titration, electrochemical sensors, optical sensors) and from various platforms (water samples from Niskin casts, ships, moored platforms, autonomous platforms such as Argo floats and gliders, etc...). National and international recommendations exist on well-known methods and sensors, such as Winkler titration or calibration of ship-based O2 profiles from electrochemical sensors such as SBE43 (McTaggart et al, 2010). However, the implementation of those recommendations often differs from one laboratory to another. There are also many ongoing R&D activities to understand the behavior of the newer optical sensors (Bittig et al., 2015a,b; Bittig et al, 2016a,b) and to define quality control procedure when those sensors are implemented on mobile platforms such as Argo floats (Takeshita et al., 2013, Thierry et al, 2016a,b, Schmechtig et al., 2016). Some of those activities are conducted as part of international projects (AtlantOS) and programs (Argo and BGC-Argo) or international working groups, such as the SCOR Working Group 142 on Quality Control Procedures for Oxygen and Other Biogeochemical Sensors on Floats and Gliders (http://www.scor-int.org/SCOR_WGs_WG142.htm). Some activities related to O2 data also occur at a more regional scale such as the intercomparison exercise on O2 measurements based on Winkler titration conducted in Brittany, France, on an annual basis and involving labs from Brest and Roscoff.

This non-exhaustive list reveals the great effort currently undertaken within the scientific community at national and international levels for the acquisition and processing of O2 data. In this context, there was a strong need to make a state of the art of the various methods implemented in the French laboratories to process O2 data acquired by fixed or mobile platforms (moorings, ship, glider and Argo floats), to share expertise and eventually to converge toward consistent and common processing methods. To achieve this objective, we organized two meetings in Paris in September 2015 and October 2016 with funding from the SOERE CTDO2.

HIGHLIGHTS OF THE DISCUSSION OF THE 2016 MEETING

Owing to the major step forward that was done in recent years on the understanding of the optode sensors (Aanderaa optode and SBE63) equipping Argo floats, a Matlab toolbox called LOCODOX was developed at LOPS/Ifremer by E. Brion and A. Piron from ALTRAN company to correct oxygen data in a routine way. Indeed, data from those sensors are always biased low and a correction needs to be applied. LOCODOX uses either reference in situ profiles (climatological atlas or ship-based O2 profiles) (Takeshita et al., 2013) or in-air measurements (Bittig et al, 2015). LOCODOX works with Argo profiles in format 3.1 and produces delayed mode NetCDF files complying with Argo 3.1 format and following the BGC/DOXY QC manuals (Thierry et al, 2016b, Schmechtig et al., 2016). This tool, which will be available to the scientific community, has already been used to correct 12 floats deployed as part of the OVIDE project [(1)]. The corrected data are now available at Coriolis DAC and in the Argo data stream. Note that oxygen data adjusted in delayed mode are available in the DOXY_ADJUSTED fields in the Argo netcdf files. Because of the systematic bias observed on the oxygen sensors equipping Argo floats [(1)], the DOXY field containing the raw data should not be used directly for scientific application.

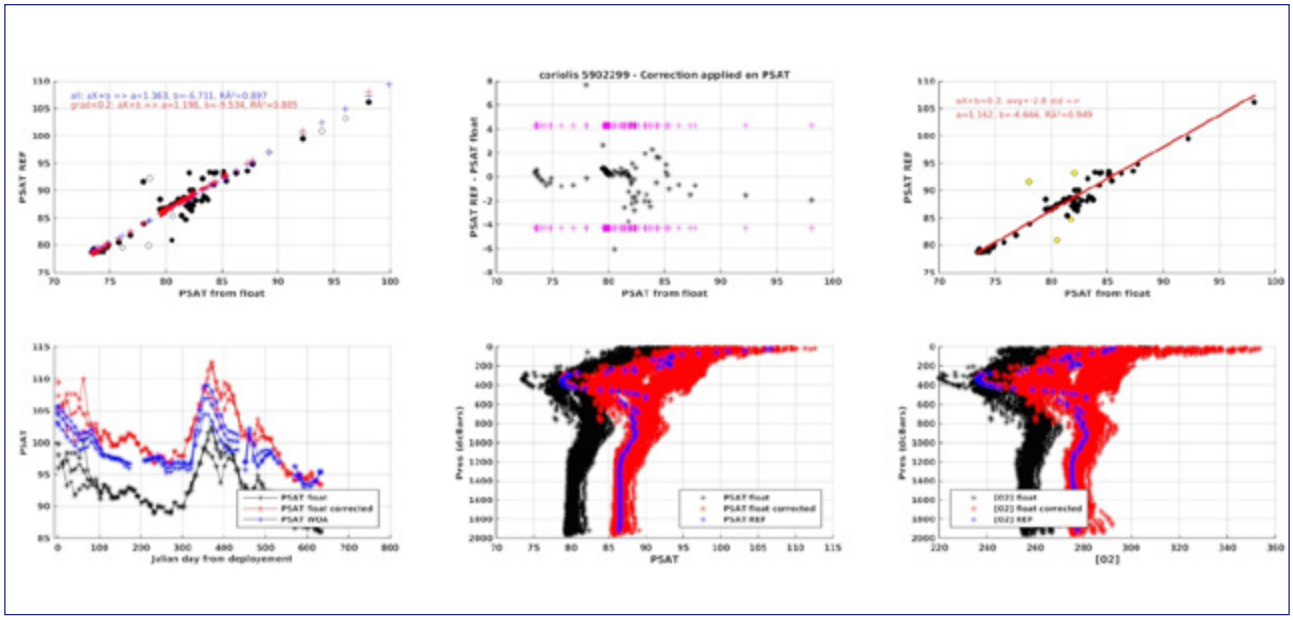


FIGURE 1

Exemple of plots produced by LOCODOX. In this case, float 5902299 was corrected based on a comparison of the first ascending profile of the float with an in situ reference profile acquired at float deployment. The correction is done in considering the percentage of saturation (PSAT).

(Upper panels) The three panels show the regression between the Argo profile and the reference profile.

(Lower left panels) PSAT in the upper 10m from the raw data (black curve) and the corrected data (red curves). PSAT estimated from the World Ocean Atlas at the float position is also provided for comparison (blue curves).

(Lower middle panel). PSAT values from the raw data (black curves), the adjusted data (red curves) and the reference profile (blue curve).

(Lower right panel) Same as the middle panel but for dissolved oxygen concentration value (DOXY et DOXY_ADJUSTED) in mumol/kg. Note that the correction is about 20 mumol/kg for this float.

The real-time procedure defined for Argo data is now implemented on glider data by Coriolis. During the 2016 meeting, we identified the need to set up a process to evaluate the quality of oxygen sensor before and after deployment in order to correct the data in delayed mode.

The behavior of the oxygen sensor (optode) on moored platforms still needs to be investigated and a procedure to qualify the data needs to be defined more precisely.

Discussions revealed that the GO-SHIP recommendations have been implemented in different ways in the French laboratories. One difference is the vertical axis (density or pressure) choice for comparing samples to the SBE43 data from the CTDO2 cast. While a comparison on density levels (as done in LEGOS for instance) is preferred in the upper oceanic layers where internal waves are dominant features, it cannot be applied in the weakly stratified deep layers where the comparison must be done on pressure levels (as in LOPS for instance). We agreed during the 2016 meeting to work on a merged method. More generally, recommendations on the quality control and correction procedures of ship-based O2 data acquired by the electrochemical SBE43 sensor need to be defined as well as a strategy for the minimal O2 sampling during oceanographic cruise required for O2 processing when the ship time is limited.

The following webpages provide further information (meeting reports and presentations):

http://www.umr-lops.fr/Recherche/Equipe-Ocean-Climat/Evenements/Atelier-O2-2015

http://www.umr-lops.fr/Recherche/Equipe-Ocean-Climat/Evenements/Atelier-O2-2016

CONCLUSIONS

The 2015 and 2016 meetings have lead to progress in providing the status on state of the art methods implemented in the French laboratories to process O2 data, to share expertise and converge toward consistent and common processing methods. We plan to maintain such activity on an annual basis. If you want to participate to the next meeting, please contact L. Coppola (coppola@obs-vlfr.fr) and V. Thierry (vthierry@ifremer.fr).

REFERENCES

Bittig, H. and Arne Körtzinger, 2015a: Tackling oxygen optode drift: near-surface and in-air oxygen optode measurements on a float provide an accurate in situ reference. J. Atmos. Oceanic Technol., 32, 1536–1543. doi:0.1175/JTECH-D-14-00162.1

Bittig H. C., B. Fiedler, P. Fietzek, and A. Körtzinger, 2015b : Pressure Response of Aanderaa and Sea-Bird Oxygen Optodes, J. Atmos. Oceanic Technol., 32, 2305-2317. DOI: 10.1175/JTECH-D-15-0108.1.

Bittig Henry, Kortzinger Arne, Emerson Steve, Gilbert Denis, Neill Craig, Thierry Virginie, Bronte Tilbrook, Uchida Hiroshi (2016a). SCOR WG 142: Recommendations on the conversion between oxygen quantities for Bio-Argo floats and other autonomous sensor platforms. http://doi.org/10.13155/45915

Bittig Henry, Kortzinger Arne, Emerson Steve, Gilbert Denis, Neill Craig, Thierry Virginie, Bronte Tilbrook, Uchida Hiroshi (2016b). SCOR WG 142: Recommendation for oxygen measurements from Argo floats, implementation of in-air measurement routine to assure highest long-term accuracy. http://doi.org/10.13155/45917

Brasseur, P., N. Gruber, R. Barciela, K. Brander, M. Doron, A. El Moussaoui, A.J. Hobday, M. Huret, A.-S. Kremeur, P. Lehodey, R. Matear, C. Moulin, R. Murtugudde, I. Senina, and E. Svendsen (2009), Integrating biogeochemistry and ecology into ocean data assimilation systems. Oceanography 22(3):206–215, doi:10.5670/oceanog.2009.80.

McTaggart K.E., G.C. Johnson, M.C. Johnson, F.M. Delahoyde, and J.H. Swift (2010) Notes on CTD/O2 data acuqisition and processing using SEA-BIRD hardware and software (as available). The GO-SHIP repeat hydrography manual : a collection of expert reports and guidelines. IOCCP Report n°14, ICPO Publication series n°134, version1.

Keeling R. F., Arne Körtzinger, and Nicolas Gruber (2010) Ocean Deoxygenation in a WarmingWorld Annu. Rev. Mar. Sci. 2:199–229

Piron A., V. Thierry, H. Mercier and G. Caniaux, 2016: Observation of basin-scale deep convection in the Irminger Sea with Argo floats in the 2011-2012 winter, Deep-Sea Research Part I, 109, 76-90, doi: 10.1016/j.dsr.2015.12.012

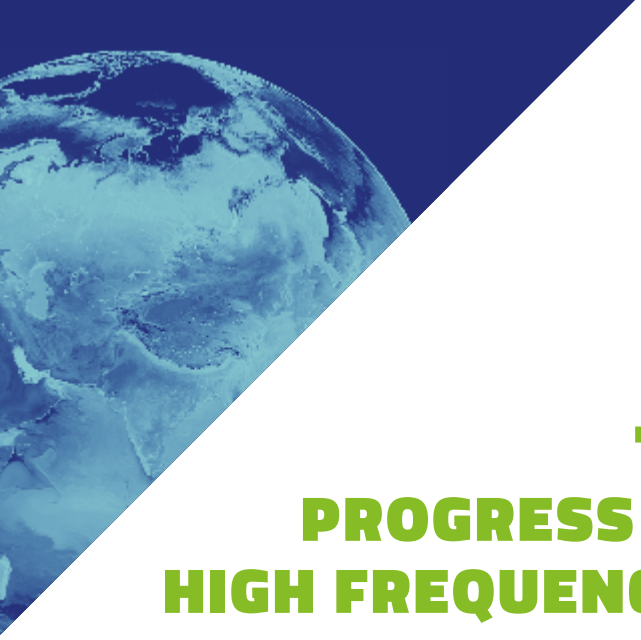
Riser, S. C., and K. S. Johnson (2008), Net production of oxygen in the subtropical ocean, Nature, 451(7176), 323–325, doi:10.1038/nature06441.

Sauzède, R., Claustre, H., Pasqueron de Fommervault, O., Bittig, H., Gattuso, J.-P., Legendre, L. and K. Johnson, Estimates of water-column nutrients concentration and carbonate system parameters in the global ocean: A novel approach based on neural networks. Frontiers in Marine Science, in revision

Thierry Virginie, Bittig Henry, The Argo-Bgc Team (2016a). Argo quality control manual for dissolved oxygen concentration. http://doi.org/10.13155/46542

Thierry Virginie, Gilbert Denis, Kobayashi Taiyo, Schmid Claudia, Kanako Sato (2016b). Processing Argo oxygen data at the DAC level cookbook. http://doi.org/10.13155/39795

Schmechtig Catherine, Thierry Virginie, The Bio Argo Team (2016). Argo quality control manual for biogeochemical data. http://doi.org/10.13155/40879



PROGRESS TOWARDS A FRENCH HIGH FREQUENCY OCEAN SURFACE WAVE RADAR NETWORK

BY

C. QUENTIN [1], B. ZAKARDJIAN [1], L.MARIÉ [2], A. RUBIO [3], A-C. BENNIS [4], F. DUMAS [5], A.SENTCHEV [9], G. SICOT [6], Y. BARBIN [1], S. JOUSSET [5], A. BONNAT [7], J. MADER [3], Y. OURMIÈRES [1], G. CHARRIA [2], S. TAROT [7], D. MALLARINO [8]

ABSTRACT

The use of High Frequency ocean Radars (HFR) as a research and monitoring tool for the observation of the coastal ocean has steadily developed throughout the world over the last decades, especially in the US where more than one hundred sites are currently maintained in a multi-purpose operational network (https://ioos.noaa.gov/project/hf-radar/) all over the country coasts. Such a networking effort is currently being built at the EU level through the JERICO-Next H2020 program in link with Euro-GOOS, with the objectives to share expertises, best practice, quality control procedures and data dissemination tool and to carry the existing EU HFR systems to the level of a unified operational HFR network. The LEFE/GMMC working group ReNHFOR (Research and Networking for High Frequency Oceanographic Radar) acts at the French level to structure the French contribution to this PanEuropean HFR network and assists the French oceanography community in the development of new applications both for academic and operational purposes. This paper presents an overview of the technology, its state of development in France, its potential as a supporting tool for ocean modeling and monitoring, and the integration of French HFR activities in the EU context.

[1] **Univ Toulon, Aix Marseille Univ, CNRS, IRD, MIO UM 110, Mediterranean Institute of Oceanography,** La Garde, France .
[2] **IFREMER, Univ. Brest, CNRS, IRD, LOPS,** Plouzané, France
[3] **AZTI Marine Research**, Pasaia, Spain
[4] **UNICAEN, CNRS, M2C,** Caen, France
[5] **Shom,** 13 Rue de Châtellier, 29200 Brest, France
[6] **ENSTA-Bretagne, CNRS, UBO, UBS, ENIB, Lab-STICC,** Brest, France
[7] **IFREMER, IMN/IDM/SISMER,** Plouzané, France
[8] **Aix Marseille Univ, CNRS, IRD, OSU PYTHEAS,** Marseille, France
[9] **Université de Lille, ULCO, CNRS, LOG,** Wimereux, France

INTRODUCTION

High Frequency Surface Wave Radar is a shore-based remote-sensing technique which permits to monitor and observe the ocean surface current field and surface oceanic transport, accurately, over large areas (typically 60 X 60 km at 12 MHz), in real-time and at high spatial (3-5 km) and temporal resolutions (10 min-1 hour). HFRs exploit the so-called Bragg scattering of decametric (5 to 45 MHz) electromagnetic waves at grazing incidence on ocean surface waves of precisely half wavelength. The phase velocity of these surface waves can be retrieved from the Doppler shift experienced upon scattering by the incident electromagnetic waves [Crombie 1955]. The dominant part of this phase velocity is due to the intrinsic motion of the surface waves, but deviations from the theoretical value are exploited to retrieve the line-of-sight projection of the surface current, also called the "radial velocity" [Steward & Joy, 1974]. Radial velocity observations from several (at least two) radar sites can be combined to provide current vector maps over the area of radar coverage.

The main scientific interest of these radars lies in their ability to characterize the rapid response of coastal circulation to wind forcing (as sea of wind, or inertial waves) and its interaction with offshore circulation (vortices and coastal/offshore current). These observations are essential for the validation of the response of coastal circulation models to these forcing events in terms of circulation structures, i.e. instabilities of currents, coastal jets and gyres over a wide spectrum of scales, that play a prime role in the dispersion or retention of pollutants, planktonic (possibly toxic) species and larvae, and more generally in cross-shelf exchanges. Real-time monitoring of surface currents over wide coastal areas is also essential for the management of crisis situations (for search and rescue support, or in case of pollution event) as well as in terms of environmental applications (stranding of stinging jellyfish, monitoring of macro-waste). Indeed, this technology is now «out» from the development laboratories, and is now a well-accepted and cost-effective technology, routinely used for real-time monitoring of ocean currents in numerous places around the world, specially in the US where more than one hundred of sites are currently maintained in a multi-purpose operational network (https://ioos.noaa.gov/project/hf-radar/) all over the country coasts. The deployment of such national and international operational radar networks provides a continuous flow of real-time data with high potential in terms of data assimilation and significant improvement in the performance of operational ocean models (e.g. Lewis et al., 1998; Oke et al 2002, Paduan and Shulman 2004, Wilkin et al., 2005, Sentchev et al., 2006, Barth et al., 2008).

Such a networking effort is currently being built at the EU level through the JERICO-NEXT (Joint European Research Infrastructure network for Coastal Observatory – Novel European eXpertise for coastal observaTories) H2020 project, in link with Euro-GOOS and EMODnet, with the objectives to share experience and expertise on the technology, improve data processing techniques, and to develop centralized data archiving, distribution and quality control infrastructures so as to get the existing EU HFR systems at the level of a unified operational HFR network. This action is reinforced by the Copernicus Marine Environment Monitoring Service (CMEMS), which supports the project Innovation and Networking for the integration of Coastal Radars into European mArine SErvices (INCREASE) for the integration into the service Evolution of all European HFR data. A first effort towards a national coordination at the French level has recently emerged, and is supported by LEFE/GMMC through the Research and Networking for High Frequency Oceanographic Radars (ReNHFOR) project, in the framework of the coastal component of the Coriolis programme (http://www.coriolis-cotier.org).

In the following sections we present first a short overview of this technology, its potential, limitations and constraints, the state of the art of the French HFR installations, some results illustrating the potential for the support of ocean modeling and, lastly, the EU framework in which the ReNHFOR project fits.

WORKING PRINCIPLE

General basis

A typical HFR system is composed of at least 2 stations, each comprising a transmitter and a receiver. These can be either collocated (monostatic configuration) or separated by a significant fraction of the observation range (bistatic configuration). The transmitter radiates vertically polarized radio waves with wavelengths between 60 m (5 MHz) and 6.6 m (45 MHz). At these frequencies, part of the transmitted energy is trapped in the "surface-wave" mode, and remains bound to the sea surface, following the Earth curvature. This allows HFRs to measure sea surface properties well beyond the optical horizon, in contrast with the more familiar microwave radars.

Surface-wave back-scattering at decametric wavelengths is dominated by the Bragg coherent scattering mechanism [Crombie 1955]. The conversion of the radio wave incident

from the transmitter into a scattered wave propagating towards the receiver is dominated by the sea state component with wavevector equal to the difference of the scattered and incident waves wavevectors. Straightforward geometry shows that in the monostatic configuration the hydrodynamic "Bragg waves" are the waves propagating along the radar line-of-sight, and in the bistatic configuration the waves propagating along the bissector of the Emitter/Observation point/Receiver angular sector. The wavelength of the hydrodynamic "Bragg waves" is given by

$$\lambda_{Bragg} = \frac{\lambda_{EM}}{2\cos\frac{\phi}{2}}$$

where λ_{EM} and ϕ are respectively the wavelength of the electromagnetic waves and the angle between the transmitter and receiver directions as seen from the observation point. In the monostatic configuration, where φ=0, the "Bragg waves" wavelength is half the electromagnetic waves wavelength.

The "Bragg waves" move along their propagation direction, approximately at the velocity given by the deep-water dispersion relation. This motion induces a Doppler shift in the received signal. A typical Doppler power spectrum is shown in Figure 1.The power spectrum of the received signal features two "first-order" peaks produced by the coherent scattering of the incident signal on Bragg waves propagating in both directions, and a broader continuum produced by "second-order" processes, such as secondary scattering of an already scattered wave, or perturbations of the Bragg waves due to their interactions with other sea state components [Barrick, 1972; Weber and Barrick, 1977; Barrick and Weber, 1977; Broche et al, 1983; Ardhuin et al, 2009].

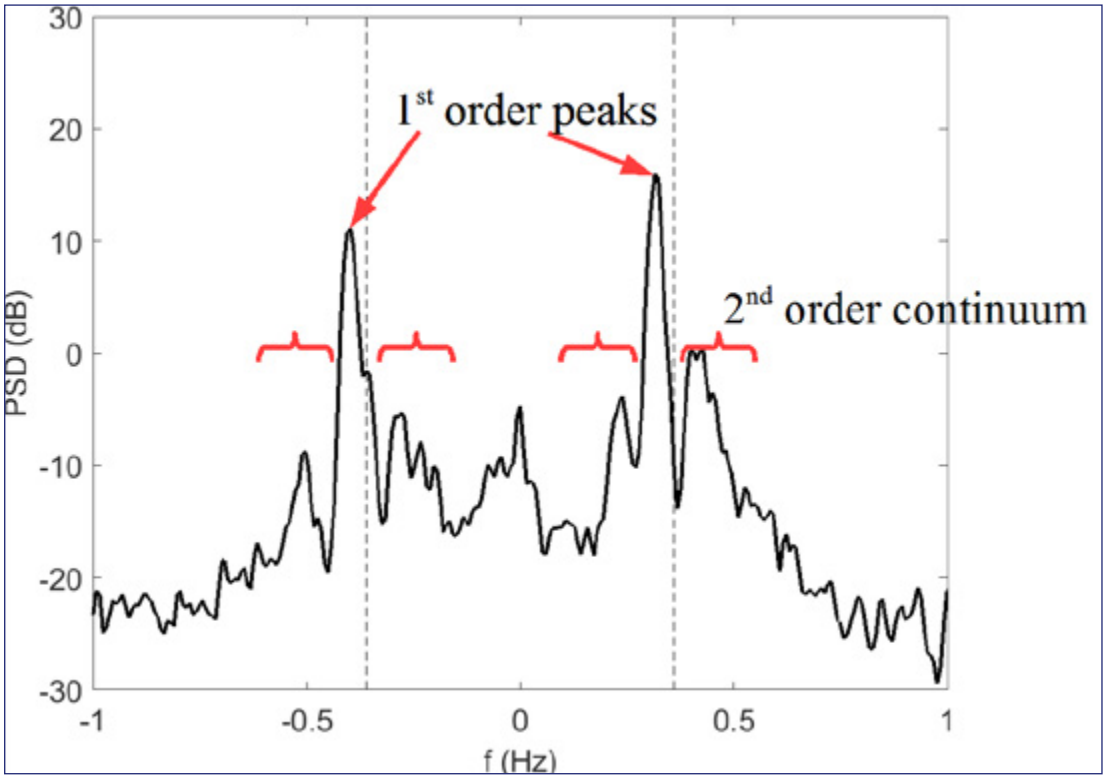


FIGURE 1: TYPICAL DOPPLER SPECTRUM

Typical Doppler spectrum of the signal back-scattered from a single range/azimuth cell, obtained by beam-forming raw data from the SHOM Garchine site on May 28, 2009, 13:39 UTC. The arrows mark the first-order peaks. The shift towards negative frequencies of the first-order peaks from the theoretical frequencies (thin vertical lines) shows that the surface current projection on the line-of-sight is directed away from the radar. The height difference between the two first-order peaks shows that Bragg waves moving towards the radar are higher than those moving away, indicating that wind is blowing towards the radar. Curly brackets mark the frequency ranges in which a 2nd-order continuum is visible. This signal can be used to retrieve sea-state information

A number of observable quantities can be retrieved from the analysis of the power spectral density of the backscattered signal:

The surface current projection in the Bragg waves propagation direction can be retrieved from the shift of the first-order peak frequencies from the theoretical value. This is the best established use of HFRs [Stewart and Joy, 1974; Gurgel et al, 1999]. The influence of shear in the surface current on the hydrodynamic dispersion relation has been studied by [Kirby and Chen, 1989], who proposed the rule of thumb that the observable quantity is the weighted average of the surface eulerian currents, with a weighting function decreasing exponentially with a depth scale of $\lambda_{EM}/8\pi$.

The ratio of the heights of the first-order peaks corresponding to forward and backward-propagating Bragg waves can be exploited to derive the angle of the wind direction with the Bragg waves propagation direction over the study area [Harlan and Georges, 1994].

Statistics of the second-order spectrum can be analysed to retrieve properties of the sea state such as significant wave height, or to invert the complete 2D sea surface elevation spectrum [Hasselman, 1970; Wyatt et al, 2003].

Comparisons of the power spectra of the backscattered signal at different radar frequencies has been used to study sea surface shear [Ivonin et al, 2004].

In order to implement this measurement principle into a working instrument, however, care must also be given to a number of practical considerations that are detailed in the following sections.

Range Resolution

Range resolution can be performed in a number of ways. Table 1 below provides a summary of the relevant figures. As a rule of thumb, obtaining a good quality Doppler signal requires each range cell to be larger than at least 150 Bragg wavelengths. Early instruments transmitted pulsed waveforms. This has the drawback that the duration of the pulses sets the range resolution of the instrument, while the energy of each pulse sets the observation range. Achieving both a usable observation range and a usable range resolution requires transmission of short (a few tens of µs), high energy (several hundred Watts peak power) pulses, causing technical and regulatory issues. In modern instruments this issue is circumvented using "pulse-compression" techniques. Low power (a few tens of Watts) radio waves are transmitted over a bandwidth broad enough (several tens of kHz) to achieve the required range resolution, but over durations of several tenths of seconds, sufficient to transmit enough energy to achieve long observation ranges. The most frequently used waveform is the linear frequency-modulated chirp, either continuous or chopped, but other techniques such as phase-coded waveforms are coming of age.

	ITU frequency bands	Radar wavelength	ocean wavelength	ocean wave period	Equivalent integration depth for current	Bragg frequency peak	Doppler offset for speed 5 cm/s	Typical minimum acquisition time	Typical sweep bandwidth	Sweep repetition frequency	Typical range resolution	Typical maximum range for current analysis	Upper Significant Wave Height Limit
	f_{em} (kHz)	λ_{EM} (m)	λ_{Bragg} (m)	T (s)	$\lambda_{EM}/8\pi$ (cm)	F_B (Hz)	δf (mHz)	>= (1/δf *3) (minutes, 60s)	Bw (kHz)	f_e (Hz)	dr (km)	R_{max} (km)	$H_{1/3}$ (m)
Long Range	4 438 4 488	67	34	4,6	267	0,22	1,5	35	25	1	12	220	25
	5 250 5 275	57	28	4,3	227	0,23	1,8	30	25	1	12	175	25
Medium range	9 305 9 355	32	16	3,2	128	0,31	3,1	16	25	1	12	80	13
	13 450 13 550	22	11	2,7	88	0,37	4,5	11	50	2	3	60	13
	16 100 16 200	19	9	2,4	74	0,41	5,4	9	50	2	3	60	13
High Resolution	24 450 24 600	12	6	2,0	49	0,51	8,2	6	150	2	1	30	7
	26 200 26 350	11	6	1,9	45	0,52	8,8	6	150	2	1	30	7
	39 000 39 500	8	4	1,6	30	0,64	13,1	4	250	4	300 m	20	3
	42 000 42 500	7	4	1,5	28	0,66	14,1	4	325	4	250 m	15	3

TABLE 1: HF RADAR PERFORMANCE VS OPERATING FREQUENCY

Each row corresponds to one of the ITU frequency bands allocated for oceanographic radar with the lower and upper band limits in frequency (fem). The radar wavelength is calculated for the centre of the frequency band. The ocean wavelength (Λ) is deduced as the half of the radar wavelength. The ocean wave period (T) and the Bragg frequency (FB) are considered for depth water and deduced from the relation of dispersion for ocean waves. The equivalent integration depth for current measurement is commonly used as to be the ocean wavelength divided by 4π ($\lambda_{Bragg}/8\pi$), simplification of Steward & Joy (1974). The additional Doppler shift is given for a radial velocity of 5cm/s, projection of the current speed in the direction of the radar. The typical minimum acquisition time refers to the integration time for calculating the Doppler spectra (time-frequency analysis by fast Fourier transform), and is not dependent of the antenna processing method. The range resolution is calculated taking into account the typical bandwidth (Bw) and the sweep repetition frequency (fe). The typical maximum range for current analysis is based on an averaged transmitted power of 40 watts and standard conductivity (temperature, salinity, sea state and radio interference noise can affect this). At the upper limit of the significant wave height, the 2nd order saturates the 1st order and no current measurements are possible.

Azimuth attribution

Azimuth attribution is currently performed in two conceptually different ways.

The "Beam-Forming" approach uses an extended "Phased-array" of receiving antennas (most current sites use linear arrays, but other geometries have also been used). Linear combinations of the signals from the individual antennas are formed at post-processing to deterministically "sweep" a beam of enhanced sensitivity over the area to be observed. A Doppler spectrum can then be computed for each range/azimuth cell. For the standard case of a linear array of equidistant antennas using the "delay-and-sum" method, ambiguous replicas of the main lobe appear when antennas are spaced more than λEM/2 apart. An N-elements array is thus at most (N-1) λEM/2 long. The usual formula for the resolving power of a linear sensor with flat weighting function yields the -3 dB main lobe half-width as 1/(N-1) x 50° for an N-antennas array, or 3.3° and 6.6° for 16 and 8 antennas devices, respectively. More sophisticated weighting functions can be used to reduce unwanted side lobes, at the expense of a slight loss in main lobe selectivity.

The "Direction-Finding" approach can provide high angular resolution observations even with compact receiving arrays for which beam forming would produce useless results. It thus provides an interesting alternative for deployments where large antenna arrays are not a practical option, due for instance to space or geometric constraints. However, it does not allow one to sweep deterministically the observation area. Instead, it relies on the sporadic appearance of strongly back-scattering patches of sea within the domain, which are localized in frequency using standard Fourier-transform approaches, then in azimuth using high-resolution goniometric techniques. Techniques of this kind, as for instance the so-called MUSIC [Schmidt, 1986; Lipa, 2006] algorithm, which relies on a subspace decomposition of the interspectral matrix of the antenna signals, do not have the same accuracy limitations as beam-forming, and can provide much better radial current azimuth resolution with the same receiving array.

On compact antenna systems only the Direction Finding algorithm can be applied. On linear-array systems, both algorithms can be used. DF processing has been applied on a subset of the Iroise Sea data collected by Shom, revealing small-scale structures of the tidal flow in the straits of the Ushant-Molène archipelago [Sentchev et al, 2013], that were out of reach of BF processing.

Doppler resolution

This step can be performed before (Direction-Finding) or after (Beam-Forming) azimuth allocation. The Beam-Forming approach produces, for each range/azimuth cell, a time series of complex back-scattered wave amplitude samples, from which the Doppler spectrum is produced using Fourier-transform techniques. A mix of coherent processing and incoherent integration is usually performed. In the Direction-Finding approach, the complex signals from the different antennas are Fourier-transformed to produce estimates of the interspectral matrix, which are then incoherently summed. This step distributes the observed high-cross-section sea patches in Doppler shift classes. For each Doppler class, azimuth allocation is then performed to progressively fill in a current line-of-sight projection map. In both cases, as resolution in Doppler shift varies as the inverse of the coherent observation segments duration, a trade-off has to be found between radial current resolution and time resolution, which usually limits the time resolution to a few tens of minutes.

Total current retrieval and gap-filling

A single HFR subsystem (transmitter Tx and receiver Rx) measures a projection of the current vector on radar beams. This projection points towards or away from the couple (Tx, Rx). For purely geometrical reasons, surface current total vectors reconstruction from HFR requires that any patch of the ocean surface is directly viewed by at least two stations forming an angle that is wider than a limit given by the level of acceptable geometric dilution of precision (GDOP). This is not always possible due to coastline geometry in the stretch in between two radar stations or in the line between two stations. Furthermore, radial velocity maps may contain gaps, due to environmental effects, like increased external noise or low signal due to ocean surface conditions. Many applications, such as Lagrangian numerical modelling of transport and dispersion of material in sea water, require continuous time series of current velocity. Gap filling in radar data is performed prior to these applications. The Totals current method is a local fitting method which reconstructs a current vector field by solving a least-squares equation (Davis, 1985) without filling the missing values. The Optimal Interpolation (OI) computes surface vector current fields from radial velocity measurements and produces a covariance of the uncertainty of the estimated vector current fields [Kim et al, 2008].The Open-boundary Modal Analysis (OMA) method is a global method which fits a series of eigenfunction modes to available radial measurements [Kaplan and Lekien, 2007]. Yaremchuk and Sentchev (2009) proposed a variational method for reconstruction of the total vector

based on the reduction of high-frequency variability of the surface vorticity and divergence fields. They more recently tested also a combined EOF/variational approach [Yaremchuk and Sentchev, 2011]. In this diversity, the choice of the best-suited method for mapping HFR sea surface currents will ultimately depend on the intended application.

Other limitations and operational constraints

The algorithms used to retrieve the geophysical quantities of interest have different requirements in terms of signal-to-noise ratios. As a rule of thumb, it can be said that current retrieval is the most robust measurement, followed by wind direction and integral measures of sea state such as significant height, with full surface elevation spectrum inversion being the most demanding in terms of measurements quality. Surface-wave signal decrease with range at HF is dominated by frequency-independent geometric decrease of amplitude and frequency-increasing ohmic attenuation, related to surface conductivity [Gurgel et al., 1999]. Low frequency systems thus tend to have longer usable measurement ranges than higher-frequency ones. High sea states tend to disperse electro-magnetic waves and reduce working range.

In contrast with microwave radars, for which thermal noise at the receiver front-end is known to be the dominant issue, noise at HF is dominated by anthropic sources of radio interferences. The situation varies with operating frequency. While local sources dominate in the 20-45 MHz frequency range, successive reflections on the ionosphere and the Earth surface can propagate interferences from sources located thousands of kilometers away in the 5-20 MHz band. The phenomenon of ionospheric reflection is limited to frequencies below the so-called "Maximum Usable Frequency" (MUF) of long-range HF communications. The MUF depends on the composition of the ionosphere. It can vary between 5 MHz and 30 MHz. It is highest near noon or in the early afternoon, and is highest during periods of greatest sunspot activity. HFRs working at frequencies below the MUF are strongly affected by interference with radio transmission from all over the world. They have to operate in a narrow bandwidth in order to realize a sufficient signal-to-noise ratio. The range cell length being inversely proportional to the frequency bandwidth, a trade-off must be found between range resolution and observation range.

Best practices and recommendations

Site maintenance quality is also crucial for a permanent, high-quality data flow. Sources of damage can be natural, such as wind and lightning activity, brush fire or animals, or human such as brushwood clearing, robbery and sacking. Good practices can be formulated to mitigate these risks, as careful electrical grounding and lightning protection, operation on uninterrupted or autonomous power supplies (UPS, solar panels, wind turbines), and also mechanical protection of cables, antennas and under grounding, air conditioning of the RADAR equipment housing. Automatic diagnostics can be devised to alert the operator of faults in the hardware components of the systems. Monthly scheduled visits at the antenna sites to check connections, cables and antennas, measure cables continuity, perform calibration routines, monitor radio interferences and switch external backup disks are however an indispensable component of site maintenance. Finally, the antenna calibration should be monitored on a regular basis, for instance at least yearly for the compact antennas of CODAR Seasonde systems, and at least every two years for phased array HFRs.

FROM LOCAL HFR SYSTEMS TO A FRENCH HF RADAR NETWORK

A summary of existing and planned HFR systems in France and the Basque country is shown in Figure 2 and detailed below.

In the Iroise sea, the Service Hydrographique et Océanographique de la Marine (Shom) has been operating since 2006 a system of two Wellen Radar (WERA), [Gurgel et al., 1999] sites to measure surface currents, wave and wind parameters over a large area off the western coast of Brittany. The area is characterized by intense tidal currents, in particular between the largest islands where it exceeds 3 m.s-1 during mean spring tides [Ardhuin et al, 2009]. This HFR system has two sites in order to measure the surface current: the Garchine site, located in Porspoder, and the Brezellec site, located in Cleden-Cap Sizun. The instrument set-ups at these two sites are identical: a rectangular array of 4 transmit antennas, and a linear equidistant array of 16 receive antennas. This system uses 100 kHz chirps in an allocated 200 kHz frequency band centered on 12.4 MHz, and has 90% current data availability up to a range slightly exceeding 100 km, depending on sea state. The operation and maintenance of these sites is subcontracted to a local company, ACTIMAR. Data dissemination and visualization have originally been performed in the framework of the PREVIMER project, devoted to the demonstation of a french pre-operational coastal ocean forecasting capability. Data are available today through the coastal Coriolis web portal (http://www.coriolis-cotier.org). As of 2017, Shom has plans to set up new HFR systems in Brittany, in the framework of the regional ROEC CPER project.

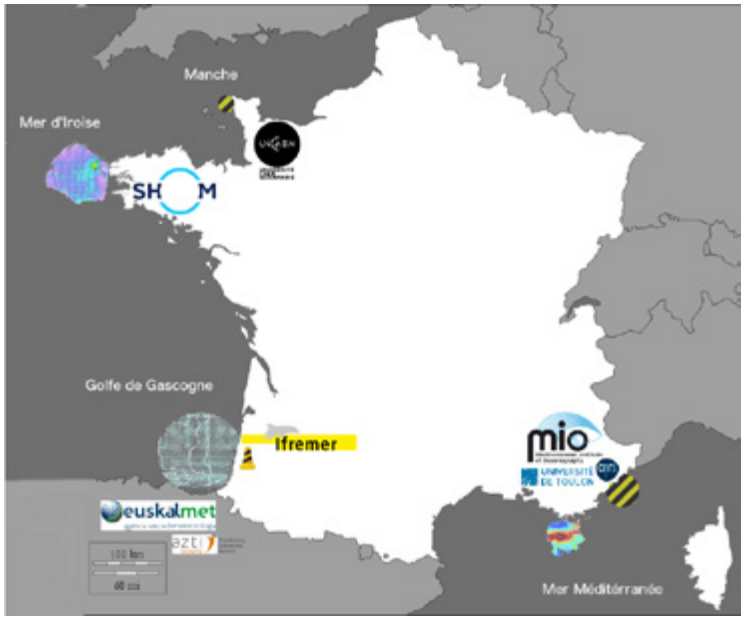


FIGURE 2: EXISTING AND PLANNED HF RADAR SITES ALONG THE FRENCH COASTS 4 AREAS OF INTERESTS

The locations of HFR systems and their actual or expected coverage for the four areas of interests (North sea, Iroise sea, Bay of Biscay, Mediterranean sea) are plotted on the map. In yellow and black stripped circle are the planned sites, in color circle are the existing sites providing real time data flows. The logos of the institutes in charge of the HFR systems or sub-systems (one, two, or three antenna) are associated for each locations.

In the Mediterranean sea, HFRs were installed as part of the Mediterranean Ocean Observing System for the Environment (MOOSE), in two areas of interest, in which instrumented lines with long time series of ocean parameters are also maintained: the Toulon/ANTARES site in the vicinity of the ANTARES infrastructure (Astronomy with a Neutrino Telescope and Abyss environmental RESearch), and the Nice/DYFAMED (DYnamics of Atmospheric Fluxes in the MEDiterranean sea) site.

The Toulon/ANTARES area, located off the coast of Toulon, is a key zone conditioning the behaviour of the Northern Mediterranean Current (NC) just upstream of the Gulf of Lions. It displays significant cross-shelf exchanges correlated to the strong northwesterlies present in the region (Mistral, Tramontane). The first HFR site is located at Cap Sicié and has been operating since 2010. It works in a quasi-monostatic configuration with a non-rectilinear, W shaped, 8-antenna receiving array and a single transmitting antenna. This unusual array configuration was the only solution to cope with the stringent space constraints. This site has been complemented in May 2012 with a fully bistatic second system, a pioneering configuration for WERA. The receiver, a regular linear 8-antenna array, is located at Cap Bénat while the transmitter, GPS-synchronized, is installed in the Porquerolles Island, 17 km away from the receiver, in order to circumvent the presence of several large islands. Theses systems are continuously working in the frequency band of 16.1 to 16.2 MHz and sweeping over a 50 kHz bandwidth, resulting in a 3 km range resolution. To provide radial velocity maps with high azimuthal resolution, a direction finding algorithm is performed on each site. The real-time data produced by this system can be visualized on a dedicated website at http://hfradar.univ-tln.fr/.

The installation of a second HFR system covering the DYFAMED area in the Ligurian sea extended the observation zone to the full coastal area between Toulon and Nice, permitting a much better coverage of the Northern Mediterranean Current. This system is composed of a pair of compact CODAR SeaSonde instruments (13.5 MHz), located at the Cap Ferrat lighthouse in Saint-Jean Cap Ferrat and the Cap Dramont semaphore in Saint-Raphaël, resulting in a 50 km baseline. These systems were set up in October 2013 and May 2014, respectively. The geological environment at the Cap Dramont site makes this instrument particularly prone to lightning damage.

In the southeastern corner of the Bay of Biscay, IFREMER and AZTI are working, in the framework of the JERICO-NEXT H2020 project, to install in 2017 a phased-array HFR site on the Landes coast to complement the two CODAR SeaSonde sites operated by AZTI and EUSKALMET (Basque government) at Cape Higer and Cape Matxitxako since 2009. These two low-frequency (4.86 MHz) sites provide observations of the meridional component of currents at ranges in excess of 200 km north of the spanish coast. At this distance from the baseline, the observation geometry is close to degenerate, the line-of-sights from both radars being essentially meridional. Complementing these two sites with a third located on the Landes coast will enhance the measurement quality over the current observation area, and allow a better retrieval of

the zonal current component in its northern part. Monitoring the along-coast current deflection in the southern part of the Bay of Biscay is an important task, as many activities (fishing, transportation, tourism) are dependent on hydrodynamical conditions in this area. This site will also permit the investigation of a number of technological and methodological questions such as the operation of an integrated cross-border system comprising radars based on different technologies, or the comparison of the observational capacities of the two types of systems.

Tidal currents in the Alderney Race in the English Channel (raz Blanchard in French) are the strongest in Europe and among the strongest in the world, ranging up to six meters per second. A number of industrial companies have planned to install tidal farms inside the Alderney Race to produce electricity from marine currents. The first two pilot tidal farm will be installed in 2018-2019. Oceanographers have been encouraged by companies and by the French ministry of the environment to lead researches on this tidal resource. Projects CPER Manche 2021 (2016-2021, PI: Unicaen) and ANR-FEM HYD2M (2016-2019, PI: M2C) aim to study the impact of wave-current interactions on the total energy. In the framework of these projects, four HFR systems were purchased by University of Caen Normandie. They should be deployed in June 2017 in La Hague to measure waves and surface currents in the Alderney Race during three years. The efficiency of HF radar measurements for tidal stream resource quantification at two other sites of high potential in France (in the Iroise Sea and in the Strait of Dover) was demonstrated recently by Thiébaut and Sentchev (2015, 2016, 2017). At the same time, X-band and VHF radar measurements will be conducted during one year. This study site implies that the HFR will be located 30 km apart and will transmit at 13.5 and 24.5 MHz in order to obtain medium and short range. The two frequencies should give additional useful information on hydrodynamics processes and their vertical shear. Phased-array systems have been chosen to allow more versatility in post-processing. Each HFR system will have 16 active monopole antenna for the receiver and 4 passive antennas for the transmitter. Data analysis will be performed first with the WERA and Seaview softwares for marine currents and sea state, respectively. Then some developments are planned to adapt these softwares to the highly energetic conditions of the Alderney Race. Lastly, HFR data will be included in coastal numerical models (e.g. Telemac or MARS 3D) using classical data assimilation methods.

POTENTIAL FOR COASTAL OCEAN MONITORING AND MODELLING: THE MEDITERRANEAN CASE STUDY

The applications of HFR measurements have spanned the fields of ocean circulation from operational oceanography to methodological development. In the Mediterranean sea, during the first step of the MOOSE-ANTARES implementation (2010-2011), when only one radar site was operational (Cap Sicié), [Marmain et al., 2011] developed and successfully tested the Vortex Identification Method (VIM) to reveal the presence of eddy-like structures through radial current velocity maps. During the PHYOCE experiment (March, 31 – April, 3 2011), the variability of the radial velocity field was used in complement to high resolution modeling (GLAZUR64, e.g. Ourmières et al., 2011), in-situ (ADCP from the NO Tethys and CTD casts) and remotely sensed measurements (SST and Chla from MyOcean) to detect and characterize a mesoscale eddy-meander structure [Guihou et al., 2013]. Two experiments were also performed in the framework of the EU-funded TOSCA (Tracking Oil Spills & Coastal Awareness network) project, gathering a comprehensive data set of surface drifter trajectories and in-situ measurements (including gliders) in the HFR area. These data have been combined into a web-based decision tool designed for authorities in charge of maritime crisis such as Search-and-Rescue and oil spills. Surface drifter trajectories have also been used to validate the HFR velocities, and to develop a new tool for blending both kinds of measurements [Berta et al., 2014]. The potential of this real-time observation lies also in the possibility of assimilating the surface current velocities in numerical ocean models. The MOOSE-ANTARES radial velocity dataset was used in [Marmain et al., 2014] who successfully showed the potential of HFR data assimilation to correct ocean boundary conditions as well as the wind forcing on a sub-regional 1/12° model configuration.

Since May 2012 the MOOSE-ANTARES system is operated in its full configuration, i.e. with two HFR sites (Cap Sicié and Cap Bénat) operating conjointly to allow a full velocity vectors reconstruction over a large area south of the Iles d'Or islands (see Figs. 2-3). The long-term continuity of this data set (3 years and half now with nearly 80% time coverage) allows to foresee new insights on the diurnal to inter-annual variability of the Northern Mediterranean Current. Here we present some examples focusing on model/data comparisons based on the Mercator Océans Atlantic Ocean PSY2V4R4 (henceforth PSY2V4) operational configuration (1/12°, 50 depth layers, ECMWF atmospheric forcing) and the sub-regional NEMO-based GLAZUR64 configuration (1/64° , 130 depth layers, PSY2V4R4 and ARPEGE-Aladin

atmospheric forcing) on the period from 01 May 2012 to 31 December 2013. Figure 3 first illustrates the time averaged surface circulation off the MOOSE-ANTARES area from the three datasets. It clearly shows a better resolution of the NC mean width (~ 20km), orientation and position (42.7-42.9 ° N south of the Porquerolles island) for the high resolution configuration by comparison with the low resolution operational configuration. The PSY2V4R4 benefiting from SSH assimilation, the clear degradation of the simulated NC in the Toulon area likely comes from its lower resolution and its associated crude bathymetry, particularly given the brutal change of the coastline orientation, the reduced shelf and the presence of several canyons along this area. In addition, both models show on average an underestimation of the maximum surface velocity in the NC core (circa 0.45 m/s in HFSRW data vs 0.25-0.23 m/s for the two models), a well known default already discussed in [Ourmières et al., 2011].

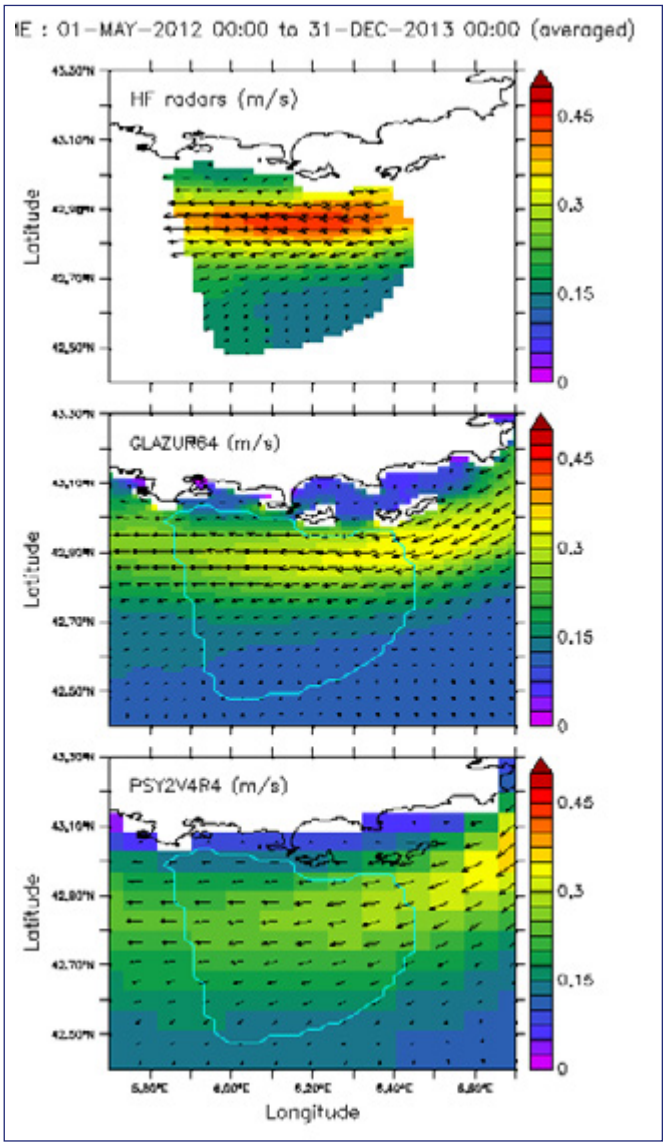


FIGURE 3:

Time averaged (May 2012 to December 2013) surface velocity fields from the HFSRW dataset (high), GLAZUR64 regional configuration (middle) and PSY2V4R4 operational configuration (low) over the Toulon/ANTARES area. The lightblue line on the both models panels shows the main domain covered by the HFSRW system.

Figure 4 shows the time series of zonal velocities south of the Porquerolles Island (6.10° E), as the main component of the NC given its average Westward orientation, for the three datasets. It confirms a better (worst) position and width in the high (low) resolution model, but also shows that the underestimation of the NC is less marked in winter in the GLAZUR64 simulation that can frequently reach velocities similar to the HFSRW measurements. Indeed, both models clearly show an intermittent and unstructured NC for the late spring and summer periods. This structural deficiency likely comes from the parent PSY2V4R4 model and persists in the child model due to a critically weak and unstable NC current during the spring to summer (May to July 2012) or summer to autumn (August to December 2013) transition periods.

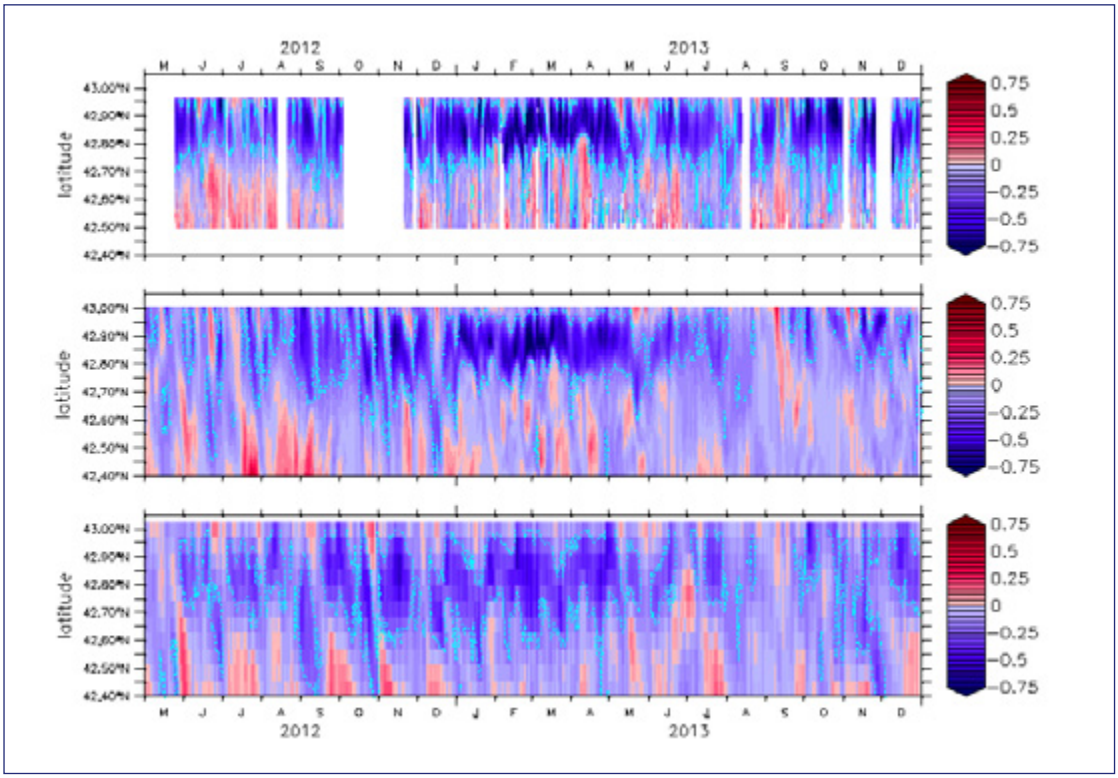


FIGURE 4:

Time series of the zonal velocities across a meridional transect at 6.10° E South of Porquerolles Island from the HFSRW MOOSE-ANTARES system (high), GLAZUR64 regional configuration (middle) and PSY2V4R4 operational configuration (low). The light blue lines on all panels corresponds to the -0.2 m/s isoline that bounds the core of the North Mediterranean Current

By contrast the August 2012 to July 2013 period offers a large variety of meteorological and oceanographic situations with much better model diagnostics. Two of those are illustrated in Figure 5 and Figure 6 for contrasted cases. The first one (Figure 5) shows the response of the surface circulation to a short term strong Mistral event on 31 August 2012, that is better resolved in the GLAZUR64 simulation than in the PSY2VR4 one, due both to the difference in wind forcing (ARPEGE-Aladin vs ECMWF) between the two simulations and a better representation of the NC position and stability in the high resolution GLAZUR64 configuration. The second example (Figure 6) focuses on meso-scale variability of the NC during winter over a 12 days period (February 15-27 2012) and confirms a better representation of the NC width and velocity in winter for the high resolution GLAZUR64 configuration, jointly to more realistic spatial patterns.

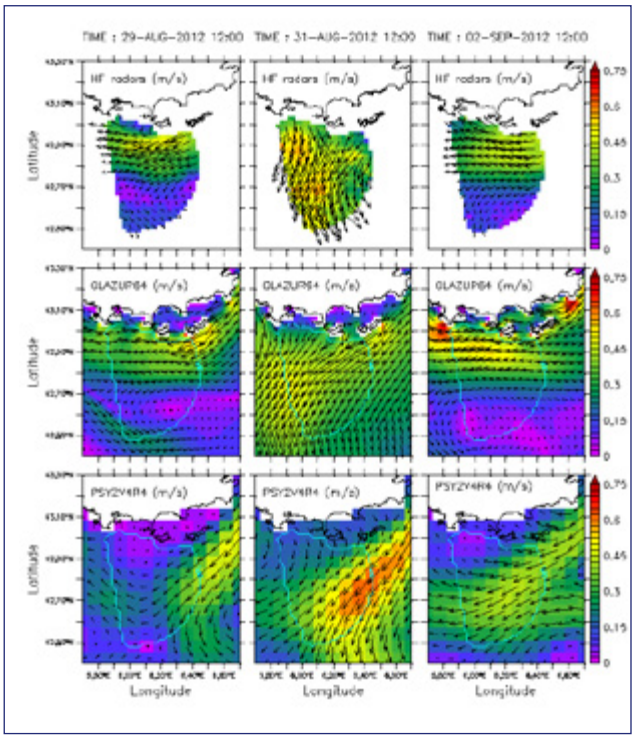


FIGURE 5

Daily averaged surface velocity fields from the HFSRW dataset (high), GLAZUR64 regional configuration (middle) and PSY2V4R4 operational configuration (low) over the Toulon/ANTARES area for August 29 to September 2, 2012. The light blue line on the both models panels shows the main domain covered by the HFSRW system at the same dates

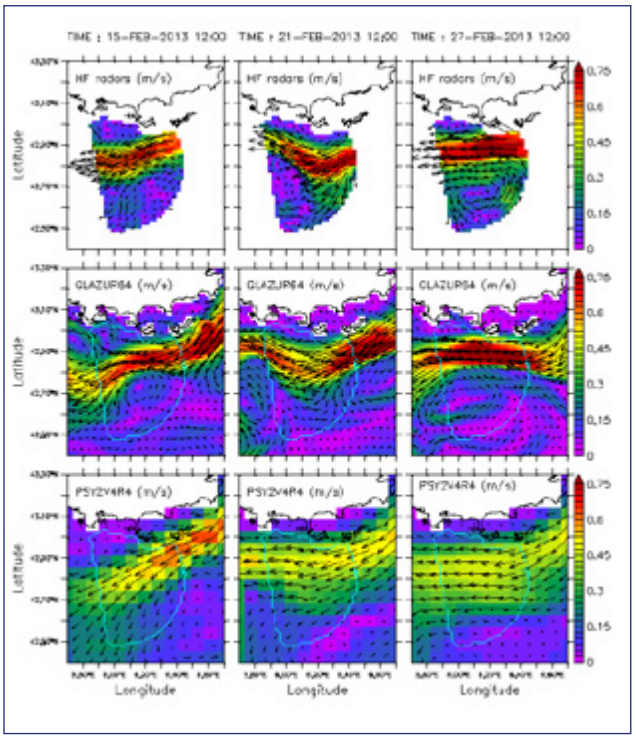


FIGURE 6

Same as Figure 5 but for February 15 to February 27, 2013

PERSPECTIVES AND ONGOING ACTIVITIES

In Europe, many actions and projects are including some tasks on HFR for their ability to monitor surface current in the coastal area. The European context and growing interest for this technology is well described by Rubio et al. (2017). Currently, more than fifty HF radars are operational along European coasts and a number are in the planning stage. EuroGOOS has launched a dedicated HF Radar Task Team (http://eurogoos.eu/high-frequency-radar-task-team/) to promote coordinated activities focussed on the existing HFR systems, and to enhance data sharing and applications. EMODnet Physics, with the support of the HF Radar Task team, has begun to develop a strategy of assembling HFR metadata and data products to make them easily accessible and interoperable. The first HF Radars Pilot Platform, showing in real-time data from 30 HFR sites, is available online at http://www.emodnet-physics.eu/. Development of homogenized quality control/quality assessment procedures and of new value-added products are being addressed in the framework of the Joint European Research Infrastructure network for Coastal Observatory - Novel European eXpertise for coastal observaTories (JERICO-NEXT, H2020, 2015-2019) project (http://www.jerico-ri.eu). Recently the Copernicus Marine Environment Monitoring Service (CMEMS) Service Evolution Call has supported the Innovation and Networking for the Integration of Coastal Radars into European mArine Services (INCREASE) project. Based on the progress of the former initiatives, INCREASE aims to begin the developments necessary for the integration of existing European HF radar operational systems into the CMEMS (http://marine.copernicus.eu) and promote the use of HF radar data for improving CMEMS numerical modelling systems. Finally, the SeaDataCloud project, which aims at advancing SeaDataNet (http://www.seadatanet.org) services and increasing their usage, will provide the further development for a pan-European infrastructure for marine and ocean data management and will contribute to the integration and long term preservation of historical time series from HFR into

the SeaDataNet infrastructure. The European efforts are aligned with initiatives at international levels, where the Group of Earth Observations is working to promote data sharing and delivery and the proliferation of HFR applications, toward the set up of a Global HFR radar network (https://rucool.marine.rutgers.edu/geohfr/).

In this context, the aim of the LEFE/GMMC ReNHFOR project is to structure and coordinate HFR-related efforts at the French level, to consolidate existing software packages devoted to a variety of processing tasks and to coordinate the development of new common tools, to support interested researchers in their effort to plan, deploy and exploit new HFR systems, and to act as an interface between the French and European research and operational HFR community. The ReNHFOR group has performed several actions in 2016, mainly two plenary meetings devoted first to a discussion of the general aspects of HFR operation and processing and secondly on high-resolution angular processing. Those also allowed to report to the members on the status of ongoing activities at the European level. As of early 2017 no data from the French coasts are available in European databases, though systems are operating in the Celtic and Mediterranean seas. Work meetings were held to solve practical issues related to the transmission of radial current data from the Mediterranean systems to IFREMER/SISMER and their archival. The follow-up issue of the transmission of the data by SISMER to the EMODNET-Physics system for aggregation/visualization/diffusion, will be addressed in 2017, when EMODNET-Physics recommendations for the meta-data fields and data files organization will converge. Among the issues likely to be discussed in 2017 meetings are data interpolation routines used to fill Direction-Finding current maps, the launch of an action about HFR assisted operational model evaluation (MENOR, PSY2V4R4...) and the various issues that will arise in the course of the Alderney Race site installation.

REFERENCES

Ardhuin, F., L. Marie, N. Rascle, P. Forget and A. Roland, "Observation and estimation of Lagrangian, Stokes and Eulerian currents induced by wind and waves at the sea surface". J. Phys. Oceanogr., 39, 2820-2838, (2009).

Barrick, D.E., "Remote Sensing of Sea State by Radar", in "Remote Sensing of the Troposphere", V.E. Derr, Editor, NOAA/Environmental Research Laboratories, (1972).

Barrick, D.E., and B.L. Weber. «On the Nonlinear Theory for Gravity Waves on the Ocean's Surface, Part II: Interpretation and Applications.» J. Phys. Oceanogr., 7, 11-21 (1977).

Barrick, D. E. and M. W. Evans, "CODAR: A coastal HF radar for real-time current mapping", U. S. Patent 4 172 255, (1979).

Barth A., Alvera-Azcarate A., Weisberg R.H., "Assimilation of high-frequency radar currents in a nested model of the West Florida Shelf", J. Geophys. Res. Oceans, 113, C08033, (2008).

Bellomo, L., et al., "Toward an integrated HF radar network in the Mediterranean Sea to improve search and rescue and oil spill response: the TOSCA project experience", J. Oper. Oceanogr., 8, 95-107, (2015).

Berta, M., et al., "Estimating Lagrangian transport blending drifters with HF radar data and models: results from the TOSCA experiment in the Ligurian Current (North Western Mediterranean Sea)", Prog. Oceanogr., 128, 15-29, (2014).

Broche, P., J.-C. De Maistre, P. Forget, "Mesure par radar décamétrique cohérent des courants superficiels engendrés par le vent", Oceanol. Acta, 6, 43-53, (1983).

Crombie, D. D., "Doppler spectrum of sea echo at 13.56 Mc./s.", Nature, 175, 681-682, (1955).

Davis, R. E., "Objective mapping by least squares fitting", J. Geophys. Res. Oceans, 90, 4773 – 4778, (1985).

Guihou, K., J. Marmain, Y. Ourmières, A. Molcard, B. Zakardjian and P. Forget, "A case study of the meso-scale dynamics in the North-Western Mediterranean Sea: a combined data-model approach", Ocean Dyn., 63, 793-808, (2013).

Gurgel, K-W., G. Antonischki, H.-H. Essen and T. Schlick, "Wellen Radar (WERA), a new ground-wave based HF radar for ocean remote sensing", Coast. Eng., 37, 219–234, (1999).

Gurgel, K-W., H.-H. Essen and S. P. Kingsley, "High frequency radars: physical limitations and recent developments", Coast. Eng., 37, 201-218, (1999).

Harlan, J. A. and T. M. Georges, "An empirical relationship between ocean surface wind direction and the Bragg line ratio of sea-echo spectra", J. Geophys. Res. Oceans, 99, 7971-7978, (1994).

Hasselman, K., "Determination of ocean-wave spectra from doppler radio return from the sea surface", Nature, 229, 16-17, (1971).

Ivonin, D., P. Broche, J.-L. Dévenon and V. I. Shrira, "Validation of HF radar probing of the vertical shear of surface currents by acoustic Doppler current profiler measurements", J. Geophys. Res. Oceans, 109, C04003, (2004).

Kaplan, D. M. and F. Lekien, "Spatial interpolation and filtering of surface current data based on open-boundary modal analysis", J. Geophys. Res. Oceans, 112, C12007, (2007).

Kim, S. Y., E. Terrill and B. Cornuelle, "Mapping surface currents from HF radar radial velocity measurements using optimal interpolation". J. Geophys. Res. Oceans, 113, C10023, (2008).

Kirby, J. and T. Chen, "Surface waves on vertically sheared flows: approximate dispersion

relations", J. Geophys. Res. Oceans, 94, 1013-1027, (1989).

Lewis, J. K., I. Shulman and A. F. Blumberg, "Assimilation of CODAR observations into ocean models", Cont. Shelf Res., 18, 541– 559, (1998).

Lipa, B., B. Nyden, D. Ullman and E. Terrill, "Seasonde radial velocities: derivation and internal consistency", IEEE J. Ocean. Eng., 31, 850-861, (2006).

Marmain, J., P. Forget and A. Molcard, "Characterization of ocean surface current properties from single-site HF/VHF radar", Ocean Dyn., 61, 1967-1979, (2011).

Marmain, J., A. Molcard, P. Forget, A. Barth and Y. Ourmières, "Assimilation of HF radar surface currents to optimize forcing in the northwestern Mediterranean Sea", Nonlin. Processes Geophys., 21, 659-675, (2014).

Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, J. A. Austin, J. A. Barth, T. J. Boyd, P. M. Kosro, and M. D. Levine, "A modeling study of the three-dimensional continental shelf circulation off Oregon. Part 1: Model-data comparisons", J. Phys. Oceanogr., 32, 1360–1382, (2002)

Ourmières, Y., B. Zakardjian, K. Béranger and C. Langlais, "Assessment of a NEMO-based downscaling experiment for the North-Western Mediterranean region: impacts on the Northern Current and comparison with ADCP data and altimetry products", Ocean Model., 39, 386-404, (2011).

Paduan, J. D., and I. Shulman, "HF radar data assimilation in the Monterey Bay area", J. Geophys. Res. Oceans, 109, C07S09, (2004)

Rubio A., et al., "HF Radar Activity in European Coastal Seas: Next Steps Towards a Pan-European HF Radar Network", Front. Mar. Sci., 4. doi: 10.3389/fmars.2017.00008, (2017).

Schmidt, R. O., "Multiple emitter location and signal parameter estimation", IEEE Trans. Antennas Propag., 34, 276-280, (1986).

Sentchev A., M. Yaremuck, F. Lyard, "Residual circulation in the English Channel as a dynamically consistent synthesis of shore-based observations of sea level and curents", Cont. Shelf Res., 26, 1884-1904, (2006).

Sentchev, A., P. Forget, Y. Barbin and M. Yaremchuk, "Surface circulation in the Iroise Sea (W. Brittany) from high resolution HF radar mapping", J. Mar. Sys., 109-110, S153-S168, (2013).

Stewart, R. H. and J. W. Joy, "HF radio measurements of ocean surface currents.", Deep Sea Res., 21, 1039-1049 (1974).

Thiebaut M. and A. Sentchev, "Asymmetry of tidal currents off the W.Brittany coast and assessment of tidal energy resource around the Ushant Island.", Renew. Energ., 105, 735-747, (2017).

Thiebaut, M. and A. Sentchev, "Tidal stream resource assessment in the Dover Strait (eastern English Channel).", Int. J. Mar. Ener., 16, 262-278, (2016).

Thiébaut, M. and A. Sentchev, "Estimation of tidal stream potential in the Iroise Sea from velocity observations by High Frequency radars.", Energ. Proc., 76, 17-26, (2015).

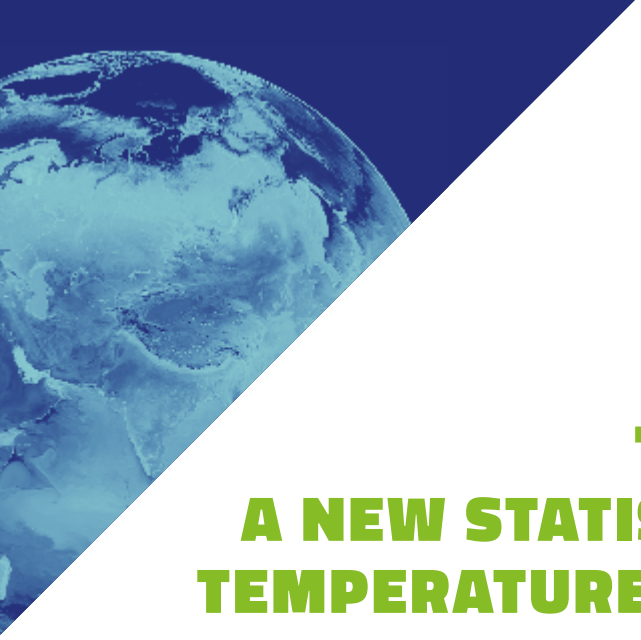
Weber, B.L., and D.E. Barrick. «On the Nonlinear Theory for Gravity Waves on the Ocean's Surface, Part I: Derivations.» J. Phys. Oceanogr., 7, 3-10 (1977).

Wilkin J. L., Hernan A.G., Haidvogel D. B., Lichtenwalner S. C., Glenn S. M. and K. S. Hedstro. "A regional ocean modeling system for the Long-term Ecosystem Observatory", J. Geophys. Res. Oceans, 110, C06S91, (2005).

Wyatt, L. R. et al., «Validation and intercomparisons of wave measurements and models during the EuroROSE experiments.», Coast. Eng., 48, 1-28 (2003).

Yaremchuk, M. and A. Sentchev, "Mapping radar-derived sea surface currents with a variational method", Cont. Shelf Res., 29, 1711-1722, (2009).

Yaremchuk, M. and A. Sentchev, "A combined EOF/variational approach for mapping radar-derived sea surface currents", Cont. Shelf Res., 31, 758-768, (2011).



A NEW STATISTICAL APPROACH FOR TEMPERATURE AND SALINITY QUALITY CONTROL BASED ON DIRECT INFERENCE OF LOCAL EXTREME VALUES

BY
J. GOURRION[1], T. SZEKELY [2], G. REVERDIN [3]

INTRODUCTION

Development of monitoring and prediction ability for the climate evolution at short and longer term has been, is, and will be for years the main challenge for the Earth Sciences community. Since a couple of decades, ocean monitoring is based on assimilation systems for which the performance depends crucially on the in-situ data quality control (QC).

Traditionally, different types of errors have been identified [Gandin, 1988], e.g. random, systematic or gross errors. For years, manual QC has been applied to detect them, but the increasing data flow makes it excessively time-consuming; the necessity of automatic, computerized QC procedures has become more and more obvious. Here, the attention is focused on systematic and gross errors which may have a dramatic impact on the model analysis. Also, for most present oceanographic observation devices, the random errors are commonly a few orders of magnitude smaller.

Basic QC procedures usually consist in checking for inconsistencies (in the date, location, displacement, platform identification, value, decoding errors, stuck values, jumps in time series, spikes). Possible horizontal inconsistencies are often addressed through comparison to local statistics from an historical reference dataset, checking that a given value is located inside some validity interval:

$$X_{min} \leq X \leq X_{max}$$

Where X stands for the relevant variable and [Xmin Xmax] defines the range of valid values. A common practice

[1] **R & D Coriolis, Ifremer, centre de Brest,** CS 10070, 29280 Plouzané, France.

[2] **R & D Coriolis, IUEM/CNRS,** Plouzané, France.
[3] **LOCEAN, CNRS,** Paris, France

consists in defining the validity interval from climatological mean and standard deviation in the neighborhood of the observation location:

$$X_{mean} - N \cdot X_{std} \leq X \leq X_{mean} + N \cdot X_{std}$$

Where N is an adjustable parameter to be defined (usually in the range 3-6), see [Gandin 1988, Boyer 1994, Carton 2000, Ingleby 2007, Cabanes 2013]. Left and right sides of inequality [Eq 2] are statistical estimators of the corresponding terms in Eq.1.

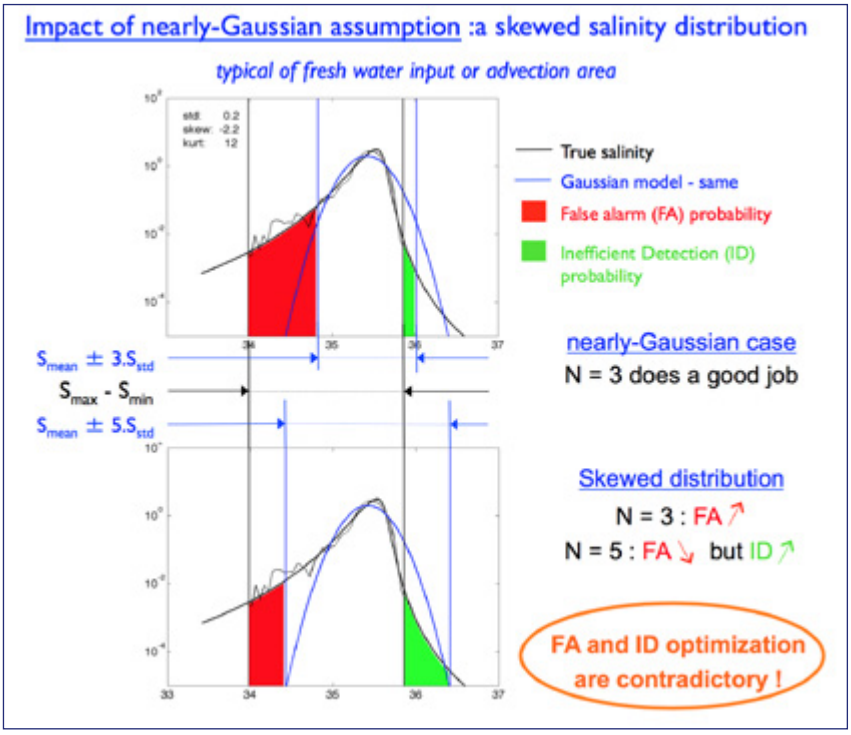


FIGURE 1

Scheme describing the impact of an asymmetric distribution on the detection efficiency..

Despite its wide use in QC procedures for hydrographical datasets, the approach described by Eq.2 implicitly makes strong statistical assumptions that lead detection inaccuracies. Indeed, when dealing with error detection, two main objectives are pursued: maximize the number of bad data detections and minimize the number of erroneous ones (or false alarms). The resulting strategy is usually the result of a particular trade-off between both objectives [Ingleby, 2007]. Figure (1) presents a schematic picture to explain this point. Clearly, under asymmetrical distribution conditions, it is impossible to simultaneously minimize the number of false alarms and the amount of errors passing through the QC procedure.

Thus, it is now clear that the classical approach does not allow accounting for asymmetry or skewness S in the local data distribution. Further, there is no a priori reason for not assigning a local value to N in Eq.2. Whenever N is defined homogeneously, a constant kurtosis K value is assumed, as long as it is claimed that the detection procedure has a homogeneous statistical significance or performance level (a further objective that any approach should pursue). Consequently, in addition to its inability to account for potential distribution asymmetry, the classical approach intrinsically lacks of flexibility in accounting for the detailed shape of the distribution in terms of peak enhancement (or flatness) and relative height of its tails. This approach is model-based, assuming that the data distribution is 1) unimodal, 2) symmetric and 3) has a constant kurtosis (as long as the N parameter is homogeneously set). It does not allow optimizing simultaneously the number of «good» and «bad» detections. Finally, when applying the method and once chosen a specific N value, extreme events may be detected as erroneous observations if they are too far from the mean, i.e. at a distance larger than N*std, even if they were observed and positively assessed in the historical dataset used to build the reference mean and standard deviation fields.

Consequently, in this paper, it is proposed to use different statistical estimators to define the boundaries of the validity interval from Eq. 1. The validity interval bounds will be directly inferred from the minimum and maximum values

present in the reference historical dataset.

The present paper is organized as follows. Section 2 presents the strategy set up to derive the Min/Max reference fields and the historical datasets used as reference. Samples of the resulting fields will be described, focusing on the innovation relative to the classical approach and their statistical robustness. In section 3, examples of improved detection using the present approach are presented. In section 4, validation of the detection method run on an independent dataset is presented. Concluding remarks come in section 5

BUILDING THE MINIMUM AND MAXIMUM REFERENCE FIELDS

Strategy for local extrema estimation

Choosing the extrema ever seen in an historical dataset is an appealing idea when defining an efficient validity interval for a given parameter. Indeed, it guarantees that an extreme event, even observed only once while judged to be realistic, will not be discarded once observed again, whatever its occurrence in the immediate neighborhood. Nevertheless, in the real world, geophysical extrema estimation turns out to be a non-trivial challenge when facing the case of in-situ observations subject to errors with various origins and magnitudes. Minimum and maximum values are extremely sensitive to measurement errors, so that an adequate strategy must be set up. The chosen procedure is manual, iterative and based on the spatial consistency of min/max maps. First, all data flags are disabled and location-independent quality checks are applied. Once data binned accordingly to longitude, latitude and pressure values, a preliminary version of min/max fields is computed, storing the information necessary to identify the measurement responsible of it. Iteratively, the field spatial discontinuities are detected, the profiles responsible of it are visually controlled, flags are potentially modified and the Min/Max fields updated up to the point where all discontinuities are judged to be realistic.

In order to reduce spatial inhomogeneity in the statistical robustness of the Min/Max estimates, reference fields are estimated onto an unstructured grid with hexagonal constant area cells. And, to increase the overall statistical robustness, a kind of spatial filtering is applied: for each cell, the min (max) values are replaced by the minimum (maximum) of the 7 values taken as those from the current cell and its 6 neighbors. This corresponds to a spatial resolution loss allowing an average increase of the number of samples by a factor 7.

Data sets

To build the reference T, S, σ minima/maxima maps, we use the full ARGO dataset. Initially, the dataset is downloaded from the global ARGO data center in early April 2014. As mentioned in the validation section, a second dataset is downloaded in September 2015 to test the impact of updating the reference fields.

Apart from ARGO, a set of CTDs from SISMER, ICES and WOD databases is included, as well as delayed-mode data from a network of CTDs mounted on sea mammals (MEOP). The original flags are not considered (even if most of these data are supposed to be already quality-controlled) and all profiles are checked using the criteria described in section 2.

Salinity Statistical parameters

Once the datasets iteratively controlled as described above, local temperature, salinity and potential density distributions are assembled over all oceanic grid cells and 20 dbar thick layers. From these distributions, Min/Max values can be derived as well as standard statistical moments.

In this section, we provide insight on the minimum/maximum parameters that are briefly described. For the sake of simplicity, results are only presented for the salinity parameter.

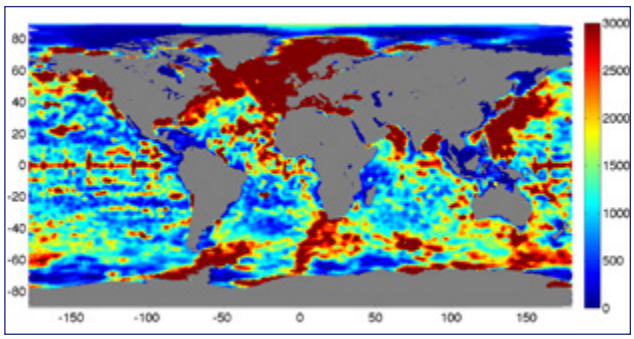


FIGURE 2

Number of samples per grid cell in the 0-20 db layer.

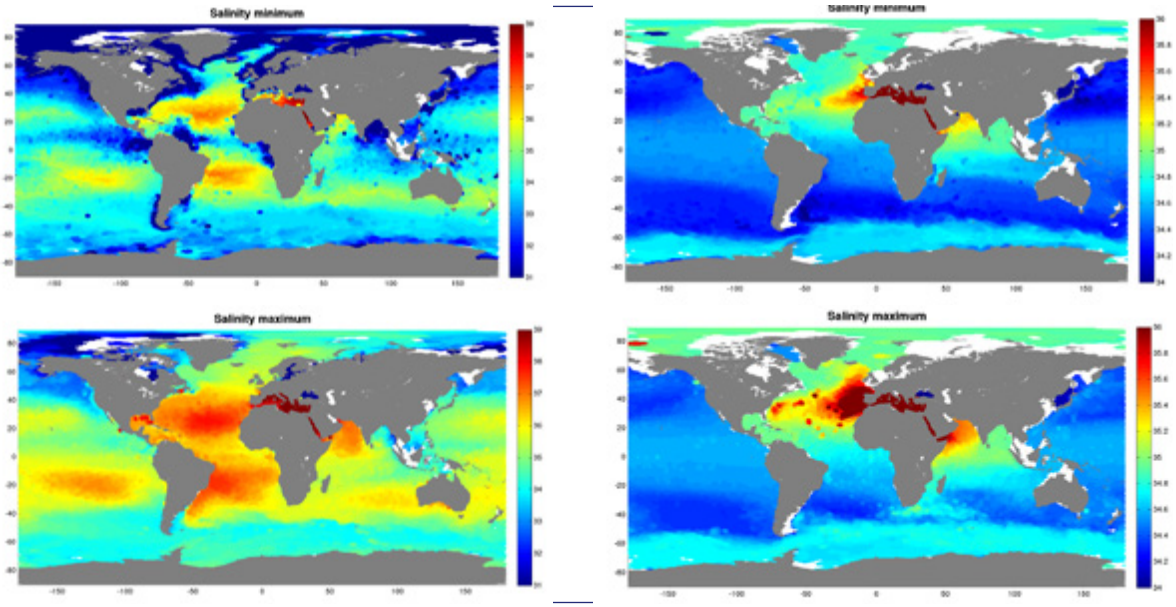


FIGURE 3

Upper (lower) panels: minimum (maximum) salinity value. Left (right) panels refer to the 0-20 dbar (500-520 dbar) layer.

Number of samples.

In figure 2, the number of samples in the surface layer is shown. Due to the ARGO array sampling, the spatial coverage is globally rather uniform, with exceptions of the more sampled Kuroshio and Gulf Stream areas and the less sampled south Atlantic and southern Ocean. The ICES dataset specifically contributes to increase density in the North Atlantic, north of 50oN. The MEOP dataset does increase the overall low sampling of the Southern Ocean. The OCL dataset contribution is particularly obvious through the zonal and meridional high density lines noticeable in the central equatorial Pacific.

Statistical moments are estimated locally and independently of time. Such parameters are used qualitatively in the following, but are not specifically presented here.

Minima and maxima

Top and bottom panels of figure 3, left (right) column, display minimum and maximum salinity fields for the surface (500 db) layer. Such fields are an alternative tool for descriptive physical oceanography that emphasizes the difference between permanent and transient features in the Ocean, with the permanent features having signature both in minimum and maximum fields, and the transient ones appearing only in one or the other. In the surface layer, the typical large scale structure consisting of salinity maxima near the tropics (associated to desertic at similar latitude over land) appears, as a permanent structure, in both minimum and maximum fields. In the 500 db layer, the presence of outflows from evaporation basins (Mediterranean Sea, Red Sea, and Persian Gulf) is a striking feature in both field types. As an example, the surface minimum salinity field displays signatures of seasonal (in an Eulerian sense) fresh water inputs such as rain in the Pacific Inter-Tropical Convergence Zone (ITCZ), or runoffs from the Amazon, Niger, Congo or Gange rivers. Other examples come from zones characteristic of mesoscale eddy activity; the Gulf stream front is clearly displaced northward when shifting from the minimum to the maximum fields; the northern westward return branch of the Southern Hemisphere supergyre appears in the deep maximum field with high salinities near 40˚S in the Atlantic (eddies generated in the Agulhas retroflection area) or oriented north-westward from the west-southern tip of Australia (water coming from the Tasman leakage, see [Rosell-Fieschi et al., 2013]).

Interval Center Shift and Interval Width Ratio

Using the Min/Max fields presented in section 2.2.2, validity intervals can now be defined for QC purposes. In this section, such intervals are presented and compared to the classical ones based on mean and standard deviation estimates for the 0-20 db layer (Fig. 4).

Left panels in figure 4 display the amplitude of the validity interval as computed from the difference Max - Min (top) or 2 N times the standard deviation (bottom), as defined in

section 1 for the classical approach. When setting N to 3, both estimates are very similar at first order both in terms of spatial structure and amplitude. Nevertheless, relative differences in the validity intervals seem to occur at second order. In the following, they are described in terms of specific statistical parameters characterizing the distribution shape: normalized interval center shift (NICS, asymmetry coefficient) and interval width ratio (IWR, tails height relative to characteristic width).

$$NICS = \frac{(S_{min}+S_{max})/2 - Smean}{2*S_{std}}$$

$$IWR = \frac{S_{max} - S_{min}}{2*S_{std}} = N_{eff}$$

The top panels in figure 4 display surface NICS (middle) and IWR (right) as estimated from the quality controlled data set obtained previously. They illustrate all the variability in the distribution shape that the present approach allows to account for, as an improvement of the classical one.

In the surface inter-tropical Pacific and Atlantic oceans, the signatures of precipitations and runoff identified in the

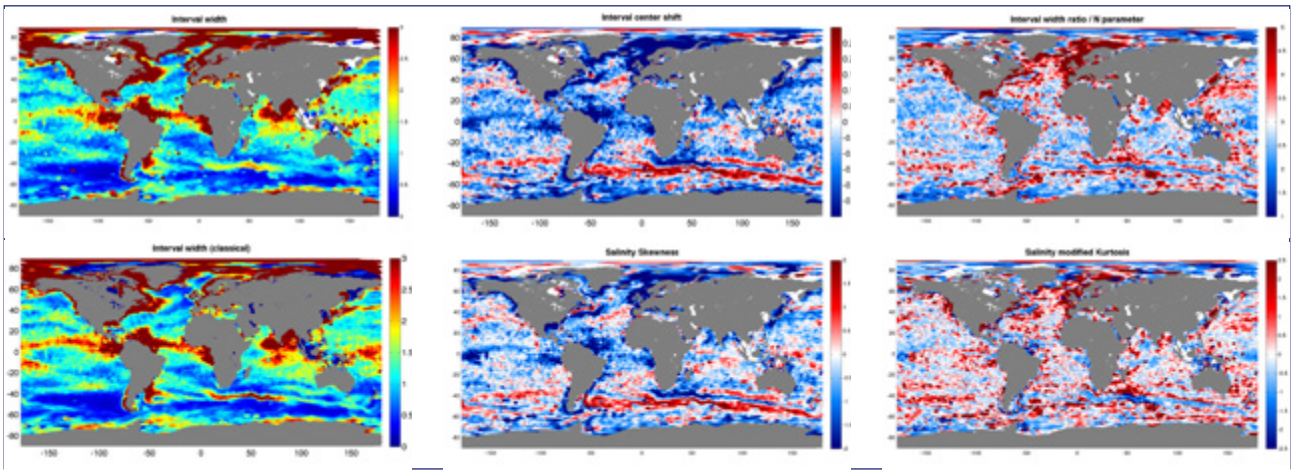


FIGURE 4

All plots refer to the 0-20 dbar layer. Upper panels: variability amplitude as estimated from (left) and (right). Middle left: normalized interval center shift as defined in Eq (3). Bottom right: skewness. Middle right: interval width ratio as defined in Eq. (4). Bottom right: kurtosis corrected from skewness.

minimum field (see figure 3) are clearly present in the NICS field: relative to the mean value, the validity interval is shifted towards negative values indicating that large fresh anomalies are expected to have larger occurrence than salty ones. In the Southern Ocean, especially east of Greenwich meridian, a striking feature, visible at crossing the subtropical front, is associated to a bimodal distribution with the presence of warm and salty South Indian Central Water (SICW) and fresher and colder SubAntarctic Surface Water (SASW).

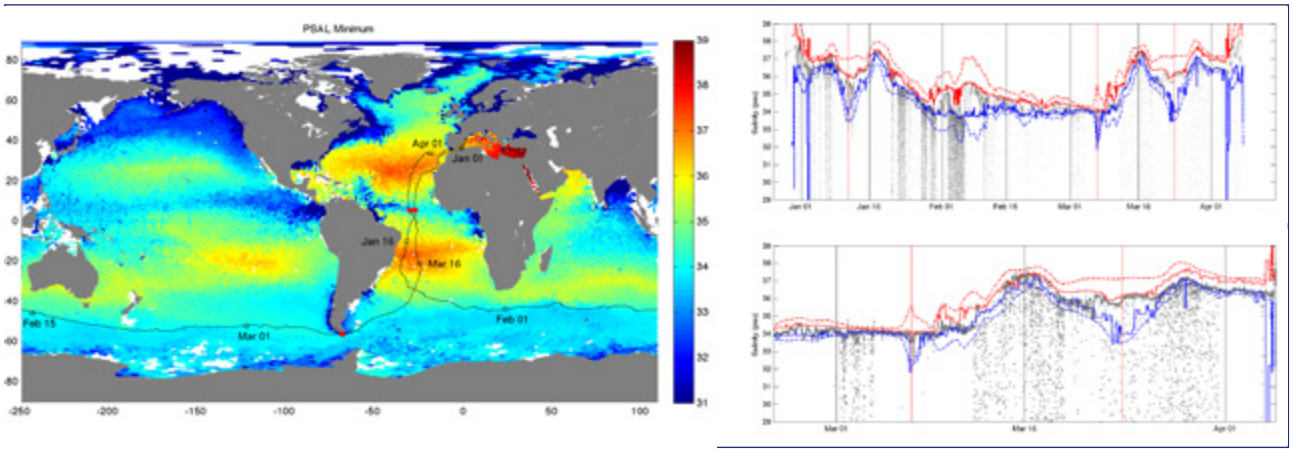


FIGURE 5

Left: Sailing ship route during the BWOR. The background color corresponds to the surface estimate. Top right: entire salinity time series during the cruise; blue (red) lines indicate the lower (higher) bound of the validity interval, with the full lines corresponding to the Min/Max approach and the dashed ones to the classical one. Vertical black lines help to locate the beginning and middle of each month; vertical red lines correspond to the events highlighted in the text. Bottom right: same as top right with a zoom onto the 2015 March-April period.

IWR field shown in the middle right panel is the ratio between the actual validity interval and twice the standard deviation, i.e. the effective N value (number of standard deviations) that the classical approach should apply to avoid misestimates of the validity interval width. Global estimates of this ratio provide an average value of 3 approximately. Systematic use of values larger than 3 leads to an overestimation of validity interval width, inducing an overall reduction in quality control efficiency through undetectable erroneous data. Focusing again on the Southern Ocean east of Greenwich meridian, a symmetric structure is observed in the cross-front direction, that can be interpreted in terms of shape departure from the Gaussian one (for which IWR = 3). At the center, values lower than 3 are observed, indicating an excessive standard deviation value compared to the tail height (sometimes referred as «heavy shoulders»), characteristic of flattened or even bimodal distribution shapes i.e. a combination of processes with similar variance but varying mean. On the front sides, IWR values larger than 3 are observed, corresponding to relatively higher distribution tails; this indicates a combination of processes with similar mean but varying variance, probably associated to variable front activity.

Similarity with S and K

Following, bottom panels of figure 4 display fields of skewness (middle) and kurtosis adjusted for skewness (footnote: and are intrinsically correlated. Here, in order to separate Skewness-induced kurtosis from other sources, Kurtosis adjusted from skewness is defined on the basis of a Pearson diagram as) (right). A remarkable similarity between NICS and skewness patterns on one side and those of IWR and kurtosis adjusted for skewness on the other side, is observed. Actually, such similarity is not a surprise when realizing that NICS and IWR definitions are analogous to robust quantile-based estimates of skewness: obviously, the skewness being a well-known indicator of the distribution asymmetry, it has to be closely related to the distribution center displacement; and the kurtosis, indicator of the distribution tail weight, has to talk with the Min/Max distance. The reader may check that the above interpretation of the NICS and IWR features in terms of distribution shape is even more obvious when the analogy with skewness and kurtosis is accounted for.

Thus, the benefit of the present approach for QC purposes is essentially related to its ability to account for high order statistical characteristics of the data distribution, where the classical approach can only account properly for second order characteristics.

It is also relevant to stress that the remarkable similarity evidenced in figure 4 is a good preliminary robustness indicator of the Min/Max and moments estimates.

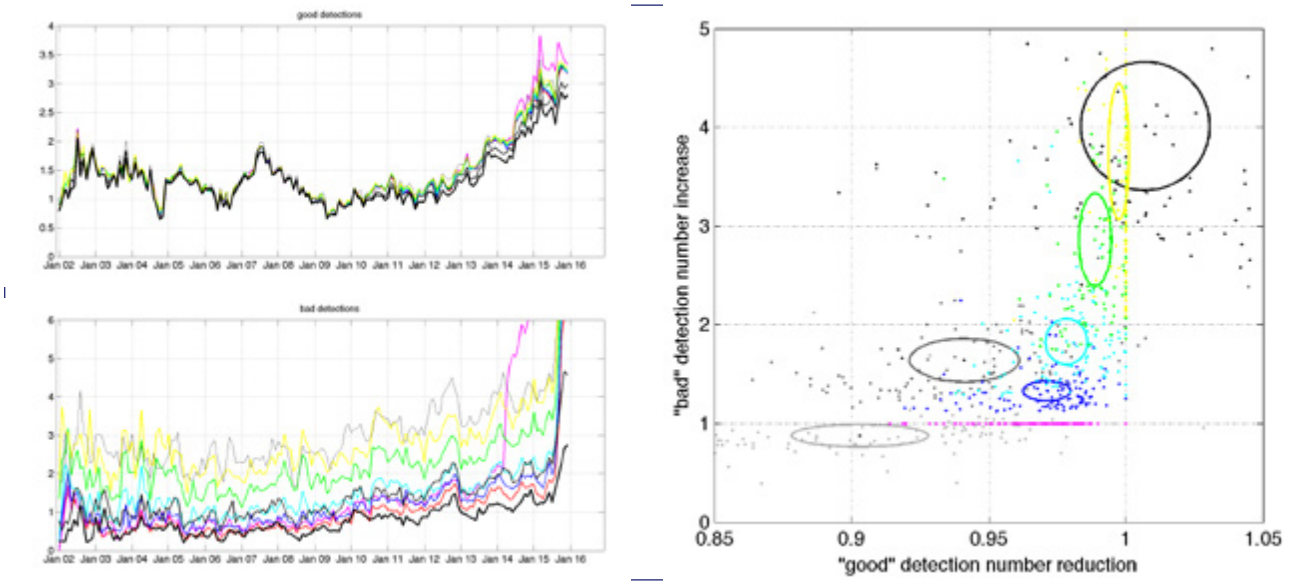


FIGURE 6

Left: monthly percentage of «good» (top) and «bad» (bottom) detections for a set of approaches and configuration: black/grey lines refer to the classical approach for N=4 (light grey), N=5 (grey) and N=6 (black); color curves refer to the Min/Max approach with different P values: 0 (red), 0.015 (blue), 0.025 (cyan), 0.04 (green), 0.05 (yellow). Right: comparative synthesis in terms of «good» detection number reduction and «bad» detection number increase relative to optimal conditions.

RESULTS

In this section, the reference Min/Max fields estimated and described in section 2 are used to illustrate their potential in characterizing the geophysical variability to be expected at the location of a given observation as deduced from available historical information.

CTD onboard a round-the-world cruise

Raw data from a CTD sensor mounted in the bilge of a sailing boat during the Barcelona World Ocean Race are now used to illustrate the potential of the Min/Max estimates for QC purpose. Figure 5 shows the trajectory during the trip, the entire salinity time series as well as a zoom onto a time period of particular interest here. Under high hull speed conditions, the boat rises above the water and air enters the sensor, resulting in low conductivity and salinity anomalies. Such anomalies need to be filtered out from the time series.

Thus, in the right panels of figure 5, the observed surface salinity time series are shown together with the local Min/Max validity interval, as well as the one derived from the classical approach estimated with N = 5 (i.e. the value used at the Coriolis data center for delayed-time quality control). From early January through to late February, the Min/Max interval is systematically thinner than its classical equivalent, suggesting that Min/Max approach has, in general, a more restrictive error detection capability; anomalous measurements have a larger probability to be detected, relatively to the classical approach i.e. the approach allows a larger number of «good» detections. Of course, such a larger relative capability would be reduced using a lower N value in the classical approach; but the price would be an increased number of «bad» detections.

In order to further inter-compare the two approaches, a few anomalies are especially relevant, e.g. with the boat crossing the ITCZ on January 11th and March 24th or sailing through a fresh water pool by the southern tip of South America on March 7th. Let's focus our attention on March 7th and March 24th events. In the first case, the low salinity anomaly encountered near Cape Horn impacts similarly the lower bound of the validity interval in both approaches. Nevertheless, for the upper bound, while the maximum value is not sensible to that fresher water, the classical upper bound is symmetrically shifted up. This is due to the fact that the method assumes a symmetrical data distribution while the mean value is not affected by the fresh anomaly, which likely has a low occurrence in the reference dataset. The consequence is that the local ability of detecting measurement errors associated to positive anomalies up to 1 psu is severely degraded. In the

second case, a similar analysis can be conducted, except that the classical upper bound of the validity interval is relatively less impacted i.e. shifted towards higher values; the fresh anomaly is partly carried by the standard deviation but also by the mean value; in this case, near 5°N, the occurrence of such a fresh event is higher as the ITCZ crosses the area twice a year, during its northward and southward migration respectively in May and November. As intuited in section 1, the performance of the classical approach depends on the occurrence of extreme events in the reference dataset, while the Min/Max approach is independent of such occurrence as long as it is larger than zero.

VALIDATION

In order to quantify the relative efficiency of the classical and Min/Max approaches at the global scale, the ARGO dataset over the 2002-2015 period is now used for validation purpose. To avoid incestuous conclusions, the dataset is modified through addition of a small perturbation to the T/S profiles: a random noise with a variance equal to some P percent of the local variability. Both QC methods are applied to the modified ARGO dataset; «good» and «bad» detections are estimated using the CORA 4.2 flags as truth, assumed to be perfect.

Results are presented in figure 6 on the basis of monthly statistics. In the left panels, color curves correspond to the Min/Max approach for different P values while black curves to the classical approach for different N values. From the upper panel, it is clear that good detections statistics are very little sensitive to the QC approach, except for the classical approach with N >= 5 where the number of good detections may decrease by up to 10 %. Contrastingly, bad detection statistics are much more sensitive to the methodology.

In the right panel of figure 6, for each approach and parameters N or P, the information from the left panel figures is synthesized and interpreted in terms of modification of good and bad detections relative to reference numbers. Under this convention, the perfect approach stands has coordinates (1,1), i.e. when good and bad detection numbers correspond to the reference ones. For similar bad detection statistics, the Min/Max approach systematically miss three times less good than the classical approach does.

As a summary, the Min/Max approach offers the best trade-off between maximal number of good detections and minimal number of bad ones.

CONCLUSION

This work has focused on the statistical inference of the validity interval values used in automatic QC procedures and specifically in the comparison to an historical knowledge of the local variability.

The success is attributed to the increased accuracy of the Min/Max statistical estimates in accounting for already observed extreme events when defining a reference validity interval.

The classical approach estimates the validity interval from the statistical moments of order 1 and 2. Here, it has been proposed to estimate directly the validity interval from the local minimum and maximum values as observed in an extensive historical dataset with global coverage and dedicated manual QC.

It has been demonstrated that, for a similar number of good detections, the new approach allows an important reduction of the number of «bad» detections. If used as an alert-raising tool combined with human QC, the number of «bad» detections can be seen as useless human time so that its reduction leads to a significant saving of human resources. The success is attributed to the increased accuracy of the Min/Max statistical estimates in accounting for already observed extreme events when defining a reference validity interval. Such an increased accuracy comes from the fact that a specific extreme event introduced in the reference dataset will be fully accounted for, while, in the classical approach, it will depend on its probability or occurrence. As a counterpart, this increased sensitivity to rare events has required an extensive and specific manual QC step. It has also been evidenced that, being more accurate, the approach may fail suddenly when observing scales of variability poorly sampled in the reference dataset; as a solution, it is claimed that the Min/Max reference fields need to be updated periodically; it is proposed that, inside an operational data production system (such as the CORIOLIS facility) it will be important that the alerts raised by automatic QC procedures but infirmed by an operator be included regularly in the reference dataset.

ACKNOWLEDGEMENTS

This work was supported by the European Commission funding through MyOcean 2, MyOcean-follow-on and Copernicus Marine Environment Monitoring Service contracts..

REFERENCES

Boyer, T., and S. Levitus, 1994: Quality control and processing of historical oceanographic temperature, salinity and oxygen data. 592 -601 pp.

Cabanes, C., and Coauthors, 2013: The cora dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements. Ocean Science, 9 (1), 1–18.

Carton, J. A., G. Chepurin, X. Cao, and B. Giese, 2000: A simple ocean data assimilation analysis of the global upper ocean 1950-95. part i: Methodology. Journal of Physical Oceanography, 30 (2), 294–309.

Gandin, L. S., 1988: Complex quality control of meteorological observations. Monthly Weather Review, 116 (5), 1137–1156.

Ingleby, B., and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles historical and real-time data. Journal of Marine Systems, 65 (1), 158–175.

Rosell-Fieschi, M., S. R. Rintoul, J. Gourrion, and J. L. Pelegri, 2013: Tasman leakage of intermediate waters as inferred from argo floats. Geophysical Research Letters, 40 (20), 5456–5460.

PROFILE CLASSIFICATION MODELS

BY
G. MAZE[1], H. MERCIER[2], C. CABANES[3]

INTRODUCTION

Ocean dynamics and the induced 3-dimensional structure and variability are so complex that it is very difficult to develop objective and efficient diagnostics of horizontally and vertically coherent oceanic patterns. However, identifying such patterns is crucial to the understanding of interior mechanisms as, for instance, the integrand giving rise to Global Ocean Indicators (e.g. heat content and sea level rise). We believe that, by using state of the art machine learning algorithms and by building on the increasing availability of ever-larger in situ and numerical model datasets, we can address this challenge in a way that was simply not possible a few years ago. This letter aims to present the principles and first results of an approach introduced by Maze et al (2017) based on what we coined a «Profile Classification Model» or PCM that focuses on vertically coherent patterns and their spatial distribution.

The goal of a PCM is to automatically extract out of a collection of profiles a synthetic statistical description, i.e. a model, of typical profiles present in the collection. Once a PCM is built, i.e. trained, one can use this model to determine, with probabilities, the typical class any new profile most resembles. Therefore, it becomes possible to assign to a given typical class of profiles appropriate parameters for a specific diagnostic (e.g.: a finely tuned density threshold for mixed layer depth computation, a depth range for a pycnocline or mode water identification), or simply to use the PCM distributions to analyze the climatology or variability of the coherent patterns in space and/or time.

Hereafter, we present the PCM method, its first results and four possible applications for a variety of ocean analysis problems.

[1] **gmaze@ifremer.fr, Ifremer, Univ. Brest, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, F-29280,** Plouzané, France.
[2] **Herle.Mercier@ifremer.fr, Ifremer, Univ. Brest, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, F-29280** Plouzané, France.
[3] **Cecile.Cabanes@ifremer.fr.fr, Ifremer, Univ. Brest, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, F-29280** Plouzané, France

METHOD

The obvious difficulty is in the construction of the PCM. The goal is to automatically determine typical profiles. This can be achieved from one simple idea: if a profile is typical, then it will be redundant (although with small variations) in a sufficiently large and heterogeneous collection of profiles. Hence, a typical profile will have a high probability of occurrence and creates a peak in the probability density function (PDF) of the collection of profiles. We thus can determine typical profiles by creating a model for the peaks of the collection PDF. To do so, we used Gaussian Mixture Models that belong to the class of unsupervised classification methods (Bishop et al 2006). It determines the most likely decomposition of a PDF into a finite sum of Gaussian modes. Each Gaussian mode property provides a model, i.e. a description, for a typical profile, including a mean profile (the center of the mode) and a spread/pattern (the squared covariance matrix of the mode). One should note that a PCM based on a Gaussian Mixture Model identifies vertically coherent patterns because the multi-dimensional Gaussian mode covariance matrix has no reason to be diagonal, which allows for complex vertical relations.

RESULTS

We trained a PCM, based on a Gaussian mixture model, with about 100,000 Argo temperature profiles located in the North Atlantic Ocean. Using a series of subset of the collection with uniform space/time distribution, we determined both objectively and through trial/error that 8 typical profiles characterize the interior large scale temperature structure of the North Atlantic between the surface and 1400m. These typical profiles, together with their spread, are shown in Figure 1. To follow the data mining vocabulary, we may also refer to typical profiles as class of profiles.

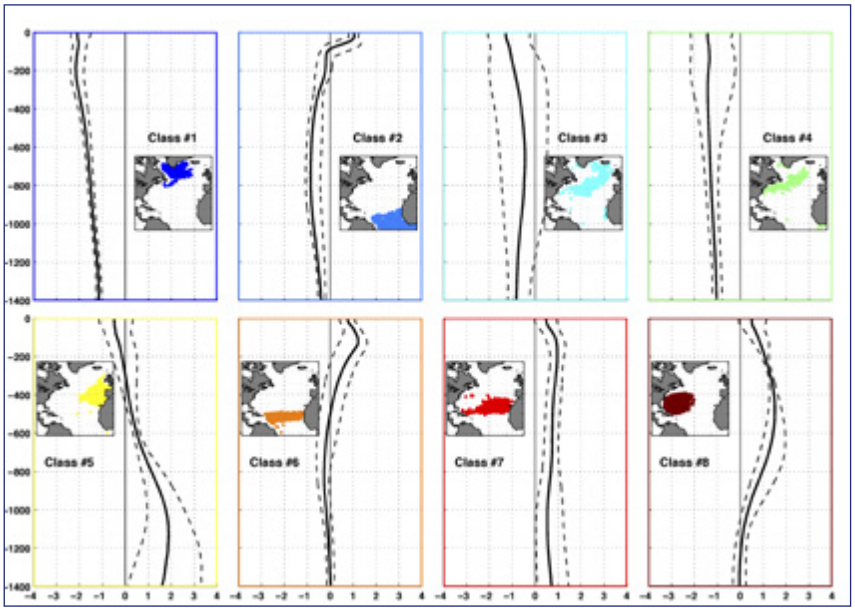


FIGURE 1

The 8 typical temperature profiles of the North Atlantic. Temperature profiles are centered/standardized at each depth level and black dashed lines indicate the 5%-95% spread of the class. Map insets show the location of profiles attributed to each class.

Two classes (#1/#4) show cold anomalies throughout the water column with amplitude decreasing with depth. One class (#3) has nearly zero anomalies, and a large spread throughout the water column. One class (#2) has warm anomalies near the surface (50m) and cold ones below 200m. The remaining four classes have warm anomalies throughout the water column, one without depth dependence (#7), the other three (#5/#6/#8) with clear maxima at different depths (1000m, 100m and 400m respectively).

Another key point of the Maze et al (2017) study is that the 8 typical profiles were identified without using the information of latitudes, longitudes and times of profile samplings. So, we furthermore investigated the locations in time and space of profile classes (note that to classify a profile, we compute the probabilities it has to be similar to each of the typical profiles and select the class maximizing these probabilities). On the one hand we found no correlation between the time of samplings and the classes (not shown). This means that whatever the season (the largest source of temporal va-

riance in the dataset) the same collection of typical profiles characterizes the dataset. On the other hand, we found a key result in locating in space the class of profiles. Figure 1 insets show the location map of profiles attributed to each class. One can see that each class delineates a specific and physically coherent region of the ocean. This is a truly remarkable result because it demonstrates objectively that a given region corresponds to a unique vertical temperature pattern. In other word, the vertical stack of water masses and thermoclines is specific to a region of the North Atlantic Ocean.

A more detailed description of typical profiles, how they relates to known water masses and thermoclines and a sensitivity analysis can be found in Maze et al (2017).

APPLICATIONS

The PCM results briefly presented above pave the way for many possible applications in data analysis and physical studies. Below we briefly review four of these promising applications.

Study of a region with natural boundaries

Let's take the class #1 that delineates the North Atlantic subpolar gyre. We can apply a PCM to a gridded interpolation of Argo data and naturally delineate the subpolar region without using rectangular boxes, complex polygons or surface data from another, possibly incoherent, source. Here we used the Argo-based PCM to classify the ISAS13 time series of optimally interpolated Argo temperature data (Gaillard et al, 2016). Figure 2-A and B show the 2002-2015 grid point average and monthly variance of the local temperature profile probabilities classified in class #1. Map A clearly shows the natural contouring of the subpolar gyre, while map B indicates that the gyre variability is mostly located along its boundaries with a narrow band in the North West Corner Region and a wider band in the Iceland Basin in the North Atlantic Current region. Furthermore, using the PCM property that the sum of profile probabilities to belong to each of the 8 classes, namely), goes to one, we can decompose the local water column heat content into 8 fractions attributed to each class **c**:

Eq.(1)

$$OHC_z(\boldsymbol{c},\theta) = \iint_{x,y}\left(p(c|x,y,t)\int_{z=0}^{z}\rho_0 C_p\theta(x,y,z,t)dz\right)dxdy$$

without losing heat because

$\sum_{c=1}^{8} p(c|x,y,t)=1$, hence $\sum_{c=1}^{8} OHC_z(\boldsymbol{c},\theta) = 1$.

Figure 2-C shows in blue the detrended interannual time series of . We won't explain here the structure of the events shown in this time series but rather focus on its decomposition into the variability arising from local temperature variations vs. gyre horizontal extent. We can indeed approximate Eq.(1) for class #1 by the sum of two terms where either the class extent or the temperature are being set to their time average:
Eq(2)

$$OHC_z(\bar{\boldsymbol{c}}=\mathbf{1},\theta) = \iint_{x,y}\left(\overline{p(\boldsymbol{c}=\mathbf{1}|x,y,t)}\int_{z=0}^{z}\rho_0 C_p\theta(x,y,z,t)dz\right)dxdy$$

Eq(3)

$$OHC_z(\boldsymbol{c}=\mathbf{1},\bar{\theta}) = \iint_{x,y}\left(p(\boldsymbol{c}=\mathbf{1}|x,y,t)\int_{z=0}^{z}\rho_0 C_p\overline{\theta(x,y,z,t)}dz\right)dxdy$$

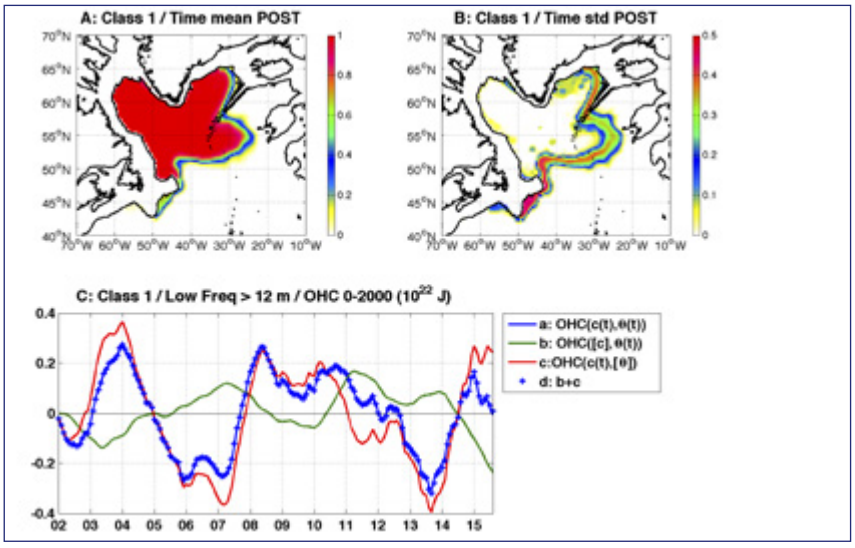


FIGURE 2

Plot A: 2002-2015 time mean probabilities of the subpolar class (#1 from Fig. 1). Plot B: 2002-2015 monthly variance of the subpolar class probabilities. Plot C: Detrended low-frequency variability of the 0-2000m OHC for the subpolar region with blue: total (Eq.1), red: due to class contour variations (Eq.2), green: due to class temperature variations (Eq.3).

Figure 2-C shows the subpolar gyre heat content variability driven by local temperature variations in green, Eq.(2), and by the gyre horizontal extent variations in red, Eq.(3). One can clearly see how the total heat content of the gyre, blue curve - Eq.(1), is driven by this later term while temperature variations appear to be anti-correlated. This means that when the gyre shrinks it also gets warmer and vice versa. We will show elsewhere how this is consistent with the diabatic and adiabatic atmospheric forcing patterns. One could furthermore show that the gyre extent defined through the class contour is consistent with the one diagnosed using the surface embedded within a fixed Sea Surface Height contour (not shown), demonstrating the relevance of the method presented here to delineate the gyre contours.

Structure of frontal regions

The PCM used in Maze et al (2017) is probabilistic, meaning that the transition between a class and another is not a step function but rather fuzzy, allowing for ambiguous profiles to be taken into account. A metric can be derived to interpret how robust is the classification of a profile. When mapped in space, robust classifications are found for profiles located in the core of the region they define (see their Fig.12). But one striking result is that highly robust classifications also appear to be located along frontal regions. This simply means that a PCM easily differentiates profiles from both sides of a front.

This is illustrated in Figure 3. We trained a 3-class PCM from temperature profiles of an eddy-resolving model simulation at 1/12° resolution in the Gulf Stream region (the DRAKKAR simulation referenced as NATL12-BAMT20, used by Maze et al, 2013 to study subtropical mode water formation). Figure 3-A shows the median and spread of class profiles. One can see how the PCM distinguishes the cold northern flank waters (blue profiles, class #a, without a clear vertical structure but the surface spread due to the seasonal cycle) from warm southern flank waters (orange profiles, class #c with almost no spread at 300m indicating the depth of the homogeneous Eighteen Degree Mode Water located above the permanent pycnocline with a larger spread). The remaining class (#b, in green) has a large spread almost throughout the water column. When mapped in space for the 5-days period centered on May 20th, 2003 (Figure 3-B), the class distribution is coherently revealing the horizontal distribution of the vertical structures identified by the PCM, i.e. the northern flank (class #a), the core of the front (class #b) and the southern flank (class #c) of the Gulf Stream. Figure 3-C and D with interior temperature and surface relative vorticity furthermore illustrates the accurate distinction being made by the PCM between these regions.

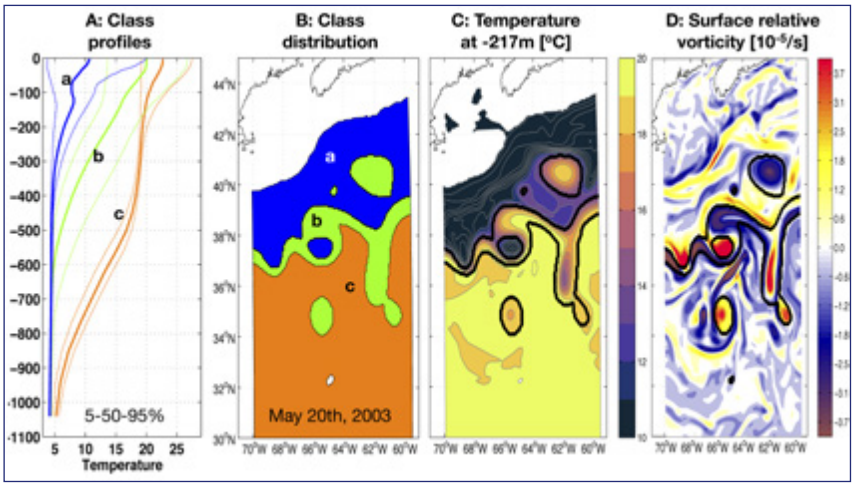


FIGURE 3

Demonstration with a 5-days averaged model output (1/12° resolution) that a PCM can distinguish the Gulf Stream front (class #b) from its flank water masses (classes #a and #c) only with ocean interior data. Plot A: 5-50-95% percentile class profiles. Plot B: Class attributed to profiles for May 20th, 2003. Plots C-D: interior temperature and surface relative vorticity superimposed with class contours (black) for the same date.

It is also of high interest to note that geographical incursions of a given class into the other coincide with meso-scale eddies. In fact, if one increases the number of classes, one could even distinguish cyclonic from anticyclonic eddies into separate classes (not shown). This PCM property will be exploited in the LEFE GMMC/IMAGO project «SOMOVAR» over the next 3 years.

Profile selection for QC in frontal regions

One can also make use of the frontal region PCM performance to improve the selection of reference profiles for Quality Control. This is illustrated in Figure 4. Again, let's take the Gulf Stream Extension region as an example. We trained a 3-class PCM from temperature profiles of the Argo reference database[3]. This is the reference database used in standard

[3] http://www.argodatamgt.org/Reference-data-base/Latest-Argo-Reference-DB

QC methods, such as OW (Cabanes et al, 2016). Figure 4-A shows the median and spread of class profiles and Figure 4-B shows the regional distribution of the class attributed to each profiles. Like in the previous case with the eddy resolving simulation, we again can distinguish the front from its flank waters. On the map we indicated the climatological location of the Gulf Stream core (black dashed line, determined from AVISO Sea Surface Height data as the latitude of the maximum zonal geostrophic velocity) to emphasize the appropriate results of the PCM.

Now imagine that we'd like to quality control a new Argo float set of profiles. A classic approach would be to take the reference collection and to compare float data with statistics from the reference. This is illustrated Figure 4-C and F for profiles #78 and #80 of the Argo float with WMO 4900136 for which the locations, and trajectory, are shown in Figure 4-E and H. From the reference database, we computed at each depth the distance weighted mean and standard deviation of temperature (for the same season as the profile to validate) using 300km and 150km decorrelation length scales in the zonal and meridional directions (results are qualitatively similar if one reduces these scales by a factor of 2). This standard reference envelop is shown in blue in Figure 4-C and F, while the float profiles are shown in red. For profile #78, data are out of the standard range from -100m to -600m depth. For profile #80, data are out of the standard range from -100m down to the bottom of the profile. Thus, using the standard reference envelop, these profiles would look suspicious and would create false alarm from automatic QC procedures.

The PCM method can, in this case, provides useful information to compare float data to the appropriate reference statistics. But first, let's examine the dynamical context of these profiles. In Figure 4-E and H we show the AVISO absolute dynamic ocean surface topography (based on all-satellites in delayed-time) for the same days as the profiles. On the

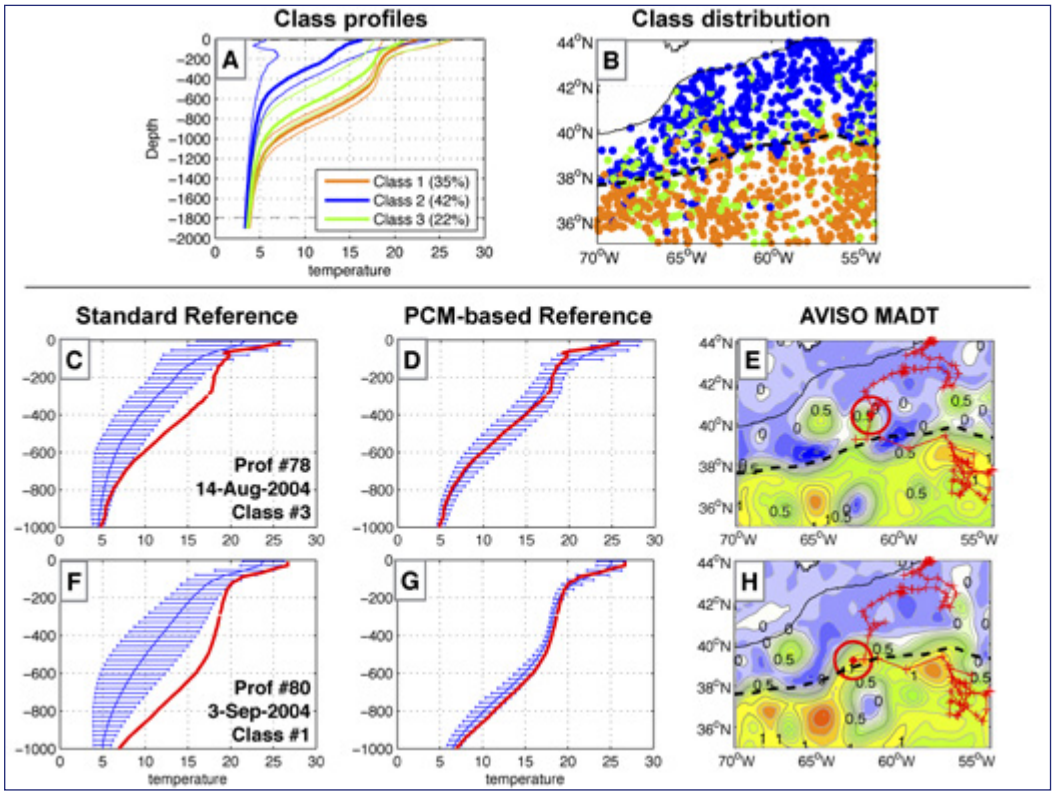


FIGURE 4

Argo-based PCM with 3 classes in the Gulf Stream Extension region. Plot A: 5-50-95% percentiles of the class temperature profiles. Plot B: Location of the profiles attributed to each class superimposed with the climatological Gulf Stream core position determined from AVISO altimetry data (black dashed). Plot C and F: Argo float 4900136 profiles #78 and #80 (in red) superimposed with the standard reference envelop (blue). Plot D and G: Argo float 4900136 profiles #78 and #80 (in red) superimposed with the PCM-based reference envelop (blue). Plot E and H: Argo float 4900136 trajectory in red with profiles #78 and #80 (red circle) superimposed with the AVISO map of Absolute Dynamic Topography height (contours every 0.1m) and climatological Gulf Stream position (black dashed). See text for details of reference envelops.

one hand, this surface dynamical context helps us to localize the profile #78 within the realm of a warm anti-cyclonic eddy located to the North of the Gulf Stream core. Using the PCM shown in Figure 4-A/B the float profile #78 was attributed to class #3, the frontal class, which is coherent with its dynamical context shown Figure 4-E. On the other hand, the float profile #80 is localized on the warm flank of the Gulf Stream that, for this particular date, is further north than its climatological position (black dashed line, Figure 4-H). Interestingly, despite the profile being localized close to the Gulf Stream core and flanked to the South by a cyclonic cold eddy, the PCM appropriately attributed it to the class #1 of southern warm waters.

This illustrates how appropriate is the PCM attribution of profiles to group with similar vertical structures, whatever their location in space. Thus, in Figure 4-D and G are shown the PCM-based reference envelop (in blue) for which the distance weighted mean and standard deviations were computed only with profiles of the reference collection attributed to the same class as the Argo float profiles to be validated. The difference with the standard reference envelops is striking. Both profiles #78 and #80 are now within the bounds of the reference statistics and would no longer raise false alarms.

One can note that these results are robust to the decorrelation length scales and to the number of profiles used to compute the standard deviation (also note that we used the same number of reference profiles to compute the standard and PCM-based statistics).

Model Evaluation

A PCM also represents an elegant method to synthesize the structural information of a profile collection. Then a PCM trained with observations can be used to evaluate numerical model realism by comparing the space/time distribution of the classes. We could also compare the two optimal PCMs trained with observations on the one hand and the numerical model on the other hand (e.g. Fig.3-A from a model compared to Fig.4-A from observations). But let's illustrate a simpler first case here. We used the Argo-based PCM of Maze et al (2017) for the North-Atlantic to evaluate a state of the art global ¼° resolution configuration of a NEMO simulation. Figure 4-A and B show the distribution of the classes attributed to model temperature profiles for the first year of the simulation (1958) and for the entire run period (1958-2015). This distribution should be compared with Fig.1 insets or Fig.11 in Maze et al (2017).

This evaluation indicates that, although the class distribution is correct at the beginning of the run (no spin-up was performed, so the first year remains close to the initial conditions based on observations), the model dynamics clearly modifies the stratification structure (not shown), which leads to a re-arrangement of the classes in space. With regard to this method, the model performs well in most of the North Atlantic Ocean, except over the Western subtropical region, south of the Gulf Stream, where class #3 (cyan) takes over class #8 (brown) that, in turn, considerably shrinks. This is due to the model dynamics in the Gulf Stream region that erodes the vertical stratification structure (the mode water and underlying permanent pycnocline, Feucher et al, 2016) and sustains a more vertically uniform structure, thus the new state is more like class #3 than class #8 (see Fig.1).

CONCLUSION

In this letter, we briefly presented a data mining statistical method recently proposed by Maze et al (2017) for physical oceanographic studies. This method is coined «Profile Classification Model» or PCM. It is based on a state of the art un-supervised classification method, a Gaussian Mixture Model, being applied to ocean profiles. They proposed several PCM applications, four are illustrated here.

This method is coined «Profile Classification Model» or PCM. It is based on a state of the art un-supervised classification method, a Gaussian Mixture Model, being applied to ocean profiles.

In Maze et al (2017), a PCM is derived from Argo temperature profiles in the North Atlantic Ocean. They showed that In this letter, we briefly presented a data mining statistical method recently proposed by Maze et al (2017) for physical oceanographic studies. This method is coined «Profile Classification Model» or PCM. It is based on a state of the art un-supervised classification method, a Gaussian Mixture Model, being applied to ocean profiles. They proposed several PCM applications, four are illustrated here.

In Maze et al (2017), a PCM is derived from Argo temperature profiles in the North Atlantic Ocean. They showed that 8 classes of profiles capture the diversity of all possible vertical structures (Figure 1). It is worth noting that this vertical structure is not only defined by a vertical mean profile but also by a square covariance matrix that contains 8 classes of profiles capture the diversity of all possible vertical structures (Figure 1). It is worth noting that this information about mode waters (no spread, homogeneity), thermoclines - or any other vertical gradients of the tracer used (large spread) or lack of vertical coherence throughout the class (frontal regions).

Maze et al (2017) showed that, although no spatial information is used to train the PCM, each of the 8 classes is co-localized in space, defining regions of the ocean with a unique vertical structure. We thus illustrated a possible application of such natural region contouring for the North Atlantic subpolar gyre (Figure 2). We furthermore examined the decomposition of its integrated heat content variability into the component driven by local temperature variations and the component driven by the gyre expansion and contraction. We found that,

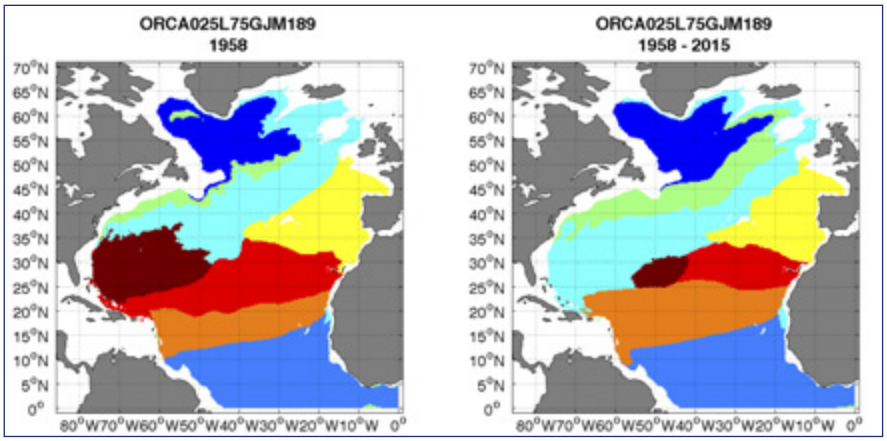


FIGURE 5

Distribution of the locally most frequent classes attributed to a GCM run, for the first year of the simulation, 1958 (left) and for the entire time series, 1958-2015 (right). The PCM used is the one trained with Argo data for which the structure and distribution are shown in Figure 1.

at the interannual time scale, the expansion component drives the gyre heat content while the temperature component is anti-correlated.

We also illustrated how a PCM can shed a new light on turbulent Western Boundary Current regions through the identification and grouping of the possible vertical structures. Using a 1/12° numerical model simulation from the DRAKKAR group (Barnier et al, 2014), we trained a PCM based on temperature data in the Gulf Stream Extension. We showed that, despite the strong seasonality of the profiles, a 3-class PCM is able to disentangle the horizontal complexity of the frontal region structure with a remarkable simplicity (Figure 3). This result will be further investigated and developed in the 2017-2019 INSU LEFE GMMC/IMAGO funded project «SOMOVAR» (LOPS, Telecom Bretagne, CERFACS). In particular, the project aims to use the PCM method for the detection of low-frequency variability and to develop a new product of in-situ gridded data in turbulent Western Boundary Current regions.

For observation data center, this latter result has a powerful direct application: a more appropriate selection of reference data for quality control procedures. This was illustrated in the Gulf Stream region for the hypothetic validation of two Argo float profiles located near the front (Figure 4). When statistics were computed (using a trivial approach of distance weighting and seasonal colocation) from a reference database to evaluate the Argo float profiles, we found the profiles to be outside of the standard reference envelop, hence raising a false alarm. But when the statistics were computed using only reference profiles with the same PCM class of the Argo float to be validated, then we found the profiles to be within the PCM-based reference envelop. The dynamical context of the profiles was provided by the AVISO altimetry data. It showed that profiles were too close to the front or within an eddy for the standard approach to be able not to bias low the reference envelop. Obviously, this can be avoided if the QC operator uses altimetry for context or more complex method, such as OW, working along isopycnal surfaces. But the PCM method, simple, automatic and in the depth/pressure space can surely help QC procedures.

Last, another application of the PCM method for model evaluation was illustrated. Indeed, a PCM is a reduced representation of the statistical properties of a collection of profiles. It thus provides the opportunity to compare two collections or to assess one with regard to the properties of the other. First case could be achieved for instance by comparing Fig.3-A from an eddy resolving simulation with Fig.4-A from Argo data. Here, we simply illustrated the later scenario in Figure 5. Using the Argo-based PCM of temperature profiles from Maze et al (2007), we classified a 1958-2015 time series of a global circulation model experiment at ¼° resolution (ORCA025, 75 vertical levels, referenced by the DRAKKAR group as GJM189). We found the model initial state to be close to observations (as expected, because the model had no spin-up) but to drift away from a realistic stratification in the Southern recirculation region of the Gulf Stream, while the other regions were stable and remained realistic. This synthetic metric provides an elegant way to assess the realism of the model state.

Our group is currently working on the PCM applications illustrated here. But, obviously, this approach can be used with other data (e.g. salinity, both temperature and salinity, density, stratification...), in other frontal regions (e.g. the ACC) and with other datasets (e.g. CORA4.2, high resolution model outputs). To foster such possible applications we made available online the Argo-based PCM (http://doi.org/10.17882/47106) and a toolbox to easily train a PCM and classify new data (https://github.com/obidam/pcm) from a collection of profiles or gridded datasets.

REFERENCES

Bishop, C. M., 2006: Pattern recognition and machine learning, Springer (738p).

Maze, G., 2017: A Profile Classification Model from North-Atlantic Argo temperature data, Seanoe, doi: 10.17882/47106.

Maze, G., Mercier, H., Fablet, R., Tandeo, P., Lopez Radcenco, M., Lenca, P., Feucher, C., and Le Goff, C., 2017: Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean. Progress in Oceanography, doi: http://dx.doi.org/10.1016/j.pocean.2016.12.008.

Feucher, C., Maze, G., and Mercier, H., 2016: Mean structure of the North Atlantic subtropical permanent pycnocline from in-situ observations. Journal of Atmospheric and Oceanic Technology, doi: 10.1175/JTECH-D-15-0192.1.

Cabanes, C., Thierry, V., and Lagadec, C., 2016: Improvement of bias detection in Argo float conductivity sensors and its application in the North Atlantic. Deep Sea Research Part I: Oceanographic Research Papers, doi: http://dx.doi.org/10.1016/j.dsr.2016.05.007.

Gaillard, F., Reynaud, T., Thierry, V., Kolodziejczyk, N., and von Schuckmann, K., 2015: In Situ–Based Reanalysis of the Global Ocean Temperature and Salinity with ISAS: Variability of the Heat Content and Steric Height. Journal of Climate, 29 (4), 1305--1323, doi: 10.1175/JCLI-D-15-0028.1.

Barnier, B., Blaker, A. T., Biastoch, A., Böning, C. W., Coward, A., Deshayes, J., Duchez, A., Hirschi, J., Sommer, J. L., Madec, G., Maze, G., Molines, J.-M., New, A., Penduff, T., Scheinert, M., Talandier, C., and Treguier, A.-M., 2014: DRAKKAR: developing high resolution ocean components for European Earth system models. CLIVAR Exchanges, 65, 18-21.

Maze, G., Deshayes, J., Marshall, J., Tréguier, A.-M., Chronis, A., and Vollmer, L., 2013: Surface vertical PV fluxes and subtropical mode water formation in an eddy-resolving numerical simulation. Deep Sea Research Part II: Topical Studies in Oceanography, 91 (0), 128--138, doi: 10.1016/j.dsr2.2013.02.026.

IMPLEMENTATION OF AN ARGO-O2 ARRAY IN THE NORTH-ATLANTIC OCEAN

BY
V. THIERRY[1], H. MERCIER[2] AND V. RACAPÉ[1]

The ocean contributes at moderating the climate change by absorbing a significant portion of the carbon dioxide emitted to the atmosphere by human activity and the excess of heat due to the enhanced greenhouse effect. Observations show a warming of the surface and deep layers of the ocean, and climate projections suggest that these trends will strengthen in the future. These changes in oceanic stratification modify the processes involved in the water masses ventilation that allow the penetration of climate signals in the ocean interior. The subpolar gyre of the North Atlantic is a privileged region for water mass transformation and greatly contributes to the formation of dense water and ventilation of the ocean interior. It is characterized by a quasi-decadal variability of both circulation and water mass properties. To better analyze multidecadal trends, it is necessary to better quantify the variability at higher frequencies and to better understand their origin.

The Labrador Sea Water (LSW) is formed by deep convection in the western North-Atlantic Ocean subpolar gyre. In the last decades several sites of LSW production were identified, each site showing large interannual to decadal variability in convection strength. In the Labrador Sea, the deepest winter mixed layers were observed between the late 80s and the mid-90s due to the recurrence of severe winters during this high North-Atlantic Oscillation (NAO) index period. They reached 2300-2400m in 1993-1995 [Kieke and Yashayaev, 2015]. Since the mid-90s, the convection was shallower due to milder winters and the maximum mixed layer depths (MLDs) ranged from 500 to 1500m, except in 2008, 2014 and 2015 when they reached 1700-1800m, respectively [Kieke and Yashayaev, 2015]. In the last decade, the Irminger Sea was also recognized as a deep convection

[1] **Ifremer, Laboratoire d'Océanographie Physique et Spatiale, UMR 6523 CNRS-IFREMER-IRD-UBO,** Plouzané, France.

[2] **CNRS, Laboratoire d'Océanographie Physique et Spatiale, UMR 6523 CNRS-IFREMER-IRD-UBO,** Plouzané, France.

site forming LSW [Piron et al., 2016, 2017]. The end of winter MLDs in the Irminger Sea reached 1000m in 2008 and 2012 [de Jong et al. 2012; Piron et al. 2016] and 1400m in 2016 [de Jong and de Steur 2016; Fröb et al. 2016]. Those changes in the atmospheric forcing and in the convection depth was accompanied with changes in the circulation and Meridional Overtuning Cell amplitudes [Mercier et al., 2015], in the water-mass properties and in the anthropogenic storage in the subpolar gyre [Pérez et al., 2013]. While the transition between the cold period observed in the early 1990s and the warm period observed in the 2000s has been well documented [e.g. Hatun et al, 2005; Thierry et al., 2008; de Boisséson et al. 2012), the transition to a possibly new cold period observed since the early 2010s still need to be documented and investigated. This includes the monitoring of the deep convection in the Labrador and Irminger Seas and the spreading of the water mass anomalies formed locally in those regions and their impacts on large-scale fields in the North Atlantic. O2 is a good tracer for the investigation of deep-convection, ventilation mechanisms and propagation of water mass anomalies. This was illustrated by (Piron et al., 2016) who used O2 data to document deep convection in the Irminger Sea during winter 2011-2012 and the exceptional deep convection that occurred in the subpolar North-Atlantic during winter 2014-2015 (Piron et al., 2017).

In the framework of the OVIDE (http://www.umr-lops.fr/Projets/Projets-actifs/OVIDE) project and Argo program, we initiated in 2010 the implementation of an Argo-O2 array in the subpolar gyre of the North-Atlantic including regular (0-2000m) and deep (0-4000m) Argo floats (Figure 1). As mentioned in Thierry et al. in this newsletter, those data will be used for many other applications. Since 2010, we deployed about 50 floats equipped with an oxygen sensors (Figure 1). The technology was not fully mature at that time and the Argo data stream was not ready to manage those data. This delayed the scientific and operational use of this dataset. A major step forward on the understanding of the O2 sensor, including the refinement of the calibration equations, has been done recently. In addition, the management and QC procedures have been defined (see Thierry et al. in this newsletter). Data of those floats have been redecoded in 2016 according to the new equations and procedures and twelve of them have already been corrected with the LOCODOX tool (see Thierry et al. in this newsletter). The remaining floats will be corrected soon and transmitted to Coriolis. We plan to maintain such array on the long term. This will be possible for the period 2016-2020 owing to the CPER Euro-Argo (Brittany region) that will fund 15 oxygen sensors/year.

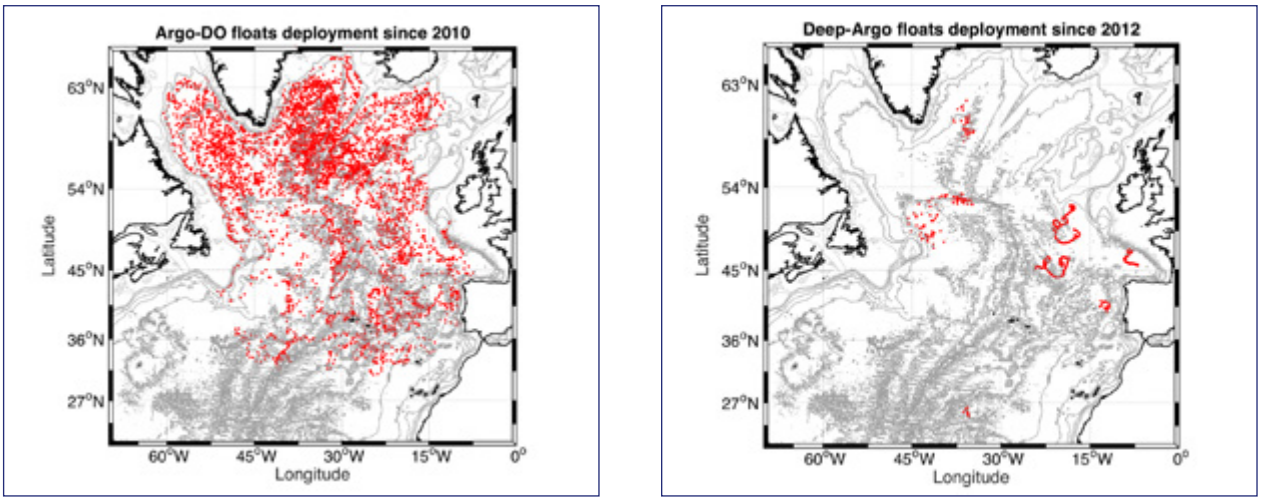


FIGURE 1,2

Positions of the profiles with O2 measurements acquired by the 48 regular Argo floats (left panel) and 10 Deep Argo floats (right panel) deployed by the LOPS in the North-Atlantic Ocean.

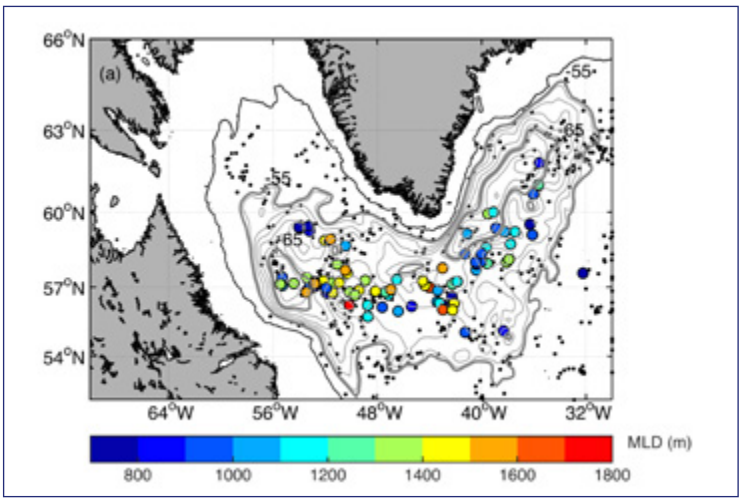


FIGURE 3

Positions of Argo profiles available over January-April 2015. (Black dots) profiles with a mixed layer depth < 700m; (Colored dots) profiles with a mixed layer depth > 700m. Superimposed are the 1000m-isobath (thin black line) and the Absolute Dynamic Topography (cm) averaged over January-April 2015 (gray contours). (Piron et al, 2017).

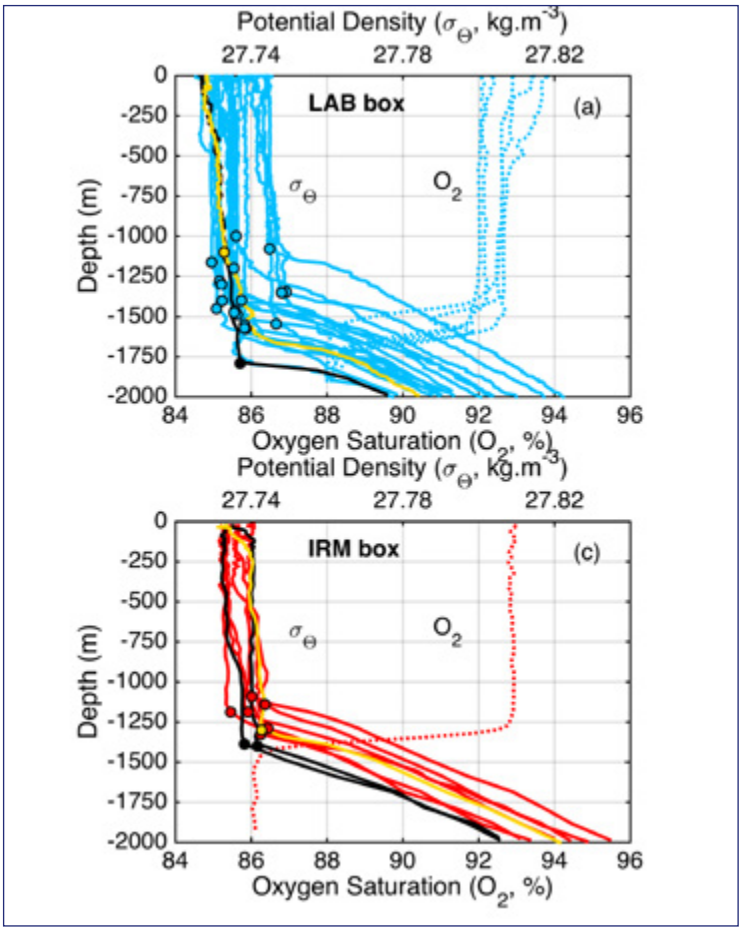


FIGURE 4

Late winter (March-April) profiles with mixed layer depth >1000m in the Labrador Sea and Irminger Seas. Thick lines: potential density (kg.m-3). Thin-dashed lines: oxygen saturation (%). Black line: profiles with the deepest mixed layers. MLD is indicated with dots on the density profiles. (d) Properties of the late winter profiles with MLD>1000m in LAB (blue dots), SCF (green dots) and IRM (red dots). The black circles identify properties of the deepest mixed layers. Potential density and oxygen of the vertical profiles were homogeneous from the surface to the base of the mixed layer, which testifies that the deep convection occurred locally. (Piron et al, 2017)

REFERENCES

de Boisséson E., V. Thierry, H. Mercier, G. Caniaux, and D. Desbruyères, 2012: Origin, formation and variability of the Subpolar Mode Water located over the Reykjanes Ridge, J. Geophys. Res., VOL. 117, C12005, doi:10.1029/2011JC007519

de Jong, M.F., H. M. van Aken, K. Våge, and R. S. Pickart (2012), Convective mixing in the central Irminger Sea: 2002–2010. Deep-Sea Res. I, 63 (1), 36–51.

de Jong M. F., L. de Steur (2016) Strong winter cooling of the Irminger Sea in winter 2014-15, exceptional deep convection, and the emergence of anomalously low SST. Geophys. Res. Lett., 43, 1717-1734, doi:10.1002/2016GL069 596.

Fröb F., A. Olsen, K. Våge, G. Moore, I. Yashayaev, E. Jeansson, B. Rajasakaren (2016), Irminger Sea deep convection injects oxygen and anthropogenic carbon to the ocean interior. Nat. Commun, 7, 13244, doi:10.1038/ncomms13244**.**

Hàtùn, H., A. B. Sando, H. Drange, B. Hansen, and H. Valdimarsson (2005), Influence of the Atlantic subpolar gyre on the thermohaline circulation, Science, 309, 1841–1844, doi:10.1126/science.1114777.

Kieke, D., and I. Yashayaev (2015), Studies of Labrador Sea Water formation and variability in the subpolar North Atlantic in the light of international partnership and collaboration. Prog. Oceanogr., 132, 220-232, doi :10.1016/j.pocean.2014.12.010

Mercier, H., P. Lherminier, A. Sarafanov, F. Gaillard, N. Daniault, D. Desbruyères, A. Falina, B. Ferron, T. Huck, V. Thierry (2015), Variability of the meridional overturning circulation at the Greenland–Portugal OVIDE section from 1993 to 2010. Prog. Oceanogr., 132, 250-261, doi:10.1016/j.pocean.2013.11.001

Pérez, F. F., H. Mercier, M. Vazquez-Rodriguez, P. Lherminier, A. Velo, P. Pardo, G. Roson, and A. Rios (2013), Reconciling air-sea CO2 fluxes and anthropogenic CO2 budgets in a changing North Atlantic. Nature Geosciences, 6, 146-152, doi:10.1038/ngeo1680.

Piron, A., V. Thierry, H. Mercier, and G. Caniaux (2016), Argo float observations of basin-scale deep convection in the Irminger sea during winter 2011–2012, Deep-Sea Res. I, 109, 76-90, doi: 10.1016/j.dsr.2015.12.012.

Piron, A., V. Thierry, H. Mercier, and G. Caniaux (2017), Gyre scale deep convection in the subpolar North-Atlantic Ocean during winter 2014-2015, Accepted for publication in Geophys. Res. Lett.

Thierry V., E. de Boisséson and H. Mercier, 2008: Interannual variability of the Subpolar Mode Water properties over the Reykjanes Ridge during 1990-2006. J. Geophys. Res., 113, doi:10.1029/2007JC004443.

Våge, K., R. S. Pickart, V. Thierry, G. Reverdin, C. M. Lee, B. Petrie, T. A. Agnew, A. Wong, and M. H. Ribergaard (2009), Surprising return of deep convection to the subpolar North Atlantic Ocean in winter 2007-2008. Nat. Geosci., 2, 67–72, doi:10.1038/ngeo382.

Yashayaev, I. (2007), Hydrographic changes in the Labrador Sea, 1960–2005. Prog. Oceanogr., 73, 242–276.

Yashayaev, I., and J. W.Loder (2009), Enhanced production of Labrador Sea Water in 2008. Geophys. Res. Lett., 36, L01606, doi:10.1029/ 2008GL036162.

INTENSE ARGO-O2 FLOATS DEPLOYMENT IN THE WESTERN MEDITERRANEAN SEA

BY
L.COPPOLA[1], C.ESTOURNEL[2] AND P.TESTOR[3]

The Mediterranean Sea is a semi-enclosed sea, sensitive to the climate and anthropogenic pressure compared to the global ocean. If the evolution of water masses is influenced by long term climate change, short time scale forcing as seasonal variability and extreme events have also an impact on the water mass properties. In the northwestern Mediterranean Sea, a dynamic basin where dense water formation occurred in winter (Gulf of Lion), deep and intermediate waters are strongly impacted and their evolution influence the biogeochemical budget. In this context, dissolved oxygen, one of the Essential Oceanic Variables (EOV), allow to trace and quantify several mechanisms affecting the water mass properties (e.g. ventilated dense water volume, recently formed dense water spreading) and the marine ecosystem (e.g. net community production related to the biological activity in surface).

In the northwestern Mediterranean Sea, dissolved oxygen data are mainly obtained thanks to the MOOSE network (national observing system) through moorings, regular monthly and annual cruises and permanent gliders sections. In support of this network, the deployment of Argo-O2 floats, initiated during the HYMEX SOP2 (Estournel et al., 2016), allowed to better constrain the oxygen dynamic in the western basin (MOOXY and MOOXY2 GMMC projects). The Argo-O2 floats are now set up as an operational network as a component of the MOOSE program and the LOP HYMEX actions (Long term Observation Period) for the water mass monitoring. Since the end of 2012, 14 Argo-O2 floats (PROVOR and ARVOR) have been deployed and 3 are still active in the north and south of western basin (Fig.1).

[1]**Sorbonne Universités (UPMC Univ. Paris 06), UMR 7093, Laboratoire d'Océanographie de Villefranche, Observatoire Océanologique,** Villefranche/mer, France

[2] **CNRS, Laboratoire d'Aérologie, UMR5560 CNRS-UPS,** Toulouse, France

[3] **Sorbonne Universités (UPMC, Univ Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory,** Paris, France

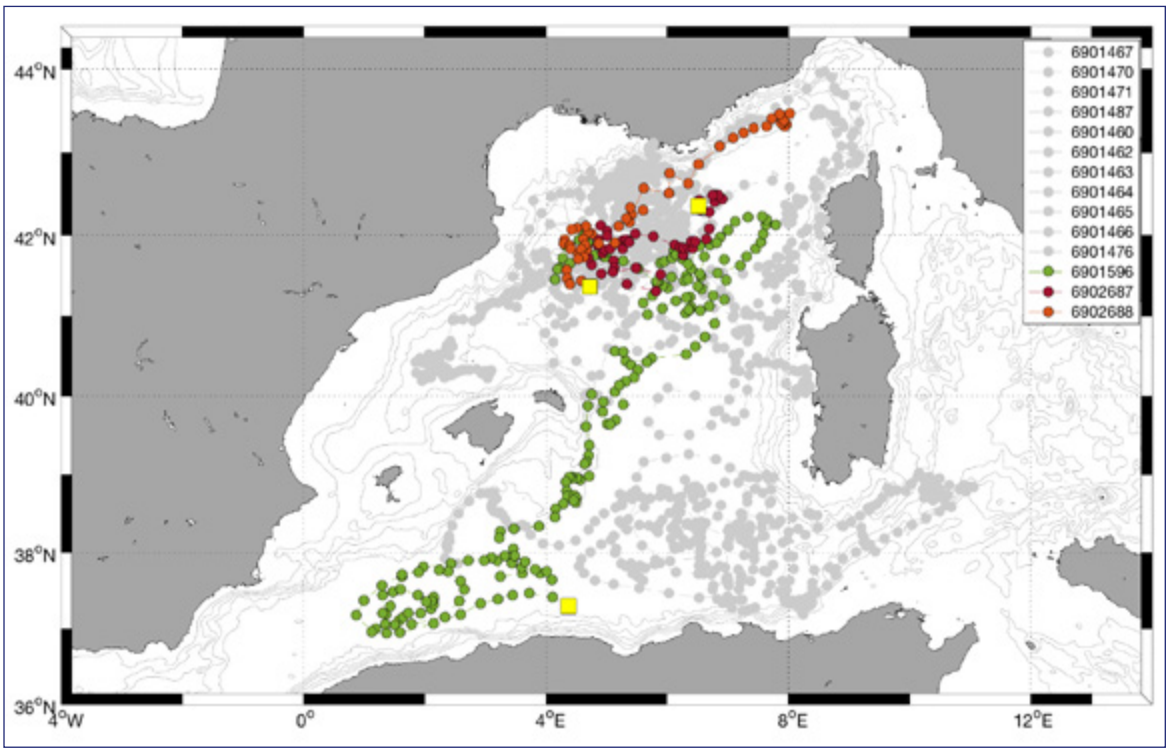


FIGURE 1

Argo-O2 floats deployed in the western Mediterranean Sea from December 2012 to January 2017. Nowadays, 3 floats are still active (color dots). Their last positions are indicated by a yellow square.

REFERENCES

Estournel, C., P. Testor, I. Taupier-Letage, M.-N. Bouin, L. Coppola, P. Durand, P. Conan, A. Bosse, P.-E. Brilouet, L. Beguery, S. Belamari, K. Béranger, J. Beuvier, D. Bourras, G. Canut, A. Doerenbecher, X. Durrieu de Madron, F. D'Ortenzio, P. Drobinski, V. Ducrocq, N. Fourrié, H. Giordani, L. Houpert, L. Labatut, C.L. Brossier, M. Nuret, L. Prieur, O. Roussot, L. Seyfried, and S. Somot. 2016. HyMeX-SOP2: The field campaign dedicated to dense water formation in the northwestern Mediterranean. Oceanography 29 (4) : 196–206, https://doi.org/10.5670/oceanog.2016.94