Argo data management report 2012 Coriolis DAC & GDAC

Data Assembly Centre and Global Data Assembly Centre Annual report September 2011 - October 2012 Version 1.0 November 6th, 2012



Status

(Please report the progress made towards completing the following tasks and if not yet complete, estimate when you expect them to be complete)

- Data acquired from floats
- Data issued to GTS
- Data issued to GDACs after real-time QC
- Data issued for delayed QC
- Delayed data sent to GDACs
- Web pages
- Statistics of Argo data usage (operational models, scientific applications, number of National PIs...)
- Products generated from Argo data ...

This report covers the activity of Coriolis data centre for a one year period from September 1st 2011 to September 30th 2012.

Data acquired from floats

These last 12 months¹, a total of 18 841 profiles from 1390 floats was collected, controlled and distributed.

The 1390 floats managed during that period had 37 versions of data format:

- APEX: 22 versions
- NEMO: 1 version
- NOVA: 1 version
- PROVOR-Arvor: 13 versions



Map of the 18 841 profiles from 1390 floats managed by Coriolis DAC this current year

Argo data management

¹ From September 2011 to October 2012

In October 2012, the first profiles from the deep ocean profiler Arvor 3500 were decoded by Coriolis data centre. The 3500 decibars temperature and salinity profiles where transmitted from mid-Atlantic through Iridium SBD transmission.



Arvor 3500: going deeper in ocean.

Data issued to GTS

All profiles processed by Coriolis are distributed on the GTS by way of Meteo-France. This operation is automatically performed. After applying the automatic Argo QC procedure, the Argo profiles are inserted on the GTS every 2 hours. Argo profiles are inserted on the GTS 365 days per year, 24 hours a day.



CORIOLIS DAC: Argo data flow

Data issued to GDACs after real-time QC

All meta-data, profiles, trajectory and technical data files are sent to Coriolis and US-GODAE GDACs. This distribution is automated.

Since August 7th 2012, Coriolis files are distributed in the Argo new format versions (profiles and trajectory version 2.3; metadata and technical data version 2.4), unfortunately, these files are rejected by the US-GODAE server file format checker.

The provision of data in the new format was dictated by a model change in the internal Coriolis database. This change was not related to Argo data-management. But, after the model change, only new Argo formats could be generated.

Data issued for delayed QC

All profile files are sent to PIs for delayed QC. Most of the Atlantic data handled by Coriolis are checked by the European project Euro-Argo.

Delayed mode data sent to GDACs

An Argo delayed mode profile contains a calibrated salinity profile (psal_adjusted parameter).

A total of 9 193 new delayed mode profiles where sent to GDACs this year.

The number of delayed mode profiles increased by 11%.

A total of 91 402 delayed profiles where sent to GDACs since 2005.

Web pages

The web site of the French DAC is available at:

<u>http://www.coriolis.eu.org/Observing-the-ocean/Observing-system-networks/Argo</u>

It provides:

- Individual float description and status (meta-data, geographic map, graphics : section, overlaid, waterfall, t/s charts)
- Individual float data (profiles, trajectories)
- FTP access
- Data selection tool
- Global geographic maps, GoogleEarth maps
- Weekly North Atlantic analyses (combines Argo data and other measurements from xbt, ctd, moorings, buoys)

Some pages of Coriolis web site are dedicated to technical monitoring:

<u>http://www.coriolis.eu.org/Observing-the-ocean/Observing-system-networks/Argo/Support-to-Data-Mgt/At-sea-monitoring</u>





Example 1: technical monitoring of Argo-France floats

Example 2: age map of floats managed by Coriolis DAC.

Data centre activity monitoring: Coriolis operators perform an activity monitoring with an online control board.

	OPERATEUR	a farme a free posterior a		Ifreme
ut (GAOA)/See all	Codec Arge (Collectes COre, COMusiens)	Ginstramenta GManuelles (GMa	Desen (Officiers /O Techniques	ice an instance (711)
IN BRAIL BRA	From TAL Automation and Aread			
the THE END OF	Argo Gdec metedata controler			
IL BA BE BLAR	Saus Book atoms mucharing			
10 415 APT 411	Cellecte Arge BAC			OR. 2011 11 ATT 16.41-962
20.811.02.00	Collects Aras Jul.C. chargement hour			VARIABLE PRESS PROFESSION
10 #11 #17 #13 O	California Arga BAC Na		• •	04. 2811 SLATTIC 29-372
O.M. MI. M. R.	Ethisian Acas SibAC Indexes		•	06. 2015 ST.BITRIJS.26.212
CB.19.49.49.49.49.07	Diffusion Argo SBAC profile detailed indexe		• •	
CO. MILLING. MIL	Editation Argan Arab		• •	
045.31	Eliftuaturi Argos graphico	•		NO JUST DI ACTORACIÓN
SO AND	Ethiston, Augu, multi, profile, Apr.		•	DR. JOHL MATTERNAL
10.45 (00.45 (0)	Ethiston Jogo multi profile pire		• •	08. 2011.11.01111.22.012
THE SAME AND AND ADDRESS OF	Elithesten, Argn. multi-profile. Jutent		•	04 2811.11 ATT16.28 8822
20.05.35	Editories and taken frances from		• •	108. 2011.11.01111.03.072
D. 03 INF MIL RE	When it least table args_index_anothen		• •	08. 2811 11 81113-82-882
10.83.87.84	Nyrecknowikation Juger, Cardella JulithC			
the local sector water	Statistics in the state of the state			the many to provide the

Example 1: distribution activity on 03/11/2011. An operator has to perform a diagnostic on anomalies of Argo data distribution (red smileys). A series of small data base incidents explains the unusual situation.



Example 2: data distribution to GDAC activity in March 2011. On 26th, a bigger than usual data distribution delayed the update of DAC files.

Statistics of Argo data usage (operational models, scientific applications, number of National Pis...)

Operational oceanography models; all floats data are distributed to:

- French model Mercator (global operational model)
- French model Previmer (regional operational)
- French model Soap (navy operational model)
- EU MyOcean models (Foam, Topaz, Moon, Noos)
- EuroGoos projects

Argo projects: this year, Coriolis data centre performed float data management for 34 Argo scientific projects and 48 PIs (Principal Investigators).

List of Principal Investigators in 2012					
Alain SERPETTE	Holger GIESE				
Alban LAZAR	Isabelle TAUPIER-LEPAGE				
Andreas STERL	Jens SCHIMANSKI				
Andreas Sterl	Jose Luis PELEGRI				
Antoine POTEAU	Juergen FISCHER				
Bernard BOURLES	Juliet HERMES				
Bert RUDELS	KP. Koltermann				
Bert Rudels	Kjell Arne MORK				
Birgit Klein	Louis PRIEUR				
BOURLES Bernard	Olaf KLATT				
C. PROVOST et N. BARRE	PASCUAL Ananda				
Cecile CABANES	Pierre Marie POULAIN				
Christine COATANOAN	Rena CZESCHEL				
Christophe MAES	Sabrina SPEICH				
Detlef QUADFASEL	Sabrina SPEICH et Michel ARHAN				
Dr Osvaldo ULLOA	Serge LE RESTE				
Einar SVENDSEN	Stephane BLAIN				
Fabien ROQUET	Sunke Schmidtko				
FONT Jordi	VELEZ BELCHI Pedro Joaquin				
Frederic VIVIER	Violeta SLABAKOVA				
Gerard ELDIN	Virginie THIERRY				
Gerasimos KORRES	Xavier ANDRE				
Gilles REVERDIN	Xavier CARTON				
Gregorio PARRILLA	Yves GOURIOU				

List of scientific projects active in 2012					
ARGO SPAIN	FLOPS				
ARGO_AWI	GOODHOPE				
ARGO_BUL	IFM				
ARGO_CHILE	IFM-GEOMAR				
ARGO_FIN	IFM2				
ARGO_NORWA	MEDARGO				
ARGO_SPAIN	MEDARGO_IT				
ASA	MFSTEP				
BIOArgo	OVIDE				
BSH	PROSAT				
CONGAS	SHOM				
CORIOLIS	TRACK				
CORIOLIS_OVIDE	TRACK2010				
DAP	WEN				
EGYPT					

Products generated from Argo data ...

Distribution of Argo oxygen observations to EU former CarboOcean project.

Once a week, all Argo floats data with oxygen observations are distributed to the German data centre Pangea using the OAI inter-operability protocol (Open Archive Initiative).

This year, 9 086 new oxygen profiles from 240 floats were distributed.

A total of 57 429 oxygen profiles from 459 floats were distributed since 2004.



Oxygen profiles collected by all Argo partners since 2004.

Sub-surface currents Atlas

Based on Argo trajectory data, Michel Ollitrault and the Ifremer team are continuously improving the "Andro" atlas of deep ocean currents.



Argo trajectories from Coriolis DAC are carefully scrutinized to produce the "Andro" atlas of deep ocean currents.

Argo data management

Delayed Mode QC

(Please report on the progress made towards providing delayed mode Argo data, how it's organized and the difficulties encountered and estimate when you expect to be pre-operational).

At the Coriolis data centre, we process the delayed mode quality control following four steps. Before running the OW method, we check carefully the metadata files, the pressure offset, the quality control done in real time and we compare with neighbor profiles to check if a drift or offset could be easily detected. As last year, we have worked on this way with PIs to strengthen the delayed mode quality control.

Some floats have been deployed from some projects, meaning a lot of PIs and a lot of time for explaining the DM procedure to all of them. A few PIs are totally able to work on DMQC following the four steps but this is not the case for most of them. Since the unavailability of the PIs leads to work by intermittence and then extend the period of work on the floats, we did the work with a private organism (Glazeo) to improve the realization of the DMQC, exchanging only with the PIs to validate results and discuss about physical oceanography in studied area. Working in this way, we have largely improved the amount of delayed mode profiles.

For a few projects, there are still no identified operators to do DMQC, for instance the first run has been done by students which have now left institutes or are not available to carry on with this work. For floats which are German floats (AWI), we found a new operator to run the DMQC. Nevertheless we have made progress with BSH (Marek Stawarz) and some floats have been processed in DMQC or are in progress (we are finalizing delayed mode QC for some floats). Only a few projects are still waiting for PI's answers.



Percentage of floats by country in the Coriolis DAC.

Codes for the countries: 06 : Germany - 15 : Bulgaria - 20 : Chili – 26 : Denmark – 29 : Spain – 34 : Finland - 35 : France – 36 : Greece - 48 : Italy – 57 : Mexico - 58 : Norway – 64 : Netherlands – 90 : Russia – CR : Costa Rica



Percentage of floats by country and by launch's year in the Coriolis DAC

Concerning the APEX floats, some progresses have been done to correct the surface pressure. Most of the APEX belongs to Germany; a lot of those German floats have been corrected by BSH. Some of the French APEX floats still need to be review in the decoding step and are in the grey list.

During the last year, 8671 new delayed mode profiles where produced and validated by PIs. A total of 90784 delayed mode profiles where produced and validated since 2005.



Evolution of the DM profiles' submission versus dates





Left: in terms of float percent and right: in terms of profile percent (DM : delayed mode – RT : real time).

The status of the quality control done on the Coriolis floats is presented in the following plot. For the two last years (2011-2012), most of the floats are still too young (code 1) to be performed in delayed mode. For the year 2010, we are working on the DMQC of those floats, which should be available for the end of this year. The codes 2 and 3 show the delayed mode profiles for respectively active and dead floats.



Status of the quality control done on profiles sorted by launch's year, code 1: young float, code 2: active float, DM done, code 3 : dead float, DM done; code 4 : DM in progress, code 5 : waiting for DM, code 6 : problems with float.

Reference database

The version CTD_for_DMQC_2012V01 is available since March 2012. A new version CTD_for_DMQC_2012V02 should be on line in November 2012.

The March's version takes into account updates from NODC as well as feedbacks from users about duplicate or invalid pair, and bug in position of some stations in not appropriate boxes. The next version will be available for November 2012, some new CTD provided by the updates of WOD2009 and by the CCHDO will be integrated as well as CTD dataset from ICES. This version will be provided in smaller tar balls, one by wmo box area (1-3-5-7): for instance, CTD_for_DMQC_2012V02_1.tar.gz for all boxes starting with wmo 1, then we will have 4 tar files.



Example of delayed mode activity

A comparison between Argo float observations with SLA and DHA (SLA, Sea Level Anomalies; DHA, Dynamic Height Anomalies) is performed on a routine mode, 4 times a year.



GDAC Functions

(If your centre operates a GDAC, report the progress made on the following tasks and if not yet complete, estimate when you expect them to be complete)

- National centres reporting to you
- Operations of the ftp server
- Operations of the www server
- Data synchronization
- Statistics of Argo data usage : Ftp and WWW access, characterization of users (countries, field of interest : operational models, scientific applications) ...

National centres reporting to you

Currently, 11 national DACs submit regularly data to the French GDAC.

The additional GTS DAC contains all the vertical profiles from floats that are not managed by a national DAC. These data come from GTS and GTSPP projects. The GTS profiles are quality controlled by the French DAC (Coriolis).

On November 5th, the following files were available from the GDAC FTP site.

DAC	metadata	increase	profile files	increase	delayed	increase	trajectory	increase
_	_files _	from last		from last	mode	from last	files	from last
	<u> </u>	year 🗾 🎽		year 🗾 🗾	profile file	year 🔄 🎽		year 📃 🔼
AOML	4 366	5 11%	535 645	16%	398 027	24%	4 236	11%
BODC	402	16%	37 274	16%	30 329	0%	383	16%
Coriolis	1 466	i 10%	127 401	16%	90 715	10%	1 382	11%
CSIO	127	27%	8 388	49%	5 879	20%	124	25%
CSIRO	533	14%	65 215	32%	35 006	11%	527	15%
INCOIS	274	16%	33 006	16%	21 061	2%	272	15%
JMA	1 144	7%	128 402	11%	83 666	10%	1 131	8%
КМА	161	. 10%	15 904	18%	9 982	0%	145	11%
KORDI	119	0%	14 142	22%	0		119	0%
MEDS	344	9%	35 639	11%	23 261	4%	336	8%
NMDIS	19) 27%	1 193	81%	0		19	27%
Total	8 955	11%	1 002 209	17%	697 926	16%	8 674	11%

One million profiles available on Argo GDAC

Since October 30th 2012, more than one million Argo floats profiles are available on the GDACs ftp servers (global data assembly centres).

The millionth profile arrived on 31/10/2012 at 16:50:07.

This profile was transmitted by University of Washington float 5901891, cycle 147.

More on: http://www.argodatamgt.org/Data-Mgt-Team/News/One-million-Argo-profiles



Number of Argo profiles on GDAC ftp server, yearly distribution

Operations of the ftp server

- Meta-data, profile, trajectory and technical data files are automatically collected from the national DACs;
- Index files of meta-data, profile and trajectory are daily updated ;
- GDAC ftp address: <u>ftp://ftp.ifremer.fr/ifremer/argo</u>

There is a monthly average of 285 unique visitors, performing 2397 sessions and downloading 1614 gigabytes.

The graphics show a stable activity on GDAC FTP during the last 12 months.

ARGO GDAC FTP statistics		_		
month	unique visitor	number of visits	hits	bandwidth GB
10/2011	445	2 304	3 242 977	1325
11/2011	356	2 240	1 239 557	1558
12/2011	197	2 138	1 502 266	1732
01/2012	210	2 242	2 454 436	1455
02/2012	197	2 166	2 346 817	1030
03/2012	292	2 656	2 570 711	1612
04/2012	302	2 550	3 085 432	1702
05/2012	350	2 624	2 659 654	1607
06/2012	369	2 540	3 159 930	1884
07/2012	209	2 292	2 574 372	1357
08/2012	209	2 382	2 498 232	2108
09/2012	284	2 630	2 503 028	1993
Average	285	2 397	2 486 451	1 614



The web site address is: <u>ht</u> t	tp://www.argodatamgt.org
--------------------------------------	--------------------------

ARGO GDAC	web statistics			
month	unique visitor	visits	pages	bandwidth MO
10/2011	384	820	2 533	283
11/2011	354	655	2 483	209
12/2011	330	652	2 795	503
01/2012	336	601	2 271	254
02/2012	745	992	2 273	231
03/2012	602	866	3 486	293
04/2012	330	594	5 838	216
05/2012	325	526	5 844	276
06/2012	398	627	5 519	199
07/2012	371	664	5 622	146
08/2012	379	764	6 348	258
09/2012	516	939	6 527	289
Average	423	725	4 295	263



Data synchronization

The synchronization with US-Godae server is performed once a day.



Example of synchronization monitoring: duration of the process in June 2010

FTP server monitoring

The Argo GDAC ftp server is actively monitored by a Nagios agent (see <u>http://en.wikipedia.org/wiki/Nagios</u>).

Every 5 minutes, a download test is performed. The success/failure of the test and the response time are recorded.

In April 2012, the FTP server was moved to a virtual server on a linux cluster. The reliability increased as well as the data transfer speed. The electrical power supply upgraded last year proved to be reliable; we did not face any electrical power supply problem.

The ftp server was available for 99.98 % of the time (compared to 99.69% last year)

The 0.02% of failure represents **1 hour 52 minutes and 54 seconds of interruption** (compared to 1 day 5 hours and 45 minutes last year).

The main problems problem occurred on May 1st 2012. The ftp server failed down, but was automatically reactivated on another node of the cluster).

Compared to last year, the new ftp server dramatically increased the files transfer time **from 100ms to 4 ms**: the files are downloaded up to 25 times faster.



Nagios monitoring: between September 2011 and October 2012



Nagios monitoring: duration of a test file download on October 2012

Argo data management

Grey list

According to the project requirements Coriolis GDAC hosts a grey list of the floats which are automatically flagged before any automatic or visual quality control.

The greylist has 927 entries (November 5th 2012), compared to 1181 entries one year ago.

Statistics of Argo data usage: Ftp and WWW access, characterization of users (countries, field of interest : operational models, scientific applications) ...





Argo GDAC : floats distribution per DAC in October 2012

Argo GDAC : profiles distribution per DAC in October 2012^2



Argo floats available from GDAC in October 2012 (This map includes active and old floats)

Argo data management

² Warning: the blue line displays the total number of active floats during a year. This total is different than the floats active at a particular day.



Active Argo profiling floats available from GDAC in October 2011



Argo GDAC : delayed-mode profiles available in October 2012



Argo GDAC : delayed-mode profiles distribution % per DAC in October 2012



Argo profiling floats with delayed-mode profiles available from GDAC in October 2012

Regional Centre Functions

(If your centre operates a regional centre, report the functions performed, and in planning)

Coriolis is involved in the North Atlantic Argo regional centre (NAARC).

This activity is managed within the European project Euro-Argo. It involves a regular monitoring of the consistency of the quality of data from various types of floats, with techniques such as objective analyses, comparison between floats and altimetry.

The CTD reference data base development is also a contribution to NAARC, as well as the floats deployment coordination.