

Argo data management report 2020

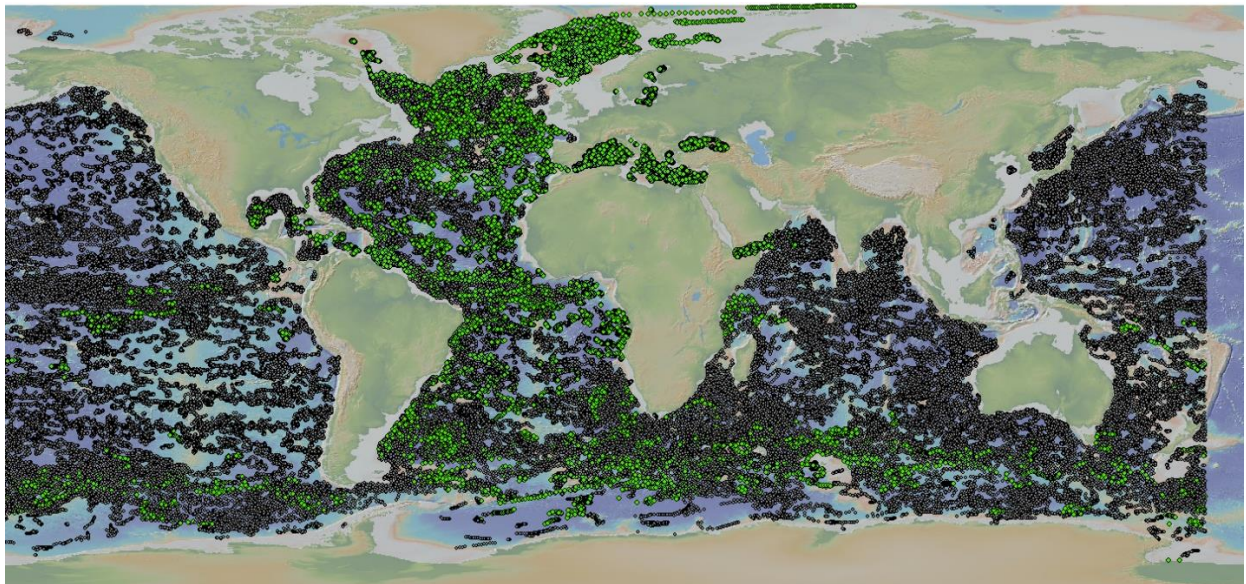
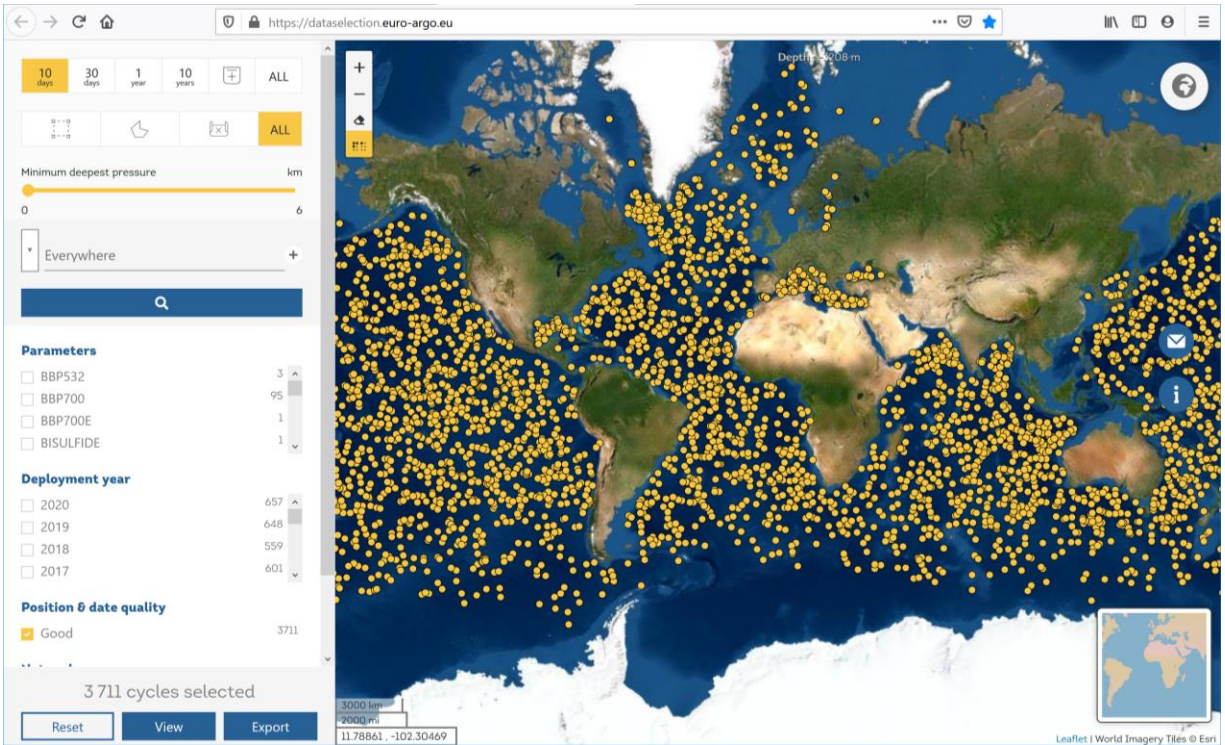
Coriolis DAC & GDAC

Data Assembly Centre and Global Data Assembly Centre

Annual report November 2020

Version 1.0

<https://doi.org/10.13155/77033>



1 DAC status

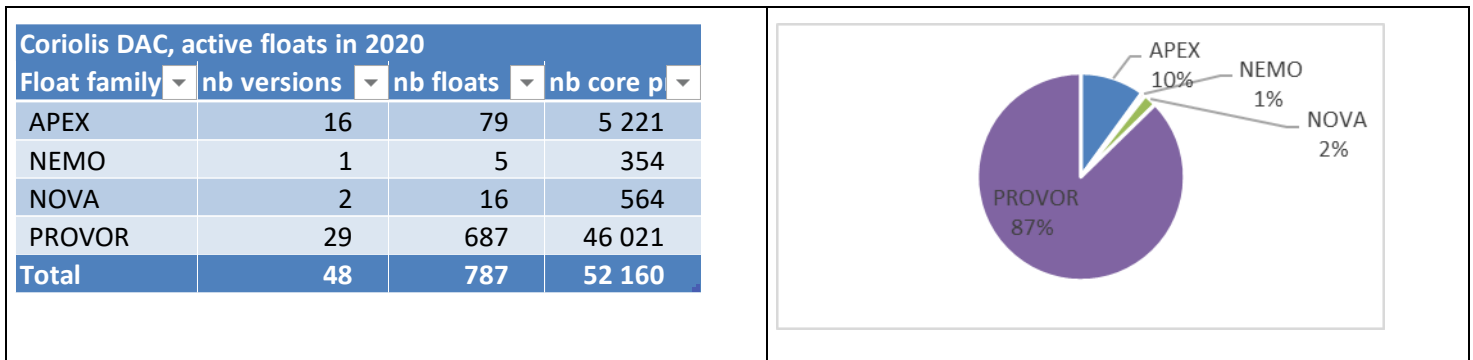
This report covers the activity of Coriolis DAC (Data Assembly Centre) for the one-year period from September 1st 2019 to October 30th 2020.

1.1 Data acquired from floats

1.1.1 Active floats for the last 12 months

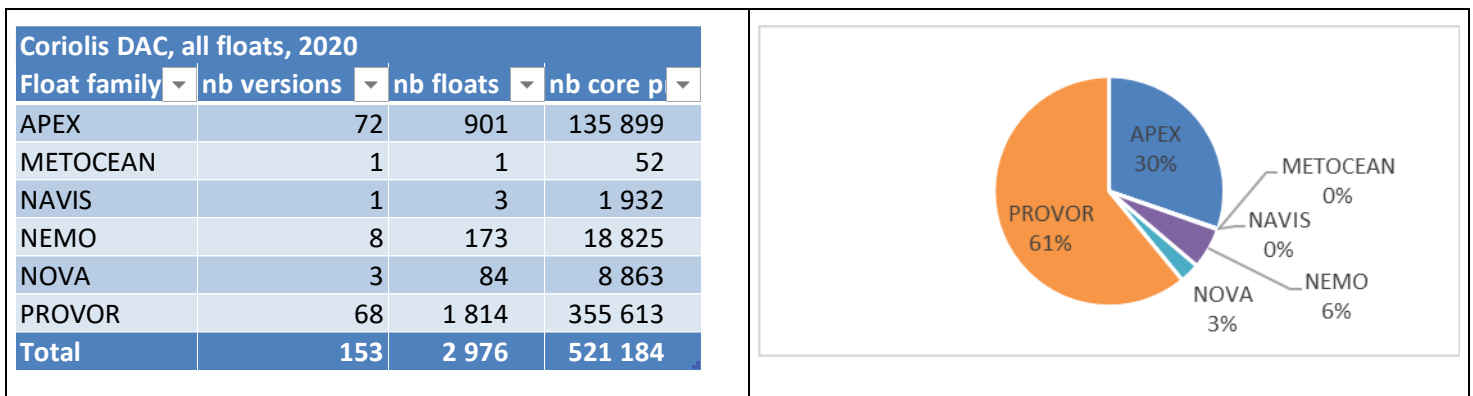
These last 12 months, **52.160 profiles from 787 active floats** were collected, controlled and distributed. Compared to 2019, **the number of profiles is significantly increasing (+49%), the number of floats decreased by 5%**. These figures illustrate a good momentum in Coriolis DAC activity, although the floats small decrease may be related to delayed deployments (maybe related to COVID-19).

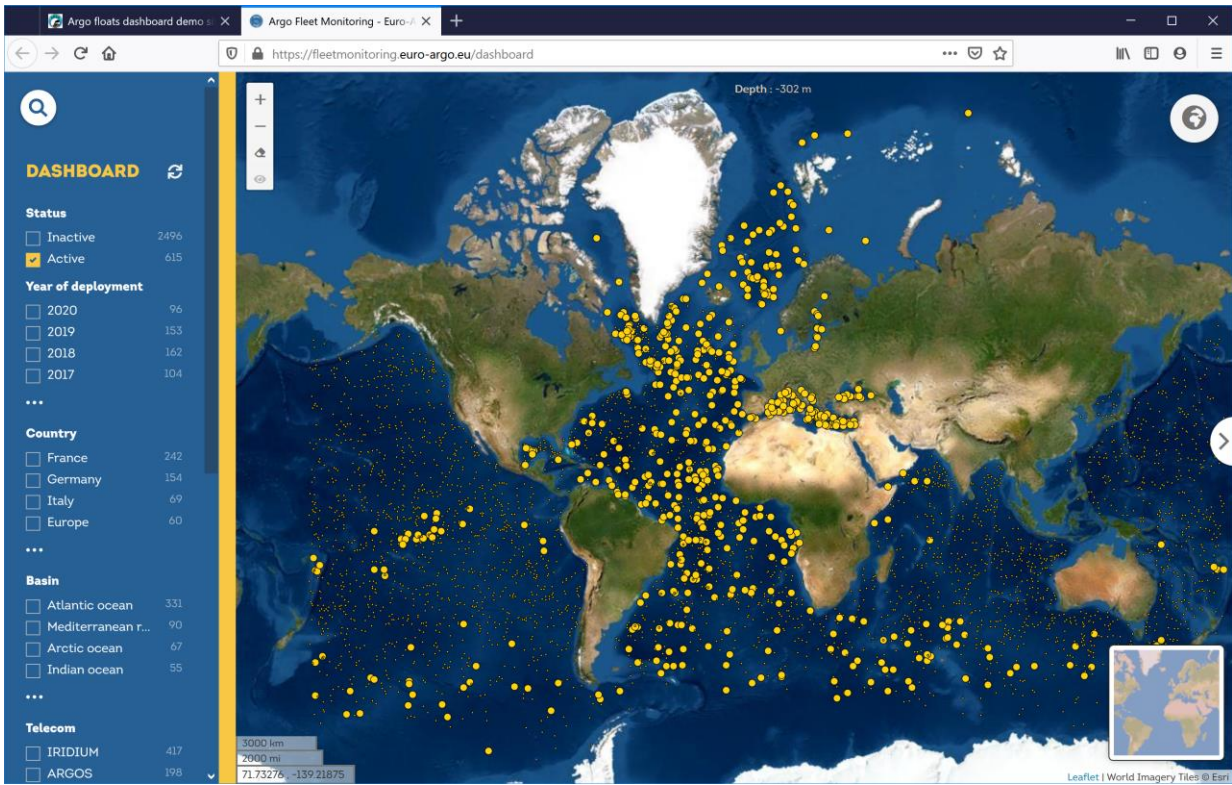
The 787 floats managed during that period had 48 versions of data formats.



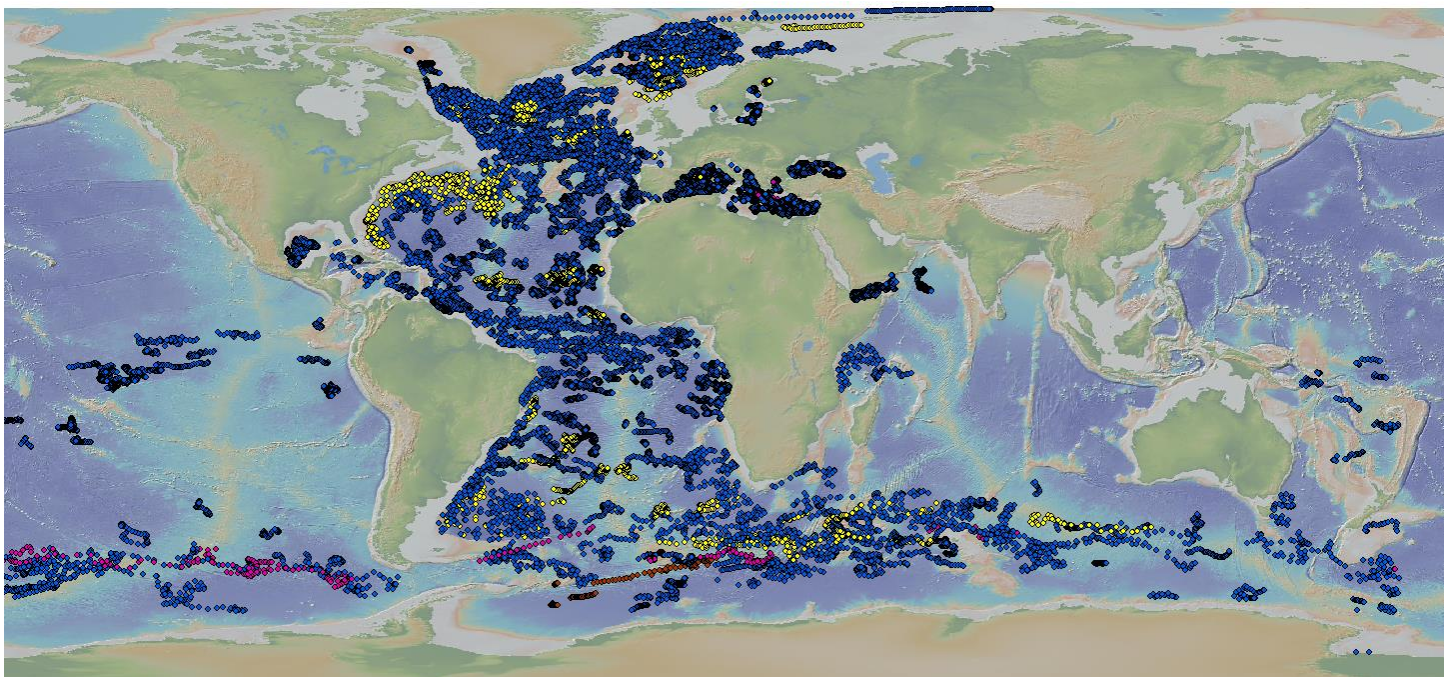
1.1.2 All floats managed by Coriolis DAC

Coriolis DAC manages a total of 2.976 floats with 153 versions, from 6 families. These floats reported 521.184core Argo vertical profiles.

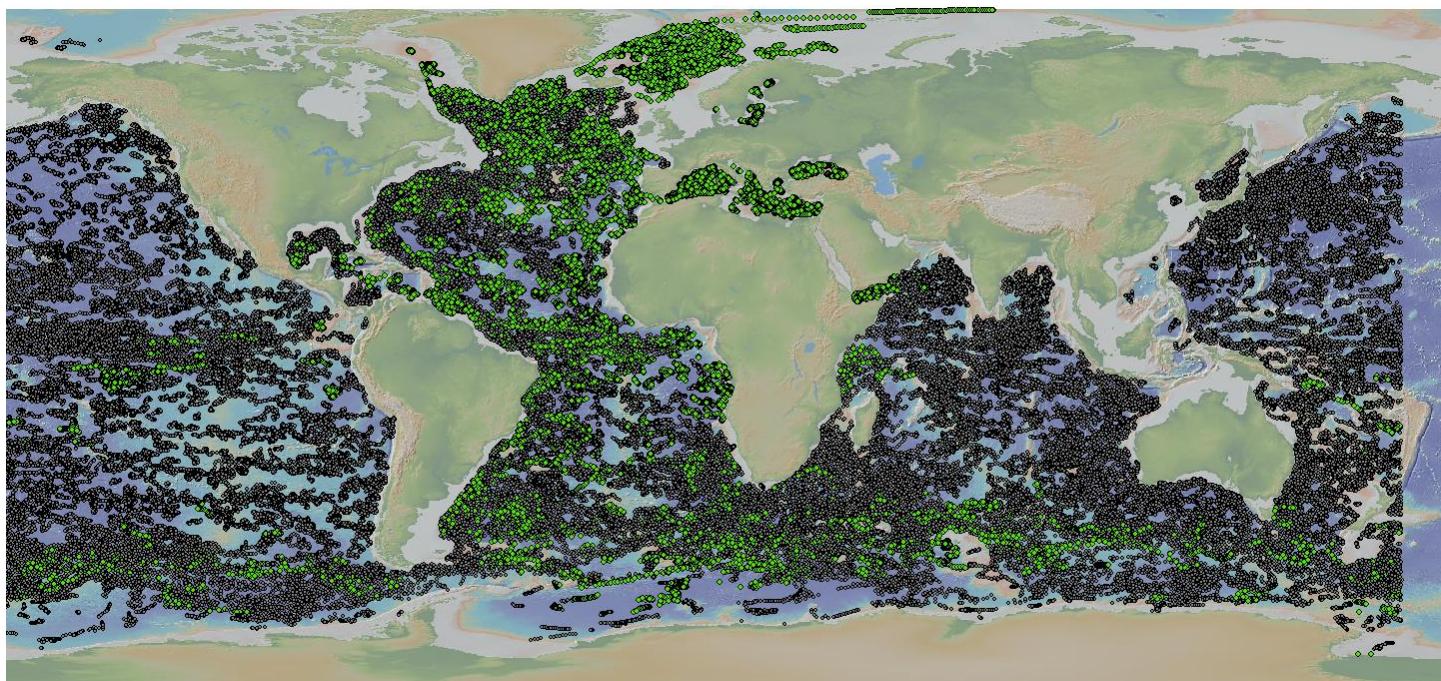




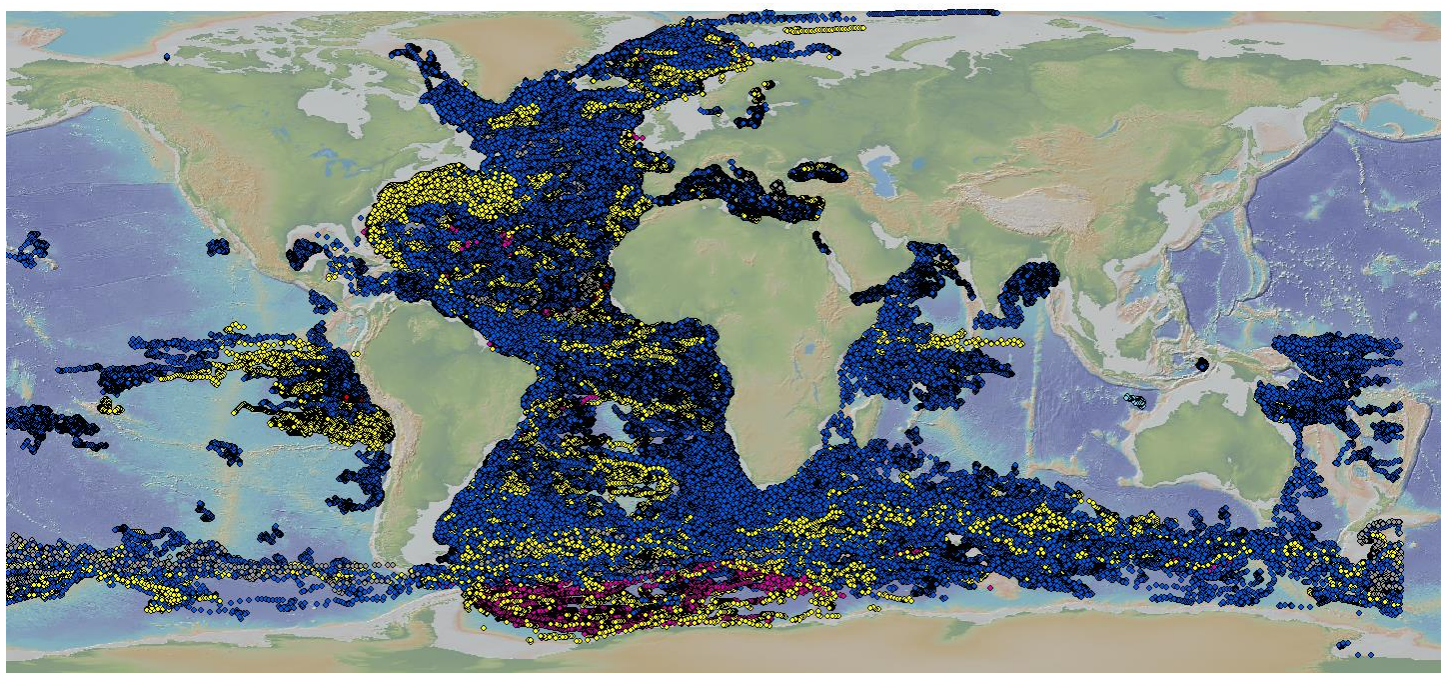
Map of the active floats on November 11th decoded by Coriolis DAC, among others DACs (small dots) as displayed on Euro-Argo floats dashboard <https://fleetmonitoring.euro-argo.eu/dashboard>



Map of the 52 160 profiles from 787 active floats decoded by Coriolis DAC this current year
Apex **Nova** **Provor** **Nemo**



Map of the profiles from active floats decoded by Coriolis DAC this current year, among the other DAC's profiles (Coriolis: green, other DACs: grey)



Map of the 521.184 profiles from 2.976 floats managed by Coriolis DAC

Apex Metocean Navis Nemo Nova Provor

1.1.3 BGC-Argo sensors on Coriolis floats

The data processing chain for data and metadata from Coriolis BGC-Argo floats is continuously improved. These are advanced types of floats performing bio-geo-chemical (BGC) measurements.

Coriolis DAC manages 513 BGC-Argo floats from 5 families. They performed 68.978 cycles.

The data processing chain is freely available:

- Coriolis Argo floats data processing chain, <http://doi.org/10.17882/45589>

Oxygen data reprocessing

In 2020, the Oxygen manual was updated: “Processing Argo oxygen data at the DAC level cookbook”
<http://doi.org/10.13155/39795>

BGC-Argo floats processed by Coriolis DAC				
Float family	nb versions	nb floats	nb profile	nb cycles
APEX	27	110	15 442	13 703
NAVIS	1	3	551	551
NEMO	1	2	297	297
NOVA	1	15	1 130	1 105
PROVOR	38	383	148 246	53 322
Total	68	513	165 666	68 978

The pie chart illustrates the distribution of BGC-Argo floats across five families. PROVOR is the most prevalent family, accounting for 75% of the total. APEX follows at 21%, NOVA at 3%, NAVIS at 1%, and NEMO at 0%.

General characteristics

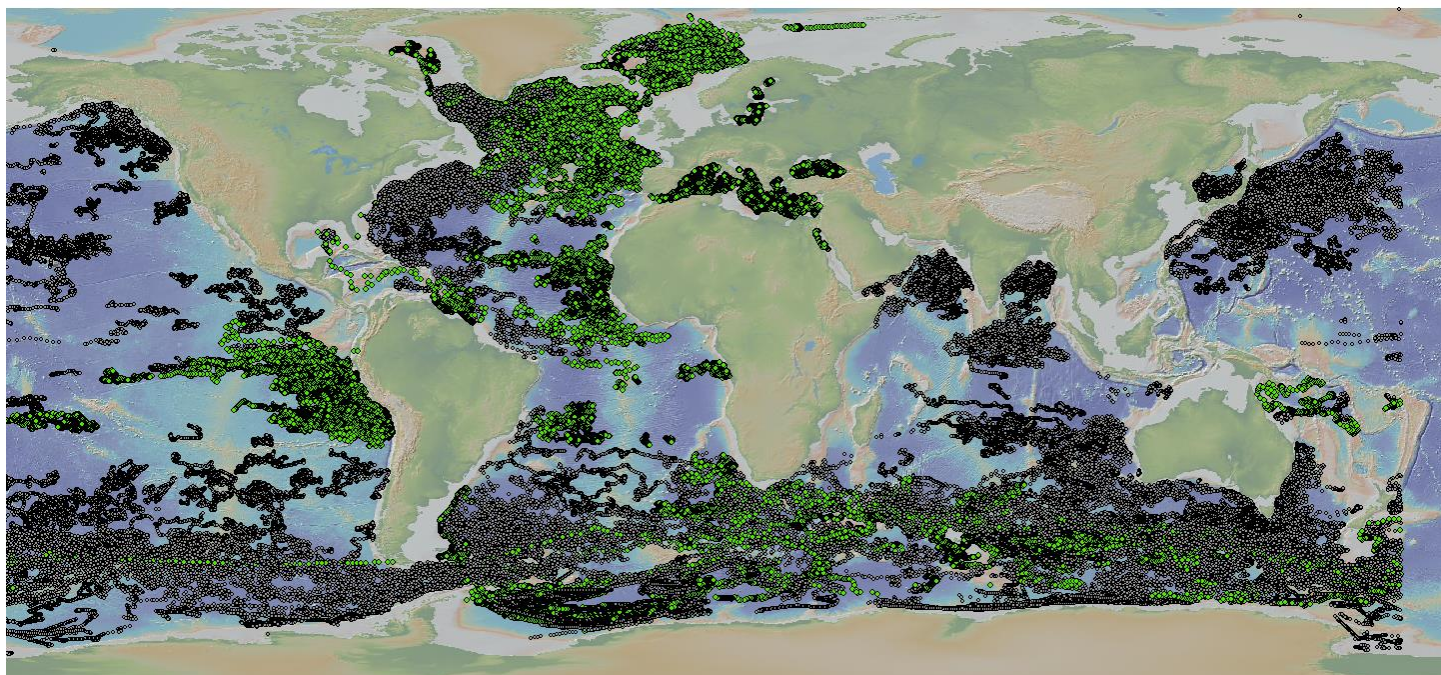
- Iridium sbd or rudics bi-directional communication or Argos
- Fourteen sensors are fitted on the floats
- Eleven BGC parameters reported

Coriolis BGC-Argo floats sensor	nb floats	nb profiles
AANDERAA_OPTODE_4330	298	39997
SATLANTIC_OCR504_ICSW	166	120504
ECO_FLBBCD	163	89820
AANDERAA_OPTODE_3830	77	10801
SUNA_V2	73	10933
SBE63_OPTODE	20	1885
ECO_FLBB_AP2	19	4982
C_ROVER	15	4449
SBE43F_IDO	13	1596
ECO_FLNTU	10	5366
SEAFET	8	409
ECO_FLBB2	4	2112
FLBB	2	616
UVP6-LP	1	30

The 14 types of sensors mounted on Coriolis BGC-Argo floats

BGC parameter	nb files
DOXY	177791
CHLA	71315
BBP700	69206
NITRATE	35052
DOWN_IRRADIANCE412	33999
CDOM	32427
PH_IN_SITU_TOTAL	13555
CP660	4849
TURBIDITY	2109
UP_RADIANCE555	619
TRANSMITTANCE_PARTICLE_BEAM_	426

The 11 main BGC parameters reported by Coriolis BGC-Argo floats



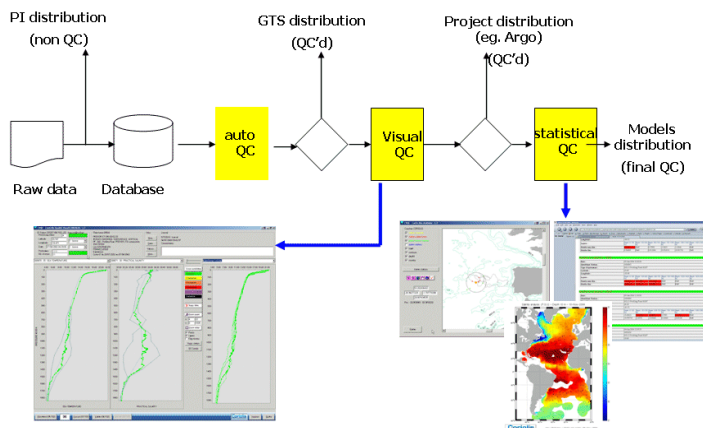
Map of the 513 BGC-Argo floats managed by Coriolis DAC (grey dots: the others DACs bio-Argo floats). They measure parameters such as oxygen, chlorophyll, turbidity, CDOM, back-scattering, UV, nitrate, bisulfide, pH, radiance, irradiance, PAR.

1.2 Data issued to GTS

Vertical profiles processed by Coriolis are distributed on the GTS by way of Meteo-France. This operation is fully automated. After applying the automatic Argo QC procedure, the Argo profiles are inserted on the GTS every hour. The profile files are sent as BUFR messages.

Vertical profiles are distributed on GTS if they are less than 30 days old. Once a day, floats data are checked with ISAS objective analysis that triggers alerts and visual inspection for suspicious observations. The corrected data are not redistributed on GTS.

In July 2019, Coriolis stopped the TESAC messages distribution; only BUFR messages are now distributed.

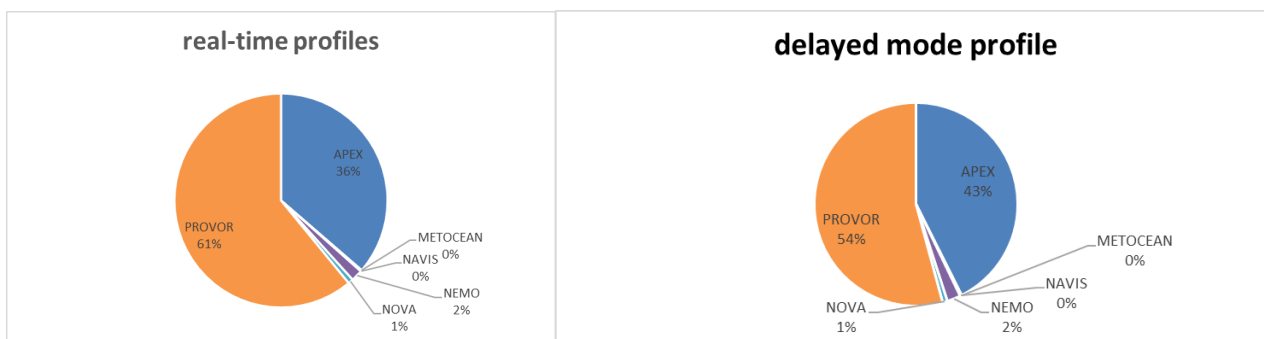
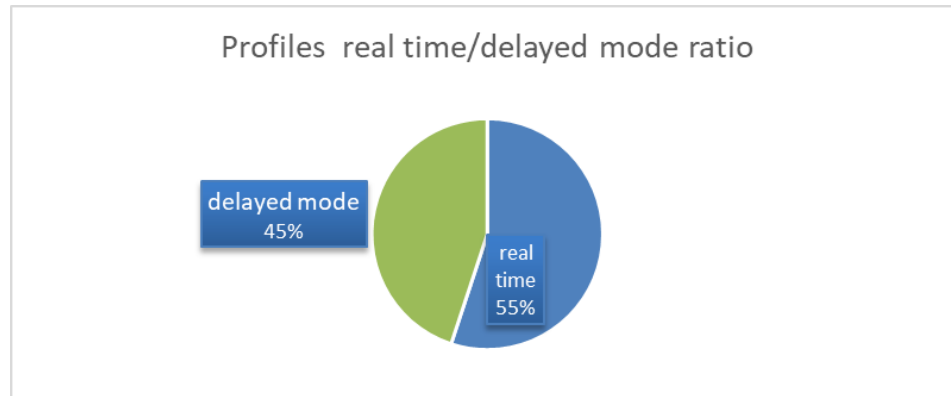


Coriolis DAC Argo data flow

1.3 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.

All Coriolis floats, number of profile files on GDAC				
Family	nb floats	nb profile	RT profile	DM profil
APEX	841	36 742	99157	94879
METOCÉAN	1	-	52	52
NAVIS	2	1 411	521	521
NEMO	102	14 084	4741	4774
NOVA	38	6 721	2142	1694
PROVOR	1337	189 212	166418	120721
Total	2 321	248 170	273 031	222 641



Distribution of Coriolis DAC real-time and delayed mode profiles

1.4 Data issued for delayed mode QC

Delayed mode profiles

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.

1.5 Delayed mode data sent to GDACs

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

- A total of **84.272 new or updated delayed mode profiles** was sent to GDACs this year.
- **A total of 222.641 delayed mode profiles** were sent to GDACs since 2005.

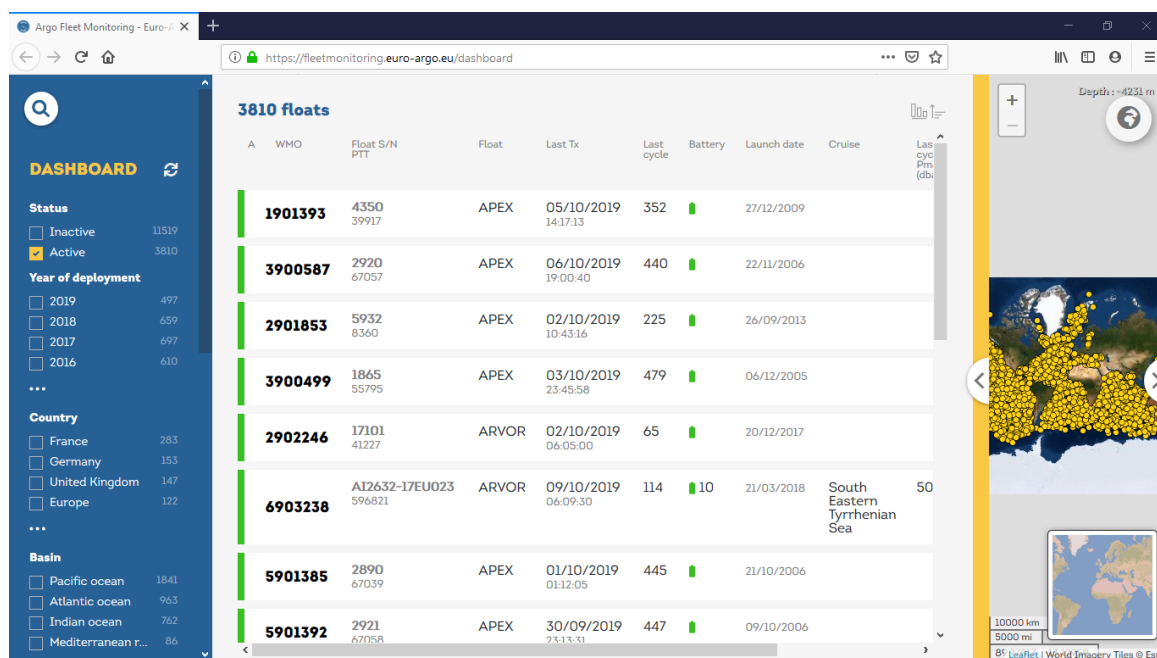
The number of delayed mode profiles decreased by 27% this year compared to 2019.

1.6 Web pages

1.6.1 Argo dashboard

The Argo floats dashboard developed in 2019 by Coriolis team is available at:

- <https://fleetmonitoring.euro-argo.eu/dashboard>

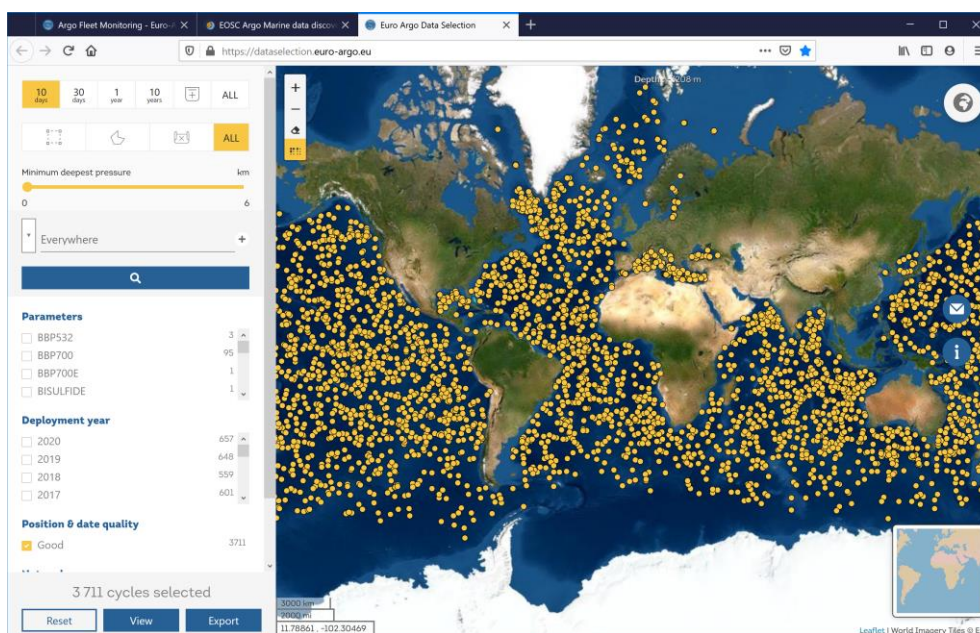


It displays all Argo floats, with faceted interrogations and instantaneous answers. The dashboard is developed on cloud and big-data techniques.

- Cloud techniques: a metadata and a data APIs, opened to internet machine to machine queries
- Big-data techniques: Argo metadata are hourly indexed in an Elasticsearch index, Argo data are hourly indexed in a Cassandra data base. Elasticsearch and Cassandra allows instant answers on dataset having billions of observations.

The Argo data selection was developed in 2020. The initial version is online at <https://dataselection.euro-argo.eu/>

It proposes data discovery with faceted search on temporal and spatial coverage, parameters, deployment years or quality codes. The selected data are downloadable in NetCDF and CSV formats.



Argo data selection, the initial version is online <https://dataselection.euro-argo.eu>

1.6.2 Argo data on EU BlueCloud

A collaboration is underway with NASA-JPL and the European Blue Cloud to use the CMC (Common Mapping Client) client as the front office of Argo dashboard to provide in situ – satellite – model integration.

- <http://bluecloud.odatis-ocean.fr>

1.6.3 Interoperability services (ERDDAP API,...)

This web page describes all Argo floats interoperability services from Coriolis:

- <http://www.coriolis.eu.org/Data-Products/Data-Delivery/Argo-floats-interoperability-services2>
 - Argo data through ERDDAP data server (www.ifremer.fr/erddap)
 - Display an individual float's data and metadata in HTML or XML format
 - Display all Argo floats, display a group of floats
 - Argo profiles and trajectories data selection (HTML or XML)
 - All individual float's metadata, profile data, trajectory data and technical data
 - Argo profiles data on OpenDAP, OGC-WCS and http
 - Argo data through Oceanotron data server
 - Argo profiles data through GCMD-DIF protocol
 - Argo data through RDF and OpenSearch protocols
 - Display Argo profiles and trajectories with GoogleEarth

1.6.4 Data centre activity monitoring

Coriolis operators perform an activity monitoring with an online control board.

Fonction	Description	Etat J	Etat J-1	Etat J-2	Etat J-3	Dernière exécution (TU)
CO-01-01-13	Synchronisation des QC de la base et du DAC Coriolis	●	●	●	●	OK 2020-11-11T12:41:31Z
CO-01-07-01-aoml	Collecte Argo DAC - aoml	●	●	●	●	OK 2020-11-11T13:01:04Z
CO-01-07-01-bodc	Collecte Argo DAC - bodc	●	●	●	●	OK 2020-11-11T13:02:05Z
CO-01-07-01-coriolis	Collecte Argo DAC - coriolis	●	●	●	●	OK 2020-11-11T13:04:02Z
CO-01-07-01-csio	Collecte Argo DAC - csio	●	●	●	●	OK 2020-11-11T13:04:03Z
CO-01-07-01-csiro	Collecte Argo DAC - csiro	●	●	●	●	OK 2020-11-11T13:05:05Z
CO-01-07-01-inctis	Collecte Argo DAC - inctis	●	●	●	●	OK 2020-11-11T13:06:11Z
CO-01-07-01-jma	Collecte Argo DAC - jma	●	●	●	●	OK 2020-11-11T13:07:09Z
CO-01-07-01-kma	Collecte Argo DAC - kma	●	●	●	●	OK 2020-11-11T13:08:02Z
CO-01-07-01-kordi	Collecte Argo DAC - kordi	●	●	●	●	OK 2020-11-11T13:09:03Z
CO-01-07-01-meds	Collecte Argo DAC - meds	●	●	●	●	OK 2020-11-11T13:10:02Z
CO-01-07-01-nmdis	Collecte Argo DAC - nmdis	●	●	●	●	OK 2020-11-11T13:11:02Z
CO-01-07-01-02	Collecte Argo DAC - table argo index profiles	●	●	●	●	OK 2020-11-11T13:11:07Z
CO-01-07-01-03	Collecte Argo DAC - resubmit files cause meta missing	●	●	●	●	OK 2020-11-10T14:51:02Z
CO-01-07-02	Synchronisation Argo GDAC US IF	●	●	●	●	OK 2020-11-11T05:09:19Z
CO-01-07-03	Collecte Argo DAC fte	●	●	●	●	OK 2020-11-11T13:09:05Z
CO-01-07-04	Synchronisation Argo GDAC BDD	●	●	●	●	OK 2020-11-10T09:19:11Z
CO-01-07-06-aoml	Collecte Argo DAC BDD - aoml	●	●	●	●	OK 2020-11-11T13:12:04Z
CO-01-07-06-bodc	Collecte Argo DAC BDD - bodc	●	●	●	●	OK 2020-11-11T13:13:03Z
CO-01-07-06-coriolis	Collecte Argo DAC BDD - coriolis	●	●	●	●	OK 2020-11-11T12:44:03Z
CO-01-07-06-csio	Collecte Argo DAC BDD - csio	●	●	●	●	OK 2020-11-11T12:45:06Z
CO-01-07-06-csiro	Collecte Argo DAC BDD - csiro	●	●	●	●	OK 2020-11-11T12:46:02Z
CO-01-07-06-dm	Collecte Argo DAC BDD - reprocessing (irrégulière)	●	●	●	●	OK 2020-11-02T18:30:30Z

Argo GDAC operations monitoring: every working day, an operator performs diagnostics and take actions on anomalies (red or orange smileys)

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

Operational oceanography models; all floats data are distributed to:

- EU Copernicus Marine service models (Mercator, Foam, Topaz, Moon, Noos, Boos)
- French model Soap (navy operational model)

Argo projects: this year, Coriolis data centre performed float data management for **72 Argo scientific projects and 52 PIs (Principal Investigators)**.

List of Coriolis scientific PIs and project names

project	nb floats
coriolis	210
argo-bsh	166
euro-argo	155
argo italy	71
mocca	61
naos wp1	56
mocca-eu	55
naos	53
pirata	30
argo spain	28

Top 10 of Coriolis DAC projects having active floats

pi	nb active flo
birgit klein	172
pierre-marie poulain	92
christine coatanoan	73
virginie thierry	52
sabrina speich	35
kjell arne mork	33
bernard bourles	33
romain cancouet	26
pedro velez	26
damien desbruyeres	20
fabrizio d'ortenzio	16
waldemar walczowski	16

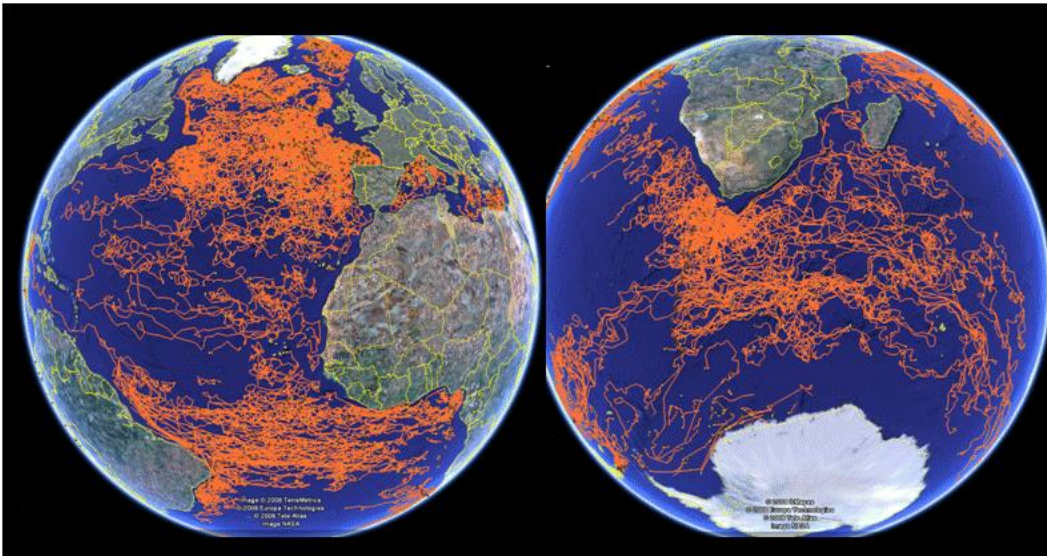
Top 10 of Principal Investigators (PI) in charge of active floats

1.8 Products generated from Argo data

Sub-surface currents ANDRO Atlas

Based on Argo trajectory data, Ifremer and CNRS team are regularly improving the “Andro” atlas of deep ocean currents. The ANDRO project provides a world sub-surface displacement data set based on Argo floats data. The description of each processing step applied on float data can be found in:

- Ollitrault Michel, Rannou Philippe (2019). **ANDRO: An Argo-based deep displacement dataset**. SEANOE. <http://doi.org/10.17882/47077>



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

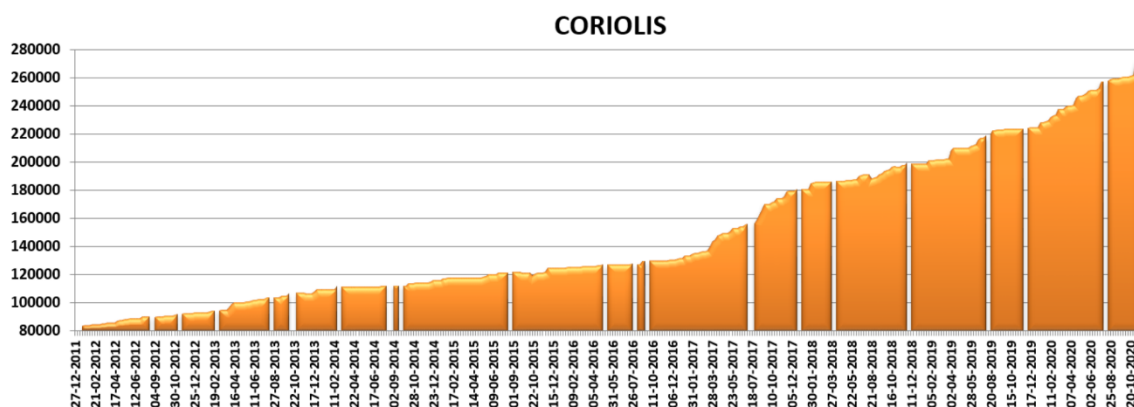
2 Delayed Mode QC

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 each 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 largely improve the amount of delayed mode profiles.

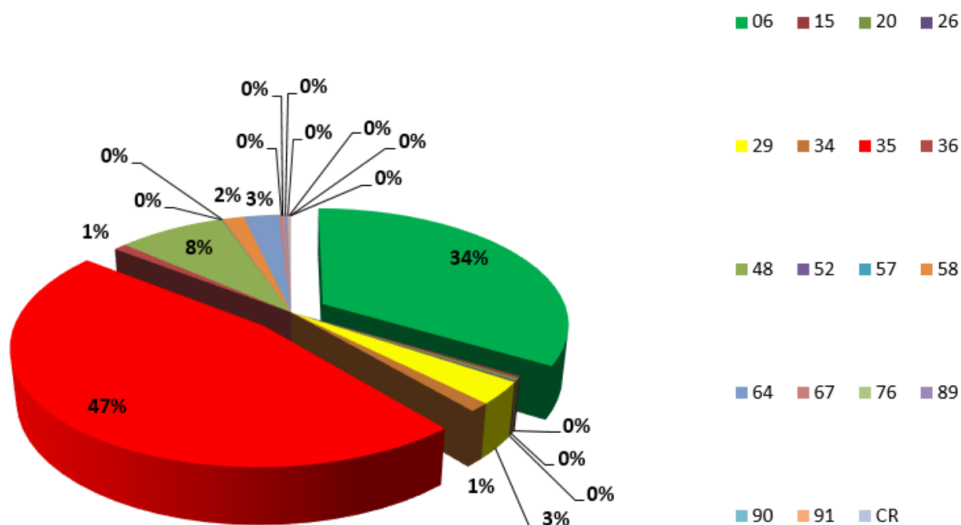
A lot of work is always done from BSH (Birgit Klein) taking into account also floats from other German institutes and OGS (Giulio Notarstefano) for the MedSea.

In the last 3 years, an important effort has been dedicated to improve the delayed mode quality control status.



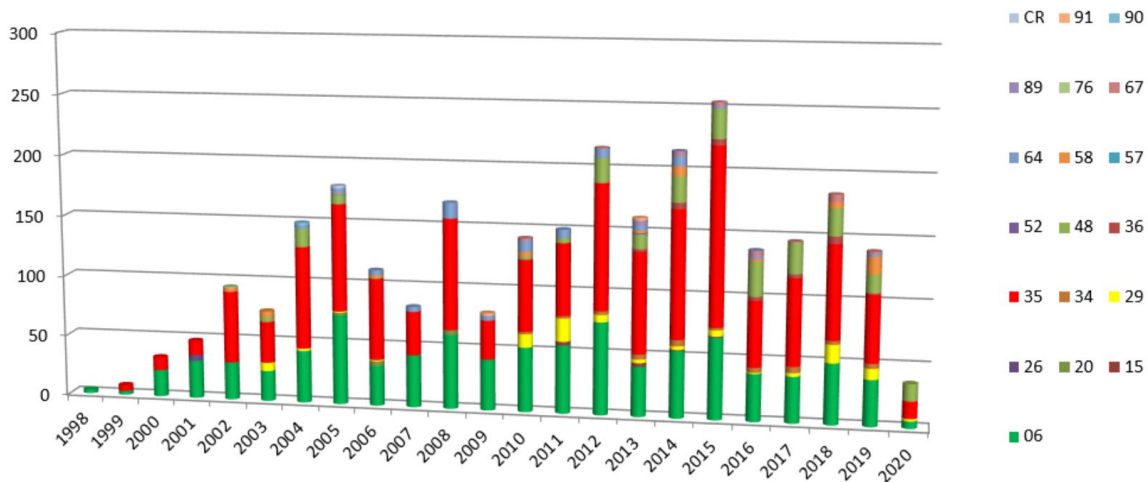
Evolution of the DM profiles' submission versus dates

Floats by country



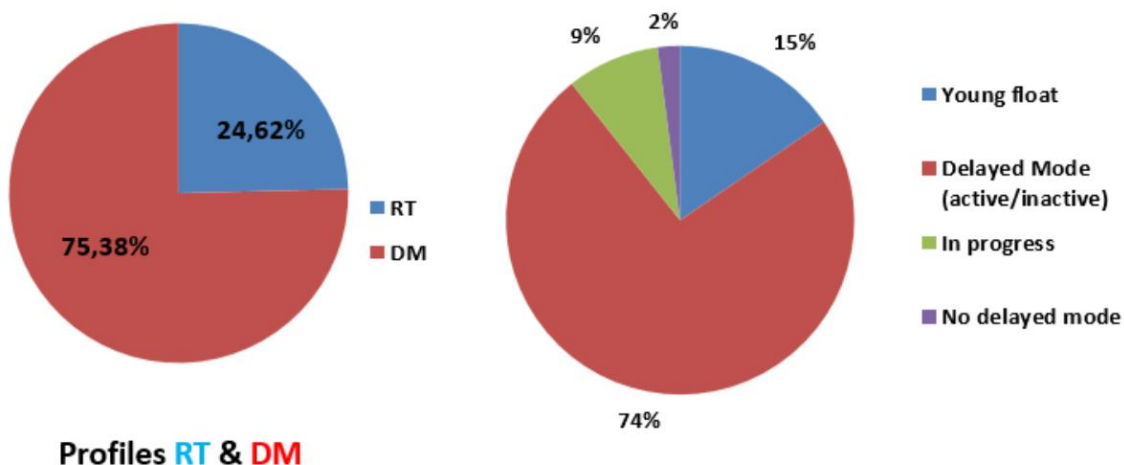
Percentage of floats by country in the Coriolis DAC.

Codes for the countries: 06 : Germany - 15 : Bulgaria - 20 : Chili - 26 : Denmark - 29 : Spain - 34 : Finland - 35 : France - 36 : Greece - 48 : Italy - 52 : Lebanon - 57 : Mexico - 58 : Norway - 64 : Netherlands - 67 : Poland - 76 : China - 89: Turkey - 90 : Russia - 91 : - South Africa - CR : Costa Rica



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

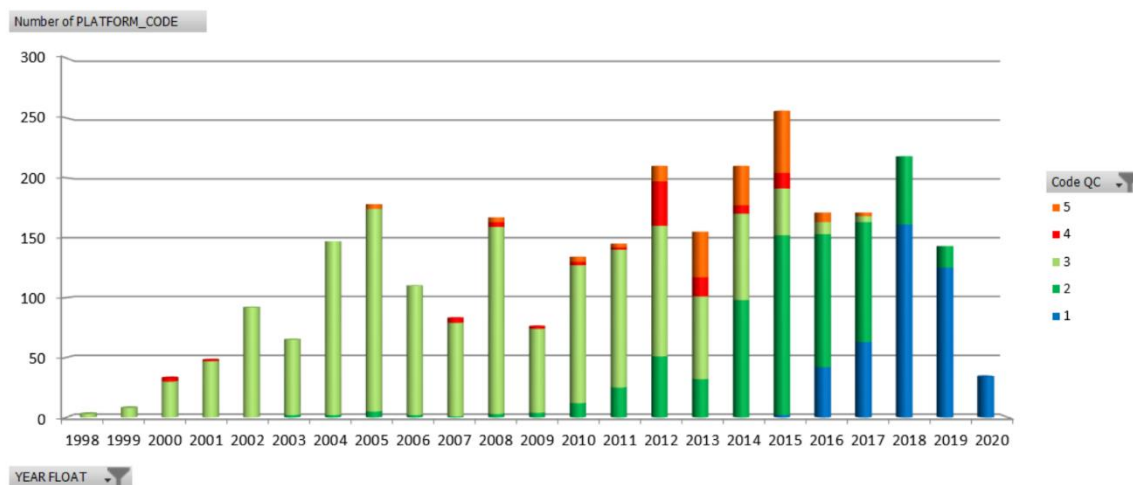
During the last year (from October 2019 to November 2020), 50641 new delayed mode profiles were produced and validated by PIs. A total of 273779 delayed mode profiles were produced and validated since 2005.



Status of the floats processed by Coriolis DAC.

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

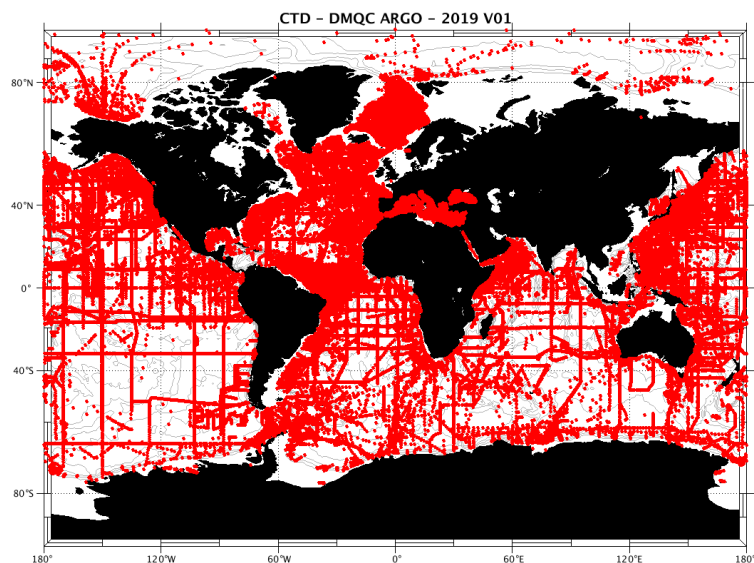
The status of the quality control done on the Coriolis floats is presented in the following plot. For the two last years (2019-2020), most of the floats are still too young (code 1) to be performed in delayed mode. For the years 2012-2013-2014, we are still working on the DMQC of some floats. The codes 2 and 3 show the delayed mode profiles for respectively active and dead floats.



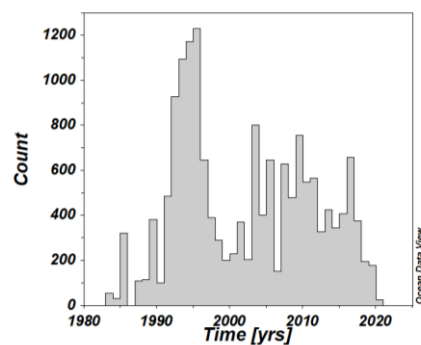
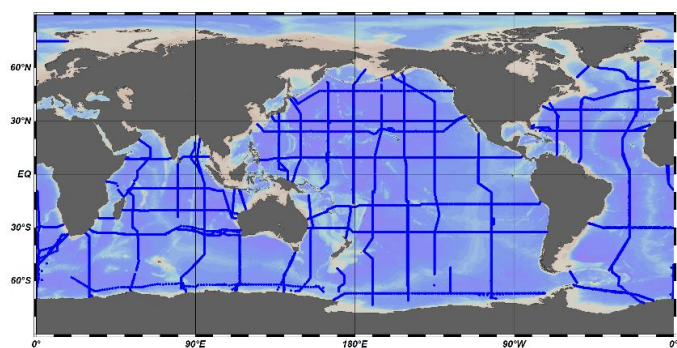
Status of the quality control done on profiles sorted by launch's year, code 1: young float, code 2: active float, DM done, code 3 : dead float, DM done; code 4 : DM in progress, code 5 : waiting for DM, code 6 : problems with float.

Reference database

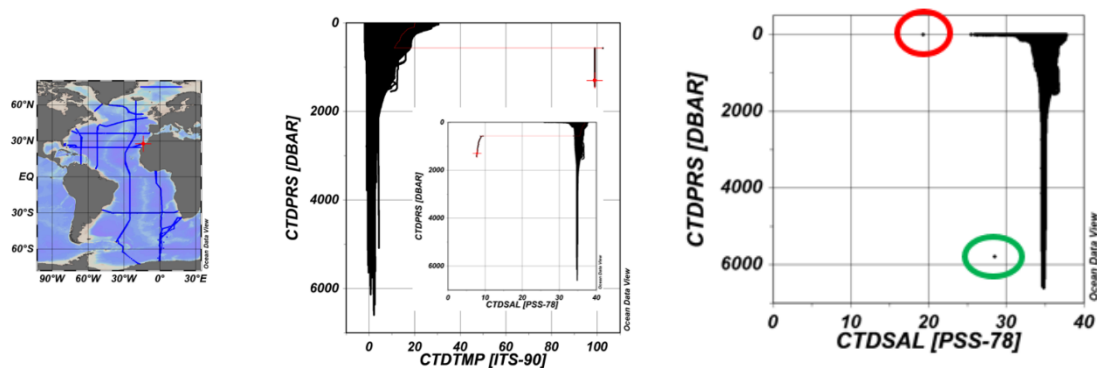
At the end of 2019, an updated version 2019V01 has been provided with some updates on a few boxes and including including CCHDO, OCL and ICES updates, CTD from PI, correction from feedbacks.



For the next version 2020V01, an action was to incorporate high QC GO-SHIP product profiles selected for DEEP-ref-DB to Core ref db. This product was made available in mid-September (16231).



A first check has shown some anomalies in the dataset that could not be integrated as is in the dataset and needed thorough quality control. Some anomalies were easy to detect and others needed more scrutinized analysis.



After the quality control, the next step that is in progress is to replace the profiles, which already exist in the Core Ref DB with the data from DEEP-ref-DB and add the new ones.

In addition, Ingrid Angel Benavides (BSH) worked on the clean-up of the CTD reference database in the Atlantic Ocean, Arctic and nordic seas, removal of out of range or incomplete samples, and duplicate checks. New mat files must be sent to Coriolis to update the reference database.

The last version is divided in smaller tar balls, one by wmo box area (1-3-5-7): for instance, CTD_for_DMQC_2019V01_1.tar.gz for all boxes starting with wmo 1, then we will have 4 tar files.

3 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)

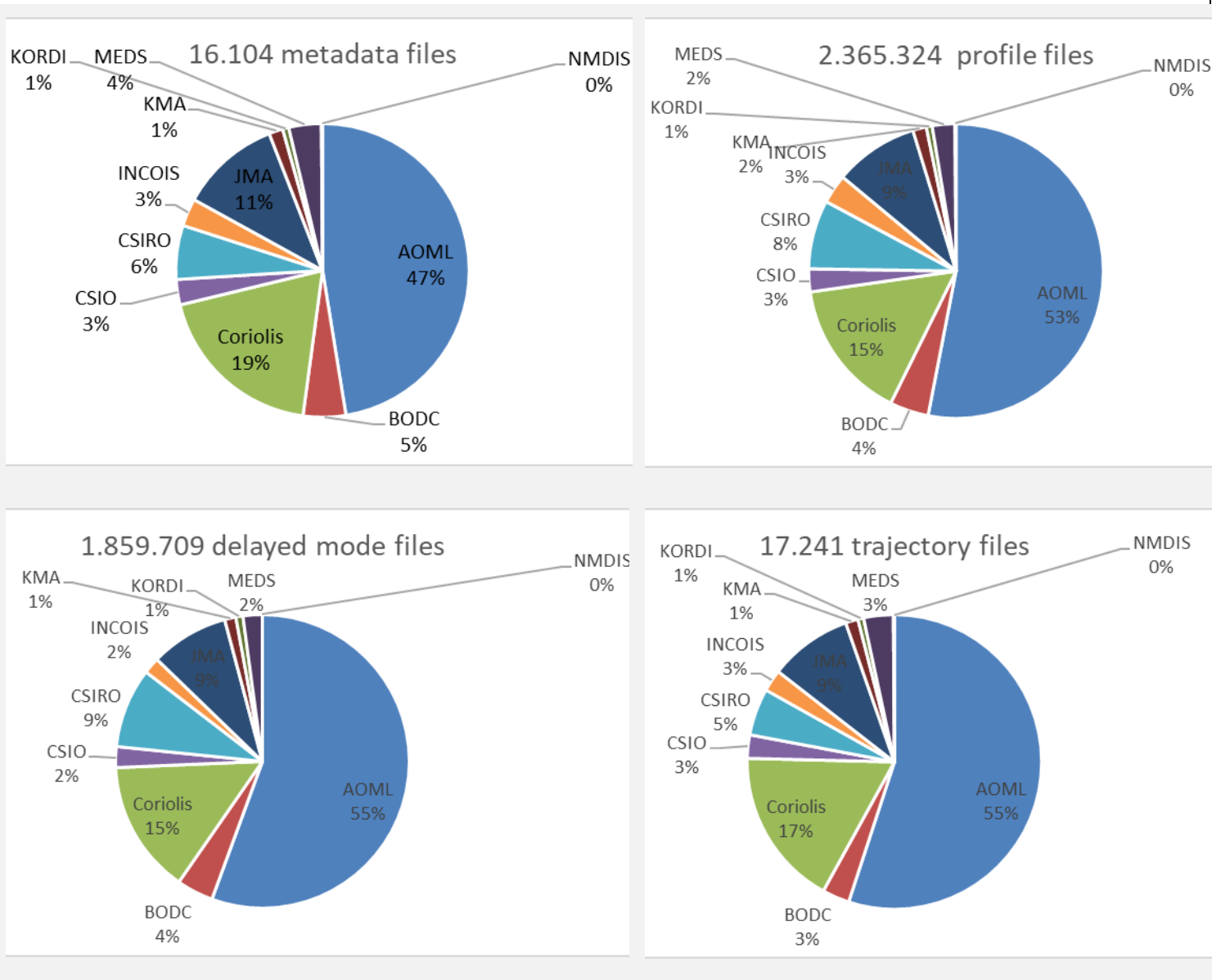
- National centres reporting to you
- Operations of the ftp server
- Operations of the www server
- Data synchronization
- Statistics of Argo data usage : Ftp and WWW access, characterization of users (countries, field of interest : operational models, scientific applications) ...

3.1 National centres reporting to you

Currently, 11 national DACs submit regularly data to Coriolis GDAC. On November 2020, the following files were available from the GDAC FTP site.

3.1.1 GDAC files distribution

DAC	metadata				delayed mode		trajectory	
	files 2020	increase	profile files	increase2	profile files	increase3	files 2019	increase4
AOML	7 640	6%	1 254 815	9%	1 033 917	17%	9 488	6%
BODC	756	6%	100 951	12%	75 251	65%	516	1%
Coriolis	3 073	5%	364 708	12%	273 031	23%	2 994	6%
CSIO	449	10%	59 538	8%	42 805	181%	443	10%
CSIRO	958	10%	179 229	9%	163 731	11%	891	8%
INCOIS	491	3%	75 564	9%	33 712	10%	413	0%
JMA	1 784	4%	218 511	6%	160 237	8%	1 571	2%
KMA	247	2%	35 325	5%	23 094	0%	238	3%
KORDI	109	0%	15 330	-3%	14 505	0%	107	0%
MEDS	578	6%	58 893	7%	39 381	3%	561	6%
NMDIS	19	0%	2 460	0%	45	-	19	0%
Total	16 104	6%	2 365 324	9%	1 859 709	18%	17 241	5%



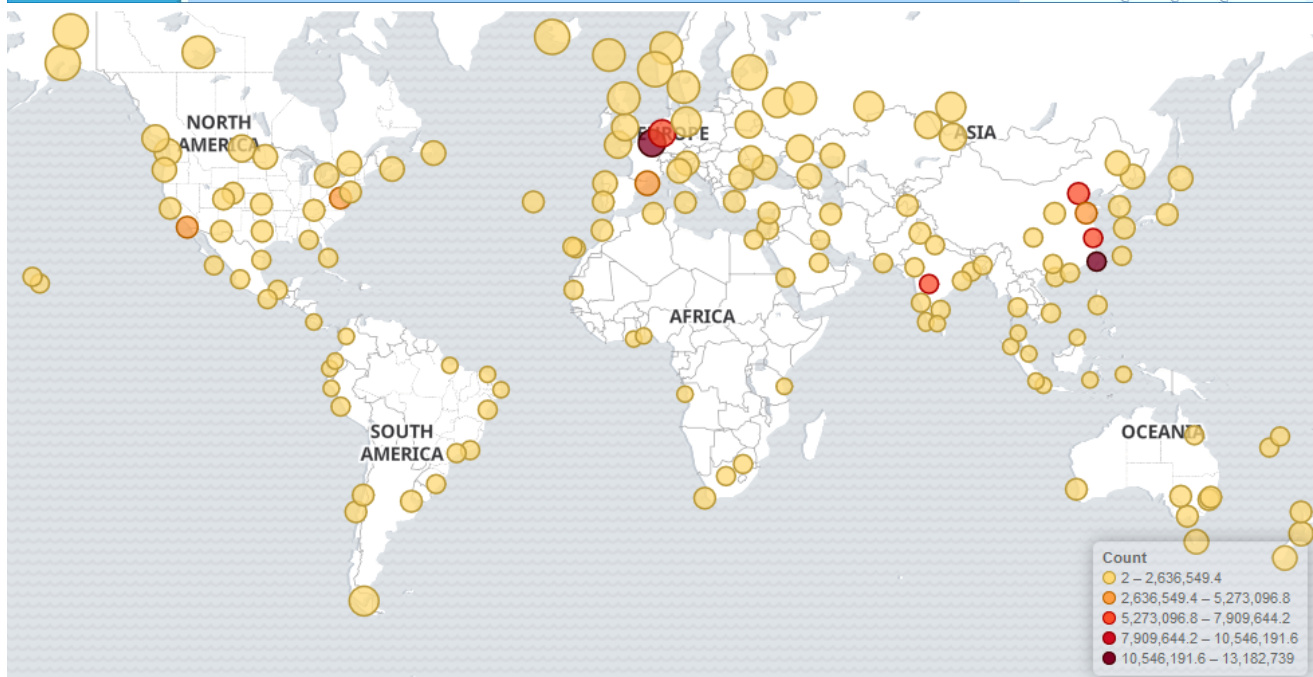
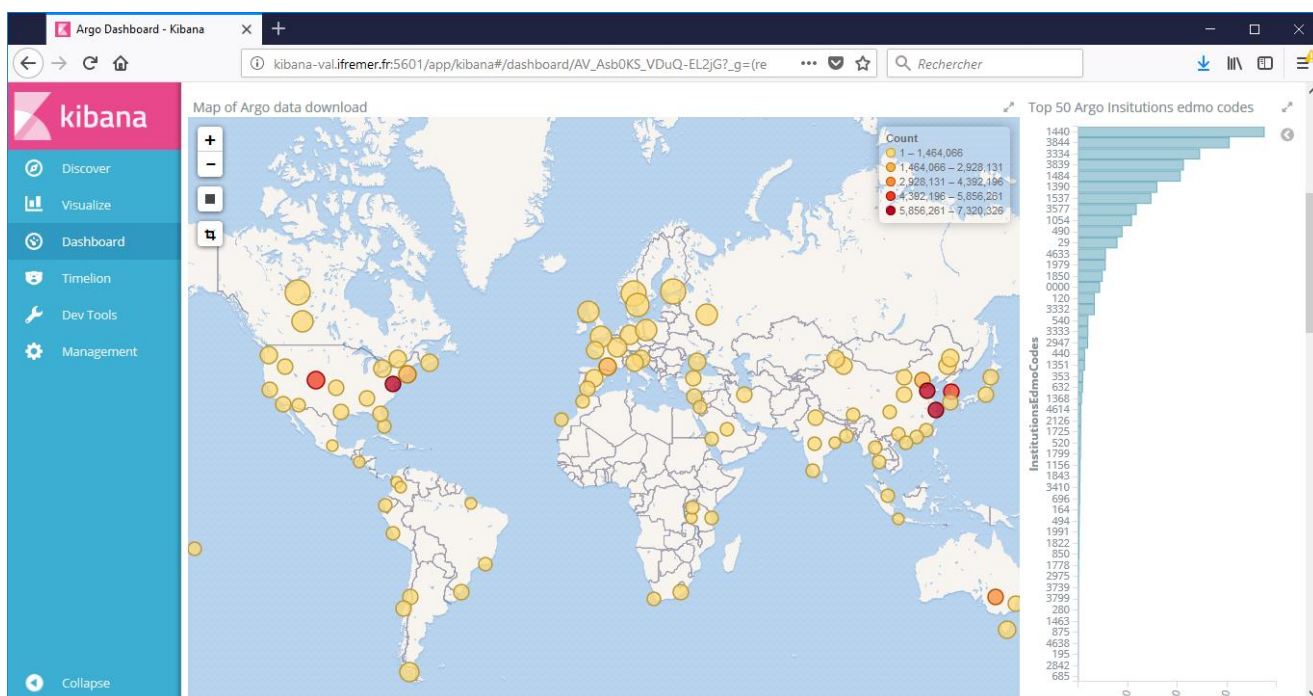
Number of files available on GDAC, November 2020

3.1.2 Argo Semaphore dashboard: give credit to data providers

Within EU AtlantOS project, Ifremer is setting up a dashboard (Semaphore) to monitor data distribution and give credit to data providers such as Argo floats.

FTP downloads log files are ingested in an Elasticsearch index. A link between downloaded files, download originators, floats included in the downloaded files and institution owners of the floats is performed. These links are displayed in a Kibana dashboard.

This dashboard will offer the possibility to give credit to Floats owner institutions such as how many data from one particular institution was downloaded, by whose data users.



Geographical distribution of GDAC ftp downloads in 2018 - 2019

The majority of users (red dots) are located in USA, China, Australia and of course Europe. The right side histogram sorts the floats institution code (1440: PMEL, 3844: WHOI, 3334: INCOIS, 3839: UWA, 1484: CSIRO, ...).



The top 50 of floats institutions downloads and the top 50 of data user's

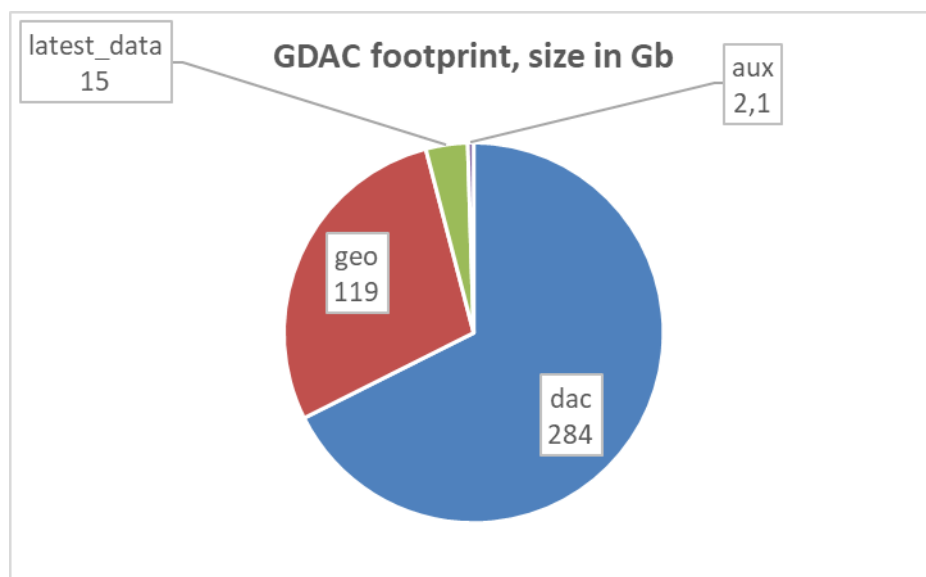
3.1.3 GDAC files size

- The total number of NetCDF files on the GDAC/dac directory was 2 867 119 (+18% in one year)
- The size of GDAC/dac directory was 284Gb (- 163%)
- The size of the GDAC directory was 572Gb (- 4%)

The decrease in size of the GDAC is related to the removal of BGC-Argo M-files, replaced by much more compact S-files. The S-files (and former M-files) are synthetic files merging core-Argo and BGC-Argo profiles.

More on : <http://www.argodatamgt.org/Data-Mgt-Team/News/BGC-Argo-M-prof-files-no-more-distributed-on-GDAC>

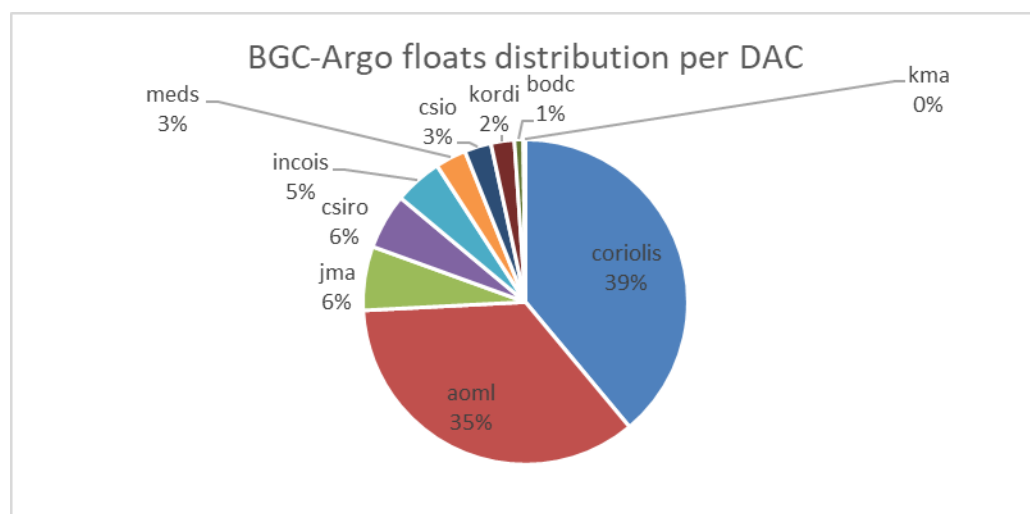
branch	GDAC size in	yearly
dac	284	-16%
geo	119	16%
latest_data	15	-29%
aux	2,1	5%
gdac total	572	-4%

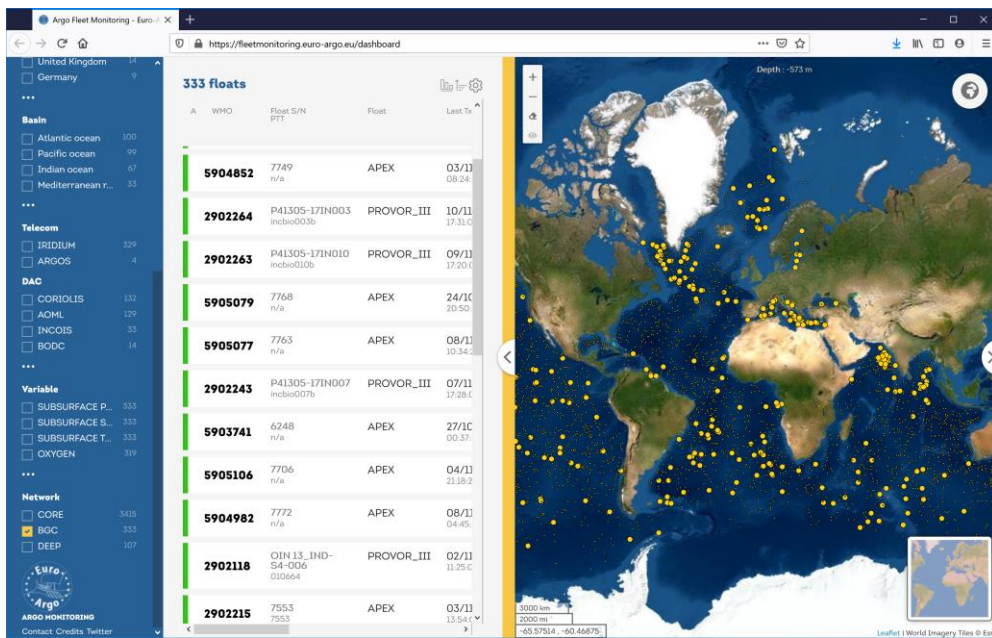


3.1.4 BGC-Argo floats

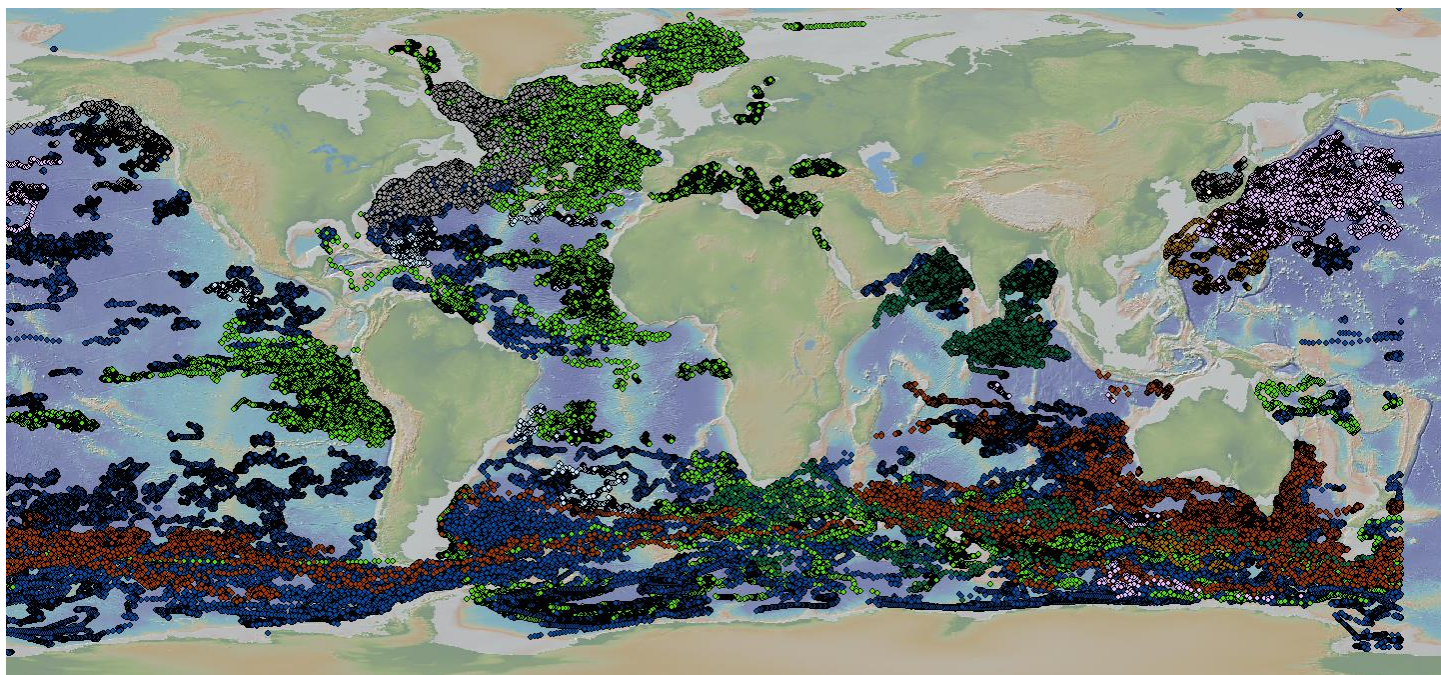
In November 2020, 225.135 BGC-Argo profiles from 1664 floats were available on Argo GDAC. This is a strong increase compared to 2019: +19% more floats and +19% more profiles.

DAC	nb bgc float	nb bgc file
coriolis	570	78 364
aoml	516	77 298
jma	93	16 193
csiro	81	21 726
incois	70	10 360
meds	45	4 537
csio	39	8 330
kordi	34	3 426
bodc	13	4 433
kma	3	468
Total	1464	225 135

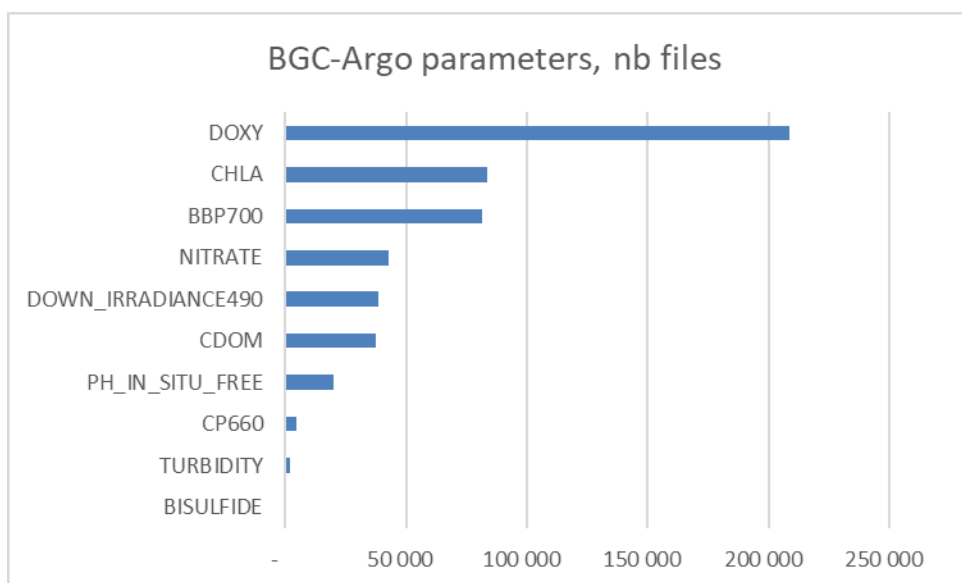




Map of 333 BGC-Argo floats (active: yellow, other: grey) from <https://fleetmonitoring.euro-argo.eu/dashboard>



BGC-Argo profiles, colored by DACs



Main BGC-Argo physical parameters, number of profiles

BGC parameter	nb files
DOXY	208 627
CHLA	83 398
BBP700	81 289
NITRATE	42 983
DOWN_IRRADIANCE490	38 412
CDOM	37 468
PH_IN_SITU_FREE	20 208
CP660	4 853
TURBIDITY	2 109
BISULFIDE	705

3.2 Operations of the ftp and web server

For each individual DAC, every 30 minutes, meta-data, profile, trajectory and technical data files are automatically collected from the national DACs. The 11 DACs are processed in parallel (one process launched every 3 minutes).

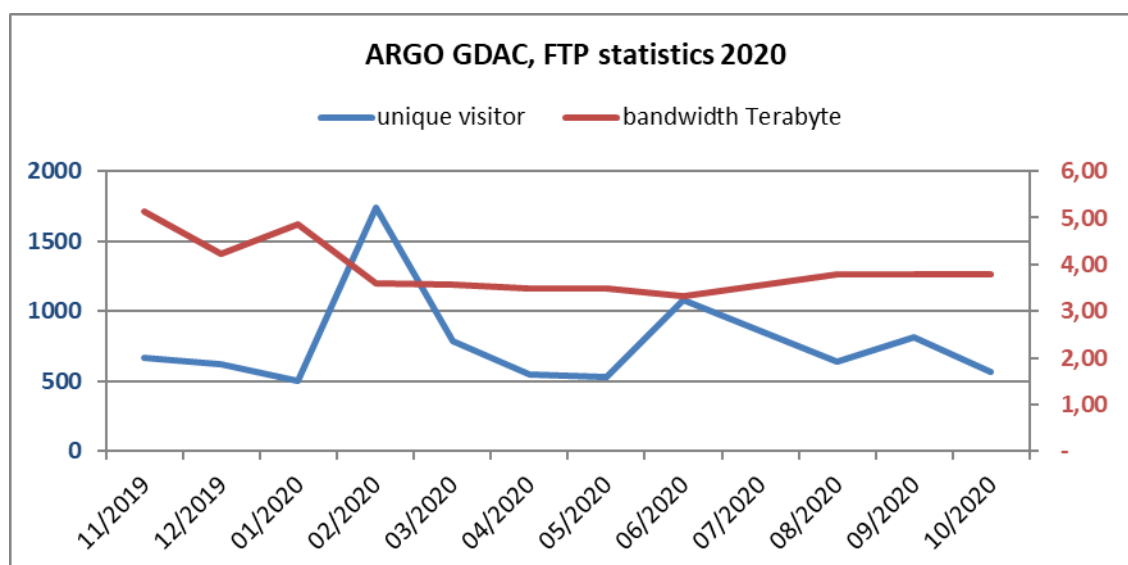
Index files of metadata, profiles, trajectories, technical and auxiliary data are hourly updated.

GDAC ftp address: <ftp://ftp.ifremer.fr/ifremer/argo>

Statistics on the Argo GDAC FTP server: <ftp://ftp.ifremer.fr/ifremer/argo>

There is a monthly average of 848 unique visitors, performing 5322 sessions and downloading 3.88 terabytes of data files.

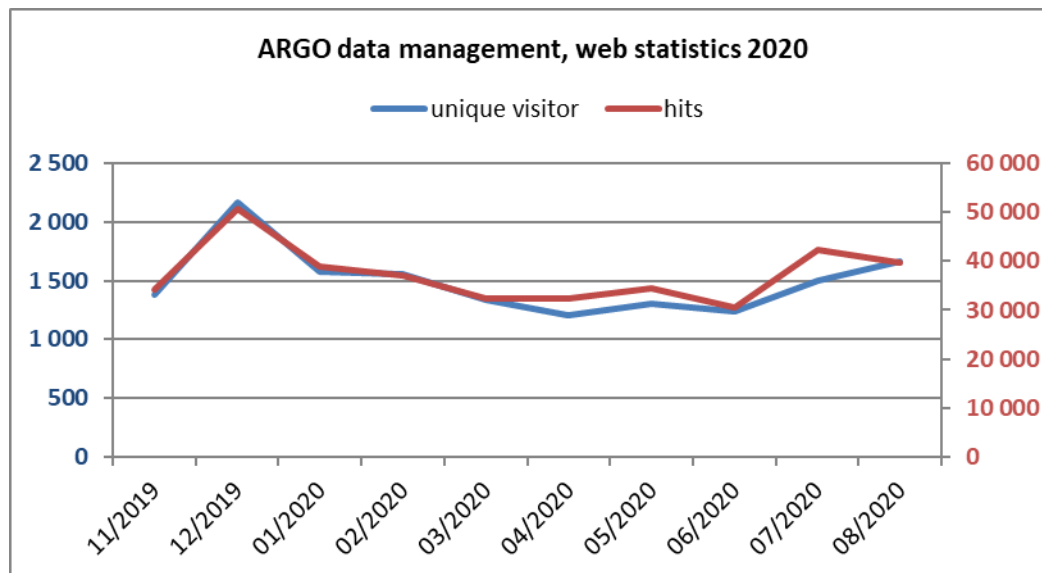
The table below shows an unusual increase of visitors in February on GDAC FTP; we do not have a specific explanation.



ARGO GDAC FTP statistics					
month	unique visitor	number o	hits	bandwidth	
11/2019	667	6 133	9 342 106	5,13	
12/2019	617	5 717	6 615 380	4,24	
01/2020	505	4 805	8 235 915	4,87	
02/2020	1739	6 560	8 738 395	3,59	
03/2020	783	3 504	10 521 897	3,56	
04/2020	551	3 610	10 022 383	3,50	
05/2020	529	3 506	13 380 715	3,49	
06/2020	1082	4 774	13 938 410	3,32	
07/2020	1685	7 538	4 723 588	3,51	
08/2020	640	6 560	2 852 686	3,80	
09/2020	811	6 172	10 189 287	3,79	
10/2020	570	4 988	8 164 053	3,80	
Average	848	5 322	8 893 735	3,88	

Statistics on the Argo data management web site: <http://www.argodatamgt.org>

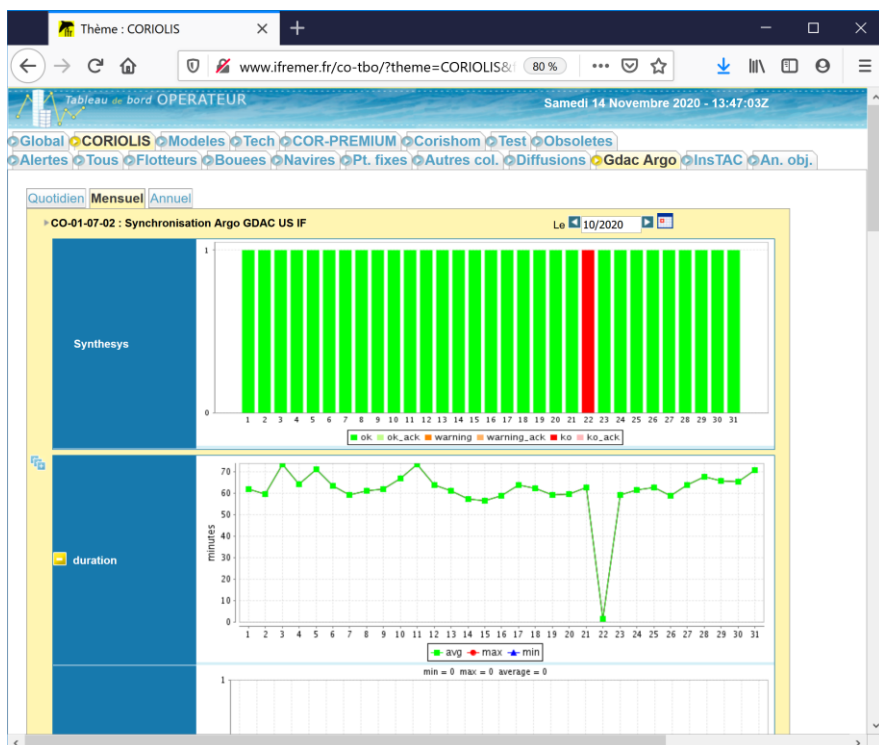
There is a monthly average of 1400 unique visitors, performing 2200 visits and 37000 hits. The graphics shows a slightly stable number of unique visitors.



ARGO GDAC web statistics						
month	unique vi	visits	pages	hits	bandwidt	
nov-19	1 689	2 488	5 171	41 499	1,1	
déc-19	1 624	2 405	4 669	36 519	1,1	
janv-20	1 385	2 070	4 186	34 275	1,6	
févr-20	2 164	3 168	6 628	50 696	1,7	
mars-20	1 572	2 729	5 664	38 943	1,5	
avr-20	1 550	2 337	4 479	37 017	1,1	
mai-20	1 336	2 081	4 298	32 350	1,0	
juin-20	1 210	1 848	4 355	32 283	0,8	
juil-20	1 304	1 919	4 084	34 536	1,3	
août-20	1 239	1 742	3 794	30 537	1,7	
09/2020	1 503	2 329	4 969	42 220	1,8	
10/2020	1 663	2 313	4 748	39 695	1,5	
Average	1 493	2 254	4 721	37 255	1,4	

3.3 GDAC files synchronization

The synchronization with US-GODAE server is performed once a day at 03:55Z



Synchronization dashboard in October 2020: the daily sync. time takes on average 1 hour, with a failure on October 22nd.

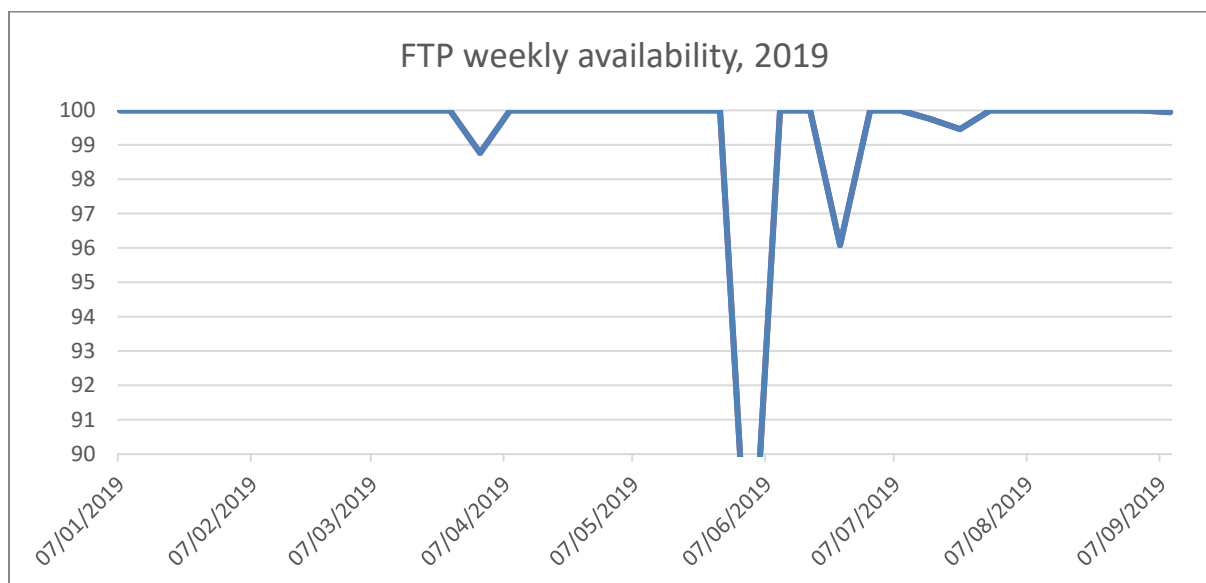
3.4 FTP server monitoring

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

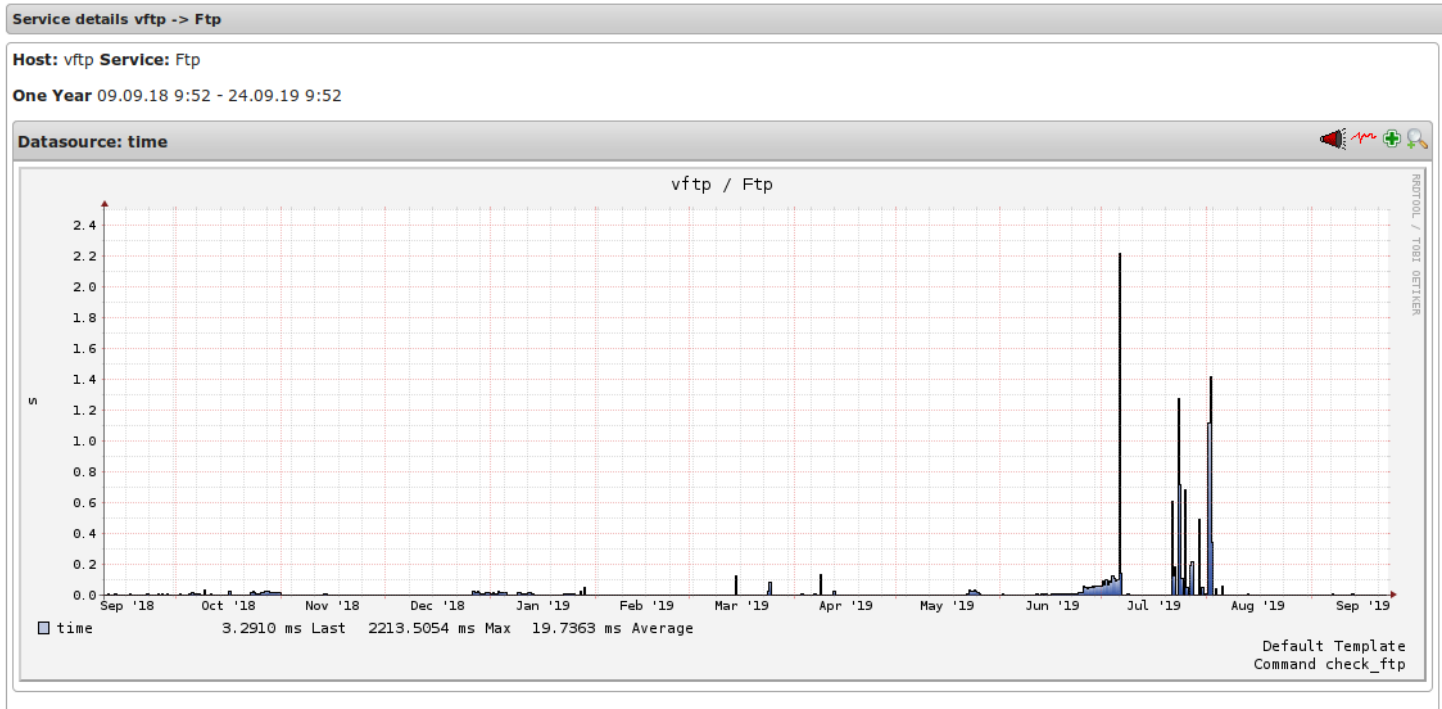
Every 5 minutes, an ftp download test and an Internet Google query are performed. The success/failure of the test and the response time are recorded. The FTP server is a virtual server on a linux cluster.

On the last 9 months, the FTP server was operational on 99.540% of time, non-operational during 1 day and 2 hours (0.421%). This is a very poor performance compared to last year (only 14 minutes non-operational in 2018). The main explanation is electricity maintenance work, which will hopefully improve the future FTP availability. The graphics below shows that the major FTP outages occurred on June 7th and then in July 6th 2019.

FTP server monitoring 01/01/2019 - 24/09/2019			
Status	percentage	duration	comment
OK	99,540%	256d 3h 7m 20s	operational
Warning	0,039%	0d 2h 10m 10s	poor performance
Unknown	0,000%	0d 0h 0m 0s	
Critical	0,421%	1d 2h 56m 22s	non operational

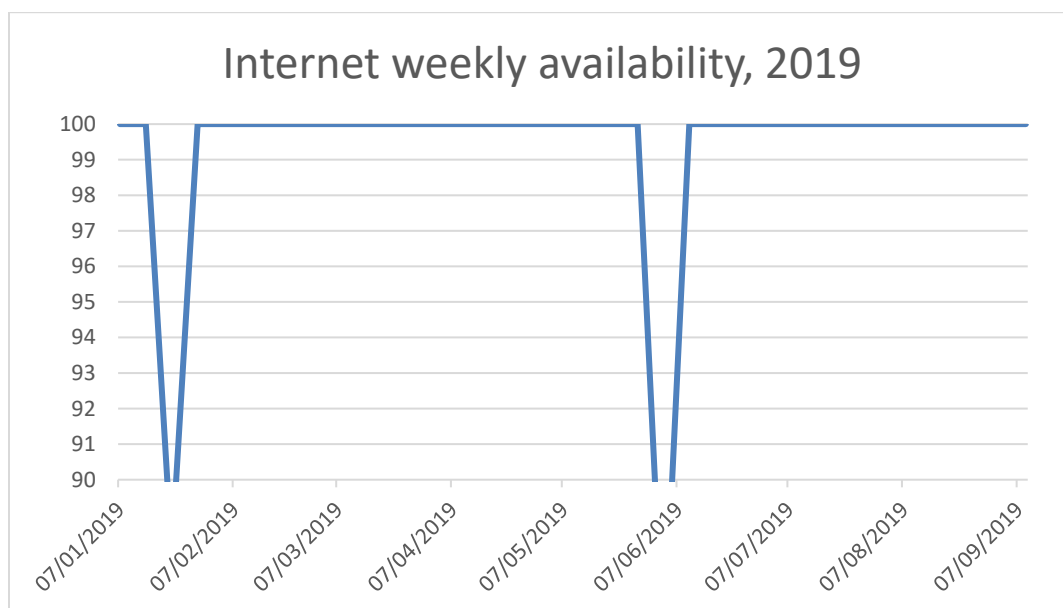


Nagios ftp monitoring: between January and September 2019



FTP server response time monitoring, poor performances end of June and in July

Internet access monitoring 01/01/2019 - 16/09/2019			
Status	percentage	duration	comment
OK	99,816%	265d 20h 33s	operational
Warning	0,000%	0d 0h 0m 0s	poor performance
Unknown	0,000%	0d 0h 0m 0s	
Critical	0,184%	0d 11h 46m 12s	non operational



Nagios Internet monitoring: between January and September 2019, poor performances in January and June.

3.5 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 2210 entries** (November 2020), compared to 2271 entries one year ago.

DAC	Nb greylist entr
aoml	966
coriolis	685
csiro	215
bodc	132
jma	103
csio	32
meds	26
incois	24
kma	18
kordi	9
Total	2210

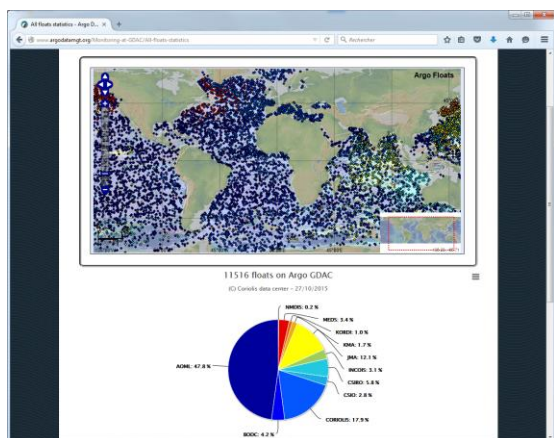
Distribution of greylist entries per DAC and per parameter

Coriolis reports many BGC greylist entries.

Parameter name	number of entries
PSAL	1274
TEMP	221
PRES	211
CDOM	140
BBP700	112
DOXY	93
CHLA	83
DOWN_IRRADIANCE380	15
DOWN_IRRADIANCE412	15
DOWN_IRRADIANCE490	15
DOWNWELLING_PAR	15
CP660	10
NITRATE	2
PH_IN_SITU_TOTAL	2
BBP532	1
PH_IN_SITU_FREE	1
Total	2210

3.6 Statistics on GDAC content

The following graphics display the distribution of data available from GDAC, per float or DACs. These statistics are daily updated on: <http://www.argodatamgt.org/Monitoring-at-GDAC>



3.7 Mirroring data from GDAC: rsync service

In July 2014, we installed a dedicated rsync server called `vdmzrs.ifremer.fr` described on:

- <http://www.argodatamgt.org/Access-to-data/Argo-GDAC-synchronization-service>

This server provides a synchronization service between the "dac" directory of the GDAC with a user mirror. From the user side, the rsync service:

- Downloads the new files
- Downloads the updated files
- Removes the files that have been removed from the GDAC
- Compresses/uncompresses the files during the transfer
- Preserves the files creation/update dates
- Lists all the files that have been transferred (easy to use for a user side post-processing)

Examples

Synchronization of a particular float

- `rsync -avzh --delete vdmzrs.ifremer.fr::argo/coriolis/69001 /home/mydirectory/...`

Synchronization of the whole dac directory of Argo GDAC

- `rsync -avzh --delete vdmzrs.ifremer.fr::argo/ /home/mydirectory/...`

3.8 Argo DOI, Digital Object Identifier on monthly snapshots

A digital object identifier (DOI) is a unique identifier for an electronic document or a dataset. Argo data-management assigns DOIs to its 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
- More on: <http://www.argodatamgt.org/Access-to-data/Argo-DOI-Digital-Object-Identifier>

Since July 2019, the DOI monthly snapshot of Argo data is a compressed archive (.gz) that contains distinct core-Argo tar files and BGC-Argo tar files. A core-Argo user can now ignore the voluminous BGC-Argo files.

Argo documents DOIs

- Argo User's manual: <http://dx.doi.org/10.13155/29825>

Argo GDAC DOI

- Argo floats data and metadata from Global Data Assembly Centre (Argo GDAC) <http://doi.org/10.17882/42182>

Argo GDAC monthly snapshots DOIs

- Snapshot of 2018 November 8th <http://doi.org/10.17882/42182#59903>
- Snapshot of 2014 October 8th <http://doi.org/10.17882/42182#42280>
- Snapshot of 2012 December 1st <http://doi.org/10.17882/42182#42250>