1. DATA RECOVERED FROM THE ANDRO PROJECT

During ADMT12 in Seoul it was decided that the data set generated for the ANDRO project will be used to populate the first delayed mode NetCDF trajectory files.
This note summarizes the work done to produce these NetCDF trajectory files.

1.1. Table of contents

1.2. The ANDRO project

The ANDRO project aims at generating a world sub-surface displacement data set based on Argo float data.

The description of each processing step applied on float data can be found in http://www.ifremer.fr/lpo/files/andro/ANDRO_JAOT_2013.pdf

The data set created for ANDRO has the following contents:

DAC name	Number of floats	Number of cycles	Last processed data date
AOML	3 123	339 129	01/01/2010
Coriolis	1 442	131 626	03/15/2013
JMA	879	90 149	01/01/2010
CSIRO	283	26 853	01/01/2010
BODC	282	24 651	01/01/2010
MEDS	268	24 318	01/01/2010
INCOIS	170	21 112	01/01/2010
KORDI	110	10 709	01/01/2010
KMA	105	10 243	01/01/2010
CSIO	62	3 496	01/01/2010
TOTAL	**6 724**	**682 286**	

Table 1: data set contents

The current version of the ANDRO atlas is available on the web (see http://wwz.ifremer.fr/lpo/Produits/ANDRO).

1.2.1. DEP data set contents

In the ANDRO project, data are stored in ASCII files called DEP files (for DEPlacement, meaning displacement in French).
We decided to gather in these DEP files all relevant information that seems useful for trajectories even if only few of it is presently used to generate the ANDRO atlas.
There is one DEP file for each processing step of each float. Thus, for a given float, the number of DEP files depends on the number of processing steps (directly linked to the float type and version) but only the final DEP file contents is used to generate the float contribution to the ANDRO atlas.

	TOTAL	Provor Argos	Provor Iridium	Arvor Argos	Arvor Iridium	APX Argos	APX Iridium (not decoded)	APX IR decoded	SOLO WHOI	SOLO SIO	NEMO	NINJA
AOML	**3 123**	8				1 413	154		715	833		
Coriolis	**1 442**	548	14	104	2	667		6			101	
JMA	**879**	90				763		8				18
CSIRO	**283**	2				263	18					
BODC	**282**	25				257						
MEDS	**268**	29				239						
INCOIS	**170**	15				155						
KORDI	**110**	21				89						
KMA	**105**					105						
CSIO	**62**	10				52						
TOTAL	**6 724**	**748**	**14**	**104**	**2**	**4 003**	**172**	**14**	**715**	**833**	**101**	**18**

Table 2: float types and transmission types in the DEP data set

At the beginning of this work (in 2007) we studied the contents of the Argo NetCDF technical and trajectory files and discovered that the data were not correctly decoded and that many useful trajectory information (surface locations, cycle timings, non-profile parameter measurements, etc..) were partially or entirely missing.
We then decided to write our own decoders and to decode anew the Argos (or Iridium) raw data provided, on demand, by the DACs.

This work has been done for all the ANDRO floats except:

- For 12 APEX Argos floats (old PALACE/APEX versions, not documented),
- For 172 APEX Iridium (we didn't succeed to get their Iridium raw data),
- For 101 NEMO floats (we used the OPTIMARE decoded files).

Consequently, for these 172 not decoded Iridium APEX floats in particular, we know that the data set can be highly improved with a correct decoding of the trajectory information provided by these floats.

1.2.2. DEP data processing

The main steps of the DEP data processing are:

- Each DEP file is created and filled with relevant data from Argo public NetCDF TRAJ/PROF/TECH/META files,
- When Argos raw data are available, the surface location set is updated with these Hex file contents,
- When both decoder and Argos (or Iridium) raw data are available, the data are decoded by our decoder and the result replaces the current DEP contents,
- A RPP (representative park pressure) (and its associated temperature if available) is computed for each cycle. This RPP is used as a realistic estimate of the float depth

during its drift at the parking pressure. These RPPs are visually checked and corrected if needed,
- The main meta-data (CYCLE_TIME, PARK/PROF_PRESSURE, REPETITION_RATE), the cycle numbers and the surface pressure measurements are visually checked and corrected,
- The grounded cycles are flagged (using a visual comparison of the RPP and the local bathymetry provided by the SRTM30+ atlas),
- For Argos APEX floats, additional cycle timings are estimated from the existing data,
- The consistency of the times gathered in the DEP file is checked.

Note that:
- Only core parameters (PRES, TEMP, PSAL) are managed in the DEP data set,
- There is no Qc in the DEP files, only ‘good’ data are preserved in the DEP file associated to each processing step,
- In the DEP files, pressures have not been adjusted for surface offset for non auto-correcting floats.

1.3. DEP file to NetCDF DM trajectory file conversion

DEP files contents can be used to populate a first version of NetCDF delayed mode trajectory files that can be submitted by the DACs to the PIs to help them generating the first Argo DM TRAJ files.
This paragraph describes what has been done to generate these trajectory files from the DEP data set.

1.3.1. Update of DEP files with Argo profile numbers

Only real time profile data are stored in the DEP files. However, to give the user a link to the DM profile data, we also store the corresponding number of the Argo profile that can be found in the GDAC FTP site.

This link is created using the expected date of the profile (computed from DEP times as defined in the DAC cookbook): ProfDate.
In the Argo NetCDF profiles of this float, we then look for a profile timely close to ProfDate: in the [ProfDate – 10 hours; ProfDate + 10 hours] interval.
If many profiles are found the timely closest one is selected.
If no profile is found, this DEP cycle cannot be linked to any Argo profile, it can be because:
- During this cycle the float has not achieved any profile,
- The profile exists in the Argo NetCDF files but it has been erroneously dated by the DAC,
- The profile has been achieved by the float but doesn’t exist in the Argo NetCDF files; it has been missed by the DAC.

Before generating trajectory DM files from DEP data set we first update this link to Argo profile numbers using the GDAC snapshot dated September 8th 2014 (DOI: 10.12770/bc3de4fa-6668-4e0e-bae3-102c6d9c8ddd , data: ftp://ftp.ifremer.fr/ifremer/argo/etc/argo-zip-archive/201409-ArgoData.tar.gz).

Note that, during this update:

- For 2.4 % of the DEP cycles, we were not able to find a corresponding NetCDF profile,
- When we succeeded to link the cycles and the profiles, the numbers differ for 5.4 % of the cycle (error in Argo profile number),
- 4 floats (46851, 49048, 51886 and 51887) from the MEDS DAC are in the DEP data set but have been removed from the Argo GDAC FTP site.

1.3.2. Creation of DEP2 files

As already said, for a given float, a new version of the DEP file is generated for each processing step and only good data are preserved in the last version of this file.
To recover the deleted data, we then have to process all the produced versions of each DEP file and to store the gathered information in a new file format called DEP2.
In a DEP2 file, DEP information are duplicated (to store original and final data) and additional information are added.

This DEP2 file, unique for each float, is used to generate its NetCDF DM trajectory file.

1.3.3. Generation of NetCDF DM trajectory files

A NetCDF trajectory file has been generated from each DEP2 file contents.
The generated trajectory file version is V3.1, compliant with the Argo User's Manual (dated July 18th 2014).
As this format version refers to the float configuration (through the CONFIG_MISSION_NUMBER variable), NetCDF meta-data files have also been generated (see 0).
Moreover we have also generated the NetCDF V3.1 profile files of the "orphan" profiles (profiles present in the DEP data set but not found in the GDAC NetCDF files, see 1.3.1).

1.3.3.1. Clock drift management

Some floats provide dated information. For these floats, the drift of the onboard clock should be estimated so that the provided times can be corrected from clock offsets.

For PROVOR/ARVOR and NINJA Argos floats, the (onboard) float time is provided in the technical message. In that case, a clock offset can be computed. As all provided float time determinations are based on technical information, if the technical message is not received, there is no float times for the concerned cycle.
Thus, in the provided trajectory DM files, **for PROVOR/ARVOR and NINJA Argos floats, the times provided by the float are always corrected from float clock offset.**

For PROVOR/ARVOR Iridium floats, the float clock is set each time the float surfaces.
Thus, **for PROVOR/ARVOR Iridium floats, the clock offset is considered to be 0.**

ANDRO SOLO (SIO and WHOI) float versions don't provide any float times.

Only APF9 versions of APEX Argos floats provide float times. Unfortunately, these float versions provide only the (onboard) float time in the test message transmitted during the

prelude phase. Consequently, only clock offset at launch can be computed for these floats and we need to estimate the clock drift (with the method presented in the cookbook) to compute the clock offset for each cycle.
Thus, in the provided trajectory DM files, **for APEX Argos floats, the times provided by the float are corrected from an estimated clock offset but only when test message data is available.**

For APEX Iridium floats the float time is transmitted each cycle and all float times can be corrected from clock offset.
Thus, in the provided trajectory DM files, **for APEX Iridium floats, the times provided by the float are always corrected from float clock offset.**

For not decoded APEX Iridium floats and for NEMO floats we have no information on the clock offset management done by the people in charge of the decoding.

1.3.3.2. Parameter adjustment management

The parameter measurements have not been adjusted in the ANDRO data set; moreover pressure offsets have not been corrected from surface offset (for non auto-correcting floats).
It was then decided at the AST15 to adjust the trajectory parameter measurements using the Argo NetCDF profile adjustments.

The rules for processing such adjustments in the trajectory files are the following:

1. If the DEP cycle is linked to an Argo nc profile:
    a. If the Argo nc profile has a 'D' mode: the PTS adjustments are used to adjust the DEP PTS measurements for this cycle ('D'),
    b. If the Argo nc has a 'A' or 'R' mode: the pressures of non auto-correcting floats are adjusted using the DEP surface pressure measurements and computing the offsets according to §2.3.1 of the Argo QC manual ('A' or 'R'),
2. If the DEP cycle is not linked to an Argo nc profile:
    a. The DEP cycle is included in the trajectory file without adjustment ('R')
    b. The DEP profile that doesn't exist in any nc Argo profile is provided to DAC who is in charge of updating the profile serie itself.

The trajectory file measurements can have different resolutions (depending on the measurement code). These resolutions have been taken into account while adjusting the data. Example: a pressure with a 1 bar resolution is adjusted with round(PresAdjInDbar/10)*10 where PresAdjInDbar is the pressure adjustment value (in dbar) found in the nc profile file.

In particular circumstances, some trajectory file measurements can be set to meta-data values (CONFIG_ParkPressure_dbar for example); these measurements have been adjusted with a 0 offset value.

1.3.3.3. NetCDF meta-data file contents

For each generated trajectory file, we have also generated its associated meta-data file, however only essential configuration parameters are provided in these files.

For PROVOR floats we provide:
- `CONFIG_CycleTime_days`
- `CONFIG_ParkSamplingPeriod_hours`
- `CONFIG_ParkPressure_dbar`
- `CONFIG_ProfilePressure_dbar`
- `CONFIG_Direction_NUMBER`
- `CONFIG_CtdCutOffPressure_dbar`
- and `CONFIG_DescentToProfTimeOut_hours` when needed

For APEX floats we provide:
- `CONFIG_CycleTime_days`
- `CONFIG_ParkPressure_dbar`
- `CONFIG_ProfilePressure_dbar`
- `CONFIG_DownTime_hours`
- `CONFIG_UpTime_hours`
- `CONFIG_DescentToProfTimeOut_hours`

For SOLO (SIO and WHOI), NEMO and NINJA floats, we provide:
- `CONFIG_CycleTime_days`
- `CONFIG_ParkPressure_dbar`
- `CONFIG_ProfilePressure_dbar`

We have also added the SANTARD_FORMAT_ID for most of the floats.

1.4. Provided files

In the current directory you can find:
- This note (*argo-andro-data_20141016.pdf*),
- The *ANDRO_floats.xls* excel file,
- The *data* directory,

The *ANDRO_floats.xls* excel file provides the list of ANDRO floats, with: its DAC, its WMO number, its ANDRO float type (see comment in ‘C1’ cell of the file), its ANDRO decoder Id (= -1 for “not decoded float”), its transmission type and its family name (proper to each DAC).

The data directory contains, for each DAC, the generated trajectory file and meta-data file and in a ‘profiles’ directory, the generated ‘orphan’ profile files.
A zipped version of these data is also provided.

If you have any question about the provided data, don’t hesitate to contact me (jean-philippe.rannou@altran.com).