

Argo data management

16 July 2002

ar-um-02-02

Comments on USER'S MANUAL

ARGO

part of the integrated global observation strategy



Argo data management

Comments on User's manual

<u>1.</u>	<u>INTRODUCTION</u>	<u>4</u>
<u>2.</u>	<u>COMMENTS ON VERSION 0.9</u>	<u>4</u>
<u>3.</u>	<u>COMMENTS ON VERSION 0.9B</u>	<u>10</u>
<u>4.</u>	<u>COMMENTS ON VERSION 0.9C</u>	<u>24</u>
<u>5.</u>	<u>ELEMENTS FOR A NEXT VERSION OF ARGO FORMATS</u>	<u>26</u>

1. Introduction

This document contains the comments received on the different versions of the argo data formats users's manual

2. Comments on version 0.9

From : Roger Goldsmith

Date : 14/12/2001

Some minor observation on release 2.0 of Argo formats.

I apologize if some seem trivial but if there is any confusion someone will get tripped up eventually.

ARGO PROFILE FORMAT

REFERENCE_DATE_TIME The DATE_TIME is 14 characters, the convention shows 16 with the inclusion of the 24 and the commented example shows 12, lacking some zeros.

TC : OK

Q_PROFILE_PSAL_QC The leading "Q_" appears to be inconsistent with the manner in which the other QC are defined.

TC : OK

ARGO TECHNICAL DATA FORMAT

DATA_TYPE It is defined here as STRING32, but defined as STRING16 in the profile and trajectory formats.

TC : string16 is ok

ARGO_HANDBOOK_VERSION Is this really different than the HANDBOOK_VERSION defined for the profile and trajectory formats?

TC : we use HAND_BOOK_VERSION (it can be applied to an other program)

ARGO META DATA FORMAT

PTT Here lower case Argos, Orbcomm is used; then uppercase in the parameter TRANS_SYSTEM. What are the conventions that will allow uniform searches?

TC : we should use uppercas

TRANS_SYSTEM Here uppercase ARGOS and ORBCOMM are used. And while not technically a transmission system, how would

GLOBALGRAM be denoted.

PLATFORM_SERIAL_NO Appropriately noted as char type so the use of "NO" and "number" may cause confusion. Perhaps "ID" and "identifier".

TC : yes but we have to wait for the next version of format

DIRECTION Here has convention for "A" and "B" while in the profile format it has "A" and "D". Perhaps all three cases "A", "D" and "B" should be declared in each format.

TC : a descending only float looks strange

START_DATE_QC long_name has "Quality on launch..." and should be "Quality on start ..." of "first descent"?

TC : OK

SENSOR_SERIAL_NO Same comment as for PLATFORM use of number versus identifier.

TC : yes but we have to wait for the next version of format

REPETITION_RATE The _FillValue should perhaps be just "99999" to match
the type "int".

TC : OK

TECHNICAL_PARAMETER_NAME Type should be lowercase char?
TECHNICAL_PARAMETER_VALUE Type should be lowercase char?

TC : yes

In various places there are defined:
STATION_PARAMETERS (profile),
PARAMETER (profile, meta),
MEASURED_PARAMETERS (trajectory),
TECHNICAL_PARAMETER_NAME/VALUE (technical, meta)

If these are all different entities then perhaps they should have more distinctive identifiers.

From : Claudia Schmid

Date : 24/01/2002

Thierry,

the document looks good.

Now some comments:

Your question on p.6:

I would use instrument (or profiling float) instead of station. I associate a station with the stopping of a ship at one particular location to obtain one (or many) observation(s), often with different instruments.

'everywhere': replace HH24 with HH (reason: we want only two numbers as hours). In definition of DATE_TIME we should specify: HH: hour of the day (0-23).

TC : OK

'everywhere': 'Universal time' is officially 'Universal Time Coordinate' and is abbreviated with UTC.

TC : OK

'everywhere': replace Exemple(s) with Example(s)

TC : OK

'everywhere': We were told not to treat the Q as a part of the WMO number. While this is not important now it will be important when we start using BUFR (I will forward the respective email to you).

TC : do we all agree ?

p.8: attributes (with s)

Do we need a reference table for DATA_TYPE?

TC : yes we do.

Do we want to discern between ARGO and non-ARGO profiling floats (AOML has several).

TC : do we or don't we ?

p.9: move PLATFORM_NUMBER to beginning of general information (also for other file formats)

p.9: ... it perform an ascending (perform needs a `s')

TC : OK

CYCLE_NUMBER: Do I understand it correct? I think what you mean is that an ascending/descending profiler has the numbers 1 1 2 2 3 3 ...
It may be good to give two examples.

p.10:

CYCLE_NUMBER: Will we ever have a 0 here? My guess is not, because you defined a cycle as descending, submerged drift, ascending, surface drift (as we also do it). I.e. cycle 0 will only consist of a surface drift.

TC : cycle 0 is the first cycle, it usually has a smaller subsurface drift.

DC_REFERENCE: Unique identifier of the profile ...

I suggest to use instrument (or profiler), since the file is designed to contain many profiles.

TC : DC reference is about a profile. One file may contain profiles from different instruments.

INST_REFERENCE: Why not INST_TYPE (consistent with long name)?

Do we need a reference table or some rules here?

p.11

JULD_LOCATION: ... is generally estimated after ... profile can be different.

TC : OK

LATITUDE/LONGITUDE: remove value

TC : OK

POSITIONING SYSTEM: GPS exists with different decoding systems. Do we want to discern between them here?

TC : do we or don't we ?

PROFILE_PARAM_QC: replace TEMP (temperature) with PARAM (parameter) in description. Similarly on p.13 (PSAL).

TC : OK

p.16: Where can we learn how the HISTORY_QCTEST values are generated? reference necessary.

TC : OK

p.17:

N_MEASUREMENT: columns not properly aligned

N_CYCLE: depends on data set

TC : OK

p.22

POSITION_ACCURACY: fill value can not be 0, because 0 is a valid ARGOS location class. Suggestion: -99

TC : I removed the value

What do we do with GPS position accuracies (also in reference tables appendix)?

Maybe we should give the accuracy in meters (or meter ranges)?

TC : OK, the range has to be defined

POSITION_QC: Here we can use the fill value 0 (consistent with the QC flags in the reference table).

TC : OK

PRES: change definition.

Pressure measurements during surface/submerged drift.

Remove the 2nd and third sentence.

What happens if a float records a negative pressure at the surface? We just store them as we get them (they may later be used to correct drifts of the pressure sensor). I.e. valid_min=0 may be problematic.

TC : maybe should we remove valid_min and valid_max

p.30

TRANS_SYSTEM: IRIDIUM is another possibility.

TC : OK

bottom of page: something is wrong with the table here (on my printout half a line, cut in the vertical, is visible)

p.32

DEPLOY_AVAILABLE_PROFILE_ID: DEPLOY_PROFILE_ID sounds better to me. I would allow "XBT 5, CTD 201" instead of just numbers (Who will remember what kind of a profile it was after 5 years?)

TC : the string256 allows these names

p.33

SENSOR_MODEL: Conductivity or salinity? I assume that the salinity is estimated on the float computer (and would thus serve as a bad example), but I don't know if that's correct. If you change it we should use the SI units dS/m (deci-Siemens per meter). The sensor resolutions will then change as well.

Concerning the data resolution: This may be affected by the way the floats store the data. Do we need a field for this (like STORAGE_RESOLUTION)?

TC : I have no opinion

WMO table 1770: Do we want to append it here, or just give the web address? I think the latter may make sense, because the table may change more often than the handbook.

TC : I inserted both.

From : Claudia Schmid

Date : 25/01/02

p.10:

CYCLE_NUMBER: Will we ever have a 0 here? My guess is not, because you defined a cycle as descending, submerged drift, ascending, surface drift (as we also do it). I.e. cycle 0 will only consist of a surface drift.

TC : cycle 0 is the first cycle, it usually has a smaller subsurface drift.

I don't understand your response. Some of our floats transmit data after deployment until they sink for the first time. The associated data is what we put into cycle 0. Therefore cycle 0 (for us) does not include any subsurface drift. Cycle 1 contains an ascending (and if applicable a descending) profile. The descending profile of a cycle, if it exists, was obtained before the ascending profile. If I understand you correctly we use different definitions for cycle 0. However, I'm not sure how your definition is. Can you clarify this?

Claudia

PS: another small suggestion. I'm a bit challenged with word documents here. We have star office on unix which can sort of read them, but it screws up the formatting. So maybe if you add your replies to comments you can do it in ascii? For the printing of the handbook itself I don't mind going to a PC.

From : Roger Goldsmith

Date : 06/02/02

Thierry Carval,
some edit for the manual.

1) There still doesn't seem to be a consensus on whether it's
POSITIONING or POSITIONNING

It is used both ways throughout the document, sometimes one way
in the variable and another in the attribute.

TC : according to harrap's, we have to use "positioning" instead of "postionning"

2) p. 32 PTT: ORBCOMM should maybe have two MM throughout?

TC : corrected

3) p. 33 LAUNCH_QC has attributes LAUNCH_DATE_QC

TC : corrected

4) p. 33 START_DATE_QC has a long name of launch date.

TC : corrected

5) p. 27 JULD_START_TRANSMISSION
the attributes have JULD_DESCENT_END

the example has DATE_TIME format rathe than the double type.

TC : corrected

From : Yasushi Talatsuki

Date : 22/02/02

[1] I found that 'NOMINAL_PROFILE_PRES', 'NOMINAL_PARKING_PRES'
and 'PARKING_PRES_QC' field were deleted in format version 2.0,
that were existed in version 1.2b.

I think these information are valuable especially for the floats
that may change parking/profiling pressure on each cycle.

Therefore, I propose to adding following definitions in
the Cycle data block;

```
int NOMINAL_PROFILE_PRES(N_CYCLE);
  NOMINAL_PROFILE_PRES:Comment = "Nominal pressure of the ascending profile for this
cycle";
  NOMINAL_PROFILE_PRES:_FillValue = 99999.f ;
  NOMINAL_PROFILE_PRES:units = "decibar";
int NOMINAL_PARKING_PRES(N_CYCLE);
  NOMINAL_PARKING_PRES:Comment = "Pressure of the submarged drift for this cycle";
  NOMINAL_PARKING_PRES:_FillValue = 99999.f ;
  NOMINAL_PARKING_PRES:units = "decibar";
char NOMINAL_PARKING_PRES_QC(N_CYCLE);
  NOMINAL_PARKING_PRES_QC:Comment = "Quality on parking pressure ofthe cycle";
  NOMINAL_PARKING_PRES_QC:Conversions="Q where Q=[0-9]";
  NOMINAL_PARKING_PRES_QC:_FillValue = "0";
```

TC : nominal values are stored in meta-data. These are not measured values that change at every cycle. They are set once before launch and recorded in meta-data.

[2] In the Profile format, we define "PSAL_CORRECTED_ERROR" to store estimated error for the salinity correction, but no definition in the trajectory format. Is it unnecessary for the trajectory format?

TC : if this parameter becomes necessary for trajectory, we shall add it when necessary.

[3] Should we treat the measurement field (PRES, TEMP, PSAL, CNDC) for the floats that do not measure during the surface/subsurface drift? (Many APEXs measure only profiles and have no drift data.)

- 1) Remain all the definition as same in profile format for the float and filled with _FillValue except for PRES? (PRES may be set to 0.)
- 2) Only defined PRES/PRES_CORRECTED... and filled with 0 for PRES?
- 3) Do not define PRES/TEMP/PSAL/CNDC and N_PARAM,TRAJECTORY_PARAMETERS for that float?

TC : the format contains the measured parameters. If a parameter is not measured, it is not defined in the file.

What do you think about above things?

From : Yasushi Takatsuki

Date : 01/02/02

I should add another thing.

According to the GTSP code table
(http://www.meds-sdmm.dfo-mpo.gc.ca/meds/databases/ocean/gtsppcodes_e.htm),
4.4 in page 44 should be
GE: BSH (Germany)
JA: JMA (Japan).
TC : ok, done

From : Yasushi Takatsuki0

Date : 31/01/02

Does Cycle 0 mean incomplete cycle (the time for the cycle may be shorter or longer than the preprogrammed cycle period) just after launch?
Therefore, in case of APEX, their cycle begin with 1 not 0, because they complete their cycle from just first descent?

The unit of conductivity in GF3 tables presented in 4.3 is "mhos/m", not "mmho/cm". Also, The unit of practical salinity is "---", not "psu".

If we use the units as listed in GF3 tables, we should change the units and "xx_format" for the parameter CNDC and PSAL. (1 mhos/s = 10 mmho/cm)

TC : in our data bases, conductivity is stored in mmho/cm . Do we change it to mhos/s ? I prefer to avoid it.

Does JULD represent either of JULD_ASCENT_END or JULD_START_TRANSMISSION in the Argo trajectory format for ascent profile?

TC : JULD is the data and time of the profile. It is the date and time at the top (surface) of the profile, wether ascending or descending.

Page 18 : in GTSP code, letter x in "QxP\$" and "QxF\$" depends on the institute where QC is performed or failed.
(see http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Databases/OCEAN/GTSPcodes_f.htm)

For example, QCP\$ and QCF\$ for MEDS, QAP\$ and QAF\$ for AOML, and so on. In Argo data format, the information of institution which performed is already presented by the field "HISTORY_INSTITUTION".

Should we use "QCP\$" and "QCF\$" for Argo data at all institute where QC is performed?

If we use "QCP\$" and "QCF\$", we must write its definition in the table in 4.7.

Page 20 : N_PARAM should be 2, 3 or 4 because we cannot explain the pressure sensor information in Float sensor information section of the meta-data format if we set N_PARAM is 1, 2 or 3.

Page 27 : Does JULD_START_TRANSMISSION represent the date/time when the float begin to transmit data even if the satellites do not receive it? Or the date/time of the first message which the satellites is received ? I intended former time. Is it right?

TC : I think it represents the date/time when the float begin to transmit data

Page 35 : I remember that "SENSOR_MODEL" associated with sensor model name (example, SBE-41 or FSI EXCELL), so we discussed that we use "SENSOR_TYPE" , which associated with sensor type such as "inductive sensor" or "thermister", rather than "SENSOR_MODEL" . Is it wrong ?

3. Comments on version 0.9b

The User manual version 0.9b was submitted to the argo-dm-format@ifremer.fr mailing list on the 5th of march 2002.

From : Roger Goldsmith

Date : 06/02/2002

Thierry Carval,
some edit for the manual.

1) There still doesn't seem to be a consensus on whether it's POSITIONING or POSITIONNING
It is used both ways throughout the document, sometimes one way in the variable and another in the attribute.

TC: we now use the right spelling is POSITIONNING.

2) p. 32 PTT: ORBCOMM should maybe have two MM throughout?

TC : ok

3) p. 33 LAUNCH_QC has attributes LAUNCH_DATE_QC

TC : ok

4) p. 33 START_DATE_QC has a long name of launch date.

TC : ok

5) p. 27 JULD_START_TRANSMISSION
the attributes have JULD_DESCENT_END
the example has DATE_TIME format rather than the double type.

TC : ok

From : Yasushi Takatsuki

Date : 22/02/2002

Dear Thierry and All,

[1] I found that 'NOMINAL_PROFILE_PRES', 'NOMINAL_PARKING_PRES' and 'PARKING_PRES_QC' field were deleted in format version 2.0, that were existed in version 1.2b.

I think these information are valuable especially for the floats that may change parking/profiling pressure on each cycle. Therefore, I propose to adding following definitions in the Cycle data block;

```
int NOMINAL_PROFILE_PRES(N_CYCLE);
    NOMINAL_PROFILE_PRES:Comment = "Nominal pressure of the ascending
profile for this cycle";
    NOMINAL_PROFILE_PRES:_FillValue = 99999.f ;
    NOMINAL_PROFILE_PRES:units = "decibar";
int NOMINAL_PARKING_PRES(N_CYCLE);
    NOMINAL_PARKING_PRES:Comment = "Pressure of the submarged drift for
this cycle";
    NOMINAL_PARKING_PRES:_FillValue = 99999.f ;
    NOMINAL_PARKING_PRES:units = "decibar";
char NOMINAL_PARKING_PRES_QC(N_CYCLE);
    NOMINAL_PARKING_PRES_QC:Comment = "Quality on parking pressure ofthe
cycle";
    NOMINAL_PARKING_PRES_QC:Convensions="Q where Q=[0-9]";
    NOMINAL_PARKING_PRES_QC:_FillValue = "0";
```

TC : nominal values are set before launching. They are located in the float cycle informations section of the meta-data format. I think that there is no need to use the word NOMINAL in this section (these parameter are here naturally nominals).

[2] In the Profile format, we define "PSAL_CORRECTED_ERROR" to store estimated error for the salinity correction, but no definition in the trajectory format. Is it unnessesary for the trajectory format?

TC : psal corrected values will probably not appear in a near future in the trajectory format. That is the reason why it is not mentioned. When corrected values will be available, they will be added.

[3] Should we treat the measurement field (PRES, TEMP, PSAL, CNDC) for the floats that do not measure during the surface/subsurface drift? (Many APEXs measure only profiles and have no drift data.)

- 1) Remain all the definition as same in profile format for the float and filled with _FillValue except for PRES? (PRES may be set to 0.)
- 2) Only defined PRES/PRES_CORRECTED... and filled with 0 for PRES?
- 3) Do not define PRES/TEMP/PSAL/CNDC and N_PARAM,TRAJECTORY_PARAMETERS for that float?

TC: I am in favour of the third option which is probably the most pragmatic.

What do you think about above things?

From : Mark Ignaszewski

Date : 22/02/2002

Question regarding the Trajectory data file format: Should the DATA_TYPE be set to something similar to "Argo trajectory data". (In the example file it

is "Argo float".)

This would seem to be more consistent with the other types where:

profile file: DATA_TYPE = "ARGO PROFILE"

technical file: DATA_TYPE = "ARGO TECHNICAL DATA"

TC : ok

From : Mark Ignaszewski

Date : 26/02/2002

Part of the GDac design is to validate the format of the data files received from the Dac. I have to show my ignorance of netCDF and ask what it means to be "correctly formatted". Specifically:

1) Does every element defined in your format description HAVE TO BE in the data file, even if there is no data for it currently? For example, if there is no "CNDC_CORRECTED" profile for a float, does it need to be defined in a Dac data file? (Should the GDac reject the file and return it to the Dac if it is not there?)

TC : the number and the name of the parameter is set in N_PARAM and PARAMETER. If a parameter exists, it has an associated corrected value. If there is no correction available, the corrected value is equal to the value.

2) (I think it is obvious but I will ask) The only "dimensions" and "variables" in a Dac data file should be those defined in the format description. Is that correct? (Should a Dac data file containing a variable not in the format definition be rejected?)

TC : if a variable of a file is not in the format, then we consider that the file format is not correct.

As I was working on some software to do the validation, I uncovered two things:

1) The Profile version 2.0 definition on your Web page contains the variable "Q_PROFILE_Psal_QC". I think it should be "PROFILE_Psal_QC"

TC : corrected

2) The example data file you sent me contains the variable "POSITIONNING SYSTEM" (notice the NN). I think it should be "POSITIONING_SYSTEM".

TC : ok, corrected

From : Bob Keeley

Date : 04/03/2002

Dear Thierry and Yasushi,

I am just back from a meeting, so have not had time to think about this.

I was hoping to see a reply from Thierry, but lacking that I thought I would give my first thoughts about this anyway. I am concerned about modifying yet again the format that we agreed to in Brest. We simply cannot

be continually doing this unless it is absolutely necessary. So, unless we have any floats either in the water or expected in the water in the next year that will modify their parking pressures from one cycle to the next,

I

would want to delay this.

If this is not the case, we have another place, perhaps, where this information could go. We do have the technical file in which we record all of the extra things for each cycle, such as battery voltages or whatever else we have. Considering that knowing the nominal parking pressures will not really influence how the data are viewed by a user, we could simply put this information in the technical file.

Regards, Bob

From : Mark Ignaszewski

Date : 06/03/2002

I noticed a few more things as I was working on the netCDF files.

1) In your "profile v2.0 specification" on the Web page all of the HISTORY_xxx variables are dimensioned as
(N_PROF, N_HISTORY, ...)

TC : corrected

They must be (N_HISTORY, N_PROF, ...)
Your example file had this correct of course.

2) In the example file the definition of HISTORY_QCTEST is:

```
float HISTORY_QCTEST(N_HISTORY, N_PROF, STRING16) ;
```

This should be "char" instead of "float", I believe.

TC : ok

From : Yasushi Takatsuki

Date : 08/03/2002

I'm sorry to delay my response.

Page 9,
Ucar web site address should be "http://www.ucar.edu/ucar/".

TC : ok

Page 10, Comment of N_PARAM
N_PARAM should takes a value of 2/3/4 instead of 1/2/3 as in other format.

TC : ok

Page 11, Comment of PI_NAME
PI means 'principal investigator', not "principle".

TC : ok

Page 22, Comment of REFERENCE_DATE_TIME
"January" instead of "january".

TC : ok

Page 24, Comment of LONGITUDE
It should be removed '-' from '-16.7222' for 16-43.1992 E.

TC : ok

Page 26-27, Comment of JULD_xxxx_xxxx
Examples should be 18833.98733.. or something else instead of "20011230090500 : December ...".

TC : ok

Page 43, Physical parameter codes (GF3)
 Concerning to the "Units", it should not be inconsistent in the same document.
 If we use "mmho/cm" for conductivity, it should be removed "Units" information from Table 4.3. But we strictly follow the GF3 manner, we should use "mhos/m" instead of "mmho/cm" in the Argo data.
 The unit used in the Argo data is not depend on the unit used in Database at each institution.
 TC : ok, we shall follow GF3 and use "mhos/m"

Page 49, http address of US-GODAE server
 Does anyone know correct address ?
 TC : ok

From : Mark Ignaszewski

Date : 08/03/2002

> Page 49, http address of US-GODAE server
 > Does anyone know correct address ?

The address is: www.usgodae.fnmoc.navy.mil
 TC : ok

Do not be alarmed that there is no mention of Argo on this Web site at this time. Links to the US Argo GDac will be on the site very soon.

From : Roger Goldsmith

Date : 12/03/2002

Re: DIRECTION: Convention ...

A descending float may look strange but the DIRECTION would seem to be Needed to indicate the direction a particular profile was taken, even though everything may be stored in conventional top to bottom. Especially true if you are keeping profiles in separate files.
 The ORBCOMM floats take a profile on the way down the first time, Profile on the way up, transmit both profiles then continue with only ascending profiles for the rest of the mission. Is this a BOTH? That seems like a good way to mess up counters. If it is ascending one has to account for the first profile somehow and it's not a separate CTD cast.
 TC : the direction of each individual profile is indicated in the profile format. The DIRECTION in the meta-data format is maybe useless. Maybe can it be removed in the next version of the format.

From : Stephen G. Loch

Date : 13/03/2002

My comments relate mostly to typos or phrasing but there are one or two substantial issues too.

Re metada format Roger's proposal (using the output from ncdump) sounds the simplest but to assure myself that it was OK I would need to see

whether the visual appearance (of STRING256) was satisfactory or accept somebody else's word for it. Are we allowed new-line characters?

Steve

S.G.Loch
BODC Systems Coordinator
www.bodc.ac.uk

Timing Issues

=====

UTC is Coordinated Universal Time (or if you like Universal Time Coordinated - note the 'd' on the end, pace Claudia). The French name, which I was tempted to think gave rise to the inversion of the natural English, is Temps Universel Coordonne (plus acute accent). So if explaining what UTC is use 'Coordinated Universal Time'.

The example given for relative Julian day number on p13, unlike p26,27, is correct but needs to go further (note: eliminate the blanks preceding the colons):

Suggest putting: 18833.8013889885 = July 25 2001 19:14:00

TC : ok

Are we required to use 1950.1.1 or can we do what we like? In the latter case you just extend the example (giving an arbitrary origin) but with the former a comment in the preamble would be useful. The obvious date to use for newly deployed floats is 2000.1.1.

TC : in version 2.0, we shall stick to 1950.1.1 . If we want to change it, we shall wait a version 2.1.

Most of the comments re-inforce the idea of using UTC. Some of the comments on DATE_TIME parameters don't though. It is important in a global setup to emphasise that UTC applies as much to the metadata as to the data.

TC : ok

Issues of Style

=====

Arguably you should capitalise 'celsius' and 'julian' but the document seems consistent here so leave it.

I guess some measure of consistency is useful on American/British spelling as you can have both on the same page or even line at the moment. Standardising on 'centre' would dispose of many of these (and underline the global nature of the enterprise) and agrees with the names of defined variables.

TC : "center" has been replaced by "centre" throughout the document.

String Dimensions

=====

Sorry, but I don't know what a null terminated (null-terminated) string dimension is. If the string length is appropriate to say STRING4, then you can store 4 characters in an instance of the NetCDF variable so why do we need to say 'null-terminated'? In C you actually need 5 characters of memory in a program because of the null (binary zero byte) that terminates a string but that has nothing to do with the NetCDF file and is not a requirement for, e.g., Fortran. If we are saying that all trailing blanks should be converted to zero then say so with a particular note to Fortran programmers. Of course this doesn't save any

space in the file. Or is it saying, as I suspect is the case, that we must reserve at least one character to allow a binary zero to be inserted at the end of the line, to make programming simpler in C (having retrieved the array no further shifting is required to produce all the strings)?

If the last interpretation is the correct one then DATA_CENTRE(String2) (p40) and perhaps other variables are underdimensioned. E.g. HISTORY_PARAMETER. The quickest way to solve the problem, if there is one, would be to add 1 to the string dimensions but keep the same names!
 TC : "null terminated" has been removed from comments.

Versions

=====

Why do have some VERSION variables as STRING4 but one or two as float? Surely better STRING4. E.g p9.

TC : yes, but we shall wait for a nex version of the format.

Typos

=====

'The GF3' crops up many times as does 'the The GF3': replace with 'the GF3'. There are important modifications to be made on p36 which until they are made prevent preclude authoritative publication.

Versions

=====

Why do have some VERSION variables as STRING4 but one or two as float? Surely better STRING4. E.g p9.

TC : ok.

Page by page

=====

p1 'Stephan Loch' --> 'Stephen Loch'

TC : ok.

p4 'Positionning' --> 'Positioning'

TC : ok.

p6 'data-format --> 'data format'

TC : ok.

p7 'An argo profiler cycle' --> 'An Argo profiler cycle'

TC : ok.

p7 'argos transmission' --> 'Argos transmission'

TC : ok.

p9 'NetCDF' appears as : 'NetCDF', 'netCDF' and 'NetCdf'. Suggest 'NetCDF'.

TC : ok.

p9 'General Informations' should be "General Information"

TC : ok.

p9 'ascii' is usually 'ASCII'

TC : ok.

p11 'acceptable data type' --> 'acceptable data types'

TC : ok.

p11 As noted above HANDBOOK_VERSION is given as float; suggest STRING4

TC : ok but in a next version.

p11 'january' --> 'January'
TC : ok.

p11 'temperature in degree' --> 'temperature in degrees'
TC : ok.

p12 'measurement occur' --> 'measurement occurs'
TC : ok.

p12 'for data center' --> 'for data centre'
TC : ok.

p12 Surely ':comment' should be ':long_name'
TC : ok but in a next version.

p12 'different identifiers schemes' --> 'different identifier schemes'
TC : ok

p12 box horizontally truncates last line of text at bottom of page on my printout
TC : I did not notice. Maybe is it a printer problem.

p13 'location of profile is': the 's' seems to be printed on top of the 'i'
TC : I did not notice. Maybe is it a printer problem.

p13 JULD and JULD_LOCATION examples need improving (see above)
TC : ok

p14 Suggest 'Each item of' be replaced by 'Each NetCDF variable in'
TC : ok

p15*2 'the The GF3' --> 'the GF3' (as noted above)
TC : ok

p15 <PARAM>_CORRECTED attributes are given for salinity. The comment therefore needs to say: 'Attributes given here are for salinity and should be changed appropriately for other variables'.
TC : ok

p16 another 'The GF3'
TC : ok

p16 'to a parameter of a profile' --> 'to a parameter in a profile'
TC : ok

p16: No examples of equations given. Surely we can manage a linear calibration?!
TC : ok

p17 'data center' --> 'data centre'
TC : ok

p17 'An history record ' --> 'A history record'
TC : ok

p17 'refernce' --> 'reference'
TC : ok

p18 There looks to be unfinished business here (remarks in italics)

TC : ok

p20 '17 :28 :34' --> '17:28:34' (as noted above)

TC : ok

p22 'january' --> 'January' (as noted by Yasushi)

TC : ok

p22 Should ARGO be capitalised. It's not an acronym, is it? Suggest Argo.

TC : ok

p22: GF3 again

TC : ok

p23: Horizontal truncation of text at foot of page

TC : I did not notice. Maybe is it a printer problem.

p24 colon spaces

TC : ok

p24 'different identifiers schemes' --> 'different identifier schemes'

TC : ok

p24 'refrence' --> 'reference'

TC : ok

p24 'Argos locations classes' --> 'Argos location classes'

TC : ok

p25 <PARAM>_CORRECTED etc. needs comment on required adjustment for other variables (see p15 comment)

TC : ok

p25 Example is inappropriate. Need to add: 18991.37872222 = December 30'th

TC : ok

p27*4 Ditto

TC : ok

p28 'Date' --> 'Date and time (UTC) of action

TC : ok

p28 HISTORY_PREVIOUS_VALUE. Are all flags numeric and are they going to stay that way? If not the type of this variable is a problem. Secondly is it unambiguous whether you are referring to the flag or the parameter? Cannot they both change at the same time? What about GF3 code 5?

TC : we may decide to change in the next version of the format.

p29 It might be helpful to say 'quality control' rather than the shorthand 'qc'

TC : ok

p31 Type of HANDBOOK_VERSION (suggest char STRING4).

TC : ok but in version 2.1

p30 Remove spaces before time colons, 'ascii' --> 'ASCII', UTC expansion

TC : ok

p31 Ditto
TC : ok

p32 'chracteristics' --> 'characteristics'
TC : ok

p32 'multi-beacons' --> multi-beacon'
TC : ok

p32*2 'is equiped' --> 'is equipped'
TC : ok

p33 'ARGO' --> 'Argo'
TC : ok

p33 'data center' --> 'data centre'
TC : ok

p33 ANOMALY. I think this is an unfortunate name as it has widespread use in physical oceanography. What's wrong with COMMENT?
TC : we may decide to change in the next version of the format.

p33 'the immesion drift' --> 'the immersion drift'
TC : ok

p34 time colon spaces
TC : ok

p34 'deployment platform' --> 'deployment platform'
TC : ok

p34 'ctd or xbt' --> 'CTD or XBT'
TC : ok

p35*2 'The GF3' --> 'the GF3'
TC : ok

p36*n 'PREDEPLOYEMENT' --> 'PREDEPLOYMENT'
TC : ok

p36 No example equations, coefficients or comments
TC : ok

p37 'Time spent to ascend' --> 'Time spent in ascent'
TC : ok

p37 'Time spent to descend' --> 'Time spent in descent'
TC : ok

p38*3 Units in this document are singular in the comments field.
Suggest 'decimal hours' --> 'decimal hour'
TC : ok

p38 Text horizontally truncated by box at bottom
TC : I did not notice. Maybe is it a printer problem.

p39 'Thes information are registred' --> 'This information is registered'
TC : ok

p39 'acii' --> 'ASCII'
TC : ok

p39 'universal time coordinate' --> 'Coordinated Universal Time',
'ascii' -->'ASCII'

TC : ok

p39 'january' -->'January'

TC : ok

p40 See comment about underdimensioning given above for DATA_CENTRE and DATA_TYPE. 'Argo technical data' is 19 characters and so STRING16 appears to be too small anyway.

TC : ok but in version 2.1

p42. On the subject of ncdump, I have recently found a bug in ncdump which can give rise to faulty listings when the C_format attribute is present. How the bug has lasted so long defeats me but perhaps because people don't report these problems. Anyway if Boulder don't fix it the patch is small and can be made available.

p43 As Yasushi has pointed out GF3 uses mho/m (surely this rather than mhos/m) and not mmho/cm.

TC : ok

p43 Should RCRD be in 4.3?

TC : sorry, I do not understand the question

p44 Shouldn't BODC feature as a data centre? How come we have RU and VL? They should be mutually exclusive so the wording for RU needs adjusting. OK, so it's a MEDS problem!

TC : ok

p45 Additional interposed blank line appearing (for 1,2 & 3). Should be removed.

TC : ok

p46 It would be helpful if DS was noted as being the concatenation of level and class.

TC : ok

p46,47 The notes are referenced by number but they are not numbered. Just add 1), 2), etc. to the start of the relevant paragraphs.

TC : ok

p47 'geospacial' --> 'geospatial' (but some dictionaries allow it)

TC : ok

p48 'occur.at' -->'occur at'. 'profile (' --> 'profile. ('

TC : ok

p48 '4.9 Positionning' --> '4.9 Positioning'

TC : ok

p48 'centers' --> 'centres'

TC : ok

p49 'delayed' --> 'delayed mode'

TC : ok

If people are unsure what the issues are relating to string dimensions, variables (and attributes) , they can look for a fuller account in section 7.15 of the NetCDF User's Guide for C, Version 3, June 1997.

Incidentally Russ Rew of Unidata accepts there is a bug in ncdump (and a couple of other places as well he thinks) and hopes to fix the problem for the next NetCDF beta release (3.5.1).

From : Mark Ignaszewski

Date : 15/03/2002

In the Profile File format:

- 1) The JULD:units and JULD_LOCATION:units both say
"days since 1995-01-01 00:00:00 UTC"

They should probably say
"days since REFERENCE_DATE_TIME"

TC : we shall discuss it for version 2.1

- 2) In the Definition of CALIBRATION_DATE, it should be
char CALIBRATION_DATE (N_PROF, etc...

TC : ok

From : Bob Keeley

Date : 25/03/2002

In the latest version of the User's Manual, certain information seemed to have been dropped. I am writing a detailed description of what changes in file contents we expect to see as a result of delayed mode QC. In the course of this, I have been checking the User's Manual and see that there is no table of Action Codes. I am sure this must have been there before.

TC : they have been transferred in the chapter 4.7 of the reference tables.

I also noted that in the description of fields, where code tables entries are required, though the comments refer to the tables, they do not refer to the exact table. I would suggest that the exact table to look at be linked to the appropriate fields. For example, the comments for HISTORY_INSTITUTION should refer to table 4.4.

TC : ok

Another note, we will be having our scientists in two different places in Canada carry out the delayed mode QC. Consequently, we will need two more entries in table 4.4 I propose to use CI for Institute of Ocean Sciences, Canada, and to use CB for Bedford Institute of Oceanography, Canada. Can you please add these entries to your list?

TC : ok

I also note that the list of GF3 codes, table 4.3, does not contain codes needed to describe dates and times in history records when these are required. Specifically we require LAT\$, LON\$ and DAT\$.

TC : is there a GF3 code for these ?

I will send a copy of what I write since I think this will be a handy thing for all of the national centres to use. We can generalize the contents for international consumption.

From : Naveenta Anand

Date : 02/04/2002

I have the three formats ready as per the differences indicated in your last email. Please run them through your program one more time. You will have to fix one thing in your program as per follows.

The DATA_TYPE in technical format has a dimension of STRING16 according to the manual. It should be STRING32 as we would like to display "Argo Technical data" -which is > 16 characters. Thus String 16 is too short for displaying the same. My program assumes a dimension of STRING32 currently. You might want to change your verification program accordingly.

TC : ok for 32 in technical data section.

From : Mark Ignaszewski

Date : 04/04/2002

Regarding the size of DATA_TYPE: The Argo Data Management community must address the issue of "flexibility" in the file formats more specifically. I

certainly don't want to put myself in the position of setting policy regarding file formats.

The published format specification defines both the size of the DATA_TYPE element and its allowed settings. It may be reasonable to allow variables dimensioned with the STRINGxx dimensions to be longer than the specification but I would like to hear the opinions of others involved in the Data Management group.

Opinions?

TC : ok, bt we shall wait version 2.1

(I too have a couple of variables I am tempted to change unilaterally but I have not yet. Specifically, I think HANDBOOK_VERSION should be "char (STRING4)" not "float" to be consistent with FORMAT_VERSION.)

TC : ok, but we shall wait version 2.1

From : Mark Ignaszewski

Date : 05/04/2002

I have a few comments regarding the formats (I apologize to Thierry if he has already seen these before):

1) The "units" attribute for all of the "JULD" related variables should be "days since REFERENCE_DATE_TIME"

not

"days since 1950-01-01 00:00:00 UTC"

TC : ok, bt we shall wait version 2.1

2) The "_FillValue" value for the "JULD" related variables should be consistent in all of the formats. It is "999999." in some and "99999." in others.

TC : the adopted fill value is "999999."

3) The "units" attribute for the TEMP variable needs to be consistent. It is defined as "degree_Celsius" in some cases and "degree_celsius" in others

TC : the unit name is "degree_Celsius"

4) The HANDBOOK_VERSION variable should probably be defined as
 char HANDBOOK_VERSION (STRING4)
 not
 float HANDBOOK_VERSION
 TC : ok, but we shall wait version 2.1

From : Denis Croize-Fillon

Date : 12/04/2002

* 'Conventions' ou 'conventions' car dans ce format, les deux ecritures sont presentes. Bien sûr, quand il fallait 'C' j'ai trouvé 'c' dans le fichier et inversement. En ce sens, le doc ne me parait pas cohérent, le format Argo Profile File Format 2.0 ne contient que 'conventions',
 * il y a un certains nombre de variables sans dimension (1 par default mais ...) : TRANS_REPETITION, CLOCCCK_DRIFT, DIRECTION, ... Pour en avoir fait l'experience dans le developpement de la lib sous matlab, on s'attend a avoir des variables parfaitement definies et non pas définies par default. Il me faut me resoudre a traiter ce cas mais il me semble que, ainsi définies, ces variables sont incohérentes,
 *un détail, CLOCK_DRIFT etant du type float, _FillValue et de type float et non pas char (les guillemets sont en trop).
 TC : we use "conventions" nstead of "Conventions"

From : Claudia Schmid

Date : 22/04/2002

Yeun-Ho found an inconsistency in the format descriptions in the user manual 01/03/02 (version 0.9b) p.26. (trajectory format):

JULD_ASCENT_START has convention YYYYMMDDHHMISS and is defined as double. I think you wanted to put the same text here than p.24 for JULD.

Same for: JULD_ASCENT_END JULD_DESCENT_START JULD_DESCENT_END
 JULD_START_TRANSMISSION

TC : JULD_xxx are declared as double. The comment "Format : YYYYMMDDHHMISS" is now deleted.

From : Argo Science team

Date : 23/03/2002

These comments were reported by Sylvie Pouliquen for the Argo Science Team (Hobart meeting, march 2002).

Page 16 : update the explanations on profile calibration. Only the best scientific calibration is stored.

TC : in version 2.0, different calibrations are stored for each profile. In version 2.1, only the best calibration will remain. The N_CALIB dimension will be removed.

4. Comments on version 0.9c

The User manual version 0.9c was submitted to the argo-dm-format@ifremer.fr mailing list on the 24th of april 2002.

From : Mathieu Belbéoch

Date : 24/04/2002

Sorry to bother you again with the float Ids but I recall that this unique WMO Id is A9nnnnn and NOT QA9nnnnn.
When BUFR will be implemented the Q will disappear ...
If it's impossible to remove the Q in your internal processing system, please think to use the standard 7 digits WMO Id on your websites.
Thanks for your comprehension,

TC : in march 2002, during IAST meeting in Hobbart it was decided to use QA9IIIIII as the standard argo float ID.

From : Claudia Schmid

Date : 24/04/2002

I agree with Matthieu concerning the treatment of Q as part of the WMO ID.
I think we should not do that. I think it's better to change it now then to wait until we start using BUFR.

TC : see above

conductivity: I'm still not sure why we use mhos/m which is not an SI unit.

TC : mhos/m is the unit of GF3 CNDC code.

Here are some editorial comments (I only looked more closely at the red text, and at UTC):

- p. 9: ... all dates and times have to be given in UTC (universal time coordinates)
- p. 13: universal time coordinates (not coordinate) (occurs twice)
- p. 16: universal time coordinates (not coordinated)
- p. 20: universal time coordinates (not coordinated)
- p. 24: universal time coordinates (not coordinate)
- p. 30: universal time coordinates (not coordinated)
- p. 39: universal time coordinates (not coordinated universal time)

TC : corrected

From : Thierry Carval

Date : 03/07/2002

We agreed on the content of argo meta-data files with argo user's manual version 0.9c .

We now have to clarify the format of these meta-data files.

1. A Netcdf format as defined in the user manual
2. An ascii format that is a ncdump of the netcdf format
3. A fixed length ascii format not defined yet

I personnaly prefer the option n.2 : a ncdump version of the netcdf file.

I join a sample generated by JMA.

TC : there is no clear opinion about what to do. JMA generates NetCDF meta-data files, MEDS generates ascii 'ncdumped files'. Let's consider that option 1 and 2 are valid until the next argo data management meeting in September 2002 (Ottawa).

From :

Date :

From :

Date :

5. Elements for a next version of argo formats

From : Roger Goldsmith

Date :

Re: DIRECTION: Convention ...

A descending float may look strange but the DIRECTION would seem to be Needed to indicate the direction a particular profile was taken, even though everything may be stored in conventional top to bottom. Especially true if you are keeping profiles in separate files.

The ORBCOMM floats take a profile on the way down the first time, Profile on the way up, transmit both profiles then continue with only ascending profiles for the rest of the mission. Is this a BOTH? That seems like a good way to mess up counters. If it is ascending one has to account for the first profile somehow and it's not a separate CTD cast.

TC : the direction of each individual profile is indicated in the profile format. The DIRECTION in the meta-data format is maybe useless. Maybe can it be removed in the next version of the format.

From : Stephen G. Loch

Date :

Are we required to use 1950.1.1 or can we do what we like? In the latter case you just extend the example (giving an arbitrary origin) but with the former a comment in the preamble would be useful. The obvious date to use for newly deployed floats is 2000.1.1.

TC : in version 2.0, we shall stick to 1950.1.1 . If we want to change it, we shall wait a version 2.1.

Versions

Why do have some VERSION variables as STRING4 but one or two as float? Surely better STRING4. E.g p9.

TC : yes, but we shall wait for a nex version of the format.

p12 Surely ':comment' should be ':long_name'

TC : ok but in a next version.

p28 HISTORY_PREVIOUS_VALUE. Are all flags numeric and are they going to stay that way? If not the type of this variable is a problem. Secondly is it unambiguous whether you are referring to the flag or the parameter? Cannot they both change at the same time? What about GF3 code 5?

TC : we may decide to change in the next version of the format.

p33 ANOMALY. I think this is an unfortunate name as it has widespread use in physical oceanography. What's wrong with COMMENT?

TC : we may decide to change in the next version of the format.

From : Mark Ignaszewski

Date :

The published format specification defines both the size of the DATA_TYPE

element and its allowed settings. It may be reasonable to allow variables dimensioned with the STRINGxx dimensions to be longer than the specification but I would like to hear the opinions of others involved in the Data Management group.

Opinions?

TC : ok, bt we shall wait version 2.1

From : Thierry Carval

Date :

The trajectory format is defined for one platform.

To disseminate argo daily trajectories (and for other purposes), we would like to use a multi-platform trajectory format.

From : Argo Science team

Date :

These comments were reported by Sylvie Pouliquen for the Argo Science Team (Hobart meeting, march 2002).

Page 16 : update the explanations on profile calibration. Only the best scientific calibration is stored.

TC : in version 2.0, different calibrations are stored for each profile. In version 2.1, only the best calibration will remain. The N_CALIB dimension will be removed.