Notes for OceanSites meeting in Vienna

Mailing list discussions

Matthias Lankhorst [mlankhorst@ucsd.edu] 25/03/2008

```
p. 14
```
Although I know the sentences about "different depth or pressure levels" were written by myself, I do not like them anymore. I propose to replace them with:
"If for some measurements it is more natural to use depth (DEPH, e.g. velocities from an ADCP), while for others it is better to use pressure (PRES, e.g. from MicroCat sensors on the mooring line), the data should be recorded in separate files."
TC 20080326: OK, done

Further suggestion:
"If PRES is used, DEPH should be provided as nominal values or as a simplified function of PRES and LATITUDE (Unesco 1983. Algorithms for computation of fundamental properties of seawater, 1983. Unesco Tech. Pap. in Mar. Sci., No. 44, 53 pp.)."
Note: I use MatLab routine "sw_dpth.m" by P. Morgan to do that.
TC 20080326: I think that we should handle only the appropriate level : PRES when pressure is recorded, DEPH is the other cases. The user will decide to convert or not.

p.15
Your comment on meta data suggests that we could include all meta data in the global attributes. I support that! It would also mean:
- no separate meta files (could save a lot of manpower)
- get rid of manual section 2.3 "General Attributes", the contents of which would be included in section 2.2 then.
TC 20080326: do we get rid of section 2.3 "General attributes" and move these items in global attributes?
I think that the global attributes section is for human readers "only".
If some information needs to be processed by software, don't you think that it is better to appear as a variable in the "General attributes section"?

Variables moved in global attributes section:
QC_MANUAL
DISTRIBUTION_STATEMENT: already in global attributes
CITATION
DATE_CREATION
DATE_UPDATE
DATE_SOURCE
PROJECT_NAME
PI_NAME
DATA_CENTRE

Variables that could remain in "General attributes"
DATA_TYPE
FORMAT_VERSION
PLATFORM_CODE

SITE_CODE
WMO_PLATFORM_CODE

pp. 8 ff., sections 2.2 and 2.3
I think we should have all global attributes of "char" type and explicitly say
so.
TC 20080326: the global attributes are always of character type. No problem to
mention that "General attributes" are of character type. Done

pp. 16-17
I am confused by the existence of reference tables 2 and 2.1. Please include a
few sentences about where they are used, e.g:
"4.2 Reference table 2: ...
... These values are used in the <PARAM>_QC variables that accompany each
measurement, cf. section 2.4."
and:
"4.2.1 Reference table 2.1: ...
... These values are used as an overall quality indicator (i.e. one
summarizing all measurements) in the attributes of each variable <PARAM>, cf.
section 2.4."
TC 20080326:OK done

p. 9, attribute "cdm_data_type"
List all possible values somewhere. The example uses "Time-series" which is
not among the options listed in the last column.
TC 20080326:OK, I also added a reference to CDM Unicar. Done

p. 21, section 5:
Should we include a recommendation or even a strict limit for max. file size
there (e.g. 100 MB)?
TC 20080326: •   <PARTX> : when an OceanSites data file size becomes
excessive, it can be splitted in smaller parts : PART1, PART2, … PARTN
Done

Suggestion for another *_QC value:
Ref. table 2 has value 7 unused. Define this as "nominal value".
Could be used for:
- Providing DEPH if instruments measure PRES.
- Position if mooring has no GPS
Coordinate this with Argo which uses similar nomenclature!
TC 20080326: I will ask/propose this use of 7 for "nominal value" to
SeaDataNet. For Argo, I think that "nominal value" is not relevant.

Reg. the parameter list:
- Should we decide to have all parameters as 4-letter abbreviations?
TC 20080326: SeaDataNet and BODC use 8-letters abbreviations
We should probably do the same

- Make sure that there are no ambiguities (e.g. SW vs. SDFA).
TC 20080326: can someone confirm that SW (shortwave radiation) is different
from SDFA (surface downwelling shortwave flux in air).
We should maybe "nominate" an OceanSites expert for each domain (e.g. :
hydrology, meteorology)

- Maybe arrange them alphabetically.
TC 20080326: ok

- Include TIME with improved short/long names.

Bill Burnett [Bill.Burnett@noaa.gov] 18/03/2008

This is brief since I'm on travel, but I'd like to spend some time
understanding who is providing data and what do we consider an OceanSITES site
to be? It seems that there are different definitions and different ways to
handle data - so I'd like to understand that.

Coriolis and NDBC need to understand how to share data - or agree to
procedures if we handle different data sets. We will back each other up, or
will we have different data at each ftp site?

Finally, I'd like to finalize the data handbook (with standards procedures).

Nan Galbraith 13/03/2008

This is getting detailed, so I apologize to anyone who would prefer not to be
on this list. Should we be using the ots-dm list for this instead?

> Nan, please include the ones that Bill used in his most recent file
> on ftp://data.ndbc.noaa.gov/data/oceansites/STRATUS/ in your
> comparison.

I ran the NDBC Stratus file through the CF-checker, available at
http://titania.badc.rl.ac.uk/cgi-bin/cf-checker.pl

The following errors were returned (the rest of the parameters passed):

WARNING (2.6.1): No 'Conventions' attribute present Checking variable: SRAD
   ERROR (3.1): Units are not consistent with those given in the standard_name
table.
Checking variable: PSAL
   ERROR (3.1): Invalid units: psu
Checking variable: FREQ
   ERROR (3.3): Invalid standard_name: frequency
   ERROR (3.1): Invalid units: Hz
Checking variable: VDEN
   ERROR (3.1): Invalid units: m^2/Hz
   ERROR (3.1): Units are not consistent with those given in the standard_name
table.

PSAL: Not too worried about the salinity units, there is a long-standing
discussion with the udunits folks about this, and I think we can live with PSU
instead of '0' for this. It is likely that PSU will eventually be added to
udunits, and if not, many datasets use the term so it's accepted and expected
by most software.

SRAD: downwelling_shortwave_radiance_in_air is the real standard name, not
shortwave_radiation, but I'm not sure why this did not return an error. I can
check with the rest of the standard names committee to see if our shorter name
is an alias. The canonical units syntax is 'W m-2' not 'W/m2'.

VDEN: (sea_surface_wave_variance_spectral_density) units should be 'm2 s', if
we are using the right standard name. The CF definition: "the variance of
wave amplitude within a range of wave frequency".

I don't see how the frequency values would be documented in that case; and I
don't see any name for 'frequency' in the CF standard name table, but there
is "sea_surface_wave_frequency"
(definition: the number of oscillations of a wave per unit time) with units of
s-1. I'm new to waves measurements, so maybe someone else can weigh in on
this problem.

Nan Galbraith 12/03/2008

There was a discussion of waves data on the CF email list back in November
2006. It is archived on the mail server starting at
http://mailman.cgd.ucar.edu/pipermail/cf-metadata/2006/001256.html.

You can use the 'next message' links to see all 8 messages in the thread. The
details may make more sense to Bill or to someone else who has worked with
waves data. I'd need to learn more about the data before I'd have any opinion
on how too handle it in our files.

Thanks -
Nan

> VDEN: (sea_surface_wave_variance_spectral_density) units should be
> 'm2 s', if we are using the right standard name. The CF definition:
> "the variance of wave amplitude within a range of wave frequency".
> I don't see how the frequency values would be documented in that case;
> and I don't see any name for 'frequency' in the CF standard name
> table, but there is "sea_surface_wave_frequency"
> (definition: the number of oscillations of a wave per unit time) with
> units of s-1. I'm new to waves measurements, so maybe someone else
> can weigh in on this problem.
>

Matthias Lankhorst, 11/03/2008

Hello, Nan (et alii),

thanks for your detailed comments. I have tried to answer the few questions
you put in.

On Tuesday 11 March 2008 08:54, Nan Galbraith wrote:
...
> > The list of parameters in the manual need to be revisited now that we
> > are using the format to encode different datasets... Would be good to
> > work by email on an updated list prior to the meeting and to only have
> > to validate it at the meeting.
>
> I'd like to work on this too. In the current version of the manual, some
> of the
> Standard Names do not match the CF standard. It would be well worth staying

> with the CF vocabulary since it's so widely used; it will make this data
> more
> accessible to other systems. I will try to go through them, add the
> ones that
> are missing and check the existing ones for compliance.

Nan, please include the ones that Bill used in his most recent file on
ftp://data.ndbc.noaa.gov/data/oceansites/STRATUS/ in your comparison. (I
guess you had that on your mind anyways, but just to make sure...)

> NetCDF can handle this fairly gracefully; you can use multiple "coordinate
> parameters" for depth. One might be an array for the stationary
> sensors, called
> depth, with dimensions (1, nsensors) and the other, for profilers,
> might be
> called profdepth with dimensions (time, nprofilers). These are assigned
> to the
> appropriate parameters when those are declared.
>
> It's true that a lot of existing software would be confused by this, so
> it might be
> preferable to store datasets like this in different files.

>From my limited perspective as a user, I would like something that my
self-written MatLab routine can load without manual intervention. For that
purpose, it would be very helpful to have consistent names, dimensions, etc.

We always need to ask ourselves why we do not have every project put their
stuff on their own ftp servers in their own plain-ascii format. I can read
any ascii file into my MatLab in a matter of minutes. What I cannot do that
way is load thirty datasets in a matter of minutes, because I have to look at
them first to figure out which column is T, S, etc. I can also not do "data
discovery" automatically that way.

> >>> 7. M. Lankhorst has developed a format checker that validates OS
> >>> file formats.
>
> Great. Is that available on line or for download?

Yes, I put it on:
http://www-pord.ucsd.edu/~mlankhorst/oceansites_formatchecker.m
It is a simple MatLab routine that checks the following (it will need major
revision soon, after we agree on attribute and variable names):
%   Based on OceanSITES User's Manual Version 1.0
%   Checks included:
%   - basic file naming convention (OS_XXX_YYY_ZZZ.nc)
%   - no unknown global attributes?
%   - all global and general attributes present, 'char' type, not
%   empty?
%   - no unknown variable names?
%   - TIME and vertical axes (DEPH or PRES) exist?
%   - *_QC exists for all variables (with some exceptions)?
%   - for every *_QC, is there a matching * ?

> >>> 9. Do we really want that many global and general attributes? I
> >>> would prefer fewer attributes, but strictly require those (as
> >>> opposed to "highly recommended" statements which I have seen for Argo).
>
> Can you point to some specific ones that you think would not be useful? It
> seems that any that are not considered required should still be part of
> the specification,
> if only to encourage people to use the same vocabularies.

"latitude", "longitude", and "sensor_depth" were designed with single-point
moorings in mind, but will be hard to define for gridded or section data.
Also, I listed these conflicts, which can be solved by removing one of each:
> >>> creation_date vs. DATE_CREATION
> >>> distrubution_statement vs. DATA_RESTRICTIONS

> >>> 11. Explicitly state that all global/general attributes are of type
> >>> "char", even those that contain numbers (or decide otherwise).
>
> That does not need to be universal in NetCDF, as it is easily dealt with in
> software. Why not leave it to the discretion of the data provider to
> use the
> most appropriate type for the information?

Let me tell you a fictional story, based on a true occurrence the details of
which I have forgotten: once upon a time, I wanted some analysis software to
look up a certain number in Argo float data files, somewhere in the
attributes. Let's say it was the "reference year" or so, and I wanted to find
a number like 1950. Some files had them stored as numeric values, which I
could use as they came. Other files had them stored as a string "1950", so I
had to find those via if-then and convert them to numeric first. Then there
were those files that had string dimensions in columns, so it read:
"
1
9
5
0
"
I selected those via if-then, flipped them into rows, and converted them to
numerics, because otherwise my "str2num" routine would always return four
individual digits rather than one number. Last not least, some files did not
have anything in this attribute because it was "highly recommended" but
unfortunately not mandatory. Although I agree that a sequence of if-thens
deals with the issue "easily", I was very frustrated!

Bluntly, I don't care what format (string, numeric, ...) it comes in, as long
as it is strictly the same in every data file!

Nan Galbraith 11/03/2008

Thanks, Matthias. My plan is to go over the standard_names for CF compliance,
and let someone else deal with the short names - does

that sound OK to all?   I'm tied up with another project at the moment
but will try to get back to this very soon.
> >> The list of parameters in the manual need to be revisited now that
> >> we are using the format to encode different datasets... Would be
> >> good to work by email on an updated list prior to the meeting and
> >> to only have to validate it at the meeting.
> > I'd like to work on this too. In the current version of the manual,
> > some of the Standard Names do not match the CF standard. It would be
> > well worth staying with the CF vocabulary since it's so widely used;
> > it will make this data more accessible to other systems. I will try
> > to go through them, add the ones that are missing and check the
> > existing ones for compliance.
> Nan, please include the ones that Bill used in his most recent file on
> ftp://data.ndbc.noaa.gov/data/oceansites/STRATUS/ in your comparison.
> (I guess you had that on your mind anyways, but just to make sure...)
>
Thanks, I had not seen that link before, having come late to the
conversation.
Or, Bill, can you just tell me the names you used? Or someone else who has
downloaded this data?

> > 7. M. Lankhorst has developed a format checker that validates OS
> > file formats.
> Great. Is that available on line or for download?
>
> Yes, I put it on:
> http://www-pord.ucsd.edu/~mlankhorst/oceansites_formatchecker.m
> It is a simple MatLab routine ...

I'm happy to hear that you're using Matlab for this work, since that's what I
use too. It can be tricky to share NetCDF code in Matlab because of the
different toolboxes, but I'll have a look. Thanks.

>>>>> 9. Do we really want that many global and general attributes? I
>>>>> would prefer fewer attributes, but strictly require those (as
>>>>> opposed to "highly recommended" statements which I have seen for Argo).
>>>>>
>> Can you point to some specific ones that you think would not be
>> useful? It seems that any that are not considered required should
>> still be part of the specification, if only to encourage people to
>> use the same vocabularies.
>>
>
> "latitude", "longitude", and "sensor_depth" were designed with
> single-point moorings in mind, but will be hard to define for gridded or
section data.
> Also, I listed these conflicts, which can be solved by removing one of each:
>
>>>>> creation_date vs. DATE_CREATION
>>>>> distrubution_statement vs. DATA_RESTRICTIONS
>>>>>

As I said (I think) this in my earlier response, you seem to be looking at
the entries in the 2 tables as separate items, but they were meant to
represent the NetCDF global atts and the metadata that would appear in a
separate metadata file, format to be determined, respectively. The fields in
the metadata file would be generated from the NetCDF attributes.

I'll start at the end of your list and work back.

The differences in the wording for the last two attributes, is, I think, the
result of the last round of editing; Thierry may know more about this since I
was not involved in that pass. I DO think the 2 concepts should be in the
global atts, and should be replicated in any stand-alone metadata
representation.
Agreed?

Sensor depth should clearly be a parameter attribute, not a global attribute,
so I agree that we can get rid of it - as long as it is used where
appropriate, at the parameter level. If sensor depth is somehow not apropos
for a particular kind of data set, we'd need to accommodate that.

As for lat/long, my understanding was that this convention was mainly for
ocean reference station data, not for gridded data or sections, but I think we
should be able to generalize enough to accommodate those if that's really what
is needed.   Would this convention work, aside from the attributes
lat/lon, for
gridded or section data? If there are a lot of other problems, we would
probably need to create a profile or extension for those data types.

If everything else in our specification is acceptable for non-station data, we
could replace the lat/long attributes with a string field "position", using
one of the existing standard vocabularies representing ocean
regions as an alternative.   The idea is that it is useful to be able to
run
ncdump -h (or its equivalent) and get overall location information; if it is
only stored as a parameter, it's not readily available as metadata.

>>>>> 11. Explicitly state that all global/general attributes are of
>>>>> type "char", even those that contain numbers (or decide otherwise).
>>>>>
>> That does not need to be universal in NetCDF, as it is easily dealt
>> with in software. Why not leave it to the discretion of the data
>> provider to use the most appropriate type for the information?
>>
> ...
>
> Bluntly, I don't care what format (string, numeric, ...) it comes in,
> as long as it is strictly the same in every data file!
>

I get your point, and I'm not adamant about maintaining this flexibility.
On the other hand, I'll gladly share some Matlab code that I picked up from my
old colleague, Chuck Denham. His library automatically detects
attribute/variable types and extracts information appropriately.
It makes all the difference when reading other people's NetCDF files - and you
can of course cast the returned attributes as strings if you need to.

Matthias Lankhorst, 11/03/2008

Bonjour, Aurelie,

Nan's message reminded me that I meant to bring this to your attention. It
seems that your script that creates the OceanSITES index file misses at least
the MOVE data. They appear nowhere in the index file, but are on the ftp
server in their "MOVE" directory. Maybe you can fix that and also verify that
it finds all the other files?

Thanks, MATTHIAS

>>> 18. ftp.ifremer.fr/ifremer/oceansites/oceansites_index.txt does not
>>> list all files that are actually on the server (e.g. MOVE).

Nan Galbraith 11/03/2008

Hi All -

Apologies for the length of this email, I tried to be brief but wanted to
address some of Matthias' comments as well as Bill's and Sylvie's.
> I don't understand your remark on excess parameters. It's not because
> we can encode a lot of parameters in the file that we will create
> empty fields... Only the parameters measured by the platform are
> recorded in the file .
Yes, one reason we went with NetCDF is that it gives us this flexibility. The
files don't need to be identical, the format is self-describing.
> ... my view is that we should provide ALL the parameters provided by
> the sites and let the user sort out the one he is more interested it
> as this may change with aplications and we know that potential use of
> OceansITES may be wider than today if data are more accessible.
Bill mentions Dominant and Average Wave Period; some buoys will have more
parameters than this and these should be included - no reason to exclude the
additional parameters we can get from directional wave packages. For the
first phase, it might be worthwhile proceeding with a subset of the parameters
that are available, to make the system easier to build, but in the (not so
long) long run we will want to include as much data as possible.
> The list of parameters in the manual need to be revisited now that we
> are using the format to encode different datasets... Would be good to
> work by email on an updated list prior to the meeting and to only have
> to validate it at the meeting.
I'd like to work on this too. In the current version of the manual, some of
the Standard Names do not match the CF standard. It would be well worth
staying with the CF vocabulary since it's so widely used; it will make this
data more accessible to other systems. I will try to go through them, add the
ones that are missing and check the existing ones for compliance.
> ----- Original Message ----- From: "Bill Burnett"
> <Bill.Burnett@noaa.gov>
>> Wave parameters - we need to agree on what wave parameters will be
>> delivered via OceanSITES. NDBC provides a Dominant Wave Period as
>> well as an Average Wave Period - should we provide both or just the
>> Average Wave Period?
>>
>> ATMP - I saw another parameter that used DRYT for Air Temperature
>> (Dry Bulb), but I can easily see people using it as Atmospheric
>> Temperature.
ATMP stands for air temperature in every dataset in which I've seen it used.
>> File sizes are a big concern. NDBC recommends that we provide
>> monthly files per year instead of having these files grow throughout
>> the years.
If we are planning to serve these with anything beyond ftp (opendap, sos,
wfs) this should not really be a problem, but I don't know if those services
are part of the plan.
>> Matthias Lankhorst wrote:
>>> 1. What should be done if depth (DEPH) parameters of different
>>> instruments in the same mooring mismatch?

>>> Example: an ADCP at the surface delivers velocity at fixed,
>>> well-known depths, while a temperature sensor along the mooring line
>>> moves up and down as the mooring moves. Potential conflict: Putting
>>> all depth information in one variable called "DEPH" may result in
>>> non-monotonic "DEPH". Potential solutions: Accept non-monotonic DEPH
>>> or put data in seperate files.
NetCDF can handle this fairly gracefully; you can use multiple "coordinate
parameters" for depth. One might be an array for the stationary sensors,
called depth, with dimensions (1, nsensors) and the other, for profilers,
might be called profdepth with dimensions (time, nprofilers). These are
assigned to the appropriate parameters when those are declared.

It's true that a lot of existing software would be confused by this, so it
might be preferable to store datasets like this in different files.
>>> 2. URGENT! Define "parameters awaiting definition" as of manual
>>> section 4.3!!!
These have standard names assigned, and choosing a short name is just a matter
of convenience. (I always use the standard name, so I'm trying not to
interfere with the choices of short names, which are more convenient for
others, except that I agree that ATMP has been mistakenly changed .)
>>> 3. URGENT! More parameters to define:
>>> Sea water velocity in XY-coordinates (eastward and northward, resp.)
>>> Air temperature (or is it ATMP ?)
Yes, we need to add current velocities, as well as several wave parameters.
>>> 4. Manual section 4.3 lists both ATMP and CAPH as air pressure, but
>>> suggests ATMP may have been confused with air temperature. Decide
>>> which one to use and remove ambiguity!
>>> 5. Manual section 4.3 references www.oceansites.org/data/units, but
>>> this page does not exist. Suggest to create this and have it include
>>> the standard parameter names and units.
We're planning on using the udunits package from unidata, as mentioned in
section 2. We could excerpt the material from that system and include it in
the Users Manual, or update the link in section 4.3 to point to the
appropriate pages on the CF/COARDS/Unidata sites.
>>> 6. How do we include height of atmospheric measurements? As negative
>>> numbers in DEPH? (OK, never mind, I think I found the solution
>>> myself: negative DEPH is ok, as are positive ones if
>>> DEPH:positive='up'.)
We actually use parameter attributes for this in our group, since a met file
will have a single depth coordinate variable (usually 0) and the various
sensors are all at different heights. We could use multiple z coordinate
variables, but that might be more confusing. I'd be interested in finding out
how others have dealt with that in NetCDF.
>>> 7. M. Lankhorst has developed a format checker that validates OS
>>> file formats.
Great. Is that available on line or for download?
>>> 8. Just what is the difference between "global attributes" (manual
>>> section 2.2) and "general attributes" (manual section 2.3)? Any
>>> technical difference? I find it very confusing that some are
>>> CAPITALIZED and others not - do we want to change this? Suggestion:
>>> all attributes non-caps, all variables and dimensions CAPS.
Yes, it is confusing to have 2 tables!

The original point of the 2 tables, briefly stated in the header for table
2.3, was that there should be separate external metadata files (useful for
data
discovery)
Every field in the external metadata file would also be in the NetCDF file.
The term "global attribute" is a part of the NetCDF specification, and these
attributes are in the NetCDF files - a subset would be in the external

file, in either SensorML, NCML or some other format.   One could omit
the information that's not useful at the discovery stage.

I'm not sure why *anything* is capitalized, but that's just a style
difference.
You're right, we should adopt a single style.
>>> 9. Do we really want that many global and general attributes? I
>>> would prefer fewer attributes, but strictly require those (as
>>> opposed to "highly recommended" statements which I have seen for Argo).
Can you point to some specific ones that you think would not be useful? It
seems that any that are not considered required should still be part of the
specification, if only to encourage people to use the same vocabularies.
>>> 10. Conflicts between global/general attributes:
>>> creation_date vs. DATE_CREATION
>>> distrubution_statement vs. DATA_RESTRICTIONS keywords_vocabulary
>>> exists, but not keywords Shouldn't we have REFERENCE_DATE_TIME?
>>> 'institution' listed twice
>>>
>>>
>>> 11. Explicitly state that all global/general attributes are of type
>>> "char", even those that contain numbers (or decide otherwise).
That does not need to be universal in NetCDF, as it is easily dealt with in
software. Why not leave it to the discretion of the data provider to use the
most appropriate type for the information?
>>> 12. Wouldn't it be better to have the same human-readable format for
>>> all dates/times in the global/general attributes? Manual v1.0 has
>>> different ones e.g. for start_date and DATE_CREATION.
Yes, these should all be consistent.
>>> 13. Do we need *_QC for TIME, LATITUDE, LONGITUDE, DEPH? If no,
>>> mention this explicitly in manual.
>>> 14. Update to manual: label and number "reference tables" more
>>> clearly (e.g. reference tables 1 and 3 each deserve a separate caption).
>>>
>>> 15. Some of the data files are rather big (>100 MB). Should we have
>>> a strict limit or at least a recommendation for file size (e.g. <50 MB)?
The document is intended to describe a NetCDF implementation, and the files
that follow that convention might be internal, served via a tool that provides
subsetting, or sent out via ftp. A size guideline probably makes a lot of
sense for files submitted/served via ftp. There might need to be a new section
or a new document describing the data sharing system.
>>> 16. Early in manual section 2, it says that "variable names are not
>>> standardized". Later on (section 4.3), they ARE standardized. I
>>> think they definitely should be standardized, and that the sentence
>>> in section 2 should be removed.
Yes, I think this is an artifact; we began with CF standard names and added
the short names - or started adding them. The point was that CF standard names
comply with an external standards organization; I don't know for sure where
the short names are coming from.
>>> 17. Similarly, there is a statement in the early parts of section 2
>>> of the manual that "dimension names are not standardized", with some
>>> suggestions on dimensions following in section 2.1. I think these
>>> dimensions should be standardized, and the suggestions of section
>>> 2.1 allow for the flexibility required. Thus again, the sentence
>>> with "not standardized" should be removed.
>>>
>>> 18. ftp.ifremer.fr/ifremer/oceansites/oceansites_index.txt does not
>>> list all files that are actually on the server (e.g. MOVE).
>>>
>>> 19. Manual should more clearly define which attributes are required
>>> for variables, e.g. *.valid_min, *.standard_name, *.QC_indicator, ...

```
>>>
>>> 20. How do I retrieve the "netcdf_version" for my software for the
>>> corresponding global attribute?
>>>
>>>
>>> 21. If variables "LATITUDE" and "LONGITUDE" are not there (they are
>>> optional acc. to manual), the corresponding "axis" attributes are
>>> also undefined (resulting in missing "X" and "Y" axes definitions).
>>> Mandatory global attributes "latitude" and "longitude" do not make
>>> sense for gridded data or section data.
>>> Suggestion: Make these variables mandatory and remove the "latitude"
>>> and "longitude" global attributes.
>>> However, keep global attributes "southernmost_latitude", etc.,
>>> because they allow for quick automatic screening and make sense for
>>> all (?) data types.
>>>
>>> 22. Likewise, require "DEPH" variable because it is the "Z"
>>> axis. (Somehow, pressure as a measured value does not make a good
>>> "axis".)
>>> Require "DEPH_QC" to allow judgement of "DEPH".
>>> Remove global attribute "sensor_depth" because it makes no sense for
>>> CTD, gridded, and section data.
If depth is not available, it seems better to allow pressure; the last thing
you want is for people to be mis-labeling their z values to force compliance.
Sensor_depth is useful for things like ADCPs, where the data depths are
specified but the actual deployment depth of the sensor may be important to
know.
>>> 23. Suggest following global attributes:
>>> author
Agreed.
>>> 24. Reg. "TIME" variable:
>>> Should have an attribute with "reference time" (e.g. 1950/1/1).
>>> Should have a "standard_name" in ref. table 3.
Yes, this would be a better solution that having the reference time in the
long name (TIME:long_name = "days since 1950-01-01 00:00:00 UTC")...
although we are *requiring* that specific reference time, so it is not a
"variable"
attribute.
>>> 25. Reg. all variables:
>>> Do we want an "accuracy" attribute? We already have "resolution" (I
>>> guess that means precision).
Yes, these are different.

Is anyone working on a revision to the users' manual? It would be great to get
some of Matthias' suggestions into the document.  I'd love to work on this,
but I'm not sure I would be able to find the time in the next couple of weeks,
and it would be nice to have a few iterations before the meeting in April.
```

Matthias Lankhorst, 10/03/2008

```
Hello, Bill,

I have looked at your file and found nice progress. Some issues remain, as you
probably expect/know already. These are:

1.
I really feel we need a standardized set of (global) attributes (mandatory and
in a well-defined format). This has less to do with "your" file but rather
with the OceanSITES definitions. We should discuss this at the Vienna meeting
```

and come up with a definite description in the "OceanSITES User Manual". Your
new file is clearly trying to fill more of the attributes than the previous
version.

2.
We will need *_QC variables for every variable. It is legal to have them in
there but all filled with zeros (meaning "no QC performed")...

3.
I like your variable names - we need an updated list of "approved" names in
the OceanSITES User Manual. I am guessing that this is on its way, and that
your names are in sync with that. You have at least one variable (VDEN) with
another axis (FREQ), which probably deserves an extra description in the
manual.

Looking forward to seeing you in Vienna,   MATT

On Monday 10 March 2008 11:12, Bill Burnett wrote:
> Matthias, Uwe, Sylvie and Thierry,
>
> NDBC OceanSITES - take 2. We received great comments from Matt and
> Aurélie, and we used the comments to reformat our Woods Hole data.
>
> Hopefully this dataset is closer to the approved OceanSITES format.
> Let me know if it is.
>
>
>http://dods.ndbc.noaa.gov:8080/opendap/oceansites/OS_32012_200712_TSM.n
>c.ht
>ml
>
> Cheers,
> Bill
>
> Matthias Lankhorst wrote:
> > Dear Bill,
> >
> > Uwe has asked me to look at the data that you posted on the ftp
> > server, i.e. the single test file "OS_32012_2008_w.nc" as of today.
> > From a quick look, I found several inconsistencies vs. the
> > OceanSITES user manual (v1.0 from
http://www.oceansites.org/docs/oceansites-user-manual.pdf).
> >
> > These seem to be common; I have been in contact with Mike McCann
> > from MBARI recently about the same issues. I have seen similar
> > problems in the past with other files on the Ifremer server, which
> > have at least partially been addressed. Let me briefly explain what I
found:
> >
> > 1.
> > The manual calls for a specific file naming convention (section 5.1).
> > Your "*_w.nc" does not match anything of reference table 1 in
> > section
> > 4.1.2 of the manual. I am guessing it should be "*_M.nc" for
> > meteorological measurements.
> >
> > 2.

```
> > For every parameter PARA in the data section, there should be a
> > quality indicator called PARA_QC, which I did not find.
> >
> > 3.
> > The parameter names ("variables" containing the data) should follow
> > a different naming convention, outlined in section 4.3 of the manual.
> > Essentially, all the names are four-letter and CAPITALIZED. I
> > totally acknowledge that many variables of your test file are yet
> > without well-defined names (cf. tables in manual section 4.3) and
> > suggest we define them asap, maybe during the April meeting in
> > Vienna - or even earlier?
> >
> > I hope you will find these comments useful.
> >
> > Kind regards, MATT LANKHORST
> >
> >
--
```

Meghan F Cronin [Meghan.F.Cronin@noaa.gov] 02/02/2008

```
I plan to attend the OceanSITES meeting 10-12 April and have submitted an
abstract to the OceanSITES session of the EGU.

In addition to the topics below, could the OceanSITES group discuss
having OceanSITES data on the GTS?   The WMO number can have a special
"84" designation that identifies it is reference site data. Operational
forecast centers are very keen on assimilating any data available and really
if these data help operational forecasts, shouldn't they be included?
Reanalyses could withhold the data so that the reference data can be
independent in assessment studies. I've found though that even when the data
are assimilated (such as TAO), they can still be used to assess systematic
biases. In anycase, this is a decision that each site with near-realtime data
needs to make.
```

Matthias Lankhorst, 31/01/2008

```
Hello all,

I have put a sample NetCDF file, which I believe to conform to the current
OceanSITES standards, on the following URL:
```

http://www-pord.ucsd.edu/~mlankhorst/OS_MOVE-V404-1_200001_TS.nc

```
Feel free to look at it, criticize, use as template, ...

Ciao, MATT
```

John Graybeal [graybeal@mbari.org] 28/01/2008

Thanks for initiatng this update, Steve. A more significant report will be drafted in coming weeks, so you'll be hearing more about this topic.

The presentations, materials, and outcomes will go onto a site, but may only include material incorporated as of the meeting -- not sure if the OceanSites work is in that category. A lot of good material and deep research was accumulated. And IODE is talking about developing a new web site that would bring together the recommended information, as well as the process (e.g., for submitting a standard for consideration).

There is some intent to go forward on standards decisions and process, as described in the second bullet, but the applicability of this to QC was a little unresolved in my mind. A number of concrete recommendations for prioritizing and consolidating QC documentation were captured (not sure when those notes will be widely available, but some outcomes will be documented soon I think).

I believe your third bullet, but interestingly can't recall that outcome at this instant -- it was a day with some personal dropouts.... That was the general tenor of the meeting and participants, though.

I tried pushing the idea on several fronts that without a QC/QA-focused site to accumulate all the existing and best practices -- QARTOD could be that but hasn't shown that ambition so far -- it will be hard to initiate a community practice. IODE may or may not step up to that plate -- they aren't exactly organized bottom-up. There is potential out there for the picking, if someone wants to do it.

John

At 9:25 AM -0800 1/28/08, Steve Hankin wrote:
>Hello all,
>
>Quick, preliminary feedback here from the Oostende meeting last week:
>(John, I hope you'll please fact check my words below.)
>
>All of the presentations and outcomes from the meeting will be posted on the Web -- probably linked from
<http://www.iode.org/index.php?option=com_content&task=view&id=74>http://www.iode.org/index.php?option=com_content&task=view&id=74.
>The OceanSites data format and procedures will, I believe, be included
>on that site as a reference document Tentatively , an outcome from the meeting will be the formation of a JCOMM/IODE "standard process" -- a Web site and committee structure for JCOMM/IODE community-wide evaluation of standards and best practices.
>There was extensive discussion of QC procedures. A game plan was outlined to begin to bring together and harmonize the QC procedures developed by separate communities on a by-platform basis (Argo, GOSUD, GTSPP, OceanSites(?), etc.). Presumably the results of such a harmonization plan (developed by a volunteer group) would be put through the JCOMM/IODE standards process for review and approval.
>Stay tuned for greater detail at the upcoming Web site.
>
>   - Steve
>
>========================
>
>Steve Diggs wrote:

```
>
>>All,
>>
>>I've known about this meeting for a while and it's interesting that there
aren't any attendees from any of the major academic oceanographic
institutions.  I know that the US institutions (RSMAS, SIO, WHOI, GSO, U/W and
TAMU) weren't aware of this meeting, which will make it challenging for any of
the standards to have traction with some of the primary producers of global
ocean data.
>>
>>That said, I agree with Nan that it would be a very good idea to see how any
of the practical solutions formulated at this month's meeting in Oostende
might be quickly incorporated into programs such as OceanSITES.  I look
forward to hearing Steve and John's synopsis of this very important meeting.
>>
>>-sd
>>
>>On Jan 7, 2008, at 2:07 PM, John Graybeal wrote:
>>
>>>Me too.
>>>
>>>Yes, one ot the topics we'll talk about at the meeting is how to bring our
'conclusions' back to the community and get feedback and engagement.  We may
have a cart-horse problem but there are a number of people attending who are
pushing hard on the community aspect.  So hopefully we can feed things back in
a (even more) graceful way.
>>>
>>>John
>>>
>>>At 12:54 PM -0800 1/7/08, Steve Hankin wrote:
>>>
>>>>Nan Galbraith wrote:
>>>>
>>>>>Hi OTS-DM members -
>>>>>
>>>>>I'm wondering if DM team members will be attending the IODE/JCOMM
>>>>>Forum on Oceanographic Data Management and Exchange Standards later
>>>>>this month.
>>>>>
>>>>Hi Nan,
>>>>
>>>>I will be there.
>>>>
>>>>It may have been my email to Sylvie and Uwe regarding this meeting, that
triggered this discussion.  I was checking to see about the existence of any
background slides on the OceanSites standards process (not the technical
standard, itself) that might be suitable as part of the overview materials for
this meeting.
>>>>
>>>>   - Steve
>>>>
>>>>=================================
>>>>
>>>>>It seems that the outline
>>>>>for that meeting includes topics that are pertinent for our
>>>>>discussion in April:
>>>>>
>>>>> * METADATA, marine profile (& meteorological profile)
>>>>> * QUALITY CONTROL
>>>>> * DOCUMENTATION
```

>>>>>
>>>>>I'm hoping we'll have access to the proceedings of that meeting for
>>>>>use in making progress on these areas; it seems like it would
>>>>>really benefit us to be able to adopt or adapt any recommendations
>>>>>they come up with, or to at least be able to discuss their approach.
>>>>>
>>>>>That meeting is posted at
>>>>><http://www.iode.org/index.php?option=com_content&task=view&id=74><
>>>>>http://www.iode.org/index.php?option=com_content&task=view&id=74><h
>>>>>ttp://www.iode.org/index.php?option=com_content&task=view&id=74>htt
>>>>>p://www.iode.org/index.php?option=com_content&task=view&id=74
>>>>>
>>>>>Cheers -
>>>>>Nan
>>>>>
>>>>
>>>>--
>>>>Steve Hankin, NOAA/PMEL --
>>>><mailto:Steven.C.Hankin@noaa.gov><mailto:Steven.C.Hankin@noaa.gov><m
>>>>ailto:Steven.C.Hankin@noaa.gov>Steven.C.Hankin@noaa.gov
>>>>7600 Sand Point Way NE, Seattle, WA 98115-0070 ph. (206) 526-6080,
>>>>FAX (206) 526-6744
>>>>
>>>>"The only thing necessary for the triumph of evil is for good men to
>>>>do nothing." -- Edmund Burke
>>>>