

Interactive comment on “The AlborEX dataset: sampling of submesoscale features in the Alboran Sea” by Charles Troupin et al.

Anonymous Referee #2

Received and published: 24 September 2018

— GENERAL COMMENTS —

Please find below my review of the manuscript entitled "The AlborEX dataset: sampling of submesoscale features in the Alboran Sea" by Troupin et al. I think the data and the paper are relatively well presented. I especially enjoyed that all the files are netCDF format. While the data are limited to a very local application (a 6-day experiment from one sub-region of the Mediterranean Sea), the data are in high-quality and may be useful for process-related studies. Overall, the manuscript may be suitable for publication after moderate reviews. This decision is detailed below.

— MAJOR COMMENTS —

My major concerns on the actual version of the paper are the following:

C1

1. I think the text is not well organized. Some info on the data is found in Section 2 (AlborEX mission) and in Section 3.3 (Data Processing). This spreading of information makes the search for information through the paper difficult. I would bring Section 3.3 earlier in the paper and avoid to spread the information for each platform in different sections. Some specific comments below are related to this problem (e.g. mention of flags even before introducing them).

2. The QC control is a weakness in this manuscript as it suggests that some QC is done, but it is not very clear on which data and how it is done. For some instruments, QC flags and their meaning are embedded in the files (e.g. float and drifters), but some doesn't (glider files). This inconsistency is not so much a problem to me as long as it is clearly stated in the paper which files contains QC flags. These quality flags should however be defined in the text. There are several mentions of "quality flags" in the text and figure caption, but little explanation is provided on these. Figure 12 has 9 quality flags that are not even described (although I see their meaning in drifters and float files). Where the QC is easy to reference (e.g. "file generated with Socib glider toolbox vX.X", or "File QC done using Socib standard procedure following a procedure described in a certain paper", etc.), it should be mentioned in the netCDF file as well.

3. Why all processing levels are not provided? The text suggests that all levels are provided (e.g. Table 3), but at the moment mostly L1 is provided. For gliders, L1 and L2 are provided. For the Float, L1 is provided for Arvor-A3 and Provor-Bio, but L0 for Arvor-C. Why? No explanation for this is provided (I think float data should be provided in L1 and L2 level as well). If some QC is applied on L1, maybe L0 should be provided as well to the future user? For glider L2 data, a choice is made regarding the vertical binning of the profiles. Which size these vertical bins are? This information should be provided somewhere.

4. Nowhere the sensor configurations are specified. I think a table gathering this information is worth it. For each platform, the list of sensors should be presented with their configuration (sampling frequency, ADCP ping-per-ensemble, ADCP vertical bin

C2

size, etc.). This should include all variables collected, for example, from the ship meteorological station from which little information (or none) is present in the text. Same for the glider where there is Chl-a and turbidity data in the files, but these were not mentioned in the text. A table gathering this information would be useful.

5. A table regrouping all the platform with their basic configuration as well as their number of casts (when it applies) should be provided (sort of extended Table 3).

— TEXT-SPECIFIC COMMENTS —

- Figure 1 too small (should take page width)
- Figure 2 too small (should take page width)
- Figure 2 caption: there is mention of "flag data equal to 1" while these flag are not introduced in the text.
- p.7, L1: The "total number of valid measurement" is not very useful. I would rather put the number of valid casts (see comment above on a new table with this info).
- p.7, L6: "a spatial interpolation is applied on the original data, leading to the so-called Level-2 data, further described in Sec. 3.3." What does 'spatial interpolation' means? Section 3.3 is not very explicit on this. I know you mean that the glider yos have been separated into downward and upward casts and then assigned to a geographical coordinate, but maybe this should be stated explicitly (and I don't think "spatial interpolation" is an accurate description). Moreover, Is there any vertical interpolation done? Because there are still some NaNs in L2 data.
- p.7, L15: "Interestingly, all the drifters exhibit a trajectory close to the front position" -> Not clear what "trajectory close to the front means". Moreover, is that really surprising that surface drifter would aggregate on a front?
- Figure 8 caption: "for the duration of the mission" -> You mean the ship mission? Or the AlborEX campaign?

C3

- Figure 10: plots on the right column are of little information here (too low resolution to mean something), I would remove.
- Table 1: "Period" should be replaced by "cycle length" as referred to in the text (Section 2.2.4).
- Table 1: netCDF file for Provor-bio indicates deployment end date 2015-04-24T12:02:59+00:00, which is different from this table.
- Figure 11 caption: "quality flag" not defined.
- Section 3.3.1: A Section on processing levels, but they are not all provided. Why? I think all levels should be provided. This is related to a previous comment.
- p.14, Level 2 (L2): "obtained by interpolating the L1 data" -> How L2 is obtained by "interpolating" L1? Isn't L1 cut into casts that makes L2?
- p.14, Level 2 (L2): "It is only provided for gliders, mostly for visualization and post-processing purposes: specific tools designed to read and display profiler data can then be used the same way for gliders." -> Is there a problem with this sentence? I don't understand it.
- Section 3.3.1 / Table 3: Is L1 level for float equivalent to L2 level for glider? For consistency, I think profiling float should have L1 and L2 data as well since these instruments have similarities on the way they profile the water column...
- p.12, L1: "This type of current measurements requires a careful processing in order to get meaningful velocities from the raw signal" -> Why? What are the limitations that makes this instrument more sensitive compare to other ones?
- p.12, L4: "Figure 12 shows the QF during the whole mission." -> How QF are calculated?
- Figure 12: Too small.

C4

- Figure 12 and text below: 9 different quality flag are presented without any introduction on how they are calculated.

- Section 3.3.2 is very short. Should be re-worked following comments above.

— COMMENTS ON DATA FILES —

The dataset consists of a relatively large number of files. I did my best but it was nearly impossible to review them all in details. Here are some comments:

- There are very large spikes in deep glider turbidity

- There are missing data for about 10h in deep glider data between May 25-26. Unless I missed it, no explanation for this are provided.

- Oxygen data for both glider seems to suffer from thermal lag problems

- Provor-bio datafile contains levels down to over 7000m. Some problems are found: 1. Why such long level dimension? 2. No good data is found below ~325m, although Table 1 suggest that the float is profiling to 1000m

- Arvor A3 data file suffers from similar problem: file contains data only down to 115m while Table 1 says 2000m

- Arvor-C data file (only L0 provided) do not contain metadata (no file attributes, etc.). In addition, missing data (at least for temperature) appears to me as very large numbers (9.969210e+36) that makes them difficult to manipulate.

- R/V Socib CTD and thermosalinograph files say that units of temperature are "C". I prefer the convention from glider files which uses "Celsius".

— MINOR COMMENTS —

- p.2; L23: "makes it possible" -> makes possible - p.2; L23: "creation and publication of aggregated datasets covering the Mediterranean Sea" -> SeaDataNet is not only about the Mediterranean - p.2; L32: "thanks due to" -> thanks to - Section

C5

2.2.1: "CTD surveys" or CTD legs? - Glider L1 files (e.g. dep0012_ideep00_ime-sldeep000_L1_2014-05-25_data_dt.nc) say that the project is "PERSEUS". Is that right? There is no mention of the AlborEX project in the file header. - p.10, L1: problems with latitude longitude degree symbol. - p.10, L5: temperature, salinity and T,S is use on the same line. Please homogenize. - p.12, L17: "Network Common Data Form (netCDF, <https://doi.org/http://doi.org/10.5065/D6H70CW6>, last accessed on August 3, 2018)" Is there a mis-placed parenthesis? - p.13, L2: problem with file name (too long for page) - p.16, L25: How stable in time the python codes made available on Github will be?

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2018-104>, 2018.

C6