ESSDD

Interactive
comment

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

Anonymous Referee #1

Received and published: 13 September 2018

This article (categorized as "review") by Troupin et al. is addressing a multidisciplinary data set collected in the western Mediterranean Sea during the AlborEX campaign. During the campaign in-situ observing devices (ships, floats, gliders, drifters. . .) have been used (described here) but also satellite data. In the manuscript some aspects of the data set are described. As it stands now I do not recommend publication in ESSD.

For the review I followed the ESSD evaluation criteria and also considered the general scope of the journal (as described on the website).
First - Is this a "review" article? ESSD defines review articles as:
". . .may compare methods or relative merits of data sets, the fitness of individual methods or data sets for specific purposes, or how combinations might be used as

more complex methods or reference data collections."
As I read it from the manuscript this is not the case. The current version of the manuscript reads more as a copy of data information from individual reports and the data section in scientific publications related to the experiment. As it stands, I do not see the criteria for a "review" type article fulfilled.

Significance
Three sub-criteria to evaluate:

- Uniqueness: It should not be possible to replicate the experiment or observation on a routine basis. Thus, any data set on a variable supposed or suspected to reflect changes in the Earth system deserves to be considered unique. This is also the case for cost-intensive data sets which will not be replicated due to financial reasons. A new or improved method should not be trivial or obvious.
  The data set is unique.
  (rating: 1 Excellent)

- Usefulness: It should be plausible that the data, alone or in combination with other data sets, can be used in future interpretations, for the comparison to model output or to verify other experiments or observations. Other possible uses mentioned by the authors will be considered.
  The current manuscript does not provide information that promote the reuse of the data set (it may for subsets). No attempt is made to provide a structured overview about the workflow that is linke dot the creation of the data set and, equally important, the QA/QC are not provided in a transparent way. For example, in the netcdf data files I see different QC flags provided – one is for example "SOCIB Quality control Data Protocol". What does that mean? This is not an international standard. A data set descrip-

tion, as envisioned in this ESSD submission, should exactly describe such
non-standard QC procedures. Which QA and QC methods were applied
(give brief description, DOIs if applicable)?
I also miss any information how/if this data is disseminated via interna-
tional data centres and how the data QC and dissemination is coordinate
with the respective observing networks (Argo, DBCP, . . .). Seadatanet is
been mentioned in the text but it is unclear which specific recommenda-
tions are given.
(rating: 4 poor)

• Completeness: A data set or collection must not be split intentionally, for exam-
ple, to increase the possible number of publications. It should contain all data that
can be reviewed without unnecessary increase of workload and can be reused in
another context by a reader.
It is difficult to evaluate this point. However, the nutrient data is not men-
tioned but is, according to Pascual et al. 2017 part of the AlborEX cam-
paign. I would expect that these data set are described here as well (and
respective QC (e.g. GO-SHIP nutrient manual??) and associated uncer-
tainty estimates
(rating: 2 to 3)

Data quality
The data must be presented readily and accessible for inspection and analysis to
make the reviewer's task possible. Even if a data set submitted is the first ever
published (on a parameter, in a region, etc.), its claimed accuracy, the instrumentation
employed, and methods of processing should reflect the "state of the art" or "best
practices". Considering all conditions and influences presented in the article, these
claims and factors must be mutually consistent. The reviewer will then apply his or her
expert knowledge and operational experience in the specific field to perform tests (e.g.
statistical tests) and cast judgement on whether the claimed findings and its factors –

individually and as a whole – are plausible and do not contain detectable faults.

I touched on that already under "Usefulness". In the manuscript no transparent QC assessment is presented. What were the methods of processing (provide key steps, DOI at least). What were, including quantification of uncertainties and qualification via flags, the results of the QA/QC procedures? Which were the major shortcomings of the data acquisition process and what could be done better in the future? For example, has the drifter data included in the European E–SurfMar data base and also in the DBCP global drifter data sets? Have the recommendations (Best Practices, Protocols) from E–SurfMar / DBCP considered? It looks like no commonly agreed standard has been used for some paramters – as "SOCIB Quality control Data Protocol" suggest?
(rating: 3)

Presentation quality
Long articles are not expected. Regarding the style, the aim is to develop stereotypical wording so that unambiguous meaning can be expressed and understood without much effort. The article should express clearly what has been found, where, when, and how. The article text and references should contain all information necessary to evaluate all claims about the data set or collection, whether the claims are explicitly written down in the article, or implicit, through the data being published or their metadata. The authors should point to suitable software or services for simple visualization and analysis, keeping in mind that neither the reviewer nor the casual "reader" will install or pay for it.
mostly OK (given the limitation outlined in the previous points). It would be useful to include a brief introduction into the "design of the experiment. Visualisation tools are not given.
(rating: 2-3)

Specific comments:

P2/l.4: I do not agree with the statement: "a perfect observational system would consist in dense array of sensors present at many geographical locations, many depths and measuring almost continuously a wide range of parameters…" – this "generalization" is trivial and useless. From an observing design point of view a "perfect" observing system must follow a design that will record only the observations that are needed to analyse the problem. As such the perfect observational system always depends on motivation for the experiment (or the problem in more general words) - in some cases a "perfect observing system" may comprise only one single sensor at one single depth at different locations if this has been found a sufficient approach for solving the problem (e.g. estimating global warming through a global tomography array). Please reformulate the statement along those lines.

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