

Interactive comment on “The CORA 5.2 dataset: global in-situ Temperature and Salinity measurements dataset. Data description and validation” by Tanguy Szekely et al.

Anonymous Referee #1

Received and published: 6 February 2019

Review of

The CORA 5.2 dataset: global in situ Temperature and Salinity measurement dataset. Data description and validation.

by T. Szekely et al.

C1

Recommendation

Minor revisions.

Synopsis

CORA is the Copernicus *in-situ* data set of temperature (T) and salinity (S). Version 5.2 covers the period 1950-2107. The data consist of vertical profiles from XBT, CTD, XCTD, Argo, and moorings as collected and validated by the Coriolis data centre. As the title implies, the paper is a thorough description of the data sources and the applied quality controls (QC). Through a cooperation with EN4 new profile data have been incorporated into the data base.

Although most data have been quality-controlled by their PI before being added to the Coriolis database, all are passed through the QC process again. The separate QC steps are described. The largest change comes from replacing Argo data that have undergone near-real time QC (automated) directly after becoming available by (semi-automatically) delayed-mode QC'ed values. The impact of the additional QC effort on the resulting data set is described. A notable result is a large reduction and homogenisation (in time) of the variability of the dataset, suggesting that no large errors remain in CORA 5.2.

Discussion

Users of a dataset or a collection of datasets should know how the data were collected and prepared, and which issues were tackled during the QC process. This information is given for CORA 5.2 in this paper, and it should therefore be published.

C2

The main data source for the recent years is Argo. Therefore, the authors should add a short paragraph explaining the relation between CORA and the recently published Argo climatology (Gouretski, V.: World Ocean Circulation Experiment – Argo Global Hydrographic Climatology, Ocean Sci., 14, 1127-1146, <https://doi.org/10.5194/os-14-1127-2018>, 2018).

At the end of sect. 5 the authors state that they do not think that their quality-controlled data are overflagged. This is a very important conclusion, and it should receive more attention. Especially, the reasons given for this conclusion should be backed by more evidence.

Detailed comments

The paper would benefit from thorough language editing.

p 2, l 5-7 three times “scientific community” - boring, please reformulate

p 2, l 23 timeseries → time scales

p 2, l 25 Baseline → baseline

p 4, l 192 barely → slightly? I am not sure what you want to say.

p 4, l 196 barely maintain a plateau at 20% → reach a plateau just below 20%? Again I am not sure what you want to say.

p 5, l 125 for each of the test described in this section a reference should be given so that the interested reader can easily find more information about the test - what does it look for, what are acceptable parameter values to be used in the test, how does it perform, etc.

C3

p 5, l 131-133 spike in what variable? From the description it seems to be a spike in $\frac{dT}{dz}$, but that's not clear from the text. Please explain.

p 5, l 141 possible measurements → possibly correct measurements?

p 5, l 148 who → which

p 6, l 179+180 °PSU → PSU

p 7, l 191 what do you mean by “hedgehog” type profile? spiky?

p 7, l 197 was → was? or is?

p 7, l 206 reinforced → increased

p 8, l 232 If there is a subsection 5.1, why isn't there a subsection 5.2?

p 8, l 239 lead in → lead to, or: result in

p 8, l 244 ocean flags → ocean quality flags into account. Apart from this, the sentence is hard to follow and should be reformulated.

p 8, eq. 1 define $l(i, j)$ and $L(i, j)$

p 9, eq. 2 define \bar{T}_p

p 9, l 295 differs from → differ by

p 9, l 301 what is GTSP? GTS is mentioned before, but what does PP stand for?

p 9, l 302 erased → deleted

p 12, l 396 insolation → insulation

p 12, l 404 contrary to what is stated here, fig. 9 only shows T

C4

p 13, 1424 as close as possible from the physical measurement → as close as possible to the physical measurement - I am not sure, perhaps better to reformulate the sentence.

p 13, 1438 something is wrong with the end of this line

Figures in all figures the labels are too small. I cannot read them.

Fig. 2 the colour scale is counter-intuitive. Low values should be blue and high values red.

Figs. 7+8 why not combine these figures into one?

Figs. 9+10 why not combine these figures into one?