

General Comments:

The manuscript presents a **novel** and **regionally important** data set: a high-resolution, four-dimensional sea-temperature climatology for the Northern Chilean Patagonia. The climatology is based on approximately three million in situ observations spanning a 75 years period (1948-2023) and provides monthly fields at 32 vertical levels from the surface to 400 m. The Data-Interpolating Variational Analysis in n dimensions (DIVAnd) software tool is used to interpolate the in situ measurements to a regular grid of an approximate horizontal resolution of 900 m. The dataset has clear potential for model validation, environmental monitoring and management, marine heatwave detection and characterization, and other potential regional applications.

The overall methodology is appropriate, as DIVAnd is a powerful, internationally recognized, and scientifically robust approach for generating climatologies from spatially and temporally irregular in situ observations. However, some further information on aspects such as the DIVAnd analysis error maps or the validation of the deeper layers, would strengthen this work.

The manuscript is generally **clear and well structured**. Some inconsistencies, presentation and language issues should be corrected.

Overall, the dataset is **scientifically valuable** and fits very well suited to the scope of ESSD. I recommend its publication after minor revisions addressing the points described below.

Specific Comments:

1. **Spatial and temporal data sampling heterogeneity:** The manuscript explains that the Comau Fjord alone accounts for approximately 76.83% of all samples, while several other regions contribute only a small fraction of the total dataset. In addition, around 70% of all observations are located within the upper 25 m of the water column. There is a strong also a strong temporal imbalance of the input data: 99% of the measurements were collected after 1990 and 55% after 2020 (if considering only non-mooring data).

Such heterogeneous distributions in space and time of the input data could potentially introduce biases into the analysis. Although the study acknowledges this heterogeneity, I think its possible influence on the resulting climatology is not discussed in sufficient detail. For example, sampling during certain winter and spring months appears to dominate in the north-western part of the study area, which could potentially influence the results. In addition, both seasonal background and monthly climatological fields may be affected by the greater availability of observations from anomalously warm or cold years.

Authors could address this issue by complementing Figs. 2 and 3 with maps showing the spatial distribution of observation counts by month and depth and by discussing, where feasible, the potential effects of the uneven spatial and temporal distribution of the input data on the resulting climatology.

2. **Validation of the results:** The product is described as four-dimensional, but the main validation table only evaluates the surface layer, between 0 and 1 m. As the dataset

is extending to 400 m and results are shown up to 100m, for the completeness of the work the authors could provide validation metrics for additional depth range(s).

3. **DIVAnd error maps:** An important advantage of the DIVAnd approach is its ability to provide analysis error maps which take into account the accuracy of the observations and their distribution, thereby helping to assess the reliability of the gridded fields and identify areas with poor observational coverage.

Given the highly heterogeneous spatial and temporal distribution of the input observations, analysis error maps would provide additional useful information on the spatial confidence of the interpolated climatological fields. Could the authors comment if DIVAnd analysis error fields were calculated and, if available, consider including these in the distributed climatology?

4. **Daily climatology:** Could the authors comment why the manuscript provides limited information and no specific evaluation of the daily climatology? At line 418 it is stated that “All climatological fields generated in this study, monthly and daily at 32 different depth levels ...”. However, the daily climatology dataset is provided at only 14 depth levels, covering the upper 0-100m. Please clarify this discrepancy.

Other comments

5. **Dataset link in Abstract:** based on the journal technical instructions for the manuscript composition a functional data set DOI and its in-text citation must be given in the abstract (<https://www.earth-system-science-data.net/submission.html#manuscriptcomposition>).
6. **Coordinates:** There are several coordinate inconsistencies in the notation of the geographic coordinates. For example, at line 81 or in the zenodo metadata (<https://zenodo.org/records/18256000>), the study area is described as extending from 41.4° S to 46.7°S and from 72°W to 74°W, whereas in line 104, Table 1 and elsewhere the coordinates are given as (-75.1°E, -72.0°E and -41.4°N, -46.7°N). Please use a consistent notation throughout the manuscript.
7. **Depths:** While negative values are used in Figure 1 and its caption, positive values are used throughout the manuscript. I suggest using a consistent depth convention throughout the manuscript.
8. **Line 14:** Change “Data Interpolating Variational Analysis (DIVAnd)” to “Data Interpolating Variational Analysis in n dimensions (DIVAnd)” to distinguish DIVAnd from the earlier DIVA method. For the same reason, replace “DIVA” to “DIVAnd” throughout the manuscript.
9. **Line 106-107:** The data repositories mentioned use different quality flag scales. Could you specify which quality flags correspond to valid data and which indicate “obviously” erroneous data?
10. **Table 1 – row 1:** Perhaps use "Entire study area" instead of "All study area"?
11. **Line 249-250:** I would suggest here to mention the format of the dataset.

12. **Line 385-388:** the current analysis produces mean monthly (and daily) climatological fields based on a dataset spanning a 75-year period from 1948-2023, but it does not assess temporal changes or shifts over this period. The fact that the data are strongly clustered within the last three decades, maybe bias the analysis towards the mean state of the 1990-2023 period, rather than providing a representative mean state for the entire 1948-2023 period against which new observations can be compared. It would be very interesting in future works to calculate decadal seasonal fields, or climatological fields for selected time periods, to identify possible changes through time. I would therefore suggest revising this sentence to better reflect the scope and any potential limitations of the current analysis.

TEXT editing and improvement

13. **Line 97:** The Jacaf Channel is not shown on the map. Shouldn't it be labelled, as are the other geographic features mentioned in the text?
14. **Figure 1 caption:** I am not a native English speaker, but I think the Figure 1 caption could be rephrased for clarity. A possible suggestion is: "Map of the study area with bathymetry. Depth contours are shown at 500 m intervals, while shallow bathymetry (i.e. values above -400 m) is indicated by color shading".
15. **Figure 2 caption:** To give more information, you could change the "non-mooring data" to "non-mooring data (profiles and trajectories)".
16. **Figure 3 caption:** please use "Black diamonds" instead of "Black diamond".
17. **Line 163:** I would rephrase the sentence to "The input to provide is a dataset of potentially irregularly distributed observations with ...".
18. **Line 180:** change the "were processes" to "were processed".
19. **Line 199, Table XX:** please use the correct table number.
20. **Line 233:** isn't austral winter "July-August-September" as defined previously at lines 168-169?
21. **Table 4 caption and content:** Change "CRMS" to "CRMSE", or use these terms consistently throughout the manuscript.
22. **Line 354:** Change "Corcovado Gulfs" to "Corcovado Gulf".

Comments on the Supplement material

23. **Line 41:** I cannot find any table. Do you mean Table 1 of the paper? Please clarify.
24. **Line 270 - S2.3 EMODnet:** The cited references correspond to reports from the Copernicus Marine In Situ service rather than links to the data themselves. Therefore, the data source appears to be CMEMS rather than EMODnet. I suggest changing the source in S2.3 from EMODnet to CMEMS, while retaining the sentence "The data set extracted from Ocean Browser ..." to clearly distinguish the data source (i.e. CMEMS)

from the data distributor (i.e. EMODnet). It would also be useful to replace references 55–59 with the corresponding data DOIs, if feasible.

However, in the first row of Table 1, CMEMS is also listed as the source of the drifter data. Are these data different from those downloaded through the EMODnet Ocean Browser? Please clarify.