We would like to thank the reviewers for their careful and constructive evaluation of our manuscript. Their comments and suggestions have helped us improve the clarity, transparency, and scientific robustness of the manuscript. We have carefully considered all comments and have revised the manuscript accordingly. In particular, we have clarified the distinction between the monthly and daily-derived climatologies, provided additional information on the validation and uncertainty of the reconstructed fields, and further discussed the spatial and temporal limitations of the observational dataset.

Below, we provide a point-by-point response to each comment.

Reviewer 1 :

Referee Report - Manuscript essd-2026-42

Manuscript Title: Climatology of the Sea Temperature from Long-Term In Situ Observations:

Northern Chilean Patagonia.

Authors: Pujol et al. 2026

1. General assessment

This paper presents a four-dimensional temperature climatology for Northern Chilean Patagonia, covering the period 1948–2023 at approximately 900 m horizontal resolution and 32 depth levels between the surface and 400 m. The interpolation is performed with DIVAnd, a well-established variational tool that accounts for bathymetric constraints and data uncertainties. The work compiles roughly 3.3 million in situ observations drawn from several public repositories and regional monitoring programmes, merging them into coherent monthly and daily-derived gridded fields.

The topic is of genuine scientific relevance. Northern Chilean Patagonia is a poorly instrumented region where persistent cloud cover and complex fjord geometry limit the utility of satellite products, and where the expansion of aquaculture makes an accurate thermal baseline increasingly important. To my knowledge, no previous study has produced a spatially continuous, in situ-based climatology for this entire region at the resolution reported here, and the potential applications listed by the authors, from model validation to marine heatwave detection, are realistic and well motivated. The manuscript is clearly written, the figures are informative, and the discussion places the results appropriately within the existing literature on Patagonian oceanography. I have a number of concerns that I believe should be addressed before publication. Most of them relate to documentation and transparency rather than to the scientific substance of the work, and

I am confident they can be resolved without difficulty. I recommend minor revisions.

2. Main concerns

2.1 Undocumented difference in depth coverage between the two products

The monthly and daily-derived climatologies differ substantially in their vertical coverage: the monthly product spans 32 depth levels from 0 to 400 m, whereas the daily product covers only 14 levels from 0 to 100 m. This distinction is correctly reflected in the Zenodo repository metadata, but it is never stated in the manuscript. The abstract and conclusions describe the study as producing a product with "32 vertical levels" and coverage "from the surface to 400 m" without clarifying that these specifications apply only to the monthly file. A reader who downloads the daily file expecting full-depth coverage will find it limited to the upper 100 m. The manuscript should clearly state, at the first description of the daily product (Section 2.3) and in the abstract, that the daily-derived climatology is restricted to 14 depth levels between 0 and 100 m, and should briefly explain the rationale for this choice.

Response: Done. Added to abtrast, mat & mat and conclusion

2.2 Absence of subsurface validation

The validation presented in Table 4 is limited to the sea surface (0–1 m depth). While this is the layer with the most observations, the product extends to 400 m, and a user intending to employ the climatology as background for subsurface assimilation or model comparison has no quantitative basis for assessing its reliability at depth. The authors do show profile comparisons in Figure 8, which is helpful, but the comparison is qualitative. I would expect, at minimum, summary validation statistics (correlation, RMSE, bias) for a representative set of depth levels, or a reference to the validation procedure for depth. If a fully stratified validation is not feasible given the data sparsity at depth, the authors should be explicit about this limitation and note the depths below which the climatology should be used with caution.

Response: We thank the reviewer for highlighting the importance of assessing the reliability of the subsurface fields. We agree that the surface validation alone does not directly quantify the accuracy of the climatology at depth. We chose not to apply the same 5% withholding procedure at depth because the number of observations decreases substantially with depth, and removing an additional 5% would further reduce the already limited observational constraint and could disproportionately affect the reconstruction. Instead, we now provide DIVAnd relative analysis error fields, which quantify the remaining analysis uncertainty relative to the background and account for the spatial distribution of observations. We have added these fields to the material and methods, Results (section 3.2.3) and discuss them as a complementary assessment of observational constraint rather than as an independent validation of accuracy. We have also clarified this distinction in the discussion.

2.3 Spatial sampling bias

Table 2 reveals that 76.8% of all observations come from Comau Fjord alone. The authors acknowledge this but do not assess its impact on the resulting climatology. The concern is whether DIVAnd's spatial smoothing adequately compensates for the extreme undersampling of other basins, particularly in the southern part of the domain (Elefantes Channel: 0.32%, Chonos Archipelago: 0.95%). The validation in Table 4 does not distinguish spatially between well-sampled and data-sparse regions, so it is not possible to tell whether the performance metrics are dominated by Comau Fjord. I would encourage the authors to provide at least a regional breakdown of the validation statistics, or to test the sensitivity of the climatology in data-sparse areas by, for instance, comparing the full climatology with one produced excluding Comau Fjord data.

Response: Although some regions account for a relatively small proportion of the total number of observations (e.g. the Chonos region, which represents 0.95% of all observations), they are nevertheless supported by a substantial number of measurements (approximately 31000 observations in the case of Chonos). Comau Fjord clearly influences the overall statistics because of its disproportionately high sampling density, however, the other regions remain well sampled. For the interpolation, a correlation length of 50 km was used, meaning that the influence of the particularly dense sampling in Comau Fjord is spatially constrained and is expected to extend primarily towards Ancud Gulf. This is also consistent with the observed hydrographic connectivity of the region. For the surface validation, we additionally calculated the statistical metrics independently for the different regions, rather than relying solely on statistics calculated over the entire study area. These results are presented in the section 3.1.

2.4 Temporal representativeness

The data span 75 years in principle, but the authors note that 99% of observations post-date 1990 and 55% post-date 2010 (lines 130–131). The climatology is therefore predominantly a representation of recent thermal conditions rather than a true long-term mean, which has implications for its use as a baseline against which to measure trends or anomalies. This is not necessarily a fatal limitation, but the authors should discuss it more explicitly, for example, whether the product is better described as a climatology of the 1990–2023 period with historical anchor points, and how the underrepresentation of earlier decades might affect the mean state.

Response: we added a sentence in the material and methods (section 2.2) " The resulting climatology should therefore be interpreted as a climatological reconstruction based predominantly on recent observations (last three decades), rather than as an equally weighted 75-year climatological mean. "

and we added a paragraph in the discussion "The temporal distribution of the observations represents an important consideration when representing the climatology. Although the dataset spans 1948-2023, the observations are strongly concentrated in recent decades, with 99% collected after 1990 and 55% after 2010. The resulting product

should therefore not be interpreted as an equally weighted representation of the mean thermal state over the entire 75-year period. Rather, it primarily reflects the thermal conditions captured by the more densely sampled recent period, while earlier observations provide historical constraints that extend the climatological reconstruction back to 1948. Consequently, the climatology is best considered a multi-decadal reference field based on observations spanning 1948-2023, rather than a strict baseline of the 75-year period. "

3. Minor concerns and technical corrections

3.1 Inconsistencies in the manuscript text

Line 199 contains the placeholder “see Table XX for details”,

Response: done

Line 319: “4-dimentional” should read “4-dimensional”.

Response: done

Line 373: “Chacao Channal” should read “Chacao Channel”.

Response: done

Line 390: “associated wit harmful” should read “associated with harmful”.

Response: done

3.2 Very high NaN fraction at depth

At 400 m, approximately 94% of the monthly grid cells contain NaN, meaning that fewer than 6% of the domain has valid interpolated values at maximum depth. This is not in itself a problem, it reflects the narrow geometry of the fjords and the rapid increase in land fraction with depth, but it is worth noting explicitly so that users are not surprised. A brief statement or a figure showing the spatial coverage as a function of depth would help.

Response: Thank you for that important comment. We choose not to put a dedicated figure but we add a small paragraph in section 3.2.2 “The proportion of the study domain covered by interpolated values decreases progressively with depth, reflecting the bathymetry characteristics of the fjord environment. The interpolation covers 100% of the domain at the surface and remains relatively high in the upper 100 m, with more than 65% of the domain containing valid interpolated values. Coverage then decreases substantially with depth, reaching approximately 30% at 200 m and only 14% at 400 m. At 400 m, the interpolated area is primarily confined to the Peru-Chile Trench west of Chiloé Island and a few deeper-micro-basins within the fjord system. ”

3.3 File format and Zenodo deposit structure

The manuscript never explicitly states the format of the distributed files (NetCDF3-Classic, as confirmed by inspection). This should be stated in Section 2.3 and in the data availability statement, along with the coordinate reference system (WGS84, assumed). Users would also benefit from a note about recommended software for reading the files and, ideally, a short code snippet or Jupyter notebook in the Zenodo repository.

Response: format added to both section 2.3 and data availability. We also provided on the same Zenodo deposit two script (MATLAB and Python Jupyter Notebook) for easy data handling and visualisation. The scripts were mentioned in the code and data availability and in section 3.1.

The Zenodo description also notes that the cited deposit (zenodo.18255999) is a subset of a larger deposit (zenodo.14845077) containing the full dataset, including the in situ observations used to construct the climatology. The manuscript does not mention this parent deposit, which seems relevant: readers who wish to access the underlying observational data should be directed to it. I recommend that the data availability statement reference both DOIs and clarify what each contains.

Response: The larger deposit does not contain the in situ observations used to construct the climatology. These observations originate from datasets provided by other organisations, which have made them freely available through their respective repositories. We therefore do not want to redistribute these datasets through our own Zenodo repository.

The larger deposit was originally associated with a previous manuscript that has since been rejected, with the recommendation to split the work into two separate studies: one focusing on the development and description of the climatology, which is the aim of the present manuscript, and a second focusing on the application of the climatology to marine heatwave detection, which will be addressed in a future paper. The larger deposit was therefore originally linked to the corresponding preprint.

We initially retained the reference to this larger deposit because it had already been disseminated online in association with the climatology. However, we agree that its inclusion may cause confusion.

We propose to remove the reference to the larger deposit from the Zenodo record.

4. Recommendation

This is a solid piece of work that will be a useful resource for the oceanographic community working in Chilean Patagonia. The methodology is sound, the regional context is well described, and the presented results are consistent with previous knowledge of the area. The concerns raised above relate primarily to documentation and transparency rather than to the scientific substance, and I expect them to be resolved without difficulty. I recommend minor revisions.

Response: Thank you.

Reviewer 2:

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).

Response: Although some regions show proportionally low quantity of observations (e.g. Chonos being 0.95% of the total amount of observations), it remains a high quantity of observations (31 000 observations in the case of Chonos). Comau Fjord clearly bias the statistics but the other regions remain well sampled. Two paragraphs were added in the discussion.

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.

Response: We now addressed this in the discussion. We are now talking about the uneven distribution of the data and how it can interfere with the reconstruction of the climatology (e.g. higher relative error in the subsurface pacific ocean where subsurface data are sparse vs Reloncaví sound which has been very well sampled whose relative error is close to 0 even at high depth). For the validation of the surface, we applied the statistics to independent regions (instead of only showing statistics for the entire study area). We found a consistency with the region-wide validation results. These results are presented in the section 3.1.

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?

Response to comments 2 and 3: We thank the reviewer for highlighting the importance of assessing the reliability of the subsurface fields. We agree that the surface validation alone does not directly quantify the accuracy of the climatology at depth. We chose not to apply the same 5% withholding procedure at depth because the number of observations decreases substantially with depth, and removing an additional 5% would

further reduce the already limited observational constraint and could disproportionately affect the reconstruction. Instead, we now provide DIVAnd relative analysis error fields, which quantify the remaining analysis uncertainty relative to the background and account for the spatial distribution of observations. We have added these fields to the material and methods, Results (section 3.2.3) and discuss them as a complementary assessment of observational constraint rather than as an independent validation of accuracy. We have also clarified this distinction in the discussion.

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.

Response: Thank you for raising that comment. It is indeed a mistake. The daily climatology does stop at 100 m to not over-interpolate the deeper layer that lay on relatively few data. This was clarified in the text in the abstract, mat & met section 2.3 and in the conclusion.

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).

Response: Thank you. It has been added.

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.

Response: Done, everything appears now as °W and °S.

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.

Response: Done. Now bathymetry appears as positive everywhere.

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.

Response: done

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?

Response: We clarified this point as follow :

"Data selection involved retaining only values flagged as valid when quality indicators were available, which means observations classified by the respective providers as “good”, “valid” or “probably good”, whereas observations flagged as “bad”, “probably bad” “suspect, “missing” or otherwise invalid were excluded. Data considered as “obviously wrong” (such as temperatures superior to 30°C) were also excluded. "

10. **Table 1 – row 1**: Perhaps use "Entire study area" instead of "All study area"?

Response: Done

11. **Line 249-250**: I would suggest here to mention the format of the dataset.

Response: the dataset format was added in section 2.2 and in the data availability

12. **Line 385-388:** the current analysis produces mean monthly (and daily) climatological fields based on a dataset spanning a 75-year period from1948-2023, 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.

Response: This issue was addressed as follow : ". Since the climatology predominantly represents the last three decades, with earlier observations providing historical constraints, the dataset therefore provides a valuable reference for characterising the spatial distribution of thermal conditions and identifying thermally distinct environments within individual basins and fjords, supporting risk mapping for temperature-..." and was modified as well in different parts of the manuscript :

we added a sentence in the material and methods (section 2.2) " The resulting climatology should therefore be interpreted as a climatological reconstruction based predominantly on recent observations (last three decades), rather than as an equally weighted 75-year climatological mean. " and added a paragraph in the discussion "The temporal distribution of the observations represents an important consideration when representing the climatology. Although the dataset spans 1948-2023, the observations are strongly concentrated in recent decades, with 99% collected after 1990 and 55% after 2010. The resulting product should therefore not be interpreted as an equally weighted representation of the mean thermal state over the entire 75-year period. Rather, it primarily reflects the thermal conditions captured by the more densely sampled recent period, while earlier observations provide historical constraints that extend the climatological reconstruction back to 1948. Consequently, the climatology is best considered a multi-decadal reference field based on observations spanning 1948-2023, rather than a strict baseline of the 75-year period. "

And, indeed, it would be very interesting to realise decadal climatologies for the last 3 decades and see if we can observe a shift… thank you for the idea :)

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?

Response: Jacaf Channel has been added to the figure 1

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".

Response : rephrased for clarity

15. **Figure 2 caption:** To give more information, you could change the “non-mooring data” to "non-mooring data (profiles and trajectories)”.

Response: done + we changed "rhombus" for "diamond" for consistency with figure 3

16. **Figure 3 caption:** please use “Black diamonds” instead of “Black diamond”.

Response: done

17. **Line 163:** I would rephrase the sentence to “The input to provide is a dataset of potentially irregularly distributed observations with … “.

Response: The input to provide is a table of potentially irregularly distributed data, with associated coordinates (longitude and latitude), depth and date -->

The required input consists of a dataset of potentially irregularly distributed observations, with associated coordinates (longitude and latitude), depth and date.

18. **Line 180:** change the “were processes” to “were processed”.

Response: done

19. Line 199, Table XX: please use the correct table number.

Response: done

20. **Line 233:** isn’t austral winter “July-August-September” as defined previously at lines 168-169?

Response: To avoid confusion, we modified the sentence in : "Correlation coefficients exceed 0.7 from June to September, when surface waters are the coldest and show less spatial variation..."

21. **Table 4 caption and content:** Change “CRMS” to “CRMSE”, or use these terms consistently throughout the manuscript.

Response: done

22. **Line 354:** Change “Corcovado Gulfs” to “Corcovado Gulf”.

Response: done

Comments on the Supplement material

23. **Line 41:** I cannot find any table. Do you mean Table 1 of the paper? Please clarify.

Response: done

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.

Response: yes, you are right. We modified the sentence as follow : "...shown in Table 1 of the manuscript" and we modified Table 1 and the supplementary table by aggregating everything as CMEMS data.

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.

Response: in the case of CMEMS CORA dataset, the bibliography about this product mentions both instruments were used. We also took the datasets found on Emodnet and can confirm that some of them were present in CORA while other were absent. In the case of repeated datasets between CORA and EMODnet datasets, our duplicate procedure has removed them.