Dear, we would like to thank you for your valuable comments and reflexions on our work. Your feedback has been very helpful and we are confident that it will contribute significantly to improving the quality and clarity of our manuscript. Please find below our detailed responses to your remarks, along with the corresponding revision made to our manuscript.

Unfortunately, while being well written, the manuscript suffers from several critical issues, and the methodology lacks both clarity and rigor—especially regarding the merging of satellite SST and in situ data. Line 205 states that "the detection of MHWs and MCSs is based on the comparison between daily climatology and daily temperature exceeding a threshold derived from the climatology. Here the climatology was calculated using in situ data, while the threshold a daily temperature were derived from satellite data". **Why do you need a climatology from in situ data**? MHWs are simply events when daily temperature exceeds a climatological threshold (typically the 90th percentile) for a few days. **How exactly are the satellite and in situ datasets integrated (which is in the title) and why? "Climatologies" sometimes refer to daily, then monthly means**, **and it is not clear if they are about the 50th or 90th percentiles**. Also, please note that, when combining datasets, careful attention must be paid to **differences in baseline periods (potentially requiring removal of temporal trends) and to the spatial variability each dataset resolves**.

Why do you need a climatology from in situ data

- We choose to construct a climatology using in situ data because it provides a more accurate representation of fjords and narrow coastal channels, which are often poorly resolved or entirely missing in satellite observations (due to high cloud cover, narrow spatial scale and land contamination that limit the effectiveness of satellite-based products in such regions, see L126-135)
- Sufficiently high resolution satellite products generally do not cover a sufficiently long time period to support robust climatological analyses
- In the context of MHW detection, the climatology is particularly crucial. Therefore, ensuring a robust and representative baseline is a key component of our approach. We will ensure to make it clear in the manuscript.

We propose to add this sentence at the beginning of the section 2.2 : "Given that the climatology forms the core of our analysis, it was essential to use a reliable and representative dataset. The extensive spatial and temporal distribution of in situ measurements makes them well-suited for this purpose." And this one at the beginning of section 2.3: "However, despite being adequate for constructing the climatology, the inhomogeneity of in situ data in space and time limits their direct use for threshold calculations. We therefore choose to used satellite data for the threshold calculation…".

- To address the specific challenges of complex coastal environments, we propose a new methodology that uses in situ data as an alternative to satellite-based approaches, especially in regions where satellite perform poorly but where in situ sampling is more frequent. We will make it clearer in the introduction.

How exactly are the satellite and in situ datasets integrated (which is in the title) and why & attention must be paid to differences in baseline periods (potentially requiring removal of temporal trends) and to the spatial variability each dataset resolves.

- In our approach, satellite and in situ datasets are not merged but rather used in a complementary manner to take advantage of their respective strengths. Specifically, the climatology is constructed from in situ data, which offers greater reliability and spatial accuracy in the complex study area, while the satellite data, which provides a more homogeneous temporal resolution, is used for daily SST values and the calculation of the MHWs thresholds
- The explanation for combining the two datasets lies in the need to use a long-term, high-quality climatology (based on in situ data, >20 years) with the high temporal resolution of satellite data for MHW detection. This setup ensures both robustness in the baseline and responsiveness in detecting events.
- We propose to add at the beginning of the section 2.4: "In this study, satellite and in situ datasets are not merged but rather used in a complementary manner. The daily climatology is derived from over 30 years of in situ data to ensure spatial accuracy and long-term reliability in the complex study area, while satellite data, with its consistent daily coverage, is employed to determine the threshold for MHW detection. This methodology thereby combines a robust baseline with high temporal resolution for effective event detection."
- To ensure consistency between the two datasets, we applied a bias correction to the satellite dataset by subtracting the mean offset relative to the in situ climatology (L201-2014). This correction aligns the two datasets and improves the reliability of the derived SST anomalies. The SST anomaly is the result from the daily SST (satellite-based product) minus the climatology (in situ-based product). Our corrected results are consistent with findings from previous studies (see section 3.2 of the manuscript).
- We will describe more precisely in the manuscript how was calculated the SST anomaly. We propose to add to section 2.4 this sentence to clarify how SST anomalies were calculated: "The SSTa are also derived from in situ and satellite-based dataset: it is the result from the climatology (in situ-based product) subtracted to the daily SST (satellite-based product)".
- We acknowledge that the term "integrating" in the title might be misleading, as it suggests a full data merge. Instead, we propose to modify the manuscript's title into "combining in situ and satellite data in complex coastal environment"

"Climatologies" sometimes refer to daily, then monthly means

We indeed calculated different types of climatologies, all derived from in situ data, to progressively refine the temporal resolution and adapt to data availability:

- Seasonal climatology: As a first step, we constructed a seasonal climatology. While in situ data are abundant, coverage is not consistent throughout the year at all locations, making it impractical to compute a daily climatology directly from raw observations. The seasonal climatology served as a preliminary baseline but is no longer used in the core analysis or mentioned further in the manuscript.
- Monthly climatology: Build on the seasonal climatology, we then developed a monthly climatology. This step again used the in situ data but benefits from the seasonal baseline as a background to improve temporal interpolation. The monthly climatology is discussed and compared with existing literature in section 3.2 of the manuscript.
- Daily climatology: The final product used for MHWs and MCSw detection is derived from the monthly climatology by applying a 90-day moving mean. This smoothing approach allowed us to estimate robust daily values despite gaps in the raw data. All MHWs and MCSs analyses presented in the manuscript (sections 3.5 to 3.8.3) are based on this daily climatology.

We will clarify it in the manuscript and explicitly state which version of the climatology is used at each step of the analysis. In particular, we will ensure it is clearly indicated that the daily climatology is the one used for MHWs and MCSs detection.

it is not clear if they are about the 50th or 90th percentiles

We do not mention the 50th percentile in the document. The climatology is based on the mean temperature, not on the median.

Regarding satellite SST, the authors correctly note limitations due to cloud cover and coastal complexity, but then proceed to use these same data to derive daily 90th percentile thresholds. The use of **DINEOF to fill satellite gaps is reasonable in principle, but it is unclear why the method was applied separately for each year**. Indeed, the **EOFs shown in Figure 7 appear to display unrealistic small-scale offshore patterns**.

unclear why the method was applied separately for each year

- We applied the method separately for each year in order to adapt the number of optimal EOFs to the specific spatio-temporal characteristics of each individual year. This approach allows for better reconstruction of the dominant patterns of variability present annually, and helps mitigate DINEOF limitations when applied over long time periods.
- Because DINEOF is based on EOFs that capture the most dominant modes of variability, applying it over multi-decadal period (e.g. 20 years) tends to emphasize persistent large-scale patterns (such as annual cycle) at the expense of high-frequency variability. As a result, interannual signals, especially those of lower amplitude, can be smoothed out or poorly reconstructed, particularly in regions with high variability (e.g. Alvera et al. (2021) that applied DINEOF year by year on chlorophyll data https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2021.707632/full). We will precise it in the manuscript L194.
- We applied a 3-days temporal filter during reprocessing to reduce noise. While this may slightly reduce the ability to resolve very high frequency events, it enhances the overall quality of the reconstruction.

EOFs shown in Figure 7 appear to display unrealistic small-scale offshore patterns

- EOFs capture only the signals present in the data, representing as well small-scale variability.
- EOF 3 structure is present almost every year. It explains only 0.3% of the explained variance, and therefore it does not influence our results.

The monthly climatology from in situ data is created by interpolating "**3 million" measurements** between 0 and 400 m onto a 900 m grid. However, **no information is provided on the spatial or temporal distribution of these observations**. Interpolating a large number of measurements into a high-resolution grid does not ensure that the resulting dataset is representative. For example, 3 million measurements could easily be generated from a thermistor with 5-minute resolution over a few sites. **The observational dataset itself is not adequately described or visualized, which is a major omission**.

no information is provided on the spatial or temporal distribution of these observations & **the observational dataset itself is not adequately described or visualized**

- We acknowledge the importance of explicitly presenting the spatial and temporal distribution of the in situ observations, as the reliability of the interpolation (and the climatology it supports) depends strongly on the coverage of the input data. Without adequate spatial and temporal sampling, any derived climatology would lack robustness.

To address this, we propose to add two new figures to the manuscripts (see at the end of this document).

Figure +1 displays the spatial distribution of all in situ observations used in the study.

Figure +2 displays the temporal distribution of observations, distinguishing contributions from mooring platforms (which offer high temporal resolution but cover mainly the last decade) and non-mooring platforms (which extend further back in time).

- Regarding the influence of mooring data, we clarify (as mentioned L156-157) that these observations are assigned lower weight in the interpolation due to their clustered nature in order to avoid bias. While moorings account for a substantial portion of the total data points, particularly in recent years, their contribution is appropriately balanced in our methodology.
- In total, they are 5880 unique stations that have been sampled in Northern Patagonia. We will add this information to the manuscript. If we round the latitude and longitude to 2 decimals, they are still 4203 unique stations.
- We now include a new version of Table 1 (referenced in section 2.2), which provides an updated summary of the datasets and supports the validation of the climatology (see at the end of this document).

It is also **unclear what analysis was performed on the in situ data at depth, since all reported MHW statistics appear to be derived from satellite data** (although this is not specified in the figure captions). **At least the authors do not incorrectly apply surface-based thresholds to subsurface data, which would be inappropriate given the stratification** evident in Figure 5.

what analysis was performed on the in situ data at depth, since all reported MHW statistics appear to be derived from satellite data

- We confirmed that while climatologies were generated at multiple depths, the analysis of MHWs presented in this study is limited to the surface layer, based on satellite-derived SST data.
- The inclusion of the full-depth climatology in the manuscript serves for two purposes:

1) To make the dataset available for future research, including by local scientists and stakeholders who may require information at depth, particularly in regions where surface data alone do not adequately represent ocean conditions.

2) To provide fundamentals for future investigations into subsurface MHWs, which we plan to pursue.

- We will put several reminders in the Results section that although we developed several climatologies for different depths, MHWs statistics reported in this work are based only on surface data.

At least the authors do not incorrectly apply surface-based thresholds to subsurface data, which would be inappropriate given the stratification

- Indeed, we did not applied surface-based threshold to subsurface since Northern Patagonia shows significant vertical variations through the water column

Sections 3.5 to 3.7 present statistics on MHWs and MCSs, but these are largely descriptive and difficult to follow due to the extensive use of local fjord names. This section reads more like a regional technical report than a manuscript suitable for an international scientific audience, and I worry that the spatial variability is due to the interpolation of too sparse data (which is evident in Figure 6). Again, interpolation is not magic and should not be abused when the initial dataset it too sparse.

Sections 3.5 to 3.7 are largely descriptive and difficult to follow due to the extensive use of local fjord names. This section reads more like a regional technical report than a manuscript suitable for an international scientific audience

- We acknowledge that sections 3.5 to 3.7 are highly descriptive. This was a deliberate choice, as Northern Patagonia exhibits pronounced spatial heterogeneity in the sea temperature, and characteristics of MHWs and MCSs, with each basin responding differently due to local oceanographic and atmospheric conditions. A generalised description would risk oversimplify this complexity and omitting important localised dynamics.
- Moreover, Northern Patagonia is a region of high socio-economic important, supporting a large portion of Chilean aquaculture, which is concentrated within fjords and channels rather than in the open coastal areas.
- While this work is a regional focus study, it also serves a broader methodological purpose. By demonstrating that our approach captures fine-scale variability within complex coastal environments, we show its applicability to other complex areas and fjord systems worldwide.
- We understand the need to balance local relevance with broader scientific readability. We therefore propose to revise sections 3.5 to 3.7 to reduce the number of local place names to improve readability for non-local readers

Spatial variability is due to the interpolation of too sparse data (which is evident in Figure 6).
We respectfully disagree with the suggestion that the observed spatial variability in Figure 6 is an artifact of interpolating sparse data. On the contrary, the variability reflects well-documented real temperature differences across Northern Patagonia. These differences arise from known physical drivers such as freshwater inputs, stratification and the enclosed nature of some basins. Various independent studies have reported this variability. Please, have a look to the different examples below that support the high variability of the region:

- Puyuhuapi Fjord: temperature ranges from <8)c in winter to >18°C in summer, as shown in Figure 5 of Schneider et al. (2014) and Figure 2 of Pérez-Santos (2017). Our own findings are consistent, as shown in f-Figure 2 of the manuscript and described L279: "Puyuhuapi Fjord's head has a temperature that falls to 8°C in winter and rises to 18°C in summer, while in its central region it varies between 9°C and 16°C, coherent with Schneider et al. (2014), Pinilla et al. (2020) and Pérez-Santos et al. (2021)'s results."
- Reloncaví Sound: Similar seasonal variability (between about 10 in winter and 17°C in summer) is reported by Pérez-Santos et al. (2021), (see their Figure 2) aligning with our results L278 "Reloncaví Sound varies from 10°C in winter to 17°C in summer"

- Comau Fjord SST anomaly : The SST anomaly we present (Figure A3, that will not be included in the manuscript but which is present at the end of the present document) closely matches the patterns shown in Figure 3C of Mardones et al. (2023), both in shape and magnitude.
- General spatial variability pattern: Figure 2 of Saldías et al. (2021) illustrates that standard deviation in SST is much higher in enclosed fjords than in more exposed basins such as Corcovado Gulf. This patterns is similar to what we presented in Figure 2 of our manuscript, confirming that our interpolation reflects known physical processes, and not data artifacts.

Section **3.8 discusses several case studies of extreme events and includes potentially interesting insights into the evolution of atmospheric drivers**. However, these claims are **unsupported by key evidence, such as a surface heat budget or time series of heat fluxes**, **wind speed**, and **sea level pressure**. Moreover, the **so-called "high-resolution" results appear to be based on ¼° reanalysis data, which is unlikely to resolve the processes of interest in such a complex coastal setting**.

unsupported by key evidence, such as a surface heat budget or time series of heat fluxes, **wind speed**, and **sea level pressure**

- We do mention the heat fluxes, winds and sea level pressure (e.g. L474, L478-482) but did not include detailed values in order to not overload the text with secondary data. We propose to put a figure in the supplementary file.

so-called "high-resolution" results appear to be based on ¼° reanalysis data, which is unlikely to resolve the processes of interest in such a complex coastal setting

- We agree that the term “high resolution” can be subjective and potentially misleading. In our case, this term was intended to refer to the resolution of the in-situ based climatology (900 m resolution) and the MHW detection framework, not the atmospheric reanalysis data. In the abstract and the conclusion, we refer to it as “very high resolution”, and suggest to change it for “high resolution”.
- The heat fluxes and atmospheric data from ERA5 are used here for a general context and not for local processes. This will be clarified in section the manuscript.

Table A1: Total number of samples collected at the sea surface (0-1m depth), separated into mooring and non-mooring sources, along with the 5% of each group set aside for validation.

Total number of samples, including those collected from mooring instruments and from non-mooring instruments (all between 0 and 1m depth) and the associated number of samples kept apart for validation (5% of the total number of samplings).

Month	Total samples	Mooring samples	Mooring validation	Non-mooring samples	Non-mooring validation
Jan	8792	8693	435	99	5
Fev	14680	13685	684	995	50
Mar	13921	13711	686	210	11
Apr	8534	8373	419	161	8
May	10081	9478	474	603	30
Jun	9409	9077	454	332	17
Jul	9825	9331	467	494	25
Aug	9022	8557	428	465	23
Sep	10289	9540	477	749	37
Oct	9659	9174	459	485	24
Nov	8712	7701	385	1011	51
Dec	9238	9020	451	218	11

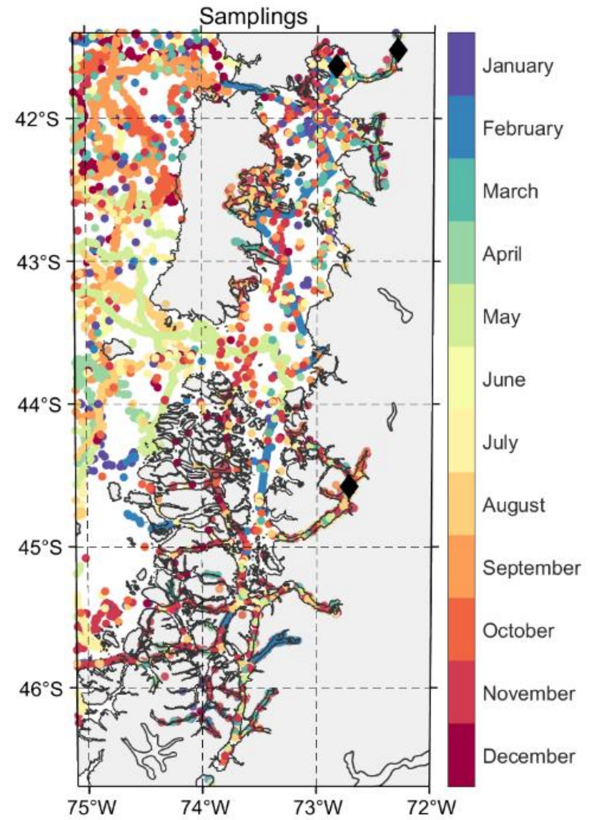

Figure A1: Spatial distribution of the in situ samples (0-400m depth), and the month during which it has been collected. The three black rhombus in Reloncaví Sound, Reloncaví Fjord and Puyuhuapi Fjord represent the localisation of the three moorings.

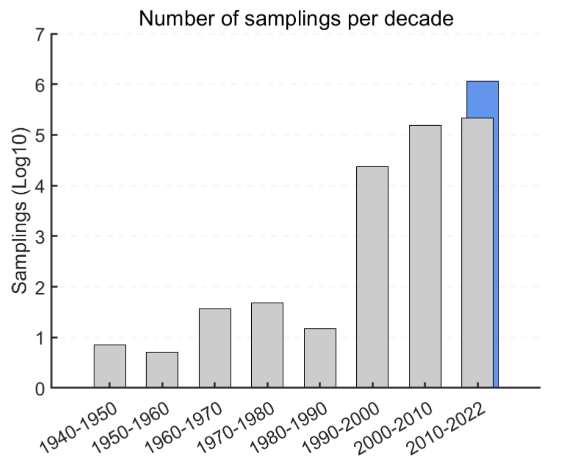

Figure A2: Temporal coverage per decade of the in situ samples in Log10. Grey: non-mooring observations, Blue: mooring observations.

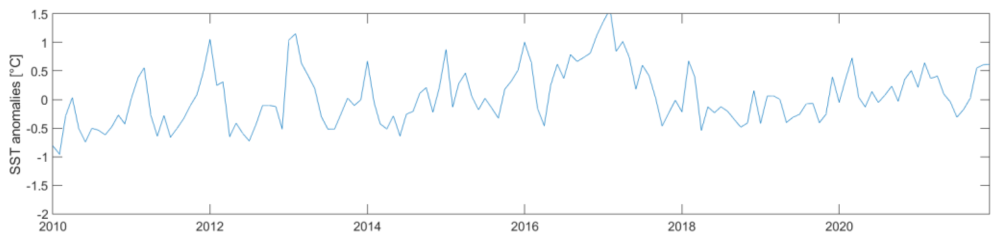

Figure A3: SST anomalies at Comau Fjord mouth