Reviewer 1

Review of Mawren et al. “Gliding through marine heatwaves: Subsurface biogeochemical characteristics on the Australian continental shelf”

Marine heatwaves (MHWs) are known to influence oxygen and biological productivity in the water column, but the majority of studies investigating them focus solely on the surface. This study addresses this important gap by analysing 16 years of glider observations on the Australian continental shelf across four contrasting coastal regions.

The authors show that summer surface marine heatwaves had a shallower mixed layer depth and enhanced stratification, confining warming to the upper ocean, whereas in other seasons there is deeper penetration under weakly stratified conditions. The study discusses the impact this has on biogeochemical variables (chlorophyll fluorescence and dissolved oxygen) across contrasting regions and also seasonally. By incorporating rigorous analysis and investigation of the vertical profiles, this study achieves the aim of showing the impact of MHWs goes beyond the surface. The authors have done a good job of summarising a large, multi-regional dataset and presenting a coherent analysis across regions and variables.

Overall, this study fits well within the scope of *Ocean Science* and is well aligned with the glider special issue theme and I recommend it for publication after the authors address specific comments below:

Response: We thank the reviewer for the positive comments and detailed review with great insights.

General comments:

In situ does not need to be italicised. (See English Guidelines in Submission on the Ocean Science webpage).

Response: These have now been changed throughout the manuscript.

There is some inconsistency in spelling conventions, with a mix of Oxford and British spelling (e.g., categorized vs characterised). Either convention is acceptable, but consistency throughout the manuscript is required. (See English Guidelines in Submission on the Ocean Science webpage).

Response: Thank you for pointing this out. We have carefully proofread the manuscript to ensure consistency.

Please ensure consistency in the use of abbreviations: define them on the first use only, and use the abbreviated form thereafter. Whilst I have highlighted a few in my specific comments, it may be worth checking through them all.

Response: Thank you, we revised the abbreviations.

Within the Results and Discussion, the manuscript refers to different areas within the defined regions such as the 'continental shelf' and 'inner shelf'. For clarity, it would be helpful to explicitly define these terms (e.g. by approximate isobaths) or annotate them in Fig. 1, particularly for readers unfamiliar with the regional shelf structure.

Response: We have reviewed our terminology for defined regions. While Figure 1 includes isobaths and glider profiles, the continental shelf is narrow in areas such that additional isobaths would be masked by the glider profiles. We have rephrased Figure 1 caption to clarify that glider profiles were over the continental shelf and shallower than 200 m.

I have a query regarding the calculated of the MHW severity. It is done using the surface based satellite data, but as the authors have pointed out, MHWs can show stronger signals in the subsurface compared to the surface. Would it be possible to justify the use of the SST data only (whether this is quality or sample size) and discuss the limitations.

Response: We thank the reviewer for this comment. We emphasize that the objective of this study is not to identify subsurface MHWs, but rather to examine the vertical structure associated with surface MHWs. We used satellite SST as it provides sufficient daily temporal resolution and long time series required to compute a climatology and percentile-based thresholds (which is impossible with gliders), as well as a high spatial resolution over the continental shelf. We acknowledge the limitations which are discussed in the manuscript at lines 778-780.

The use of surface MHW detection is important in this context to ensure consistency with existing MHW studies and to allow intercomparison across regions and datasets, particularly given our a focus on the biogeochemical response to surface MHWs on the Australian continental shelf.

Specific comments:

Introduction

Lines 63 – 65: Consider explicitly linking MHW characteristics to stratification here. This relationship recurs throughout the manuscript and may benefit from being introduced earlier.

Response: We thank the reviewer for this suggestion. In the Introduction at lines 65-71, we have included several sentences to introduce the role of stratification in controlling MHW characteristics with supporting references.

Line 71: The emphasis on coastal areas could be clarified. I appreciate the next paragraph addresses this, but as one of the examples of subsurface MHWs is in the North Pacific, the coastal comment here lacks context and the importance of coastal regions could be emphasised here.

Response: We agree with the reviewer and have added the importance of coastal regions by highlighting how strong stratification, complex circulation and shallow bathymetry can amplify subsurface temperature anomalies. Please see lines 75-79 in the manuscript.

Lines 85-87: This is where it is important to understand the link between MHWs and stratification.

Response: We have strengthened this section by linking MHW to stratification through mixed layer dynamics, vertical mixing and thermocline variability in modifying heat storage at lines 69-71.

Line 111: Please indicate the temporal limits of glider coverage and comment on how this sampling frequency relates to the duration of extreme events, particularly MHWs.

Response: Unlike a fixed mooring which provides a continuous time series at a single location, gliders offer flexible deployments: seagliders can operate for several months at a time while slocum gliders typically conduct missions lasting 3-5 weeks (see line 185). Consequently, glider sampling is intermittent and extreme events such as MHWs may not always be fully captured. Due to this limitation, we do not attempt to quantify the full duration of MHWs. Instead the analysis focuses on the localised impact of these events on the vertical physical and biogeochemical structure, which can be robustly assessed using these glider profiles.

Lines 113 – 114: Sensors on gliders do not directly measure stratification or phytoplankton, it may be better to list the variables they do measure (e.g. temperature, salinity, oxygen, chlorophyll fluorescence) and then indicate that from these stratification and phytoplankton dynamics can be inferred.

Response: We have retained the sentence stating the variables measured by gliders at lines 119-121 and lines 123-124 stating that they can be used to infer stratification and phytoplankton dynamics.

Lines 117 – 120: Were any of the cited studies specifically focused on MHWs? If so, please clarify. Additionally, please define what is meant by “short-term” (e.g. hours, days).

Response: To our knowledge, only a limited number of studies have explicitly examined MHWs using glider observations around Australia. We have cited the relevant study by Benthuysen et al. (2018) who captured a specific MHW off northeastern Australia and have removed short-term as the measurements from IMOS are continual monitoring aimed at capturing mean conditions over specific regions. Please see lines 131-134.

Methods

Line 151: Extra space between “is” and “used”.

Response: Corrected in the revised manuscript.

Line 163 – 164: IMOS has already been defined on line 123

Response: This has been modified accordingly.

Line 194 – 197: What was the quenching depth? Is it worth retaining daytime measurements from this depth and below to limit the removal of valuable data? Particularly given several hypotheses rely on chlorophyll interpretation.

Response: We thank the reviewer for this comment. A fixed quenching depth was not defined as quenching varies with different factors such as light conditions, stratification, and mixed layer depth particularly over shallow continental shelf regions. Given our focus being on bathymetry ~ 40-100m depth depending on the region, we used a more conservative approach and retained only night-time chlorophyll observations, following Schaeffer et al., (2016b), rather than applying an uncertain depth based on daytime data.

Line 202: Consider using “few” instead of “rare”.

Response: Corrected accordingly.

Line 217: There appears to be a typesetting issue here and the link doesn’t work.

Response: We have included a reference for the Best Practice Manual at line 203.

Line 229 – 230: Is excluding negative seasonal anomalies standard practise when identifying MHWs? Please clarify whether this refers to identifying events or calculating the severity index.

Response: The MHW identification followed the standard methodology by Sen Gupta et al. (2020), based on severity >1, without excluding negative anomalies. Previously, we only retained temperature anomalies above the seasonal mean to highlight the warm anomalies linked to MHW conditions when calculating the averaged profiles in Figures 6-9. Although results were maintained, we recomputed the averaged profiles using all available MHW profiles (Figs. 6-9), irrespective of whether some profiles fell below the seasonal mean to keep methodological consistency.

Line 238: Add “defined” before “as”

Response: Corrected accordingly.

Line 328: Potentially an extra space between “sampled” and “off”.

Response: Corrected in the revised manuscript.

Lines 334 – 341: Where possible, figure captions should appear on the same page as the figure.

Response: Corrected in the revised text.

Results

Lines 369 – 370: How do you define the continental shelf and offshore waters? Also, clarify the area spoken about here is in the QLD region (if I have understood correctly).

Response: We define the continental shelf as waters shallower than 200m, following the United Nations Convention on the Law of the Sea (UNCLOS). Offshore waters are thus deeper than 200m. The 200m continental shelf boundary is now mentioned in the updated Figure 1 caption.

Lines 370 – 372: Not sure what is meant here by “In agreement with the higher frequency”, please clarify the link between frequency and duration.

Response: We have removed “In agreement with the higher frequency”.

Line 372: GBR has not yet been defined.

Response: Thank you for pointing this out. A thorough check has been performed to make sure all abbreviations are defined at first use.

Lines 379 – 381: It could be helpful to briefly clarify what types of quality control criteria led to the exclusion of these profiles (e.g. sensor issues, incomplete dives), and whether this could bias the representation of the most extreme conditions.

Response: We thank the reviewer for this comment. The gliders shown in Figure 4 represent the full duration of each deployment after quality control. Some individual profiles were excluded due to the criteria outlined in section 2.4, including data outside the depth range, or daytime measurements etc. These criteria are methodological and were applied consistently across all deployments, rather than selectively during MHWs. While this may result in the exclusion of some profiles coinciding with the most extreme conditions, the number of removed profiles is relatively small to the total dataset.

Lines 389 – 391: Could more frequent summer and autumn events reflect increased glider sampling during these seasons?

Response: The IMOS Ocean Glider program includes routine deployments which are spread over the seasons, as well as targeted MHW deployments, which tends to be biased towards summer/autumn when the ocean temperatures are high. We have rephrased this sentence at lines 402-408 to avoid confusion.

Figure 4: Can the axis labels and colourbar labels be larger for readability?

Response: Corrected accordingly.

Line 410: DOX has already been defined on line 270.

Response: Corrected accordingly.

Line 415: For clarity, refer to the MLD rather than “both layers”.

Response: Modified accordingly.

Lines 434 – 435: Could you elaborate on how the regional and seasonal regimes do affect the multi-modal structure?

Response: Thank you for this comment. Because the sampling is insufficient to robustly resolve the full region, season, and severity interaction, we examined mean dissolved oxygen profiles separately by season and by region to assess whether particular regimes may be shaping the aggregated multimodal structure. Please see the new Supplementary Figure S2. The results suggest that both seasonal and regional variability likely plays a role, as spring and summer show higher dissolved oxygen compared to autumn, particularly in TAS and NSW regions, while QLD contributes more strongly to intermediate and lower DO ranges. We added some explanation at lines 451-454 in the updated manuscript.

Line 476: Missing comma after Holbrook and Bindoff.

Response: Corrected accordingly.

Line 501: Should this refer to oxygen saturation rather than DOX?

Response: Modified accordingly at lines 523-528.

Lines 504 – 510: Caption placement should ideally be on the same page as the figure.

Response: Modified accordingly.

Line 525: Consider using “DCM” here for consistency.

Response: Corrected in the revised text.

Line 546: GBR should be defined here rather than later on line 648.

Response: Modified accordingly.

Lines 566 – 569: Could this pattern also reflect autumn stratification breakdown and nutrient entrainment from deeper waters?

Response: We thank the reviewer for this suggestion. We agree that the elevated chlorophyll concentrations observed during autumn may also reflect seasonal stratification weakening and associated nutrient entrainment from deeper waters. We have revised the manuscript at lines 594-598 to acknowledge this.

Lines 576 – 578: Could nutrient limitation toward the end of the bloom period also contribute?

Response: We agree that with the observed seasonal evolution of stratification and chlorophyll pattern, nutrient limitation toward the end of the bloom period could contribute to lower DO observed in autumn. We added additional explanations at lines 606-609.

Line 604: Is oxygen saturation intended rather than DOX?

Response: We thank the reviewer for pointing this out. We clarified this at lines 637-639 that the oxygen saturation is above 100% throughout the water column and higher than non-MHW periods.

Discussion

Since results consistency with hypotheses (2) and (4) is discussed, it may be useful to explicitly link findings back to hypotheses (1) and (3) as well, either here or in the conclusions.

Response: Thanks for pointing this out. We have revised the Discussion section throughout to explicitly restate each hypothesis when referring to it.

Line 668: Do you have a suggestion for why it is different for NSW?

Response: Thank you for this comment. The hydrography of the EAC dominates the NSW shelf year-round. Unlike other regions, stratification in NSW is not driven purely by seasonal heating/cooling but rather modulated by shelf encroachment of the EAC, mesoscale eddies and current instabilities. As shown in Schaeffer et al, (2015), the seasonality of phytoplankton in NSW reflects the persistent influence of the EAC, with summer-spring biomass maxima and reduced winter abundance. We added this explanation at lines 683-688.

Line 716 – Earlier referred to as the Eastern Australian current on line 473 – be consistent with East or Eastern

Response: Modified according to East Australian Current in the manuscript.

Supplementary

Line 6: Add a space after “Panels”.

Response: Modified accordingly.

Fig S6: Also doesn’t have the colour coding Fig. 10 has; consider mentioning this in the caption.

Response: Thank you for the suggestion. We would like to clarify that the colors in Fig. 10 refer to the Pearson correlation coefficients as shown in the colorbar, whereas the values in Fig S8 refer to the number of data points. We clarified it in the caption.

Reviewer 2

Review of "Gliding through marine heatwaves: Subsurface biogeochemical characteristics on the Australian continental shelf" by Mawren et al.

General comments

This study aims to characterize the subsurface structure of chlorophyll-a and dissolved oxygen during marine heatwaves by analyzing the 16-year glider observations from four Australian continental shelf regions. Results are presented in two parts: 1) the compilation of all profiles but separating them into non-MHW and MHW cases and also above and below MLD (Section 3.2); and 2) the comparison between non-MHW and MHW cases for each region and season (Section 3.3). Key findings are that: 1) the responses to MHWs can be generally characterized by shallower MLD in summer and deeper MLD in winter, and deeper and stronger subsurface chlorophyll maxima; 2) dissolved oxygen above and below MLD had a bimodal distribution when all profiles were compiled together; and 3) regional responses may be related to local environmental setting.

First of all, it is great to see an effort of compiling 16-year glider observations equipped with biogeochemical sensors. Glider observations are a valuable tool especially for understanding fine-scale and subsurface coastal processes. While I appreciate this great effort, I find several issues in the manuscript mainly due to the insufficient number of sampling. There are 16 years in total, but there are four regions and four seasons to consider because of strong regional and seasonal variability. Hence, for some regions and seasons, the data are limited to robustly characterize MHW impacts (Fig. 4). My detailed feedback is provided below, followed by specific comments and technical corrections for improvements.

Response: We are grateful to the reviewer for the thorough review and helpful suggestions.

Limitation of the study. I think it is important to acknowledge the limitation of the study due to insufficient sampling. Even though 16 years of glider observations are wonderful, there are regions, depths, and seasons in which samples are insufficient in number to consider that the findings are robust. These regions, depths, and seasons are clearly shown in Figure 2. For example, the subsurface properties during spring MHWs are likely not robustly characterized for TAS (Fig. 6 second column) because these composite profiles are based on two MHW events only (Fig. 2b). There could be a paragraph in the Discussion section to acknowledge this

limitation and identify regions and seasons that require further sampling for enabling more robust analysis in the future.

Response: We thank the reviewer for this suggestion. The Discussion has been revised at lines 769-784 to explicitly acknowledge limitations related to insufficient and uneven sampling, with reference to specific regions and seasons in the study.

Seasonal climatology. For the same reason above, referring to the seasonal composite averages as “seasonal mean” can be misleading in Section 3.3. I suggest deleting black lines which the authors refer to as seasonal means, which are the averages of all profiles (non-MHW and MHW profiles). These would be representative of a true seasonal climatology if there is a sufficient number of samples. Considering the MHW definition, there should be at least 90 profiles of non-MHW cases for every 10 MHW profiles. However, there are regions like QLD in which nearly half of the profiles are of MHW cases (Fig. 2e). In this case, seasonal averaging will give a seasonal climatology that is biased towards MHW cases. This is clearly reflected on Fig. 8a, where black is closer to red than to blue. I suggest only presenting and comparing non-MHW (blue) vs. MHW (red) composite means for Figs. 6-9).

Response: Thank you for this constructive suggestion. We agree that referring to the seasonal composite (computed from all profiles) as a “seasonal mean” may be misleading, especially in regions or seasons where the proportion of MHWs is higher than non-MHWs. To assess the magnitude of this potential bias, we performed an additional evaluation using collocated satellite-glider SST (please see figure below). Specifically, we extracted SST from satellite at the nearest position and time of each glider profile and compared seasonal composites using three approaches:

- All Profiles
- Non-MHW profiles only
- Random subsampling (90% non-MHWs, 10% MHWs)

The comparison shows that satellite and glider seasonal composites agree very closely when evaluated along the glider tracks and over the glider time period (see figure below). More importantly, differences among the three sampling approaches are small, indicating that the seasonal structure is robust and not strongly sensitive to the proportion of MHW profiles included.

We acknowledge the reviewer’s concern regarding terminology and have therefore removed the black “seasonal mean” lines from Figures 6-8. We now use non-MHW profiles to define the seasonal composites for anomaly calculations in Figure 10, Supplementary Figures 4-7.

SST glider-satellite comparison by region and season

		Composite mean (all profiles)			Composite mean (non-MHW profiles)			Composite mean (90% non-mhws, 10% mhws)			Seasonal mean (1985-2026)
		A			B			C			D
Region	Season	Glider	Satellite	Diff	Glider	Satellite	Diff	Glider	Satellite	Diff	Satellite clim
TAS	DJF	17.37	17.49	**-0.12**	16.96	17.05	**-0.09**	17.02	17.11	**-0.09**	16.38
	MAM	16.88	16.82	**0.06**	16.63	16.57	**0.06**	16.66	16.60	**0.06**	16.12
	JJA	12.37	12.21	**0.16**	12.35	12.19	**0.16**	12.36	12.19	**0.17**	13.22
	SON	12.85	12.73	**0.12**	12.81	12.67	**0.14**	12.81	12.68	**0.13**	13.13
NSW	DJF	22.49	23.30	**-0.81**	22.29	23.08	**-0.79**	22.32	23.11	**-0.79**	23.45
	MAM	23.02	23.36	**-0.34**	22.73	23.08	**-0.35**	22.77	23.12	**-0.35**	22.86
	JJA	20.27	20.48	**-0.21**	20.27	20.48	**-0.21**	20.29	20.48	**-0.19**	19.38
	SON	20.68	21.08	**-0.4**	20.54	20.94	**-0.4**	20.55	20.96	**-0.41**	19.87
QLD	DJF	29.35	29.27	**0.08**	29.02	28.93	**0.09**	29.07	28.99	**0.08**	28.68
	MAM	27.94	27.86	**0.08**	27.46	27.43	**0.03**	27.60	27.56	**0.04**	27.60
	JJA	24.66	24.67	**-0.01**	24.02	24.07	**-0.05**	24.13	24.18	**-0.05**	24.95
	SON	26.63	26.57	**0.06**	26.51	26.44	**0.07**	26.53	26.46	**0.07**	26.06
SW WA	DJF	21.67	22.02	**-0.35**	21.49	21.86	**-0.37**	21.52	21.88	**-0.36**	21.43
	MAM	22.24	22.11	**0.13**	22.09	21.97	**0.12**	22.10	21.98	**0.12**	22.24
	JJA	19.68	19.75	**-0.07**	19.63	19.72	**-0.09**	19.63	19.72	**-0.09**	19.92
	SON	19.60	19.49	**0.11**	19.56	19.45	**0.11**	19.57	19.45	**0.12**	19.08

A: Seasonal SST composites from glider profiles (all profiles), co-located with nearest satellite pixel in space and time

B: Seasonal SST composites from glider profiles (all non-MHW profiles), co-located with nearest satellite pixel in space and time

C: Balanced seasonal composite (90% MHW, 10% MHW) using repeated equal-sample resampling of MHW/non-MHW profiles (~100 iterations)

D: Seasonal mean satellite SST averaged over all glider locations and the full glider study period (2009-2025)

The MHW depth extent. The definition of the MHW depth extent is too broad that it would include any warm/positive anomaly (L238). In the surface, MHWs are defined as above 90th percentile. Should the subsurface MHWs be defined similarly? How did previous studies on subsurface MHWs define? Once again, it would be difficult to define the percentile threshold for some regions/seasons due to sampling biased towards MHWs. I think it is difficult to define the MHW depth extent with the data available.

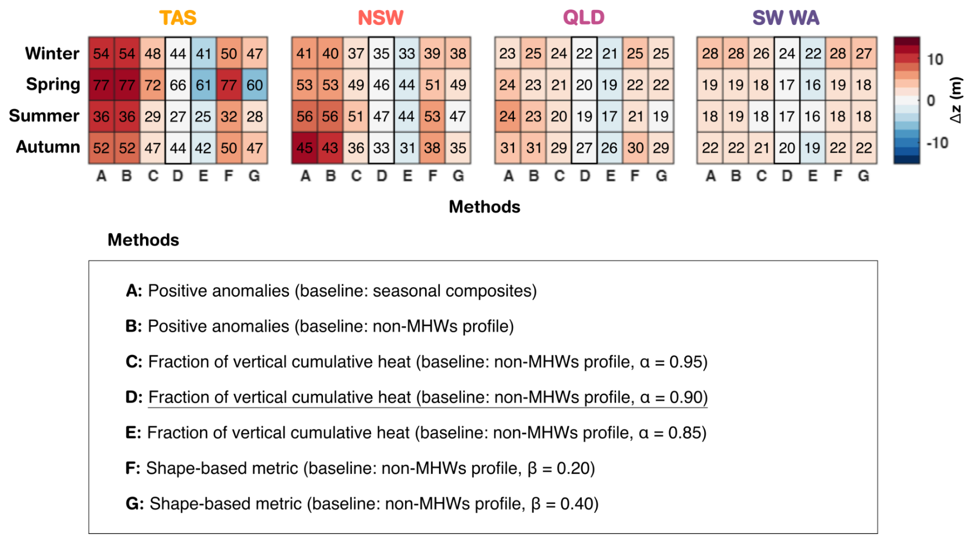
Response: This study does not aim to identify subsurface MHWs or apply percentile-based thresholds at depth, due to the lack of consistent temporal coverage in the glider dataset.

MHWs here are only identified at the surface using satellite SST following the definition from Hobday et al. (2016). The glider observations are then used to examine the subsurface temperature structure and biogeochemical responses associated with these surface-detected MHWs. In other words, the calculation of MHW depth extent is applied exclusively for glider profiles categorized as MHW profiles.

To address the reviewer's concern regarding the MHW depth extent definition, we revised our original approach. Instead of using seasonal composites as a reference, we now define temperature anomalies relative to the mean non-MHW profile. This provides a physically consistent background state and avoids potential bias introduced by uneven sampling of MHW and non-MHW profiles.

We evaluated the sensitivity of the MHW depth extent to both the choice of baseline and the calculation method. The figure below (also included in the Supp. Materials, Fig. S7, L88) shows the mean MHW depth extent for each region and season across all methods (absolute values in each cell), while colours indicate the deviation relative to Method D (non-MHW baseline with α=0.90), which we adopt as our reference following Elzahaby and Schaeffer (2019). Method A corresponds to the previous seasonal-composite baseline, and Method B uses the non-MHW baseline with a zero-crossing criterion. We also included the shape-based metric (Methods F and G), defined as the deepest depth where the positive temperature anomaly exceeds a fraction β of the maximum anomaly (ΔT_{max}) in the profile. These provide an independent estimate based on the vertical structure of the anomaly, providing values consistent with the other methods and, therefore, extra support for our chosen reference method, i.e. Method D.

Overall, differences between methods are relatively small, varying between -5 m and 12 m but mostly around ±2 m depth, which is minor compared to the absolute depth extent. This suggests that our main conclusions are not sensitive to the exact definition used. Greater differences are mostly tied to the TAS and NSW regions, which could be related to subsurface-intensified warm anomalies (e.g., eddy cores or intrusions along the thermocline), which affect presence-based metrics more than cumulative-heat approaches. In QLD, despite the imbalance between MHW and non-MHW profiles, differences between Methods A and B are minimal, suggesting that the MHW depth extent calculation is not strongly sensitive to baseline selection.



Section 3.2. This section is based on the compilation of profiles from all regions and seasons. The findings from this section take up a large part of the abstract (L35-41). However, I am not sure how robust these findings are given the strong regional and seasonal variability. From this analysis, the authors report a bimodal distribution of dissolved oxygen but without an explanation for what this represents. Is this possibly a seasonal cycle (high DO in winter and low DO in summer) or a pattern dominated by two contrasting regions? Similarly, there was no explanation for the difference between the impacts of moderate vs. strong MHWs provided, although their difference was reported (L37). I suggest more careful analysis to come up with a conclusive explanation for the bimodal distribution as well as the MHW severity dependence on the impact on the bimodal distribution. If the seasonal cycle or regional variability was the driver, I suggest focusing on the findings from Section 3.3.

Response: We appreciate the comment and suggestion. The abstract was rewritten accordingly. Following the reviewer's comment on the bimodal DO distribution in Section 3.2, we explored whether this pattern could be primarily driven by seasonal or regional variability. Given that we don't have sufficient sampling to robustly separate region × season × severity, we looked at mean DO profiles by season and by region separately (see figures below and added Fig. S2). The depth in the x-axis was rescaled, analogous to Malan et al. (2025), since the averaged MLD could be misleading due to large seasonal and regional differences. The analysis helps to qualitatively assess whether specific seasons or particular regions stand out and could be dominating the aggregated distributions. The results suggest that both seasonal and regional variability likely plays a role, as spring and summer show higher DO compared to autumn as well as the QLD region - which also presents the highest number of sampling profiles - also stands out from the other regions. Please see lines 451-454 in the manuscript. This analysis is not meant as a definitive attribution, but rather as supporting evidence to help interpret the bimodal distributions.

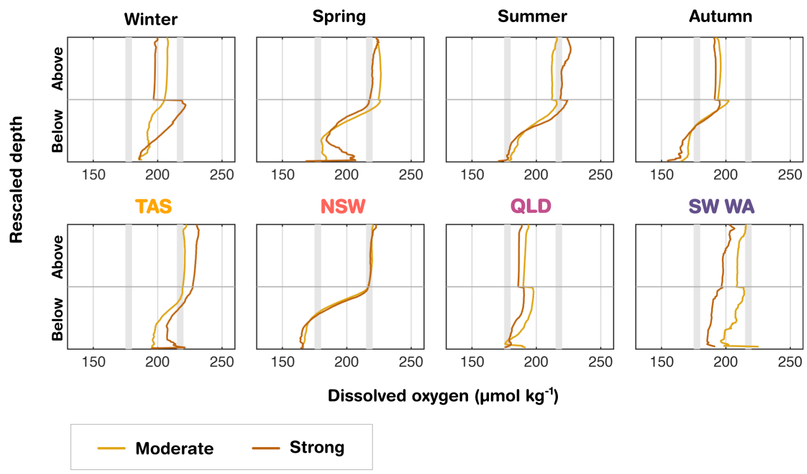


Figure 10. This figure needs to be included in the results section instead of the discussion section, because it is a result based on the analysis of this study. Likewise, this result needs to be described in the results section first instead of L673 onwards. In this figure, anomalies are used for temperature, chlorophyll-a, and dissolved oxygen. I guess they are anomalies relative to the seasonal means, but for the reason mentioned above (## Seasonal climatology), I suggest not using such anomalies.

Response: As noted in our responses above (see comments on seasonal climatology and MHW depth extent), we have revised the methodology used in Figure 10. Specifically, temperature, chlorophyll and dissolved oxygen anomalies are now calculated related to seasonal composites derived from non-MHW profiles only, rather than all the profiles. We have also moved Figure 10 from the Discussion to results section as suggested.

Consistency for the presentation order. It was difficult to follow the manuscript when the presentation order among the four regions was inconsistent throughout the manuscript. For example, Fig. 2 (NSW, TAS, QLD, SW WA in panel i vs. TAS, NSW, QLD, SW WA in the earlier panels), Fig. 3 (TAS, NSW, QLD, SW WA), Fig. 4 (TAS, SW WA, NSW, QLD), Section 3.3 (TAS, NSW, QLD, SW WA), Fig 10 (NSW, TAS, QLD, SW WA). Please carefully reorganize the presentation order for all figures and tables as well as the paragraphs.

Response: Thank you for highlighting this issue. In the revised manuscript, we have carefully reorganized the order of regions (TAS, NSW, QLD, SW WA) throughout the text, tables, figures and captions.

Specific comments and technical corrections

L34: Two issues with this sentence. (1) Instead of “Stratification favoured …” please be more specific that this is referring to specific conditions (i.e., Enhanced stratification during summer MHWs”). (2) It may not be intuitive to follow that enhanced stratification (with shallower MLD; L32) favoured deeper DCM. If you think of light limitation, shallower MLD should lead to photosynthesis at shallower depth. Please consider adding a few words to briefly explain the possible mechanism so that readers will not get stuck with this sentence.

Response: We have revised the sentence and have added a brief clarification of the mechanism in the revised manuscript at lines 39-43.

L54: Is it correct to claim that marine heatwaves shoal MLD and alter subsurface BGC? Are MHWs the driver or are they the consequence of some other drivers?

Response: Thank you for this comment. We have rephrased our sentence at lines 54-55, and we describe the association rather than a causal mechanism.

L63: Please consider removing "long-lasting" as MHWs can be as short as 5 days according to the Hobday definition used in this study.

Response: Modified accordingly at lines 63-65.

L70-71: Please consider moving this topic statement to the beginning of the next paragraph (L73).

Response: Modified accordingly at line 81.

L84: Please considering replacing "latitude" with "light" or similar, as latitude is not a direct factor for phytoplankton growth.

Response: Modified accordingly at line 96.

L192-193: It was unclear at first how the outliers were identified. After looking at Fig S1, it became clear to me (if I understand correctly) that the outliers were identified (1) for each glider mission and that (2) the sliding window of 1000 points moves along the glider sampling (so it goes ups and downs). Please re-write the sentence so that these two points will be clear to readers who are not familiar with glider observations/methods.

Response: Modified accordingly in section 2.4.

L237: The definition of the MHW depth extent is too broad. Anything above the seasonal mean implies above 50th percentile, whereas the surface MHW is defined as above 90th percentile. Should the subsurface MHW follow the same percentile criterion as the surface MHW?

Response: As previously discussed on the #Seasonal climatology and #MHW depth extent comments, we modified our methodology on the MHW depth extent definition, following Elzahaby and Schaeffer (2019). We used the baseline of mean non-MHWs profile for the corresponding region and season. For each MHW profile, positive temperature anomalies are vertically integrated, and the MHW depth extent is taken as the depth containing 90% of the cumulative anomaly.

Elzahaby, Y., and Schaeffer, A.: Observational insight into the subsurface anomalies of marine heatwaves, Frontiers in Marine Science, 6, doi:10.3389/fmars.2019.00745, 2019.

L270-271: DOX and CPHL are used as abbreviations for dissolved oxygen and chlorophyll, respectively. Are they commonly used abbreviations for these? Please consider DO and CHL as alternatives, which are more commonly used in the field.

Response: Modified accordingly throughout the manuscript.

L283: Should "0.005" be "0.05"? p-value of < 0.05 is a commonly used threshold for statistical significance.

Response: We thank the reviewer for catching this typographical error. The p-value used is 0.05 and we have corrected it in the manuscript at line 298.

L283: What is wrong with the well-mixed profiles? As long as it is anomalously warm (>90th percentile), I do not think there is an issue to consider that the entire profile is a MHW.

Response: Using only stratified profiles imply that we exclude homogeneous profiles which can inflate correlations and can give bias analyses based on the depth of the profile. By filtering these profiles, we ensure that the relationships we note reflect meaningful correlations rather than artefacts.

L290: Suggest deleting "in surface MHWs" because this section describes both MHW and non-MHW glider missions.

Response: Modified accordingly.

L405-406: The four selected lon/lat coordinates can be indicated either on Fig. 1 or 3.

Response: We have included the coordinates in Figure 3.

L428: Please provide a possible explanation for the two oxygen regimes. What are they?

Response: As we illustrated in the new Supplementary 2, both seasonal and regional variability likely played a role: spring and summer showed higher DO compared to autumn. Regions such as TAS/NSW also showed higher DO in the mixed layer compared to low-oxygen waters in QLD.

L445: Yes. If the results simply reflect the seasonal or regional differences, they should not be key findings in the abstract. Please carefully check.

Response: We thank the reviewer for this comment. We have modified the abstract accordingly.

L455: Fig. 4 shows that there are glider observations when the MHW severity index was above 3 (Category 3; e.g., Fig. 4d March 2017), but this figure only shows two categories. Why?

Response: There were only 11 profiles with Category 3 (Severe) and restricted to the TAS region. Therefore, there was not enough representative sampling to make conclusions out of the CHL and DO distributions.

L471: “eddy-rich” is mentioned in the section title, but is not mentioned at all in this section. Instead of much of the discussion is attributed to the influence of the EAC extension. I suggest either renaming the title or including the discussion on the influence of eddies.

Response: We thank the reviewer for the comment. We included previous literature on the association between EAC extension and eddy presence in the region, as well as their influence on the biogeochemical processes and MHW development (L493-497).

L500: I do not think that mixing would be the case here, because mixing would break the DCM and the subsurface oxygen maximum.

Response: We thank the reviewer for this insightful comment. We agree and we have modified this statement accordingly at line 522.

L531-533: A possible explanation for this vertical structure in DO is that: 1) reduced DO in the upper 20 m due to reduced solubility; and 2) increased DO below 50 m, as a result of increased respiration due to the increased sinking of organic matter from the intermediate depth.

Response: We appreciate this helpful comment and we have elaborated this explanation in the revised manuscript at lines 557-561.

L545: incomplete section title. Missing a word like “area”

Response: Modified accordingly.

L549-550: Suggest rephrasing. “The *seasonal* evolution … followed the *seasonal* cycle.”

Response: Modified accordingly.

L568: “following the summer bloom period”. Are you sure this is sequential? This is based on composite means, so it could be that the summer and autumn DCMs seen in this figure represent different MHW events. It is hard to believe that a DCM shoals from summer to autumn.

Response: We agree that our original wording implied a sequential evolution from summer to autumn which is not supported by our seasonal composite based analysis. We have revised the text at line 596-598 to remove the temporal progression between the two seasons.

L571: Could “contradiction” be a better word than “anomaly” for emphasis here?

Response: We changed text in line 601 from "... thermodynamic anomaly…" to "… counterintuitive thermodynamic behaviour…", we appreciate the suggestion.

L572-573: Chlorophyll-a during MHWs is lower in the upper 20 m, hence it does not support the argument here.

Response: Thank you for pointing this out. We have considerably revised this paragraph at lines 594-596.

L625: Salinity is missing in the list of properties?

Response: Modified accordingly.

L630: This and the subsequent three paragraphs may be deleted because they are just introductions rather than discussions.

Response: Modified accordingly.

L667-668: "deeper MLD in winter". I do not understand why this would be the case. Please add an explanation for this mechanism.

Response: Thanks for pointing this out. While one might expect winter MHWs to also shoal the MLD like summer MHWs, this is not observed in all regions. Winter MHWs are sometimes characterized by a deeper MLD compared to normal winter conditions because the water column is already weakly stratified and warming alone does not generate a strong density gradient to shoal the MLD. In addition, the anomalous processes associated with winter MHWs such wind-driven mixing and horizontal advection, can further deepen the MLD relative to typical winter conditions. In some regions, such as Western Australia, Leeuwin-Current-driven MHWs produce deep warming which can result in a deep MHW structure (Zhang et al., 2023), where warmer water penetrates to greater depths, leading to a deeper MLD. It is an interesting point and we have included that explanation in the revised manuscript at lines 678-683.

Zhang, Y., Du, Y., Feng, M., & Hobday, A. J. (2023). Vertical structures of marine heatwaves. *Nature Communications*, *14*(1), 6483.

L683: Please specify the seasons for clarity instead of "particular seasons".

Response: Modified accordingly.

L683: Here and elsewhere in this section, it would be good to restate what the hypothesis was, so that readers will not have to go back to the earlier section.

Response: We thank the reviewer for this helpful suggestion. We have revised the Discussion section throughout to explicitly restate each hypothesis when referring to it. This has improved clarity and readability.

L685: It is incorrect to say that light penetrates deeper. The penetration depth does not change, but due to the shoaling of MLD, phytoplankton have access to stronger light intensity.

Response: Thank you for pointing this out. We have rephrased the sentence at lines 709-711.

L686: “Hayashida and Strutton, 2020” should be “Hayashida et al. 2020” (missing Matear).

Response: Modified accordingly.

L710: “these mechanism” This is the first sentence of a paragraph. As such, it is unclear what “these” refers to.

Response: Modified accordingly.