Review of the manuscript: https://doi.org/10.5194/egusphere-2024-4079

Title: The global ocean mixed layer depth derived from an energy approach

Reviewer #1

We appreciate all your comments. After a careful revision that incorporated your comments and those of the other two reviewers, this version has greatly improved. We trust that you will find the revisions substantial and consider our study a valuable contribution to mixed-layer depth research. Below is a description of the major modifications made to the manuscript:

- The methodology (subsection 2.1) was extended by including details concerning the connection of the work done by buoyancy (WB) with the turbulent kinetic energy budget.
- We revised our definition of the mixed layer and how to calculate its depth (subsection 2.2. Defining the mixed layer). The mixed layer depth (MLD) was not calculated with a unique WB equipotential.
- Since we adjusted the MLD calculation, we rewrote the Results and Discussion sections entirely. The Conclusions and Abstract were also rewritten accordingly.
- The Supplement was rewritten to provide a detailed comparison of our MLD methodology with the common ones; we included global monthly climatologies of the MLD and the WB at the MLD.
- The dataset containing the monthly climatology of mixed layer depth and the derived variables calculated in this study is now publicly available at SEANOE (https://doi.org/10.17882/106181).

Responses to specific comments are shown below.

<u>Summary</u>

Comment 1: The MLD is not an objective quantity, it requires defining timescales and spatial scales since the ocean surface boundary layer is under consistent and variable external forcing. Furthermore, identifying a mixed layer invokes some qualitative analysis to quantify what is meant by "mixed". This ambiguity makes the present goal of deriving a "global ocean" mixed layer depth challenging, and I acknowledge the efforts the authors have devoted to this matter. I find the proposed WB method for estimating MLD is potentially interesting, but the paper presently relies heavily on ad-hoc and empirical arguments. This limitation makes it difficult to judge the value of the WB approach versus other approaches. The study would therefore benefit significantly if it can identify quantifiable metrics to bolster its claims.

Answer 1: We agree with you on the challenges in defining the MLD. In this version, we clarify and better describe the aspects of the mixed layer addressed in our study in *subsection 2.2 Defining the mixed layer (page 5, lines 124-129)*.

We really appreciate that you find our approach to estimating the MLD interesting. We acknowledge that in the prior version of the manuscript, we derived our results based on empirical and non-robust arguments. In *subsection 2.2*, we proposed a quantitative procedure to calculate the MLD to fix that (*page 5, lines 129-133 and pages 7-8, lines 156-184)*.

Due to the adjusted definition of the mixed layer, the new results are completely different from those of the prior version of the manuscript. We conducted a detailed analysis (qualitative and quantitative) of the resulting MLD climatology and the WB threshold that defines it. We also compared our methodology with others on a global scale, considering different metrics. We consider that we greatly improved the significance and interpretation of our results (*Results and Discussion sections).*

Comment 2: The desire for a MLD metric led to proposing PE anomaly in the Reichl et al. (2022) analysis of MLDs, as it quantifies the energetic distance of a column of sea water from being well mixed. This naturally leads to the PE anomaly as a possible basis for identifying the MLD, but it is not obvious that the same energetic distance should define the MLD in all regions or on all timescales. WB itself may provide another interesting metric for quantifying mixed layers, but it differs in important ways from the PE anomaly. WB provides an integrated measure of stratification of the column by considering the displacement of a particle from the mixed layer base. This yields some insight into the stratification, but less information than if you evaluated the buoyancy work associated with displacement of all particles within the mixed layer. It is also not obvious that a single WB quantity should define the MLD at all locations and on all timescales.

Answer 2: You are right; WB differs from the PE anomaly in that they describe different physical variables. However, we consider that WB is physically suited to calculate the MLD (*subsections 2.1 and 2.2*). Via the WB profile WB(z), we actually provide the work done by buoyancy to vertically displace any water parcel within the mixed layer to the top of the mixed layer. We used the WB value at the MLD as the energy threshold characterizing the mixed layer.

We agree that it is not obvious that a unique WB equipotential should determine the MLD globally year-round, which was obtained from Eq. 8 and represented in Fig. 2; please refer to the detailed analysis of the WB metric in *subsections 2.1 and 2.2*. Results of the WB threshold characterizing the MLD are shown in *Figs. 4 and 11*.

Major concerns

Comment 3: The WB method for estimating MLD has merit over a density threshold method since it invokes a dynamical quantity. However, I am unsure that this dynamical quantity truly addresses two main shortcomings of the threshold method. It would help if this paper better identified what it does and doesn't improve upon other methods.

- WB does not offer any physical guidance for choosing a threshold value, so the threshold remains ad-hoc. The paper argues that 20 J/m3 is a universally applicable value, but this was derived empirically based on arguments about its vertical gradient over limited regions of the ocean. There is no physical significance offered for the integrated buoyancy work of 20 J/m3, which would justify it from first principles for identifying the upper ocean mixed layer in all seasons and locations.
- The new method is still sensitive to the details of the choice of the threshold value and the reference depth. Interestingly though, some of this sensitivity is reduced since WB is an integrated quantity and therefore less responsive to noise in a profile.

Answer 3: Considering this and the previous comment, we conducted a detailed analysis of the WB metric and its physical significance (*subsections 2.1 and 2.2*). We also revised our definition of the mixed layer and proposed a quantitative procedure for calculating its depth, including details concerning the density inhomogeneity along the mixed layer, the WB threshold, and the reference depth (*subsection 2.2*). The *Discussion section* extends the analysis of our MLD methodology and its downsides.

Comment 4: I did not agree with the claim in the text: "There is a correspondence between our methodology and that of Reichl et al. (2022), suggesting that our energy-based methodology is consistent with the turbulence approach of the mixed layer formation". I do not find the correspondence between the WB and PE anomaly to be obvious, other than both using energy based criteria. PE anomaly is a column integrated energy value that quantifies the distance of the column from being perfectly homogenous, hence it has units of J/m2. WB only considers the energetics of a single parcel advected through the column, hence it has units of J/m3. One could construct different idealized profiles that give the same WB MLD yet have different PE anomalies, which was identified as a shortcoming of other MLD methods in Reichl et al. 2022. Perhaps the WB method has less PE anomaly sensitivity than other methods, but an evaluation of WB in terms of PE anomaly was not attempted in this paper. These important differences from the PE anomaly do not support the statement that this method connects to boundary layer turbulence in the same way as argued for PE anomaly.

Answer 4: We are very sorry for this inaccuracy. Now, in *subsection 2.1 An energy measure of the vertical homogeneity of the water column*, we show an alternative expression of WB that allows us to appreciate its connection with the turbulent kinetic energy budget and, thus, with the physics of boundary layer turbulence. We did not provide further analysis of such a connection because that is beyond the scope of this paper and is proposed for future research (*pages 4-5, lines 93-113)*.

Via the WB profile WB(*z*), we actually provide the work done by buoyancy to vertically displace any water parcel within the mixed layer to the top of the mixed layer. We used the WB value at the MLD as the energy threshold characterizing the mixed layer.

Comment 5: The article claims the WB MLD with an energy value of 20 J/m3 is "accurate, robust, and of global applicability". This claim is based on looking at "challenging regions", which are defined where a suite of existing methods disagree with each other in MLD magnitude. One issue is that this approach appears to weight the analysis to deep MLD regions (usually convective regions), but this appears to deemphasize shallow MLD regions (e.g., the Arctic Ocean (more on this below) and summer time mixed layers in biologically productive regions) that are just as valuable. Furthermore, I did not find a convincing and quantitative argument in section 3.2.2 for why the WB method estimates a better and/or more consistent MLD than other methods in these regions; it mostly relies on ad-hoc arguments.

Answer 5: We agree with you that local results can not support global conclusions. To circumvent this, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics (*subsection 3.2. MLD methodologies intercomparison*). We removed the limitation of focusing only on "challenging regions". We found that there is a lack of consistency among the methodologies in determining the MLD, but also that all the methodologies perform well under the oceanographic conditions for which they were built, according to the parameter being addressed. Since it is almost impossible to ensure that the true value of the MLD in any location and time is known, the accuracy of any methodology can not be determined, that is, the closeness of any MLD estimation to the true value. When comparing MLD methodologies, we can only evaluate their precision: the closeness between the different MLD estimates. Thus, we evaluated the precision of our methodology on a global scale and removed ad-hoc arguments from the description of our results.

Comment 6: I found the analysis of the climatology lacked a significant new/novel result that advances beyond previous work on mixed layer climatologies. It would be useful if issues in previous climatologies could be identified and to discuss why the WB method was able to provide a more meaningful estimate of a MLD climatology (or for what contexts it may be more meaningful).

Answer 6: Our study's significance relies on the energy definition of the mixed layer, which represents an important advancement in this topic (*subsections 2.1 and 2.2*). In the revised version of the manuscript, we carried out an intercomparison of MLD methodologies (*subsection 3.2. MLD methodologies intercomparison*) that showed to what extent our MLD estimation could be more meaningful than others; such an analysis was further extended in the *Discussion section*.

Comment 7: The article states that it offers globally applicable mixed layer depths and provides a mixed layer depth climatology. Constructing a gridded mixed layer depth climatology requires many decisions, and little technical detail on how this is done was provided. Nor is the resulting climatology data made available, which would severely limit any potential impacts of this work. The climatology is produced only using Argo data, which neglects a lot of additional oceanic profile data that exists. This also leads to coastal oceans and the entire Arctic Ocean being absent from this work.

Answer 7: In the revised version of the manuscript, we rewrote how we constructed the MLD climatology, the data we used, and its limitations in spatial coverage (*subsections 2.3 and 2.4, pages 8-9, lines 185-214*), hoping to be clearer. We acknowledge the limitations in the spatial coverage of Argo data; this study could not explore the MLD in coastal zones, and the robustness of the findings in the subpolar oceans may be limited. For future research, we propose incorporating additional observational datasets covering the regions not extensively mapped by Argo to expand the scope and robustness of this study.

The dataset containing the monthly climatology of mixed layer depth and the derived variables calculated in this study is now publicly available at SEANOE (https://doi.org/10.17882/106181). This information was added to the *Data availability section*.

Other Comments

Comment 8: What reference pressure is used to define the potential density in equation 3? Can a fixed reference pressure be used as a suitable approximation?

Answer 8: We are sorry for not providing that information. According to Stewart (2008), the potential density referred to as 0 bar is adequate for vertical displacements not exceeding a few thousand meters (*Page 3, lines 77-79*).

Comment 9: The WB method may offer some physical insights into when and why different MLD methods differ, and it is closely related to the threshold method such that it may better ground the threshold values used in present studies. One can rewrite the WB criteria (equation 3, by rewriting the 2nd terms RHS w/ H defined as the MLD and _Hmean defined as the depth integral mean of potential density over the MLD): WB = g*H*[_Hmean - rho_ref], which is now expressed similarly to the threshold method: delta = [rho(z) - rho_ref]. This makes two notable differences of WB from the threshold method more apparent: (1) the density difference is now from the average of the density over the column instead of at the MLD, and (2) the density difference is effectively weighted by the

MLD (and gravity). These have some interesting implications: (1) integrating the potential density makes the MLD less sensitive to noise in the profile, and (2) weighting the threshold by the MLD reduces the depth of deeper MLDs with the same energy value. The first effect seems advantageous for practical use. The second effect makes sense from the perspective that it takes more energy to raise a parcel from a deeper depth, but since deeper MLDs often experience more turbulence and mixing, it is not straightforward to me that the second effect is obviously "better" for MLD identification. The ocean does not experience uniform levels of energy inputs to turbulence. (This is a main reason I am not convinced there should be a single universal energy value for either the WB method or the PE anomaly method.) Furthermore, the 2nd point above about MLD weighting also offers an interpretation of why shallower depths are associated with larger density differences, e.g. Figure 7.

Answer 9: We agree with you that our methodology is a threshold method. The procedure to determine the MLD in terms of WB is based on a threshold of the density inhomogeneity along the mixed layer (*subsection 2.2. Defining the mixed layer*). By considering the vertical integration of density, our methodology could represent an improved or well-founded version of the threshold density criterion proposed by de Boyer Montégut *et al*. (2004) to define the MLD. One of our results showed that MLD methodologies based on density thresholds of about 0.125 kgm-3 along the mixed layer produce overestimated MLDs and are inadequate to define a well-mixed layer in energetic terms. Our methodology serves to better support the threshold values of existing MLD methodologies. However, its main feature is that it represents a unique and complete methodology to estimate the MLD.

We rewrote WB as you suggested and conducted a detailed analysis of WB from the resulting expression (*Eq. 8* in the new version of the manuscript). The features you mentioned were included and properly described (*pages 5-7, lines 135-155*). The second effect you mentioned reflects the relationship between the density inhomogeneity and the associated WB along the vertical, but does not negatively affect the MLD estimation. In fact, this relationship (*Eq. 8*) was used to propose the procedure to calculate the MLD.

We agree with you that a unique WB equipotential should not determine the MLD globally year-round. A detailed analysis of the WB metric (*subsection 2.2 Defining the mixed layer*) and our new results (*Figs. 4 and 11*) confirmed your observation. Turbulence and its associated energy levels are spatially and temporally variable on a global scale; thus, the water column's stratification and vertical homogenization are not spatially uniform through the seasons.

Comment 10: L336: "It is also easy to implement numerically." Easy compared to what?

Answer 10: We are sorry for not being clear in that sentence. The numerical implementation of our MLD methodology only requires the potential density profile referred to 0 dbar, which is easily obtained from simple survey ocean data or numerical data. The script to compute the MLD is very short, and its formulae are not complex. In that regard, our methodology is easy to implement numerically. We rewrote this paragraph, hoping to be clearer (*page 25, lines 435-437)*.

Review of the manuscript: https://doi.org/10.5194/egusphere-2024-4079

Title: The global ocean mixed layer depth derived from an energy approach

Reviewer #2

We appreciate all your comments. After a careful revision that incorporated your comments and those of the other two reviewers, this version has greatly improved. We trust that you will find the revisions substantial and consider our study a valuable contribution to mixed-layer depth research. Below is a description of the major modifications made to the manuscript:

- The methodology (subsection 2.1) was extended by including details concerning the connection of the work done by buoyancy (WB) with the turbulent kinetic energy budget.
- We revised our definition of the mixed layer and how to calculate its depth (subsection 2.2. Defining the mixed layer). The mixed layer depth (MLD) was not calculated with a unique WB equipotential.
- Since we adjusted the MLD calculation, we rewrote the Results and Discussion sections entirely. The Conclusions and Abstract were also rewritten accordingly.
- The Supplement was rewritten to provide a detailed comparison of our MLD methodology with the common ones; we included global monthly climatologies of the MLD and the WB at the MLD.
- The dataset containing the monthly climatology of mixed layer depth and the derived variables calculated in this study is now publicly available at SEANOE (https://doi.org/10.17882/106181).

Responses to specific comments are shown below.

<u>Summary</u>

Comment 1: Many methods have been proposed to identify the mixed-layer depth (MLD), which can be broadly classified into two categories, threshold method and energy method. This article is part of the overall effort to determine the MLD by using an energy based method (EBM). This choice is valuable because scalar threshold-based methods are subjective and not necessarily consistent with the physics of vertical mixing. This is not the case for the EMB method, because it is based on the work of the buoyancy force (WB) on the vertical.

Answer 1: We really appreciate that you find our approach to estimating the MLD significant.

<u>Major points</u>

Comment 2: Basics of this paper come from Reichl et al. (2022), who derive MLD from the concept of potential energy. However I have not clearly identified the links between your approach and the Reichl et al. (2022) one. Therefore I suggest to discuss the common points, differences, put into perspective, your approach with that of Reichl et al. (2022) in Section 2.1. It would be interesting to propose a mathematical development, which shows that the WB threshold is a potential energy barrier for water parcels. I do not see this correspondence in the article. In the same spirit, the stratification index (SI = R z 0 N 2zdz, where N 2 = −g ρ∂ρ∂z) could also have been used.

Answer 2: Thank you for the comment. Although WB and the PE anomaly of Reichl et al. (2022) share some similarities, our work is not based on Reichl's. In *subsection 2.1 An energy measure of the vertical homogeneity of the water column* (*pages 4-5, lines 93-113)*, we show an alternative expression of WB that allows us to appreciate its connection with the turbulent kinetic energy budget and, thus, with the physics of boundary layer turbulence. We did not provide further analysis of such a connection because that is beyond the scope of this paper and is proposed for future research.

WB represents an energy measure of the vertical homogeneity of the water column. In that respect, the WB values at different depths quantify the energy barriers for the vertical displacement of water parcels. Additional mathematical development (*Eqs. 4-8*) further describes the meaning and significance of WB. Comparing WB with other stratification indexes (including the buoyancy frequency) is very interesting; however, such an analysis is beyond the scope of this study and is proposed for future research.

Comment 3: The original contribution of this paper is to propose a WB threshold equal to 20W.m−2 to define global MLDs. However this constant WB threshold was derived from 6 transects in the Pacific ocean and I am wondering why this threshold should be constant everywhere in space and time? I think this point needs to be argued because ”local” conclusions; i.e. along transects; might not be valid at the global scale. For example, we see in Figure 4 that EBM does not always provide the ”best” MLD estimates compared to other methods, knowing that ”best” in this paper (see Section 3.2.2) is a visual criterion on the density profile, which is finally the classic scalar threshold. This last point shows that references have to be provided to define what is ”best”. Probably the definition of MLD depends on the question to be addressed. Is the EBM-derived MLD close to the density threshold-derived mixed-layer or to the turbulent kinetic energy threshold-derived mixing-layer (turbocline)?

Answer 3: Thank you for the comment. We agree with you that local results can not support global conclusions. To circumvent this, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics (*subsection 3.2. MLD methodologies intercomparison*). We removed the limitation of focusing only on “challenging regions”. We found that there is a lack of consistency among the methodologies in determining the MLD, but also that all the methodologies perform well under the oceanographic conditions for which they were built, according to the parameter being addressed. Since it is almost impossible to ensure that the true value of the MLD in any location and time is known, the accuracy of any methodology can not be determined, that is, the closeness of any MLD estimation to the true value. When comparing MLD methodologies, we can only evaluate their precision: the closeness between the different MLD estimates. Thus, we evaluated the precision of our methodology on a global scale and removed ad-hoc arguments from the description of our results.

Our methodology is a threshold method. The procedure to determine the MLD in terms of WB is based on a threshold of the density inhomogeneity along the mixed layer (*subsection 2.2: page 5, lines 129-133 and pages 7-8, lines 156-184*). By considering the density vertically integrated, our methodology could represent an improved or well-founded version of the threshold density criterion proposed by de Boyer Montégut *et al.* (2004) to define the MLD. Thus, we could conclude that our methodology is closer to the density threshold than to the turbulent kinetic energy threshold derived from the turbulence theory.

Minor points

Comment 4: Line 82 : Replace the lower bound of the integral zref by zeq in Equation 3.

Answer 4: Done.

Comment 5: Line 95 : This sentence is confusing. We see in Figure 1 that all density variations are associated with an increase in WB. Please clarify.

Answer 5: Suggestion accepted (*page 5, lines 120-122*).
"... the stratified profile (Fig. 1d) has a larger density variation than the winter profile (Fig. 1b), but the WB variation is larger in the winter than in the stratified profile: large density variations do not always correspond to large WB values."

Comment 6: Line 146 : Describe in few words the methods of Holte and Talley (2009) and Romero et al. (2023).

Answer 6: Suggestion accepted (*page 9, lines 208-212*).
"The third common MLD methodology is the multi-criteria method of Holte and Talley (2009), which calculates possible MLDs derived from threshold and gradient methods to select a final MLD estimate based on physical features in the profile. The recent MLD methodology is the sigmoid function fitting method of Romero et al. (2023), which computes the MLD and the maximum thermocline depth by evaluating the fit of the sigmoid function to the temperature profile."

Comment 7: Lines 177-181 : Sorry but I do not understand the procedure to select the WB threshold. That is what I understand. For a given W Bz you obtain a family of W Bi with i = 1, ..., n and each pair (i, i + 1) corresponds to a depth variation Δzi. The minimum Δzi selects zi and finally the W B threshold. Please clarify. In fact you could identify the structural change in W B just from W Bz . Figure 2 shows smaller depth differences between W Bz = 1.5 and W Bz = 2.5 than between W B = 10 and W B = 35, suggesting that W Bz = 2 might be a good threshold. So my questions are : Why is WB a better metric than WBz? Why do you need WBz to get a WB threshold? Please correct me and rewrite the lines 177-181 to clarify this important point.

Answer 7: We are sorry for the lack of clarity in that part of the procedure. However, in the revised version of the manuscript, we no longer use the first derivative of WB to find the MLD. We hope you find the new approach adequate (*subsection 2.2: page 5, lines 129-133 and pages 7-8, lines 156-184*).

Comment 8: Line 217 : Please give a reference for the inter-quartile formula, which detect outliers. Note that this formula is valid for a Normal distribution. Is it the case here?

Answer 8: You are right that the formula to detect outliers is valid for normal distributions. In the revised version of the manuscript, the intercomparison of different MLD methodologies was made on a global scale. We removed the limitation of focusing only on "challenging regions"; thus, the detection of outliers was no longer needed.

Comment 9: Line 250-254 : "EBM better identifies the relatively homogeneous upper ocean layer ..." This sentence is quite subjective, "best" is relative to the eye's estimation. For example on Figure 4c, I find HT09 better than EBM.

Answer 9: Please refer to answer 3 for a response to this comment.

Comment 10: Section 3.3 : It would be instructive to compare the EBM monthly MLD climatology shown Figure 5 with that of Boyer Montégut.

Answer 10: Thank you for your suggestion, which we implemented. In *subsection 3.2 MLD methodologies intercomparison*, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics. In the *Supplement*, we have maps of MLD climatologies obtained with different methodologies, which further support the comparison.

Comment 11: Line 275-279: MLDs around 600-700 m in the Greenland, Labrador, Iceland, Norway Seas, southern Pacific and Indian Oceans seem shallow. Another climatology would be helpful (Boyer Montégut?)

Answer 11: Thank you for your suggestion. In *subsection 3.2 MLD methodologies intercomparison*, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics. In the *Supplement*, we have maps of MLD climatologies obtained with different methodologies, which further support the comparison. The new MLD estimates in those regions are larger than the previous ones. The MLD can reach values of up to 945 m in the south of Iceland and the Labrador Sea, 1074 m in the Greenland, Iceland, and Norwegian seas region, and 614 m in the South Pacific and South Indian Oceans (*Page 11, lines 241-245)*.

Comment 12: Line 285-286 : Mention that the deepest MLDs are collocated with strong winds in these regions, add references.

Answer 12: Thanks for the comment. We are sorry for the omission. That sentence was removed in the new version of the manuscript. However, we are very careful not to omit relevant references throughout the manuscript.

Comment 13: Figure 6 : What is the unit of density in the lower panel ? I do not understand what is plotted, is it the density difference between the mixed-layer base and z = 10 m? Higher/lower densities occur in summer/winter. I would expect the opposite. But ok if it is the difference. Please check.

Answer 13: Of course. We referred to the histogram (expressed in density of probability). In the new version of the manuscript, we plotted cumulative density functions (CDFs).

Figures 7 and 8 in the prior version of the manuscript are now *Figs. 6 and 7*. In them, we plotted the absolute differences in potential density and conservative temperature from the reference depth of 10 m to the MLD, respectively. In constructing the mixed layer definition, the density variations along the mixed layer throughout the year were established; the majority of the world ocean should have potential density differences in the interval (0.0150, 0.0300) kgm-3. The differences in conservative temperature from the reference depth of 10 m to the MLD shown in *Fig. 7* are heterogeneous in space and change over time with a type of seasonal variation. The temperature differences are generally large for large MLDs and vice versa; however, the temperature differences do not have the same structure or seasonal variation as those of the MLD. We found that the energy-based mixed layer is very close to quasi-homogeneity in density and temperature: almost 100% of the world ocean has density differences of less than 0.03 kg m-3, and 95% of the world ocean has temperature differences of less than 0.2°C throughout the year (*page 15, lines 275-285 and page 16, lines 286-293)*.

Comment 14: Line 288-289 : How do you explain the bimodal distribution?

Answer 14: Thank you for the comment; this is a very good observation. This bimodality is because of the region we used to compute the MLD histograms (*Fig. 6* in the prior version of the manuscript). In this region, we consistently have a region with large MLDs and a region to the north with consistently smaller MLDs. That structure in the MLD is the origin of the observed bimodality, which could be substantially reduced if we restrict the analysis to latitudes south of 50°S. In the new version of the manuscript, the Southern Ocean analysis is no longer present; we extended the analysis to a global scale, considering the different seasons (*subsection 3.2 MLD methodologies intercomparison*).

Comment 15: Line 291 : MLD variances are not shown, mention ”not shown”.

Answer 15: We are sorry for the omission. In the new version of the manuscript, the Southern Ocean analysis is no longer present; we extended the analysis to a global scale, considering the different seasons (*subsection 3.2 MLD methodologies intercomparison*).

Comment 16: Line 292-294 : The skewness of the distributions are not consistent with that of Johnson and Lyman (2022). Please explain why.

Answer 16: Thank you. This is a very good observation. The MLD should have persistence throughout the year; thus, the skewness of the MLD distribution is expected to remain with the same sign throughout the year. We have no insight into the result obtained by Johnson and Lyman (2022), but it would be interesting to investigate it. In the new version of the manuscript, the Southern Ocean analysis is no longer present; we extended the analysis to a global scale, considering the different seasons (*subsection 3.2 MLD methodologies intercomparison*). In the new version of the manuscript, we compared different MLD methodologies using different statistical metrics on a global scale during each season (*Table 1*). For each methodology, the skewness remains with the same sign throughout the year, which supports our prior claim.

Comment 17: Figure 7: Upper panel, another climatology would be useful for comparison. Lower panel, what is the variable and its unit on the ”y” axis?

Answer 17: Thank you for your suggestion, which we implemented. In *subsection 3.2 MLD methodologies intercomparison*, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics. In the Supplement, we have maps of MLD climatologies obtained with different methodologies, which further support the comparison.

Regarding the units on the “*y*” axis, we referred to the histogram (expressed in density of probability). In the new version of the manuscript, we plotted cumulative density functions (CDFs).

Comment 18: Figure 8: Same remarks as Figure 7.

Answer 18: Thank you for your suggestion, which we implemented. In *subsection 3.2 MLD methodologies intercomparison,* we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics. In the Supplement, we have maps of MLD climatologies obtained with different methodologies, which further support the comparison.

Regarding the units on the "*y*" axis, we referred to the histogram (expressed in density of probability). In the new version of the manuscript, we plotted cumulative density functions (CDFs).

Comment 19: Lines 397-400 : These density and temperature thresholds are not so far from the usual ones. Does that means that your EBM-derived MLD climatology does not depart significantly from other climatologies? As mentioned above (point 5) it would be instructive to compare the Figure 5 with another climatology.

Answer 19: Thank you for your suggestion, which we implemented. In *subsection 3.2 MLD methodologies intercomparison*, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics. In the Supplement, we have maps of MLD climatologies obtained with different methodologies, which further support the comparison. When comparing MLD methodologies, we evaluated the precision of our methodology, i.e, the closeness between the MLD estimates obtained with different methodologies. We found that our methodology is precise.

Review of the manuscript: https://doi.org/10.5194/egusphere-2024-4079

Title: The global ocean mixed layer depth derived from an energy approach

Reviewer #3

We appreciate all your comments. After a careful revision that incorporated your comments and those of the other two reviewers, this version has greatly improved. We trust that you will find the revisions substantial and consider our study a valuable contribution to mixed-layer depth research. Below is a description of the major modifications made to the manuscript:

- The methodology (subsection 2.1) was extended by including details concerning the connection of the work done by buoyancy (WB) with the turbulent kinetic energy budget.
- We revised our definition of the mixed layer and how to calculate its depth (subsection 2.2. Defining the mixed layer). The mixed layer depth (MLD) was not calculated with a unique WB equipotential.
- Since we adjusted the MLD calculation, we rewrote the Results and Discussion sections entirely. The Conclusions and Abstract were also rewritten accordingly.
- The Supplement was rewritten to provide a detailed comparison of our MLD methodology with the common ones; we included global monthly climatologies of the MLD and the WB at the MLD.
- The dataset containing the monthly climatology of mixed layer depth and the derived variables calculated in this study is now publicly available at SEANOE (https://doi.org/10.17882/106181).

Responses to specific comments are shown below.

<u>Summary</u>

Comment 1: The mixed layer is an important oceanographic parameter, playing a significant role in understanding upper-ocean dynamics and air-sea interactions. Therefore, accurately calculating the mixed layer depth (MLD) is crucial. Over time, many researchers have proposed various methods to estimate the MLD, which can generally be categorized into three types: threshold methods, energy-based methods, and geometric shape methods. However, due to the very nature of the mixed layer, there is no perfect method. How can we quantitatively define "mixing"? It always requires some reference values, and the use of such reference values inherently leads to spatial or temporal dependence in the applicability of different methods.
This work approaches the problem from an energy perspective, proposing a new method for calculating MLD based on the amount of buoyancy work. The authors demonstrate through practical applications that this method performs reasonably well. Overall, I find this to be a very interesting study. It not only helps us better understand the mixed layer, but also provides an additional option for calculating its depth.

Answer 1: We really appreciate that you find our approach to estimating the MLD interesting. We agree with you that in MLD studies, it is necessary to clarify what aspects of the mixed layer are being studied and the considerations taken. That information was properly described in *subsection 2.2 Defining the mixed layer (page 5, lines 124-129)*.

Comments

Comment 2: I believe the authors may have overstated the performance of their proposed method. In the abstract, they claim that this method provides "a robust criterion based on physical principles." However, it should be noted that: 1) The authors still use a threshold value (20 J/m³) as the criterion for defining the mixed layer. This value is derived only from a few observational sections, and whether it is applicable globally remains questionable. 2) The authors' evaluation of "good" or "bad" performance seems to focus on regions where traditional methods do not perform well. Is this method better across all seasons and global ocean regions? In areas where density profiles change gradually, the MLD itself is inherently ambiguous—how should a "good" standard be defined in such cases? 3) The authors mention in the introduction that "in regions with vertically compensated layers, the density threshold may overestimate the MLD." Can the proposed method in this paper avoid this issue? Considering that the method is still based on density, it seems unlikely that this problem can be completely avoided.

Answer 2: Thanks for the comment. We acknowledge that in the prior version of the manuscript, we derived our results based on non-robust arguments. To fix that, we proposed a quantitative procedure to calculate the MLD; it is applicable in regions where density profiles change gradually. We agree that it is not obvious that a unique WB equipotential should determine the MLD globally year-round, which was obtained from Eq. 8 and represented in Fig. 2; please refer to the detailed analysis of the WB metric in *subsections 2.1 and 2.2*. Results of the WB threshold characterizing the MLD are shown in *Figs. 4 and 11*. Thus, the MLD was no longer calculated with the WB equipotential of 20 J/m³. The new results are completely different from those of the prior version of the manuscript due to the adjusted definition of the mixed layer.

We agree with you that local results can not support global conclusions. To circumvent this, we compared (qualitatively and quantitatively) our methodology with others on a global scale, considering different metrics (*subsection 3.2. MLD methodologies intercomparison*). We removed the limitation of focusing only on "challenging regions". We found that there is a lack of consistency among the methodologies in determining the MLD, but also that all the methodologies perform well under the oceanographic conditions for which they were built, according to the parameter being addressed. Since it is almost impossible to ensure that the true value of the MLD in any location and time is known, the accuracy of any methodology can not be determined, that is, the closeness of any MLD estimation to the true value. When comparing MLD methodologies, we can only evaluate their precision: the closeness between the different MLD estimates. Thus, we evaluated the precision of our methodology on a global scale and removed ad-hoc arguments from the description of our results.

For vertically compensated layers, our methodology may overestimate the MLD, in a similar way that the density threshold criterion of de Boyer Montégut et al. (2004); nonetheless, using WB, we can measure the degree of inhomogeneity of the water column associated with the compensated layer and investigate whether it is intense enough to suppress mixing. While our methodology may not provide a better or more meaningful MLD estimate than other methodologies, it does measure the water column inhomogeneity in terms of energy, a unique feature that other methodologies lack (*pages 21-23, lines 389-393)*.

Comment 3: Although the authors express the buoyancy work integral in the form of Equation 3, in essence, this equation is the depth integration of density anomalies. This means the method still heavily depends on the choice of threshold value—especially in summer with shallow mixed layers when density increases slowly with depth. In such cases, different thresholds could lead to significantly different results. I recommend the authors conduct a deeper discussion and analysis of the limitations of their method, rather than focusing solely on its advantages.

Answer 3: We agree with you that our methodology is a threshold method. The procedure to determine the MLD in terms of WB is based on a threshold of the density inhomogeneity along the mixed layer (*subsection 2.2 Defining the mixed layer*). By considering the density vertically integrated, our methodology could represent an improved or well-founded version of the threshold density criterion proposed by de Boyer Montégut *et al*. (2004) to define the MLD. As you suggested, we carried out a deeper discussion and analysis of the limitations of our methodology. We conducted a detailed analysis of the WB metric and its physical significance (*subsection 2.1 An energy measure of the vertical homogeneity of the water column*) and evaluated its performance by comparing it with other methodologies (*subsection 3.2. MLD methodologies intercomparison*). The *Discussion section* extends the analysis of our MLD methodology and its downsides.

Comment 4: Although there is no universally accepted standard method for determining MLD, there is a parameter that can reflect the effectiveness of a given method to some extent—the Quality Index (https://doi.org/10.1029/2003JC002157). I suggest the authors consider using this parameter to evaluate the performance of their method.

Answer 4: We really appreciate your suggestion, which we implemented (*page 11, lines 255-267*). In the analysis of our MLD climatology, we adapted the Quality Index of Lorbacher et al. (2006) to WB and calculated it (*Fig. 5*). Results showed that, according to the quality index, our methodology performs very well in almost all the world ocean year-round: 96.72% of the world ocean has quality index values larger than 0.8 and in only 0.03% of the world ocean we have values less than 0.5. There is room for improvement in the use of the quality index in MLD studies; however, it was useful to evaluate the general performance of our methodology.