Author Response to Anonymous Referee #1
Submitted on 14 Jan 2026

Dear Anonymous Referee #1,

Thank you for your thoughtful feedback. We have addressed all comments below sequentially.

Best Regards,
Liz, on behalf of all co-authors

The manuscript titled "A Revised Temperature-Dependent Remineralization Scheme for the Community Earth System Model (v1.2.2)" has been revised in response to the reviewers' comments. However, the revisions are limited, and the manuscript still requires substantial improvement and scientific robustness before it can be considered for publication in my opinion. Several issues previously raised remain unaddressed. I provide additional comments below.

MAJOR COMMENTS

1. In Figure 9, as noted previously, the labels for PI and Tdep remain unchanged. The scatter corresponding to PI still shows higher R2 and lower cRMSE, which contradicts the conclusions stated in the manuscript. This discrepancy needs to be clarified; otherwise, the results suggest that the skill of the new remineralization parameterization is worse than that of the previous scheme.

Thank you for noting the Figure 9 labels. In our previous response, we suggested that the labels could have been reversed. However, upon review, we confirmed that they were correct and that no change was necessary. This clarification is now stated explicitly in the revised manuscript in the Figure 9 caption comments.

The objective of the temperature-dependent remineralization (Tdep) implementation was to maintain upper-ocean nutrient concentrations comparable to the control simulation while improving transfer efficiency as the primary metric. Figure 9 shows that the PI control exhibits slightly lower cRMSE and higher R^2 than the Tdep simulation. However, the magnitude of the cRMSE difference is comparable to the characteristic inter-product mismatch between WOA and GLODAP for global phosphate concentrations, reported to be ~0.03 for the full water column. Differences in cRMSE of this magnitude are therefore interpreted as marginal and within bounds of the observational product disagreement. Clarification of this point was added in line 285.

Similarly, R^2 values for the Atlantic, Southern Ocean, and Indian Ocean are nearly identical between configurations, while the global mean R^2 differs by ~0.02, which is considered a small difference in pattern agreement. The largest R^2 difference occurs in the Pacific basin (~0.1). This deviation is considered meaningful and is addressed explicitly in expanded Discussion (~line 376) on the role of tropical nutrient distributions and regional sensitivity.

In addition, Figure 9 is not well suited for evaluating model performance, as the two metrics shown are independent of each other. Instead, I recommend using a Taylor diagram, which is a standard tool for assessing climate model performance in terms of correlation, root-mean-square error, and the ratio of variances (Taylor, 2001).

Taylor,K.E.(2001). Summarizing multiple aspects of model performance in a single diagram. JGR: Atmosphere, 106 (D7), 7183–7192

We agree that Taylor diagrams are a standard and widely used tool for evaluating climate model performance (Taylor, 2001). The metrics shown in Figure 9 (cRMSE and R^2) correspond to two of the three components summarized in a Taylor diagram. While these metrics diagnose different aspects of model performance, they are not statistically independent, as cRMSE is mathematically related to correlation and standard deviation within the Taylor framework. We therefore interpret cRMSE and R^2 as complementary diagnostics of model accuracy and pattern agreement.

We intentionally did not include the standard deviation ratio, as it is not central to the scientific question addressed here. Our focus is on model–data agreement in upper-ocean nutrient distributions and residual error magnitude, for which cRMSE and R^2 provide a sufficient and appropriate assessment of model performance.

2. In Figure 11, the authors average over the IAO region, which appears to combine the Indian Ocean and the Arctic Ocean. Is there a specific reason for grouping these two regions into a single category? These regions represent very different oceanic environments—for example, the Indian Ocean is a warm, predominantly tropical basin, whereas the Arctic Ocean is a cold, polar system.

In addition, the authors state that the regional definitions follow Weber et al. (2016); however, that study does not combine the Indian and Arctic Oceans into a single region. Finally, even if the ocean-region names are adopted from Weber et al. (2016), it is strongly recommended that all abbreviations be explicitly defined (as was done for AAZ and ETP) to avoid ambiguity.

The regional definitions used in Figure 11 were taken directly from the basin mask files provided by Thomas Weber. You are correct that the IAO region combines the Indian Ocean and the Arctic Ocean. In Weber et al. (2016), this combined region is defined in the basin mask files but is excluded from their subsequent analysis (also noted in personal correspondence with Thomas Weber), with the focus placed instead on the remaining eight regions.

While we acknowledge that the Indian and Arctic Oceans represent very different oceanographic environments, we chose to retain the IAO region in this analysis in order to provide a complete, global assessment of basin-scale responses. Including this region allows us to evaluate whether trends observed elsewhere, particularly improvements in transfer efficiency with the inclusion of temperature-dependent remineralization, are also evident when considering the full global ocean. We have clarified this choice in the revised text (line 326) to avoid ambiguity regarding consistency with Weber et al. (2016).

We agree that explicit definition of all acronyms is important for clarity. A complete list of regional abbreviations has now been added to the Figure 11 caption, consistent with the definitions provided elsewhere in the manuscript (e.g., AAZ and ETP).

3. The authors stated in their response that clarification regarding the use of the last 30 years would be provided in Section 2.3; however, the manuscript does not currently include such an explanation. As I understand it, the analysis involves interannual variability, with results presented as 30-year averages. Nevertheless, climate models inherently exhibit year-to-year variability arising from internal variability and model-specific characteristics, particularly in CESM. I therefore recommend that the associated uncertainties be explicitly represented by showing the ranges of interannual variability, for example using standard deviations. Specifically, Figures 5 and 6 could include

latitude-dependent shading to indicate variability ranges, and Figure 11 could present uncertainty ranges (e.g., error bars) for each bar.

Apologies for the confusion regarding the previously stated addition of 30 year averages in Section 2.3. This clarification was actually added in Section 2.1 (Line 94) in the previous track-changes version.

The analyses presented here are intended to characterize equilibrium or climatological differences between model configurations rather than interannual variability. For this reason, we use the 30-year average to suppress seasonal and year-to-year fluctuations and to represent the equilibrated mean state. This averaging period is standard practice for CESM equilibrium analyses and is sufficient to isolate the signal of interest. However, visualizing the range of interannual variability captured in these averages is a fair concern. To address this, we have added a 1-σ band to the global zonal TE in Figure 5, and commented further in both the figure caption text and main text noting the addition.

Figure 11 depicts the residuals of TE in PI and Tdep vs those reported in Weber. Since the Weber product is statistically derived based on both empirical data and modeling products, these data do have interannual variability for comparison.

4. Finally, Figure 7 requires additional clarification by providing quantitative performance metrics, such as the RMSE between PI and Tdep for each location. Aside from the two Equatorial Pacific regions, PI appears to show better agreement with observations at several sites (e.g., ALOHA, Peru, Arabian Sea). Therefore, further information is needed to clearly demonstrate that the new temperature-dependent remineralization parameterization represents a genuine improvement over the previous formulation, rather than a degradation in model performance.

We appreciate the reviewer's comment. As mentioned in response to #1 above, the primary goal of this analysis was to improve transfer efficiency performance with the inclusion of temperature-dependent remineralization (Tdep), and this is the main metric that was targeted for improvement. Following this, POC flux attenuation and upper-ocean nutrient content (noted above) were used as secondary diagnostics to confirm that the model behavior remained largely consistent in both the control and Tdep simulations. While transfer efficiency improves at some sites, such as in the Equatorial Pacific, other regions like Peru show more modest differences, as noted ~ line 266.

We used methodology following Laufkötter et al. (2017) for the POC flux attenuation analysis, applying it to an expanded data compilation. This approach is a widely accepted benchmark for assessing model performance, and although there are alternative methods, we chose this one due to its established use in similar studies for validating the behavior of biological carbon pump models.

Author Response to Anonymous Referee #2
Submitted on 14 Jan 2026

Dear Anonymous Referee #2,

Thank you for your thoughtful feedback. We have addressed all comments below sequentially.

Best Regards,
Liz, on behalf of all co-authors

1. **In the response letter, the authors mentioned additional discussion on why the flux attenuation in the Southern Ocean (Fig. 7) is almost identical in both PI and Tdep, while their transfer efficiency in Fig. 6 is vastly different. This difference is noted in the manuscript (line 275), but the discussion would benefit from a few more details. I suggest adding an explanation for this discrepancy.**

Thank you for highlighting this important point. Indeed, there is a marked difference in the transfer efficiency of the Southern Ocean region, as shown in Figure 6. A key detail is with the depth bands used for the transfer efficiency metric, which is at a fixed 100 meters and 1000 meters. Looking at the flux at only 1000 meters for the Southern Ocean site, the modeled PI flux is lower (around 25%) than the Tdep (around 30%). Given similar export at 100 meters, this would account for the higher transfer efficiencies. This subtlety, related to the 100 meter and 1000 meter fixed depths used by the community for TE, is something that we feel is an area of further discussion and potential for development of a new standardized metric for TE that can account for upper column (0-100m) adjustments of the pump (see e.g., line 355), perhaps by considering integrated depth bands versus an individual depth similar to discussed in Buesseler et al., 2020.

In terms of consistent trends in cold water regions for both TE and POC flux as you note, we have also added further discussion in the body of the text (line 266).

2. **Similarly, the discussion regarding Fig. 8 (see previous comment) would also benefit from more details: I recommend to include 1-2 sentences in the discussion about *why* the phosphate concentration in the Tdep simulation does not perform as well as in the PI simulation despite the improvements in simulating transfer efficiency. This is an important point for potential future improvements.**

This is a great point and a topic of further discussion. We have added additional discussion (line 376).