We thank Reviewer #1 for the constructive comments and suggestions, which greatly help to improve the quality of our manuscript. We have made revisions and replied to all the comments. Please find the point-by-point responses to the comments below. Our responses are shown in "Blue" and the changes in the manuscript are shown in "Red". The line numbers correspond to those in the clean version of our revised manuscript.

Response to the comments from Reviewer #1

General Comment:

This manuscript presents the implementation of a newly developed four-dimensional ensemble variational (4DEnVar)-based Weakly Coupled Ocean Data Assimilation (WCODA) system within the E3SMv2 framework. The study provides a comprehensive evaluation of how assimilating monthly ocean temperature and salinity improves the model's ability to capture both interannual and decadal climate variability.

Response:

Thank you very much for providing us with very useful comments. Each of the specific comments raised below has been carefully considered, and the manuscript has been revised accordingly. Please find our point-by-point responses to each of your specific comments.

Comment#1:

The manuscript is motivated by the critical need for "optimal initial conditions" for seasonal-to-decadal (S2D) predictions. While I agree that a necessary condition for a good initial state is its ability to phase climate variability with observations, it is crucial to distinguish between assimilation skill and initialization quality. High correlation in an assimilation run does not inherently guarantee an optimal initial condition for forecasts. For example, a strong nudging approach might yield high skill during the assimilation cycle but can trigger significant initialization shocks and rapid drift in subsequent forecasts due to dynamical inconsistencies induced by physically inconsistent updates. Since the current study primarily evaluates performance within a continuous assimilation cycle (ASSIM) without conducting actual hindcast experiments to verify forecast initialization quality, the frequent framing of the results as "optimal" or "realistic" initial conditions feels somewhat premature. I suggest that the authors moderate the tone of these assertions rather than definitively claiming them as "optimal initial conditions" at this stage.

Note to Authors: If this point is robustly reinforced with diagnostic evidence, the current framing of "optimality" in the manuscript may be justified. However, if this remains as a purely speculative discussion, I suggest that the tone of the primary claims in Comment 1 and the Introduction be appropriately moderated to reflect a more cautious interpretation.

Response:

We thank the reviewer for this constructive comment regarding the critical distinction between assimilation skill and forecast initialization quality. The same 4DEnVar methodology used in this study (specifically, the DRP-4DVar approach) was previously employed by He et al. (2017)

to construct an ocean data assimilation system within a different coupled climate model (FGOALS-g2). He et al. (2017) have already demonstrated that this DRP-4DVar approach effectively reduces the initialization shock that arises in the early years of initialized hindcasts. Unlike nudging methods, observational information is assimilated along the continuous trajectory of the model solution within the four-dimensional assimilation window, thereby generating dynamically consistent initial conditions.

To further verify this behavior within E3SMv2, additional hindcast experiments have been conducted and are now presented in the new Figure A4. As shown in Fig. A4, the ensemble mean hindcast experiments initialized from this 4DEnVar-based WCODA system at five-year intervals from 1981 to 2006 closely track the observed annual mean global average surface air temperature anomalies, without the occurrence of significant initialization shock in the early years. Consistent with the findings of He et al. (2017), this result indicates that the 4DEnVar-based WCODA system produces dynamically consistent initial conditions and substantially reduces the initialization shock in E3SMv2 as well.

Based on this comment, we have added the reference to He et al. (2017) to explicitly illustrate the capability of the DRP-4DVar method to generate dynamically consistent initial conditions and reduce early-stage initialization shock (L135-137). In addition, we have added a new Figure A4 in the Discussion to indicate that the 4DEnVar-based WCODA system in this study substantially reduces the initialization shock in E3SMv2 (L499-504). Furthermore, we have moderated the tone in the abstract by removing the assertion of "realistic initial conditions" and replacing it with a more measured statement that the WCODA system lays the foundation for seasonal-to-decadal prediction applications (L33-35).

L135-137: In particular, the DRP-4DVar approach has been shown to generate dynamically consistent initial conditions that effectively mitigate the initialization shock during the early years of initialized hindcasts (He et al., 2017).

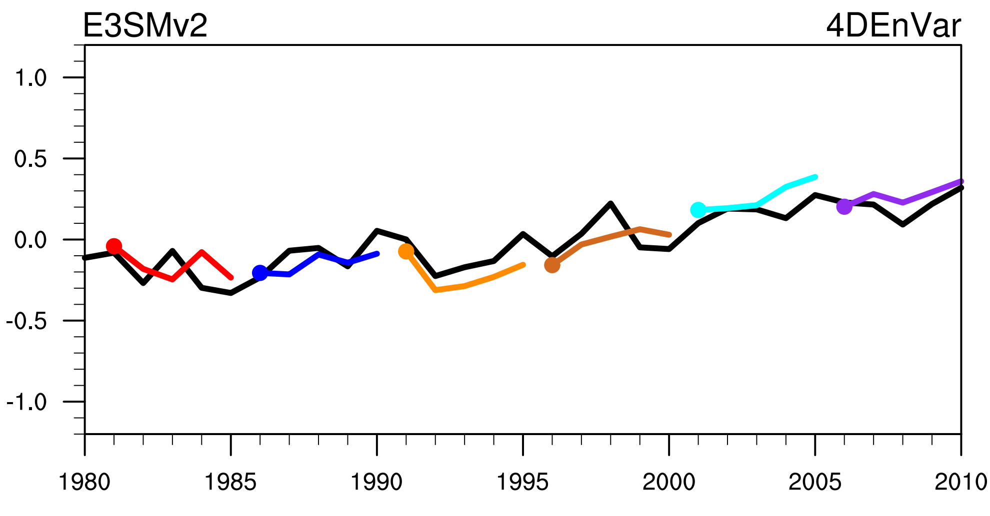

Figure A4. Time series of annual mean global average surface air temperature anomalies (units: °C) from 1980 to 2010. The black line denotes the observation, while the colored lines denote the ensemble mean hindcast experiments produced by E3SMv2 initialized with the 4DEnVar-based weakly coupled ocean data assimilation system. Each hindcast is initialized at the start of a different year (1981, 1986, 1991, 1996, 2001, and 2006; marked by filled circles) and is then integrated freely for five years. For each target initialization year, three ensemble members are initialized in the preceding September and November and in January of the target year. The ensemble mean is then computed from these three members throughout the entire five-year hindcast period.

L499-504: To assess the quality of the resulting initial conditions, additional hindcast experiments were conducted at five-year intervals from 1981 to 2006 (Figs. A4 & A5). As shown in Fig. A4, the ensemble mean hindcasts initialized from the WCODA system closely track the observed annual mean global average surface air temperature anomalies, with no appreciable initialization shock during the early years of the hindcasts.

L33-35: thereby generating dynamically consistent initial conditions that support more accurate simulations of Earth system variability and lay the foundation for seasonal-to-decadal prediction applications.

Comment#2:

To further highlight the advantages of the proposed 4DEnVar-based WCODA system, I suggest the authors expand their discussion on the dynamical consistency of the assimilation products, which is a core interest for the readers of Geoscientific Model Development. The current manuscript focuses heavily on statistical metrics (e.g., correlations with observations). However, the true merit of WCODA may reside in its potential to better preserve physical balances in variables not directly assimilated (e.g., atmospheric circulation or land states) by allowing the model's internal physics to mediate the adjustment during the coupled forecast step of the DA cycle. Specifically, I recommend the following: a) Please discuss how the spatial patterns and variability of atmospheric and land components serve as evidence that the system is not merely "fitting to observations" but is maintaining a dynamically balanced state. In contrast to overconstrained methods (such as strong nudging) that often lead to "initialization shocks," the successful propagation of ocean information into the atmosphere without disrupting teleconnection patterns (as shown in Figure 14) may be a significant technical strength of this WCODA implementation.

Response:

Thank you very much for this insightful suggestion. We fully agree that the principal merit of the WCODA framework lies not in merely fitting observations but in its ability to preserve physical balance in variables that are not directly assimilated, and we appreciate the opportunity to elaborate on this point.

The successful propagation of ocean information into the atmosphere through the model's internal physics is evidenced by the well-reproduced ENSO teleconnection patterns of winter surface air temperature and precipitation in ASSIM (Fig. 14). These improvements are achieved without any direct assimilation of atmospheric or land observations, indicating that ocean information is propagated into the atmosphere through the model's internal physical processes during the coupled forecast step of the assimilation cycle. To further demonstrate this point in actual hindcast experiments, the ENSO teleconnection patterns reproduced in the initialized hindcasts are examined in the new Figure A5. As shown in Fig. A5, the regression patterns of winter surface air temperature and precipitation onto the Niño 3.4 index in the hindcasts closely resemble the observed and ASSIM patterns, further demonstrating the dynamical consistency of the WCODA-generated initial conditions. Together with the mitigation of initialization shock demonstrated in Fig. A4, these results suggest that the 4DEnVar-based WCODA system is capable of producing dynamically consistent initial conditions in which ocean information is transferred to other Earth system components through multi-component interactions, and the hindcasts initialized from these dynamically balanced initial conditions preserve similar teleconnection patterns as in the assimilation run.

In response to this suggestion, we have added a new Figure A5 to examine the ENSO teleconnection patterns of winter surface air temperature and precipitation reproduced in the initialized hindcast experiments. In addition, we have expanded the discussion (L504-512) in the revised manuscript to elaborate on the dynamical consistency of the WCODA-generated initial conditions, with reference to both the assimilation experiment (Fig. 14) and the initialized hindcasts (Fig. A5).

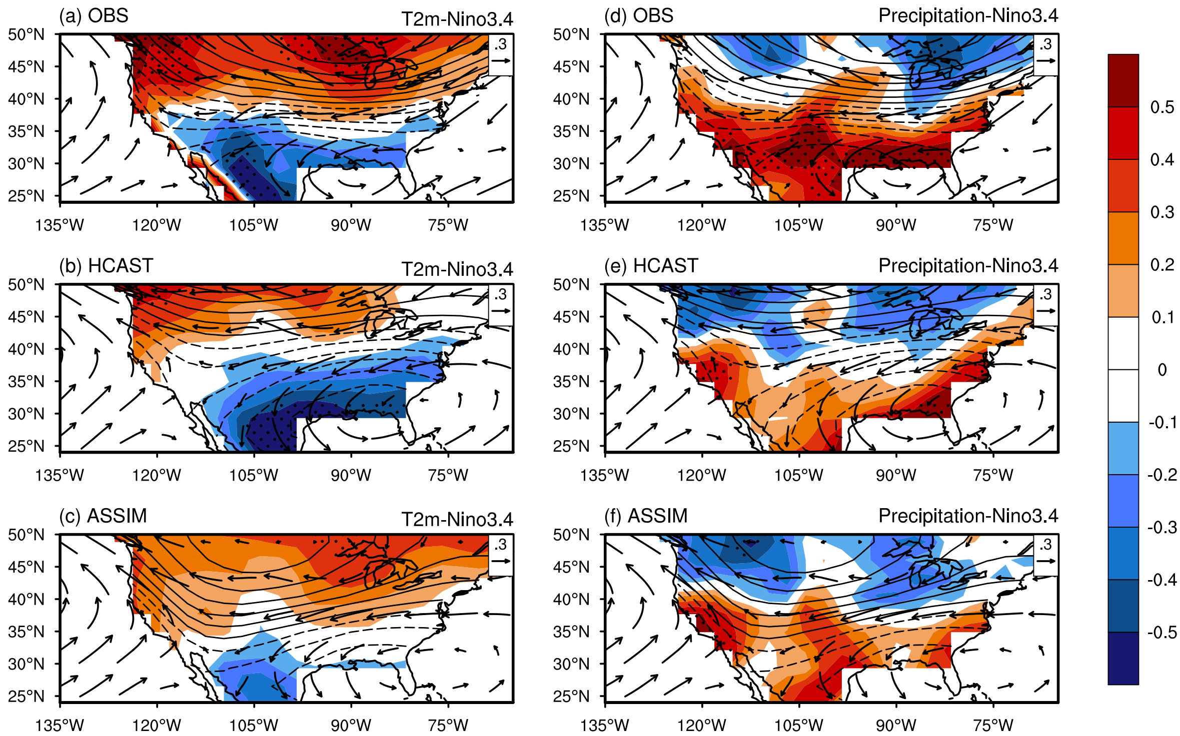

Figure A5. Regression patterns of winter surface air temperature (left column; shaded) and precipitation (right column; shaded), 500hPa geopotential height (contours), and 850hPa winds (vectors) on the standardized winter Niño 3.4 index from 1981 to 2010 for (a, d) the observation, (b, e) HCAST, and (c, f) ASSIM. The HCAST results are based on the ensemble mean hindcasts initialized at five-year intervals from 1981 to 2006, as described in Fig. A4. The successive five-year segments are then concatenated to form a continuous time series from 1981 to 2010, without any bias correction applied. Dotted areas indicate regions where the regressions are statistically significant at the 95% confidence level.

L504-512: The dynamical consistency of the WCODA-generated initial conditions is further supported by the well-reproduced ENSO teleconnection patterns in both the assimilation experiment (Fig. 14) and the initialized hindcasts (Fig. A5). Specifically, the regression patterns of winter surface air temperature and precipitation onto the Niño 3.4 index in the hindcasts (Fig. A5) closely resemble the observed and ASSIM patterns. These results suggest that the 4DEnVar-based WCODA system is capable of producing dynamically consistent initial conditions in which ocean information is transferred to other Earth system components through multi-component interactions, and the hindcasts initialized from these dynamically balanced initial conditions preserve similar teleconnection patterns as in the assimilation run.

Comment#3:

b) While a direct numerical comparison with other DA methods (like simple nudging or uncoupled DA) within the same E3SM framework would be ideal, I recognize this may be beyond the current scope. Instead, I suggest the authors include a more robust discussion comparing their results with documented issues in other decadal prediction systems that utilize ocean-only initialization. For instance, according to Supplementary Table 3 in Meehl et al. (2021), various prediction systems (e.g., BCC-CSM-MR, FGOALS-g3/f3, MRI-ESM2, CESM1, and NorCPM1) as well as studies by Chikamoto et al. (2019) and Kim et al. (2025) have documented challenges with ocean-only initialization. Please clarify how the authors' approach potentially mitigates the uncoupled-related shocks found in those systems. Elaborating on how this method leads to a more stable and physically sound initial state would significantly strengthen the initial motivation regarding "optimal initialization" (as mentioned in Comment 1) and demonstrate the system's potential to enhance actual S2D predictability.

Response:

We agree that contrasting the 4DEnVar-based WCODA implementation with documented challenges in ocean-only initialized systems would strengthen the core motivation of this study. A well-recognized limitation of ocean-only initialization is that ocean initial conditions are generated independently of the atmosphere and land, resulting in dynamical imbalances across components. These imbalances can trigger significant initialization shock and rapid drift that degrade forecast skill in the early years of the hindcasts, as documented across a range of prediction systems (Meehl et al., 2021; Chikamoto et al., 2019; Kim et al., 2025).

The 4DEnVar-based WCODA system developed in this study is designed to address these limitations. By performing data assimilation within the fully coupled E3SMv2 framework, ocean observational information is propagated to the atmosphere and land through multi-component interactions during the coupled forecast step, thereby generating more dynamically consistent initial conditions across components. Furthermore, the DRP-4DVar increments are derived by fitting the oceanic analysis within the four-dimensional assimilation window along the trajectory of the fully coupled model rather than the standalone ocean model, ensuring that the resulting ocean state is dynamically consistent with the atmosphere and land. As demonstrated in Figs. A4 and A5, the hindcast experiments initialized from this WCODA system exhibit no appreciable initialization shock in the early forecast years and reproduce physically coherent ENSO teleconnection patterns, suggesting that the coupled initialization framework adopted here offers a practical pathway for mitigating the initialization shock.

Based on this comment, we have added a new paragraph to the Discussion (L513-526) for contrasting the 4DEnVar-based WCODA system with traditional ocean-only initialized decadal prediction systems, with reference to documented challenges in such systems (Chikamoto et al., 2019; Meehl et al., 2021; Kim et al., 2025). Additionally, we have also revised the abstract to include hindcast-based evidence (Figs. A4 and A5) for suggesting the dynamical consistency of the coupled initialization framework (L26-29).

L513-526: The capacity of this 4DEnVar-based WCODA system to mitigate initialization shock and maintain cross-component dynamical consistency is particularly noteworthy when contrasted with traditional ocean-only initialized decadal prediction systems. A well-recognized limitation of ocean-only initialization is that ocean initial conditions are generated independently of the atmosphere and land, resulting in dynamical imbalances across components that can trigger pronounced initialization shock and rapid drift in the early years of the hindcasts (Chikamoto et al., 2019; Meehl et al., 2021; Kim et al., 2025). In contrast, by performing data assimilation within the fully coupled E3SMv2 framework, ocean observational information is propagated to other components through multi-component interactions during the coupled forecast step. Furthermore, the DRP-4DVar increments are derived by fitting the oceanic analysis within the four-dimensional assimilation window along the trajectory of the fully coupled model rather than the standalone ocean model. These features promote more dynamically consistent initial conditions across model components, thereby offering a practical pathway for reducing the initialization shock.

L26-29: Additional hindcast experiments initialized from the WCODA system exhibit no appreciable initialization shock in the early years and reproduce physically coherent ENSO teleconnection patterns, suggesting the dynamical consistency of the coupled initialization framework.

References:

Chikamoto, Y., Timmermann, A., Widlansky, M. J., Zhang, S., and Balmaseda, M. A.: A drift-free decadal climate prediction system for the Community Earth System Model, Journal of Climate, 32, 5967–5995, https://doi.org/10.1175/JCLI-D-18-0788.1, 2019.

He, Y., Wang, B., Liu, M., Liu, L., Yu, Y., Liu, J., Li, R., Zhang, C., Xu, S., Huang, W., Liu, Q., Wang, Y., and Li, F.: Reduction of initial shock in decadal predictions using a new initialization strategy, Geophysical Research Letters, 44(16), 8538–8547, https://doi.org/10.1002/2017GL074028, 2017.

Kim, Y.-Y., Lee, J.-Y., Chikamoto, Y., Timmermann, A., Lee, S.-S., Kwon, E. Y., Park, W., Hasan, N. A., Bethke, I., Fransner, F., Karwat, A., Subrahmanian, A. R., and Franzke, C.

L. E.: Robust estimates of Earth system predictability of the first kind using the CESM2 multiyear prediction system (CESM2-MP), Bulletin of the American Meteorological Society, 106, 2087–2109, https://doi.org/10.1175/BAMS-D-24-0329.1, 2025.

Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., Donat, M. G., England, M. H., Fyfe, J. C., Han, W., Kim, H., Kirtman, B. P., Kushnir, Y., Lovenduski, N. S., Mann, M. E., Merryfield, W. J., Nieves, V., Pegion, K., Rosenbloom, N., Sanchez, S. C., Scaife, A. A., Smith, D., Subramanian, A. C., Sun, L., Thompson, D., Ummenhofer, C. C., and Xie, S.-P.: Initialized Earth System prediction from subseasonal to decadal timescales, Nature Reviews Earth & Environment, 2(5), 340–357, https://doi.org/10.1038/s43017-021-00155-x, 2021.