Authors' Response to Reviews of

Improving Thermodynamic Nudging in the E3SM Atmosphere Model Version 2 (EAMv2): Strategy and Hindcast Skills on Weather Systems

Shixuan Zhang and Ruby Leung (on behalf of all authors)
Geoscientific Model Development (GMD), 10.5194/egusphere-2025-4277

RC: ***Reviewers' Comment***, AR: Authors' Response, ☐ Manuscript Text

We would like to thank the anonymous reviewer for their useful comments to help us improve the manuscript. Our responses to individual comments from the reviewer are included below.

1. Comments from the Anonymous Referee #1

The manuscript is overall well written, logically structured, and presents a useful and innovative nudging strategy with clear relevance to atmospheric modeling and constrained simulations. The methodological development is meaningful, and the results are generally convincing. I only have several minor comments and requests for clarification that I believe will further strengthen the manuscript.

RC: *Equation3. It seems that "1 Pa" in the equation may actually refer to 1 hPa. Please verify whether this is a typographical error.*

AR: Thank you for pointing this out. This was indeed a typographical error: the expression should read 1 hPa rather than 1 Pa. We have corrected this in the revised manuscript.

RC: *In addition, how sensitive are the results to the choice of P_0 ? It would be beneficial if the authors could provide guidance, or at least discuss strategies, for choosing P_0 , especially for readers who might apply the method in different models.*

AR: Thank you for the valuable suggestions. We agree with reviewer, and have added a more detailed discussions on the choice of P_0 in our revised manuscript (lines 119–136):

> We emphasize that P_0 and Z_b in Eq. (4) are, by design, tunable parameters. Sensitivity experiments using different combinations of P_0 and Z_b provide a general pathway for identifying optimal values for specific modeling configurations or scientific applications; however, such experiments can be computationally demanding, especially for long-term (10 years or longer) global simulations with modeling systems such as E3SM. In this study, we avoid such extensive tuning by setting Z_b equal to the diagnosed PBL height, following guidance from previous studies (e.g., Lo et al., 2008; Vincent and Hahmann, 2015). For P_0, a practical rule of thumb is to choose a threshold such that nudging is restricted to the altitude range where high-quality, observationally constrained reanalysis data exist, thereby avoiding regions that require extrapolation. Consequently, as summarized in Table 2, we specify different values of P_0 for different nudging variables, following the guidance from the nudging implementation in the GFDL FV3 model (NOAA GFDL, 2020). These values are suitable for our application because, for ERA5, extrapolation is triggered when the model levels extend above roughly 1 hPa, which is near the top of the reanalysis domain. Therefore, the P_0 values listed in Table 2 are all

> chosen to exceed 1 hPa, with additional expert judgment applied to select threshold values for each model state variable that are large enough to ensure that the nudging strength is smoothly reduced in layers that would otherwise depend on extrapolated reanalysis fields. The primary goal of this study is to explore the benefits of the modified weighting function in Eq. (4) for nudged E3SM simulations, particularly in the context of long-term model analysis and evaluation applications. Although a full set of sensitivity experiments to optimize these tuning parameters would be useful, it is beyond the scope of this study. Furthermore, as discussed in Section 4, these specific choices of P_0 and Z_b provide a reasonable configuration for the proposed scheme, and their implementation in E3SM improves the robustness and overall fidelity of simulations nudged to reanalysis.

RC: *Line 133-135. The nudging tendency term is calculated and applied at different locations in the model. Could the authors explain the reasoning behind this choice? Is there a specific numerical or physical advantage to computing and applying the tendency at different places?*

AR: Thank you for this thoughtful question. The rationale behind this design has been examined in detail in our previous work on nudging in the E3SM model series (Zhang et al., 2022). In brief, computing and applying the nudging tendency at different stages of the time-integration loop improves numerical consistency with the native model physics and helps avoid numerical artifacts in nudged simulations. Specifically, we found in Zhang et al. (2022) that when the model was nudged toward EAMv1's own meteorology, the nudged simulation still exhibited non-negligible deviations from the corresponding free-running simulation, particularly over subtropical marine stratocumulus and trade-cumulus regions. These deviations were traced to where in the integration sequence the nudging tendency was calculated and applied. By computing the tendency at a point in the loop that is more consistent with how prognostic variables are updated, and applying it later in coordination with the physical parameterizations, the nudged solution becomes more dynamically consistent with the free-running baseline, and the spurious deviations are substantially reduced.

To address the reviewer's question, we have revised the manuscript text to briefly explain this rationale and to refer the readers to Zhang et al. (2022) for full methodological details (lines 156–158 in revised manuscript):

> This implementation strategy follows Zhang et al. (2022), who showed that computing the nudging tendency and applying it at these specific locations in the time-integration loop improves numerical consistency with the native model physics and reduces spurious deviations in the nudged simulations.

Reference:

- Zhang, S., Zhang, K., Wan, H., and Sun, J.: Further improvement and evaluation of nudging in the E3SM Atmosphere Model version 1 (EAMv1): simulations of the mean climate, weather events, and anthropogenic aerosol effects, *Geosci. Model Dev.*, 15, 6787–6816, 2022. `https://doi.org/10.5194/gmd-15-6787-2022`

RC: *Line 304-307. The speculation here is interesting. Could the authors be more specific about what type of microphysics tuning that is applied in the free-running configuration but cannot be to the nudged simulations? Relatedly, this raises a broader question about other sub-grid tuning (e.g., CAPE relaxation time in the convection scheme). The authors may wish to expand their discussion on how such tuning parameters might interact with or influence nudging behavior.*

AR: Thank you for these valuable questions and comments. Our intention in this discussion was to emphasize that certain physics parameterizations, such as microphysics, often require parameter tuning that is implicitly tied to the model's mean-state climate. Because nudging modifies the mean state, the optimal tuning may no longer be the same as in a free-running configuration. For example, in the bulk microphysics parameterization used in E3SM, the autoconversion rate depends on the mean cloud liquid water content (q_c) and cloud droplet number concentration (N_c) (e.g., Ma et al., 2022; Song et al., 2024). Tuning parameters associated with the autoconversion scheme are therefore implicitly tuned to the model's climatological distributions of q_c and N_c, which in turn depend on the mean total water content of the atmospheric column. In our experiments, the column-integrated total precipitable water (TPW) in the free-running simulation (CTRL; Fig. 6a1) differs substantially from that in the nudged simulations (e.g., NDGUVTQ_SRF1; Fig. 6a4), because the latter are strongly constrained by humidity nudging. This alteration of the background moisture state plausibly shifts the cloud-water and droplet-number regimes sampled by the microphysics, so tuning derived for the free-running configuration may not remain optimal under nudging.

We also agree that this raises a broader question regarding other subgrid tuning choices (e.g., the CAPE relaxation time in the convection scheme). In principle, one could consider re-tuning certain parameters in the context of nudged simulations for climate models such as E3SM, since nudging tends to align the large-scale fields (U, V, T, Q) more closely with the observational reanalysis, which represents our best estimate of the real atmosphere. However, a comprehensive assessment of such interactions is beyond the scope of the present study, whose primary focus is to present and evaluate the nudging strategy itself. To address the reviewer's concerns, we extended our discussion in lines 304-307 in the revised manuscript (lines 360–379):

> Furthermore, the biases in large-scale moisture convergence (LSMC; Figure C3) in NDG-UVT, NDG-UVTQ, and NDG-UVTQ_SRF1 do not show corresponding changes in total precipitable water (TPW; Figs. 6a2–a4) or precipitation (PRECT; Figs. 6c2–c4). This lack of a coherent response between LSMC and TPW suggests that the sensitivity of precipitation to nudging strategies does not primarily arise from changes in large-scale moisture transport, but rather from adjustments in subgrid physics processes (e.g., microphysics) that may be triggered by the nudging configurations. This also helps explain the relatively weak improvement in precipitation biases in NDG-UVTQ_SRF1 (Fig. 6c4) compared with the other nudging simulations (Figs. 6c1–c3), because precipitation is determined not only by realistic large-scale moisture convergence and surface conditions, but also by how nudging interacts with the physics parameterizations. This interpretation is supported by the strong sensitivity of cloud properties and by systematic changes in both shortwave and longwave cloud radiative effects across the simulations (Figure C4). As shown in prior parameterization-tuning studies within the E3SM modeling system (e.g., Ma et al., 2022; Song et al., 2025), several physics parameterizations (e.g., microphysics and deep convection) were tuned in the biased mean climate of the free-running simulations (i.e., CTRL), whereas nudging shifts the large-scale thermodynamic state toward the reanalysis constraints. As a consequence, parameterizations that perform well in free-running mode may exhibit unintended sensitivities when nudging is applied—especially when temperature and humidity fields are constrained—because the background cloud-water and instability environments sampled by microphysics and deep convection are altered. Therefore, the increased biases in cloud-related quantities from the nudged simulations relative to CLIM, as shown in Figures 6 and C3, should not necessarily be viewed as degradations in hindcast skill. Instead, they likely reflect that some degree of retuning may be required for the physics parameterizations to achieve optimal performance under nudged configurations. A systematic evaluation of how these tuning parameters behave under nudged configurations would be valuable, but such an analysis lies beyond the scope of the present study.

References:

- Ma, P.-L., Harrop, B. E., Larson, V. E., Neale, R. B., Gettelman, A., Morrison, H., Wang, H., Zhang, K., Klein, S. A., Zelinka, M. D., Zhang, Y., Qian, Y., Yoon, J.-H., Jones, C. R., Huang, M., Tai, S.-L., Singh, B., Bogenschutz, P. A., Zheng, X., Lin, W., Quaas, J., Chepfer, H., Brunke, M. A., Zeng, X., Mülmenstädt, J., Hagos, S., Zhang, Z., Song, H., Liu, X., Pritchard, M. S., Wan, H., Wang, J., Tang, Q., Caldwell, P. M., Fan, J., Berg, L. K., Fast, J. D., Taylor, M. A., Golaz, J.-C., Xie, S., Rasch, P. J., and Leung, L. R.: Better calibration of cloud parameterizations and subgrid effects increases the fidelity of the E3SM Atmosphere Model version 1, Geosci. Model Dev., 15, 2881–2916, https://doi.org/10.5194/gmd-15-2881-2022, 2022.

- Song, X., Zhang, G. J., Terai, C., & Xie, S. (2025). Enhanced convective microphysics scheme and its impacts on mean climate in E3SM. Journal of Advances in Modeling Earth Systems, 17, e2024MS004656. https://doi.org/10.1029/2024MS004656.

RC: *Figure 1. For DNDG-UVQ and DNDG-UVTQ, the reported PCC values (0.79 and 0.82) are already quite high. It may help to comment on whether such high correlations are expected or what they imply about the baseline model behavior.*

AR: Thank you for the helpful comment. The relatively high PCC values in DNDG-UVQ and DNDG-UVTQ are expected, because PCC mainly reflects the spatial and temporal structure of precipitation systems. Humidity nudging improves the large-scale moisture distribution, which tends to enhance the realism of the precipitation structure even when biases remain in precipitation intensity. Following the reviewer's comment, we have extended the discussions related to Figure 1 in our revised manuscript (lines 220–229):

> In contrast, substituting temperature with humidity (DNDG-UVQ; panel e) results in a noticeable suppression of precipitation intensity, despite improved spatial correlation. This degradation becomes more pronounced when both temperature and humidity are nudged alongside winds (DNDG-UVTQ; panel f), revealing a key trade-off in the traditional nudging formulation: nudging Q improves the location and timing of precipitation (higher PCC) but substantially worsens its magnitude relative to observations (higher RMSE). As shown in previous studies (Sun et al., 2019), in addition to the realistic large-scale circulation provided by wind (U,V) nudging, further nudging of humidity (Q) improves the large-scale moisture distribution and leads to a more realistic spatial organization of precipitation systems, thereby increasing PCC. At the same time, those studies also demonstrated that improved pattern correlation does not necessarily guarantee a more realistic precipitation intensity, likely because humidity nudging can introduce unrealistic feedbacks between the large-scale moisture field and other model components (e.g., convection and microphysics) through changes to the condensational heating and vertical motion.

RC: *Figure 2. What does the "STRESS MAG" stand for?Moreover, it does not appear to be listed in Table B1. please clarify or include the relevant information.*

AR: Thank you for pointing this out. "STRESS_MAG" refers to the magnitude of the surface wind stress, computed as $\sqrt{\mathrm{TAUX}^2 + \mathrm{TAUY}^2}$, where TAUX and TAUY are the zonal and meridional surface wind stress components, respectively. The definitions of the diagnostic variables used for the metrics in Figure 2 are provided in Table B1 (Appendix B). Following the reviewer's suggestion, we have revised Table B1 to explicitly include the definition of STRESS_MAG for clarity:

> Table B1. List of physical quantities used for diagnostics in Fig. 4 and Fig. C2. The fifth column lists the reference datasets used to derive the error metrics, primarily sourced from ERA5 reanalysis (Hersbach et al., 2020), GPCP (Version 2.3; Adler et al., 2003), and CERES_EBAF (Version 4.1; Kato

et al., 2018). Note that all diagnostic variables are directly saved in the E3SM model output, except that STRESS_MAG is computed as $\sqrt{\mathrm{TAUX}^2 + \mathrm{TAUY}^2}$, where TAUX and TAUY are the zonal and meridional surface wind stress components from the model output, respectively.

RC: *Figure 5. Why is the analysis limited to December 1, 2010 to November 30, 2011? A short justification would help the reader understand this choice.*

AR: Thank you for the question. Figure 5 is intended to provide an illustrative example of the daily evolution of precipitation systems. We selected a single representative year (December 2010–November 2011) to keep the figure readable and to make the differences in precipitation structures easier to visualize. This example is complementary to Figure 4, which already presents a comprehensive 10-year evaluation of global precipitation RMSE and pattern correlation across seasons. Similar conclusions were obtained when examining other years, as the nudged simulations were run continuously from 2007 to 2017 (Section 2). We have added a brief clarification before the discussion to avoid confusion in our revised manuscript (lines 325–328):

> As a complement to the 10-year global statistics presented in Figure 4, Figure 5 further presents Hovmöller diagrams of the daily precipitation rate (PRECT; units: mm day−1) for a representative evaluation year (December 2010–November 2011) over selected tropical (panels a1–e1) and midlatitude regions (panels a2–e2), comparing the free-running and nudged simulations listed in Table 1.

RC: *Figure 6. The NDG-UVQT_SRF1 experiment does not seem to substantially improve PRECT. Could the authors explain why this might be the case?*

AR: Thank you for the good question. The discussion in Section 4.2 (paragraphs 3–4 following paragraph 2) was intended to explain why NDG-UVQT and NDG-UVQT_SRF1 do not yield large improvements in PRECT and other surface quantities shown in Figure 6. However, we recognize that the original wording may not have made this sufficiently clear, which likely triggered the reviewer's question. We have revised the manuscript with additional clarification and explanations (lines 360–379):

> Furthermore, the biases in large-scale moisture convergence (LSMC; Figure C3) in NDG-UVT, NDG-UVTQ, and NDG-UVTQ_SRF1 do not show corresponding changes in total precipitable water (TPW; Figs. 6a2–a4) or precipitation (PRECT; Figs. 6c2–c4). This lack of a coherent response between LSMC and TPW suggests that the sensitivity of precipitation to nudging strategies does not primarily arise from changes in large-scale moisture transport, but rather from adjustments in subgrid physics processes (e.g., microphysics) that may be triggered by the nudging configurations. This also helps explain the relatively weak improvement in precipitation biases in NDG-UVTQ_SRF1 (Fig. 6c4) compared with the other nudging simulations (Figs. 6c1–c3), because precipitation is determined not only by realistic large-scale moisture convergence and surface conditions, but also by how nudging interacts with the physics parameterizations. This interpretation is supported by the strong sensitivity of cloud properties and by systematic changes in both shortwave and longwave cloud radiative effects across the simulations (Figure C4). As shown in prior parameterization-tuning studies within the E3SM modeling system (e.g., Ma et al., 2022; Song et al., 2025), several physics parameterizations (e.g., microphysics and deep convection) were tuned in the biased mean climate of the free-running simulations (i.e., CTRL), whereas nudging shifts the large-scale thermodynamic state toward the reanalysis constraints. As a consequence, parameterizations that perform well in free-running mode may exhibit unintended sensitivities when nudging is applied—especially when temperature and humidity fields are constrained—because the background cloud-water and instability environments sampled by microphysics and deep convection

are altered. Therefore, the increased biases in cloud-related quantities from the nudged simulations relative to CLIM, as shown in Figures 6 and C3, should not necessarily be viewed as degradations in hindcast skill. Instead, they likely reflect that some degree of retuning may be required for the physics parameterizations to achieve optimal performance under nudged configurations. A systematic evaluation of how these tuning parameters behave under nudged configurations would be valuable, but such an analysis lies beyond the scope of the present study.

We hope that these revisions make clearer why NDG-UVQT_SRF1 does not produce substantially larger improvements in PRECT and related surface quantities.

RC: *Figure 10 seems to indicate that nudging temperature and humidity does not significantly change the result. Could the authors elaborate on why the impact is relatively small here?*

AR: Thank you for the question. Figure 10 shows the seasonal cycle of extratropical cyclone (ETC) track density derived from sea-level pressure (PSL) fields during 2008–2017. The tracking algorithm identifies cyclone centers primarily from PSL minima, meaning that the diagnosed ETC statistics are most sensitive to the large-scale dynamical fields (e.g., vorticity, baroclinicity, and surface pressure gradients). These quantities are already strongly constrained once the horizontal wind components (U, V) are nudged, because wind nudging effectively constrains the synoptic-scale circulation patterns and storm-track variability. By contrast, nudging temperature and humidity mainly modifies the thermodynamic structure of the atmosphere and has a less direct influence on PSL. As a result, adding T and Q nudging does not substantially alter the large-scale pressure patterns that determine cyclone detection, and therefore produces only modest additional changes in ETC track density. In other words, cyclone track statistics in our configuration are primarily circulation-controlled, and the circulation is already well captured once wind nudging is applied.

In response to the reviewer's comment, we have made revisions with additional clarifications in the revised manuscript (lines 461–474):

> Further analysis of two additional high-impact weather phenomena, namely extratropical cyclones (ETCs) and atmospheric rivers (ARs), reveals that the benefits of the proposed nudging strategy extend beyond tropical cyclones. Figure 10 presents the seasonal cycle of ETC occurrence frequency and the subset of intense ETCs (defined as those with central pressure below 990 hPa), derived from TempestExtremes tracking of 6-hourly EAMv2 output from the experiments listed in Table 1. The results are averaged over the Northern (panels a–b) and Southern (panels c–d) Hemispheres, respectively, as defined according to the climatological distribution of extratropical cyclones (see Fig. C6). Relative to the free-running simulation (CLIM), wind nudging alone (NDG-UV) already brings PSL fields and storm-track characteristics closer to ERA5, reflecting the strong control of large-scale circulation on ETC statistics. However, further nudging of temperature and humidity (NDG-UVT and NDG-UVTQ) still produces additional—though generally modest—improvements in ETC statistics, especially for intense ETCs. Specifically, NDG-UVT and NDG-UVTQ reduce the overestimation of storm frequency during boreal summer (JJA) and correct the underestimation during boreal winter (DJF) in the Northern Hemisphere, while also partially mitigating the underrepresentation of intense ETCs in the Southern Hemisphere. These results highlight the influence of large-scale thermodynamic fields on ETC climatology and support the utility of full-field nudging for improving storm-track variability and extremes in low-resolution models.