

Comments on “**Iterative run-time bias corrections in an atmospheric GCM (LMDZ v6.3)**”
by Gerhard Krinner, Aude Champouillon, Juliette Blanchet, and Frédérique Chéruy

This manuscript proposes and evaluates a variant of the traditional reanalysis nudging (Newtonian relaxation) approach, which is widely used in general circulation models to improve the representation of the real atmosphere as estimated by weather reanalysis products. Unlike classical nudging, which applies time-varying relaxation toward the reanalysis during model integration, the study focuses on a form of *climatological nudging–based bias correction*, in which nudging tendencies diagnosed from a prior nudged simulation are saved and subsequently applied as additional correction terms in free-running model simulations. The main innovation of the study lies in the iterative application of this correction strategy: the corrected model wind states (U and V components) are repeatedly nudged toward the reanalysis to derive accumulated correction tendencies. The authors find that this recursive nudging correction improves model performance in terms of agreement with the reanalysis targets, with optimal performance achieved after two iterations, while additional iterations lead to a degradation of performance. **While the investigation represents an interesting extension of conventional nudging implementations targeting on correcting long-term systematic errors in a dynamical atmospheric model, several aspects of the methodology and its interpretation would benefit from further clarification and discussion. I recommend a Major revision for current version of manuscript, with detailed comments listed below:**

1. The introduction in its current form does not sufficiently articulate the motivation for the proposed approach or clearly explain its potential impacts on model performance and predictive skill. For example, the study by **Guldberg et al. (2005)**, which is cited in this manuscript, along with the related literature therein, provides an important foundation on empirical model correction and the reduction of long-term systematic errors. A more explicit discussion of this prior work in the introduction would help readers better understand the scientific context, value, and fundamental advances underlying the exploration presented in this study. In addition, I would emphasize that it is important to clearly distinguish the target of the bias-correction problem addressed in this study—namely, the correction of long-term systematic errors—from the weather-scale bias correction typically associated with classical nudging. Clarifying this problem setup in the introduction would help readers better understand the intended scope of the proposed method and avoid potential confusion regarding its relationship to standard nudging approaches. Also, as will be mentioned in comments below, some existing similar studies should be mentioned as an background for the introduction of this study.
2. As described in Section 2.1, the nudging-based bias-correction framework in this study consists of two key steps: the construction of *cyclostationary climatological nudging increments* using a classical nudging procedure that constrains the model toward a reanalysis product, and the subsequent application of these increments, either directly or recursively, to free-running model simulations to reduce long-term systematic biases in the simulated climate. Although the primary focus of this study is on the second step, the methodology used to generate the nudging increments in the first step is essential in determining the behavior of the recursive nudging approach, including the number of iterations required to achieve optimal performance. In particular, the choice of the nudging relaxation timescale τ directly affects the strength of the diagnosed increments, raising the question of whether stronger nudging (e.g., $\tau = 6\text{h}$ instead of $\tau = 24\text{h}$) would require fewer iterations to reach performance comparable to the two-iteration case shown here, while weaker nudging (e.g., $\tau = 48\text{h}$) might require more iterations to achieve optimal performance. More generally, as discussed in many previous studies (e.g., Zhang et al., 2014; Sun et al., 2019; Zhang et al., 2023), a key challenge in nudging formulations such as Eq. (2) is determining an optimal choice of τ that is strong enough to reduce model biases while remaining weak enough to avoid undue interference with the model’s intrinsic physics and dynamics. The appropriate choice often depending on the scientific purpose of the nudged simulations. Similarly, for the recursive approach proposed in this study, if the number of iterations required for optimal performance is highly sensitive to the nudging configuration, this may pose a challenge for the

generalization of the method to other applications and model and raise questioning on the value of the method to the modeling community proposed by this study. From a machine-learning perspective, the iterative nudging framework can be viewed as a form of repeated residual correction, in which correction tendencies are successively accumulated to reduce systematic error. In this context, the sensitivity of performance to parameters such as the nudging timescale and iteration number is closely related to issues of stability, regularization, and generalization, and therefore merits further discussion. Overall, I think that the current version of the manuscript would benefit from additional discussion addressing the points raised above, that an expanded discussion in this regard would help strengthen the value and impact of the proposed method in this study.

References:

- Zhang, K., Wan, H., Liu, X., Ghan, S. J., Kooperman, G. J., Ma, P.-L., Rasch, P. J., Neubauer, D., and Lohmann, U.: Technical Note: On the use of nudging for aerosol–climate model intercomparison studies, *Atmos. Chem. Phys.*, 14, 8631–8645, <https://doi.org/10.5194/acp-14-8631-2014>, 2014.
 - Sun, J., Zhang, K., Wan, H., Ma, P.-L., Tang, Q., & Zhang, S. (2019). Impact of nudging strategy on the climate representativeness and hindcast skill of constrained EAMv1 simulations. *Journal of Advances in Modeling Earth Systems*, 11, 3911–3933. <https://doi.org/10.1029/2019MS001831>
 - 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, <https://doi.org/10.5194/gmd-15-6787-2022>, 2022.
3. Closely related to the above comments, the approach discussed in this study appears to be conceptually equivalent to the **tendency bias correction (TBC)** framework described in Chang et al. (2019). It would be helpful for the authors to explicitly acknowledge this connection and clarify similarities and differences between the two approaches, particularly in terms of methodology, assumptions, and intended applications. The primary distinction between TBC and the approach proposed in this manuscript appears to lie in the method used to estimate the climatological tendency-bias correction terms. In TBC, the correction tendencies defined as **climatological 6-hourly mean differences** between the model and observations (reanalysis), such definition by itself is general for any model systems and in absent of the dependence on the empirical nudging relaxation time scale τ . To me, the approach presented in this study,

References:

Chang, Y., S. D. Schubert, R. D. Koster, A. M. Molod, and H. Wang, 2019: Tendency Bias Correction in Coupled and Uncoupled Global Climate Models with a Focus on Impacts over North America. *J. Climate*, 32, 639–661, <https://doi.org/10.1175/JCLI-D-18-0598.1>.

Another factor that may influence the performance of the correction is the length of the correction simulations. By definition, the nudging-based tendency bias correction targets slowly evolving, long-term systematic errors, and therefore may be sensitive to the length of the period used to estimate the correction terms. For example, a 40-year continuous correction simulation may outperform a 20-year simulation, as a longer integration provides a more robust estimate of the climatological tendency biases and allows the cumulative effects of the correction to be more fully realized. A brief discussion of this sensitivity would help clarify the robustness of the proposed correction approach. Some analysis and

discussions on this regard can be useful to enhance the value of the study. For example, the authors could compare the mean bias reductions for the first 10 years and the last 10 years of the free-running correction period (2001–2020) to assess whether the correction effectiveness show differences for two 10-year periods. Finally, the climatological mean bias-correction terms are estimated over the period 1981–2000, but are applied to correct the free-running model simulations during 2001–2020. However, the climatological states in both ERA5 and the model may differ between these two periods. How might such differences affect the effectiveness of the iterative bias-correction approach and the determination of the optimal number of iterations?

4. The design of the iterative run-time bias correction is somewhat confusing. Based on Section 2.1 and Table 1, my understanding is the following:
 - a. The N_x groups always employ Eq. (4), which corresponds to classical nudging using ERA5 during 1981–2001, with an additional climatological mean bias-correction term defined as the average of the nudging tendencies over 1981–2001 from the previous iteration (N_{x-1}).
 - b. The C_x groups always employ Eq. (5), in which a climatological mean bias-correction term (independent of the model equation and estimated from the N_x simulations) is applied directly to the model equations.

Is this interpretation correct? In addition, it is unclear whether the N_x and C_x experiments are performed as independent integrations (each initialized from the same initial condition and differing only in the applied correction term) or whether subsequent iterations (e.g., N_2) are continued from the end state of the previous iteration (N_1). This distinction is important because it affects how the iterative procedure should be interpreted. I suggest that the authors provide a clearer and more explicit explanation of the experimental design and the iterative bias-correction framework to help reader better understand. Moreover, it is hard for me to understand what the purpose of the N_1 , N_2 and N_3 experiments and the resulting nudging tendency. The N_0 was the classical nudging towards reanalysis (i.e. ERA5), which was used to estimate the climatological bias correction tendency terms C_0 to be used and applied to the free running simulations. Here, C_0 can be interpreted as the **mean tendency correction required to offset the model's systematic drift relative to reanalysis**. However, for simulations starting from N_1 , C_0 is added in addition to the ERA5 nudging terms. In this configuration, it becomes unclear how the subsequently diagnosed correction terms (e.g., C_1 derived from N_1) should be interpreted. As noted above, C_0 is diagnosed from the nudging tendency term, $-\frac{1}{\tau}(X - X_{\text{ERA5}})$. The explicit dependence on the relaxation time scale τ implies that C_0 may vary with different choices of τ , and therefore may not represent an accurate or intrinsic estimate of the climatological mean model tendency bias. This may leave room for the subsequent iterations (N_1 , N_2 , etc.) to act toward a progressively improved estimate of the climatological bias-correction tendency through the iterative procedure. However, such convergence is not guaranteed, given the nonlinear and online interactions between the nudging terms and the model's physical and dynamical processes. Overall, the manuscript would benefit from a clearer and more systematic explanation of the theoretical basis and practical implementation of the iterative bias-correction framework, including its physical interpretation, potential deficiencies or limitations, and the key considerations underlying its implementation.

Other specific comments:

- Line 9: The phrase “However, while ...” sounds somewhat awkward. Please consider revising the sentence.
- Line 35: Consider replacing “more perfect” with “more effective.”

- Line 104: “root-mean-square error (RMSE)” is defined multiple times throughout the manuscript. In addition, the terms “root-mean-square error,” “RMSE,” and “root-mean-square error (RMSE)” are used inconsistently. Please consider using a consistent format throughout the manuscript.
- Section 3.1: Could you elaborate on how the results in Figure 1 should be interpreted in terms of the nudging procedure? In addition, you state that “Instead, this ratio has distinctive spatial structures that vary in space and time throughout the annual cycle (not shown), indicating that the iterative procedure is not identical to a simple uniform amplification of the initial correction.” What is the underlying reason for this behavior? Finally, I am confused by Figure 2, why the ratio of $C3/C0$ are all above 1.5, given that the magnitude in Figure 1d are all smaller than those in Figure 1a?
- Section 3: The authors present their discussion using only metrics focused on the U and V components. It would be valuable to include analyses of other fields, such as temperature, humidity, precipitation, and sea level pressure etc., which are not directly nudged but may still be significantly affected by the U and V nudging. This will be also valuable supports to your discussions in Section 4.
- Section 4: Some of the discussion in Section 4.4 could be made more meaningful by linking it more directly to the results of this study (e.g. section 4.1 and section 4.2), rather than relying on speculative or high-level arguments. I also suggest that the authors focus the discussion more closely on parameters and methodological choices related to the proposed approach, instead of emphasizing broader aspects such as model resolution (e.g. section 4.3). To me, broader considerations could be better briefly summarized in a short paragraph in the conclusion.