# Replies to comments by Reviewer 3

Gerhard Krinner, Aude Champouillon, Juliette Blanchet and Frédérique Chéruy

April 2, 2026

We sincerely thank all three reviewers for their thoughtful comments and suggestions which we have taken into account in the revised version of this article.

## Overall comments

**Main comment #1:** My main difficulty with the manuscript is that most of the metrics presented and parts of the methodologies are loosely defined textually, leaving significant room for uncertainty in interpretation.

**Reply :** In the subsequent minor comments, the reviewer points to specific issues where clarification of definitions and methodologies is required. We will follow the requests made there.

**Main comment #2:** Can you mention why you chose a nudging time scale of one day? This choice does seem reasonable. Still, one might expect qualitatively but not quantitatively similar results for a different choice of time scale, as dycores tend to fight nudging tendencies to different degrees depending on this time scale (e.g. Kruse et al. 2022). If you chose, say, a nudging time scale of half a day, do you think your results on variability (short-term in particular), would still hold?

**Reply :** This point has also been raised in similar terms by the two other reviewers. We are currently finalizing work that compares the different ERBC approaches ("classical", CAB-COR, iterative, and an implementation of conditional ERBC) that also comprises a detailed evaluation of the effects of nudging time constants in the "classical" and iterative approach with the LMDZ model. In addition, the reviewers have indicated references to relevant published studies that, together with our experience, provide a rather complete picture of the question, including the finding that optimal nudging timescales can be strongly model- and application-dependent. In response to these comments, we have substantially expanded section 3.1 and added a new section 4.2 in the discussion, specifically addressing the question of nudging timescales. Concerning the question about shorter nudging timescales than one day, we have not tested such values here, but an older paper (Krinner et al., 2019) uses a time constant of 6 hours, and in that study, no particularly detrimental effect on short-term variability was seen. However, we also note that a recent study by Scinocca and Kharin (2024) showed strong degradation of ERBC results for small time constants under certain conditions, which lead us to refrain from simulations with small nudging time constants $\tau$. We mention this now in section 4.2:

> However, in a study of the effect of "classical" ERBCs on the simulation of the Antarctic climate (Krinner et al., 2019), a nudging timescale of $\tau = 6h$ was used, substantially less than that used here ($\tau$ = 1 d), and that in that study, improvements of the simulated climate were shown across all timescales, including high-frequency variability. We note in this respect that the iterative procedure used here can be seen as a method that leads to a stronger effective nudging, as shown by the fact that the amplitude of the combined correction terms $G_{0+...+i}$ for higher numbers $i$ of iterations increases (see Figure 2). Additional systematic tests of the effects of smaller time constants $\tau$ ¡ 1 d in LMDZ are planned for future studies.

**Main comment #3:** Just a comment, not a request: It'd be very interesting to repeat this analysis using two different models, or versions of models, that have different mean states (i.e. are "biased" relative to each other). Much stronger statistics could be achieved. Conclusions on the influence of such a method on short-term, smaller-scale variability might be much stronger.

**Reply :** We agree. We mention this need in the conclusions of the article, as a consequence of model-dependency of such results, as an outlook:

> More generally, because ERBC implementation choices and thus the corresponding results are at least partly model-dependent, it would be interesting to see the effect of iterative ERBC in other models, possibly in the framework of a coordinated multi-model intercomparison of various ERBC methods.

## Minor comments

**Minor comment #1:** Line 52: It'd be very helpful to the readers if you define "cyclostationary time average". Is $G_0$ a function of x,y,z,t then?

**Reply :** Yes. We indicate this clearly now after the equation that first defines $G$:

> ...where $G_0$ is a function of space (latitude, longitude and vertical level) and time (day of the year).

**Minor comment #2:** Line 92-94: There's no cancellation between $G_0$, $G_1$, ...? Not too surprising, this but cannot be inferred from Fig. 1 since the absolute value is presented. If you were to further iterate, would you expect the sum of $G_i$ to converge?

**Reply :** The reviewer is right. We analyze this in more detail in section 3.1 in the revised version:

> Do these successive correction terms converge towards some "final" correction term? Figure 1d shows that the nudging increments in the third iterated nudging step $N_3$ do not vanish, although they are substantially weaker than in $N_0$ (Figure 1a). The global mean of the absolute zonal wind nudging tendencies (in January, to be consistent with Figure 1) is 0.50 m/s/day for $N_0$, 0.35 m/s/day for $N_1$, 0.29 m/s/day for $N_2$, and 0.26 m/s/day for $N_3$. This means that the intensity of the remaining nudging tendencies decreases at higher iterations, but convergence towards potentially vanishing final nudging tendencies is still far away after 3 iterations. The combined correction terms arising from the sum of these absolute zonal wind nudging tendencies have global mean values of 0.50 m/s/day for $G_0$ (because $G_0$ is identical to the mean nudging tendencies of $N_0$), 0.83 m/s/day for $G_{0+1}$, 1.09 m/s/day for $G_{0+1+2}$, and 1.30 m/s/day for $G_{0+\ldots+3}$, and are thus somewhat lower than the corresponding sums of the global mean of the absolute zonal wind nudging tendencies (which would be 0.5, 0.85, 1.14 and 1.40 m/s/day, respectively), indicating that some local-scale compensation occurs between different iterations, as already shown by Figure 2.

**Minor comment #3:** Fig. 2: are you plotting $C_3/C_0$ (i.e. total model tendencies?) or $G_3/G_0$? I assume the latter, but this is not completely clear from the text or figure caption. An inline equation would be very helpful.

**Reply :** The reviewer is right to state that we have not been precise enough here. We modified the figure label to clearly indicate that the figure shows the ratio between the sum of $G_i$ (i=0,...,3), which is used in $C_3$, and $G_0$ (used in $C_0$). We also specify this in the caption of the figure now:

> Ratio of the zonal mean absolute zonal wind bias correction terms (averages for January 1981 to 2000) between those used in $C_3$ (bias correction terms $G_{0+\ldots+3}$) and those used in $C_0$ (bias correction terms $G_0$). ...)

**Minor comment #4:** Fig. 3: Are these biases in the corrected, but free-running runs (i.e. 2001-2020)?

**Reply :** Yes, these are the biases of the corrected, "free-running" simulations. We refrain from using "free-running" in the revised version because one reviewer found this expression confusing.

**Minor comment #5:** Fig. 4: What is plotted is not stated in the figure caption nor the text! Is this RMSE 2001-2020 time averaged u and v between the bias-corrected, free-running runs and ERA5? That is, the RMSE of a spatial bias?

**Reply :** We clarify this in the text:

> Figure 3 displays the 1981–2000 annual mean bias of the zonal wind component at 500 hPa for the different simulations. Figure 4 displays, in each of its two panels, the ratio of the 1981–2000 root mean square error (RMSE) in each of the corrected simulations $C_i$ ($i = 0, \dots, 3$) and the RMSE of the uncorrected simulation M : RMSE($C_i$)/RMSE(M), for the annual mean of the zonal (a) and merdidional (b) wind speed, where the RMSE is calculate with respect to ERA5.

**Minor comment #6:** Line 117: Do you mean middle and upper troposphere and not middle-atmosphere (i.e. stratosphere/mesosphere)?

**Reply :** This was not clear indeed. We now write:

> In this case the iterative run-time bias correction procedure leads to further improvement in most parts of atmosphere above about 900 hPa [...]

**Minor comment #7:** Table 2: Again, some quantitative definition of these metrics would be helpful. I think I understand your metrics, but ideally there wouldn't be uncertainty in interpretation.

**Reply :** We revised the text to be more explicit:

> The dominant patterns of interannual variability of monthly circulation structures are indeed more realistically depicted in the corrected simulation $C_0$ than in the uncorrected control simulations M, and even more so in simulations with iterated run-time bias corrections ($C_{1-3}$). This can be seen in Table 2 which displays the seasonal squared spatial Pearson correlation coefficient $r^2$ between simulated (LMDZ) and "observed" (ERA5) dominant modes of monthly 500 hPa geopotential anomalies, as identified by principal component analysis for the different corrected model runs $C_i$, relative to the corresponding squared spatial correlation coefficient for the uncorrected control simulations M: $R = r^2_{C_i}/r^2_M$.

And we also made the table caption more explicit on the definitions:

> Squared spatial correlation coefficient $r^2$ between the corrected LMDZ runs $C_i$ ($i$=0,...3) and ERA5, for the first EOF of monthly extratropical (30-90° latitude) variability of the 500 hPa geopotential height ($\phi_{500\text{hPa}}$) for selected seasons and hemispheres, 2001-2020, relative to the corresponding squared correlation coefficient $r^2_M$ of the uncorrected simulation M: $R = r^2_{C_i}/r^2_M$.

We also inverted the order of the table columns (and adapted the caption) such as to have the main results in the first columns, while the last column serves as a reference. We hope that this makes the table much easier to understand.

**Minor comment #8:** Table 3: Again, it's very difficult to be sure I understand all of the individual variability metrics. Could you please define these metrics (not ratios) quantitatively?

**Reply :** We defined the metric more properly in the text. For the sea-level pressure variability, we write:

> The winter short-term (2.5–8 day) mid-latitude (40-60°N and 40-60°S) sea-level pressure variability, diagnosed from the temporal standard deviation of sea-level pressure from its seasonal mean ($\sigma_{2.5-8\text{d}}$, using a bandpass filtering in the frequency domain, and limited to grid points with surface height below 1000 m) ...

The definition of the blocking frequency score was indeed missing. We now write:

> Blocking anticyclones are a significant atmospheric phenomenon, although there is not one standard accepted definition (Lupo, 2021). Here we follow the widely used definition by Davini et al. (2012): Blocking events are diagnosed at a given point in time and space based on the reversal of the meridional gradient of geopotential height measured at 500 hPa, extending over at least 15° of continuous longitude, and persisting within a 5° latitude × 10° longitude box centered on a given grid point for at least 5 days. We then calculate the annual mean frequency of these events for each grid point, and calculate the spatial RMSE $E$ of these frequencies compared to ERA5. We then calculate, for each corrected simulation, the ratio between $E_{C_i}$ for that simulation and $E_M$ of the uncorrected simulation, yielding a score $f_B = E_{C_i}/E_M$.
>
> While the iterated procedure seems to improve the model score for the 30-75°N annual and spatial mean blocking frequency, (i.e., $f_B$ (30-75°N) $< 1$), with best results obtained for 2 and 3 iterations, the bias-corrected simulations show degraded performance compared to the uncorrected control simulations (i.e., $f_B$ (30-75°S) $> 1$) in the Southern Hemisphere (30-75°S). However, in the Southern Hemisphere, the iterated bias correction procedure leads to less degraded performance compared to a single bias correction in $C_0$.

# References


Davini, P., Cagnazzo, C., Gualdi, S., and Navarra, A.: Bidimensional diagnostics, variability, and trends of Northern Hemisphere blocking, Journal of Climate, 25, 6496–6509, 2012.

Krinner, G., Beaumet, J., Favier, V., Deque, M., and Brutel-Vuilmet, C.: Empirical run-time bias correction for Antarctic regional climate projections with a stretched-grid AGCM, Journal of Advances in Modeling Earth Systems, 11, 64–82, 2019.

Lupo, A. R.: Atmospheric blocking events: A review, Annals of the New York Academy of sciences, 1504, 5–24, 2021.

Scinocca, J. F. and Kharin, V. V.: Climatological adaptive bias correction of climate models, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 563, 2024.