https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Intercomparison of run-time bias correction methods in LMDZ v6.3

Aude Champouillon[1,2], Gerhard Krinner[1], and Juliette Blanchet[1]

[1]Univ. Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, 38000 Grenoble, France
[2]École nationale des ponts et chaussées, Institut Polytechnique de Paris, 77455 Marne-la-vallée, France

Correspondence: Aude Champouillon (aude.champouillon@univ-grenoble-alpes.fr)

Abstract. Despite progress in physical development and calibration, climate models still exhibit biases with respect to historical observations. As an alternative way to reduce them, run-time bias correction approaches have been developed, which consist in adding empirical tendency adjustment terms to the prognostic equations of some key physical variables. Although their ability to effectively reduce atmospheric circulation biases has been demonstrated, information is still missing regarding which method for estimating the adjustment terms is best suited for a given application. In this study, we implement a set of these methods in the atmospheric general circulation model LMDZ: nudging-based bias correction (the basis approach, a state-dependent version, and an iterative version), and the so-called climatological adaptive bias correction. Applying run-time bias correction on horizontal winds only, using these methods and varying some of their parameters, nine "bias-corrected versions" of the model are created. They are evaluated using aggregate scores of global mean errors in circulation, temperature, and precipitation, as well as mid-latitude atmospheric variability features. A more regional perspective is also adopted, and a large region covering Europe and the North-Atlantic serves as a case study. It is found that, when evaluated on global aggregate scores, some versions outperform others. We also show that this does not prejudge the outcome on mid-latitude atmospheric variability features or at regional scale. No strict recommendation can be made regarding the optimal methodological choice, and great caution is advised. The choice should be guided by the model user's needs and priorities.

1 Introduction

Climate models are primary tools to improve our understanding of the Earth's climate and to quantify future climate change under anthropogenic forcing from global to local scales. Constant efforts are made to improve their performance. However, they still exhibit biases in the climatological mean and the variability of key variables such as temperature, precipitation, and humidity, as well as in the representation of large-scale circulation features (Eyring et al., 2021; Fernandez-Granja et al., 2021). The amplitude of these biases sometimes reaches or exceeds the order of magnitude of the projected changes.

For many use cases, bias reduction proves beneficial if not necessary. As complementary tools to a posteriori bias adjustment methods, i.e., performed on simulation outputs (see Maraun, 2016; Dinh and Aires, 2023; Menapace et al., 2025, for reviews), run-time bias correction methods have been developed and experimented with in atmospheric general circulation models (AGCMs). In the run-time approach, the equations of the model's prognostic equations are altered by adding a corrective term, as originally described in a simplified AGCM by D'Andrea and Vautard (2000) or for seasonal forecasting by Guldberg et al. (2005), and later by Kharin and Scinocca (2012) and Chang et al. (2019). Run-time bias correction in AGCM

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

for climate simulations has since then been implemented with various ways of estimating the corrective term (Chang et al., 2019; Krinner et al., 2019, 2020; Chapman and Berner, 2024; Scinocca and Kharin, 2024). It typically reduces biases in the mean of targeted variables by 30 to 50% on a global scale.

In addition to the reduction of large-scale biases, some studies show the regional impacts of run-time bias correction in a global climate model using a stretched grid (Beaumet et al., 2021; Balhane et al., 2024) or not (Koskentausta et al., 2025). Scinocca et al. (2025) and Labonté et al. (2025) also show that using a bias-corrected global simulation as boundary conditions for a regional climate model (RCM) leads to substantial bias reduction at continental and regional scale in the RCM simulation.

Run-time bias correction arises as a way to reduce climate models' biases while avoiding some of the drawbacks of a posteriori bias correction, such as the potential resulting physical inconsistency between circulations and the corrected variables (Hall, 2014; Maraun et al., 2017). It also has the ability to modify the model's climate change responses, unlike typical a posteriori bias-adjustment which often are designed to preserve the model's climate change tendency. However, the use of all empirical bias-correction methods in projection simulations, run-time or a posteriori, relies on the assumption of stationarity of biases through climate change. Using a "perfect model experiment", Krinner and Flanner (2018) show that under strong climate change, large-scale climate model bias patterns reveals stationary to a high degree, giving confidence in correcting large-scale circulation errors in climate models.

Previous studies conducted detailed comparisons methods to estimate corrective terms. Estimation of corrective terms, also referred to as bias correction increments hereafter, is a two-step procedure: a "learning" simulation (also termed "adaptation" simulation) is used to derive bias correction terms to be used in a bias-corrected simulation. The learning simulation can typically be a simulation with nudging, i.e. relaxation, towards a reanalysis (as in Guldberg et al., 2005), while the bias correction terms are derived as a cyclostationnary mean of nudging tendencies. Scinocca and Kharin (2024), using he Canadian Center for Climate Modeling and Analysis' fifth generation atmospheric climate model (CanAM5.1), compares the "classical" approach based on nudging and an approach based on an iterative procedure using a reference climatology, as well as their sensitivity to hyperparameters. Chapman and Berner (2024), using the Finite Volume Community Atmosphere Model version 6 (CAM6-FV), compares the nudging-based approach to a more sophisticated assimilation method, as well as the effect of introducing stochastic terms.

Here, we propose an extension of these previous works and produce an intercomparison of a set of bias correction increments estimation's methods, although far from exhaustive, with the LMDZ6 AGCM (Hourdin et al., 2020a) at low resolution (96 (longitude) $\times$ 95 (latitude) horizontal grid points and 79 vertical levels). We develop a consistent framework for evaluation, ranging from mean aggregated variables to aspects of atmospheric variability. We also take a regional perspective and focus on the North Atlantic-Europe region.

The next section describes the different run-time bias correction methods compared in this paper, as well as the experimental set-up. Section 3 then presents the results focusing first on global means, on atmospheric variability features and on the effect of the correction at a continental scale. The implications of these results and their potential caveats are discussed in Section 4. The last section summarizes this work and suggests possible further working pathways.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

2 Methods

2.1 General principle of empirical run-time bias correction

Empirical run-time bias correction (ERBC) consists of altering the prognostic equations of selected variables by adding a corrective term:

$$\frac{\partial X}{\partial t} = F_{\mathrm{M}}(t) + G(t)$$

where X is a prognostic variable of the model, F_{M} describes its evolution equation in the model and G is the corrective term. Typically, G depends at least on time of the year and location within the model grid. Its estimation can be achieved by various means, all relying on a "learning simulation" also termed "adaptation simulation". It is built primarily to reduce biases in the mean annual cycle of a variable.

The following sections describe the different methods used in this study. Exact details of their implementation are given in Appendix A.

2.2 Nudging-based ERBC

2.2.1 Base methodology

Nudging refers to a Newtonian relaxation towards a reference. With X a prognostic variable, if X_{M} is the result of the time integration of $\frac{\partial X}{\partial t} = F_{\mathrm{M}}(t)$, the nudging operation consists of incrementing X_{M} by the relaxation term, such that the prognostic equation of X is now:

$$\frac{\partial X}{\partial t} = F_{\mathrm{M}}(t) + \frac{X_{\mathrm{R}}(t) - X_{\mathrm{M}}(t)}{\tau}$$

where X_R is the value of X in the reference dataset and τ is a time constant, describing the "strength" of the relaxation.

ERBC based on nudging makes use of a nudged simulation of a sufficient duration (typically multiple decades) and computes $G(t)$ as a cyclostationary time average of the relaxation terms' time series :

$$G_{\mathrm{N}}(t) := \overline{\frac{X_{\mathrm{R}}(t) - X_{\mathrm{M}}(t)}{\tau}}^{C}$$

where the $\overline{}^{C}$ operator is a cyclostationary time average.

2.2.2 Iterative nudging-based ERBC

The iterative procedure of ERBC based on nudging is first introduced and described by Krinner et al. (2025). The model, altered by the addition of $G_{\mathrm{N}}(t) := G_{\mathrm{N}_0}(t)$ for the selected prognostic variables, is again nudged towards the reference. A cyclostationary time average of the resulting relaxation terms' time series yields a second corrective term : $G_{\mathrm{N}_1}(t)$. The sum $G_{\mathrm{N}_0}(t) + G_{\mathrm{N}_1}(t)$ forms an updated corrective term. This iterative procedure can be repeated multiple times. Krinner et al. (2025) showed that two iterations allows further bias reduction while avoiding "over-correction".

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

2.2.3 State-dependent nudging-based ERBC

As stated in subsection 2.1, the corrective terms G depends at least on the time of the year and the location within the model grid. Inspired by some a posteriori bias adjustment methods (Verfaillie et al., 2017; Addor et al., 2016) and machine learning-based run-time bias correction (Watt-Meyer et al., 2021; Chapman and Berner, 2025, e.g.), one could attempt to make it as well conditional on the current weather situation. Multiple choices must be made regarding the definition of "current weather situation", starting with the region considered around a given grid point. In this study, the Earth has been divided in 12 fixed regions (see Fig. A1 for their definition) in which 3 typical weather situations are defined for each month using self-organizing maps (SOMs) (Kohonen, 2013) on geopotential height at 500 hPa. For a given grid point, the corrective term G depends now on the time of the year, the location within the model grid and the identified weather situation within the region which this grid point belongs to. The procedure to infer this state-dependent corrective term is detailed in Appendix A2. The choices made regarding the number of regions, their definition, the number of weather types and the procedure to define them are discussed in Section 4.

2.3 Climatological Adaptive ERBC

Scinocca and Kharin (2024) introduced an alternative way of deriving correction terms to implement ERBC, named Climatological Adaptive Bias Correction (CABCOR) method. In this method, a long simulation is run with a yearly evolving tendency adjustment term $C_n(t)$:

$$\frac{\partial X}{\partial t} = F_{\mathrm{M}}(t) + C_n(t)$$

$C_n(t)$ is updated each year, such that for the year $n+1$, $C_{n+1}(t)$ is equal to

$$\frac{X_{\mathrm{R}}(t) - \overline{X(t)}^{C_n}}{\tau}$$

where $\overline{}^{C_n}$ refers to the cyclostationary time average of the n first years of simulation.

It is assumed that the series of $C_n(t)$ will converge to a cyclostationary term, which should then be used as ERBC corrective term $G_{\mathrm{CA}}(t)$ in an ERBC run:

$$G_{\mathrm{CA}}(t) = C_N(t)$$

where N is the length of the adaptation simulation. Here, the adaptation simulation is 110 years long and we used BC terms given by adjustment terms of the 30th, the 70th and the 110th years.

2.4 Model and simulations

Simulations in this study are performed using the LMDZ v6.3 AGCM (Hourdin et al., 2020b) at low resolution (96 (longitude) $\times$ 95 (latitude) horizontal grid points and 79 vertical levels).

Control (CTRL) simulations refer to simulations in which no alteration of the prognostic equations is performed.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Nudged simulations refer to simulations with nudging towards the ERA5 reanalysis (Hersbach et al., 2020) (using the first year as spinup). Using a hyperbolic tangent damping coefficient on nudging tendencies, nudging is performed only between $P = 0.85P_s$ and $P \approx 0.06P_s$ (where P and P_s are the atmospheric pressure at a given level and at the surface, respectively). The planetary boundary layer is let unperturbed in order to avoid unwanted interactions with paramaterizations, while the upper part of the atmosphere remains free for model stability reasons. They are run over the 1980-2000 period with AMIP boundary conditions, the first year being used as spinup. The period 1981-2000 is used as "learning period", from which nudging terms are extracted to build ERBC terms.

CABCOR adaptation runs are longer simulations, with repeated annual-cycle forcings (as in Scinocca and Kharin, 2024), set to the year 1991. The yearly update procedure of the increments is conducted using ERA5 1981-2010 climatological means of monthly meridional and zonal winds. As in nudged simulations, the planetary boundary layer and the upper stratosphere remain unaltered. A 20-year spinup for a slow convergence of the time constant τ is applied as described in Scinocca and Kharin (2024).

All bias-corrected (BC) simulations span the 2000-2020 period, the first year being used as spinup. They are run with the following forcings:

- AMIP boundary conditions,
- historical greenhouse gas and aerosols concentrations from 2000 to 2014,
- greenhouse gas concentrations from the SSP2-4.5 scenario from 2015 to 2020,
- repeated 2014 aerosols concentration from 2015 to 2020.

In this study, only zonal and meridional winds are corrected, unlike other studies like Scinocca and Kharin (2024), that also corrected temperature, humidity and surface pressure. LMDZ6 having a rather large and uniform cold bias within the free troposphere, adding a "warming" increment proves to be detrimental in LMDZ6 on convective processes representation and on the surface energy balance. Therefore, as in Krinner et al. (2025) we refrain to use the above-described methods on temperature. Humidity and surface pressure are also left unaltered to avoid interfering with global mass and water balances.

Table 1 lists all simulations mentioned in the present paper. Adaptation simulations are named *XXnd* with *XX* being "N" for nudged simulations and "CA" for CABCOR adaptation simulations, and *n* indicating the value in days of the time constant τ (e.g. 1d for τ =1 day). BC simulations are named *BC(xx-yy)* where *xx* refers to the corresponding adaptation simulation and *yy* gives details on the estimation method. Note that some simulations have been run twice with varying initial conditions, as indicated in the column "Nb of simulated years". For figures that display maps, only one of these two-simulation ensembles is used.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Name	Details	Nb of sim. years
CTRL	control simulation, defautl model	2×20
N1d	Nudged sim., $\tau = 1$ day	20
N2d	Nudged sim., $\tau = 2$ days	20
N3d	Nudged sim. , $\tau = 3$ days	20
BC(N1d)	ERBC, derived from N1d	2×20
BC(N2d)	ERBC, derived from N2d	20
BC(N3d)	ERBC, derived from N3d	20
BC(N1d-it2)	Iterative nudging-based ERBC derived from nudged sim. with $\tau = 1$ day iteration 2	2×20
BC(N3d-it2)	Iterative nudging-based ERBC derived from nudged sim. with $\tau = 3$ days iteration 2	20
BC(N1d-cond)	State-dependent nudging-based ERBC derived from nudged sim. with $\tau = 1$ day 3 weather types per month	20
CA1d	CABCOR adaptation sim., $\tau = 1$ day	110
BC(CA1d-y30)	CABCOR ERBC from year 30 of CA1d	20
BC(CA1d-y70)	CABCOR ERBC from year 70 of CA1d	20
BC(CA1d-y110)	CABCOR ERBC from year 110 of CA1d	20

Table 1. Name and description of all simulations mentioned in this work. Adaptation simulations are named *XXnd* with *XX* being "N" for nudged simulations and "CA" for CABCOR adaptation simulations, and *n* indicating the value in days of the time constant τ (e.g. 1d for τ =1 day). BC simulations are named *BC(xx-yy)* where *xx* refers to the corresponding adaptation simulation and *yy* gives details on the estimation method. Note that some simulations have been run twice with varying initial conditions, as indicated in the column "Nb of simulated years".

2.5 Reference datasets and evaluation criteria

Evaluation is carried out over the 2001-2020 period against ERA5 reanalysis, except for precipitation for which we used the Global precipitation Climatology Dataset (GPCP) (Huffman et al., 2024) (monthly mean precipitation averaged over the 2001-2020 period).

Biases of time means are computed as global (Section 3.2.1) or regional (Section 3.5.1) root mean squared error (RMSE) for the following variables:

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

- $\mathbf{wind_{200}}$, $\mathbf{wind_{500}}$, $\mathbf{wind_{850}}$: horizontal wind vector at 200 hPa, 500 hPa, 850 hPa. The squared error is computed as the sum of the squared error of zonal wind and the squared error of meridional wind.
- $\text{slp}_{\text{ocean}}$: sea level pressure (slp) over ocean only.
- T_{200}, T_{500}, T_{850}: air temperature at 200 hPa, 500 hPa, 850 hPa.
- $T_{\text{2m,land}}$: air temperature at 2 meters above ground over land and sea ice only.
- precip: total precipitation.

Two key midlatitude atmospheric variability features are evaluated: blocking frequency and storm tracks. Blocking frequency is diagnosed on an annual and seasonal timescale using the definition of Davini et al. (2012) that uses gradients of 500 hPa geopotential height. Storm tracks are diagnosed by applying a band-pass time filter of 2 to 6 days to the 500 hPa geopotential height field and computing the seasonal standard deviation of the resulting field.

At regional scale, monthly to interannual variability is described by computing empirical orthogonal functions (EOF) of the monthly 500 hPa geopotential height fields grouped by season (December-January-February: DJF, March-April-May: MAM, June-July-August: JJA, September-October-November: SON). It is evaluated by computing the correlation between the model's first component and the reanalysis's first component.

3 Results

3.1 Correction terms

All derivation methods lead to similar bias correction increment patterns as shown in Fig. 1 for the 500 hPa level: an eastward increment at 30°S, stronger between 60°W and 120°E, and an eastward increment in the North Pacific at 30°N are the clearest signals. All methods also lead to large corrective terms over the Andes and the Himalayas whose direction is rather chaotic. In some regions however, patterns can differ. For example, in the North Atlantic at about 30°N, where all other methods result in a south-eastward increment, BC(N1d-it2) gives a rather strong north-eastward increment. Overall, BC(N1d-it2) terms stand out by their amplitude, which is much larger. Nudging-based BC terms, where only τ varies, yield increments that are very similar – but not identical – in shape. However, the amplitude of the BC terms decreases with increasing value of τ, by a factor that is not equal to $\frac{1}{\tau}$. BC(N1d-cond) terms are almost identical to those of BC(N1d). This is not surprising considering that the weather types on which the correction terms of BC(N1d-cond) depend are roughly equiprobable. So, when averaged over time, the conditional part of the correction terms cancels out, despite it being contrasted depending on the weather type. Interestingly, corrective terms from CA1d are very similar no matter the year from which they are extracted. This is further discussed in Section 4.1.

When examining the zonal means of the meridional and zonal wind increments (see Fig. B1 and Fig. B2), similar conclusions can be drawn. Overall, the corrective terms are very similar, with differences mainly in amplitude depending on the number of iterations, the value of τ, or – but only marginally – the length of the adaptation run for CABCOR terms.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

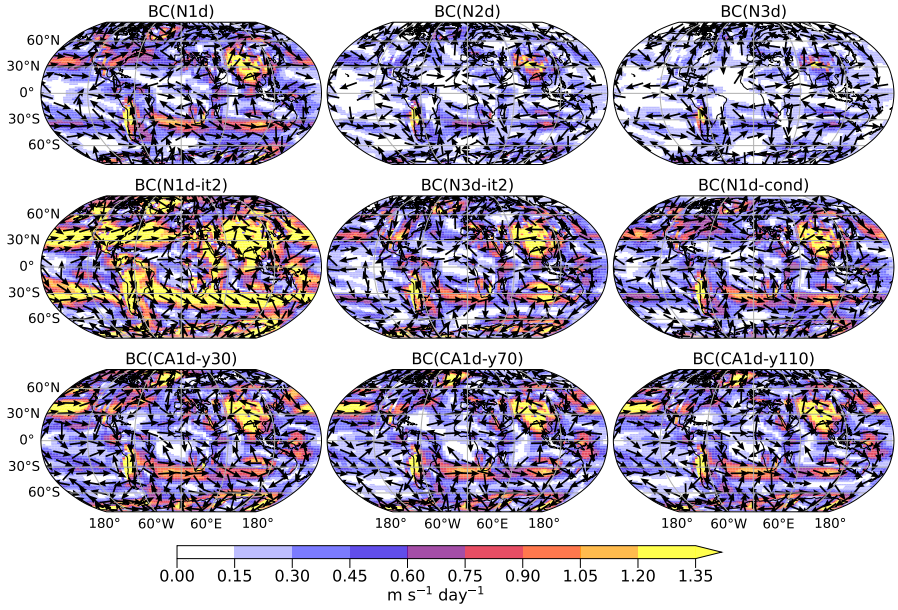

Figure 1. Annual mean of wind correction terms at 500 hPa. Color shading is the norm (in m s^{-1} day^{1}) while arrows indicate the direction. Only one out of four grid points carry an arrow to enhance readability. No arrows are drawn where the norm is less than 0.1 m s^{-1} day^{1}.

It is noteworthy that BC terms derived from nudging increments and those obtained with the CABCOR method are very similar. This is further commented on in Section 4.4.

3.2 Global biases of time means

3.2.1 Global scores

Global mean biases of wind and temperature at different pressure levels and mean sea level pressure (over ocean only) are evaluated using their root mean squared error (RMSE) with respect to ERA5. Global total precipitation RMSE is computed with respect to GPCP. Figure 2 shows an intercomparison of global performance of the methods in terms of bias reduction. Scores (represented by colors) are defined for each criterion (each column as):

$$\frac{\mathrm{RMSE} - \mathrm{RMSE}_{\mathrm{min}}}{\mathrm{RMSE}_{\mathrm{max}} - \mathrm{RMSE}_{\mathrm{min}}}$$

where $\mathrm{RMSE}_{\mathrm{max}}$ (resp. $\mathrm{RMSE}_{\mathrm{min}}$) is the highest (resp. the lowest) RMSE for this criterion among the evaluated simulations. Hence, a simulation with a score equal to 0 means it has the lowest RMSE and a score equal to 1 means it has the highest RMSE. CTRL and N1d results give respectively a lower and upper bound of the expected performance, as they are, for most variables, the simulations with the highest and lowest RMSE. This particular score, normalized by the maximum improvement for a given variable, was preferred over a simple ratio of RMSE over the CTRL RMSE, to account for the difference in expected effect of

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

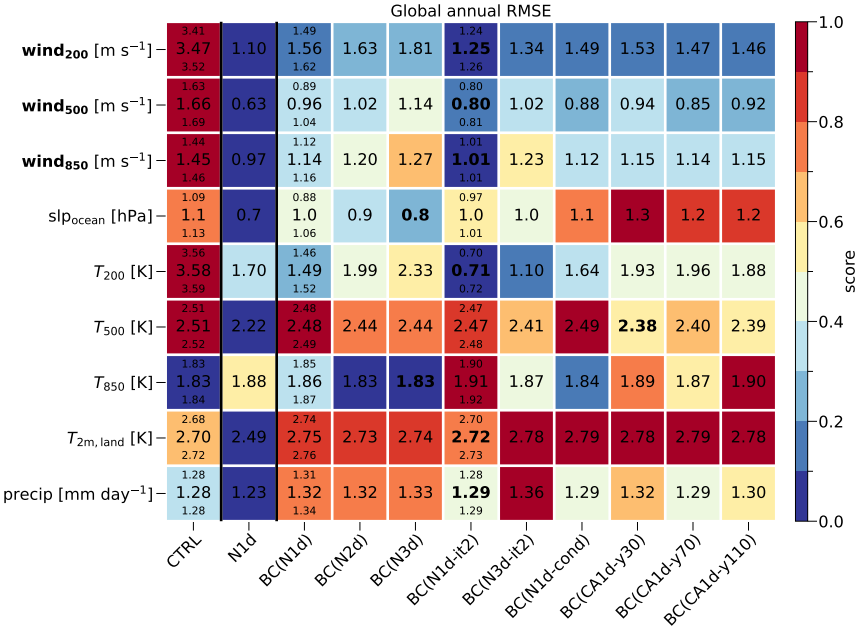

Figure 2. RMSE of global mean variables and corresponding scores. Numbers correspond to the RMSE while color shading corresponds to the associated score. Bold font is used for the lowest RMSE for a given variable among BC simulations. For simulation types for which multiple runs were available, the minimum and maximum RMSE are indicated below and above the mean RMSE. The score's and variables' definitions are in the main text.

BC. Effects on wind are much larger than on temperature, when compared to CTRL. One must, however, read Fig. 2 carefully, as in some cases, large variations in score, and therefore in colors, are actually related to very limited RMSE differences.

Among nudging-based BC simulations without iterations, for all wind-related scores, the best results are obtained with $\tau = 1$ day in the nudged simulation. Sea level pressure biases are, however, best improved when $\tau = 3$ days. Temperature biases are strongly reduced at 200 hPa, but only marginally at lower altitudes. For temperature at 200 hPa, BC(N1d) again performed the best among these three.

Making corrective terms state dependent, as done in BC(N1d-cond) does not substantially improve the performance in wind scores, and does not reduce the global sea level pressure bias compared to CTRL simulations: RMSEs in BC(N1d-cond) are indeed smaller than mean RMSEs of BC(N1d), but this difference is most likely attributable to internal variability (see lower RMSEs of BC(N1d)). For temperature also, no clear improvement or degradation can be observed compared to BC(N1d). However, precipitation is improved compared to BC(N1d), although to a limited extend, which significance is not guaranteed given global precipitation decadal variability.

After two iterations, with $\tau = 1$ day, biases of winds and temperature at 200 hPa are even more strongly reduced and T_{200} biases are even lower than in N1d. Sea-level pressure biases and temperature biases at other altitudes remain however in

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

the same range. When iterating with $\tau = 3$ days performances are improved when compared to BC(N3d), but they do not outperform BC(N1d) (and hence also not BC(N1d-it2)) except for $\mathbf{wind_{200}}$ and T_{200}.

There is no major difference among the three CABCOR simulations. Temperature biases at 200 hPa are well reduced, as for the nudging-based BC, and temperature at lower altitudes is not degraded when compared to CTRL simulations. Wind scores are comparable to those of BC(N1d), some slightly better, but not outperforming BC(N1d-it2). Sea-level pressure biases are not reduced with this method. Temperature biases remain in the same range as with the other methods. There is no major difference among these three CABCOR simulations.

All BC simulations have degraded precipitation global bias compared to CTRL simulations, but not N1d. Note however that in the CTRL simulation, precipitation bias is already large. The model is indeed known to overestimate precipitation globally, and to have a rather important global mean bias, similarly to other AGCM (Hourdin et al., 2020a). Here, the degradation of precipitation bias is also of a similar order of magnitude to that of switching from one version of the model to another (Hourdin et al., 2020a). Examining maps of annual precipitation (see Fig. B5), it can be concluded that BC tends to increase positive precipitation biases and over-compensates for negative bias in the tropical Pacific, leading to a degraded global RMSE. BC does however reduce bias in the tropical Atlantic. The effect of bias-correction on precipitation might reveal a complex entanglement between run-time bias correction and model tuning. It must also be considered carefully, as precipitation observational data are themselves subjected to biases.

All BC simulations also have slightly degraded $T_{\mathrm{2m,land}}$ bias compared to CTRL simulations, but again, not N1d. Examining maps of annual $T_{\mathrm{2m,land}}$ (see Fig. B4) reveal that the "unwanted" effects of BC are a strong warming of North-America ($\approx$ +2 K at the most), a cooling of Antarctica (up to $\approx$ -2K locally) and a slight cooling of Africa ($\approx$ -0.5 K).

3.2.2 Spatial analysis: describing the spatial effects of the ERBC methods

Figure 3 shows the mean $\mathbf{wind_{500}}$ bias of BC simulations and the CTRL simulations. Whereas nudging reduces $\mathbf{wind_{500}}$ biases without changing the overall bias patterns (not shown), bias-corrected runs have bias patterns that differ from the CTRL simulations. For instance, BC runs all exhibit an eastward $\mathbf{wind_{500}}$ bias at 60°S circling around the South Pole, with varying amplitude, opposite to the original bias in the CTRL simulations. The North Pacific remains a region of strong wind biases in nudging-based BC simulations ; this is less the case but less of CABCOR simulations. Wind biases in the inter-tropical region are stongly attenuated in BC simulations compared to CTRL. Iterations of nudging-based BC seems to allow reducing persisting biases, i.e., biases pattern that were present in CTRL as well as in nudging-based BC simulations with no iteration (see for example, the North Pacific in BC(N1d) and in BC(N1d-it2)). CABCOR simulations have similar residual bias patterns than nudging-based BC simulations, but stand out in some regions: less bias in the North Pacific, and more zonal bias in the Southern midlatitudes between 60°W and 180°E, for example.

Figure 4 shows the mean slp bias over ocean of BC simulations and CTRL. CTRL runs present annual slp biases mostly above 30° of latitude, with an amplitude that regionally reaches 2 hPa to 4 hPa. BC(N1d), BC(N2d) and BC(N3d) have very similar slp biases patterns. Noteworthy are the negative bias around Antarctica, the rather strong positive bias in the North Pacific, and the improvement compared to CTRL simulations in the North Atlantic. Iterating enhances slp bias reduction in

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

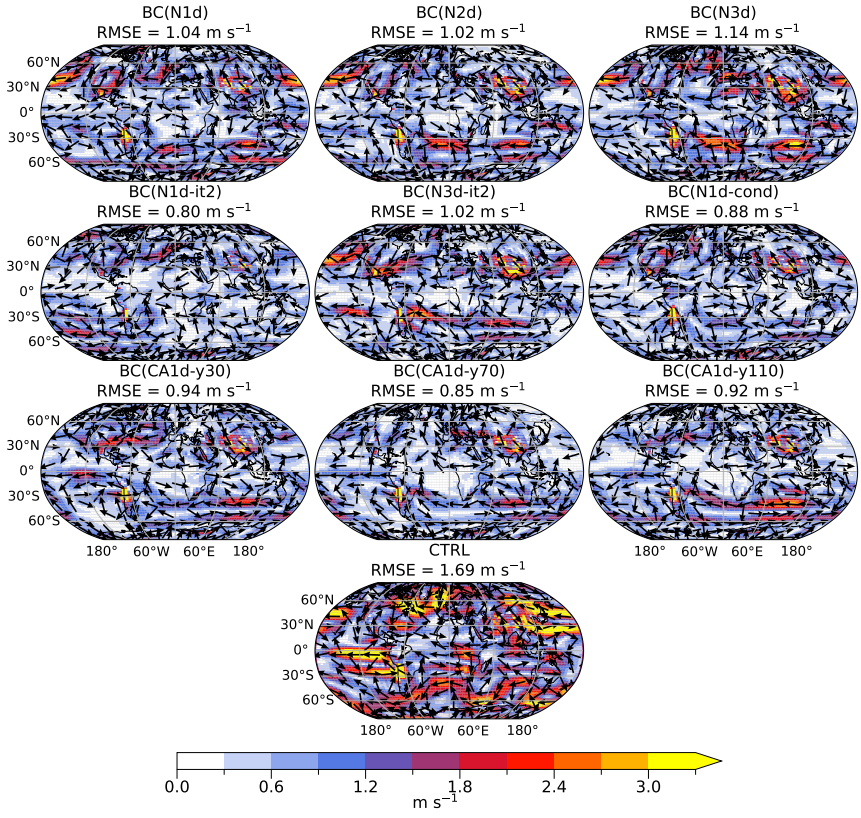

Figure 3. Mean wind bias at 500 hPa. Corresponding RMSEs are indicated on top of each panel. Color shading corresponds to the norm of the wind bias vector, while arrows indicate its direction. Only one out of three grid points carry an arrow to enhance readability. No arrows are drawn where the norm is less than 0.3 m s^{-1}.

the North Pacific and around Antarctica (see BC(N1d-it2)). BC(N1d-cond) performs poorly in terms of slp biases, due to large negative biases at high latitudes. BC(CA1d-y110), BC(CA1d-y70) and BC(CA1d-y30) have strong negative bias at high latitudes as well, but extending further South in the Northern hemisphere.

As described before, the strongest impact on temperature is found above 250 hPa, and a zonal mean analysis (see Fig. B3) shows that the warming occurs at latitudes higher than 30° approximately. A slight warming is also present for all methods between 750 hPa and 300 hPa at approximately 30°N and 30°S. However, all BC simulations present a slight cooling of the free troposphere at high latitudes, hence accentuating the cold bias of the model.

Global climate models are known for their equatorward bias in the mid-latitude jet streams (Harvey et al., 2020; Hourdin et al., 2020a). This is also the case of the uncorrected LMDZ. Figure 5 shows that ERBC overall reduces this positional bias. These zonal means of wind biases also reveals that all BC procedures give rise to a positive zonal wind bias extending through

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

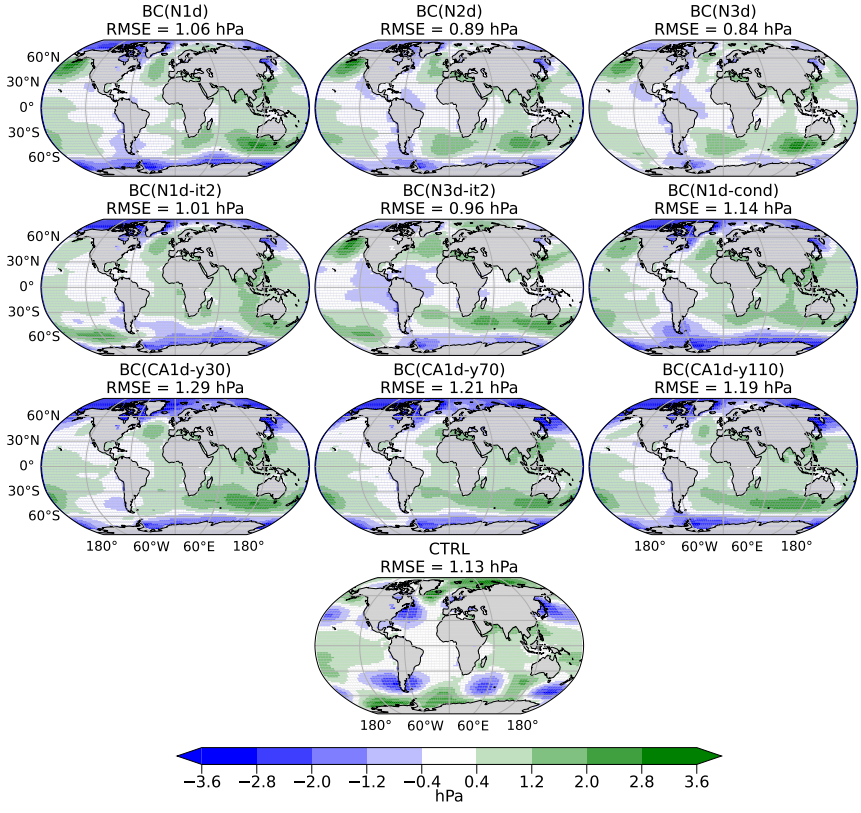

Figure 4. Mean slp bias over ocean only. Corresponding RMSEs are indicated on top of each panel.

the whole atmosphere at about 60°S. This is consistent with what is observed at 500 hPa on Fig. 3. No major differences are observable between the BC methods.

3.3 Midlatitude atmospheric variability features

Run-time bias correction are designed to reduce biases in climatological means of targeted variables, in practice here, winds. They however have the potential to also improve aspects of the atmospheric variability. Here, we focus on two important physical properties that are known to influence greatly the climate at midlatitudes.

3.3.1 Blockings

Midlatitude atmospheric blocking are often studied in climate models as they have a great influence on midlatitude weather and climate (Lenggenhager and Martius, 2019; Mohr et al., 2019). Climate models are also known to often underestimate their frequency (Davini and D'Andrea, 2020). Kleiner et al. (2021) shows that correcting the mean atmospheric state could improve the blocking frequency. Numerous ways to identify blockings have been developped and here we use the rather simple

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

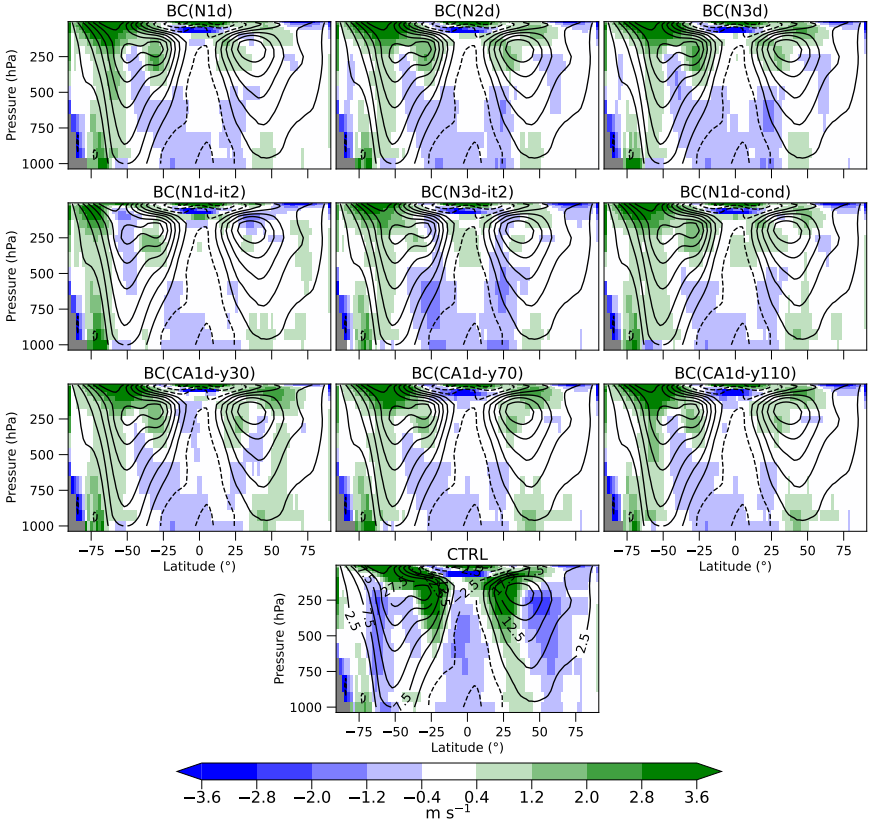

Figure 5. Zonal mean of zonal wind bias. Zonal mean of zonal wind in ERA5 is represented by black contours, with values in m s^{-1} indicated only on the lowest panel.

definition of Davini et al. (2012) based on the daily 500 hPa geopotential field that characterizes each grid-point each day as blocked or not. When averaged over a long period, this yields a map of blocking frequencies.

Figure 6 shows the effect of BC on annual blocking frequency in the Northern Hemisphere. In the Northern Hemisphere, blockings are most frequent North of Europe and over a region covering Alaska and Western Siberia. The CTRL version of the model underestimates the frequency over these regions by about 3 to 5 points. It tends to overestimates it by 1 to 2 points over the North of American continent and over continental Siberia. All BC versions reduce the overall RMSE of blocking frequency by increasing slightly the frequency where the negative bias was the strongest. However, they all tend to overcompensate for the positive bias North of America creating a negative bias. In the Southern Hemisphere (see Fig. B6), the annual frequency of blockings is overall smaller than in the Northern Hemisphere, and so are the biases. Similarly to the Northern Hemisphere, the CTRL model underestimates the frequency in regions where the frequency is the highest in the reference. Bias-corrected versions of the model however do not necessarily correct this, and even tend to degrade the frequency in many regions, resulting in a larger RMSE for all methods in the Southern Hemisphere.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

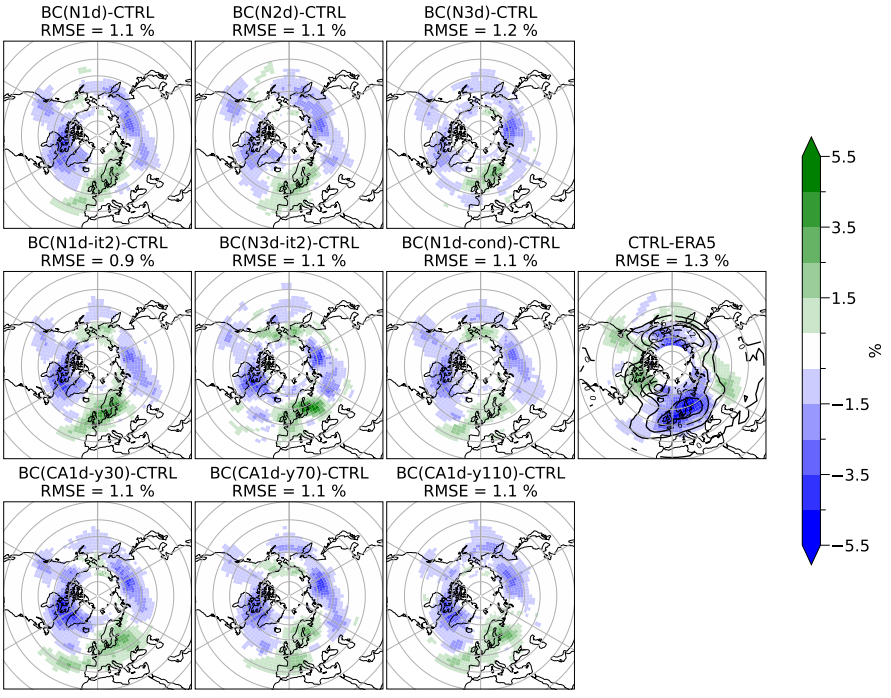

Figure 6. Effect of BC on annual mean blocking frequency in the Northern Hemisphere. Aside from the extreme right one, panels show the difference in annual mean frequency of blockings between BC simulations and the CTRL simulations. The extreme right panel shows the annual mean bias in frequency of blockings of a CTRL simulation with respect to ERA5 as well as ERA5' frequency represented by black contours. RMSEs indicated on top of each panel correspond to the RMSE of the simulation with respect to ERA5.

3.3.2 Storm tracks

Extratropical storm activity is much intertwined with midlatitude weather and climate. Accurately position storm tracks remains however quite a challenge for climate models (Priestley et al., 2020). Here we use seasonal standard deviation of daily 500 hPa geopotential field, after applying a band-pass filter of 2 to 6 days, as a metric for storm activity.

Figure 7 shows the DJF storm tracks bias of the CTRL simulation as well as the actual storm tracks in ERA5 (panel at the extreme right) in the Northern Hemisphere. For BC runs, it shows the difference with the CTRL run, hence showing the effect of BC. Without BC, the model underestimates storm activity almost everywhere in the Northern Hemisphere. With BC, this underestimation is partly compensated for, leading to reduced RMSEs. All methods produce a similar effect when compared to the CTRL simulation, and the differences in pattern might not be significant. It seems however that the nudging-based BC with iteration gives worse results than its non-iterated counter part (see RMSEs). CABCOR runs have residual biases comparable to nudging-based BC runs, with a minimum in RMSE achieved by BC(CA1d-y70).

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

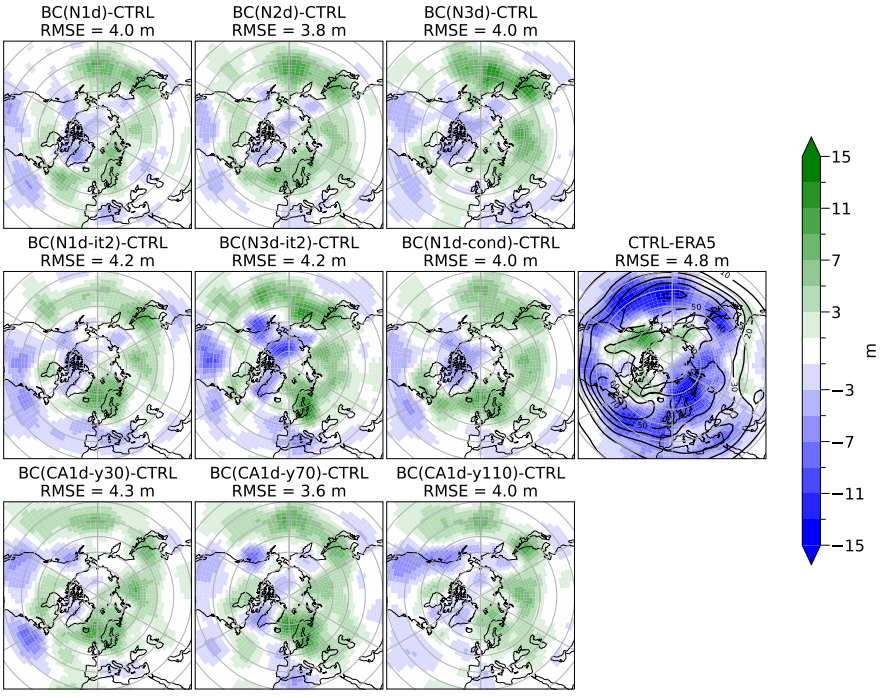

Figure 7. Effect of BC on DJF storm tracks in the Northern Hemisphere. Aside from the extreme right one, panels show the difference in DJF storm tracks between BC simulations and the CTRL simulation. The extreme right panel shows the mean bias in DJF storm tracks of a CTRL simulation with respect to ERA5 as well as ERA5' storm tracks represented by black contours. RMSEs indicated on top of each panel correspond to the RMSE of the simulation with respect to ERA5.

3.4 Regional view: do the methods perform better in some regions than others?

Before focusing on a specific region to evaluate the contribution of the various methods, we addressed two questions:

1. At what size of region is it reasonable to expect a reduction in average biases?
2. Are there differences in the contribution of the correction between regions?

To address these questions, we defined a "moving RMSE": at each grid-point, the RMSE of a given variable with respect to the reference is computed within a circular region centered at that grid-point, and this with different radii. In the following, we focus on $\mathbf{wind}_{500}$. Figure 8 shows the resulting map for a radius of 4,000 km in the CTRL simulation (lower panel). Other panels on Fig. 8 show the ratio of the moving RMSE in BC simulations over the moving RMSE in the CTRL simulation. We performed the same analysis with other radii (2,000 km, 3,000 km, 6,000 km, 8,000 km, see Fig. B7, Fig. B8) and it was concluded that 4,000 km is the smallest radius for which no region, or only very small ones, in BC simulations has degraded RMSE compared to CTRL; which answers question 1, at least for this variable. At this scale, only BC(N3d-it2) exhibits a very

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

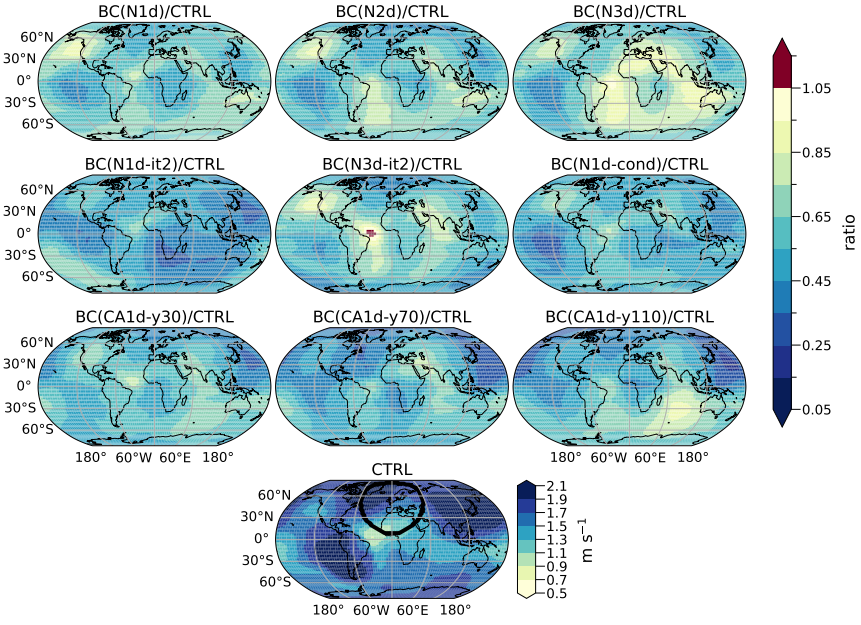

Figure 8. Ratio of moving RMSE of wind at 500 hPa of BC simulations over the CTRL one (top 6 panels) and moving RMSE of wind at 500 hPa of a CTRL simulation (lower panel). RMSE in a grid point corresponds to the RMSE within the circle with a radius of 4,000 km around it. A circle centered at 0°W-45°N, with radius 4,000 km is drawn on the lowest panel.

small region of slight performance degradation, corresponding to a region where the RMSE in the CTRL simulations is very low.

By examining Fig. 8, we can now try to answer question 2. Even though, bias reduction is expected everywhere, it is very limited in some regions for some methods. BC(N3d) exhibits a large region covering Europe, Africa, the Indian Ocean, the South Pacific and the ocean Southern of Africa where bias reduction is only of about 10%. These maps also reveal that iterating the nudging-based BC (as in BC(N1d-it2) and BC(N3d-it2)) improve bias reduction in some regions and degrade it in others, indicating that despite a global improvement of the bias, such an improvement is not guaranteed in all regions.

The same analysis could be performed for other variables. Similar conclusions on the contrasted bias reduction depending on the method and the region are expected.

One might want to select the method that gives the smallest residual biases in their region of interest. Figure 9 displays the method leading to the lowest moving RMSE for wind and temperature at different altitudes, within circular regions with radius 4,000 km. The first observation is that the results depend on the variable. For wind, there is nevertheless a large region for which, no matter the altitude, BC(N1d-it2) gives the lowest residual biases. There is also a quite extensive region where BC(CA1d-y110) leads to the lowest residual biases of $\mathbf{wind_{200}}$ and $\mathbf{wind_{500}}$. BC(N1d-it2) gives the lowest residual biases of $\mathbf{wind_{850}}$ over a region covering almost all the mid to low latitudes. Interestingly, BC(N1d-cond), which does not stand out at global scale, gives the lowest residual $\mathbf{wind_{500}}$ biases in the South-Pacific. For temperature, one must first remember, that

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

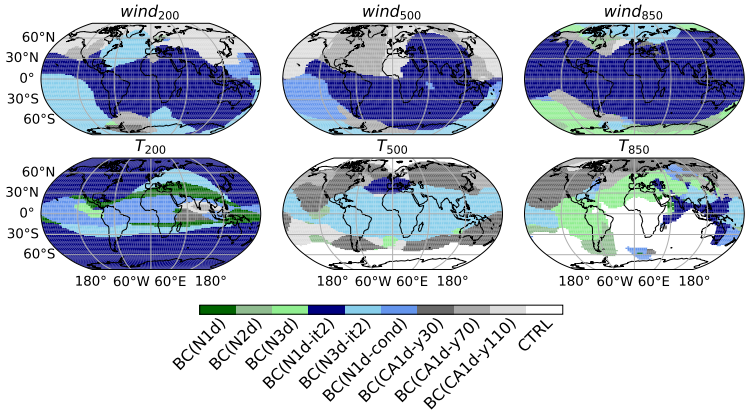

Figure 9. Method leading to the lowest moving RMSE for various variables within circular regions with radius 4,000 km. RMSE in a grid point corresponds to the RMSE within the circle with a radius of 4,000 km around it

with bias-correction applied solely on winds, only T_{200} biases were found to be reduced. BC(N1d-it2) leads to the lowest T_{200} biases for most of the extratropical regions, where BC(N1d-cond) does in most of the tropical region. At 500 hPa, temperature biases are the lowest in the CTRL simulation at the Poles, and in BC(N3d-it2) in the tropical region. In the extratropical regions, CABCOR simulations rather dominates, mostly BC(CA1d-y30). At 850 hPa, temperature biases are the lowest when the model is unmodified (CTRL) or only slightly altered (BC(N3d)) for a very large portion of the globe. BC(CA1d-y30) gives however the lowest residual biases in the Northern high latitudes and in the North-Atlantic, whereas iterated nudged-based BC (BC(N1d-it2) and BC(N3d-it2)) does in South-Asia and part of the tropical Pacific.

Figure 9 illustrates to what extent the answer to the question "which method performs best" depends on the region and the variable of interest. We now dive deeper into the analysis of one specific region.

3.5 Focus on a specific region: example of the Europe-Atlantic region

This subsection focuses on a region covering Europe and a large part of the North Atlantic (60°W, 60°E, 30°N, 65°N). Note that at 50°N, the width of this region (120°) is close to 8,000 km, hence has the same width than a circular region with a radius of 4,000 km as was used to compute moving RMSEs on Fig. 8.

3.5.1 EU-A Annual scores

Figure 10 shows annual performance scores for this region. Unlike on a global scale, in this region, it is not clear that a lower value for τ in nudging-based methods is optimal. In fact, for wind, BC(N2d) gives the best results among all three BC(Nxd) and BC(N3d-it2) gives the best results of all nudging-based methods. For slp, the simulation with the less wind bias reduction gives the best results (BC(N3d)). For temperature, which methods perform the best really depends on the altitude.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

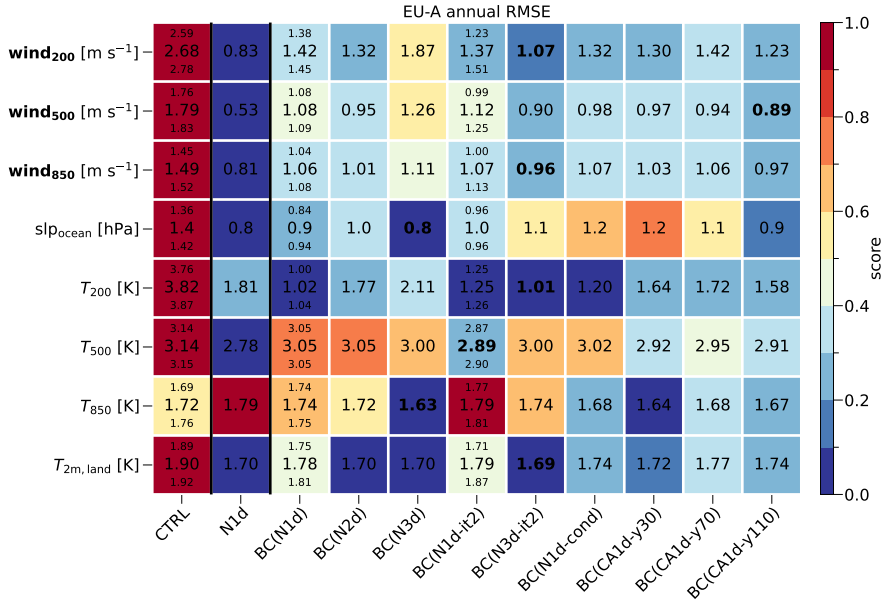

Figure 10. RMSE of EU-A mean variables and corresponding scores. Numbers correspond to the RMSE while color shading corresponds to the associated score. Bold font is used for the lowest RMSE for a given variable among BC simulations. For simulation types for which multiple runs were available, the minimum and maximum RMSE are indicated below and above the mean RMSE. The score's and variables' definitions are in the main text.

The same scores and RMSE as in Fig. 10 were computed by season (DJF, MAM, JJA, SON) (not shown). From these results it becomes apparent that on a seasonal regional scale, none of the methods can be declared optimal for all variables examined. This indicates yet again that a method performing best at global annual scale is not necessarily optimal for a specific region, season and/or sets of variables.

Figure 11 shows maps of mean $\mathbf{wind_{500}}$ bias (it is a close-up of Fig. 3 for the region of interest). One can see that ERBC allows a strong reduction of the large negative bias over the North-Atlantic, and to a lesser extent of the positive bias over the Alps. When the length of the adaptation simulation is increased for CABCOR simulations, the RMSE decreases and the bias pattern evolves. BC(CA1d-y30) and BC(CA1d-y110) have similar bias patterns, different from BC(CA1d-y70). It is either a sign of internal variability or the indication that the correction terms are still evolving in the adaptation simulation. Iterating the nudging-based methods gives lower RMSEs of $\mathbf{wind_{500}}$ in this region, though in the case where $\tau = 1$ day the better score might be attributable to variability (see minimum-to-maximum intervals in Fig. 10). The residual bias pattern is not substantially modified after the iterations. When $\tau = 3$ days, a strong residual bias in BC(N3d) is greatly reduced in BC(N3d-it2). With state-dependent ERBC (BC(N1d-cond)), $\mathbf{wind_{500}}$ bias in this region is more reduced than in BC(N1d), but the residual bias pattern remains close.

Note that all these conclusions must be considered carefully as strong internal variability is expected at this regional scale.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

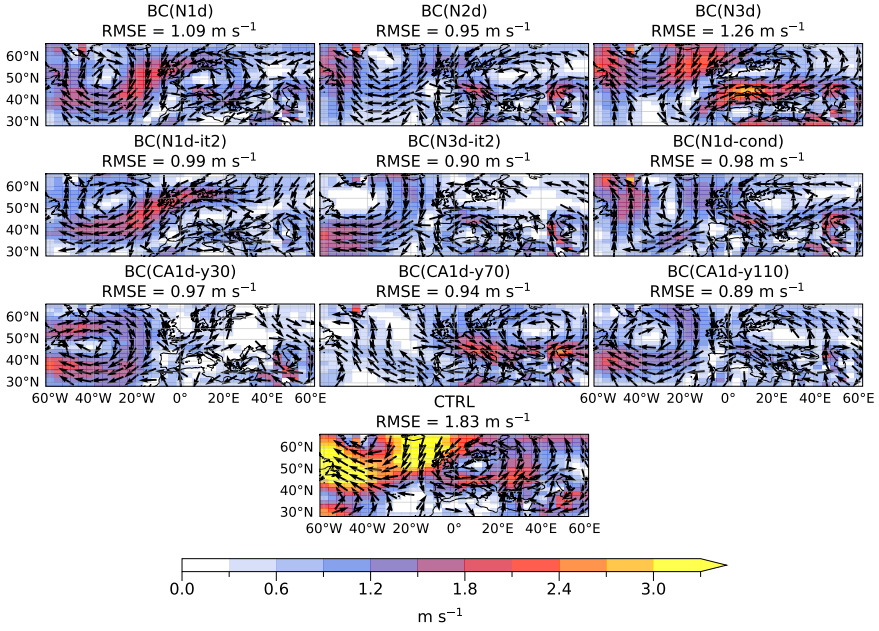

Figure 11. Mean wind bias at 500 hPa in EU-A. Close-up of Fig. 3. Corresponding RMSEs are indicated on top of each panel. Color shading corresponds to the norm of the wind bias vector, while arrows indicate its direction. Only one out of two grid points carry an arrow to enhance readability. No arrows are drawn where the norm is less than 0.3 m s^{-1}.

3.5.2 Monthly to interannual circulation variability

Monthly to interannual circulation variability is measured over this region by doing, for each season, an empirical orthogonal function (EOF) analysis on monthly geopotential height at 500 hPa. Performances of the model are evaluated by computing the correlation between the model first component and the reanalysis first component. Results are shown in Table 2. In some cases, the model first component resembles more one of the other reanalysis components (based on RMSE), and not the first. Only values corresponding to cases where the model first component actually matches the reanalysis first component are shown.

In DJF, BC simulations are at least as good as the CTRL simulations, even better for most of them except BC(CA1d-y70). In MAM, JJA and SON, CTRL simulations first pattern do not match the reanalysis first pattern. In MAM, BC seems to allow improvement, except nudging-based method with $\tau \neq 1$ day and BC(CA1d-y70). In JJA, only BC(N2d), BC(N1d) and BC(CA1d-y110) get the match, with rather low correlations, except BC(CA1d-y110). In SON, almost all BC simulations do better than the CTRL simulation. Highest correlations are obtained with BC(N1d) and BC(CA1d-y110).

Note that BC(CA1d-y70) gets the first pattern to match for neither of the seasons, and that only BC(CA1d-y110) and BC(N1d) do for every season.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Simulation	DJF	MAM	JJA	SON
CTRL	0.82 [0.78-0.85]			
N1d	1.0	1.0	0.99	1.0
BC(N1d)	0.83	0.78 [0.76-0.80]	0.60	0.87 [0.93-0.80]
BC(N2d)	0.86		0.58	0.72
BC(N3d)	0.94			0.70
BC(N1d-it2)	0.88 [0.83-0.92]	0.84 [0.82-0.87]		0.44
BC(N3d-it2)	0.83			
BC(N1d-cond)	0.83	0.74		
BC(CA1d-y30)	0.88	0.64		0.61
BC(CA1d-y70)				
BC(CA1d-y110)	0.70	0.64	0.82	0.81

Table 2. Correlation between model's first EOF component and ERA5's first component. Only values corresponding to cases where the model first component actually matches the reanalysis first component are shown. Minimum-to-maximum intervals are indicated in brackets for configurations for which multiple simulations were available

Note also that in the case of BC(N1d) and BC(N1d-it2), for which two simulations are available, in certain seasons, only one of the two actually gets the first pattern to match. This reveals great variability between simulations with the same BC method used.

Finally, BC(N1d-cond) does not particularly stand out and does not outperform BC(N1d) for this metric, even though one could expect state-conditional BC to prove beneficial for variability aspects.

4 Discussion

This study presents a comparison of different empirical methods for estimating run-time bias correction increments using the LMDZ6 AGCM. We examined the impact of various methodological choices on the model's performance, evaluating it from a global scale – using aggregated scores of global mean errors in circulation, temperature, and precipitation – to mid-latitude atmospheric variability, and finally to a regional scale focusing on the model's performance over a large area covering Europe and the North Atlantic.

4.1 Optimal set of parameters at global scale

With LMDZ, we found that a nudging time constant of 1 day gives the best results in nudging-based bias-correction simulations for most criteria, while Scinocca and Kharin (2024) found however with CanAM5.1 that the optimal time constant value is 3 days (when no spatial filtering is applied to the corrective terms). This difference confirms the intuition that the optimal sets of parameters for the estimation of corrective terms is model-dependent.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Our work here confirms that iterating the nudging-based method reduces the mean bias of bias-corrected variables compared to simulations with nudging-based ERBC without iteration, as was shown in Krinner et al. (2025). It also confirms that it does not necessarily improve the representation of atmospheric circulation variability on interannual and shorter timescales. Here, we also tested this method with a different value of nudging time constant ($\tau = 3$ days), while Krinner et al. (2025) only tested $\tau = 1$ day. We thus showed that the effect of iterating is similar with this value, and that $\tau = 1$ day, as well as for the method without iteration, gives better result on mean variables at a global scale.

Scinocca and Kharin (2024) assess the impact of the length of the adaptation simulation for CABCOR simulations. They use lengths ranging from 110 years to 510 years, whereas here, we limited our analysis to 30 to 110 years. This restriction was motivated by the fact that we did not observe a major effect of prolonging the adaptation simulation from 30 to 110 years. So, unlike Scinocca and Kharin (2024), we do not conclude on an optimal length of the adaptation simulation. We simply observe that with LMDZ, with $\tau = 1$ day, we obtain very satisfactory results regarding bias reduction, with an adaptation simulation of length ranging from 30 to 110 years. We do observe contrasted results at regional scale.

4.2 Which variables should be bias-corrected

Other studies (Scinocca and Kharin, 2024; Chapman and Berner, 2024; Chang et al., 2019) employ ERBC not only on horizontal winds but also on temperature, humidity and surface pressure. Here, we applied ERBC on horizontal winds only. Regarding humidity and surface pressure, we believe that it is preferable to not interfere with the water and mass balances of the model: applying ERBC might indeed introduce artificial water or mass sources, although we did not evaluate rigorously these potential detrimental effects. For temperature, as LMDZ has a substantial cold bias in the whole troposphere at every latitude (see Fig. B3), it is very tempting to apply ERBC. Doing so, we were able to make the following observations. When a compensatory term is added to the temperature advection equation, as with the nudging procedure or with the bias correction, this cold bias is properly reduced. However, a new global equilibrium is quickly established in which oceanic evaporation is strongly reduced and convection is weakened. The evaporation deficit causes a surface energy imbalance reaching +20W/m^2 on average over 20 years of simulation. On average, temperature nudging tendencies or BC tendencies are of the same order of magnitudes than global mean temperature convective tendency of a CTRL simulation. In simulations with temperature nudging or BC, the convective tendency is attenuated strongly enough to keep a global temperature balance, despite the artificial warming tendency. The deficit in evaporation is consistent with an accumulation of humidity in the planetary boundary layer due to a lack of convective mixing. Because of these major side effects, we consider that adjusting temperature with run-time bias correction in LMDZ is not advisable. In other models, with less pronounced temperature biases, or which biases have different origins, bias-correcting temperature might not have such detrimental effects. In the other studies, which applied ERBC to temperature (e.g. Scinocca and Kharin, 2024; Scinocca et al., 2025; Chapman and Berner, 2025; Chang et al., 2019), nothing of the sort has indeed been mentioned.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

4.3 State-dependent bias correction

In this study, we chose to include one simulation with state-dependent bias correction. It relies on the definition of weather types in each of the 12 regions of the world defined Appendix A2 for each month using a self-organizing maps algorithm. The number of weather types per month is a free parameter of the method and in the simulation used in this study, it was set to 3. Three weather types per month and per region was indeed identified in a preliminary work as possible clustering of the nudging terms, in the sense that it creates clusters of nudging terms whose averages are significantly different than the average of all nudging terms. The lack of effect of making correction state dependent with this method may be the sign that our methodology is not adequate and must be adjusted. It was indeed assumed that correction terms should depend on weather types defined by the 500 hPa geopotential field, but it is a strong and maybe too simplistic assumption that could be questioned. It may also be the sign of a sampling problem: the length of the nudged-simulations might be too short to extract a significant state-dependent signal. Further investigations must be conducted to deepen our analysis.

4.4 Similarity between nudging-based correction terms, CABCOR correction terms and biases of the model

It is noteworthy that BC terms derived from nudging increments and those obtained with the CABCOR method are very similar, though they are obtained by quite different methods. BC terms derived from nudging increments are, up to a constant multiplier, the mean initial tendency errors from an observed state. CABCOR BC terms on the other hand, are, up to a constant multiplier, the residual climatological biases of the "adapted" model (i.e. adjusted by the yearly-evolving adjustment terms of the CABCOR procedure).

There are also similarities in patterns between correction terms and the climatological biases of the default model. This can be seen, for example, by comparing $\mathbf{wind_{500}}$ biases in the CTRL simulation (Fig. 3) and correction terms à 500 hPa (Fig. 1). This is also evident by comparing zonal wind biases in the CTRL simulation (Fig. 5) and the zonal mean of the zonal component of correction terms (Fig. B1). For example, around 200 hPa at 30°S and 30°N the positive biases can clearly be matched with the negative correction terms. There are, however, also major differences, for example in the intertropical region.

When nudging-based correction terms and biases of the initial biases are similar (up to a negative constant multiplier), it can be interpreted as the sign that long-term climatological biases are actually very rapidly formed and result mostly from the initial tendency erros of the model.

4.5 Completeness of the study

In this study, we delivered an intercomparison of multiple methods and their variations, but it remains far from exhaustive. In particular, we did not evaluate the effect of varying τ in the CABCOR method. Scinocca and Kharin (2024) show that the choice of the value for τ influences the overall performance of the correction. We chose to use $\tau = 1d$ only for our study, as this value gives the best results with nudging-based BC with LMDZ. We hence limited our study to the comparison between CABCOR and nudging-based methods and to the effect of the adaptation simulation's length.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Scinocca and Kharin (2024) also study the influence of spatial smoothing of the correction terms. As preliminary tests showed no significant effect, we decided to not investigate this further.

Regarding state-dependent ERBC, only one set of parameters is presented here, resulting from a preliminary analysis (see Section 4.3). In particular, the number of weather types is a very important parameter. Other tests were performed with various numbers, but as these choices had only limited influence on the model's performance, for the sake of brevity, they are not included in this study. Other sensitivity tests, as regions' or weather types' definition, could also be performed but they are beyond the scope of this work.

4.6 Evaluation period

Simulations were evaluated against reference datasets over the 2001-2020 period. The choice of this period has been guided by the need to evaluate simulations with bias correction outside of the period used to estimate the empirical correction terms (1981-2000). It can be argued that 2001-2020 is a period with a rather rapid and significant global climate change which could add to the "out-of-sample" effects. However, the main objective of developing ERBC is to use it in climate change simulations. Therefore, evaluating its effect on a period with climate change is not necessarily deleterious. We are also convinced that this choice should not influence greatly our conclusions.

5 Conclusions

This study produces an intercomparison of bias correction increments estimation's methods with the LMDZ6 AGCM at low resolution (96 (longitude) × 95 (latitude) horizontal grid points and 79 vertical levels). It explores the effect of methodological choices on various aspects of the model's performance. From aggregate scores of global mean errors in circulation, temperature and precipitation, we delve into mid-latitude atmospheric variability features. A more regional perspective is finally adopted, a large region covering Europe and the North-Atlantic serving as case study.

The following conclusions can be drawn:

- CABCOR and nudging-based corrections increments are strikingly similar. This results in rather similar effects of the bias correction, and similar model performance.
- No beneficial contribution was observed from making nudging-based BC state dependent, in the way it was done for this study.
- For nudging-based BC, a value of 1 day for the time constant τ is the best choice in view of most wind-related scores at global scale. However, it does not perform the best for all variables and performance criteria, as the "side effects" of BC (e.g. on sea level pressure) might also be amplified.
- Iterating the nudging-based BC has mostly positive effects on global wind biases, but not necessarily on other variables biases, on atmospheric variability or at a regional scale.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

In summary, no strict recommendation can be made regarding the optimal methodological choice. Global aggregated performance scores do not prejudge the outcome on other aspects of the model's performance. The choice of an ERBC method appears to be similar to the choice of a driving GCM for downscaling exercises. A number of criteria, and ways to aggregate them can be developed, but they all involve subjectivity. Plus, the user might choose the model (or its configuration) that gives the best results in their region of interests or on a particular aspects, with no guarantee that the selection criteria were relevant.

Ongoing work to further reduce model bias includes new tuning strategies and methods, continuous improvement of physical parameterizations, as well as the development of state dependent ERBC based on machine learning methods. Further work is also needed to assess the contribution to ERBC in AGCM in the case of dynamical downscaling. First studies showed the benefit of using run-time bias-corrected boundary conditions in regional climate models, and these results must be confirmed for more regions and more assessment criteria. It also remains an open question how much it would reduce the need of a posteriori bias-correction for impact studies and/or improve its performance, for example on extreme events.

Code and data availability. Datasets used for the model validation are:

- Copernicus Climate Change Service (2023): ERA5 monthly averaged data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.6860a573 (Hersbach et al., 2023)
- Copernicus Climate Change Service, Climate Data Store, (2024): ERA5 post-processed daily-statistics on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.50314f4c (Muñoz Sabater et al., 2024)
- Copernicus Climate Change Service (2021): precipitation monthly and daily gridded data from 1979 to present derived from satellite measurement. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.c14d9324 (Huffman et al., 2024)

Configuration files and model code changes, scripts and model output data used to plot the graphics are available on Zenodo under doi: 10.5281/zenodo.18955446 (Champouillon, 2026a). The modipsl code infrastructure required to compile the codes, including the full model code (in particular LMDZ v6.3 but also ORCHIDEE and XIOS), is available on Zenodo under doi:10.5281/zenodo.17381281 (Champouillon, 2026b). Note that the full LMDZ code can be downloaded via https://lmdz.lmd.jussieu.fr/ and that Champouillon (2026a) contains the additional and modified Fortran code needed for the iterative bias correction method as well as the state-dependent bias correction method.

Appendix A: Details of the procedure to derive BC terms

A1 Cyclostationary time average of nudging increments

In the nudged simulations, the model is relaxed every 6 hour towards the reanalysis and increments δ are saved. To preserve the diurnal cycle, the four moments of the day are treated separately. A multi-year daily mean is applied to the time series of increments, followed by a 20-day moving mean to smooth the resulting annual cycle. This yields for a day of the year D and an hour of the day H at a grid point p:

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

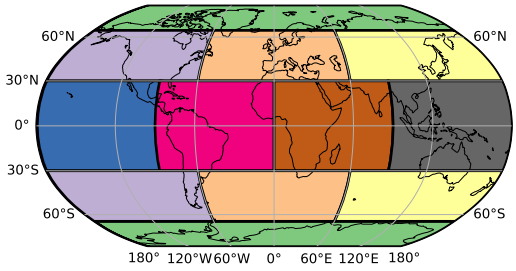

Figure A1. Definition of the 12 regions of the Earth used to determine weather types.

$$G_N(D,H,p) = \frac{1}{20}\sum_{k=-9}^{10}\frac{1}{n_y}\sum_{t\in\Gamma(D+k,H)}\delta(t,p)$$

where :

- n_y is the length in years of the nudged simulation
- $\Gamma(D,H) = \{t \mid d(t) = D,\ h(t) = H\}$ where $d(t)$ (resp. $h(t)$) is the day of the year (resp. hour of the day) associated to time step t.

A2 State conditional bias correction

The Earth is divided in 12 regions, as shown by Fig. A1, each treated independently. Based on the 500 hPa geopotential daily field, three weather types per month and per region are defined using a neural-network-based tool called self-organizing maps (SOMs) (Kohonen, 2013). Within a region R, nudging increments of a month M are clustered according to the weather type identified at the moment their were added. A BC "anomaly" is then constructed for each month and each weather type within a region as the difference between the mean of nudging increments of month M falling into cluster W and the mean of all nudging increments of month M :

$$g_{\text{cond}}(M,W) = \frac{1}{n_{M,W}}\sum_{t\in M\cap W}\delta(t) - \frac{1}{n_M}\sum_{t\in M}\delta(t)$$

where n_M and $n_{M,W}$ are the number of increments respectively in the whole month M and in the cluster of weather type W of month M.

In the BC simulation, before each addition of the correcting increments, the weather type is identified and the BC term is computed accordingly. For a given time t, corresponding to the day of the year $d(t)$, the hour of the day $h(t)$ and the month $m(t)$, and a grid-point p, the weather type $w(t,p)$ being identified, the BC term is given by:

$$G_N(d(t),h(t),p) + g_{\text{cond}}(m(t),w(t,p))$$

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

A3 Climate Adaptative Bias Correction

The CABCOR method requires a long adaptation simulation with increments updated each year of simulation. We implemented the procedure as follows:

- The first year, the models runs with no modification.
- At the end of year n, the model climatology is computed: $\overline{X}^{C_n}$
- During year $n+1$, a tendency adjustment given by:

$$C_{n+1} = \frac{X_{\mathrm{r}} - \overline{X}^{C_n}}{\tau}$$

  is added to the prognostic equation of the targeted variables.

X_{r} is the reference climatology, in this case the multi-year monthly mean of ERA5. The model climatology at the end of year n is computed as the multi-year monthly mean of the past n years. During the first 20 years of adaptation simulation, $\frac{1}{\tau}$ evolves linearly from $\frac{1}{1\times 104}\mathrm{day}^{-1}$ to its target value.

This adaptation simulation was run with repeated 1991 AMIP forcings. It was run 110 years and BC increments were given by C_{30}, C_{70}, C_{110}.

Appendix B: Additional figures

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

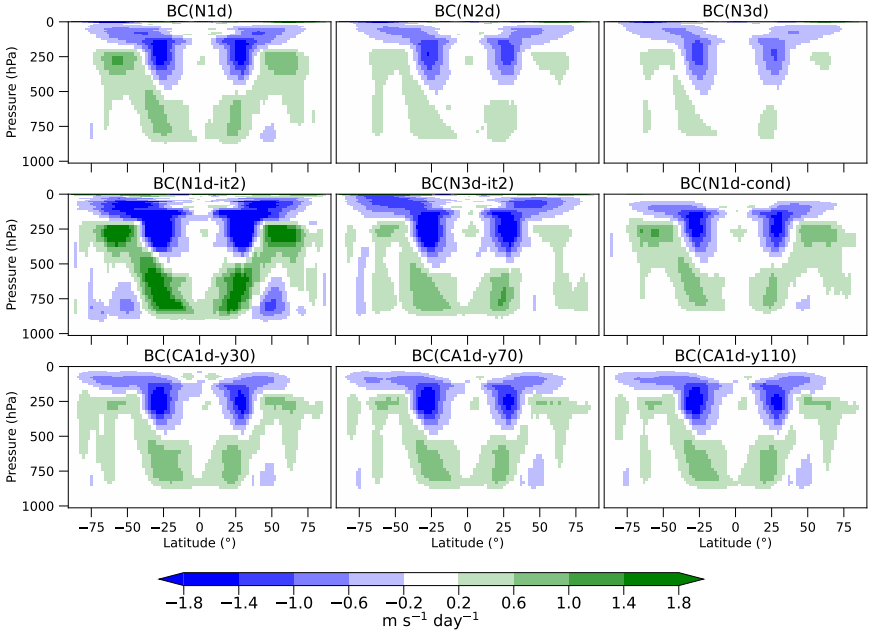

Figure B1. Annual zonal mean of zonal wind correction terms.

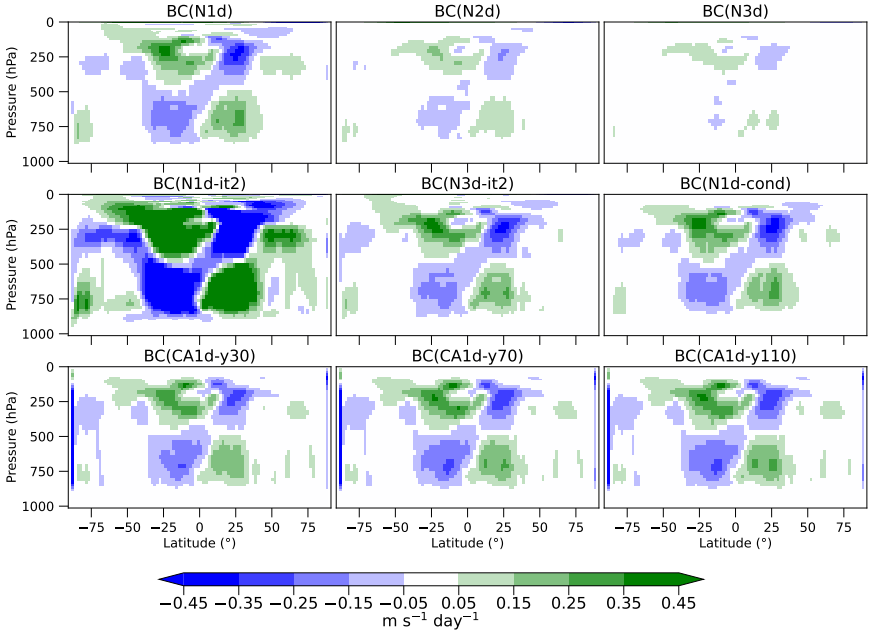

Figure B2. Annual zonal mean of meridional wind correction terms.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

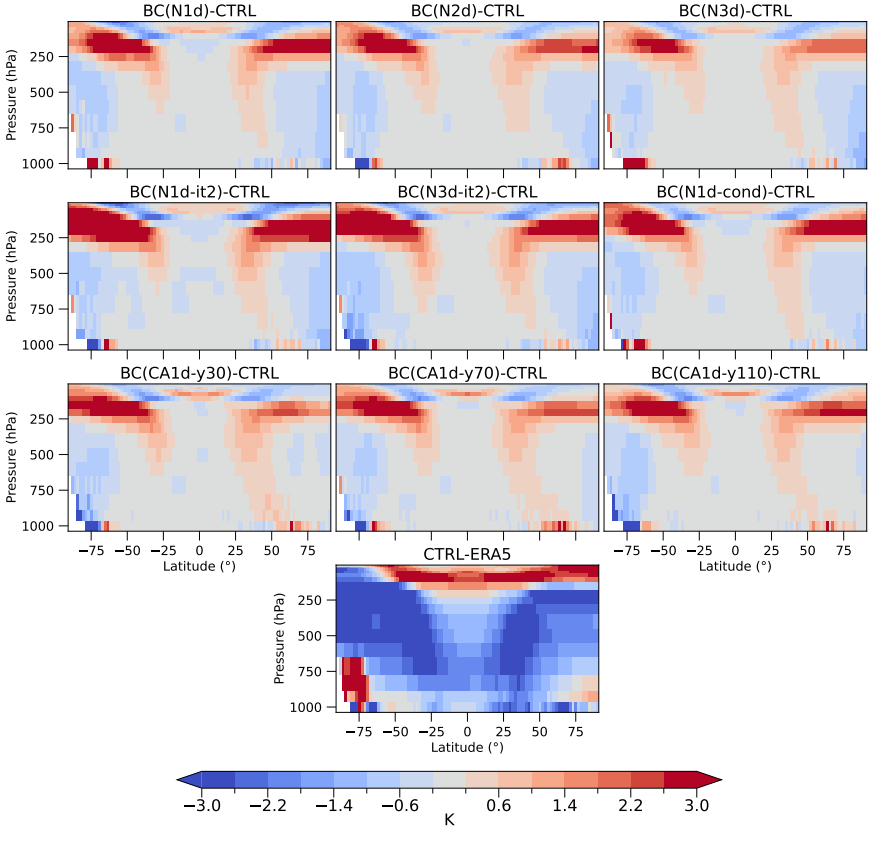

Figure B3. Effect of BC on mean zonal temperature. Only the lowest panel represents temperature bias with respect to ERA5 in the free simulation. The rest shows the difference with the free simulation.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

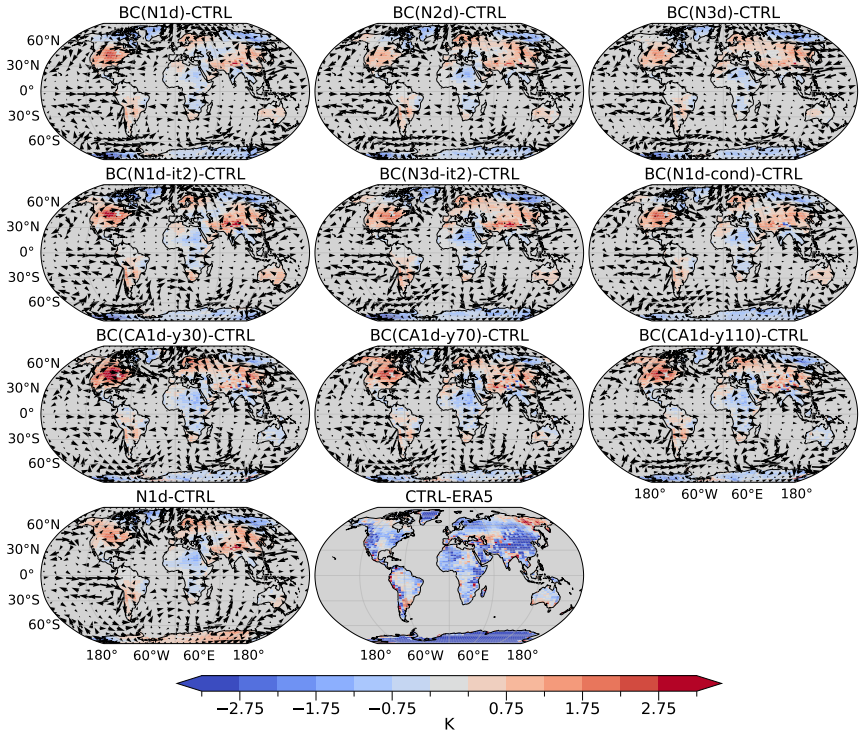

Figure B4. Effect of BC on $T_{2m,land}$. Except for the free simulation, maps show the difference in $T_{2m,land}$ (shading) and the difference in wind vector at 500 hPa (arrows) between BC simulation and free simulation. Only one out of three arrows are drawn. Arrows lengths are proportional to the norm of the vector represented. For the CTRL simulation, the t2m bias with respect to ERA5 is represented.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

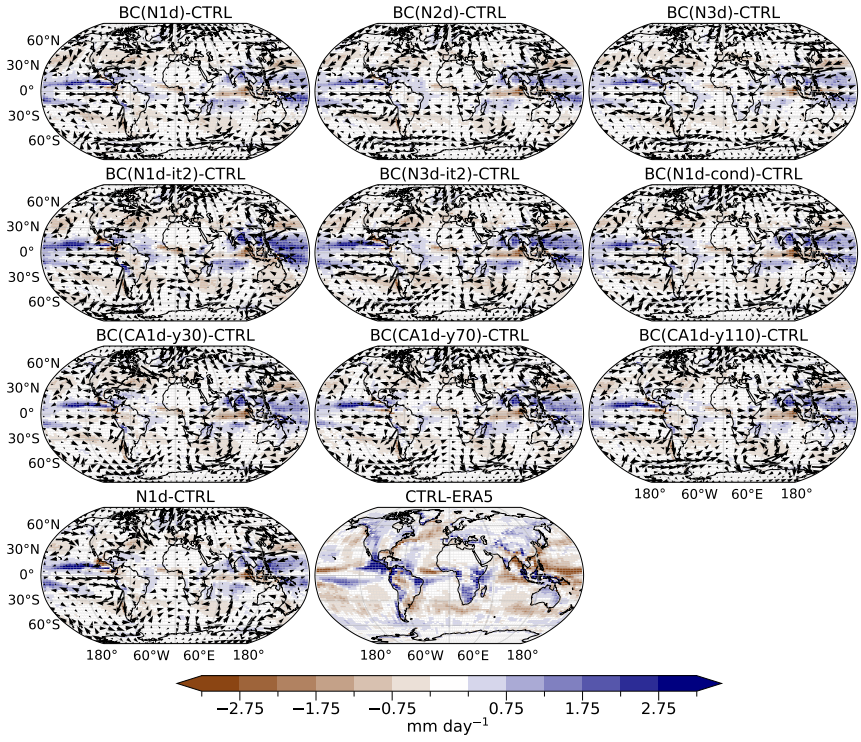

Figure B5. Effect of BC on precipitation. Except for the free simulation, maps show the difference in precipitation (shading) and the difference in wind vector at 500 hPa (arrows) between BC simulation and free simulation. Only one out of three arrows are drawn. Arrows lengths are proportional to the norm of the vector represented. For the CTRl simulation, the t2m bias with respect to ERA5 is represented.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

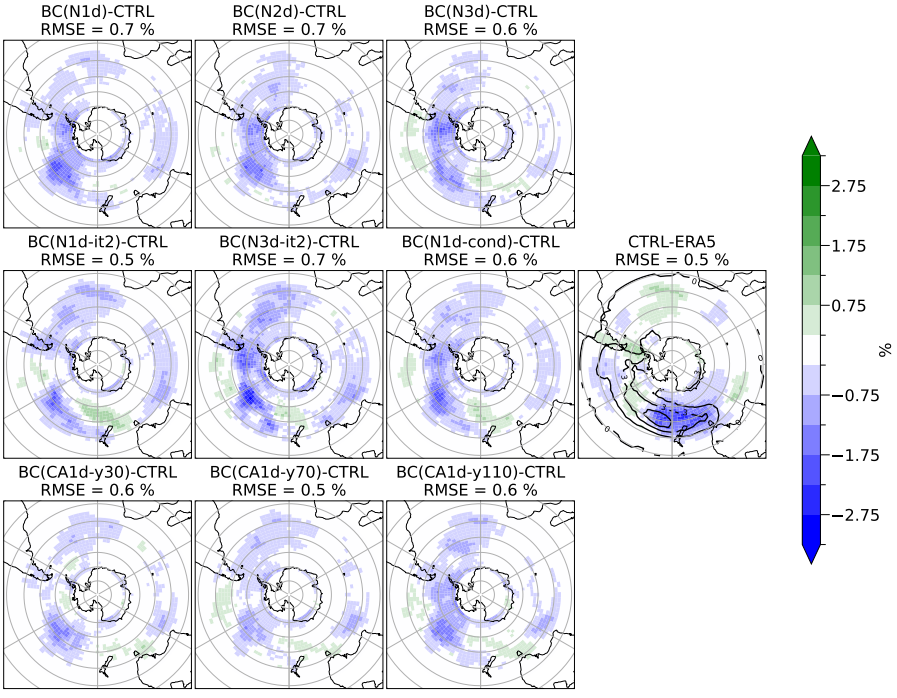

Figure B6. Effect of BC on annual mean blocking frequency in the Southern Hemisphere. Aside from the extreme right one, panels show the difference in annual mean frequency of blockings between BC simulations and the free simulations. The extreme right panel shows the annual mean bias in frequency of blockings of a CTRL simulation with respect to ERA5 as well as ERA5' frequency represented by black contours. RMSEs indicated on top of each panel correspond to the RMSE of the simulation with respect to ERA5.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

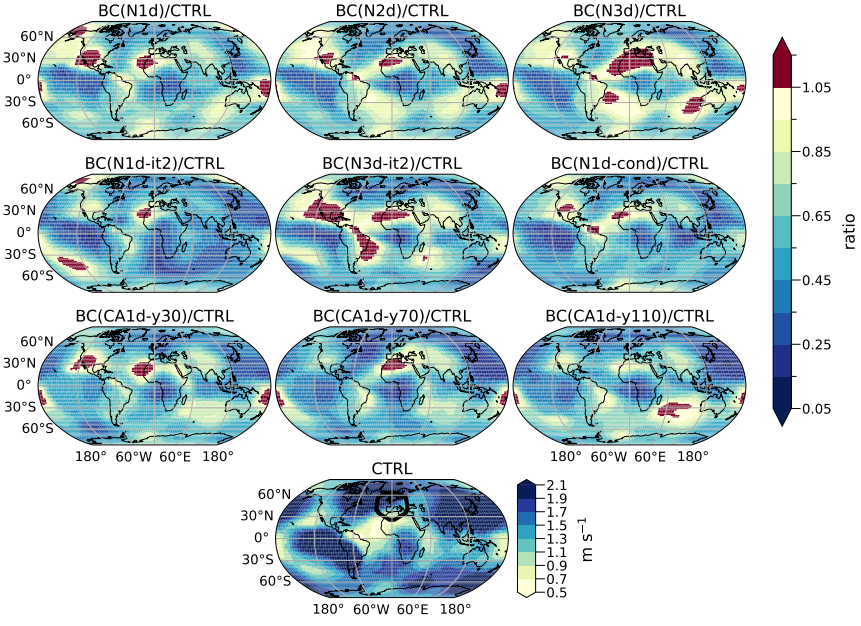

Figure B7. Ratio of moving RMSE of wind at 500 hPa of BC simulations over the CTRL one (top 6 panels) and moving RMSE of wind at 500 hPa of a CTRL simulation (lower panel). RMSE in a grid point corresponds to the RMSE within the circle with a radius of 2,000 km around it. A circle centered at 0°W-45°N, with radius 2,000 km is drawn on the lowest panel.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

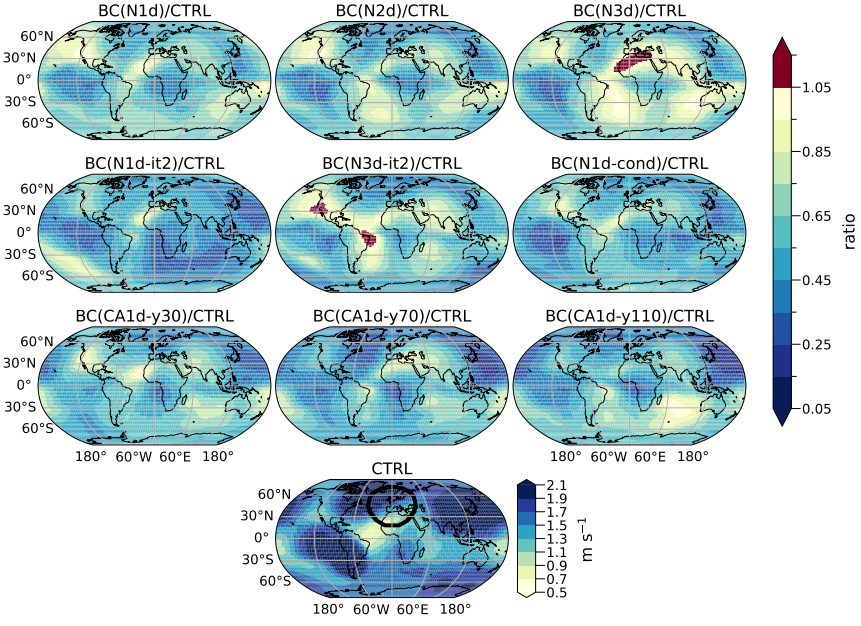

Figure B8. Ratio of moving RMSE of wind at 500 hPa of BC simulations over the CTRL one (top 6 panels) and moving RMSE of wind at 500 hPa of a CTRL simulation (lower panel). RMSE in a grid point corresponds to the RMSE within the circle with a radius of 3,000 km around it. A circle centered at 0°W-45°N, with radius 3,000 km is drawn on the lowest panel.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Author contributions. AC, GK and JB designed the study and contributed to the analysis and writing. GK and AC carried out the simulations. AC wrote the first draft of this article.

Competing interests. The authors declare no competing interests

Acknowledgements. This study has received funding from Agence Nationale de la Recherche - France 2030 as part of the PEPR TRACCS programme under grant numbers ANR-22-EXTR-0005, ANR-22-EXTR-0008, ANR-22-EXTR-0010, ANR-22-EXTR-0011. It was provided with HPC computing and storage resources by GENCI at TGCC under the grants 2024-A0180116219 and 2024-AD010101523R on the Joliot Curie supercomputer (SKL and ROME partitions). A. Champouillon acknowledges the French Ministry for the Environmental Transition for funding her PhD.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

References

Addor, N., Rohrer, M., Furrer, R., and Seibert, J.: Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment, Journal of Geophysical Research: Atmospheres, 121, 2075–2089, https://doi.org/10.1002/2015JD024040, 2016.

Balhane, S., Cheruy, F., Driouech, F., El Rhaz, K., Idelkadi, A., Sima, A., Vignon, É., Drobinski, P., and Chehbouni, A.: Towards an advanced representation of precipitation over Morocco in a global climate model with resolution enhancement and empirical run-time bias corrections, International Journal of Climatology, 44, 1691–1709, https://doi.org/10.1002/joc.8405, 2024.

Beaumet, J., Déqué, M., Krinner, G., Agosta, C., Alias, A., and Favier, V.: Significant additional Antarctic warming in atmospheric bias-corrected ARPEGE projections with respect to control run, The Cryosphere, 15, 3615–3635, https://doi.org/10.5194/tc-15-3615-2021, 2021.

Champouillon, A.: Intercomparison of run-time bias correction methods in LMDZ v6.3 - LMDZ configuration files, output data and analysis scripts, Zenodo Dataset, https://doi.org/10.5281/zenodo.18955446, 2026a.

Champouillon, G.: Intercomparison of run-time bias correction methods in LMDZ v6.3 - model infrastructure and code, Zenodo Dataset, https://doi.org/10.5281/zenodo.20410139, 2026b.

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

Chapman, W. E. and Berner, J.: Deterministic and stochastic tendency adjustments derived from data assimilation and nudging, Quarterly Journal of the Royal Meteorological Society, 150, 1420–1446, https://doi.org/10.1002/qj.4652, 2024.

Chapman, W. E. and Berner, J.: Improving Climate Bias and Variability via CNN-Based State-Dependent Model-Error Corrections, Geophysical Research Letters, 52, e2024GL114 106, https://doi.org/10.1029/2024GL114106, 2025.

Davini, P. and D'Andrea, F.: From CMIP3 to CMIP6: Northern Hemisphere Atmospheric Blocking Simulation in Present and Future Climate, Journal of Climate, 33, 10 021–10 038, https://doi.org/10.1175/JCLI-D-19-0862.1, 2020.

Davini, P., Cagnazzo, C., Gualdi, S., and Navarra, A.: Bidimensional Diagnostics, Variability, and Trends of Northern Hemisphere Blocking, Journal of Climate, 25, 6496–6509, https://doi.org/10.1175/JCLI-D-12-00032.1, 2012.

Dinh, T. L. A. and Aires, F.: Revisiting the bias correction of climate models for impact studies, Climatic Change, 176, 140, https://doi.org/10.1007/s10584-023-03597-y, 2023.

D'Andrea, F. and Vautard, R.: Reducing systematic errors by empirically correcting model errors, Tellus, 52, https://doi.org/10.3402/tellusa.v52i1.12251, 2000.

Eyring, V., Gillett, N., Achuta Rao, K., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., Cassou, C., Durack, P., Kosaka, Y., McGregor, S., Min, S., Morgenstern, O., and Sun, Y.: Human Influence on the Climate System, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., et al., book section 3, pp. 423–551, Cambridge University Press, Cambridge, UK and New York, NY, USA, https://doi.org/10.1017/9781009157896.005, 2021.

Fernandez-Granja, J. A., Casanueva, A., Bedia, J., and Fernandez, J.: Improved atmospheric circulation over Europe by the new generation of CMIP6 earth system models, Climate Dynamics, 56, 3527–3540, https://doi.org/10.1007/s00382-021-05652-9, 2021.

Guldberg, A., Kaas, E., Déqué, M., Yang, S., and Vester Thorsen, S.: Reduction of systematic errors by empirical model correction: impact on seasonal prediction skill, Tellus A: Dynamic Meteorology and Oceanography, 57, 575–588, https://doi.org/10.3402/tellusa.v57i4.14707, 2005.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Hall, A.: Projecting regional change, Science, 346, 1461–1462, https://doi.org/10.1126/science.aaa0629, 2014.

Harvey, B. J., Cook, P., Shaffrey, L. C., and Schiemann, R.: The Response of the Northern Hemisphere Storm Tracks and Jet Streams to Climate Change in the CMIP3, CMIP5, and CMIP6 Climate Models, Journal of Geophysical Research: Atmospheres, 125, e2020JD032 701, https://doi.org/10.1029/2020JD032701, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., C., S., Dee, D., and Thépaut, J.-N.: ERA5 monthly averaged data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.6860a573, 2023.

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L., Foujols, M.-A., Mellul, L., Traore, A.-K., Dufresne, J.-L., Boucher, O., Lefebvre, M.-P., Millour, E., Vignon, E., Jouhaud, J., Diallo, F. B., Lott, F., Gastineau, G., Caubel, A., Meurdesoif, Y., and Ghattas, J.: LMDZ6A: The Atmospheric Component of the IPSL Climate Model With Improved and Better Tuned Physics, Journal of Advances in Modeling Earth Systems, 12, e2019MS001 892, https://doi.org/10.1029/2019MS001892, 2020a.

Hourdin, F., Rio, C., Jam, A., Traore, A.-K., and Musat, I.: Convective boundary layer control of the sea surface temperature in the tropics, Journal of Advances in Modeling Earth Systems, 12, e2019MS001 988, https://doi.org/10.1029/2019MS001988, 2020b.

Huffman, G. J., Behrangi, A., Bolvin, D., and Nelkin, E.: GPCP Version 3.3 Satellite-Gauge (SG) Combined Precipitation Data Set,, https://doi.org/10.5067/MEASURES/GPCP/DATA306, 2024.

Kharin, V. V. and Scinocca, J. F.: The impact of model fidelity on seasonal predictive skill, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL052815, 2012.

Kleiner, N., Chan, P. W., Wang, L., Ma, D., and Kuang, Z.: Effects of Climate Model Mean-State Bias on Blocking Underestimation, Geophysical Research Letters, 48, e2021GL094 129, https://doi.org/10.1029/2021GL094129, 2021.

Kohonen, T.: Essentials of the self-organizing map, Neural Networks, 37, 52–65, https://doi.org/10.1016/j.neunet.2012.09.018, 2013.

Koskentausta, J., Karpechko, A. Y., Köhler, R. H., Levine, X., Wijngaard, R. R., and Sinclair, V. A.: The Role of Model Biases in the Simulated North Atlantic Jet Stream Response to Global Warming, Journal of Climate, 38, 5007–5023, https://doi.org/10.1175/JCLI-D-24-0648.1, 2025.

Krinner, G. and Flanner, M. G.: Striking stationarity of large-scale climate model bias patterns under strong climate change, Proceedings of the National Academy of Sciences, 115, 9462–9466, https://doi.org/10.1073/pnas.1807912115, 2018.

Krinner, G., Beaumet, J., Favier, V., Déqué, 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, https://doi.org/10.1029/2018MS001438, 2019.

Krinner, G., Kharin, V., Roehrig, R., Scinocca, J., and Codron, F.: Historically-based run-time bias corrections substantially improve model projections of 100 years of future climate change, Communications Earth & Environment, 1, 1–7, https://doi.org/10.1038/s43247-020-00035-0, 2020.

Krinner, G., Champouillon, A., Blanchet, J., and Chéruy, F.: Iterative run-time bias corrections in an atmospheric GCM (LMDZ v6.3), EGUsphere, pp. 1–20, https://doi.org/10.5194/egusphere-2025-3553, 2025.

Labonté, M.-P., Matte, D., Paquin, D., Scinocca, J. F., Kharin, V. V., and Jiao, Y.: Runtime bias corrected driving data for regional climate models: regional-scale impacts, Climate Dynamics, 63, 406, https://doi.org/10.1007/s00382-025-07911-5, 2025.

https://doi.org/10.5194/egusphere-2026-1380
Preprint. Discussion started: 29 May 2026

Lenggenhager, S. and Martius, O.: Atmospheric blocks modulate the odds of heavy precipitation events in Europe, Climate Dynamics, 53, 4155–4171, https://doi.org/10.1007/s00382-019-04779-0, 2019.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nature Climate Change, 7, 764–773, https://doi.org/10.1038/nclimate3418, 2017.

Menapace, A., Dhawan, P., Dalla Torre, D., Kaffas, K., Crespi, A., Larcher, M., Righetti, M., and Cannon, A. J.: Review of bias correction methods for climate model outputs in hydrology, Journal of Hydrology, 660, 133 213, https://doi.org/10.1016/j.jhydrol.2025.133213, 2025.

Mohr, S., Wandel, J., Lenggenhager, S., and Martius, O.: Relationship between atmospheric blocking and warm-season thunderstorms over western and central Europe, Quarterly Journal of the Royal Meteorological Society, 145, 3040–3056, https://doi.org/10.1002/qj.3603, 2019.

Muñoz Sabater, J., Comyn-Platt, E., Hersbach, H., B., B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., Cagnazo, C., and Cucchi, M.: ERA5-land post-processed daily-statistics from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.e9c9c792, 2024.

Priestley, M. D. K., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., and Lee, R. W.: An Overview of the Extratropical Storm Tracks in CMIP6 Historical Simulations, Journal of Climate, 33, 6315–6343, 2020.

Scinocca, J. F. and Kharin, V. V.: Climatological Adaptive Bias Correction of Climate Models, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 563, https://doi.org/10.1029/2024MS004563, 2024.

Scinocca, J. F., Kharin, V. V., Matte, D., Jiao, Y., Labonté, M.-P., Qian, M., Paquin, D., Akingunola, A., and Lazare, M.: Runtime bias correction of regional climate model driving data and its continental-scale impacts, Climate Dynamics, 63, 462, https://doi.org/10.1007/s00382-025-07814-5, 2025.

Verfaillie, D., Déqué, M., Morin, S., and Lafaysse, M.: The method ADAMONT v1.0 for statistical adjustment of climate projections applicable to energy balance land surface models, Geoscientific Model Development, 10, 4257–4283, https://doi.org/10.5194/gmd-10-4257-2017, 2017.

Watt-Meyer, O., Brenowitz, N. D., Clark, S. K., Henn, B., Kwa, A., McGibbon, J., Perkins, W. A., and Bretherton, C. S.: Correcting Weather and Climate Models by Machine Learning Nudged Historical Simulations, Geophysical Research Letters, 48, e2021GL092 555, https://doi.org/10.1029/2021GL092555, 2021.