Simulating the variations of carbon dioxide in the global atmosphere
on the hexagonal grid of DYNAMICO coupled with the LMDZ6 model

Zoé Lloret¹, Frédéric Chevallier¹, Anne Cozic¹, Marine Remaud¹*, Yann Meurdesoif²

¹Laboratoire des Sciences du Climat et de l’Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
²currently at Faculty of Science, A-LIFE, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, the Netherlands

Correspondence to: Zoé Lloret (zoe.lloret@lsce.ipsl.fr)

Abstract. Efforts to monitor the emissions and absorptions of atmospheric carbon dioxide (CO₂) over the globe and to understand their varying regional patterns with greater accuracy have intensified in recent years. This study evaluates the performance of a new model coupling, ICO, built around the Laboratoire de Météorologie Dynamique atmospheric general circulation model (LMDZ) for simulating CO₂ transport. ICO utilizes the new icosahedral hydrostatic dynamical core called Dynamico running on an unstructured grid, which enables potential improvements in spatial resolution at the Equator while removing artificial distortions and numerical filters at the poles. Comparisons with a reference configuration using a structured latitude-longitude grid reveal that ICO well captures seasonal variations in CO₂ concentrations at surface stations. While not significantly enhancing the capture of complex seasonal patterns, ICO maintains comparable accuracy. Both configurations exhibit similar vertical CO₂ concentration profiles and display a consistent bias in the lower stratosphere relative to observational data. ICO demonstrates advantages in computational efficiency and storage, thanks to its reduced cell count per level and a homogeneous grid structure. It holds promise for future developments, including with the LMDZ offline model and associated inversion system, which contribute to the Copernicus Atmosphere Monitoring Service. Overall, the ICO configuration showcases the efficacy of utilizing an unstructured grid for the physics, and the capability of Dynamico in accurately simulating CO₂ transport. This study emphasizes the importance of advanced modeling approaches and high-resolution innovative grids in enhancing our understanding of the global carbon cycle and refining climate models.

1 Introduction

The key role of carbon dioxide (CO₂) in climate change has motivated increasing efforts in recent decades to monitor its variations in the global atmosphere. Sources and sinks of this trace gas are found primarily on the Earth’s surface. They induce the highest CO₂ gradients in the boundary layer, for example around anthropogenic emission hotspots, while their direct influences gradually mix over time at all altitudes to contribute to the overall CO₂ background. The distribution of CO₂ in the atmosphere therefore spans a wide range of spatial and temporal scales, mainly combining influences from surface sources, surface sinks and meteorology. This complexity is sampled by growing high-quality observation networks on the ground, in the atmosphere (aircraft, balloons, drones) and in space (e.g., Ciais et al., 2014; Crisp et al., 2018). It is also simulated, more or less well, by Atmospheric General Circulation Models (GCMs) and dedicated tracer transport models (e.g., Remaud et al., 2018; Basu et al., 2018; Agustí-Panareda et al., 2022). Many uncertainties in the model input data (boundary conditions, meteorology) and the model equations (advection schemes, subgrid parameterizations) still limit these simulations. However, there is a strong incentive towards higher spatial resolutions in order to benefit from an increased realism for orography, coastlines, and known emission or absorption hot-spots, and to reduce any artificial smoothing of the 3D fields (Agustí-Panareda et al., 2019).
However, this wish is tempered by the need to carry out long simulations of this long-lived tracer, typically several years, that may be massively repeated in the case of inverse modeling. Increasing the resolution without affecting the time-to-solution leads to revisiting the numerical efficiency of models in order to gain computing time margins. Porting codes on Graphical Processing Units (GPUs) may largely contribute to this effort (Chevallier et al., 2023), but not enough to close the gap with, e.g., the kilometric resolution of the current space-borne observations. In particular, models running on a regular longitude-latitude grid face scaling limitations due to advection at the poles requiring significant data communication to solve the problem of resolution clustering. This data exchange can create a computing bottleneck on supercomputers using large amounts of processors (Staniforth & Thuburn, 2012). In this paper, we are addressing this specific issue for the simulation of CO₂ using an unstructured quasi-uniform grid made of non-quadrilateral grid cells. Such a solution has been explored by few models so far, either for use in Earth system models or directly for atmospheric inversion (Niwa et al., 2017; Giorgetta et al., 2018; Sakaguchi et al., 2020).

We build on the dynamical core Dynamico (Dubos et al., 2015), which has recently been integrated into LMDZ which is the GCM (Hourdin et al., 2020) of the Laboratoire de Météorologie Dynamique. LMDZ has been used as the atmospheric component of the Institut Pierre-Simon-Laplace (IPSL) Earth system model (Sepulchre et al., 2020) and for the Climate Model Intercomparison Project (CMIP) with its traditional regular longitude-latitude grid. For comparisons with real observations, e.g., for inverse modeling, it is nudged to horizontal wind fields obtained from a numerical weather forecast reanalysis.

This paper evaluates the ability of this new coupling of the LMDZ GCM using the Dynamico dynamical core to transport a long-lived tracer like CO₂. Dynamico has been extensively compared to other dynamical cores (Ullrich et al., 2017), but the same has not been done yet for its coupled configurations. Coupling it to the LMDZ GCM also represents the first step towards the use of Dynamico for inverse modeling. We compare it to the previous equivalent version running on a regular longitude-latitude grid and to various observations of CO₂ mole fractions over a 40-year period.

Section 2 describes the two configurations of our GCM, the experiments we ran to compare them, and the method for our study. Section 3 presents the results of the direct comparison between our models and the observations. Section 4 concludes the study.

2 Presentation of the model and experiments
2.1 Model description
2.1.1 LMDZORINCA - Regular longitude-latitude configuration
Our general circulation model configuration consists of the coupling between the LMDZ model of Hourdin et al. (2013, 2020) itself, an aerosol and reactive chemistry model called INteractions between Chemistry and Aerosols (INCA, Hauglustaine, et al., 2004) and the Organizing Carbon and Hydrology in Dynamic Ecosystems land surface model (ORCHIDEE, Krinner et al., 2005). ORCHIDEE simulates the water and energy exchanges between the soil and the atmosphere, but yearly land cover maps were used here instead of simulating vegetation dynamics. In the following, we will refer to this LMDZ-INCA-ORCHIDEE coupled model on the regular latitude-longitude grid as REG for simplicity.

Tracers, such as CO₂, are modeled by INCA, and their transport is calculated and synced with the LMDZ GCM physics timestep every 15 minutes. Chemical processes are also calculated every 15 minutes by computing differential equations to update the atmospheric concentration fields of each cell. Using tracers from INCA instead of only having them in LMDZ allows interaction between chemical reactions and the tracer transport process, which is crucial for some tracers such as CH₄, although it has no impact on CO₂.
The latest version of LMDZ physics is described in Hourdin et al. (2020). Most notably for tracer transport, dry and cloudy shallow convection is separated from deep convection. Shallow convection is unified and combines the Mellor and Yamada (1974) diffusive approach for small-scale turbulence with a thermal plume model (Rio & Hourdin, 2008) for the boundary layer. Deep convection uses a modified version of the mass-flux formulation of Emanuel (1991) (Grandpeix et al., 2004, Rochetin et al., 2014). Longwave radiation is modeled using the Rapid Radiation Transfer Model (RRTM; Mlawer et al., 1997), and shortwave radiation uses a 6-band code derived from Fouquart and Bonnel (1980).

The dynamical core of LMDZ is a mix of a finite difference and finite volume discretization on the sphere of the primitive equations of meteorology and of transport equations (Hourdin et al., 2006, 2013). Water and other tracers are advected with a scheme from Van Leer (1997), and angular momentum is conserved numerically. This full configuration was previously evaluated for CO₂ transport by Remaud et al. (2018).

The dynamical core of LMDZ is parallelized in latitude using distributed memory with the Message Passing Interface standard (MPI) and in the vertical with shared memory using the Open Multi-Processing interface (OpenMP). A longitudinal filter near the poles avoids the use of very small time steps, but limits the efficiency of any parallelism along the longitudes.

The parallelization of the physical parameterizations within LMDZ follows a different approach. It utilizes a combination of MPI and OpenMP processes with shared memory by splitting a single vector that runs through the entire horizontal grid into independent domains. This is possible due to the fundamental 1-D nature of the LMDZ physical parameterizations that only compute vertical transfers. The performance of the model is optimized by using domain decomposition parallelism on the horizontal layer with MPI and shared memory parallelism with OpenMP. For LMDZ at our resolution, the optimal compromise between resources and performance is achieved by using 71 MPI processes with 8 OpenMP threads running on 568 cores (Hourdin, 2020).

2.1.2 ICOLMDZORINCA - unstructured grid configuration

ICOLMDZORINCA is a novel configuration of our coupled model that integrates the previously described ORCHIDEE land surface model, the INCA chemistry model, and the physics module of the LMDZ model. The previous dynamical core in LMDZ has been replaced by a new one, known as DYNAMICO (Dubos et al., 2015), which operates on a quasi-uniform icosahedral C-grid for its horizontal mesh. In the following, we will refer to this DYNAMICO-LMDZ-INCA-ORCHIDEE coupled model as ICO for simplicity.

The hydrostatic and shallow-atmosphere non-hydrostatic Euler equations can be solved using the DYNAMICO dynamical core (Ullrich et al., 2017). The mesh is based on a tessellation of the sphere into triangles, which when joined, creates the primal hexagonal-pentagonal mesh. A quasi-uniform grid avoids any singularity at the poles, thereby improving the load balancing on parallel computers. By construction, this grid has a coarser resolution than a regular longitude-latitude grid in the high latitudes, even when accounting for the polar filter (Herrington et al., preprint). Figure 1 provides an example of a visual representation of the icosahedral C-grid.

Kinematics and dynamics were separated as much as possible so that transport equations do not use any information from the momentum equations. The kinematics handle the transport of mass, potential temperatures, and tracers using the mass fluxes computed by the dynamics. The vertical transport uses a slope-limited Van Leer’s scheme (Van Leer, 1977) and does not differ
from the vertical transport of LMDZ. The fully discrete finite volume horizontal advection scheme is described in Dubey et al. (2015). It uses a flux-corrected transport approach to stay positive-definite rather than slope limiters.

To achieve efficient parallelism in the horizontal dimension, the ICO configuration partitions the mesh into rhombi, whose sides pass through the centers of some of the hexagons. The hexagons covered by a rhombus are processed together and the rhombi can be processed in parallel. This parallelization strategy is implemented with a combination of OpenMP and MPI.

The vertical parallelization is identical to the one in the REG configuration.

Figure 1: Unstructured grid of the ICO configuration used here.

2.2 Description of the simulations

For each configuration (REG and ICO), we have run a simulation from 1979 to 2020. The first year is used for spin-up and is not analyzed. In both configurations, the model was not let free, but its large-scale atmospheric circulation was kept in the vicinity of the observed one by nudging its 6-hourly horizontal winds toward the ERA5 reanalysis (Hersbach et al., 2020) with a relaxation time of 3 hours. The nudging drives the large-scale atmospheric circulation of the model. Initial atmospheric CO₂ mixing ratios values were set using the Copernicus Atmosphere Monitoring Service (CAMS) atmospheric inversion, version 20r2 (Chevallier et al., 2005; https://atmosphere.copernicus.eu/, access 31 May 2023). This same product prescribed the CO₂ surface fluxes every 3 hours. These surface fluxes carry some imprint from the REG model with a regular grid since CAMS uses an older REG model version at coarser spatial resolution. Still, after Remaud et al. (2018) who tested a distinct set of surface fluxes for their model evaluation within a similar framework, we consider that this imprint hardly affects our conclusions.

The boundary files used for the two simulations were identical. However, for the simulation running on ICO, the boundary files were either interpolated or recreated onto the new grid ahead of time to fit the unstructured grid or interpolated during execution.
The initial total mass of CO$_2$ in the atmosphere had a difference of only 0.01% between the two simulations because of these operations.

We had an hourly model output for all variables. This high frequency output was chosen in order to well assess the differences in synoptic variability of tracer transport between our two model configurations.

We ran REG on a horizontal grid of 144 points in longitude $\times$ 143 grid points in latitude, which corresponds to a resolution of 2.5° in longitude and 1.27° in latitude, equivalent to 278 km by 140 km at the equator. We use 79 vertical layers going up to 80 km in altitude, with around 25 layers dedicated to the first 2 km. The complete grid configuration is described in more detail in Hourdin et al. (2020).

We ran ICO on a horizontal grid of 16002 cells, and the same 79 vertical layers. This gives an horizontal resolution at the equator of around 2.5° in longitude and 1.25° in latitude, each cell has the area of an hexagon of side 110 km in order to have similar resolution at the equator to the equivalent longitude-latitude grid from REG. With this setup, ICO has 22% less cells than REG.

2.3 Observational data

To compare our simulated tracer concentrations to observations, we sampled the concentration fields at the nearest cell center, model level and timestamp for each data point. We used the high-quality measurements of the CO$_2$ GLOBALVIEWplus v8.0_2022-08-27 ObsPack database (Schuldt et al., 2022).

In this dataset, observations were calibrated according to the WMO CO$_2$ X2019 scale (Hall et al., 2021). Like for inverse modeling with LMDZ (Chevallier et al., 2010), only afternoon non flagged data from 12:00 to 16:00 local time were selected for continuous in-situ surface stations under 1000 m above sea level (a.s.l.), and only night time data from 00:00 to 4:00 local time were kept for in-situ stations above 1000 m a.s.l. All flask data, and all upper-air data (aircraft data and AirCore measurements) were kept. This selection accounts for the usual failure of transport models to well represent the accumulation of tracers at low altitude during the night as well as the inability to model the phenomenon in mountain stations where air masses are advected during daytime through updrafts on the sun-exposed slopes (Geels et al., 2007).

We divided the observations into three groups: surface in situ and flask data, aircraft observations and observations from AirCore flights. We used the aircraft measurements and AirCore data to obtain vertical profiles of CO$_2$ concentration. AirCore (Karion et al. 2010) is an atmospheric sampling system consisting of an open ended steel tube launched from an aerial platform and that collects many successive samples of the ambient air when descending. For surface data, 107 stations have been selected from the Obspack dataset out of the original 222 stations. Surface stations with less than 5000 measurement points that passed the initial data selection described above over the entire duration of the study were excluded from the analysis. For aircrafts, we have selected 33 sites and campaigns out of a possible 51, only keeping those with more than 2000 measurement points. For the AirCore data, we kept all observations. The full list of sites and datasets used is presented in Table 1 and Table 2 as a supplement.

The uncertainty of the reference CO$_2$ mixing ratio measurements used here is on the order of 0.1 ppm (see, e.g., Crotwell et al., 2020, for the systematic errors and Hazan et al., 2016, for the standard deviation). It is negligible compared to the model uncertainty due to transport error which is on the order of 1 ppm under 3000 m (Lauvaux et al., 2009) and is not further discussed in the following. Collection altitude determination error from AirCore measurements can be high and depends on the
altitude, and is on the order of 250 m below 20 km and up to 1 km above that altitude (Wagenhäuser et al., 2021). We discuss the potential impact of these uncertainties on our model evaluation in section 3.5.2.

2.4 Evaluation methodology

2.4.1 Surface stations

For surface stations with continuous measurements, we used a curve-fitting method on both the model and observations CO₂ mixing ratios time series to extract the annual mean, the seasonal cycles and the synoptic variations. A smoothed function consisting of a second-order polynomial and eight harmonics was used to fit the time series over the 19080-2020 period. The polynomials were used to calculate the annual trend and growth rate, while the harmonics were used to get the seasonal cycle. The synoptic variations are obtained from the difference between the raw data and the fitted smooth curve.

To evaluate the two model configurations performance between each other and compared to observations we use metrics which we will describe in the following subsections.

2.4.2 Annual gradient between stations

We use the measurements from South Pole station (SPO), which is far from any major CO₂ source or sink, to validate the simulated background growth rate of CO₂ concentrations. Then, we study the cross-site gradients by calculating the yearly growth rate at each site relative to SPO. To do so, we average the annual growth rate of the CO₂ concentration over the 1980-2020 period for each site and subtract the value at SPO. Comparing the observed and modeled values of this variable informs us on both the growth rate of the CO₂ concentration at each site, and on concentration gradients of our transport model which are key for use in an inverse system. To study the interannual variation of these growth rates, we calculate their standard deviation for both measurements and models. We normalize the average model’s standard deviation by dividing it by the measurement standard deviation. This gives us information on how well the model captures the magnitude and direction of these variations.

We compute the yearly growth rate for each year of the 1980-2020 period using the smooth curved fit described above, before averaging it. To evaluate this variable, we then look at the mean bias and the root-mean-square error (RMSE) of the CO₂ concentration gradient for each station relative to SPO.

2.4.3 Seasonal cycle

We evaluate the capacity of our model to represent the CO₂ seasonal cycle by comparing the phase and amplitude of the harmonics of their smoothed fitted curve to the one of the measurements at each station. At each measurement site we calculated the Pearson correlation coefficient between measurements and model time series to evaluate the phase of the seasonal cycle. And we evaluated the amplitude of the seasonal cycle by looking at the ratio between peak-to-peak amplitudes of the harmonics. We normalized this variable by dividing the values of the model’s seasonal cycle peak-to-peak amplitude at each station by the ones from the observations.

2.4.4 Synoptic variability

To evaluate our model ability to represent the phase of the CO₂ synoptic variability we again used the Pearson correlation coefficient between the residual from the smoothed fitted curve of the model and the measurements. The amplitude of the synoptic variations at each station were evaluated by the normalized standard deviation.
2.4.5 ERA5

To compare the simulated temperature with the ERA5 reanalysis, we divided the output into seasons and then into bins of 30° latitudes. For each bin, we averaged the data for each model level for each season. We then did an identical operation on the ERA5 reanalysis data before comparing the two.

2.4.6 Aircraft measurements

The aircraft measurements have been binned into 1 km altitude bins, and then averaged for each hour and over each bin for each site or campaign. Then the data was averaged over all sites and campaigns. This process was done for each season and for the whole year.

2.4.7 AirCore measurements

For measurements from AirCore, we binned and averaged the data into 50 altitude bins, from the ground to the maximum altitude of the data (27 km) to get an average vertical profile of CO₂ concentrations.

3 Results and discussions

3.1 Mass conservation

Conservation of mass is closely examined for the simulation of long-lived tracers as it directly supports the simulation of the tracer's global growth rate. In inverse systems, it makes it possible to infer surface fluxes far from observations, far in space as well as in time. In practice, numerical approximations may make the model lose or gain tracer mass (Houweling et al., 2010).

In this section, we evaluate mass conservation in our models by calculating the total mass of CO₂ at the beginning and at the end of the simulations. To do that, we multiply the CO₂ mole fraction with the air mass in each cell and sum it over the whole globe.

We then compare the difference between the CO₂ mass at the end and beginning with the total amount of prescribed surface CO₂ fluxes. The difference between these two values is the amount of CO₂ lost or gained by our model over time.

\[ \Delta M_{CO_2} = M_{CO_2}^e - M_{CO_2}^i = \sum_{n=1}^{N_{cell}} m_{air}^{n,x} \times c_{CO_2}^{n,x} - \sum_{n=1}^{N_{cell}} m_{air}^{n,i} \times c_{CO_2}^{n,i} \]  

\[ M_{CO_2}^{emi} = \sum_{t=1}^{T} \sum_{n=1}^{N} emi^{t,n}_{CO_2} \times area \]  

\[ M_{loss} = M_{CO_2}^{emi} - \Delta M_{CO_2} \]

For REG, the difference is equal to -0.13 % of the CO₂ mass emitted over the 1979 - 2020 period. For ICO it is -0.28 % for this same period.

Therefore, while our models do not exactly conserve mass, they lose only around 0.014 GtC integrated over 10 years for REG, and 0.027 GtC for ICO.

The total amount of CO₂ in each model also depends on the prescribed surface CO₂ fluxes described in section 2.2 (\( M_{CO_2}^{emi} \)) which are interpolated on the 2 different grids and therefore, not strictly identical either for each configuration. However, the average difference in yearly emitted CO₂ between the two model configurations is 0.0006 % only.
3.2 Computational efficiency

Simulations were run on the Skylake partition of the Joliot Curie, a BullSequana X1000 supercomputer operated since 2017 by Très Grand Centre de Calcul (TGCC, Bruyère-le-Châtel, France). This partition is composed of 1656 nodes, each of which has an Intel Skylake 8168 dual-processor. We used the Intel Fortran compiler version 20.0.0.

For our simulations REG used 47 MPI processes and 8 OpenMP threads for a total of 384 CPU cores, while ICO employed 80 MPI processes and 4 OpenMP threads for a total of 336 CPU cores. On average, over the whole simulation, REG achieved a wall-clock-time of 2594 seconds and consumed 2767 CPU hours per month simulated, while ICO executed in 2238 seconds and consumed 2089 CPU hours per month simulated. These results indicate that ICO provides a speedup of 14% in total CPU hours consumed over REG, with 22% less cells.

These results are highly dependent on the given output frequency of our simulation, in our case outputting large amounts of variables every hour greatly increases the execution time and becomes a computational bottleneck.

To better compare the configurations scaled up and in their ideal state, speed tests were run with identical numbers of CPUs: 71 MPI processes and 8 OpenMP threads with an additional 8 CPU used for XIOS servers for a total of 576 CPU cores. XIOS is a tool used for reading the input files in parallel and we chose 8 servers to ensure that this operation does not become a computational bottleneck for our models. Only monitoring files tracking the progress of the simulations were output, no physical variables were saved in order to avoid comparing the time it takes to write the files on disk. To avoid variability due to individual node performance, the tests were performed multiple times over several days, and outlier months caused by node performance issues were removed.

The average monthly time to completion in this setup for REG was 823 seconds (132 CPU hours), and for ICO 662 seconds (106 total CPU hours). This shows that for identical computational setups, ICO is on average 20% faster than REG.

This speedup is comparable to the reduced number of cells in ICO. For our spatial resolution, it seems that other differences such as the absence of a polar filter for ICO did not significantly improve the computational speed.

3.3 Vertical temperature profiles

To get a first idea of the differences between REG and ICO simulations, we consider atmospheric temperature and compare it with ERA5 values. Note that our models are nudged toward ERA5 horizontal winds (Section 2.2), but do not use the ERA5 temperature fields. Figure 2 shows the vertical profiles of the average temperature over the year 2000 for different zonal cuts in 60° latitude increments. We can already see that REG and ICO differ on several aspects for different altitudes. The tropopause height, as identified by the change in the vertical temperature gradient, is the same in both configurations, but its temperature varies between 2.5 K to 5 K for each configuration outside of the tropics. At the stratopause, a difference of up to 10 K for the yearly temperature average in high latitudes is observed between the simulations from REG and ICO (Not shown on the figure).
Figure 2: Vertical profile of zonal temperatures averaged over the year 2000 for the two model configurations and the ERA5 reanalysis, with REG in blue, ICO in red and ERA5 in yellow.

Looking at the temporal change of the temperature rather than yearly averages reveals a different pattern. We can see on Fig. 3 that the large temperature difference at the stratopause between our configurations is only present during winters for high latitudes. During summers, both configurations have much more similar temperatures in these latitudes, and all year around in the tropics. This is explained by the fact that during summers, the polar stratopause is mainly driven by ozone, whereas in winter it is driven by gravity-waves (Hitchman et al., 1989). The difference in parametrization and tuning of gravity waves in DYNAMICO used in our new configuration ICO compared to the previously used and much-tested REG version likely explains the observed differences in temperature of the stratopause. This large difference in temperature in the stratosphere also affects temperature lower in the troposphere, as has been shown for the stratospheric dynamics of the LMDZ GCM in Lott et al. (2005).
Future versions of the ICO configuration will contain a better parametrization of gravity waves as well as the introduction of a so-called “sponge layer” (Shepherd et al., 1996) to nudge high atmospheric winds towards zonal averages, which was already present in the REG configuration but not in ICO yet.

Figure 3: Time series of the average difference in zonal temperature at the stratopause (53 km) between the two model configurations REG and ICO for the year 2000, divided in 3 latitude zones of 60°.

We now turn to CO$_2$ concentrations to see how the different models affect tracer transport.
3.4 Seasonal analysis

3.4.1 Annual gradient

Figure 4: Annual gradients of CO$_2$ mixing ratio compared to SPO averaged over the 1980-2020 period for every station (a) or only stations at high latitudes (> 70°N/S) (b). Blue circles are the model outputs for the REG configuration, and red circles for the ICO configuration. The dotted lines correspond to the linear fitted lines of the corresponding colored configurations, and the black dotted lines correspond to the 1:1 relation.

Figure 4 shows the annual gradients of surface stations compared to SPO averaged over the 1980-2020 period, and for the two model configurations the differences between the modeled and observed values of this gradient. We find an average yearly growth rate of CO$_2$ mixing ratio at SPO of 1.79 ppm per year from observations, and of 1.74 ppm per year for both the REG and ICO configurations. This difference of 0.05 ppm between our models and observations shows that the background growth rate of CO$_2$ concentration is well modeled and within the small uncertainty range of the observations.

When looking at all surface stations (a), the ICO configuration exhibits a slightly lower overall bias, but an identical spread as seen by the root mean square error (RMSE). Both configurations show a positive bias of less than 0.1 ppm per year compared to observations. The two model configurations therefore successfully model the annual gradients between surface stations over the globe.
Figure 5: Normalized standard deviation of the annual gradient for both configurations for each station. Blue circles are the model outputs for the REG configuration, and red circles for the ICO configuration. The dotted line corresponds to the ideal normalized standard deviation of 1.

The average normalized standard deviation of the interannual variation in the annual gradient for the REG configuration is 1.43 and 1.3 for the ICO configuration. ICO therefore better captures the temporal variations of this gradient, but both configurations show a good agreement in magnitude of these variations for the majority of stations.

Since the biggest change regarding the grid and resolution takes place at the poles, we also checked the statistics and linear fit of these gradients restricted to stations at high latitudes (higher than 70°N and lower than 70°S). ICO performs just as well as REG for these stations in terms of both general bias and spread (Fig. 4(b)). Even though the effective resolution is much coarser for the ICO configuration at these latitudes, it has not significantly affected the simulation of long term trends of CO₂ concentrations. This shows that forced resolution clustering at the poles of the regular latitude-longitude grid is not necessary for properly resolving tracer transport.
## 3.4.2 Seasonal cycle

![Seasonal cycle](image)

The seasonal cycles at most surface stations are well captured by both configurations, with regards to both phase and amplitude, as illustrated in Fig. 6. Some stations exhibiting more complex and higher frequency patterns of CO₂ concentrations variation throughout seasons have a lower correlation coefficient. This pattern is observed for both configurations. However, almost all stations that are adequately modeled by the REG configuration with regards to seasonal cycles (correlation coefficient higher than 0.8) are equally well represented in the ICO configuration, as shown in Fig. 7 (a). Out of the 107 stations analyzed, only 12...
stations did not exceed a correlation of 0.8 with the REG configuration, and 11 with the ICO configuration. Additionally, no station exhibited satisfactory performance with REG but not with ICO, while the opposite was true for one station. The amplitude of the seasonal cycle is also well captured for almost all stations, as shown in Fig. 7 (b).

Figure 7: Pearson correlation coefficient (a) and normalized standard deviation (b) of the seasonal cycle for all surface stations studied averaged over the period 1980-2020, with blue circles for REG and red circles for ICO, the gray line is the difference between the two. The stations are ordered on the abscissa by increasing correlation coefficient for REG.

3.4.2 Synoptic variability
Figure 8: Pearson correlation (a) and normalized standard deviation (b) of the daily average residue between our modeled and measured CO$_2$ concentrations at the surface stations described in section 2.5 for the period 1980-2020. The model output from the REG and ICO configurations are in blue and red respectively.

To study the synoptic variability modeled by our two model configurations, we look at the correlation and the normalized standard deviation (NSD) of the daily averaged residue of our seasonal analysis for each surface station. This gives us information on the accuracy of our simulation for higher frequency than the seasonal cycles. Both configurations have correlation coefficients over 0.57 for 25% of all stations and a mean value of 0.47 for REG and ICO. The ICO configuration has a lower mean NSD of 1.06 compared to the one of REG of 1.20. And stations that offer a good correlation also tend to exhibit a better spread of the synoptic variability characterized by the NSD. These results are in line with what can be expected of a simulation at these resolutions as shown in Agustí-Panareda et al. (2019).

3.5 Vertical profiles of CO$_2$ concentrations

3.5.1 Troposphere

Figure 9: Seasonal and annual means of the difference in CO$_2$ vertical profile between the two model configurations of the model and aircraft measurements. The data has been binned into 1 km altitude bins for each season of the 1980-2020 period, then averaged per hour, and finally averaged across all aircraft sites and campaigns. The blue line represents the difference between REG and the measurements, while the red line represents the difference between ICO and the measurements.

In the troposphere, we studied CO$_2$ vertical profiles using various aircraft measurements described in section 2.4.2. Figure 9 shows the differences between the simulated and observed values for our two model configurations, REG and ICO. Only a small number of aircraft campaigns reach high altitudes above 15 km and not all seasons are covered. Both configurations show very similar vertical profiles up to 15 km altitude, before diverging above. Both configurations show a general negative bias compared to measurements. The variations in vertical gradients are almost identical for all altitudes, but the extent of the differences between model’s output and measurements differ at high altitudes. REG has much greater variations in CO$_2$ concentrations while ICO has an increased negative bias at high altitudes. This is similar to the results in the next section 3.5.2.
3.5.2 Low stratosphere

We utilized data from AirCore flights to compare the CO₂ mixing ratios of our model with observed data and obtain vertical profiles extending to the low stratosphere, in order to investigate the potential effects of the change in dynamics on vertical mixing within a column. However, since these measurements were only conducted in latitudes higher than 30°N and lower than 30°S, information about vertical tracer transport in the tropics was not obtained. As shown in Fig. 10, both model configurations exhibit an excess of CO₂ concentrations around the 12.5 km range. However, the REG configuration has an additional peak in CO₂ concentrations at 20 km, followed by a gradient change and a subsequent decrease in concentration at higher altitudes. In contrast, the ICO configuration does not display the same peak at 20 km, but a similar gradient change is observed above this altitude. This finding suggests that tracer vertical transport is inadequate between the low and high stratosphere at these latitudes, and CO₂ accumulates at lower levels than expected for the REG configuration. While the ICO configuration employs the same van Leer vertical transport scheme, differences in the vertical temperature profiles (see Fig. 2 and Fig. 3) discussed in section 3.4 could explain the disparity in the vertical profile at the stratosphere. The attribution of this systematic error to a particular process is complicated by the high potential collection altitude determination error of AirCore measurements, which can be on the order of a kilometer in the stratosphere as discussed briefly in section 2.3 (Wagenhäuser et al., 2021). The previously discussed conclusions however are independently verified by the aircraft measurements that do not suffer from the altitude determination error and show similar differences in CO₂ concentrations at 20 km (Fig. 9).
Figure 10: Difference in CO$_2$ mixing ratio vertical profile between the two model configurations of the model and AirCore measurements. The blue line represents the difference between the REG model output and measurements, while the red line indicates the difference between the ICO model output and measurements. The fitted lines were generated by averaging the data over 50 altitude bins.
Conclusion

As demonstrated in the previous section's results, the configuration ICO based on the new dynamical core Dynamico, using an unstructured grid is just as effective as the reference configuration that used a structured latitude-longitude grid for modeling atmospheric CO₂ transport when the dynamics was nudged to horizontal winds of an ERA5 reanalysis. Both configurations accurately capture the seasonal variations in CO₂ concentrations at most surface stations, and while the ICO configuration did not better capture more complex seasonal patterns, it did not worsen it either. A low percentage of station's seasonal cycles are properly captured by only one of the two model configurations. The annual gradient between stations exhibit slightly higher overall bias with ICO than with REG, but ICO has a smaller dispersion compared to observations. Regarding synoptic variability, the ICO configuration generally exhibits a lower correlation but a smaller standard deviation in comparison to observations. Nevertheless, both configurations provide an inadequate modeling of synoptic variability, as the local high-frequency emissions are poorly constrained.

Additionally, both configurations offer comparable vertical CO₂ concentration profiles and exhibit the same bias in the lower stratosphere relative to observational data. Temperature profiles in the tropopause and stratosphere seem to indicate that gravity waves still need to be tuned in the new ICO configuration (Lott et al., 2005). Their impact on atmospheric transport of CO₂ at lower altitudes has not been specifically evaluated but is expected to be minimal given the small differences shown between the two model configurations. Tuning of the climatology of the LMDZ - Dynamico coupling in general is still an ongoing process.

The new ICO configuration offers new opportunities in terms of development. Its use of fewer cells per level for a comparable resolution at the equator results in faster computation times of around 20% in our tests and easier-to-store outputs thanks to their smaller size on disk. Unlike regular latitude-longitude grids, ICO does not require a polar filter, whereas these filters generally parallelize badly, on both CPU and GPU. It gives a more homogeneous grid compared to the higher resolution at the poles of a regular latitude-longitude grid which is not always needed for modelisation at a global scale.

While running, REG and ICO can archive specific meteorological variables like air mass fluxes which can then be read by an offline version of the model dedicated to tracer transport. This economical transport model forms the basis of the inversion system of Chevallier et al. (2005) to generate the CO₂ and N₂O inversion products of the Copernicus Atmosphere Monitoring Service of the European Commission (CAMS service, (https://atmosphere.copernicus.eu/ghg-services). Our next task is the implementation of Dynamico in this offline model in order to prepare future resolution increases, while damping the induced increase of the code time-to-solution.

Code and data availability

The source code for the REG and ICO configurations is freely available online via the following address: https://forge.ipsl.jussieu.fr/igcmg/browser/CONFIG/publications/ICOLMDZORINCA_CO2_Transport_GMD_2023 under the CeCILL v2 Free Software License (http://www.cecill.info/index.en.html, last access: 11 September 2023, CECILL, 2020). The exact version of the model used to produce the results used in this paper is archived on Zenodo, as are input data and scripts to run the model and produce the plots for all the simulations presented in this paper (Lloret et al., 2023).
Author contributions
ZL designed, ran and analyzed the simulations and prepared the paper under direction and advice of FC. AC contributed to the development and preparation of the simulations. MR was also involved in the development of the simulations. YM supported the development of the model. All authors contributed to the text.

Disclaimer
The authors declare that they have no conflict of interest.

Acknowledgements
This work was granted access to the HPC resources of TGCC under the allocation A0130102201 made by GENCI. It was funded by the Copernicus Atmosphere Monitoring Service, implemented by ECMWF on behalf of the European Commission (Grant: CAMS2 55). We thank Olivier Boucher for fruitful discussions on vertical temperature profiles and the LMDZ model, and Sakina Takache for discussions on Dynamico. We also thank Julien Derouillat for his help in the implementation of XIOS for our application. We also gratefully acknowledge the many people who contributed atmospheric observations (aircraft, AirCore and surface air sample measurements). Measurements at Jungfraujoch were supported by ICOS Switzerland, which is funded by the Swiss National Science Foundation, in-house contributions and the State Secretariat for Education, Research and Innovation. We thank Kathryn McKain for her comments and suggestions on the best way to describe aircraft and AirCore data.

References


