Sweep Interpolation: A ~~Fourth-Order Accurate~~ Cost Effective Semi-Lagrangian Scheme in the Global Environmental Multiscale Model

Mohammad Mortezazadeh[1], Jean-François Cossette[2], Ashu Dastoor[1], Jean de Grandpré[1], Irena Ivanova[1], Abdessamad Qaddouri[2]

[1]Air Quality Research Division, Environment and Climate Change Canada, Dorval, H9P 1J3, Canada
[2]Meteorological Research Division, Environment and Climate Change Canada, Dorval, H9P 1J3, Canada

Correspondence to: Ashu Dastoor (ashu.dastoor@canada.ca)

Abstract. The interpolation process is the most computationally expensive step of the semi-Lagrangian (SL) approach for solving advection which is commonly used in numerical weather prediction (NWP) models. It has a significant impact on the accuracy of the solution and can potentially be the most expensive part of model integration. The sweep algorithm, which was first described by Mortezazadeh and Wang (2017), performs SL interpolation with the same computational cost as a third order polynomial scheme but ~~with~~ at the accuracy of ~~a~~ fourth order ~~interpolation scheme~~. This improvement is achieved by using two 3rd-order backward and forward polynomial interpolation schemes in two consecutive time steps. In this paper, we present a new application of the sweep algorithm within the context of global forecasts produced with Environment Climate Change Canada's Global Environmental Multiscale (GEM) model. Results show that the SL scheme with sweep interpolation ~~scheme~~ is computationally more efficient compared to a conventional SL scheme with fourth order polynomial ~~scheme~~interpolation, especially when a very large number of passive tracers are advected.~~especially evident for increasing number of several advected several passive tracers.~~ An additional advantage of this new approach is that its implementation in a chemical and weather forecast model~~s~~ requires minimum modifications of the interpolation weighting coefficients. An analysis of the computational performance for a set of theoretical benchmarks as well as a global ozone forecast experiment show that up to 15% reduction in total wall clock time is achieved. Forecasting experiments using the global version of the GEM model and the new interpolation show that the sweep interpolation can perform very well in predicting ozone distribution, especially in the tropopause region where transport processes play a significant role.

1 Introduction

Following their development nearly 4 decades ago, SL schemes have been widely used by NWP models such as the Global Environmental Multiscale Model (GEM), the Integrated Forecasting System (IFS), and the Global Forecast System (GFS) (Robert, 1980; Staniforth and Côté, 1991; Ritchie et al., 1995; Girard et al., 2014; Husain and Girard, 2017; Husain et al., 2019; Mukhopadhyay et al., 2019). One of the main reasons for ~~its~~ their popularity is the fact that ~~it~~ these schemes~~y~~ can use large time steps without suffering from the same stability issues as the Eulerian methods, which are constrained by the Courant-Friedrich-Lewy condition (McDonald and Bates, 1987; Côté and Staniforth, 1988; Ritchie, 1988). SL schemes can be derived from an integral along the path that links a flow trajectory's departure point to its arrival point located on a fixed

Eulerian grid (Smolarkiewicz and Pudykiewicz, 1992). From this perspective, the flow is considered as a group of discrete fluid particles following their characteristic curves along the Lagrangian coordinate (Mortezazadeh and Wang, 2017). The SL scheme consists of ~~two main~~three distinct steps. First, it finds the departure point by performing backward integration of the kinematic relationship describing the characteristic path of the fluid particles (McDonald and Haugen, 1992; Hortal, 2002). Second, an interpolation scheme maps the data from the Eulerian grid to the position of the departure point (Williamson and Rasch, 1989; Purser and Leslie, 1991). Third, contributions from forcings are taken into account as integrals along the flow trajectories (e.g. see Cossette et al., 2014 and the references therein).

Choosing a suitable interpolation scheme plays an important role in determining the accuracy and computational cost of the SL approach (Yabe et al., 2001). In particular, ~~Low~~ low order interpolation is susceptible to high dissipation error and as a result it may induce large conservation and shape preservation errors in the solution (Zerroukat, 2010; Mortezazadeh and Wang, 2017). Thus, a high order interpolation scheme is preferable to reduce the dissipation error. Various high order interpolation schemes have been proposed to improve the accuracy of the SL approach such as Lagrange and Hermite polynomial interpolations, as well as Weighted Essentially Non-Oscillatory (WENO) interpolation ~~such as cubic polynomial interpolation scheme, Hermite interpolation, Weighted Essentially Non-Oscillatory (WENO).~~ (Shapiro and Hastings, 1973; Girard et al., 2014; Nakao et al., 2022; Petras et al., 2022). One of the common polynomial interpolation schemes that is used in weather and air quality forecasting systems is the cubic Lagrange interpolation scheme (Aires et al., 2004; Girard et al., 2014~~).),~~ ~~This scheme~~which can provide sufficient accuracy for tracer transport and advection without being prohibitively computationally expensive.

Nonetheless, a significant amount of information from neighboring grid points is required to perform any high-order polynomial interpolation. For instance, a fourth-order-accurate 3D Lagrange interpolation scheme uses 64 grid points, i.e. 4 grid points ~~s~~ in each direction. Because of the significant computational cost of the high-order interpolation, the transport of tracers in atmospheric models can significantly affect the computational performance, especially when considering the evolution of large number of tracers (Bradley et al., 2022). This issue is particularly important with the extensive development of operational Air Quality prediction systems which attempt to resolve a comprehensive set of physicochemical processes involving a growing ~~number~~number of atmospheric chemical constituents (Im et al., 2015; Makar et al., 2015). Depending on the application, the number of tracers used in atmospheric air quality models can vary from few to hundreds, thereby increasing the overall cost of tracer transport.~~increasing computational cost of tracer transport.~~ For example, Golaz et al. showed that in an atmospheric model with 40 tracers, 75% of the dynamical core wall clock time is taken by the tracer transport solver (Golaz et al., 2019).

By contrast to the standard fourth-order-accurate interpolation, the sweep interpolation scheme can reduce the cost of the interpolation process by using fewer neighborhood grid points but keeping almost the same accuracy as fourth order interpolation. This scheme is based on the alternate use of third order backward and forward polynomial interpolation schemes (Mortezazadeh and Wang, 2017). This procedure is applied in two successive time steps, so that at the end of each two time steps the truncation error is approximately equal to the error of the cubic interpolation scheme. Performance of the

sweep interpolation scheme was investigated in 1D, 2D, and 3D benchmarks, a wave function and cavity flow in a 2D and 3D squares, respectively, and complete error analysis showing the accuracy of the scheme was reported (Mortezazadeh and Wang, 2017).

In this paper, the sweep interpolation scheme is implemented into the Global Environmental Multiscale (GEM) model and its accuracy and computational performance in tracer transport are assessed. The paper is organized as follows. In the first section we present the formulation of the sweep interpolation scheme. Next, three benchmarks are used to evaluate the performance of sweep interpolation. Finally, the accuracy of sweep interpolation is evaluated by performing an ensemble of ozone forecasts over a ~~10~~ 10 ~~weeks~~week period within the ECCC NWP global forecasting system based on the GEM model. Additionally, we investigate the performance of the new interpolation scheme by using different numbers of passive tracers. ECCC's air quality forecast model, GEM-MACH (Global Environmental Multiscale – Modelling Air quality and CHemistry) (Zhou et al., 2021; Stevens et al., 2022), is an in-line chemistry-meteorology model based on the GEM model. The benefits of sweep interpolation scheme in reducing computational cost of the transport of chemical species in GEM-MACH will be tested in a future study.

2 Methodology

GEM is used for meteorological forecasting at all scales from 15 [km] to 2.5 [km]. It solves the elastic Euler equations using a two-time-level Crank-Nicholson temporal discretization with SL treatment of the advection terms (Girard et al., 2014). The current version of GEM uses a finite difference discretization on a Yin-Yang grid (Qaddouri and Lee, 2011). The governing equations are formulated using spherical coordinates together with a log-hydrostatic pressure type terrain following vertical coordinate (Husain et al., 2020). ~~Equations~~ They are discretized on an Arakawa C grid (Arakawa, 1988) in the horizontal, whereas ~~and~~ in the vertical direction, they are discretized using a Charney–Phillips grid~~a hydrostatic pressure coordinate is used and the equations are discretized on the Charney–Phillips grid~~. Tracer transport is accomplished by first solving the advection equation for a passive tracer and then by adding contributions from physics forcings in split mode. The current interpolation scheme in GEM is fourth-order-accurate cubic Lagrange interpolation. It is used to calculate the variables at the departure point, as well as to perform the exchange of data on the boundaries of the two sub-grids of the global Yin-Yang grid. In this study, we document the impact of using sweep interpolation for the advection of tracers as well as for the exchange of data between Yin and Yang sub-grids in GEM. ~~In the horizontal direction, spherical coordinates are used to solve the equations.~~

GEM solves a system of four prognostic equations, which include momentum (Eq. (1)), energy (Eq. (2)), mass conservation (Eq. (3)), and ideal gas law (Eq. (4)):

$$\frac{d\mathbf{V}}{dt} + f\mathbf{k} \times \mathbf{V} + \frac{1}{\rho}\boldsymbol{\nabla} p + g\mathbf{k} = \mathbf{F}, \tag{1}$$

$$\frac{dT}{dt} - \frac{1}{\rho c_p}\frac{\partial p}{\partial t} = \frac{Q}{c_p}, \tag{2}$$

$$\frac{d\ln\rho}{dt} + \nabla.\mathbf{V} = 0, \tag{3}$$

$$\rho = \frac{p}{RT}, \tag{4}$$

where $\mathbf{V} = (\mathbf{V}_h, w)$, T, p, and ρ are the velocity vector, temperature, pressure, and density, respectively. Here, **F** and Q are the source terms for friction and heat, respectively. g is the gravitational force, and c_p and R are the thermodynamic parameters. The general form of nonlinear term in the prognostic equations and tracer transport equations can be written as follows:

$$\frac{d\emptyset}{dt} = P, \tag{5}$$

where $\emptyset$ represents a prognostic variable or tracer scalars, and P is the forcings. When performing tracer advection ~~In~~ in GEM, ~~by using the split approach,~~ we solve Eq. (5) based on two steps. At first, we solve the pure advection equation and calculate the intermediate solution $\emptyset_A^*$.

$$\frac{d\emptyset}{dt} = \frac{\partial\emptyset}{\partial t} + \boldsymbol{u}\cdot\frac{\partial\emptyset}{\partial \boldsymbol{X}} = 0, \tag{6}$$

where $\boldsymbol{X} = (x, y, z)$ and t specify spatial and temporal coordinates. Then by using the semi-Lagrangian scheme, Eq. (6) is written as:

$$\emptyset_A^* = \emptyset_D, \tag{7}$$

Here, A: $(\mathbf{X}, t)$ and D: $(\mathbf{X} - \Delta\mathbf{X}, t - \Delta t)$ correspond to the arrival and departure positions of flow trajectories, respectively (see Figure 1).

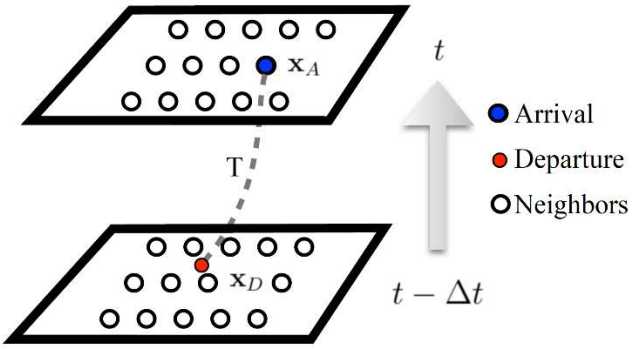


Figure 1. Procedure of time marching in semi-Lagrangian scheme.

In the second step, we add the forcing term to the intermediate solution to find the results at the present time step:

$$\emptyset_A = \emptyset_A^* + \Delta t P, \tag{8}$$

In the semi-Lagrangian method, the most challenging part is calculating the value of the variables at the departure point $\boldsymbol{X}_D = (x_D, y_D, z_D)$. To accomplish this, first we need to find the position $\boldsymbol{X}_D$ by moving back along the flow ~~trajectories~~ trajectory T arriving ~~T~~ at $\boldsymbol{X}_A$:

$$\boldsymbol{X}_D = \boldsymbol{X}_A - \int_{t-\Delta t}^{t} \mathbf{V}(\mathrm{X}(\tau), \tau)\, d\tau, \tag{9}$$

By using an averaging procedure for the integral of velocity on the right-hand-side of the Eq. (9), this equation can be written as follows:

$$\boldsymbol{X}_D = \boldsymbol{X}_A - (BV_A + (1-B)V_D), \tag{10}$$

where $B \geq 0.5$ is an off-centering weight. Because the position of the departure point ($\boldsymbol{X}_D$) does not necessarily coincide with a grid point, the use of an interpolation scheme is required to transfer data from the Eulerian grid to the departure point~~is not necessarily on the grid points, use of an interpolation scheme is unavoidable to transfer the data from Eulerian grids to the position of the departure point~~ (Williamson and Rasch, 1989; Purser and Leslie, 1991). Lagrange interpolation is used in GEM~~:~~.

$$\emptyset_A^* = \sum_l w_l \emptyset_l^{t-\Delta t}, \tag{11}$$

where l represents the index of neighbor grid points surrounding the departure point and w_l are the interpolation weights. The cubic interpolation currently used in GEM uses 64 neighbor cells, i.e. 4 grids in each direction, to achieve fourth-order spatial accuracy ~~In the current version of GEM, cubic interpolation is available. In this method, for doing the interpolation, 64 neighbor cells, i.e. 4 grids in each direction, are considered to achieve the fourth order of accuracy in space~~. Below~~In the following study~~, we explain the Mortezazadeh and Wang (2017) method to generate a high-order accurate semi-Lagrangian scheme by using only 27 neighbor cells that has the same accuracy as a SL method based on the standard cubic interpolation and which can also lead to a faster interpolation process.

~~Here, for the sake of simplicity, we discretize the equations and investigate the numerical approach based on 1D advection ($\boldsymbol{X} \rightarrow x$) and a uniform grid with the grid spacing size of Δx (see Figure 2).~~ For the sake of illustrating how the sweep method works, we use the 1D discretization of the equations based on the uniform grid spacing size of Δx that is shown in Figure 2.

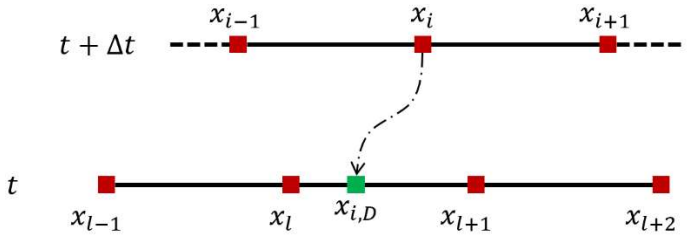


Figure 2. 1D computational stencil diagram of neighborhoods near the departure point.

As mentioned before, cubic Lagrange interpolation scheme uses 4 grid neighborhoods, along the 1D direction, to calculate $\emptyset_D$ at the departure point $x_{i,D}$, i.e. x_{l-1}, x_l, x_{l+1}, and x_{l+2}. Here i is the index of arrival cells or Eulerian grid points. Note that here using the Lagrange interpolation scheme, we can construct a prediction function $\widehat{\emptyset}(x_{i,D})$ from the exact solution of $\emptyset(x_{i,D})$ within the cell $[x_l, x_{l+1}]$. The Error error in cubic interpolation is can be expressed as follows:

$$E_{Cubic} = \emptyset(x_{i,D}) - \widehat{\emptyset}(x_{i,D}) = \frac{1}{4!}\emptyset^{(4)}(x_{i,D}) \times \prod_{l-1}^{l+2}(x_{i,D} - x_l) = O(\Delta x^4), \tag{12}$$

Sweep interpolation is an integration of two 3rd-order backward and forward interpolation steps that are used successively in two time steps. In the following, we show how sweep interpolation can achieve almost the similar accuracy as cubic interpolation by using fewer neighborhoodsgrid neighbors. Equation (13) shows the numerical errors in 3rdd-order backward (E_B) and forward (E_F) polynomial interpolation scheme:

$$\begin{aligned} E_B &= \frac{1}{3!}\emptyset^{(3)}(x_{i,D}) \times \prod_{l}^{l+2}(x_{i,D} - x_l) = -O(\Delta x^3), \\ E_F &= \frac{1}{3!}\emptyset^{(3)}(x_{i,D}) \times \prod_{l-1}^{l+1}(x_{i,D} - x_l) = O(\Delta x^3), \end{aligned} \tag{13}$$

From Eq. (13) we will find that the numerical error for the two interpolation steps is almost the same but with opposite sign. While use of either backward or forward interpolation scheme after two time steps can generate $E_{B,2\times\Delta t} = 2 \times E_B$, using 3rd-order backward and forward interpolation successively in two time steps can eliminate the 3rd-order numerical errors $O(\Delta x^3)$ and produce an and convert the total interpolation error to $O(\Delta x^4)$ interpolation error.

$$E_{2\times\Delta t} = E_B + E_F \approx O(\Delta x^4), \tag{14}$$

Although Here the analysis of the numerical errors is shown by assuming a constant Δx (uniform grids) in the above equations, but it is also valid for non-uniform grids (Mortezazadeh and Wang, 2017). Although, as shown above, the truncation errors of two backward and forward steps nearly cancel out to produce a total truncation error with a leading term that is proportional to the fourth power of the grid increment, the actual numerical accuracy of its implementation in SL schemes needs to be assessed. In the next section, we will seeassess how this the sweep interpolation approach can control the error after every two time steps in the SL scheme in GEM. In this study, sweep interpolation is used not only for semi-Lagrangian advection, but also for exchanging the data between Yin and Yang grids in GEM. In the following, we investigate the performance of the model by solving three benchmarks and an ozone prediction case.

3 Validation

The objective of this evaluation is to illustrate the feasibility of implementing the sweep interpolation within the GEM model to replicate the forecast results achieved by the GEM configuration based on cubic interpolation. The evaluation has been done by running one 2D case and three 3D cases to represent the accuracy and performance of the proposed method. Error

analysis is performed to~~and contours~~ show that semi-Lagrangian with sweep interpolation can provide similar results as cubic interpolation in a more cost-effective manner.

3.1 2D vortex

The 2D vortex is a conventional benchmark which is used here to compare the accuracy of the sweep interpolation with the cubic interpolation. This case consists of a scalar field that is deformed by two static vortices centered at the geographical poles (Nair and Machenhauer, 2002). The flow field utilized in this benchmark is positive definite. The spatial resolution used in both horizontal directions is approximately 105 [km] and a time step of 7200 [s] is employed, which yields a maximum value of Courant number ~~is~~ of 0.85426~~Here, maximum value of Courant number is 0.85426 ($\Delta t = 7200\ [s]$)~~. By considering (λ, φ) as the rotated coordinate system, North Pole location: λ_0, φ_0, and an angular velocity: ω, we can write the rotation of the scalar field as:

$$\frac{d\lambda}{dt} = \omega, \tag{15}$$

$$\frac{d\varphi}{dt} = 0,$$

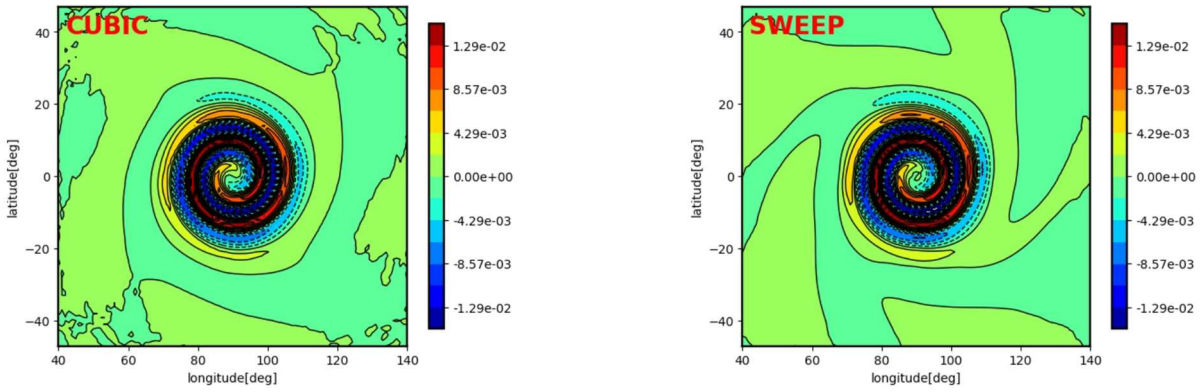
The normalized tangential velocity of the vortex is defined as:

$$V_t = \frac{3\sqrt{3}}{2}\mathrm{sech}^2(r)\tanh(r), \tag{16}$$

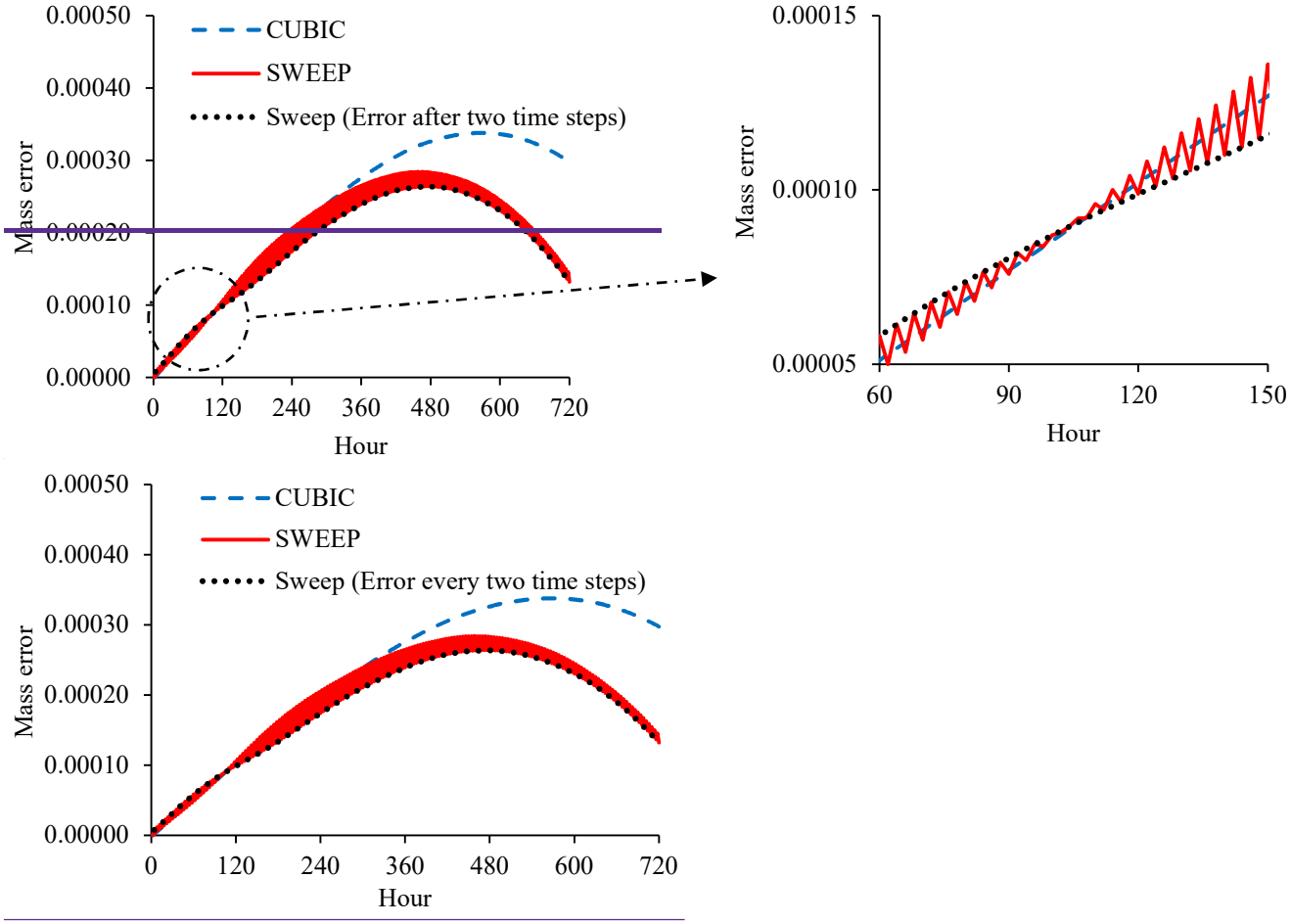
where $r = r_0 \cos\varphi$ is the radius of the vortex and r_0 is a constant. Then, the angular velocity ω is defined as:

$$\omega(\varphi) = \begin{cases} 0 & if\ r = 0 \\ \frac{V_t}{r} & if\ r \neq 0 \end{cases}, \tag{17}$$

The analytical solution of this case at time t is available in the reference (Nair and Machenhauer, 2002). Figure 3(a) shows error in the tracer at day 12. It can be seen that the error distribution inside the domain is almost similar for both schemes.



(a)



(b)

Figure 3. 2D vortex: (a) Horizontal cross section of ~~Errors~~ errors in tracer distribution at day 12, (b) Mass conservation error (in %) over 1 month period.

Figure 3(b) shows the mass error over 1 month. This figure shows an oscillation in mass error, calculated by sweep interpolation, which comes from the inherent behavior of sweep interpolation to control the growth of numerical error over two successive time steps~~the iterations~~. The black dotted line in the figure 3(b) shows the mass error after each two time steps, which represents the impact of sweep interpolation to reduce the error. This line represents the impact of sweep interpolation to reduce the error. The main reason of choosing the 2D case was showing the oscillation in the mass error for sweep interpolation. In this case, the oscillation is~~was~~ obvious and ~~could~~ helps ~~us better to~~ explain the behavior of~~in~~ sweep interpolation. The same behavior has been seen in the other test cases (see next sections). For this case, the normalized infinity norm error ($E_\infty = \frac{\max |cubic-sweep|}{\max |cubic|}$)=0.001. Figure 3(a) shows that the error distribution associated with the sweep interpolation is less noisy compared with the cubic interpolation error, especially outside the vortex region, which could be

explained by the fact that the lower order Lagrange interpolation used in the sweep algorithm generates less spurious oscillations compared to the standard cubic interpolation. ~~Figure 3(a) shows the error distribution is less noisy, especially outside the vortex region. Using lower order Lagrange interpolation may minimize the noises inside the domain which can be the reason of this minor difference in the error distributions.~~

3.2 Hadley-like meridional circulation

The Hadley-like Meridional Circulation test case is used to emphasize the solution of horizontal-vertical coupling (Ullrich et al., 2013). The tracer field consists of a single layer, which is deformed over the duration of 24 hours. At the end, the tracer field returns to its original configuration. Meridional and vertical velocity fields for this test case are specified as:

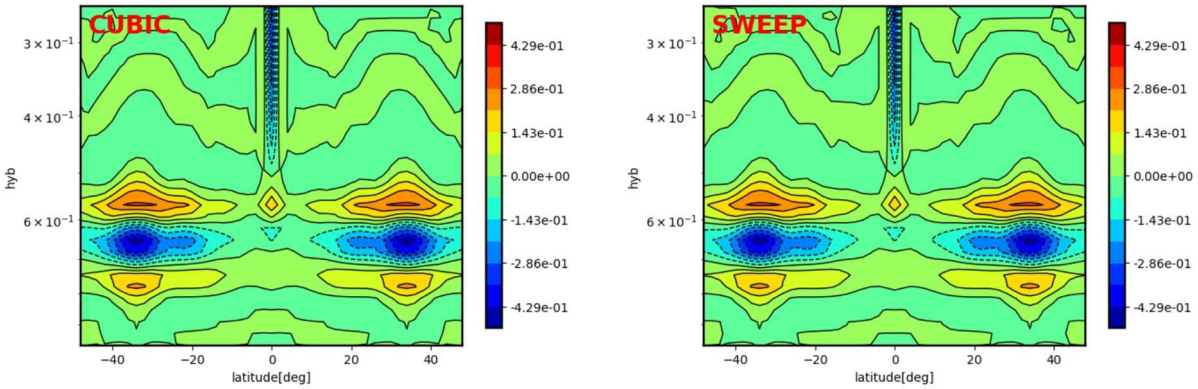
$$u(\lambda, \varphi, z, t) = u_0 \cos\varphi, \tag{18}$$

$$v(\lambda, \varphi, z, t) = \frac{a w_0 \pi \rho_0}{K z_{top} \rho} \cos\varphi \sin(K\varphi) \cos\left(\frac{\pi z}{z_{top}}\right) \cos\left(\frac{\pi t}{\tau}\right), \tag{19}$$

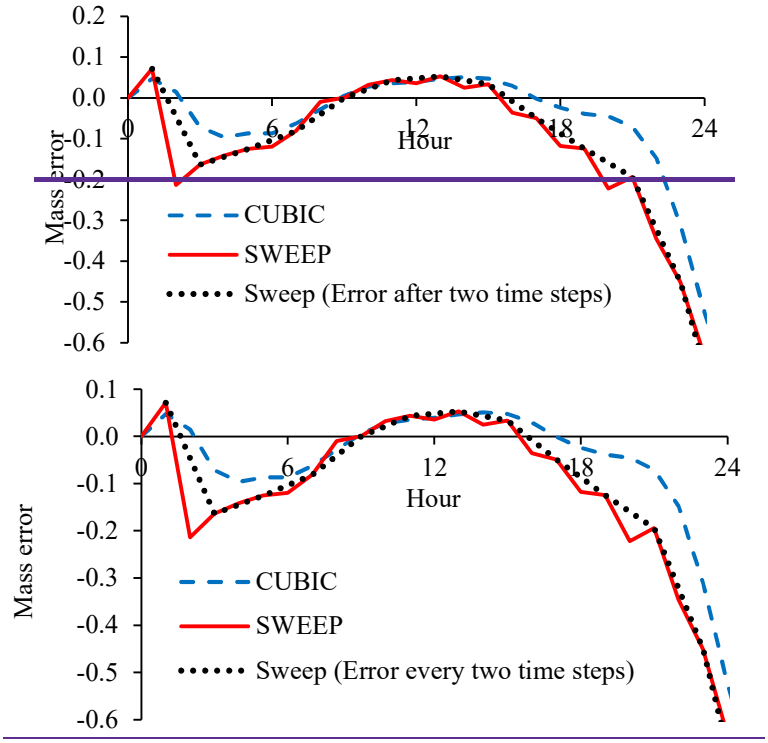
$$w(\lambda, \varphi, z, t) = \frac{w_0 \rho_0}{K\rho} (-2\sin(K\varphi)\sin\varphi + K\cos\varphi\cos(K\varphi)) \sin\left(\frac{\pi z}{z_{top}}\right) \cos\left(\frac{\pi t}{\tau}\right), \tag{20}$$

where ρ_0, $K = 5$, $z_{top} = 12000\ [m]$, $w_0 = 0.15\ [m\ s^{-1}]$, $u_0 = 40\ [m\ s^{-1}]$, a and $\tau = 86400\ [s]$ are the density at the surface, number of overturning cells, height position of the model top, reference vertical velocity, reference zonal velocity, reference earth radius, and period of motion. The horizontal spatial resolution and time step used in this example are, respectively, 205 km and ~~dt=~~3600s, which yields a horizontal Courant number (CFL) of 5.0 ~~The maximal Courant number $CFL = 5.0$ ($\Delta t = 3600\ [s]$) and the spatial resolution is 205 [km]~~.

~~Figure 4(a) presents the meridional cross sections of the errors in the tracer solution after a 24 hours integration period. The results show that sweep interpolation generates a solution (right panel) that is similar to the one generated with the fourth-order accurate cubic interpolation (left panel). Figure 4(b) shows the evolution of the error in the mass conservation (in %) over the same 24 hour period corresponding to both interpolations. The evolution in both solutions remain almost the same for both sweep and cubic interpolations. Note that here we can see the oscillations in error for sweep interpolation scheme. As explained in the previous section, this shows how sweep interpolation can control the error growth over each two time steps.~~



(a)



(b)

Figure 4. Hadley-like Meridional Circulation: (a) Vertical cross section of ~~Errors~~ errors in tracer at day 1, (b) Mass conservation error (in %) over 1 day.

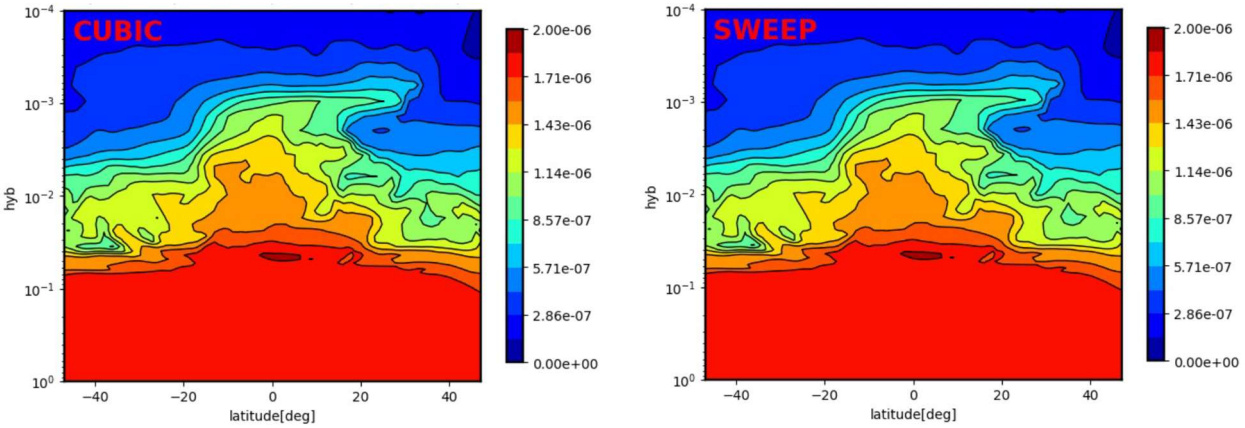
Figure 4(a) presents the meridional cross sections of the errors in the tracer solution after a 24-hour integration period. The results show that sweep interpolation generates a solution (right panel) that is similar to the one generated with the fourth-

order-accurate cubic interpolation (left panel). Figure 4(b) shows the evolution of the error in the mass conservation (in %) over the same 24-hour period corresponding to both interpolations. The evolution in both solutions remain almost the same for both sweep and cubic interpolations. Note that here we can see the oscillations in error for sweep interpolation scheme. As explained in the previous section, this shows how sweep interpolation can control the error growth over each two-time steps. Here, the normalized infinity norm error is $E_\infty = 0.03$.

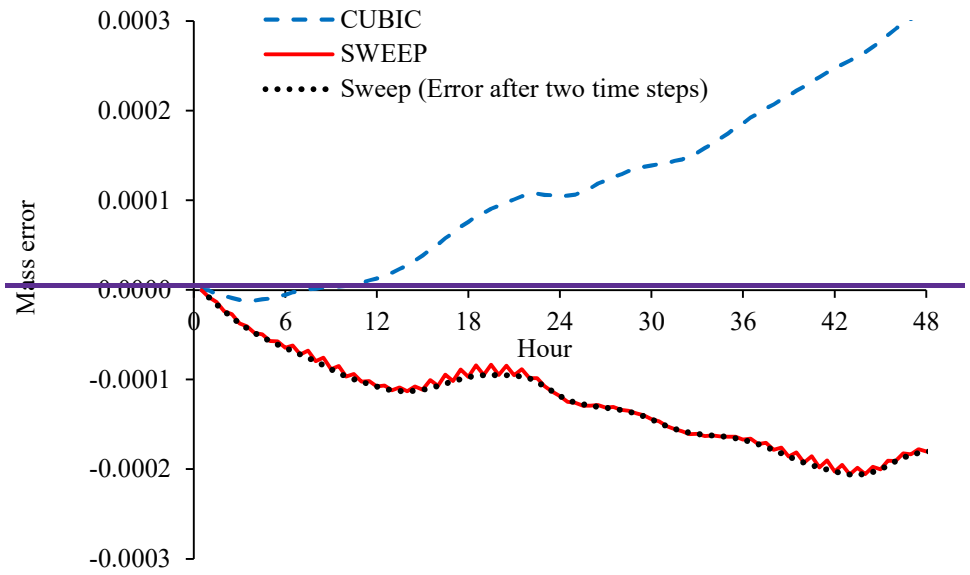
3.3 Atmospheric methane-like tracer

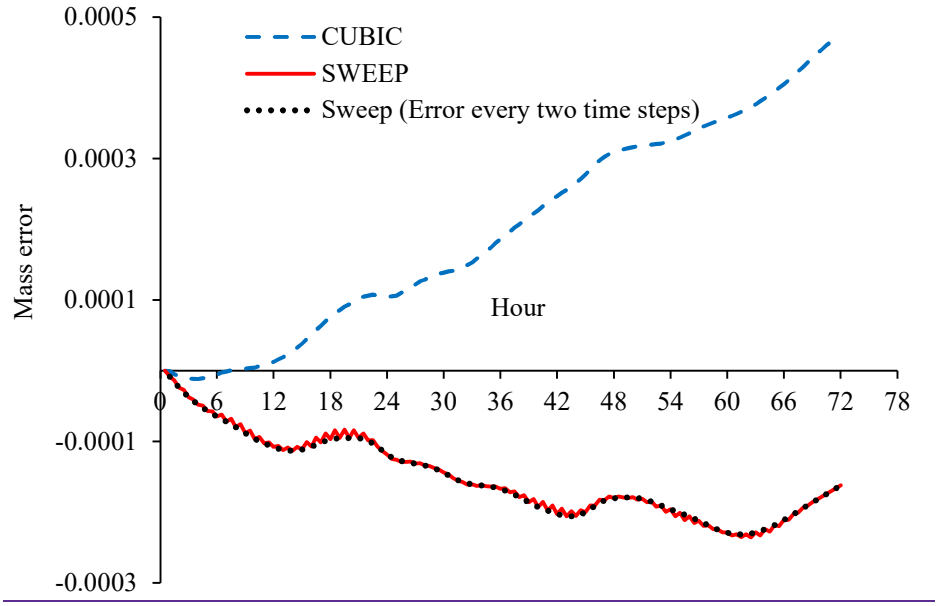
In this test case, we compare ~~48 hour~~48-hour forecasts of atmospheric methane (CH_4) like passive tracer (without chemical productions and sinks) using ~~different interpolation schemes,~~ sweep interpolation and cubic interpolation. These experiments were performed with the global version of GEM NWP model using ~~with~~ a ~~30 minute~~30-minute time step and ~~a~~ 105 [km] horizontal resolution ~~at 84 levels resolution in the vertical direction to give a maximal courant number of 4.7. The grid distribution along the vertical direction is non-uniform. The horizontal grid resolution is 115 km~~resulting in a maximal courant number of 4.7. The height of the model top was chosen to be at 0.1 hPa and 84 vertical levels were used. The vertical grid resolution is non-uniform as a result of the choice of vertical coordinate which is based on the logarithm of the hydrostatic pressure (Husain et al., 2020). The methane-like experiment was initialized from a climatology based on a multi-year simulation performed with the GEM model. The model employs a simplified approach, in which methane production and loss are predetermined based on present-day conditions (Prather et al., 2012). Figure 5(a) presents meridional cross sections of CH_4 at the end of day 1. Solutions from both interpolators look qualitatively the same, and the sweep interpolation provides acceptable results in comparison with cubic interpolation. Figure 5(b) shows the mass error over 24 hours. It shows that both cubic and sweep interpolations could control the error and keep its range below 0.005% after 48 time steps of simulation. Here, the normalized infinity norm error is $E_\infty = 0.018$.

Although sweep interpolation was able to better control the mass error growth over the simulation time compared to the cubic interpolation for this case, it is not necessarily expected to perform better in all cases. Based on our discussion in the previous section, we expect sweep interpolation to provide almost the same accuracy as cubic interpolation. This is supported by Fig 5(b), which shows that sweep and cubic interpolations produce mass errors that are of the same order of magnitude. However, since both methods rely on different finite difference approximations, we expect to see differences in the evolutions of their respective error trends, which is confirmed by the results of Fig. 5(b).


CUBIC
SWEEP
hyb
latitude[deg]
2.00e-06
1.71e-06
1.43e-06
1.14e-06
8.57e-07
5.71e-07
2.86e-07
0.00e+00

(a)


CUBIC
SWEEP
Sweep (Error after two time steps)
Mass error
Hour
0.0003
0.0002
0.0001
0.0000
-0.0001
-0.0002
-0.0003
0
6
12
18
24
30
36
42
48



(b)

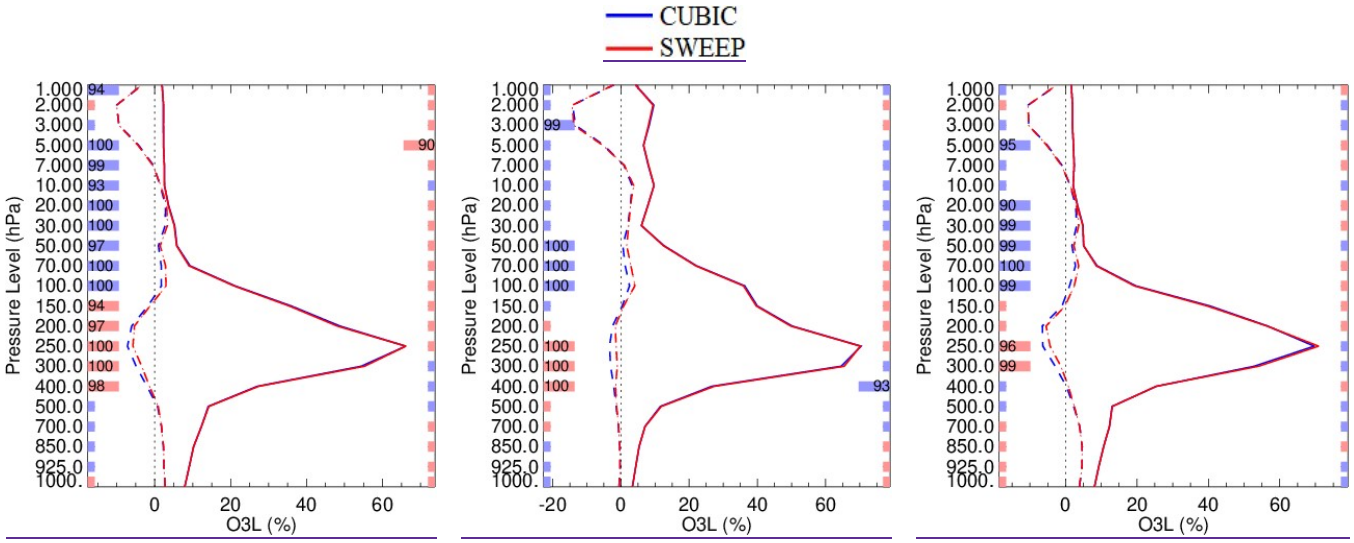
Figure 5. CH_4 Forecasts: (a) Meridional cross sections at ~~180 degree~~180-degree longitude at the end of day 1, (b) Normalized mass error (in %) over ~~2 day~~3-day forecast.

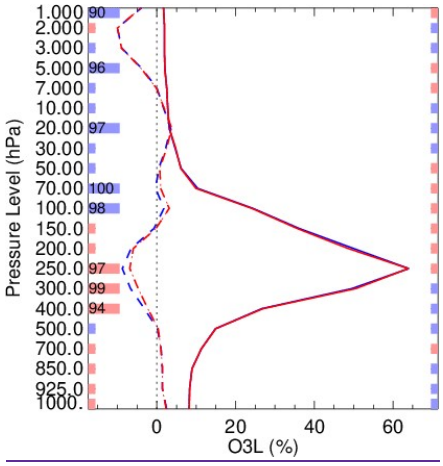
~~Although sweep interpolation could better control the mass error growth over the simulation time for this case, it is not necessarily expected to perform better in all cases. Based on our discussion in the previous section, we expect sweep interpolation to provide almost the same accuracy as cubic interpolation.~~

3.4 Medium range forecast~~s~~ of ozone

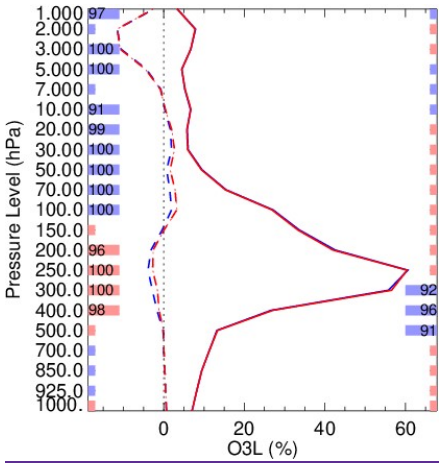
In the UTLS region, ozone is mainly driven by transport processes which makes it a valuable tracer for assessing the impact of the new interpolation scheme on the transport of chemical constituents. The implementation of sweep in the ECCC GEM NWP model allowed evaluation of its impact on the ozone predictability in the context of an operational NWP system. In this experiment, an ensemble of 10-day ozone forecasts were launched every 12 hours from June 13th to August 31st 2019 using the global version of the GEM NWP model. These forecasts were performed at 15 [km] horizontal resolution on 84 vertical levels with a lid at .1 hPa and using a 450 seconds timestep. The model included a prognostic representation of stratospheric ozone based on a linearized chemistry scheme which did not interact with the radiation (de Grandpré et al., 2016). Ozone forecasts have been evaluated at different lead-times against operational analyses generated by the Global Deterministic Prediction System (Charron et al., 2012). This evaluation was done throughout the tropospheric and stratospheric regions in various areas including Northern/Southern extratropics, North America, Asia, ~~Arctic~~Arctic, and Antarctic.

~~The results from two interpolations, i.e. cubic and sweep interpolations, at different locations over the whole earth experiments demonstrate a good agreement between sweep interpolation and cubic interpolation. The two main metries of the statistical accuracy, bias and standard deviation, are used to show the differences between the two experiments. In the following figures (Figure 6 and 7), solid and dashed lines represent the standard deviation and bias metrics, respectively. Figure 6 represents the average of all 10 day lead time forecasts in the simulation time period. This figure shows that the standard deviation for both experiments are almost the same, but some minor deviations can be seen in the bias. The bias for ozone variation along the vertical direction is similar for both experiments in the lower part of troposphere (>500 hPa) and in the upper stratosphere (<50 hPa). The smaller concentration of ozone in these two regions accounts for the absence of sharp gradients in the vertical error profiles. The large maximum in the standard deviation around 250 hPa is associated with the transport of ozone across the tropopause. This process is sensitive to the choice of interpolation scheme and the results highlight its impact on stratosphere troposphere exchange processes throughout the entire upper troposphere and lower stratosphere region between 400 and 50 hPa. In this region, a minor deviation of bias error can be seen between sweep interpolation and cubic interpolation, but this deviation is small and not significant (see Figure 6).~~

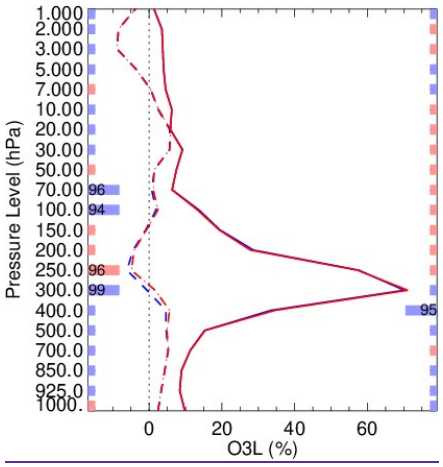




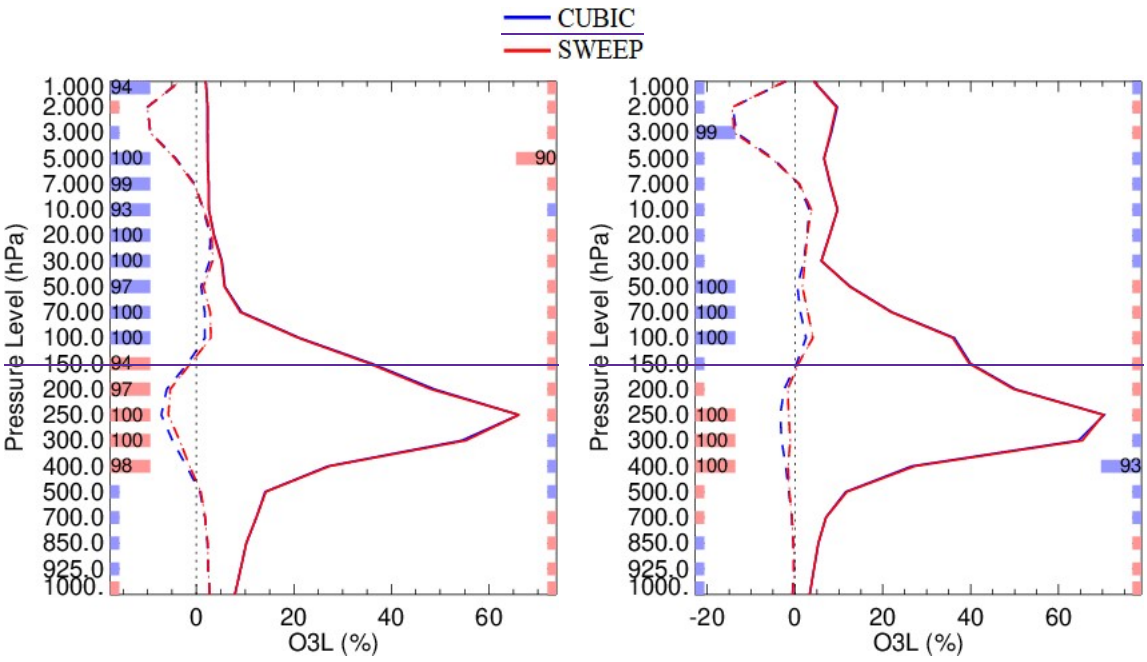
Asia



World

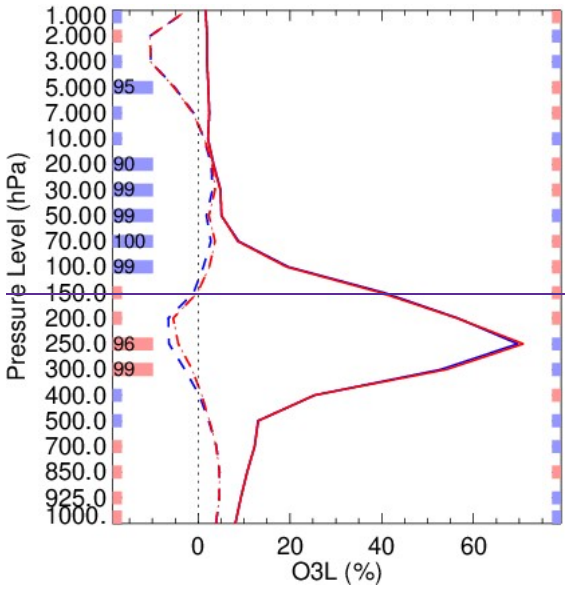


North Pole

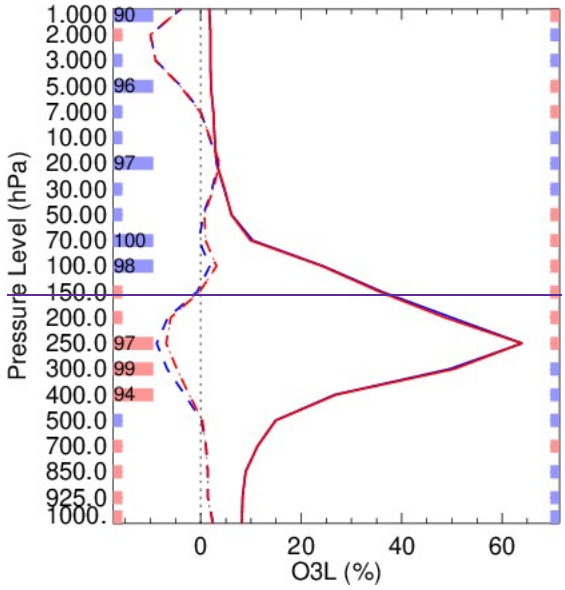


~~Northern extratropics~~

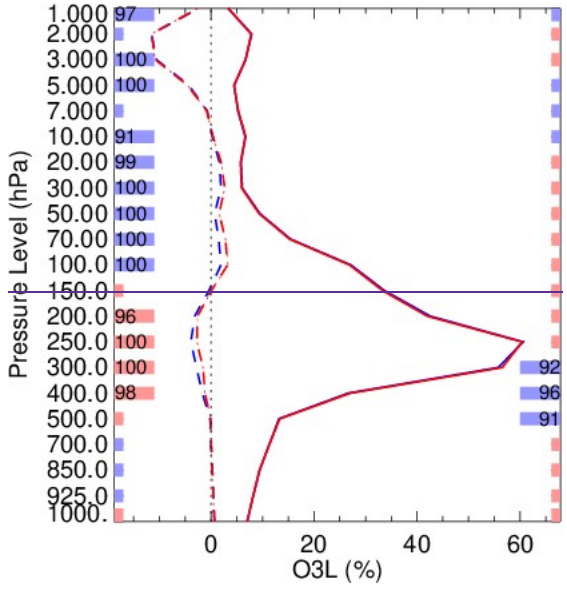
~~Southern extratropics~~



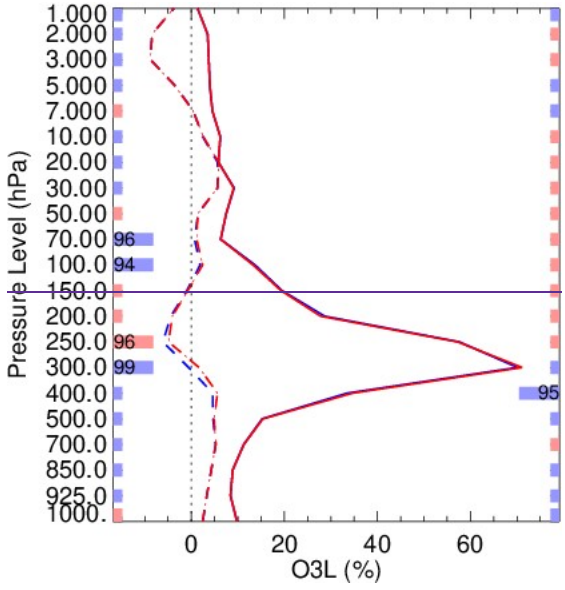
North America



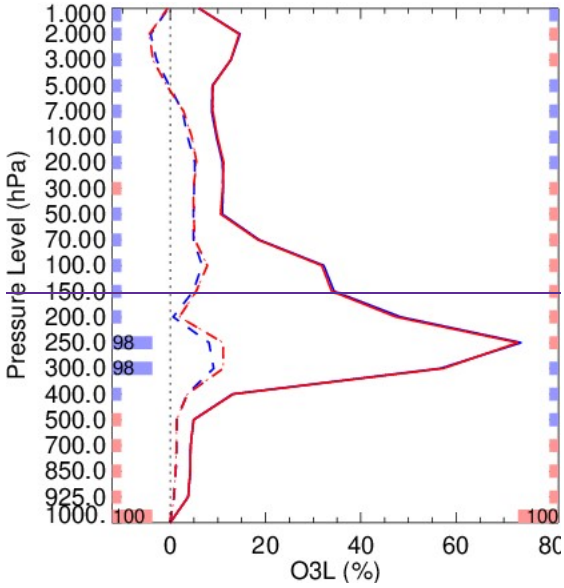
Asia



World



North Pole



~~South Pole~~

Figure 6. The bias (dashed line) and standard deviation (solid line) relative to GDPS analyses (%) for ozone forecasts at 240 hr lead-time from June 13th to August 31st 2019. Boxes on the left (right) indicate statistical significance levels for biases (standard deviation). Red (blue) boxes mean that the experiment which includes the sweep interpolation is better (worse). Small boxes mean that differences between both experiments are not statistically significant at the 90% confidence level.

The results from two interpolations, i.e. cubic and sweep interpolations, at different locations over the whole earth experiments demonstrate a good agreement between sweep interpolation and cubic interpolation. The two main metrics of the statistical accuracy, bias and standard deviation, are used to show the differences between the two experiments. In the following figures (Figure 6 and 7), solid and dashed lines represent the standard deviation and bias metrics, respectively. Figure 6 represents the average of all 10-day lead-time forecasts in the simulation time period. This figure shows that the standard deviations for both experiments are almost the same, but some minor deviations can be seen in the bias. The bias for ozone variation along the vertical direction is similar for both experiments in the lower part of troposphere (>500 hPa) and in the upper stratosphere (<50 hPa). The smaller concentration of ozone in these two regions accounts for the absence of sharp gradients in the vertical error profiles. The large maximum in the standard deviation around 250 hPa is associated with the transport of ozone across the tropopause. This process is sensitive to the choice of interpolation scheme and the results highlight its impact on stratosphere-troposphere exchange processes throughout the entire upper troposphere and lower stratosphere region between 400 and 50 hPa. In this region, a minor deviation of bias error can be seen between sweep interpolation and cubic interpolation, but this deviation is small and not significant (see Figure 6).

Figure 7 shows the evolution of the bias and (Charron et al., 2012) standard deviation as a function of lead-time for the North America region at the 200 hPa level, which corresponds to the region where both quantities vary the most abruptly along the

vertical direction. The upper panel shows the standard deviation and bias errors over 10 days prediction. It demonstrates that sweep interpolation is characterized by the same error growth as the cubic interpolation. The figure also shows some improvement in the bias error that was obtained by using sweep interpolation. ~~by using sweep interpolation.~~ One possible explanation for this phenomenon is that the lower order Lagrange interpolations used in the backward and forward steps of the sweep interpolation tend to generate less oscillation and dispersion errors than the standard fourth-order accurate cubic interpolation. ~~One possible explanation about this phenomenon is that lower order Lagrange interpolations tend to generate less oscillation and dispersion errors.~~ To confirm this explanation, further study needs to be done which is out of the scope of this paper.

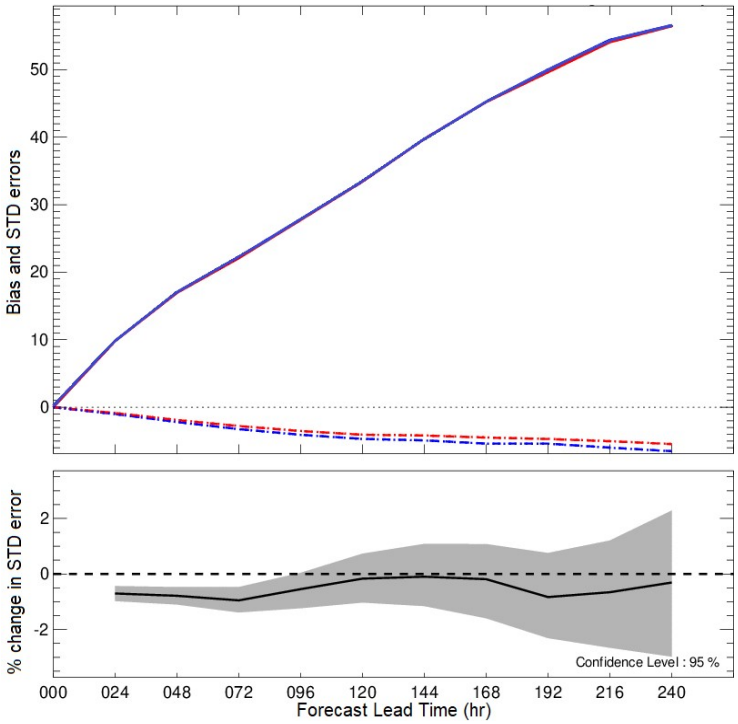


Figure 7. Standard deviation (solid line) and bias (dashed line) errors over North America for ozone at the level 200 hPa.

4 Impact of number of tracers on computational time

In the following, we investigate the impact of increasing the number of tracers on the timing of cubic interpolation process compared to sweep interpolation. We use the benchmark test case 2 (i.e., 24 hours forecasts of chemical CH_4 tracer) from the previous section and add additional tracers to the experiment. Note that in GEM we use cubic interpolation not only for advection solver, but also for exchanging the data between Yin and Yang grid (Qaddouri & Lee, 2011). In this study, we also implemented sweep interpolation into Yin and Yang exchange of data. Figure 8 shows the impact of sweep interpolation on the timing of the advection process, the exchange of data between the Yin and Yang sub-grids, and the total timing of the simulation. The numbers shown in ~~the figure~~Fig. 8 correspond to (time_cubic - time_sweep)/time_cubic x100. It is demonstrated that by increasing the number of tracers, the percentage of computational time saving increases for the semi-Lagrangian operation, Yin and Yang exchange, and the total timing of the simulation. When using only 20 tracers, the performance of sweep interpolation is almost the same as cubic interpolation. However, when the number of tracers is increased to 230, sweep interpolation can reduce the computational cost of simulation in the advection step by 25%. The sweep interpolation reduction in Yin and Yang exchange is about 18.5%, and its impact on the total timing of the simulation is about 16%.

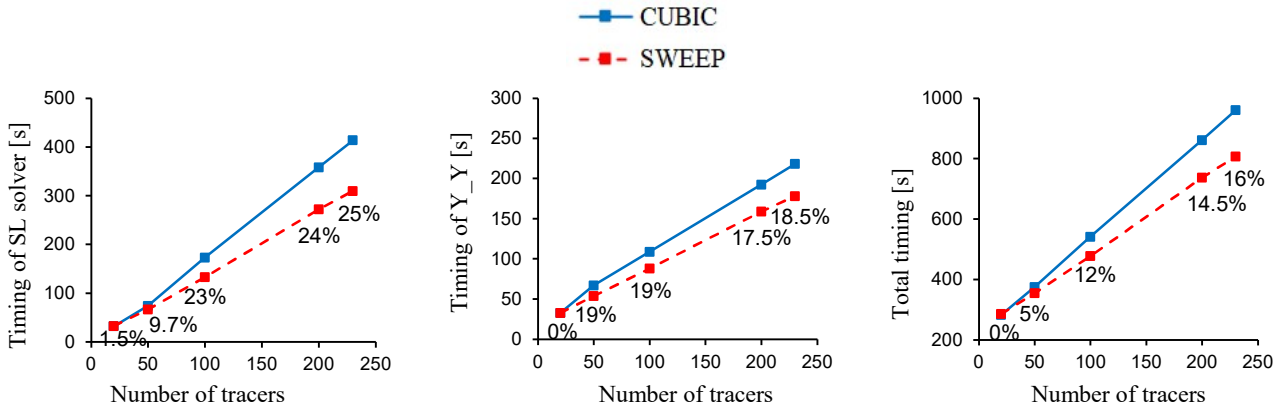


Figure 8. Computational time analysis based on the number of tracers.

5 Conclusion

In this study, we implemented the sweep interpolation scheme (an integration of two 3rd-order backward and forward interpolation steps) into ECCC's GEM model. Two theoretical benchmarks and a global ozone forecast experiment were used to investigate the performance of the model. Based on the results, the sweep interpolation results are in a good agreement with those from the standard cubic interpolation, but ~~begin~~ achieved at a lower computational cost~~sweep interpolation has a good agreement with cubic interpolation with lower computational cost~~.

The main benefit of using sweep interpolation in GEM is its lower computational cost compared to the cubic interpolation. By using this interpolation, we can reduce the timing of semi-Lagrangian interpolation, Yin-Yang exchange, and the total wall clock time of the simulation up to 25%, 19%, and 15%, respectively, for the experiments with around 250 tracers. Additionally, this replacement is very simple with minimum modifications required in the model. Sweep interpolation is suitable to be implemented in other NWP models, which also rely on a SL approach and polynomial interpolation schemes.

Although sweep interpolation is developed based on 3rd~~d~~-order backward and forward Lagrange interpolation schemes, it can control the growth of the mass conservation error over the forecast lead-time in a way that agrees very closely to the error growth of the standard cubic interpolation. The ~~it can control the growth of error over the forecast lead-time comparable to the cubic interpolation.~~ Ozone forecasting experiment using the global version of the GEM model shows that the sweep interpolation has a small impact on the ozone distribution along the vertical direction. ~~–~~The impact is ~~particularly important~~the most significant in the tropopause region where transport processes play a significant role. The overall impact of the new scheme on model biases and standard deviations was evaluated at different lead-times, which shows that the overall performance of the ozone forecasting system has not suffered from the use of a faster interpolation approach. In some cases, we found that sweep interpolation can perform better and reduce the numerical errors compared to those of the standard cubic interpolation. The impact on biases in ozone experiments generally increase with longer lead-time and is larger over some regions driven by transport processes. This improvement is likely related to the lower oscillations generated

by the lower order of Lagrange interpolation schemes used in sweep compared to the standard cubic interpolation, but this needs to be further investigated in a future study.

Code availability. GEM (Global Environmental Multi-scale), the Environment and Climate Change Canada's online weather prediction model, is a free software which can be redistributed and/or modified under the terms of the GNU Lesser General Public License as published by the Free Software Foundation. The GEM model is available to download from: https://github.com/ECCC-ASTD-MRD/gem/. The git repository contains the instructions that explain how to install the code, download the input data and run the model. The modified interpolation code can be downloaded from the Zenodo website: https://doi.org/10.5281/zenodo.8246831. The model output requires a large amount of memory space in a binary format specific to Environment and Climate Change Canada's modelling systems. The conversion to other formats is possible by an email request to Ashu Dastoor (ashu.dastoor@canada.ca).

Author contributions. In this research work, MM developed the model, wrote the code, performed the simulation, and was responsible for analysis of the results, and writing the manuscript. JFC was responsible for writing the code, analysis the results, and writing the manuscript. AD was responsible for the study concept, analysis of model simulations, and writing the manuscript. JDG and II were responsible for GEM-O3 set-up and error analysis of the model. AQ was responsible for writing the manuscript. All authors contributed to the editing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank our ECCC colleagues Rabah Aider, Andrei Ryjkov, and Kenjiro Toyota for their help with running GEM and visualizing the results.

References

Aires, F., Catherine, P., and Rossow, W. B.: Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions, J. Geophys. Res.-Atmos., 109, D04313, doi:10.1029/2003JD003527, 2004.

Arakawa, A.: Finite-difference methods in climate modelling, In Physically-Based Modelling and Simulation of Climate and Climatic Change, 243, 79–168, doi: 10.1007/978-94-009-3041-4_3, 1988.

Bradley, A., Bosler, P., and Guba, O.: Islet: interpolation semi-Lagrangian element-based transport, Geos. Model Dev., 15, 6285–6310, doi: 10.5194/gmd-15-6285-2022, 2022.

Charron, M., Polavarapu, S., Buehner, M., Vaillancourt, P. A., Charette, C., Roch, M., Morneau, J., Garand, L., Aparicio, J. M., MacPherson, S., Pellerin, S., St-James, J., and Heilliette, S.: The Stratospheric Extension of the Canadian Global

Deterministic Medium-Range Weather Forecasting System and Its Impact on Tropospheric Forecasts, Mon. Weather Rev., 140(6), 1924–1944, doi: 10.1175/MWR-D-11-00097.1, 2012.

Cossette, J.-F., Smolarkiewicz, P.K., and Charbonneau, P.: The Monge-Ampère trajectory correction for semi-Lagrangian schemes, J. Comput. Phys., 274, 208-229, doi:10.1016/j.jcp.2014.05.016 , 2014.

Côté, J., and Staniforth, A.: A two-time-level semi-Lagrangian semi-implicit scheme for spectral models, Mon. Weather Rev., 2003-2012, doi: 10.1175/1520-0493(1988)116<2003:ATTLSL>2.0.CO;2, 1988.

Courant, R., Isaacson, E., and Rees, M.: On the solution of nonlinear hyperbolic differential equations by finite differences, Commun. Pure Appl. Math., V, 243–255, doi: 10.1002/cpa.3160050303, 1952.

de Grandpré, J., Tanguay, M., Qaddouri, A., Zerroukat, M., and McLinden, C.: Semi-Lagrangian Advection of Stratospheric Ozone on a Yin–Yang Grid System, Mon. Weather Rev., 1035–1050, doi: 10.1175/MWR-D-15-0142.1, 2016.

Girard, C., Plante, A., Desgagné, M., Mctaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P. A., and Zadra, A.: Staggered vertical discretization of the canadian environmental multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type, Mon. Weather Rev., 1183–1196, doi: 10.1175/MWR-D-13-00255.1, 2014.

Golaz, J., Caldwell, P., Van Roekel, L., Petersen, M., Tang, Q., Wolfe, J., ... and Baldwin, S.: The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution, Adv. in Modelling Earth Syst., 2089–2129, doi: 10.1029/2018MS001603, 2019.

Hortal, M.: The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model, Q. J. R. Meteorol. Soc., 128, 1671–1687, doi: 10.1002/qj.200212858314, 2002.

Husain, S. Z., and Girard, C.: Impact of consistent semi-Lagrangian trajectory calculations on numerical weather prediction performance, Mon. Weather Rev., 4127–4150, doi: 10.1175/MWR-D-17-0138.1, 2017.

Husain, S. Z., Girard, C., Qaddouri, A., and Plante, A.: A new dynamical core of the Global Environmental Multiscale (GEM) model with a height-based terrain-following vertical coordinate, Mon. Weather Rev., 2555–2578, doi: 10.1175/MWR-D-18-0438.1, 2019.

Husain, S. Z., Girard, C., Separovic, L., Plante, A. and Corvec, S.: On the progressive attenuation of finescale orography contributions to the vertical coordinate surfaces within a terrain-following coordinate system. Mon. Weather Rev., 148(10), 4143–4158, doi: 10.1175/MWR-D-20-0085.1, 2020.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., ... and Galmarini, S.: Evaluation of operational online-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part II: Particulate matter, Atmos. Sci., 115, 421–441, doi: 10.1016/j.atmosenv.2014.08.072, 2015.

Makar, P., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., ... and Galmarini, S.: Feedbacks between air pollution and weather, Part 1: Effects on weather, Atmos. Environ., 442–469, doi: 10.1016/j.atmosenv.2014.12.003, 2015.

McDonald, A., and Bates, J. R.: Improving the estimate of the departure point position in a two-time level semi-Lagrangian and semi-implicit scheme, Monthly Weather Review, 737–739, doi: 10.1175/1520-0493(1987)115<0737:ITEOTD>2.0.CO;2, 1987.

McDonald, A., and Haugen, J.: A two-time-level, three-dimensional semi-Lagrangian, semi-implicit, limited-area gridpoint model of the primitive equations, Mon. weather rev., 2603–2621, doi: 10.1175/1520-0493(1992)120<2603:ATTLTD>2.0.CO;2, 1992.

Mortezazadeh, M., and Wang, L.: A high-order backward forward sweep interpolating algorithm for semi-Lagrangian method, Int. J. Num. Methods in Fluids, 84, 584–597, doi: 10.1002/fld.4362, 2017.

Mukhopadhyay, P., Prasad, V., Krishna, R., Deshpande, M., Ganai, M., Tirkey, S., ... and Mahakur, M.: ~~(2019).~~ Performance of a very high-resolution global forecast system model (GFS T1534) at 12.5 km over the Indian region during the 2016–2017 monsoon seasons, Earth Sys. Sci., 128, 1–18, doi: 10.1007/s12040-019-1186-6, 2019.

Nair, R. D., and Machenhauer, B.: The mass-conservative cell-integrated semi-Lagrangian advection scheme on the sphere, Monthly Weather Review, 649–667, doi: 10.1175/1520-0493(2002)130<0649:TMCCIS>2.0.CO;2, 2002.

Nakao, J., Chen, J., and Qiu, J.: An Eulerian-Lagrangian Runge-Kutta finite volume (EL-RK-FV) method for solving convection and convection-diffusion equations, J. Comput. Phys., 470, 111589, doi: 10.1016/j.jcp.2022.111589, 2022.

Petras, A., Ling, L., and Ruuth, S.: Meshfree Semi-Lagrangian Methods for Solving Surface Advection PDEs, J. Sci. Comput., 93, 11, doi: 10.1007/s10915-022-01966-w, 2022.

Prather, M. J., Holmes, C. D., and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophys. Res. Let., 39, 9, doi: 10.1029/2012GL051440, 2012.

Purser, R., and Leslie, L.: An efficient interpolation procedure for high-order three-dimensional semi-Lagrangian models, Mon. Weather Rev., 119, 2492–2498, doi: 10.1175/1520-0493(1991)119<2492:AEIPFH>2.0.CO;2, 1991.

Qaddouri, A., and Lee, V.: The Canadian global environmental multiscale model on the Yin-Yang grid system, Quarterly J. Royal Meteorol. Soc., 137, 1913–1926, doi: 10.1002/qj.873, 2011.

Ritchie, H. Application of the semi-Lagrangian method to a spectral model of the shallow water equations, Mon. Weather Rev., 1587–1598, doi: 10.1175/1520-0493(1988)116<1587:AOTSLM>2.0.CO;2, 1988.

Ritchie, H., Temperton, C., Simmons, A., Hortal, M., Davies, T., Dent, D., and Hamrud, M.: Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF forecast model, Mon. Weather Rev., 489–514, doi: 10.1175/1520-0493(1995)123<0489:IOTSLM>2.0.CO;2, 1995.

Robert, A. (1980). A stable numerical integration scheme for the primitive meteorological equations. Atmos. Ocean, 19(1), 35–46, doi: 10.1080/07055900.1981.9649098, 1980.

Shapiro, M., and Hastings, J.: Objective cross-section analyses by Hermite polynomial interpolation on isentropic surfaces, App. Meteorol. and Climatol., 753–762, doi: 10.1175/1520-0450(1973)012<0753:OCSABH>2.0.CO;2, 1973.

Smolarkiewicz, P. K., and Pudykiewicz, J. A.: A Class of Semi-Lagrangian Approximations for Fluids, Atmos. Sci., 49(22), 2082–2096, doi: 10.1175/1520-0469(1992)049<2082:ACOSLA>2.0.CO;2, 1992.

Staniforth, A., and Côté, J.: Semi-Lagrangian integration schemes for atmospheric models—A review, Mon. weather rev., 119, 2206–2223, doi: 10.1175/1520-0493(1991)119<2206:SLISFA>2.0.CO;2, 1991.

Stevens, R., Ryjkov, A., Majdzadeh, M., and Dastoor, A.: An improved representation of aerosol mixing state for air quality–weather interactions, Atmos. Chem. Phys., 22, 13527–13549, doi: 10.5194/acp-22-13527-2022, 2022.

Ullrich, P. A., Jablonowski, C., Kent, J., Lauritzen, P. H., Nair, R. D., and Taylor, M. A.: Dynamical Core Model Intercomparison Project (DCMIP) Test Case Document. DCMIP Summer School, 2013.

Williamson, D. L., and Rasch, P. J.: Two-dimensional semi-Lagrangian transport with shape-preserving interpolation, Mon. Weather Rev., 117, 102–129, doi: 10.1175/1520-0493(1989)117<0102:TDSLTW>2.0.CO;2, 1989.

Yabe, T., Tanaka, R., Nakamura, T., and Xiao, F.: An exactly conservative semi-Lagrangian scheme (CIP–CSL) in one dimension, Mon. Weather Rev., 129, 332–344, doi: 10.1175/1520-0493(2001)129<0332:AECSLS>2.0.CO;2, 2021.

Zerroukat, M.: A simple mass conserving semi-Lagrangian scheme for transport problems, J. Comp. Phys., 229, 9011–9019, doi: 10.1016/j.jcp.2010.08.017, 2010.

Zhou, J., Obrist, D., Dastoor, A., Jiskra, M., and Ryjkov, A.: Vegetation uptake of mercury and impacts on global cycling, nat. rev. earth & environ., 2, 269–284, doi: 10.1038/s43017-021-00146-y, 2021.