Response to Reviewer 5

We are incredibly grateful for your efficient review process. Your insightful comments have provided valuable guidance for revising the manuscript. We have revised the manuscript according to your suggestions and will respond to your comments paragraph by paragraph. The comments are given below in black, our responses are in blue, and proposed changes to the manuscript are in red. Additional references are provided at the end of this document. The final revisions and specific locations corresponding to the manuscript will be marked uniformly after receiving feedback from other reviewers.

Specific comments:

- The algorithm description needs more attention and description in the text. By comparison, the description of MPAS seems much better detailed and it could be argued that much of it is less important. I recommend finding ways to enhance the content of section 2.1.

Thank you for your valuable feedback. The algorithm itself is proposed by Møller (1965), and it only apply in the step-by-step integration of ordinary differential equations, how to apply it in numerical models is core content, so to address this, we will:

1. Add the clear description of the solution method for the equations including temporal integration scheme and spatial discretization scheme in section 2.2. Moreover, the algorithm applied here primarily addresses the rounding error compensation between large and small numbers in addition. Currently, it is only applicable to the time integration process and has not been implemented in the spatial discretization process. I apologize for not mentioning that part in the manuscript. I will revise and provide the supplement. See 2.2 additional content below.
2. Enhance the algorithm description in Section 2.3. Specify at which step the quasi double-precision algorithm is applied within the computation process in section 2.3 and replace the figure (corresponding to Figure 3 in the manuscript) and explain this process using formulas and explanations. See 2.3 additional content below.
3. we will select a representative equation and include a detailed algorithmic description of its iterative calculation process in the Supplement (see Figure 1). We believe this addition will be beneficial for readers who wish to gain a deeper understanding of the computational aspects of our method.

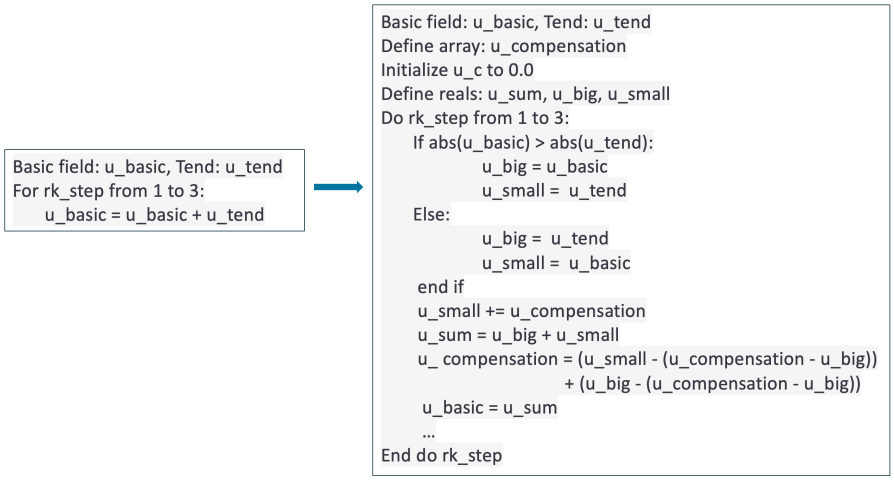

Figure 1. The pseudo-code for variable of U.

3.2 Additional content

The MPAS-A solves the fully compressible, nonhydrostatic equations of motion (Skamarock et al. 2012). The spatial discretization uses a horizontal (spherical) centroidal Voronoi mesh with a terrain-following geometric-height vertical coordinate and C-grid staggering for momentum. The temporal discretization uses the explicit time-split Runge–Kutta technique from Wicker and Skamarock (2002) and Klemp et al. (2007).

The algorithm applied here primarily addresses the rounding error compensation between large and small numbers in addition. Currently, it is only applicable to the time integration process and has not been implemented in the spatial discretization process. Therefore, this section will provide a detailed introduction to the time integration scheme. For the spatial discretization scheme, please refer to Skamarock et al. (2012), and it will not be introduced upon here.

The formulation of the scheme can be considered in on dimension as equation Wicker and Skamarock (2002):

$$\frac{\partial \emptyset}{\partial t} = RHS_{\emptyset} \tag{1}$$

The variable Ø represents any prognostic variable in the prognostic equations, while RHS represents the right-hand side of the prognostic equations (i.e., the spatial discretization equation). In MPAS-A, a forward-in-time finite difference is used, and it can be written as Eq. (2):

$$\frac{{\emptyset_i}^{n+1} - {\emptyset_i}^{n}}{\Delta t} = RHS_{\emptyset} \tag{2}$$

Where superscript represent the time step, and subscript represent the position of grid zone.

The two-order Runge-Kutta time scheme is used in MPAS-A as described in Gear et al. (1971):

$$\emptyset^{*} = \emptyset^{t} + \frac{\Delta t}{2} * RHS(\emptyset^{t}) \tag{3}$$

$$\emptyset^{**} = \emptyset^{t} + \frac{\Delta t}{2} * RHS(\emptyset^{*}) \tag{4}$$

$$\emptyset^{t+\Delta t} = \emptyset^{t} + \Delta t * RHS(\emptyset^{**}) \tag{5}$$

3.3 Additional content

According to Equation Eq. (3), (4) and (5), it can be observed that in the time integration scheme, each step involves the process of adding tends on the basic field $\emptyset^{t}$. In numerical models, the basic field is generally much larger than the tends, which aligns with the principles of numerical computation regarding the addition of large and small numbers, as well as the time integration process. It is important to note that the quasi double-precision algorithm currently only addresses time integration and has not been validated during the spatial discretization process. The spatial discretization primarily involves subtraction, specifically the subtraction of a small number from a large number or the subtraction of two close values. Whether this algorithm is applicable in spatial discretization remains uncertain, therefore, we will not apply it in this context.

Based on the application principles of the algorithm, which involve the processes of adding large and small numbers as well as the time integration process, we have established a strategy for applying the quasi double-precision algorithm within the MPAS-A. Specific improvements are provided based on the predictive equations:

$$\frac{\partial \boldsymbol{V}_H}{\partial t} = -\frac{\rho_d}{\rho_m}\left[\nabla_\zeta\left(\frac{p}{\zeta_z}\right) - \frac{\partial z_H p}{\partial \zeta}\right] - \eta \boldsymbol{k} \times \boldsymbol{V}_H - \boldsymbol{\nu}_H \nabla_\zeta \cdot \boldsymbol{V} - \frac{\partial \Omega \nu_H}{\partial \zeta} - \rho_d \nabla_\zeta K - eWcos\alpha_r - \frac{\nu_H W}{r_e} + \boldsymbol{F}_{\boldsymbol{V}_H} \tag{6}$$

$$\frac{\partial W}{\partial t} = -\frac{\rho_d}{\rho_m}\left[\frac{\partial p}{\partial \zeta} + g\tilde{\rho}_m\right] - (\nabla \cdot \boldsymbol{\nu} W)_\zeta + \frac{uU+vV}{r_e} + e(Ucos\alpha_r - Vsin\alpha_r) + F_W \tag{7}$$

$$\frac{\partial \Theta_m}{\partial t} = -(\nabla \cdot \boldsymbol{V}\theta_m)_\zeta + F_{\Theta_m} \tag{8}$$

$$\frac{\partial \tilde{\rho}_d}{\partial t} = -(\nabla \cdot \boldsymbol{V})_\zeta \tag{9}$$

The meaning of each variable in the equations exactly follows Skamarock et al. (2012), so that we don't repeating explanation. For a numerical model, the most crucial variables are the prognostic variables. Therefore, In the MPAS-A model we applied the quasi double-precision algorithm to the time integration process of these prognostic variables, including horizontal momentum ($\boldsymbol{V}_H$) , dry air density ($\tilde{\rho}_d$), potential temperature (Θ_m) and vertical velocity (W) , that is , the process in red of Eq. (6), (7), (8) and (9). (Only the predictive equations for the dynamic core are presented here, without the scalar transport.)

- For clarification, is the QDP runtime is 2% more expensive compared to the DBL runtime or the SGL runtime? Also how much reduction in runtime was there for the SGL version compared to the DBL version? This can be used to give context to what benefits are achieved for the more noticeable errors in SGL.

Thanks for your suggestion, following your insightful comment, we found the previously reported figures were indeed rough estimations, so we have re-evaluated the exact computational performance using a measurement tool to determine the runtimes. So, we will:

1. Add a dedicated section 3.4 in the paper to describe the computational performance. The size of the computational performance will be represented in terms of runtime, and we will discuss the runtime for each case in tabular form.
2. Revise the description of computational performance in the abstract to reflect these updated and more accurate measurements.

3.4 Computational performance

In comparison with the SGL, although there is a slight increase in runtime, it is minimal, at only 6% (Jablonowski and Williamson baroclinic wave), 0.3% (Super-cell), 2% (Real data with resolution of 120km) and 18% (Real data with resolution of 240km) (Table 1). This slight increase is attributed to the addition of a small number of global variable arrays when using quasi double-precision. And compared to DBL, QDP demonstrated relatively better performance across different cases, reducing the runtime by 29% (Jablonowski and Williamson baroclinic wave), 29% (Super-cell), 21% (Real data with resolution of 120km) and 6% (Real data with resolution of 240km) (Table 1).

Table 1. Elapsed time of DBL, SGL and QDP test (unit:s).

Case name	DBL	SGL	QDP
Jablonowski and Williamson baroclinic wave	1768	1191	1263
Super-cell	1507	1073	1077
Real data with resolution of 120km	19126	14765	15092
Real data with resolution of 240km	1397	1118	1317

Abstract

The content ‘The round-off error of surface pressure is reduced by 68%, 75%, 97%, 96% in cases, the memory has been reduced by almost half, while the computation increases only 2%, significantly reducing computational cost.’ will be revised to ‘The bias of surface pressure are reduced respectively by 68%, 75%, 97% and 96% in cases, the memory has been reduced by almost half, while the computation increases only 6%, 0.3%, 2%, and 18% in cases, significantly reducing computational cost.’

- Given errors appear to become noticeably larger beyond 10+ days, how are these figures dependent

with different spatial time average lengths? Do they all only look drastically different because of errors after round off error becomes significant around day 10+? Is this error tied to simulation length or is this more determined by the number of time steps taken?

Thank you for your insightful questions. As you noted, errors become significantly larger after approximately 10 days due to factors such as round-off errors arising from reduced numerical precision and energy loss during the propagation process. Around day 10, the accumulation of rounding errors becomes pronounced, continuing to accumulate over time, which can ultimately lead to model instability.

In this study, we did not conduct a comparative analysis involving different time step lengths; therefore, we cannot definitively conclude whether the observed errors are related to the chosen time step sizes. However, round-off errors do accumulate with the simulation length.

We appreciate your suggestion, and in future research, we will focus on whether integration step lengths have an impact on error growth.

- In some experiments, the reduction of error is a large percentage but feels like a small error. In general to the reader, it's not very clear that reducing these errors will matter that much. For example, while the error is reduced using QDP in a lot of cases, reduction of the surface pressure (hPa) error by a RSME of 2.8x10-1 to 1.4x10-1 seems like a small amount. Likewise for other sources of error and that this all needs context. It is important to consider these errors relative to other sources of model errors.

Thank for your suggestion. The core of this study is the introduction of error compensation methods (quasi double-precision). By using error compensation methods (quasi double-precision), we can maintain integration stability comparable to that applying double precision scheme while significantly reducing memory requirements by lowering the numerical precision of all variables and improved the accuracy comparable to that applying the single precision. This approach not only reduces communication pressure but also allows for substantial increases in computational speed through vectorization optimization. So, we will add a section to describe it (see 3.4 Computational performance)

And we will add an analysis of other sources of error in Section 3.1. The content will include the following:

3.1 Additional content

The sources of unpredictability, as noted by Bauer et al. (2015), include instabilities that inject chaotic ‘noise’ at small scales and the upscale propagation of their energy. For the cases examined, both SGL and QDP begin to exhibit errors after 10 days of integration. These errors arise from factors such as rounding errors due to reduced numerical precision and energy loss during the propagation process. The quasi double-precision algorithm can reduce the impacts of these errors.

While we acknowledge other potential sources of uncertainty, such as initial condition errors, we have not conducted an in-depth study on them in this research. Our primary focus remains on evaluating the improvements provided by the compensation algorithm in addressing rounding errors.

- Additionally, I was wondering if the authors could leverage the test case where grid resolution was also incorporated as the model discretization is a source of model error compared to the model error due to precision differences.

Thank you for your insightful comments. As you mentioned, there are multiple factors contributing to the errors observed in our study. In addition to round-off errors associated with floating-point arithmetic, the choice of grid resolution also has a significant impact on bias. We will add a description in Section 3.3, you can see as follows:

3.3 Additional content

In this research, we focus on the processes of summing the basic field and trends. When the resolution is increased, the basic field remains relatively unchanged; however, the trends become smaller. This characteristic aligns with the nature of adding large and small numbers, making the advantages of the quasi double-precision algorithm more pronounced. Thus, it is evident from Figure 8 that as the resolution increases, the improvement achieved by quasi double-precision algorithm also enhances.

On the other hand, it is important to note that the propagation of rounding errors is not immediately apparent over short time scales. However, as the number of iterations increases, these errors can become more significant. The quasi double-precision algorithm employs compensation mechanisms that help mitigate the propagation of these errors.

- For the cases in 3.3, it is not clear why the decision to have the higher resolution simulation of 120 km x 120 km have the same time step of 720 seconds as the 240 km x 240 km case. This seems rather unusual as typically a reduction in grid cell size is accompanied by a proportionate reduction in time step size to maintain a consistent Courant number.

Thank you for your question. As you mentioned, it is common for time step sizes to decrease proportionally with grid resolution. However, in our experiments, we found that the simulation could run stably at the finer resolution (i.e., 120 km) without requiring a reduction in the time step size. To minimize errors arising from differences in resolution during data processing, we chose to maintain the same time step across both simulations.

Minor comments:

- Figure 1, Figure 2 and a lesser extent Figure 3: The algorithm are written in a way that is difficult to read. I suggest that it is better distinguished the text (i.e. "evaluation") from variables (i.e. s,u,v). In Figure 2, everything is appears quite mathematical while Figure 3 is quite the opposite. It would be beneficial to have these all looking the same.

Thank you for your valuable feedback. We have made the necessary modifications. For details on these changes, please refer to the responses provided in the first specific comments.

- Line 150: "so we close the scalar transport in all cases". Does this mean it's turned off/not solved?

Unless "closed" is a common term in the MPAS community, I suggest the authors consider a term closer to "disable/disabled" instead of "close/closed" when talking about model processes that are not absent in the simulation.

Thank you for your insightful comment. We will revise it from "closed" to "turned off".

- Line 179: Has this set of simulation parameters been utilized in a previous published work as a test case? This would be preferred to stating it was on a website that is subject to change.

Thank you for your suggestion. We will revise this section to clarify the source of our simulation parameters. The modified content will state:

The experimental configuration is consistent with the test case presented by Jablonowski and Williamson (2006).

- Line 195 states the resolution is 84 km x 84 km where I believe this is actually the total domain size and not resolution. Assuming my understanding is correct, how many grid cells are within this domain and/or what is the average grid cell size in this simulation? Is it the ~500 m in Klemp 2015?

Thank you for your question. As you pointed out, 84 km × 84 km refers to the total domain size. We will revise this in the manuscript. Within this domain, there are a total of 28,080 grid cells, which is not the 500 m that you mentioned.

- Line 197: While the errors in Figure 6(a) are indeed quite small, it may be beneficial to say how small relative to the actual values of kinetic energy? And I think a more clear statement explaining this figure's behavior would help clarify the plot as it's not the expected behavior throughout the manuscript that QDP does better in terms of error.

Thank you for your insightful comment regarding Figure 6(a) of section 3.1. According to RC1, we carefully reviewed our code based on your comments and identified that the issue was a problem with the data processing and plotting. We have corrected this issue, and the revised Figure 2 is provided below, it can be seen, the average bias between DBL and QDP is smaller than DBL and SGL.

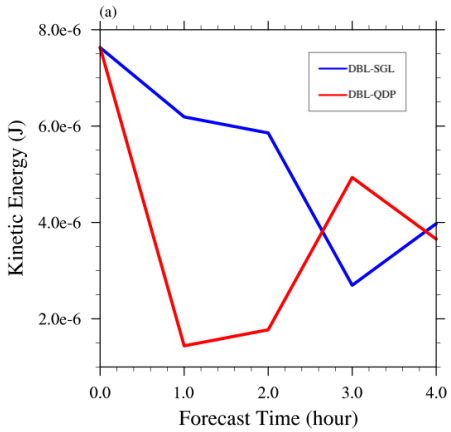

Figure 2. The temporal evolution of spatially averaged difference of kinetic energy between DBL and SGL, as well as difference between DBL and QDP in case of super-cell.

- Figure 7: The location of the Y (km) should be moved up near the axes (and should probably be labeled X?) and not under the color bar and Z is used which is typically thought of as a vertical direction is on the y axis. Color bar should probably have different scaling with a smaller range. Additionally why are the circles different colors between the different panels?

Thank you for your valuable suggestions. We have made the necessary modifications as follows, as seen in Figure 3 (which corresponds to Figure 7 in the manuscript):

1. The axis labels have been updated.
2. The position of the color bar has been adjusted.
3. The range of the color bar has been modified for better scaling.

Furthermore, we have added a note in the caption to clarify the differences in circle colors between the various panels.

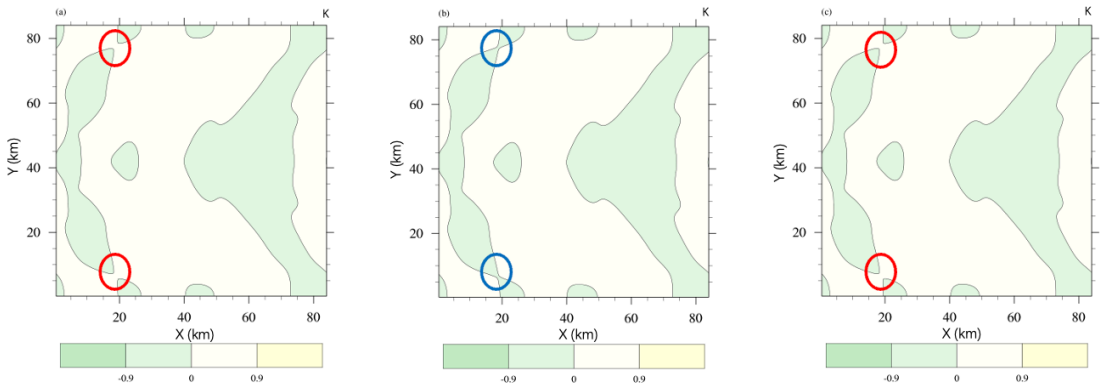

Figure 3. Perturbation theta in super-cell development at 5400s in the (a) DBL simulation, (b) SGL simulation and (c) QDP simulation (bias has reduced), unit: K, the circle represents the pattern bias (the same color means the consistent value).

- Figure 9 is one of the only figures that has color bar limits change between the SGL and QDP panels. Given that quite often the color bar limit have been fixed ranges in most figures, a note in the caption may be beneficial for this.

Thank you for your suggestion. We will include a note in the caption of Figure 9 indicating that the color bars used in panels (a) and (b) are different. This will help clarify the distinction for readers.

- Figure 10(b) uses the same color bar as Figure 10(a) but reveals no detail because the same color bar limits are too large. I suggest that a new color bar range is selected (and it is then noted in the caption as being different in an and b as suggested for Figure 9).

Thank you for your suggestion. I believe you may be referring to Figure 11 rather than Figure 10. I will revise the color bar for Figure 11 to improve clarity and detail, as well as update the caption accordingly. Please see below for Figure 4 (corresponding to Figure 11 in the manuscript).

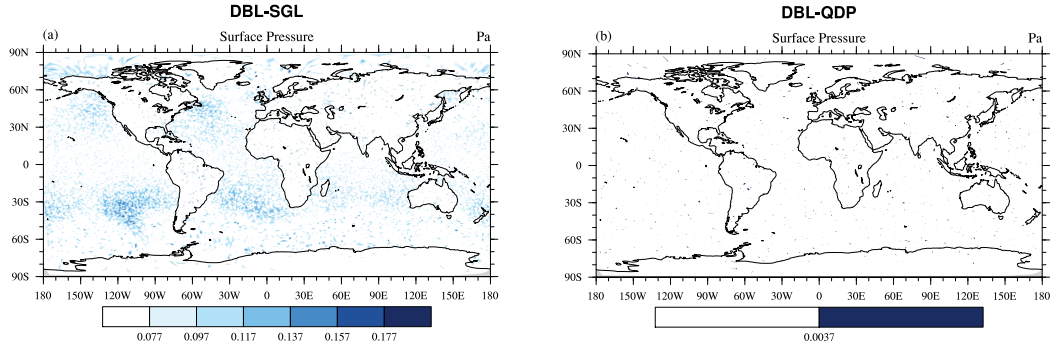

Figure 4. distributions of averaged (1-15days) difference of surface pressure (units: Pa) between DBL and (a) SGL simulation, (b) QDP simulation (resolution: 120 km × 120 km) (round-off error has reduced). The RMSE of surface pressure between DBL and (a) SGL simulation is 6.33×10^{-2} Pa, (b) QDP simulation is 2.25×10^{-2} Pa. (The color bars in (a) and (b) are different)

- For code and data availability section, it is mentioned that "model code and plotting data" is available While I was able to find the versions of the model code, the simulation inputs and plotting scripts, I was unable to locate actual data. It's unclear if this is intended as the wording indicates "plotting data" and we should rather anticipate running the model to produce the data.

Thank you for your valuable feedback. Due to the large size of the output data generated by the model, we have not uploaded the actual output files. If you would like to obtain the processed output data, you can compile and run the model following the instructions in the README file, which will allow you to generate the output data.

To enhance clarity, I will revise the description in the "Code and Data Availability" section accordingly. The updated wording is provided below.

Code and data availability. Model code and plotting data related to this manuscript is available at:

https://doi.org/10.5281/zenodo.13765422. Details regarding the code structure and instructions for running the code are provided in the supplementary material, which can be downloaded and viewed in Fig. S1. This figure provides a visual overview of the code organization. The information of steps how to execute the simulations can be found in README file in each test case folder.

Technical comments:

- Negative exponents have been routinely miswritten throughout the manuscript such as 10-2 rather than 10-2.

Thank you for your suggestion. We have corrected these errors throughout the manuscript.

- Line 218: "The Spatial RMSE of with 120 km x 120km.." is missing the variable name, which appears to be surface pressure.

Thank you for your observation. I appreciate your attention to detail. I have corrected this in the revised manuscript by including the variable name, which is indeed surface pressure.

JABLONOWSKI C, WILLIAMSON D L.: A baroclinic instability test case for atmospheric model dynamical cores[J/OL]. Quarterly journal of the Royal Meteorological Society, 132(621C): 2943-2975. DOI:10.1256/qj.06.12, 2006.