Reply to Referee#2

We appreciate your careful reading of the manuscript and thoughtful comments for improving the manuscript. For your concern about English, we would like to inform you that the revised manuscript will be edited by professional English editor. Please find below a point-by-point response to each of the comments. The original comments are in italics.

Specific comments:

1) Motivation: what is a typical situation in which the direction splitting would be of advantage over unsplit SEM or finite-difference approaches? At least a reference would be useful to better assess the scope of the paper here.

=> In models using the unsplit spectral element method (SEM), existing physics packages cannot be plugged in directly to the models. On the other side, in case of models using the unsplit (higher order) finite-difference method (FDM), the models have limited scalability in many-core computers. Therefore, the direction splitting method (horizontally SEM vertically FDM) could have an advantage of both good scalability and existing physics packages.

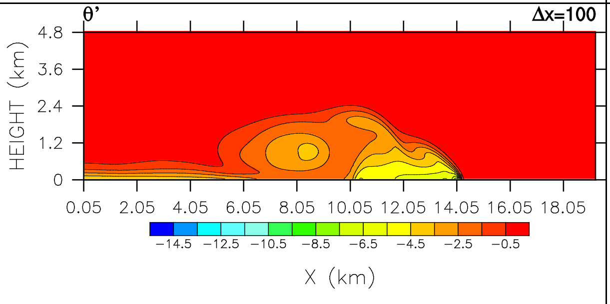
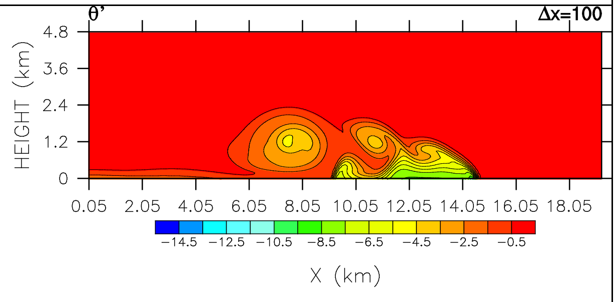
2) Time discretization: which order of accuracy is the employed discretization? According to Skamarock and Klemp (2008) the method should be second-order for nonlinear equations, but no discussion is present in Section 3.2 of the manuscript about this point. Furthermore, split-explicit models usually employ divergence damping for stability reasons. Is that the case for the discretization proposed in this paper as well?

=> Because our model uses the time integration technique of Skamarock and Klemp (2008), the time integration technique is third-order accurate for linear equations and second-order accurate for nonlinear equations. The divergence damping is used too with the damping coefficient of 0.1 which is the same as it used in WRF. We will revise accordingly.

3) Some pointwise comparison with WRF model on the same tests at a given resolution would be helpful to better assess the results, see e.g. figures 4 and 5 of Skamarock and Klemp (2008). If an accuracy gain is achieved, this should be explained and documented. Otherwise it is difficult for the reader to understand the advantage given by the proposed discretization over the existing ones.

=> We believe that 1) one of the advantages of the discretization in this study is a computational efficiency in petascale computers with many cores (i.e., scalability) and 2) it was described in the manuscript already. Dennis et al. (2012) show better performance results from the spectral element method than that from the finite volume method. Although the scalability of an extended global version of this 2DNH should be examined in the future, we can expect the high efficiency as Dennis et al. (2012). In the manuscript, we mentioned our ultimate objective that is to build a 3D global NH model. In a global modeling, a higher scalability is more important.

Actually it is hard to compare an accuracy gain by comparison between our model's solution and WRF's. In case of WRF, the density current simulation can be reproduced properly only by using spatial diffusion in physical (x,z) space, not by using diffusion along coordinate surfaces (see Auxiliary FIG. 1). While our model uses diffusion along coordinate surfaces. Also our model can increase accuracy by increasing the order of the basis function. Therefore it would be better not to touch the comparison to WRF.

By using spatial diffusion along coordinate surfaces	By using spatial diffusion in physical (x,z) space
	

Auxiliary FIG. 1. WRF simulations of the density current test with $\Delta\overline{x} = 100$ m, $\Delta\overline{z} = 64$ m, and $\Delta t = 0.3$ s at 900 seconds; (right) the simulation using diffusion along coordinate surfaces (left) the simulation using diffusion in physical (x,z) space.

4) The employed time steps are not reported in the description of the numerical tests. For the sake of reproducibility the employed time step sizes / Courant numbers should be reported for each simulation.

=> We will add the time step sizes for all test cases in the manuscript. These are as follows:

Experiment	Resolution (m)	Time step size (sec)
Linear hydrostatic mountain wave	(5th order basis function) $\Delta\overline{x} = 2000$ and $\Delta\overline{z} = 375$	$\Delta t = 20$
Schär mountain gravity wave	(5th order basis function) $\Delta\overline{x} = 300$ and $\Delta\overline{z} = 250$	$\Delta t = 3$
2-D density current	(5th and 8th order basis function) $\Delta\overline{x} = 400$ and $\Delta\overline{z} = 64$ $\Delta\overline{x} = 200$ and $\Delta\overline{z} = 64$ $\Delta\overline{x} = 100$ and $\Delta\overline{z} = 64$ $\Delta\overline{x} = 50$ and $\Delta\overline{z} = 64$	$\Delta t = 0.3$
Inertia-gravity wave	(8th order basis function) $\Delta\overline{x} = 1250$ and $\Delta\overline{z} = 250$ $\Delta\overline{x} = 500$ and $\Delta\overline{z} = 250$ $\Delta\overline{x} = 250$ and $\Delta\overline{z} = 250$ $\Delta\overline{x} = 125$ and $\Delta\overline{z} = 250$	$\Delta t = 1$

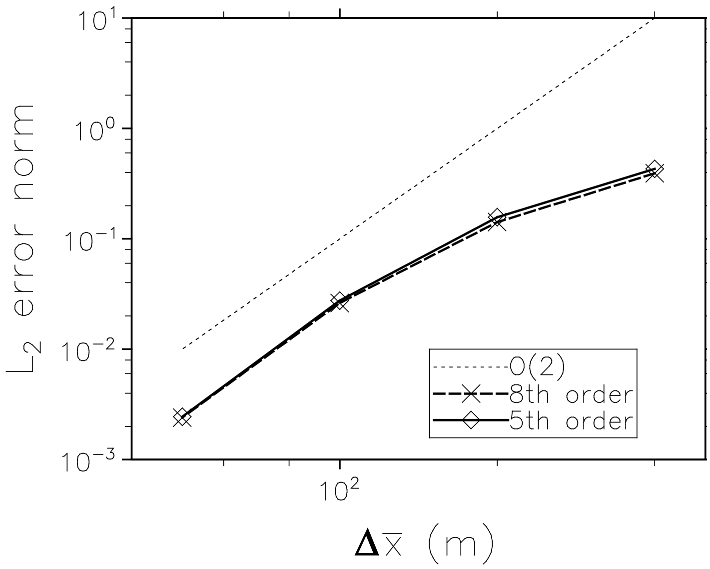
Rising thermal bubble	(5th order basis function) $\Delta\bar{x} = 20$ and $\Delta\bar{z} = 20$	$\Delta t = 0.2$
	$\Delta\bar{x} = 10$ and $\Delta\bar{z} = 10$	$\Delta t = 0.1$
	$\Delta\bar{x} = 5$ and $\Delta\bar{z} = 5$	$\Delta t = 0.05$

5) The model is effectively shown to compare well with published solutions. It would be interesting to assess the model behaviour in a convergence test. For instance, in the case of the density current, a self-convergence test like the one documented in Figure 6.7 of Restelli and Giraldo (2009) could be performed.

=> It's a good suggestion. The self-convergence test is carried out. For this test, a reference solution at spatial resolution 25 m and $\Delta t = 0.1$ s is used. The model solutions of the four spatial resolutions 400m, 200m, 100m, and 50m are obtained by fixing the time step $\Delta t = 0.3$ s. Because our model uses Gauss-Lobatto-Legendre (GLL) points and a pressure-based vertical coordinate, the all model solutions are interpolated to the equi-distance grid of $\Delta x = 400$ and $\Delta z = 50$, and then is used to evaluate errors. Here, we evaluate the error by using the relative L_2 error defined by

$$\left\| q_{simulation} \right\|_{L_2} = \sqrt{\frac{\int_\Omega \left(q_{ref} - q_{simulation} \right)^2 d\Omega}{\int_\Omega q_{ref}^2 d\Omega}},$$

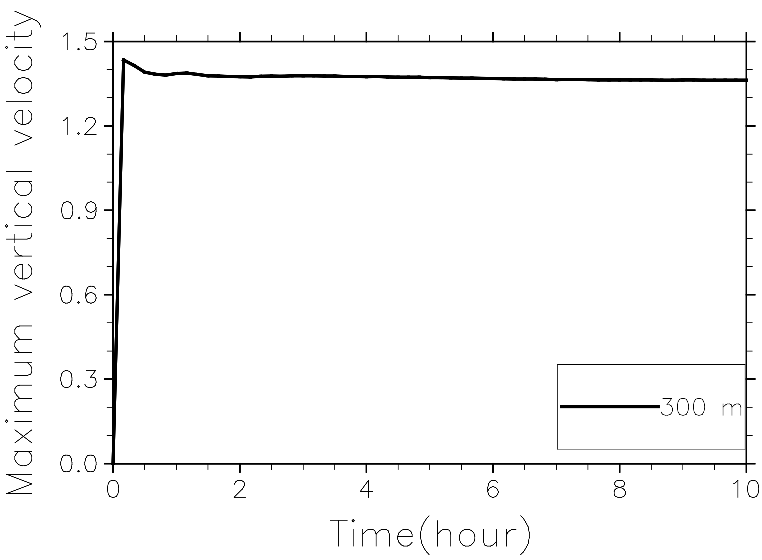
where $q_{simulation}$ and q_{ref} represent the model solution and reference solution, respectively. The resulting L_2 norm of the error in the potential temperature perturbation θ' is plotted in Auxiliary Fig.2. It is noted that at the highest resolution of 50m, the experimental convergence rate reaches the theoretical convergence rate 2. Also it is depicted that the error of the solutions of 8th order basis function is slightly smaller than that of 5th order basis function.



Auxiliary FIG. 2. Self-convergence test for the density current test; Relative L_2 error norms of the potential temperature perturbation θ' as functions of the space resolution $\Delta\bar{X}$ are shown. The reference solutions for these computations were made with $\Delta\bar{X} = 25$ m, $\Delta\bar{Z} = 64$ m, and $\Delta t = 0.1$s. The dotted line represents second-order convergence.

6) As already noted in a previous comment in the discussion, it would be interesting to evaluate the maximum vertical velocities generated by the model in a long-time simulation of a resting atmosphere above orography.

=> Following Edior's suggestion, we conducted the Schär mountain gravity wave. The model was integrated with a grid resolution of $\Delta\bar{X} = 300$ m using 5th order basis polynomials per element and $\Delta\bar{Z} = 250$ m using 80 levels for 10-hour without any diffusion or viscosity. The time step is 3 s. A more detail configuration of the model can be referred to the previous reply. Auxiliary Fig.3 shows the time evolution of the maximum vertical velocities for 10 hours. We observe that the maximum vertical velocity reaches a state of equilibrium after 1hour.



Auxiliary FIG. 3. Time evolution of the maximum vertical velocities (m/s) for 10 hours in the Schär Mountain test case.

7) The language needs to be revised as sometimes the phrasing is unclear and the reading flow is lost. This is especially true for, but not limited to, the Test cases Section 4, and subsections 4.3 and 4.4 in particular.

=> The revised manuscript will be edited by professional English editor.

Technical corrections:

1) Page 3718, line 9: what does "quadrature" refer to in a finite difference context?

=> We will change it to "vertical integration". The vertical integration is based on the centered finite difference that is $\int q d\eta \approx \sum_{k} q_{k+1/2}\left(\eta_{k+1}-\eta_{k}\right)$.

2) Page 3718, lines 9-12: "The Euler equation ... in this model". The reader is left *guessing which kind of vertical coordinate is used in which model. Please rephrase in a clearer way.*

=> We will revise accordingly.

Before	After
The Euler equations used here are in a flux form based on the hydrostatic pressure vertical coordinate, which are the same as those used in the Weather Research and Forecasting (WRF) model, but a hybrid sigma-pressure vertical coordinate is implemented in this model.	The Euler equations used here are in a flux form based on the hybrid sigma hydrostatic pressure vertical coordinate. There equations are similar to those used in the Weather Research and Forecasting (WRF) model which are based on the sigma hydrostatic pressure vertical coordinate.

3) Page 3718, line 12: "verified"-> validated

=> We will change it.

4) Page 3718, line 26: Graldo -> Giraldo.

=> We will change it.

5) Page 3719, line 18: " an attractive alternatively" -> alternative

=> We will revise accordingly.

6) Page 3719, line 22: "conservative flux-form finite-difference method" -> is it a finite volume or finite difference method? Please clarify.

=> Here, we will change it simply to "finite-difference method".

7) Page 3720, line 18-20: the sentence is not clear, what does "in which" refer to?

=> We will revise accordingly.

Before	After
In the next section we describe the governing equations with a definition of the prognostic and diagnostic variables used in our model, in which we present essential changes from SK08.	In the next section we describe the governing equations with a definition of the prognostic and diagnostic variables used in our model. In this section, we focus on essential changes from SK08.

8) Page 3720, line 26: "reported by PK13", is the coordinate introduced in PK13? Otherwise please include a reference to the work where the hybrid coordinate is first used.

=> "reported by" will be changed to "introduced in".

9) Page 3721, line 16: it is not immediately obvious that equation (4) is in flux form, given that μ_d is variable. Moreover, in the last term of the first line of equation (4) it is not clear whether ∇_η is the gradient only of ϕ or includes the bracket as well.

=> The original continuous equations are in a flux (conservative) form. In order to reduce truncation errors in the horizontal pressure gradient calculations in the discrete solver, we recast the equations using perturbation variables, which results in equation (4) in the paper. A more detailed description can be found in Skamarock and Klemp (2008). We will revise the description to make it clear.

And $\nabla_\eta \phi \left(\frac{\partial p'}{\partial \eta} - \mu'_d \right)$ will be changed to $\left(\frac{\partial p'}{\partial \eta} - \mu'_d \right) \nabla_\eta \phi$.

10) Page 3722, line 7: are the overbars needed above z as well?

=> Yes, the overbars are needed. It means the reference height in an atmosphere at rest.

11) Page 3722, lines 9-16. The sentence is too long and includes two formulas. Please rephrase.

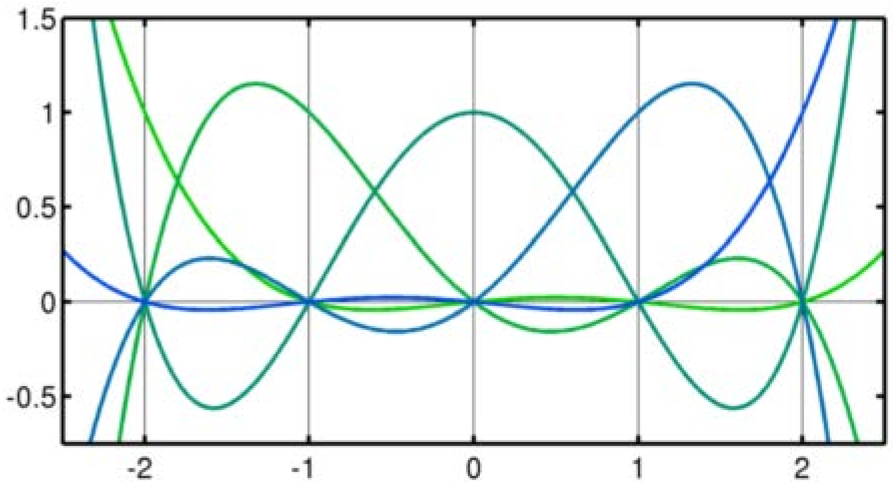
=> We will revise accordingly.

12) Page 3723, line 11: "(X-Z) slice framework", is there a reason why x and z are capitalized here?

=> No, there is not. We will revise accordingly.

13) Page 3723, line 19: the text in the bracket is somehow confusing. Surely the basis functions cannot be constant?

=>The basis functions are constant. In terms of a continuous function, the basis functions oscillate between nodal points. For better understanding, I captured the following figure from the web page (http://www.cb.uu.se/~cris/blog/index.php/archives/113) .



Although the 5 equidistance nodal points are used in the above figure, it shows well the property of the basis functions we used. Each function has zoro at all nodal points except a specific point.

14) Page 3724, line 10, "basis function": please refer to ψ_k *here as well.*

=> We will revise accordingly.

15) Page 3724, line 16, "The right-hand sides is evaluated" -> are evaluated.

=> Thank you for your correction.

16) Page 3725, line 13: the index k is used in Section 3.1.1 for the formulation of the horizontal discretization and in Section 3.1.2. in the vertical. The authors may consider using a different index to improve clarity.

=> We changed the "index k" in Section 3.1.1 to "index i".

17) Page 3725, lines 16 and 17: are the brackets encompassing the derivative terms needed in the inline formulas?

=> We included the following formula to describe the derivative terms.

$\left.\dfrac{\partial g}{\partial \eta}\right|_k = \dfrac{g_{k+1/2} - g_{k-1/2}}{\Delta \eta}$, where g corresponds to the variable such as p' and φ.

18) Page 3725, lines 22-24: which kind of quadrature rule is actually used in the vertical?

=> The vertical quadrature (vertical integration) is based on the centered finite difference that is $\int q d\eta \approx \sum_{k} q_{k+1/2}\left(\eta_{k+1}-\eta_{k}\right)$.

19) Page 3726, lines 1-8. The sentence is too long and hard to follow.

=> We will revise accordingly.

Before	After
For integrating the equations, we use the time-split RK3 integration technique following the strategy of SK08, in which low-frequency modes due to advective forcings are explicitly advanced using a large time step of the RK3 scheme, but high-frequency modes are integrated over smaller time steps using an explicit forward-backward time integration scheme for the horizontally propagating acoustic/gravity waves and a fully implicit scheme for vertically propagating acoustic waves and buoyancy oscillations (Klemp et al. 2007) .	For integrating the equations, we use the time-split RK3 integration technique following the strategy of SK08. In the time-split RK3 integration, low-frequency modes due to advective forcings are explicitly advanced using a large time step of the RK3 scheme, but high-frequency modes are integrated over smaller time steps. Among the high-frequency modes, the horizontally propagating acoustic/gravity waves are advanced using an explicit forward-backward time integration scheme, and vertically propagating acoustic waves and buoyancy oscillations are advanced using a fully implicit scheme (Klemp et al. 2007).

20) Page 3727, line 22: "center of the profile". You can add "xc" afterwards to define it.

=> We will revise accordingly.

21) Page 3728, line 14: “The extrema ... is” -> are.

=> We will revise accordingly.

22) Section 4.1. Please report the information about the use of 5th order polynomials for this test case as detailed in the caption of Fig. 2 in the text as well. Moreover, it would be helpful to report at which resolution the referred studies are running this test case.

=> We will revise the sentence as follows:

“The model is integrated with a grid resolution of $\Delta\bar{x} = 2$ km using 5th order basis polynomials per element and $\Delta\bar{z} = 375$ m for a nondimensional time of $\frac{\bar{u}t}{a} = 60$ which corresponds to 8.33 hours without diffusion or viscosity. ”

We also add the description of resolutions used in Durran and Klemp (1983) and Giraldo and Restelli (2008). Durran and Klemp (1983) uses $\Delta x = 2$ km and $\Delta z = 200$ m, and Giraldo and Restelli (2008) uses $\Delta\bar{x} = 1.2$ km and $\Delta\bar{z} = 240$ m using 10th order basis polynomials.

23) Section 4.2. In the original study of Straka et al. (1993), the 15 K perturbation is actually on the temperature, not on the potential temperature, see also Müller et al. (2013) for corroboration. This results in an initial potential temperature perturbation of -16.63 K in the center of the cold bubble (see the caption of Figure 1 page 4 of Straka et al. (1993)).

=> Yes, you are right. In this study, however, 15 K potential temperature perturbation is adopted similar to Giraldo and Restelli (2008) and Li et al. (2013) in order to compare our model’s solution to their solutions.

We will add the following sentence for clear description.

"Note that in this study, the potential temperature perturbation of $\theta_c = -15$ K is adopted similar to Giraldo and Restelli (2008) and Li et al. (2013) for comparison, which corresponds to -16.63 K in the center of the cold bubble. Straka et al. (1993) originally use -15 K temperature perturbation."

24) Page 3729, line 7: how is the viscosity term discretized? I appreciate the authors have replied to another comment about the bubble section regarding the discretization of the diffusion term. If the same discretization for the diffusion term is used both in Section 4.2 and in Section 4.4, it might be a good idea to anticipate the description to the first time it is mentioned, i.e. in Section 4.2.

=> We explained about the viscosity in the previous reply. We will revise the manuscript accordingly.

25) Page 3730, lines 8-13: the sentence is long and hard to read, "of which" at line 12 appears to refer to the Table but should be clarified. Same for "relieved" at lines 15-17.

=> We will revise accordingly.

Before	After
In addition to the profiles, the front location (-1K of potential temperature perturbation at the surface), and the extrema of the pressure perturbation and potential temperature perturbation agree well with each other (Table 1), of which the numbers are comparable to those of GR08.	In addition to the profiles, the front location (-1K of potential temperature perturbation at the surface), and the extrema of the pressure perturbation and potential temperature perturbation agree well with each other (Table 1). The numbers in Table 1 are comparable to those of GR08.

Before	After
That of 8th order polynomials, however, tends to be relieved from the deviation from the converged solution (Fig. 6c).	That of 8th order polynomials, however, tends to be closer to the converged solution (Fig. 6c).

26) Page 3730, line 23: the phrasing "the perturbation diverges" is prone to misunderstanding. Please reformulate

=> We will revise accordingly.

Before	After
This initial perturbation diverges to the left and right symmetrically	This initial potential temperature perturbation θ' radiates to the left and right symmetrically

27) Equation (25): shouldn't the bracket in the denominator be squared as in Skamarock and Klemp (1994) eq. (16)?

=> Thank you. We do not realize the typo.

Before	After
$\theta'(x,z) = \theta_c \dfrac{\sin\left(\dfrac{\pi z}{z_c}\right)}{1+\left(\dfrac{x-x_c}{a_c}\right)}$	$\theta'(x,z) = \theta_c \dfrac{\sin\left(\dfrac{\pi z}{z_c}\right)}{1+\left(\dfrac{x-x_c}{a_c}\right)^2}$

28) Page 3732, lines 1-5: The first sentence of the page is hard to follow, please rephrase.

=> We will revise accordingly.

Before	After
It is noted that all experiments give almost the same values for potential temperature perturbation where these values in the range $\theta' \in \left[-1.52\times10^{-3}, 2.83\times10^{-3}\right]$ are comparable to other studies (e.g., GR08 and Li et al. 2013). GR08 give the ranges of $\theta' \in \left[-1.51\times10^{-3}, 2.78\times10^{-3}\right]$ from the model based on the spectral element and discontinuous Galerkin method. Also Li et al. (2013) show $\theta' \in \left[-1.53\times10^{-3}, 2.80\times10^{-3}\right]$ using the high-order conservative finite volume model which are similar to our results.	It is noted that all experiments give almost the same values for potential temperature perturbation which is in the range $\theta' \in \left[-1.52\times10^{-3}, 2.83\times10^{-3}\right]$. These values are comparable to other studies. For example, GR08 give the ranges of $\theta' \in \left[-1.51\times10^{-3}, 2.78\times10^{-3}\right]$ from the model based on the spectral element and discontinuous Galerkin method. And Li et al. (2013) show $\theta' \in \left[-1.53\times10^{-3}, 2.80\times10^{-3}\right]$ using the high-order conservative finite volume model.

29) Page 3732, line 26 to page 3733, line 5. To facilitate readability it would be a good idea to split the long sentence into two sentences.

=> We will revise accordingly.

Before	After
It should be noted that an explicit second-order diffusion on coordinate surfaces is used with a viscosity coefficient of $\nu = 1$ m^{2}s^{-1}	It should be noted that explicit second-order diffusion on coordinate surfaces was used with a viscosity coefficient of $\nu = 1$ m^{2}s^{-1}

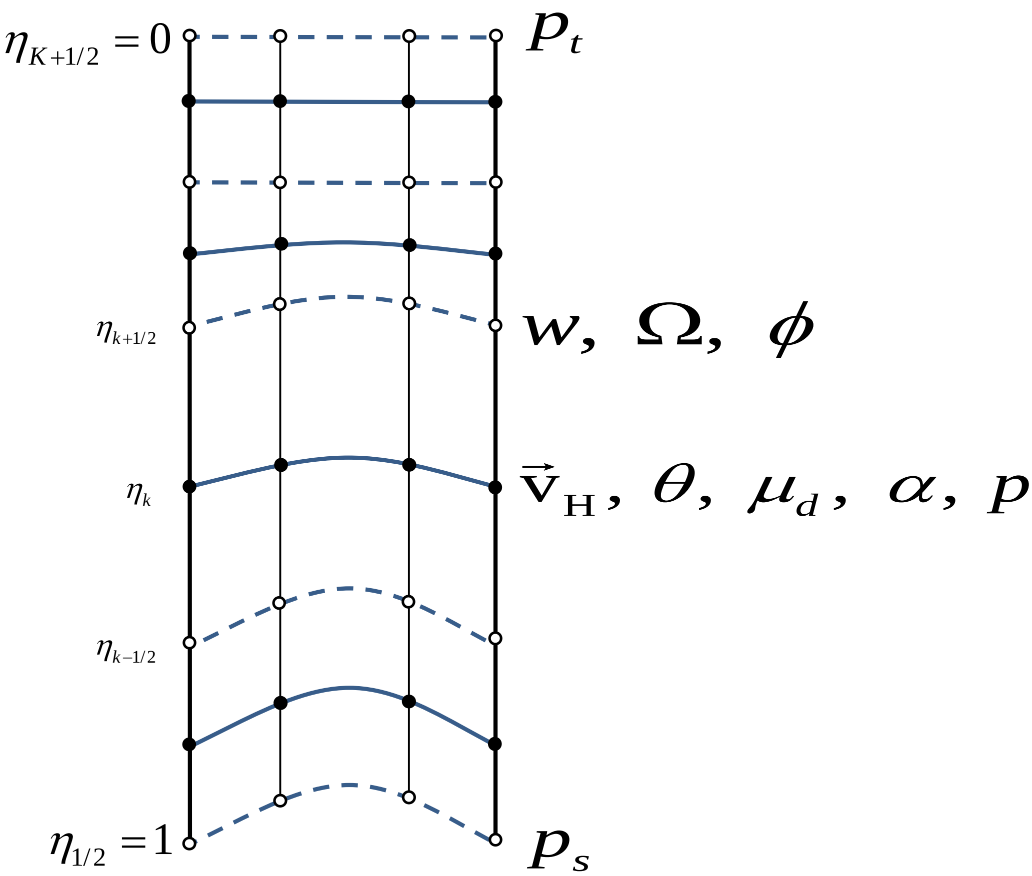
for all simulations of this test. The numerical diffusion is applied for momentum and potential temperature along the horizontal and vertical directions so that it eliminates the erroneous oscillations at the small scale – while this amount of diffusion might seem excessive, it has been chosen because it allows the model to remain stable even after the bubble hits the top boundary.	for all simulations of this test. The numerical diffusion was applied for momentum and potential temperature fields in horizontal and vertical directions to eliminate erroneous oscillations at the small scale. While this amount of diffusion might seem excessive, it has been chosen because it allows the model to remain stable even after the bubble hits the top boundary.

30) Page 3732, lines 9 and 12. What do you mean by "perfectly simmetric" and "concaving contours"?

=> "perfectly symmetric distribution" should be changed to "perfectly symmetrical distribution". "concaving contours" can be changed to "concave lines".

31) Figure 1: is there a reason why ps is indicated but pt is not?.

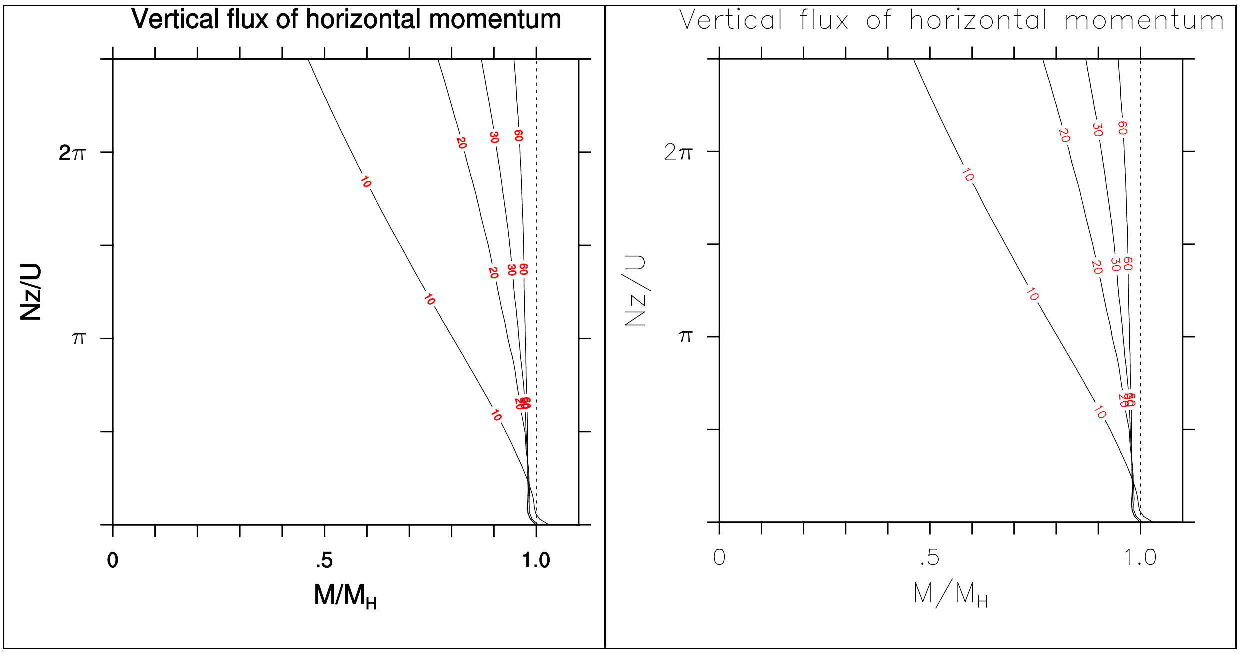
=> We will revise the figures accordingly, for example as shown below.



32) Please beware that in a printed version of the article some of the text in the figures appears so bold that it becomes unreadable, notably in the axis labels in Figures 2 to 7, 9 and in the contour information in figures 2, 3 and 9.

=> We will revise the figures accordingly, for example as shown below.

Before	After



33) It would be a good idea to give the contour interval values in the captions of Figures 4, 5 and 7.

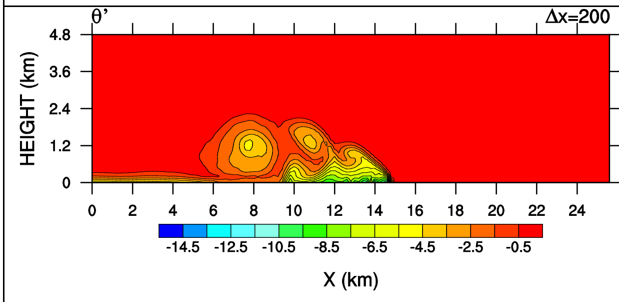
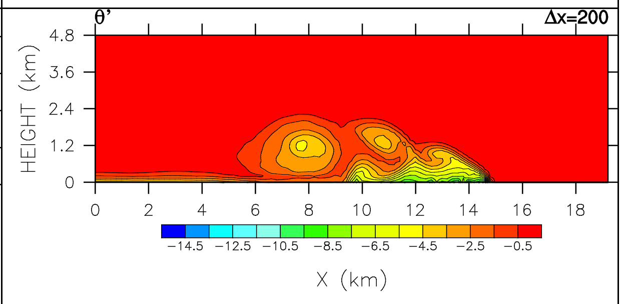
=> We will revise accordingly, for example as shown below.

"FIG. 4. Potential temperature perturbation after 900 s using (a) $\Delta \overline{X} = 400$ m, (b) $\Delta \overline{X} = 200$ m, (c) $\Delta \overline{X} = 100$ m, and (d) $\Delta \overline{X} = 50$ m grid spacing with 5th order basis polynomials per element for the density current. All simulations use $\Delta \overline{Z} = 64$ m grid spacing. The contour values are from -14.5 to -0.5 with an interval of -1.0."

"FIG. 7. Potential temperature perturbation at the (left) initial time and (right) time 3000s for $\Delta \overline{X} = 250$ m using 8th order basis polynomials per element and $\Delta \overline{Z} = 250$ m for the inertia-gravity wave. The contour values are from 0 (-0.0015) to 0.009 (0.0025) with an interval of 0.001 (0.0005) for the initial time (time 3000s)."

34) Figures 4 and 5: in most references reporting the right branch of the density current (e.g., Figure 4 of Skamarock and Klemp (2008) and Figure 7 of Giraldo and Restelli (2008)), the range in the x axis is limited to x = 19200 m.

=> We will revise the figures accordingly, for example as shown below.

Before	After
	

References

Dennis, J. M., J. Edwards, K. J. Evans, O. Guba, P. H. Lauritzen, A. A. Mirin, A. St-Cyr, M. A. Taylor, and P. H. Worley, 2012: CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model, *International Journal of High Performance Computing Applications, 26(1) 74-89, DOI: 10.1177/1094342011428142*

Skamarock, W. C. and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. Journal of Computational Physics, 227, 3465-3485.