The author is very grateful to the reviewer for his/her review which is really in-depth. The author was very encouraged to know that expert in the field finds the approach advocated in the paper worthwhile. The review is so helpful and goes so far beyond what usual reviews do that, in the case the paper is accepted for publication, I would be happy to acknowledge contribution by the reviewer by name, if he/she is willing to open it.

Practically all Editorial Considerations were accepted. Three exemptions are comments 18, 20, 29. Comment 3: Eq. (3) seems to be correct; comment 20: the arrows denoting vectors look OK in my computer; comment (29): line 217 seems to be OK.

As Editorial comment 28 suggested, two new subsections: 3.1 and 3.2 were introduced.

Below follow detailed answers to the Technical Considerations put forward by the Reviewer.

1. The author agrees that more consistent "dot" notation for the scalar product is preferable. Corresponding modifications will be made throughout the paper, in particular in Eqs. (1,7,8,9,16,17,20,23) and in a few unnumbered equations. The typo in the continuity equation will be fixed.
2. The rotation term in Eq. (7) is a scalar. It becomes obvious when dot is added between vector $\vec{\Omega}$ and the following parenthesis.
3. The reason for absence of the factor $\sum C_{ij}$ in the boundary term is that the divergency term is calculated in the $\vec{r}$ - and not $\vec{a}$ -coordinates:

$$-\int \sum_{i,j=1}^{3} C_{ij} \frac{\partial}{\partial a_j}\left(P\delta\xi_i\right)d\vec{a} = -\int \sum_{i=1}^{3} \frac{\partial}{\partial r_i}\left(P\delta\xi_i\right)d\vec{r} = -\int \left(\vec{N}\cdot d\vec{\xi}\right)Pd\Sigma$$

since from the expression for C_{ij} in line 98 one has

$$\sum_{j=1}^{3} d\vec{a}C_{ij}\frac{\partial}{\partial a_j} = \sum_{j=1}^{3} d\vec{a}\left|\frac{\partial\vec{r}}{\partial\vec{a}}\right|\frac{\partial a_j}{\partial r_i}\frac{\partial}{\partial a_j} = d\vec{r}\frac{\partial}{\partial r_i}$$

4. Dots will be added to indicate corresponding inner products what makes the LHS of (9) manifestly scalar.
5. Integration with respect to time of the term written down by the Reviewer in his/her comment gives:

$$\int d\vec{a}\int_{t_0}^{t_1} dt\,\partial_t\left[\delta\vec{\xi}\cdot\left(\partial_t\vec{\xi}+\vec{\xi}\times\vec{\Omega}\right)\right] = \int d\vec{a}\left[\delta\vec{\xi}\cdot\left(\partial_t\vec{\xi}+\vec{\xi}\times\vec{\Omega}\right)\right]_{t_0}^{t_1}$$

Variations $\delta\vec{\xi}$ at $t=t_0$ and $t=t_1$ are assumed zero throughout the medium (i.e., for all $\vec{a}$) what is possible since they are arbitrary. As a result, correct equations of motion follow.

6. A hydrostatic state is an exact solution of Eqs. (6), (25) because in this state the potential energy has a minimum; i.e.:

$$\frac{\delta}{\delta\vec{\xi}}\left[\rho_0(\vec{a})\left(E\left(\alpha,s_0(\vec{a})\right)+g\vec{n}\cdot\vec{\xi}\right)\right]=\frac{\delta}{\delta\vec{\xi}}\rho_0(\vec{a})E\left(\alpha,s_0(\vec{a})\right)+\rho_0(\vec{a})g\vec{n}=0$$

For this reason, $\vec{\xi}_k=0,\ \vec{p}_k=0$ is a (hydrostatic) solution of Eq. (25) for all k (i.e., for all vertices). Thus Eq. (25) allows linearization around the hydrostatic state. No modifications of the Lagrangian or Hamiltonian are necessary.

It is worth mentioning here that in the 2D case considered in Sec. 4, to ensure that the hydrostatic equilibrium is a solution of governing equations, one has to introduce a "calibration" function $\Psi(s)$. For this reason, treatment of the dynamics in terms of Lagrangian coordinates (in a "hydrodynamic" sense of the word) seems to be preferable.

7. The author apologizes for the typo; the tilde sigh in the unnumbered equation following Eq. (16) should be absent. The following sentence will be added following Eq. (4) for clarity: "*The "star" index in L_* is introduced to distinguish the total Lagrangian L_* from its (spatial) density $L=dL_*/d\vec{a}$ where $d\vec{a}$ means an element of volume with respect to $\vec{a}$ -coordinates. The same pertains to the Hamiltonian in the equations below.*"
8. Calculation of the variation of the internal energy term with respect to $\vec{\xi}$ includes a non-trivial transition between $\vec{a}$ - and $\vec{r}$ - coordinates, what leads to extra factors in the term following the second equality sign in the Reviewer derivation. The pressure gradient term in the unnumbered equation below (16) follows from Eq. (8) (for points inside the volume the surface term doesn't contribute).
9. Two corresponding dots will be added.
10. The following sentence will be added after Eq. (19): "*The same equation holds also in 2D case with $1/4$ factor replaced by $1/3$ and summations proceeding from 1 to 3."*
11. Two corresponding dots will be added.
12. This point does need a clarification. The following sentence will be added after Eq. (20): *"Here and below vector symbols (arrows on top) indicate 3D vectors, and indices correspond to vertices. Dots between vectors denote scalar products with respect to 3D vector coordinates, correspondingly."*
13. Yes, this is correct. Index τ will be added to $\tilde{H}$ in Eq. (23): $\tilde{H}\to\tilde{H}_\tau$.
14. Yes, this is correct; each vertex in Eq. (23) appears only once and there is no restrictions on the number of tetrahedra sharing a vertex.
15. It seems so. However, not being an expert in modeling atmospheric dynamics, I avoided the term "Lagrangian remapping" not being sure if it assumes more than just setting initial conditions.

    My understanding is that the initial conditions in Eulerian and Lagrangian descriptions are in fact the same: at $t=0$ to each point in space a velocity vector corresponding to the fluid particle located at this point is assigned. Then either the shift vector of the same fluid particle as a function of time is chosen to represent the evolution, or the velocity vector at the same spatial point. Since we are using the Lagrangian description (the shift vector $\vec{\xi}$) within the time step, to perform a next time step we have to formulate the initial condition

at $t = \Delta t$. For the Eulerian description this is just the velocities obtained at the end of the previous timestep. For the Lagrangian description this is not that straightforward and requires figuring out first what particle has arrived to a given vertex at $t = \Delta t$, and then assigning to corresponding vertex the particle's velocity (or momentum). If this is in fact "Lagrangian remapping", I will use this term in the paper.

16. No, τ *s* in Eq. (26) are not barycentric coordinates (in the latter case there should be four of them.*)*
17. Since τ s are not barycentric coordinates, the condition $\sum_i \tau_i = 1$ doesn't hold.
18. Since (28) is a vector equation, we have in fact three equations for three scalar parameters $\left(\tau_1, \tau_2, \tau_3\right)$. The reviewer's statement is correct that solution of Eq. (29*) per se* doesn't determine from which tetrahedron the particle came: this is done due to condition (27). The most straightforward approach is to solve linear set (27) for all tetrahedra containing corresponding vertex, although different plausible "first guesses" could be suggested to save computation time (which is not too big for solution of the third order linear set anyway).
19. What is meant here that the shifts $\vec{\xi}$ should be small as compared to the tetrahedra sizes. Then the tetrahedra do not degenerate within a time step and the field of shifts $\vec{\xi}\left(\vec{a}\right)$ ensures one-to-one correspondence between the original and shifted positions of fluid particles. As a result, one and only one fluid particle arriving to each vertex can always be found, although it is generally not known in advance form which tetrahedron.
20. My (speculative) vision of the parallelization option is as follows. Each available core handles a set of connected tetrahedra. Data exchange between cores will include only tendencies for the "boundary vertices", i.e. vertices which belong to tetrahedra which are handled by different cores. Since the boundary vertices constitute generally a small share of all vertices, corresponding data fluxes will hopefully be not too intense. I would imagine one central core which receives tendencies for the boundary vertices from all cores and redirects these tendencies to corresponding counterpart cores (note, that due to splitting/recombination of tetrahedra the set of boundary vertices may also vary, and the central core has to track corresponding changes).
    Regarding computer realization of the code I would like to mention the following. Essentially all job will be done by a single function which handles a single tetrahedron. This function takes at the input $\left(\vec{\xi}, \vec{p}, \rho_0, s_0\right)$ parameters at the four vertices ($(3+3+2)\cdot 4 = 32$ parameters altogether) and at the output provides by calculating corresponding Hamiltonian (given by Eq. (32)) and its derivatives 24 tendencies $\left(\partial_t \vec{\xi}, \partial_t \vec{p}\right)$ at the same vertices. This function is run within a single loop over all tetrahedra. Accumulation of the (weighted by volumes) tendencies at every vertex provides total tendencies for all vertices $\left(\partial_t \vec{\xi}_k, \partial_t \vec{p}_k\right)$. These tendencies represent input for any suitable time-integration scheme.

21. The author agrees that comparison of the numerical results obtained in Sec. 4 vs. established algorithm would be highly desirable. Such undertaking is, however, too hard for a single researcher who doesn't have an access to the existing dynamical cores nor experience in running them. Thus, the only test of validity of the numerical solution in Sec.4 provides conservation of mass. On the other hand, the main purpose of the paper is not to prove correctness of the developed code but to demonstrate feasibility of a code based on the suggested approach including an option of splitting/recombining tetrahedra.
22. This seems to be the case for at least Reynold stresses. Another important factor is heating/cooling of the air which can be taken into when reassignment of the entropy at the vertices is done. This is all pretty much speculative at this stage, though. Still, if a value the approach advocated in this paper is demonstrated in the simplest conservative case, there seems to exist ways to accommodate dissipative processes.