INVITED MANUSCRIPT

A mixed finite-element, finite-volume, semi-implicit discretization for atmospheric dynamics: Cartesian geometry

Thomas Melvin[1] | Tommaso Benacchio[1†] | Ben Shipway[1] | Nigel Wood[1] | John Thuburn[2] | Colin Cotter[3]

[1]Dynamics Research, Met Office, Exeter, UK
[2]Mathematics Department, University of Exeter, UK
[3]Mathematics Department, Imperial College, London, UK

Correspondence
Thomas Melvin, Met Office, FitzRoy Road, Exeter EX1 3PB, UK.
Email: Thomas.Melvin@metoffice.gov.uk
†Present address: MOX —Modelling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Italy

Funding information
Natural Environment Research Council through grants NE/K006762/1 and NE/K006789/1,

To meet the challenges posed by future generations of massively parallel supercomputers, a reformulation of the dynamical core for the Met Office's weather and climate model is presented. This new dynamical core uses explicit finite-volume type discretizations for the transport of scalar fields coupled with an iterated-implicit, mixed finite-element discretization for all other terms. The target model aims to maintain the accuracy, stability and mimetic properties of the existing Met Office model independent of the chosen mesh while improving the conservation properties of the model. This paper details that proposed formulation and, as a first step towards complete testing, demonstrates its performance for a number of test cases in (the context of) a Cartesian domain. The new model is shown to produce similar results to both the existing semi-implicit semi-Lagrangian model used at the Met Office and other models in the literature on a range of bubble tests and orographically forced flows in two and three dimensions.

KEYWORDS
dynamical core, mimetic discretization, spatial discretization, temporal discretization

1 | INTRODUCTION

The dynamical core of a weather and climate prediction model is responsible for simulating those fluid dynamical aspects that are resolved on the mesh that is chosen for the spatial discretisation. The governing fluid dynamical equations are well known and presented in section 2. The principal properties required of a dynamical core are the accuracy, numerical stability and efficiency with which those equations are numerically approximated (e.g. Lauritzen *et al.*, 2011 gives a discussion of various such aspects). The current trend of supercomputer architectures is towards a greatly increasing number of processors, together with an increasingly complex hierarchy of heterogeneous processors and memories. In terms of the efficiency of a dynamical core, this trend shifts interest from optimizing the number of calculations towards optimizing the management of memory and communications between processors (Lawrence *et al.*, 2017). This has led to a renewed interest in the choice of mesh used and

in particular to a desire, for those still using it, to move away from a latitude–longitude mesh, (Staniforth and Thuburn, 2012). However, as noted by Staniforth and Thuburn (2012), the latitude–longitude mesh confers a number of advantages over many of the alternative meshes. The challenge then is to use spatial discretizations that retain the same accuracy and stability on the alternative meshes as obtained with latitude–longitude meshes.

The mixed finite-element approach of Cotter and Shipton (2012); Cotter and Thuburn (2014) and Thuburn and Cotter (2015) is attractive as it achieves numerical consistency without relying on the orthogonality of the mesh. Those authors focused on the shallow-water equations and developed a scheme that shares many of the beneficial properties of the C-grid finite-difference scheme, in particular good wave dispersion properties together with the necessary conditions to avoid spurious computational modes. This approach is very general in terms of the order of accuracy; arbitrarily high-order schemes can be straightforwardly defined. In line

with most current dynamical cores, here only second-order accuracy for the non-transport aspects (i.e. those aspects principally responsible for wave propagation) is sought. Therefore the lowest-order version of the mixed finite-element scheme is used here but extended to the three-dimensional Euler equations (section 4). This is achieved by extending the hierarchy of finite-element spaces to include the particular temperature space proposed by Guerra and Ullrich (2016); Natale *et al.* (2016) and Melvin *et al.* (2018). At lowest order, this space resembles a finite-difference Charney–Phillips staggering of temperature. In particular it gives a finite-element scheme that has good wave dispersion properties for vertically propagating waves, as well as having the necessary conditions to avoid spurious computational modes in three dimensions.

A critical component of a weather and climate prediction system is the numerical scheme used for the transport of scalar quantities, e.g. the semi-Lagrangian scheme (Staniforth and Côté, 1991) has proven to be a very effective scheme and is used by many operational weather and climate prediction centres. However, an important weakness of virtually all such schemes is that they do not conserve the quantity they transport. Here the aim is to retain the good properties of the semi-Lagrangian scheme, namely an upwind scheme with small dispersive errors and scale-selective damping and with flexibility in the order of the scheme, but to additionally provide a flux-form scheme, at least for the density field, and hence exact conservation. Since the lowest-order finite-element scheme is second-order, this means using a transport scheme that is separate from the rest of the dynamical core (as is the case generally for most models and in particular for semi-implicit semi-Lagrangian schemes). Whilst the transport scheme could be a finite-element scheme, here (following Thuburn and Cotter, 2015) a finite-volume scheme is used (presented in section 5) as that would seem a straightforward way of retaining the desired properties. Specifically a method of lines approach is used in which third-order upwind polynomial reconstructions are coupled with a third-order Runge–Kutta temporal discretization.

The temporal discretization scheme described here takes as its starting point the dynamical core described in Wood *et al.* (2014) (referred to there as the "standard SISL" version) which forms the basis of the weather and climate prediction system described by Williams *et al.* (2015) and Walters *et al.* (2017). In particular it is desired to retain the good temporal accuracy and long time-step stability of that model. So while this work describes the replacement of the spatial discretization and transport schemes of Wood *et al.* (2014), the present scheme uses a similar iterated-implicit temporal discretization, an overview of which is given in section 3 and a more detailed treatment in section 6. The current scheme differs from that of Wood *et al.* (2014) due to the presence, in the mixed finite-element method, of non-diagonal mass matrices, which complicates the solution of the implicit system of equations.

Although the development of this scheme is motivated by its application to non-orthogonal global meshes on the sphere, it is important that it has good accuracy on Cartesian meshes; without this, its performance on the global meshes is unlikely to be acceptably good and it certainly will not be acceptable as the basis for regional modelling. Therefore, example computational results from a variety of essentially two-dimensional and also some three-dimensional test cases are reported in section 7 before a concluding discussion in section 8.

2 | CONTINUOUS EQUATIONS

The Euler equations for a perfect gas in a rotating frame are

$$\frac{\partial \mathbf{u}}{\partial t} = -\boldsymbol{\xi} \times \mathbf{u} - 2\boldsymbol{\Omega} \times \mathbf{u} - \nabla (K + \Phi) - c_p \theta \nabla \Pi, \quad (1)$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u}), \quad (2)$$

$$\frac{\partial \theta}{\partial t} = -\mathbf{u} \cdot \nabla \theta, \quad (3)$$

together with the nonlinear equation of state

$$\Pi^{\left(\frac{1-\kappa}{\kappa}\right)} = \frac{R}{p_0} \rho \theta, \quad (4)$$

where $\mathbf{u}$ is the velocity vector, $\boldsymbol{\xi} \equiv \nabla \times \mathbf{u}$ is the relative vorticity,
$\boldsymbol{\Omega}$ is the rotation vector, $K \equiv \mathbf{u} \cdot \mathbf{u}/2$ is the kinetic energy per unit mass,
Φ is the geopotential such that $\nabla \Phi = -\boldsymbol{g}$ where $\boldsymbol{g}$ is the acceleration due to gravity, c_p is the specific heat at constant pressure,
θ is potential temperature related to temperature through $T = \theta \Pi$,
$\Pi = (p/p_0)^\kappa$ is the Exner pressure with p pressure and p_0 a constant reference pressure, R is the gas constant per unit mass, $\kappa \equiv R/c_p$ and ρ is density.

These equations are solved subject to the boundary condition of zero mass flux through the boundaries of the domain.

Note that, following Cotter and Shipton (2012); Cotter and Thuburn (2014) and Thuburn and Cotter (2015), the velocity equation is written in the vector-invariant form. In the shallow-water form of the equations, this allows the discretized version of the equations to retain some of the mimetic properties discussed by Staniforth and Thuburn (2012). The scheme presented here targets most of the desirable criteria outlined in Staniforth and Thuburn (2012), but notably it does not target conservation of energy or axial angular momentum.

3 | OVERVIEW OF THE SPATIO-TEMPORAL DISCRETIZATION

The temporal discretization of the equations is inspired by the iterative-semi-implicit semi-Lagrangian discretization such as that used in Wood *et al.* (2014). In that scheme all advective terms are handled using a semi-Lagrangian scheme. In

the implementation of Wood *et al.* (2014), the advected quantities are the start of time-step fields whilst, at convergence of the iterative scheme, the advecting velocity is averaged in time. All other terms are handled using an iterative-implicit temporal discretization.

The same basic discretization is targeted here but with the following differences:

1. To achieve good conservation properties, instead of a semi-Lagrangian scheme, an explicit Eulerian flux-form scheme is used for the continuity equation.
2. A similar scheme is also used for potential temperature but in advective form (this is to achieve good wave dispersion properties by avoiding the need to average the vertical wind).
3. The self-advection terms in the velocity equation (which in the vector invariant form of the equation are manifested in the kinetic energy ∇K and vorticity $\xi \times \mathbf{u}$ terms of Equation 1) are averaged in time and evaluated iterative-implicitly.
4. A mixed finite-element spatial discretisation is used in place of the finite-difference one used in Wood *et al.* (2014).

First consider the velocity Equation 1. This can be written as

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{S}, \tag{5}$$

where

$$\mathbf{S} \equiv -\xi \times \mathbf{u} - 2\mathbf{\Omega} \times \mathbf{u} - \nabla\left(K + \Phi\right) - c_p \theta \nabla \Pi. \tag{6}$$

This is first integrated in time to give

$$\frac{\mathbf{u}\left(\mathbf{x}, t + \Delta t\right) - \mathbf{u}\left(\mathbf{x}, t\right)}{\Delta t} = \frac{1}{\Delta t}\int_t^{t+\Delta t} \mathbf{S}\, \mathrm{d}t, \tag{7}$$

where $\mathbf{x}$ is a fixed position. The right-hand side time integral is then approximated using a, possibly off-centred, trapezoidal rule to give

$$\delta_t \mathbf{u} = \overline{\mathbf{S}}^{\alpha}, \tag{8}$$

where, for a generic scalar or vector variable F,

$$\delta_t F \equiv \frac{F^{n+1} - F^n}{\Delta t}, \tag{9}$$

and

$$\overline{F}^{\alpha} \equiv \alpha F^{n+1} + (1 - \alpha) F^n. \tag{10}$$

The parameter α is an off-centring parameter which takes the value $1/2$ for a centred scheme. The superscripts $n+1$ and n denote the time-level of a variable.

Using the same notation, the density Equation 2 and potential temperature Equation 3 are discretized as

$$\delta_t \rho = -\nabla \cdot \left[\mathcal{F}\left(\rho^n, \overline{\mathbf{u}}^{1/2}\right)\right], \tag{11}$$

$$\delta_t \theta = -\mathcal{A}\left(\theta^n, \overline{\mathbf{u}}^{1/2}\right), \tag{12}$$

where $\overline{\mathbf{u}}^{1/2}$ indicates that the advecting velocity is a centred average in time, $\mathcal{F}\left(\rho^n, \overline{\mathbf{u}}^{1/2}\right)$ is the time-averaged flux of density and $\mathcal{A}\left(\theta^n, \overline{\mathbf{u}}^{1/2}\right)$ is

the time-averaged advection tendency of the potential temperature.

All terms are discretized in space using the mixed finite-element scheme described in section 4, except for $\mathcal{F}$ and $\mathcal{A}$ which are discretized using the finite-volume scheme described in section 5.

Equations 8–12 and 4 represent a set of coupled, nonlinear equations. These are solved using a quasi-Newton method that is detailed in section 6.

4 | THE MIXED FINITE-ELEMENT DISCRETIZATION

4.1 | Subdivision of domain

The three-dimensional model domain (D, with boundary ∂D) is partitioned into a mesh consisting of a number of cells (C), each cell having a number of faces (F), edges (E) and vertices (V). Where appropriate, subscripts D, C, F, E and V, respectively, will be used to denote evaluation of quantities over objects of these types. Here, in three dimensions, hexahedral cells are used that are aligned in columns in the vertical (with their lateral faces having normals that are perpendicular to gravity).

4.2 | Function spaces

To form the finite-element function spaces in three dimensions, the sequence of finite-element spaces of Natale *et al.* (2016) is used. This sequence is the natural extension of the two-dimensional function spaces used for the shallow-water equations presented in Cotter and Shipton (2012); Thuburn and Cotter (2015).

There are four principal function spaces, denoted by $\mathbb{W}_i$, $i = 0, 1, 2, 3$, each of which has specific attributes and, in particular, varying degrees of continuity across cell boundaries. These spaces are related by the de Rham complex (Bott and Raoul, 1982):

$$\mathbb{W}_0 \xrightarrow{\nabla} \mathbb{W}_1 \xrightarrow{\nabla \times} \mathbb{W}_2 \xrightarrow{\nabla \cdot} \mathbb{W}_3. \tag{13}$$

These function spaces at order l for hexahedral elements correspond to:

$\mathbb{W}_0$ The Q_{l+1} space of scalar functions;
$\mathbb{W}_1$ The Nédélec N_l space of vector functions;
$\mathbb{W}_2$ The Raviart–Thomas RT_l space of vector functions;
$\mathbb{W}_3$ The Q_l^{DG} space of scalar functions.

Further details on these spaces can be found in Boffi *et al.* (2013). This de Rham complex is complemented by the additional function spaces:

$\mathbb{W}_\theta$ The space of scalar functions based on the vertical part of $\mathbb{W}_2$; as discussed in section 1, this is used to avoid vertical averaging in the coupling between the vertical momentum

and potential temperature equations and hence to obtain good numerical wave dispersion properties;

$\mathbb{W}_\chi$ The Q_m^{DG} space of scalar functions, where m may be different from l; this allows the representation of the coordinate field to be decoupled from the choice of the other finite-element spaces.

Details on the structure of these spaces can be found in Appendix A. As noted in the Introduction, the choice of lowest-order elements is made here, i.e. $l = 0$. Details of the form of the basis functions for this choice can be found in Appendix B.

4.3 | Variable expansions and weak forms

Each variable is assigned to a function space that is consistent with its physical nature. Specifically:

$\Phi \in \mathbb{W}_0$, pointwise scalar functions;
$\boldsymbol{\xi} \in \mathbb{W}_1$, vector functions corresponding to circulations;
$\mathbf{u} \in \mathbb{W}_2$, vector functions corresponding to fluxes; and
$\rho \in \mathbb{W}_3$, scalar functions corresponding to volume integrals.

Additionally, Π is placed in $\mathbb{W}_3$ and the form of the rotation vector $\boldsymbol{\Omega}$ is assumed to be known analytically and so will be computed where needed. In order to obtain good wave dispersion properties equivalent to the use of a Charney–Phillips grid and to avoid the computational mode of the Lorenz grid, $\theta \in \mathbb{W}_\theta$. Melvin *et al.* (2018) give details on this choice.

Each variable is expanded in the trial functions associated with its function space. Finally, the (time-dependent) coefficients of this expansion are chosen such that the projection of each equation onto test functions vanishes for all test functions from the appropriate function space, i.e. the error is orthogonalized to the test function space. This is the discrete weak form of the equation.

For example, for a prototypical discrete equation, $f = 0$, for some variable f, the weak form of this equation is given by multiplying the equation by a test function h and integrating over the domain

$$\int_{\mathrm{D}} hf \, dV = 0, \tag{14}$$

and then requiring that this equation holds for all test functions g in the appropriate function space. Equation 14 can be concisely written as

$$\langle h, f \rangle = 0. \tag{15}$$

The Galerkin method is followed in which the test functions are chosen from the same space as the trial functions. These functions are given in Appendix B. For each of $\mathbb{W}_0, \mathbb{W}_1, \mathbb{W}_2, \mathbb{W}_3$, these functions are usually denoted respectively by γ, $\mathbf{c}$, $\mathbf{v}$, and σ. The test function for $\mathbb{W}_\theta$ is denoted by w.

Note that the boundary condition,

$$\mathbf{n}_{\partial \mathrm{D}} \cdot \mathbf{u} = 0, \tag{16}$$

on the boundary of the domain, $\partial \mathrm{D}$, (where $\mathbf{n}_{\partial \mathrm{D}}$ is the outward normal to that boundary) is enforced in the expansion of

the velocity, $\mathbf{u}$, in the trial functions for $\mathbb{W}_2$. This boundary condition is only valid on the top and bottom of the domain, where it is appropriate to apply a no-flux boundary condition. For all the cases considered here (which use periodic boundary conditions in the horizontal), and more generally for spherical domains, no extra boundary conditions are needed. However, for bounded domains (such as used for local-area modelling), extra flux integral boundary conditions would arise when integrating by parts (as done in the next section).

4.4 | Weak form of the equations

4.4.1 | Velocity equation

Using Equation 6, multiplying Equation 8 by test functions from $\mathbb{W}_2$ and integrating over the domain D gives the weak form of the velocity equation as

$$\langle \mathbf{v}, \delta_t \mathbf{u} \rangle = -\overline{\langle \mathbf{v}, \boldsymbol{\xi} \times \mathbf{u} + 2\boldsymbol{\Omega} \times \mathbf{u} \rangle}^\alpha \\ -\overline{\langle \mathbf{v}, \nabla K + \nabla \Phi + c_p \theta \nabla \Pi \rangle}^\alpha. \tag{17}$$

However, various quantities in this equation do not have the required continuity to permit the required vector operations to be evaluated.

- Exner pressure, Π, is discontinuous between cells, so $\nabla \Pi$ is not defined at cell boundaries. Additionally, the potential temperature, θ, is only continuous between cells in the vertical direction. This prevents direct evaluation of Equation 17. Instead, the procedure is: first split the integration over the domain into the sum of integrations over cells; then integrate by parts over each cell, introducing boundary integrals over the cell faces; and then rewrite as global integrals, i.e.

$$\langle \mathbf{v}, c_p \theta \nabla \Pi \rangle = \sum_{\mathrm{C}} \langle c_p \theta \mathbf{v}, \nabla \Pi \rangle_{\mathrm{C}} \\ = \sum_{\mathrm{F}} \langle\!\langle [\![c_p \theta \mathbf{v}]\!]_{\mathrm{F}}, \{\Pi\}_{\mathrm{F}} \rangle\!\rangle_{\mathrm{F}} \\ - \sum_{\mathrm{C}} \langle \nabla_{\mathrm{C}} \cdot (c_p \theta \mathbf{v}), \Pi \rangle_{\mathrm{C}} \\ = \langle\!\langle [\![c_p \theta \mathbf{v}]\!], \{\Pi\} \rangle\!\rangle - \langle \nabla_{\mathrm{C}} \cdot (c_p \theta \mathbf{v}), \Pi \rangle, \tag{18}$$

where $\langle \cdot \rangle_{\mathrm{C}}$ indicates integration over a cell; $\nabla_{\mathrm{C}} \cdot$ indicates that the divergence is evaluated only within the interior of the cell, excluding the faces; $\langle\!\langle \cdot \rangle\!\rangle_{\mathrm{F}}$ indicates integration over a face; and $\langle\!\langle \cdot \rangle\!\rangle$ indicates the sum of integrations over all faces (because of the boundary condition (16) this sum is in fact only non-zero for interior faces).

The notation $[\![\cdot]\!]_{\mathrm{F}}$ indicates the jump in its argument across a face and $\{\cdot\}_{\mathrm{F}}$ indicates the value of its argument on a face; the choice made here is to use the average of the discontinuous values on the face. The appearance of $[\![\cdot]\!]_{\mathrm{F}}$ and $\{\cdot\}_{\mathrm{F}}$ without subscripts indicates the sum of integrations over all faces. Appendix C gives details of these operators.

- Since $\mathbf{u} \in \mathbb{W}_2$, $\boldsymbol{\xi} = \nabla \times \mathbf{u}$ is not defined at cell boundaries. Therefore the vorticity $\boldsymbol{\xi} \in \mathbb{W}_1$ is obtained as the weak
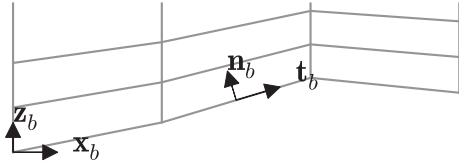
FIGURE 1 Unit vectors in terrain-following coordinates. $\mathbf{z}_b$ is parallel to gravity and $\mathbf{x}_b$ is normal to gravity. $\mathbf{n}_b$ is normal to the model layers and $\mathbf{t}_b$ is parallel to model layers

curl of $\mathbf{u}$ by requiring that

$$\langle \mathbf{c}, \boldsymbol{\xi} \rangle = \langle \nabla \times \mathbf{c}, \mathbf{u} \rangle - \langle\!\langle \mathbf{c}, \{\mathbf{u}\} \times \mathbf{n} \rangle\!\rangle, \qquad (19)$$

for all $\mathbf{c} \in \mathbb{W}_1$ where $\mathbf{n}$ is the outward normal to each face. Note that, due to the uniqueness of $\{\mathbf{u}\}$ on faces, the boundary term, $\langle\!\langle \cdot \rangle\!\rangle$, in Equation 19 is only non-zero on the top and bottom boundaries of the domain. At these points, consistently with an assumed free-slip condition on the velocity field, the value of $\{\mathbf{u}\}$ on the domain boundaries takes the value just inside the domain. This choice results in the horizontal components of the vorticity being zero on the top and bottom boundaries of the domain.

- The kinetic energy, K, does not in general have any continuity between cells and its gradient is not defined. Therefore, this term is integrated by parts to give

$$\langle \mathbf{v}, \nabla K \rangle = -\langle \nabla \cdot \mathbf{v}, K \rangle, \qquad (20)$$

where the boundary condition (16) has been used to eliminate the boundary integral.

Therefore, with the addition of a Rayleigh damping term, Equation 17 becomes

$$\langle \mathbf{v}, \delta_t \mathbf{u} \rangle = -\overline{\langle \mathbf{v}, \boldsymbol{\xi} \times \mathbf{u} \rangle}^\alpha - \overline{\langle \mathbf{v}, 2\boldsymbol{\Omega} \times \mathbf{u} \rangle}^\alpha + \overline{\langle \nabla \cdot \mathbf{v}, K \rangle}^\alpha - \overline{\langle \mathbf{v}, \nabla \Phi \rangle}^\alpha$$
$$- \overline{\langle\!\langle [\![c_p \theta \mathbf{v}]\!], \{\Pi\} \rangle\!\rangle}^\alpha + \overline{\langle \nabla_C \cdot (c_p \theta \mathbf{v}), \Pi \rangle}^\alpha$$
$$- \overline{\left\langle \mathbf{v}, \mu \left(\frac{\mathbf{u} \cdot \mathbf{n}_b}{\mathbf{z}_b \cdot \mathbf{n}_b} \right) \mathbf{z}_b \right\rangle}^1. \qquad (21)$$

For the Rayleigh damping term:

- μ is a damping profile that varies with height above the surface;
- the overbar $\overline{\langle \ \rangle}^1$ indicates that the term is evaluated fully implicitly in time, i.e. at time level $n+1$;
- $\mathbf{z}_b$ is the basis vector of $\mathbb{W}_2$ aligned with the vertical direction; and
- $\mathbf{n}_b$ is the basis vector of $\mathbb{W}_1$ aligned with the direction normal to the vertically facing cell face (Figure 1).

4.4.2 | Continuity equation

Multiplying Equation 11 by test functions from $\mathbb{W}_3$ and integrating over the domain D gives the weak form of the continuity equation as

$$\langle \sigma, \delta_t \rho \rangle = -\left\langle \sigma, \nabla \cdot \mathcal{F}\left(\rho^n, \bar{\mathbf{u}}^{1/2} \right) \right\rangle, \qquad (22)$$

where it is assumed that the transport scheme used to evaluate $\mathcal{F}$ returns a vector in $\mathbb{W}_2$ so that its divergence is defined everywhere.

4.4.3 | Thermodynamic equation

Multiplying Equation 12 by test functions from $\mathbb{W}_\theta$ and integrating over the domain D gives the weak form of the thermodynamic equation as

$$\langle w, \delta_t \theta \rangle = -\left\langle w, \mathcal{A}\left(\theta^n, \bar{\mathbf{u}}^{1/2} \right) \right\rangle. \qquad (23)$$

4.4.4 | Equation of state

Multiplying Equation 4 by test functions from $\mathbb{W}_3$ and integrating over the domain D gives the weak form of the equation of state as

$$\left\langle \sigma, \Pi^{\frac{1-\kappa}{\kappa}} \right\rangle = \left\langle \sigma, \frac{R}{p_0} \rho \theta \right\rangle. \qquad (24)$$

4.5 | Transformation to a reference cell

It is possible to evaluate the various integrals required for the weak formulation (i.e. Equations 21, 22, 23 and 24) directly in physical space. However, for any mesh other than one that consists of identical cells, this approach would require the evaluation of a number of integrals that are specific to each cell (for example the evaluation of the integral of the product of various basis functions). It is generally accepted that a more efficient approach is to transform the equations for each physical cell into a single, reference cell (Rognes *et al.*, 2009). Then only one set of basis functions and one set of quadrature points are needed, rather than different sets being required for each cell.

Therefore, consider a mapping $\boldsymbol{\phi} : \widehat{\mathsf{C}} \to \mathsf{C}$ between a reference cell $\widehat{\mathsf{C}}$ with coordinates $\widehat{\boldsymbol{\chi}} = \left(\widehat{\chi}_1, \widehat{\chi}_2, \widehat{\chi}_3 \right)$ and a physical cell C with coordinates $\boldsymbol{\chi} = (\chi_1, \chi_2, \chi_3)$ such that $\boldsymbol{\chi} = \boldsymbol{\phi}\left(\widehat{\boldsymbol{\chi}} \right)$. Variables and operators in the reference cell are denoted with a $\widehat{\ }$ to differentiate them from the undressed variables and operators used to indicate evaluation in the physical cell.

It is important that the transformations between the physical and reference cells preserve the various geometric properties of the mixed finite-element discretization. This would happen automatically if the metric tensor of the reference cell were the transformation of the metric tensor of the physical cell, but this would reintroduce a dependency in the reference cell on the physical cell it is mapped with. Instead a Cartesian metric tensor is assumed for the reference cell independently of the physical cell. Therefore, preservation of the required properties is achieved by using a specific collection of transformations that are specific to each function space. These transformations are designed to preserve the hierarchy of the function spaces by preserving:

(a) the appropriate continuity between cells, specifically maintaining continuity of vector components that are tangential to cell faces for vectors in $\mathbb{W}_1$ and vector components that are normal to cell faces for vectors in $\mathbb{W}_2$; and

b) the integrals appropriate to each space, i.e. pointwise evaluation for $\mathbb{W}_0$, line integrals for $\mathbb{W}_1$, area integrals for $\mathbb{W}_2$, and volume integrals for $\mathbb{W}_3$.

The transformations will collectively be referred to as Piola transformations. They are given below for each of the function spaces. More details can be found in Brezzi and Fortin (1991); Monk (2003) and Rognes *et al.* (2009). Furthermore, it is assumed that the physical space uses the same Cartesian metric tensor as the reference space.

- For scalar fields in $\mathbb{W}_0$, which represent pointwise scalars (0-forms), the transform is

$$\gamma(\chi) \equiv \gamma\left[\phi\left(\widehat{\chi}\right)\right] = \widehat{\gamma}\left(\widehat{\chi}\right), \qquad (25)$$

which satisfies

$$\nabla\gamma = \mathbf{J}^{-\mathrm{T}}\widehat{\nabla}\widehat{\gamma}, \qquad (26)$$

where the Jacobian $\mathbf{J} \equiv \partial\phi\left(\widehat{\chi}\right)/\partial\widehat{\chi}$ and $\mathbf{J}^{-\mathrm{T}} \equiv \left(\mathbf{J}^{-1}\right)^{\mathrm{T}}$.

- For vector fields in $\mathbb{W}_1$, which represent circulation vectors (1-forms), the covariant Piola transform is

$$\mathbf{c}(\chi) \equiv \mathbf{c}\left[\phi\left(\widehat{\chi}\right)\right] = \mathbf{J}^{-\mathrm{T}}\widehat{\mathbf{c}}\left(\widehat{\chi}\right), \qquad (27)$$

which satisfies

$$\nabla \times \mathbf{c}(\chi) = \frac{\mathbf{J}}{\det\mathbf{J}}\widehat{\nabla} \times \widehat{\mathbf{c}}\left(\widehat{\chi}\right). \qquad (28)$$

- For vector fields in $\mathbb{W}_2$, which represent flux vectors (2-forms), the contravariant Piola transform is

$$\mathbf{v}(\chi) \equiv \mathbf{v}\left[\phi\left(\widehat{\chi}\right)\right] = \frac{\mathbf{J}\widehat{\mathbf{v}}\left(\widehat{\chi}\right)}{\det\mathbf{J}}, \qquad (29)$$

which satisfies

$$\nabla \cdot \mathbf{v}(\chi) = \frac{1}{\det\mathbf{J}}\widehat{\nabla} \cdot \widehat{\mathbf{v}}\left(\widehat{\chi}\right). \qquad (30)$$

- For scalars in $\mathbb{W}_3$, which represent volume averaged quantities (3-forms), the transformation would naturally be

$$\sigma(\chi) \equiv \sigma\left[\phi\left(\widehat{\chi}\right)\right] = \frac{\widehat{\sigma}\left(\widehat{\chi}\right)}{\det\mathbf{J}}. \qquad (31)$$

However, use of Equation 31 would result in the weak form of the divergence transforming as

$$\int_{\mathrm{D}}\sigma\nabla \cdot \mathbf{v}\,\mathrm{d}V = \int_{\widehat{\mathrm{D}}}\frac{\widehat{\sigma}}{\det\mathbf{J}}\widehat{\nabla} \cdot \widehat{\mathbf{v}}\,\mathrm{d}\widehat{V}, \qquad (32)$$

where $\mathrm{d}\widehat{V}$ denotes the transformation of the physical volume element, $\mathrm{d}V$, and is given by $\mathrm{d}V/\det\mathbf{J}$. For non-affine cells (which in the context of the present hexahedral cells means cells that are not parallelepipeds), $\det\mathbf{J}$ is not constant within a cell and therefore Equation 32 cannot be integrated exactly using numerical integration, in fact it cannot even accurately represent a constant, so it is not even first-order accurate. The solution applied here is

TABLE 1 Transformations between physical space χ and computational space $\widehat{\chi}$, using the mapping $\phi\left(\widehat{\chi}\right) = \chi$ and $\mathbf{J} \equiv \partial\phi\left(\widehat{\chi}\right)/\partial\widehat{\chi}$

Space	Function	Differential of Function
$\mathbb{W}_0$	$\gamma = \widehat{\gamma}$	
$\mathbb{W}_1$	$\mathbf{c} = \mathbf{J}^{-\mathrm{T}}\widehat{\mathbf{c}}$	$\nabla\gamma = \mathbf{J}^{-\mathrm{T}}\widehat{\nabla}\widehat{\gamma}$
$\mathbb{W}_2$	$\mathbf{v} = \mathbf{J}\widehat{\mathbf{v}}/\det\mathbf{J}$	$\nabla \times \mathbf{c} = \mathbf{J}\widehat{\nabla} \times \widehat{\mathbf{c}}/\det\mathbf{J}$
$\mathbb{W}_3$	$\sigma = \widehat{\sigma}$	$\nabla \cdot \mathbf{v} = \widehat{\nabla} \cdot \widehat{\mathbf{v}}/\det\mathbf{J}$

rehabilitation (Bochev and Ridzal, 2010) in which the $\mathbb{W}_3$ mapping Equation 31 is modified to

$$\sigma(\chi) \equiv \sigma\left[\phi\left(\widehat{\chi}\right)\right] = \widehat{\sigma}\left(\widehat{\chi}\right). \qquad (33)$$

Equation 32 now becomes

$$\int_{\mathrm{D}}\sigma\nabla \cdot \mathbf{v}\,\mathrm{d}V = \int_{\widehat{\mathrm{D}}}\widehat{\sigma}\widehat{\nabla} \cdot \widehat{\mathbf{v}}\,\mathrm{d}\widehat{V}. \qquad (34)$$

The rehabilitation method is designed so that the order of accuracy of the scheme is maintained on arbitrary meshes. This though comes at the expense that the divergence operator applied to a vector field in $\mathbb{W}_2$ no longer maps to $\mathbb{W}_3$. However, Natale *et al.* (2016) showed that for the kind of meshes looked at here (terrain-following in a Cartesian domain), the coordinate mapping is close enough to affine that both the rehabilitated and non-rehabilitated method have similar accuracy. Therefore, the loss of the property that the divergence of a vector field in $\mathbb{W}_2$ maps into $\mathbb{W}_3$ is not expected to impact the properties of the scheme.

Table 1 summarizes the spaces and transformations for functions in each space $\mathbb{W}_0$ to $\mathbb{W}_3$. The additional function spaces $\mathbb{W}_\theta$ and $\mathbb{W}_\chi$ use the same transformations as $\mathbb{W}_0$ (i.e. those appropriate for pointwise scalars).

Additionally, Equations 14 and 15 become

$$\int_{\widehat{\mathrm{D}}}hf\det\mathbf{J}\,\mathrm{d}\widehat{V} \equiv \langle h, f\det\mathbf{J}\rangle, \qquad (35)$$

where the angle bracket notation still denotes the domain volume integral but now with respect to the reference cell coordinates.

4.6 | Discrete equations using the reference cell

Applying the coordinate transformations to Equation 21 gives:

$$\left\langle \mathbf{J}\widehat{\mathbf{v}}, \frac{\mathbf{J}\delta_t\widehat{\mathbf{u}}}{\det\mathbf{J}}\right\rangle = -\overline{\left\langle \mathbf{J}\widehat{\mathbf{v}}, \left(\mathbf{J}^{-\mathrm{T}}\widehat{\xi}\right) \times \left(\frac{\mathbf{J}\widehat{\mathbf{u}}}{\det\mathbf{J}}\right)\right\rangle}^{\alpha}$$
$$- \left\langle \mathbf{J}\widehat{\mathbf{v}}, 2\boldsymbol{\Omega} \times \frac{\mathbf{J}\widehat{\mathbf{u}}}{\det\mathbf{J}}\right\rangle$$
$$+ \overline{\left\langle \widehat{\nabla} \cdot \widehat{\mathbf{v}}, \frac{1}{2}\left(\frac{\mathbf{J}\widehat{\mathbf{u}}}{\det\mathbf{J}}\right) \cdot \left(\frac{\mathbf{J}\widehat{\mathbf{u}}}{\det\mathbf{J}}\right)\right\rangle}^{\alpha}$$
$$- \overline{\left\langle \widehat{\mathbf{v}}, \widehat{\nabla}\widehat{\Phi}\right\rangle}^{\alpha} - \overline{\langle\!\langle [\![c_\theta\widehat{\nabla}\widehat{\mathbf{v}}]\!], \{\widehat{\Pi}\}\rangle\!\rangle}^{\alpha}$$
$$+ \overline{\left\langle c_p\widehat{\theta}\widehat{\nabla}_{\mathrm{C}} \cdot \widehat{\mathbf{v}} + \widehat{\mathbf{v}} \cdot \widehat{\nabla}_{\mathrm{C}}\left(c_p\widehat{\theta}\right), \widehat{\Pi}\right\rangle}^{\alpha}$$

$$-\overline{\left\langle \mathbf{J}\widehat{\mathbf{v}}, \mu \left(\frac{\widehat{\mathbf{u}} \cdot \widehat{\mathbf{n}}_b}{\widehat{\mathbf{z}}_b \cdot \widehat{\mathbf{n}}_b} \right) \frac{\mathbf{J}\widehat{\mathbf{z}}_b}{\det \mathbf{J}} \right\rangle}^1, \qquad (36)$$

where $\langle\!\langle \cdot \rangle\!\rangle$ now denotes the surface integrals over the collection of all cell faces evaluated in the reference space and $[\![\cdot]\!]$ is defined in terms of normal vectors evaluated in the reference space. Also, the fact that the $\mathbb{W}_2$ contravariant Piola transformation (29) preserves surface integrals over cell faces has been used.

The vorticity $\widehat{\boldsymbol{\xi}}$ is obtained by evaluating

$$\left\langle \mathbf{J}^{-T}\widehat{\mathbf{c}}, \mathbf{J}^{-T}\widehat{\boldsymbol{\xi}} \det \mathbf{J} \right\rangle = \left\langle \mathbf{J}\widehat{\boldsymbol{v}} \times \widehat{\mathbf{c}}, \frac{\mathbf{J}\widehat{\mathbf{u}}}{\det \mathbf{J}} \right\rangle$$
$$- \langle\!\langle \mathbf{J}^{-T}\widehat{\mathbf{c}}, \mathbf{J}\{\widehat{\mathbf{u}}\} \times \mathbf{J}^{-T}\widehat{\mathbf{n}} \rangle\!\rangle. \quad (37)$$

For Equation 22 the coordinate transformation gives:

$$\left\langle \widehat{\sigma}, \det \mathbf{J}\delta_t\widehat{\rho} \right\rangle = -\left\langle \widehat{\sigma}, \widehat{\nabla} \cdot \widehat{\boldsymbol{\mathcal{F}}}\left(\widehat{\rho}^n, \overline{\widehat{\mathbf{u}}}^{1/2} \right) \right\rangle, \qquad (38)$$

where the fact that the transport scheme gives a flux that is in $\mathbb{W}_2$ has been used.

For Equation 23 it gives:

$$\left\langle \widehat{w}, \det \mathbf{J}\delta_t\widehat{\theta} \right\rangle = -\left\langle \widehat{w}, \det \mathbf{J}\widehat{\mathcal{A}}\left(\widehat{\theta}^n, \overline{\widehat{\mathbf{u}}}^{1/2} \right) \right\rangle. \qquad (39)$$

Here $\widehat{\mathcal{A}}$ is constructed to be in the $\mathbb{W}_\theta$ reference space (section 5.3.2) and is defined continuously as the transformation of $\mathcal{A} \equiv \mathbf{u} \cdot \nabla\theta$, i.e.

$$\widehat{\mathcal{A}} = \frac{1}{\det \mathbf{J}}\widehat{\mathbf{u}} \cdot \widehat{\nabla}\widehat{\theta}. \qquad (40)$$

And for Equation 24 it gives:

$$\left\langle \widehat{\sigma}, \det \mathbf{J}\widehat{\Pi}^{\frac{1-\kappa}{\kappa}} \right\rangle = \left\langle \widehat{\sigma}, \det \mathbf{J}\frac{R}{p_0}\widehat{\rho}\widehat{\theta} \right\rangle. \qquad (41)$$

Solutions are sought such that these equations hold for all the test functions. To achieve this, as discussed in section 4.3, each of the prognostic variables is expanded as the product of temporally varying degrees of freedom with the spatially varying set of trial functions. Thus:

$$\widehat{\mathbf{u}} = \sum_j \widetilde{u}_j(t)\widehat{\mathbf{v}}_j(\widehat{\chi}), \qquad (42)$$

$$\widehat{\rho} = \sum_j \widetilde{\rho}_j(t)\widehat{\sigma}_j(\widehat{\chi}), \qquad (43)$$

$$\widehat{\theta} = \sum_j \widetilde{\theta}_j(t)\widehat{w}_j(\widehat{\chi}), \qquad (44)$$

$$\widehat{\Pi} = \sum_j \widetilde{\Pi}_j(t)\widehat{\sigma}_j(\widehat{\chi}), \qquad (45)$$

where the sum is over all the trial functions of the appropriate space, and each of $\widetilde{u}_j, \widetilde{\rho}_j, \widetilde{\theta}_j$ and $\widetilde{\Pi}_j$ represents a vector of the degrees of freedom associated with the respective trial function $\widehat{\mathbf{v}}_j, \widehat{\sigma}_j, \widehat{w}_j$ and $\widehat{\sigma}_j$ (Appendix B gives details).

Let $\widetilde{u}$ denote the vector made up of all the coefficients $\widetilde{u}_j$, i.e. $\widetilde{u} = [\widetilde{u}_1, \widetilde{u}_2, ...]^T$, and similarly for $\widetilde{\rho}, \widetilde{\theta}$, and $\widetilde{\Pi}$. Substituting the expansions (42)–(45) into Equations 36, 38, and 39 leads

to

$$M_2\delta_t\widetilde{u} + M_\mu\overline{\widetilde{u}}^1 = \overline{R_{\mathbf{u}}}^\alpha, \qquad (46)$$

$$M_3\delta_t\widetilde{\rho} = R_\rho^F, \qquad (47)$$

$$M_\theta\delta_t\widetilde{\theta} = R_\theta^A, \qquad (48)$$

where, with the exception of the damping layer term in the momentum equation, $\overline{R_{\mathbf{u}}}^\alpha$, R_ρ^F and R_θ^A are defined to be the vectors obtained from the right-hand sides respectively of Equations 36, 38, and 39, and the components of the mass matrices are defined as

$$(M_2)_{ij} \equiv \left\langle \mathbf{J}\widehat{\mathbf{v}}_i, \frac{\mathbf{J}\widehat{\mathbf{v}}_j}{\det \mathbf{J}} \right\rangle, \qquad (49)$$

$$(M_3)_{ij} \equiv \left\langle \widehat{\sigma}_i, \det \mathbf{J}\widehat{\sigma}_j \right\rangle, \qquad (50)$$

$$(M_\theta)_{ij} \equiv \left\langle \widehat{w}_i, \det \mathbf{J}\widehat{w}_j \right\rangle, \qquad (51)$$

and

$$(M_\mu)_{ij} \equiv \left\langle \mathbf{J}\widehat{\mathbf{v}}_i, \mu \left(\frac{\widehat{\mathbf{v}}_j \cdot \widehat{\mathbf{n}}_b}{\widehat{\mathbf{z}}_b \cdot \widehat{\mathbf{n}}_b} \right) \frac{\mathbf{J}\widehat{\mathbf{z}}_b}{\det \mathbf{J}} \right\rangle. \qquad (52)$$

4.7 | Calculation of the Jacobian

For various calculations the Jacobian of the coordinate transformation from the reference cell to each physical cell, along with its determinant is required. This is achieved by setting $\boldsymbol{\phi}(\widehat{\chi}) = [\phi_1(\widehat{\chi}), \phi_2(\widehat{\chi}), \phi_3(\widehat{\chi})]$ and placing each ϕ_i in $\mathbb{W}_\chi$, for $i = 1, 2, 3$. In the interior of the computational domain, the coordinates are continuous fields. However, in a bi-periodic domain the coordinates are discontinuous across the computational "edges" of the domain where they jump by the length of the domain. Therefore $\mathbb{W}_\chi$ is chosen to be a discontinuous version of $\mathbb{W}_0$ (but for which the interior degrees of freedom are in fact continuous). The Jacobian can then be calculated everywhere it is needed and in particular it will have the appropriate values across the computational "edges" of a bi-periodic domain. Appendix D gives details of the computation of the Jacobian.

4.8 | Quadrature

In order to numerically evaluate the various spatial integrals that are required, Gaussian quadrature is used with quadrature weights denoted by λ_i and quadrature points denoted by $\widehat{\chi}_i$. Therefore, integrals of the form (35) are approximated by

$$\int_{\widehat{D}} hf \det \mathbf{J} \, d\widehat{V} \approx \sum_{i=1}^n \lambda_i h(\widehat{\chi}_i) f(\widehat{\chi}_i) \det \mathbf{J}(\widehat{\chi}_i), \qquad (53)$$

where the volume element has been absorbed into the quadrature weights, and n denotes the total number of quadrature points. Since $\det \mathbf{J}$ is polynomial (Appendix D), if both h and f are also polynomial then, provided a suitable quadrature rule with enough quadrature points is used, Equation 53 is exact. For the lowest-order ($l = 0$) elements used here a three-point Gaussian quadrature rule is used in each direction. This is exact up to fifth-order polynomials and is chosen

to ensure that, provided det $\mathbf{J}$ is constant, then all terms are integrated exactly. However, in the presence of orography, det $\mathbf{J}$ is a non-constant polynomial (Appendix D). Since in general g and f may be such that the integrand contains factors proportional to $(\det \mathbf{J})^{-1}$ the approximation to the integral will not then be exact. Results (not shown) using a five-point Gaussian quadrature are visually indistinguishable from those presented using a three-point rule, indicating that the errors due to inexact quadrature are indeed small.

5 | FINITE-VOLUME TRANSPORT DISCRETIZATION

5.1 | Method of lines advection

To complete the discretization, expressions for mass flux $\widehat{\mathcal{F}}$ and the advection of potential temperature $\widehat{\mathcal{A}}$ are required. A method-of-lines approach is used in which the temporal and spatial aspects are treated separately. The temporal aspects are handled using an explicit Runge–Kutta scheme described in section 5.2 while the spatial aspects are handled by finite-volume upwind polynomial reconstruction described in section 5.3.

The first step is to map from the finite-element degrees of freedom to finite-volume degrees of freedom. For the finite-volume degrees of freedom, a C-grid staggering in the horizontal and a Charney–Phillips staggering in the vertical is chosen. A consequence then of using the lowest-order mixed finite-element spaces described here is that there is a one-to-one correspondence between the finite-volume and finite-element degrees of freedom and the mapping between them is trivial.

Although the advecting velocity $\overline{\mathbf{u}}^{1/2}$ is updated as part of the overall scheme, it is not updated within the transport scheme itself.

5.2 | Temporal aspects

Consistent with using a finite-volume approach $\widehat{\mathcal{F}}$ and $\widehat{\mathcal{A}}$ are approximations to their average value over a time step. Let the pair of generic variables y and f denote either the pair $\widehat{\rho}$ and $\widehat{\nabla} \cdot \widehat{\mathcal{F}}$, or the pair $\widehat{\theta}$ and $\widehat{\mathcal{A}}$. Then, to evaluate the time-averaged value of f, denoted by $\overline{f}$, the equation

$$\frac{\partial y}{\partial t} = f, \tag{54}$$

is solved using an explicit Runge–Kutta scheme and $\overline{f}$ is obtained as the weighted sum of values used in the final stage of that scheme. Therefore, if the m-stage Runge–Kutta scheme is written in terms of some known coefficients a_{ij} and b_k as

$$y^{(i)} = y^n + \Delta t \sum_{j=1}^{i-1} a_{ij} f\left(y^{(j)}\right), \quad i = 1, \ldots, m, \tag{55}$$

$$y^{n+1} = y^n + \Delta t \sum_{k=1}^{m} b_k f\left(y^{(k)}\right), \tag{56}$$

then $\overline{f}$ is given by

$$\overline{f} \equiv \sum_{k=1}^{m} b_k f\left(y^{(k)}\right). \tag{57}$$

Note that, for the case of $f = \widehat{\nabla} \cdot \widehat{\mathcal{F}}$, the divergence operator used in this calculation is the same as that used in the finite-element scheme.

The specific Runge–Kutta scheme used here is the third-order, three-stage, strong stability preserving Runge–Kutta scheme (Gottlieb, 2005).

5.3 | Spatial aspects

All calculations are performed using the reference cell and its neighbours together with the specification of a uniform mesh.

5.3.1 | Mass flux

The mass flux $\widehat{\mathcal{F}}$ on a cell face is evaluated as the product of the normal component of velocity on that face with an estimation of the density on that face. The normal velocity is obtained directly as the appropriate degree of freedom of the finite-element velocity field $\widehat{\mathbf{u}}$. The value of density on the face is obtained by first constructing a one-dimensional polynomial representation of the density field as a function of the reference coordinate in the direction normal to the cell face, and then evaluating the polynomial at the cell face.

The polynomial, of even order p, is constructed using a stencil of $p+1$ cells. Noting that $p+1$ is odd, this stencil is centred about the cell that is immediately upwind of the target face. The coefficients of the polynomial are obtained by requiring that the volume integral of the polynomial over any cell in the stencil is equal to the mass in that cell. Specifically, the polynomial in powers of $\widehat{\chi}$ is given by

$$\widetilde{\rho}\left(\widehat{\chi}\right) = \sum_{i=0}^{p} a_i \widehat{\chi}^{\,i}, \tag{58}$$

with the constraint

$$\int_{\widehat{c}} \widetilde{\rho} \, d\widehat{V} = \int_{\widehat{c}} \widehat{\rho}_i \, d\widehat{V} \equiv \widehat{\rho}_i, \tag{59}$$

over all cells i in the stencil. Here $\widehat{\rho}_i$ is the value of the density field in cell i, and the fact that the volume of the reference cell is chosen to be 1 has been used. From Equation 58 the reconstructed value of $\widetilde{\rho}$ at the flux point $\widehat{\chi}_F$ can be obtained as

$$\widetilde{\rho}\left(\widehat{\chi}_F\right) = \sum_{i=0}^{p} \alpha_i \widehat{\rho}_i, \tag{60}$$

where the new coefficients α_i are linear combinations of the a_i and also depend upon $\widehat{\chi}_F$. Near the vertical boundaries of the domain, where there are not enough points to construct the polynomial, the order is reduced. This is done in steps of two in order to retain an upwind bias.

5.3.2 | Potential temperature advection

To compute $\widehat{\mathcal{A}}$ a similar method to that for the mass flux is used in that it uses a one-dimensional polynomial reconstruction but of odd order p. However, in this case the polynomial is obtained by matching directly the values of the degrees of freedom of $\widehat{\theta}$ in an appropriate stencil. The stencil has an even number $p + 1$ of points and is biased in the upwind direction. The gradient of $\widehat{\theta}$ is obtained by differentiating the polynomial and evaluating the derivative at the position of the desired degree of freedom of $\widehat{\theta}$. This derivative is multiplied by the evaluation of $\widehat{\mathbf{u}}$ at that point (where $\{\widehat{\mathbf{u}}\}$ is used where $\widehat{\mathbf{u}}$ is discontinuous).

The process is repeated for each of the three directions to provide a full three-dimensional update. Finally, this update is divided by det $\mathbf{J}$ to obtain $\widehat{\mathcal{A}}$ as defined by Equation 40.

5.3.3 | Consistent metrics

As identified by Klemp *et al.* (2003), in order to obtain accurate solutions for stratified flow over fine-scale orography, a consistent discretization of certain metric terms that arise due to the terrain-following vertical coordinate is needed.

Consider a domain with terrain-following coordinates as set out in Figure 1, where orthogonal physical unit vectors $\mathbf{x}_b$ and $\mathbf{z}_b$ are normal and parallel to gravity and orthogonal terrain following unit vectors $\mathbf{t}_b$ and $\mathbf{n}_b$ are parallel and normal to the model levels. The evolution of a scalar quantity q that is preserved over Lagrangian trajectories in this two-dimensional domain is governed by

$$\frac{\partial q}{\partial t} = -\dot{X}\frac{\partial q}{\partial X} - \dot{Z}\frac{\partial q}{\partial Z}, \tag{61}$$

where X is aligned with $\mathbf{x}_b$, $\dot{X}$ denotes the material derivative of X, Z is aligned with $\mathbf{z}_b$, and $\dot{Z}$ denotes the material derivative of Z. If q is vertically stratified ($q = q(Z)$) and the flow is horizontal ($\dot{Z} \equiv 0$) then $\partial q/\partial t = 0$.

Consider now what is required for this result to hold in terrain-following coordinates (ζ, η) where ζ is aligned with $\mathbf{t}_b$ and η with $\mathbf{n}_b$. In these coordinates

$$\frac{\partial q}{\partial t} = -\dot{\zeta}\frac{\partial q}{\partial \zeta} - \dot{\eta}\frac{\partial q}{\partial \eta}. \tag{62}$$

By inverting the transformation from (ζ, η) to (X, Z), it is found that

$$\dot{\zeta} = \frac{1}{J}\left(\frac{\partial Z}{\partial \eta}\dot{X} - \frac{\partial X}{\partial \eta}\dot{Z}\right), \tag{63}$$

$$\dot{\eta} = \frac{1}{J}\left(-\frac{\partial Z}{\partial \zeta}\dot{X} + \frac{\partial X}{\partial \zeta}\dot{Z}\right), \tag{64}$$

where J is the Jacobian of the coordinate transformation. Substituting these into Equation 62 gives

$$\frac{\partial q}{\partial t} = -\frac{1}{J}\left(\dot{X}\frac{\partial Z}{\partial \eta} - \dot{Z}\frac{\partial X}{\partial \eta}\right)\frac{\partial q}{\partial \zeta}$$
$$-\frac{1}{J}\left(-\dot{X}\frac{\partial Z}{\partial \zeta} + \dot{Z}\frac{\partial X}{\partial \zeta}\right)\frac{\partial q}{\partial \eta}. \tag{65}$$

Therefore, for q to remain constant in time when $\dot{Z} = 0$ and $q = q(Z)$, it is required that

$$-\dot{X}\frac{1}{J}\frac{dq}{dZ}\left(\frac{\partial Z}{\partial \eta}\frac{\partial Z}{\partial \zeta}\Big|_A - \frac{\partial Z}{\partial \zeta}\frac{\partial Z}{\partial \eta}\Big|_A\right) = 0, \tag{66}$$

where the subscript A indicates terms computed by the advection operator. For the terrain-following coordinate transformations used here, the dominant term is $\partial Z/\partial \zeta$ and hence the term $\partial Z/\partial \zeta|_A$ computed by the advection operator along model layers needs to match the metric term $\partial Z/\partial \zeta$ contained in the advecting velocity $\dot{\eta}$ normal to model layers.

As in Melvin *et al.* (2010), this is achieved by modifying Equation 64 such that the $\partial Z/\partial \zeta$ term is computed by the advection operator. In practice this is achieved by modifying the velocity vector used in the advection scheme according to

$$\widehat{u}_i^* = \widehat{u}_i + \widehat{\mathbf{u}} \cdot \widehat{\nabla}\chi_i - \widehat{\mathcal{A}}\left(\chi_i, \widehat{\mathbf{u}}\right)\det(\mathbf{J}), \quad i = 1, 2, 3, \tag{67}$$

where $\widehat{u}_i$ is the ith component of $\widehat{\mathbf{u}}$ and $\widehat{\mathcal{A}}\left(\chi_i, \widehat{\mathbf{u}}\right)$ is the advection operator applied to the ith component of the coordinate field χ. This modification means that the metric term $\left(\widehat{\nabla}\chi_i\right)$ component of $\widehat{u}_i$, as computed by the finite-element scheme, is replaced by that computed by the high-order upwind scheme $\widehat{\mathcal{A}}(\chi_i, \cdot)$. This then means that both the metric term components and the advected components of χ are computed using the same scheme and so there is no inconsistency. Although in principle this modification can be applied to the velocity vector for all advection terms, in order to remove the distortion over fine-scale orography, it is only necessary to apply it to the potential temperature advection term, such that $\widehat{\mathbf{u}^*}^{1/2}$ is used to advect $\widehat{\theta}^n$ in Equation 40. Furthermore, for all the examples presented here a uniform mesh in the χ_1 and χ_2 directions is used such that Equation 67 only results in modifying the vertical component of the velocity $\widehat{u}_3$.

6 | SOLUTION PROCEDURE

6.1 | Notation

The convention followed below is that calligraphic $\mathcal{R}$s indicate residuals of the equations to be solved. Gothic $\mathfrak{R}$s denote linear combinations of these residuals. The italic Rs of section 4.4 indicate right-hand sides of the equations.

6.2 | Overview

The governing Equations 46, 47, 48, and 41, can be compactly written as

$$\mathcal{R}\left(\mathbf{x}^{n+1}\right) = 0, \tag{68}$$

where $\mathbf{x}^{n+1} \equiv \left[\widehat{\mathbf{u}}^{n+1}, \widehat{\rho}^{n+1}, \widehat{\theta}^{n+1}, \widehat{\Pi}^{n+1}\right]^{\mathrm{T}}$ is the sought-after state vector at the next time step. A full Newton method would solve this equation iteratively as

$$\mathcal{J}\left(\mathbf{x}^{(k)}\right)\mathbf{x}' = -\mathcal{R}\left(\mathbf{x}^{(k)}\right), \tag{69}$$

where $\mathbf{x}' \equiv \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$ is the increment to the state vector, $\mathcal{J}$ is the Jacobian of $\mathcal{R}$ with respect to $\mathbf{x}$, and superscript (k) indicates the iteration index of the Newton loop. However, $\mathcal{J}$ is a large matrix and its inverse is dense, therefore a quasi-Newton method is used in which the Jacobian is approximated by a simpler linear system

$$\mathcal{J}\left(\mathbf{x}^{(k)}\right)\mathbf{x}' \approx \mathcal{L}\mathbf{x}'. \tag{70}$$

Note that, since both $\mathcal{J}$ and $\mathcal{L}$ operate on increments to the solution, at convergence of the iterative solution, the same full, nonlinear Equation 68 is solved. The choice of whether or how to approximate $\mathcal{J}$ affects whether, and how quickly, the iterative scheme convergences and also the computational efficiency of the scheme.

Following the approach of Wood *et al.* (2014), the choice of the linear operator $\mathcal{L}$ is inspired by the linearization of $\mathcal{R}$ about some reference state $\mathbf{x}^*$ to obtain $\mathcal{L}(\mathbf{x}^*)$ and then solve

$$\mathcal{L}(\mathbf{x}^*)\mathbf{x}' = -\mathcal{R}\left(\mathbf{x}^{(k)}\right). \tag{71}$$

The spatially continuous form of $\mathcal{L}$ in physical space is given by

$$\mathcal{L}\left(\mathbf{x}^*_{\text{phys}}\right)\mathbf{x}'_{\text{phys}} = \begin{cases} \mathbf{u}' - \mu\left(\frac{\mathbf{n_b}\cdot\mathbf{u}'}{\mathbf{n_b}\cdot\mathbf{z_b}}\right)\mathbf{z_b} \\ \quad +\tau_u\Delta t c_p\left(\theta'\nabla\Pi^* + \theta^*\nabla\Pi'\right), \\ \rho' + \tau_\rho\Delta t\nabla\cdot\left(\rho^*\mathbf{u}'\right), \\ \theta' + \tau_\theta\Delta t\mathbf{u}'\cdot\nabla\theta^*, \\ \frac{1-\kappa}{\kappa}\frac{\Pi'}{\Pi^*} - \frac{\rho'}{\rho^*} - \frac{\theta'}{\theta^*}, \end{cases} \tag{72}$$

where $\tau_{u,\theta,\rho}$ are relaxation parameters. Applying the mixed finite-element discretization presented above to these operators results in:

$$M_2^\mu\widetilde{u}' - P_{2\theta}^{\Pi^*}\widetilde{\theta}' - G^{\theta^*}\widetilde{\Pi}' = -\mathcal{R}_{\mathbf{u}}, \tag{73}$$

$$M_3\widetilde{\rho}' + D\left(\widehat{\rho}^*\widetilde{u}'\right) = -\mathcal{R}_\rho, \tag{74}$$

$$M_\theta\widetilde{\theta}' + P_{\theta 2}^{\theta^*}\widetilde{u}' = -\mathcal{R}_\theta, \tag{75}$$

$$M_3^{\Pi^*}\widetilde{\Pi}' - M_3^{\rho^*}\widetilde{\rho}' - P_{3\theta}^*\widetilde{\theta}' = -\mathcal{R}_\Pi. \tag{76}$$

Note that, at convergence of the iterative procedure, primed quantities vanish and Equation 68 is solved, which, given the definitions (78)–(81) below, is equivalent to solving Equations 46–48 and 41. In these expressions M_2^μ is the operator formed by combining the $\mathbb{W}_2$ mass matrix with the operator arising from the Rayleigh damping:

$$M_2^\mu \equiv M_2 + \Delta t M_\mu. \tag{77}$$

Additionally, a number of definitions have been used here. First, from Equations 46–48 and 41 the residuals from the current estimate of the solution are defined as

$$\mathcal{R}_{\mathbf{u}} \equiv \Delta t\left(M_2\delta_t\widetilde{u} + M_\mu\overline{\widetilde{u}}^{-1} - \overline{R_{\mathbf{u}}}^\alpha\right), \tag{78}$$

$$\mathcal{R}_\rho \equiv \Delta t\left(M_3\delta_t\widetilde{\rho} - R_\rho^F\right), \tag{79}$$

$$\mathcal{R}_\theta \equiv \Delta t\left(M_\theta\delta_t\widetilde{\theta} - R_\theta^A\right), \tag{80}$$

and

$$\mathcal{R}_\Pi \equiv \left\langle\widehat{\sigma}, \det\mathbf{J}\left[\left(\widehat{\Pi}^{(k)}\right)^{\frac{1-\kappa}{\kappa}} - \frac{R}{p_0}\widehat{\rho}^{(k)}\widehat{\theta}^{(k)}\right]\right\rangle, \tag{81}$$

where the time level $n+1$ variables in the definitions (9) for δ_t and (10) for $\overline{}^\alpha$ have been replaced by the latest iterates, denoted by superscript (k). Second, various operators are defined as

$$D \equiv \tau_\rho\Delta t\left\langle\widehat{\sigma}, \widehat{\nabla}\cdot\widehat{\mathbf{v}}\right\rangle, \tag{82}$$

$$G^{\theta^*} \equiv \tau_u\Delta t\left\langle c_p\widehat{\theta}^*\widehat{\nabla}_C\cdot\widehat{\mathbf{v}} + \widehat{\mathbf{v}}\cdot\widehat{\nabla}_C\left(c_p\widehat{\theta}^*\right), \widehat{\sigma}\right\rangle$$
$$-\tau_u\Delta t\langle\!\langle[\![c_p\widehat{\theta}^*\widehat{\mathbf{v}}]\!], \{\widehat{\sigma}\}\rangle\!\rangle, \tag{83}$$

$$P_{2\theta}^{\Pi^*} \equiv \tau_u\Delta t\left\langle c_p\widehat{w}\widehat{\nabla}_C\cdot\widehat{\mathbf{v}}, \widehat{\Pi}^*\right\rangle$$
$$+\tau_u\Delta t\left\langle\widehat{\mathbf{v}}\cdot\widehat{\nabla}_C\left(c_p\widehat{w}\right), \widehat{\Pi}^*\right\rangle$$
$$-\tau_u\Delta t\langle\!\langle[\![c_p\widehat{w}\widehat{\mathbf{v}}]\!], \{\widehat{\Pi}^*\}\rangle\!\rangle, \tag{84}$$

$$M_3^{\rho^*} \equiv \left\langle\widehat{\sigma}, \frac{\widehat{\sigma}}{\widehat{\rho}^*}\det\mathbf{J}\right\rangle, \tag{85}$$

$$M_3^{\Pi^*} \equiv \frac{1-\kappa}{\kappa}\left\langle\widehat{\sigma}, \frac{\widehat{\sigma}}{\widehat{\Pi}^*}\det\mathbf{J}\right\rangle, \tag{86}$$

$$P_{\theta 2}^{\theta^*} \equiv \tau_\theta\Delta t\langle\!\langle[\![\widehat{w}\widehat{\mathbf{v}}]\!], \{\widehat{\theta}^*\}\rangle\!\rangle$$
$$-\tau_\theta\Delta t\left\langle\widehat{\nabla}_C\cdot\left(\widehat{w}\widehat{\mathbf{v}}\right), \widehat{\theta}^*\right\rangle, \tag{87}$$

$$P_{3\theta}^* \equiv \left\langle\widehat{\sigma}, \frac{\widehat{w}}{\widehat{\theta}^*}\det\mathbf{J}\right\rangle. \tag{88}$$

The subscripted operators P_{ij} denote projections that map from $\mathbb{W}_j$ to $\mathbb{W}_i$ (Appendix E gives the derivation of $P_{\theta 2}^{\theta^*}$); D is a divergence operator that maps from $\mathbb{W}_2$ to $\mathbb{W}_3$, and G^{θ^*} is a gradient operator that maps from $\mathbb{W}_3$ to $\mathbb{W}_2$.

The system of Equations 73–76 is solved using an iterative Krylov method that is preconditioned by an approximate Schur complement of the equations for the pressure increment. The approximate Schur complement is formed by using lumped forms of the M_2^μ and M_θ mass matrices. The right-hand side terms are then updated using the latest estimates for the prognostic variables. This includes the FV transport terms through the updated advecting wind field; Table 2 gives details. This process is iterated a number of times. For all results presented here, four iterations are used.

7 | COMPUTATIONAL EXAMPLES

In the following, the results are presented of model runs on standard Cartesian benchmarks of atmospheric dynamics, drawing on the suite considered in Melvin *et al.* (2010) for the vertical slice tests, with additional 3D tests. The boundary set-up has doubly periodic boundary conditions in the horizontal and zero flux on the top and bottom boundaries. The test parameters are summarized in Table 3. As noted, while in principle the finite-element methodology affords flexibility on the polynomial order, here the focus will be on results in the

TABLE 2 Outline of the iterative solution procedure used within a time step

do $n = 1, N$ (begin time-step loop)

– Given the solution $\mathbf{x}^n \equiv \left(\widehat{\mathbf{u}}, \widehat{\rho}, \widehat{\theta}, \widehat{\Pi}\right)^n$, let the first

estimate for $\mathbf{x}^{n+1} \equiv \left(\widehat{\mathbf{u}}, \widehat{\rho}, \widehat{\theta}, \widehat{\Pi}\right)^{n+1}$ be $\mathbf{x}^{(0)} = \mathbf{x}^n$.

– Set $\mathbf{x}^* \equiv (\rho^*, \theta^*, \Pi^*) = \left(\widehat{\rho}, \widehat{\theta}, \widehat{\Pi}\right)^n$ and compute

the operators in Equation 73–76 for $\mathcal{L}(\mathbf{x}^*)$.

– Compute the time-level n components of $\left(\mathcal{R}_\mathbf{u}, \mathcal{R}_\rho, \mathcal{R}_\theta, \mathcal{R}_\Pi\right)$.

do $k = 1, K$ (Newton iteration)

– Set the advecting wind $\widehat{\overline{\mathbf{u}}}^{1/2} \equiv \left(\widehat{\mathbf{u}}^{(k-1)} + \widehat{\mathbf{u}}^n\right)/2$.

– Compute the advection terms $\mathcal{F}\left(\widehat{\rho}^n, \widehat{\overline{\mathbf{u}}}^{1/2}\right), \mathcal{A}\left(\widehat{\theta}^n, \widehat{\overline{\mathbf{u}}}^{1/2}\right)$.

– Compute the time-level $n + 1$ components of $\left(\mathcal{R}_\mathbf{u}, \mathcal{R}_\rho, \mathcal{R}_\theta, \mathcal{R}_\Pi\right)$.

– Solve Equation 71 to obtain the increments $\mathbf{x}' \equiv \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$

and hence the updated estimate $\mathbf{x}^{(k)}$ for the time-level $n + 1$ fields.

end do

end do

lowest-order case. Additionally a number of simplifications and specifications are made:

- All tests are in a non-rotating frame such that $\mathbf{\Omega} \equiv 0$.
- There are no monotonicity constraints applied to the model.
- All meshes are uniform and orthogonal in the χ_1 and χ_2 directions; this results in a simplified form of the coordinate Jacobian $\mathbf{J}$ (Appendix D).
- The coordinate fields are obtained from

$$\chi_1 = (1 - \varepsilon_1)\chi_1^- + \varepsilon_1\chi_1^+, \tag{89}$$

$$\chi_2 = (1 - \varepsilon_2)\chi_2^- + \varepsilon_2\chi_2^+, \tag{90}$$

$$\chi_3 = (1 - \varepsilon_3)\chi_3^-(\chi_1, \chi_2) + \varepsilon_3\chi_3^+, \tag{91}$$

where $\left(\chi_1^-, \chi_1^+\right)$ and $\left(\chi_2^-, \chi_2^+\right)$ are the constant minimum and maximum values of χ_1 and χ_2 respectively. $\chi_3^-(\chi_1, \chi_2)$ is the orographic profile, χ_3^+ is the constant height of the domain top and $\varepsilon_i, i = 1, 2, 3$ is a parameter that takes values between zero and one.

- The semi-implicit scheme is centred in time so that $\alpha = 1/2$ and the relaxation parameters $\tau_{u,\rho,\theta} = 1/2$. Additionally, four iterations of the Newton loop are used.

- Following Wood *et al.* (2014), the reference profiles $\mathbf{x}^*$ are taken to be based upon the start of time-step fields $\mathbf{x}^* \equiv \mathbf{x}^n$, however in contrast to Wood *et al.* (2014) there is no further modification of the profiles (i.e. static adjustment applied to θ^* or recomputation of ρ^*).
- A quadratic reconstruction of the density is used for the mass flux $\widehat{\mathcal{F}}$ and a quadratic reconstruction of the gradient term in the advective update $\widehat{\mathcal{A}}$.
- Where applied, the Rayleigh damping profile takes the same form as used in Melvin *et al.* (2010), i.e.

$$\mu(z) = \begin{cases} 0, & z < z_B, \\ \overline{\mu}\sin^2\left[\frac{\pi}{2}\left(\frac{z - z_B}{z_T - z_B}\right)\right] & z \geq z_B, \end{cases} \tag{92}$$

where z_B is the height at which the Rayleigh starts, z_T is the top of the model domain and the parameter $\overline{\mu}$ is specific to each test case.

For the Cartesian domain used in these examples, the general coordinate χ used previously is replaced by the standard Cartesian coordinates $\mathbf{x}$ so that $(\chi_1, \chi_2, \chi_3) \equiv (x, y, z)$. In this section w is used to denote the vertical component of the velocity $\mathbf{u}$ (i.e. $w = D\chi_3/Dt = Dz/Dt$) as distinct from the test function for $\mathbb{W}_\theta$.

7.1 | Non-hydrostatic gravity waves

First, the model is tested on the non-hydrostatic gravity wave test in Skamarock and Klemp (1994). In a two-dimensional domain, $(x, z) \in [-150, 150]\,\text{km} \times [0, 10]\,\text{km}$, a potential temperature perturbation of the form:

$$\theta' = \frac{\theta_0 \sin(\pi z/H)}{1 + [(x - x_c)/a]^2}, \tag{93}$$

with $\theta_0 = 0.01\,\text{K}$, $x_c = 0$, $a = 5\,\text{km}$, $H = 10\,\text{km}$, is superposed on a background atmosphere with constant buoyancy frequency $N = 0.01\,\text{s}^{-1}$ and a horizontal wind $U = 20\,\text{m/s}$. The initial potential temperature perturbation spreads out in the form of gravity waves (Figure 2). The final perturbation

TABLE 3 Model parameters for each test. The Courant number is given by $C \equiv U\Delta t/\Delta x$: for cases where $U = 0$, C has been calculated using the largest value of $\mathbf{u}'$. The acoustic Courant number (using a representative speed of sound $C_s = 343$ m/s) varies from 3 up to 40. For all cases the surface pressure (away from, or in the absence of, orography) is $p_{surf} = 1\,000$ hPa $\Rightarrow \pi_{surf} = 1$. For 3DBH and 3DRB, $\Delta y = \Delta x$ and the domain size in the y direction is 40 and 1 km respectively

Test	Δx (km)	Δz (m)	Δt (s)	Domain (km×km)	Background Initial State	T_{surf} (K)	U (m/s)	C
NHGW	1	1,000	12	300 × 10	$N = 0.01\,\text{s}^{-1}$	300	20	0.24
DC	0.025–0.4	25–400	0.25–4	51.2 × 6.4	Isentropic $\theta = T_{surf}$	300	0	≈ 0.4
HMW	2	250	20	240 × 50	Isothermal $T = T_{surf}$	250	20	0.2
NHMW	0.4	250	5	144 × 35	$N = 0.01\,\text{s}^{-1}$	300	10	0.125
SH	0.5	300	8	100 × 30	$N = 0.01\,\text{s}^{-1}$	288	10	0.16
3DBH	0.2	200	4	60 × 16	$N = 0.01\,\text{s}^{-1}$	293.15	10	0.2
3DRB	0.01	10	1.25	1 × 1.5	Isentropic $\theta = T_{surf}$	300	0	≈ 0.36

NHGW = Non-Hydrostatic Gravity Waves; DC = Density Current; HMW = Hydrostatic Mountain Waves; NHMW = Non-Hydrostatic Mountain Waves; SH = Schär Hill; 3DBH = Three-Dimensional Bell-shaped Hill; 3DRB = Three-Dimensional Rising Bubble
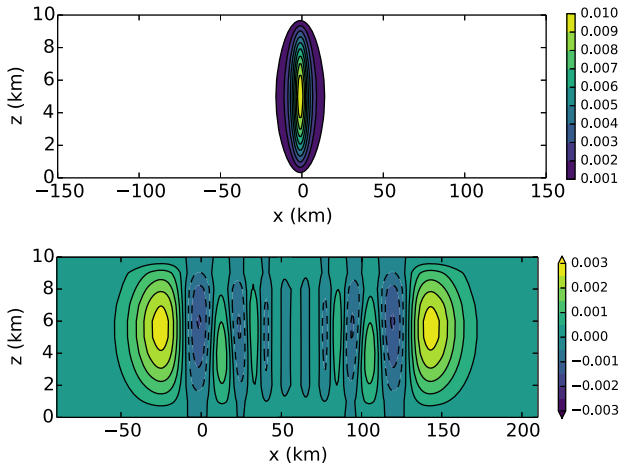
FIGURE 2 Potential temperature perturbation (a) at the initial time (contours every 10^{-3} K) and (b) at time $t = 3,000$ s (contours every 5×10^{-4} K) for the non-hydrostatic inertia-gravity wave test [Colour figure can be viewed at wileyonlinelibrary.com]

is in line with results in the literature, and in particular with Melvin *et al.* (2010).

7.2 | Density current

Next, the case of a falling cold air bubble in Straka *et al.* (1993) is considered. A negative thermal perturbation:

$$T' = \begin{cases} 0 \text{ K} & \text{if } r > 1, \\ -\frac{15}{2}[1 + \cos(\pi r)] \text{ K} & \text{if } r < 1, \end{cases} \quad (94)$$

where $r = \left\{[(x - x_c)/x_r]^2 + [(z - z_c)/z_r]^2\right\}^{0.5}$,
$x_c = 0$ km, $x_r = 4$ km, $z_c = 3$ km and $z_r = 2$ km, is superposed on a $[-25.6, 25.6]$ km $\times [0, 6.4]$ km motionless isentropic atmosphere with constant background $\theta = T_{\text{surf}} = 300$ K. This test includes an additional artificial diffusion term applied to the potential temperature and the components of the velocity vector of the form $\nu\nabla^2\phi$, where ϕ is the prognostic variable, and $\nu = 75$ m^2/s . Driven by its negative buoyancy, the bubble falls, hits the bottom boundary and moves outward, developing vortices (Figure 3). Convergence with increasing resolution is evident from the final potential temperature distribution. At the finest resolution and final time, the minimum perturbation value and the front location match the results in Melvin *et al.* (2010) to within less than one percent (Table 4).

7.3 | Linear hydrostatic/non-hydrostatic flow over a hill

The ability of the model to simulate orographically driven flow is tested with idealized profiles of increasing slope. A height-based terrain-following coordinate (Equation 91) is used as in Melvin *et al.* (2010), and the thermodynamic variables are initially hydrostatically balanced.

In the first two tests, the bottom boundary is described by the function:

$$z_S = \frac{h_m}{1 + (x/a)^2}. \quad (95)$$
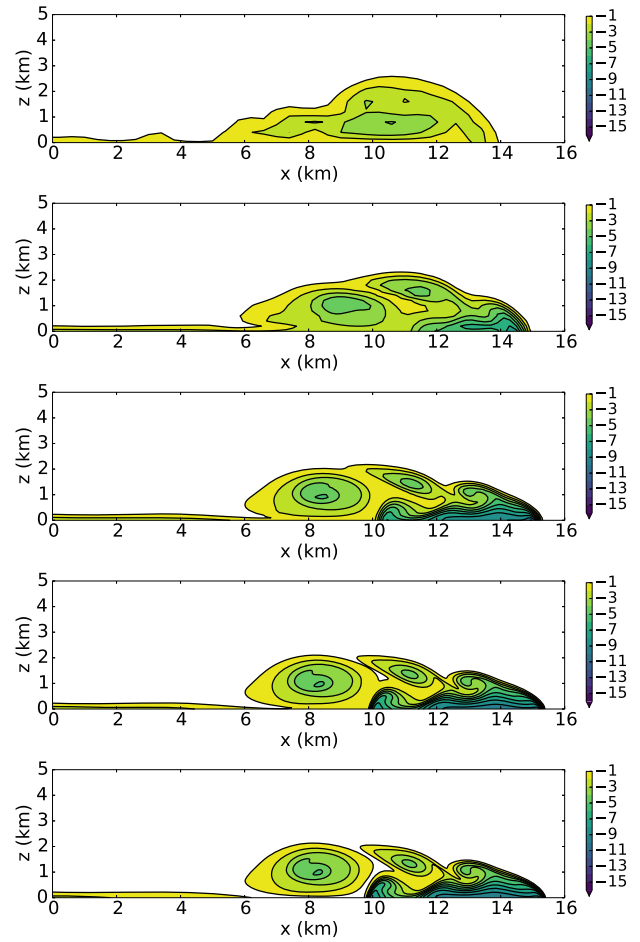
FIGURE 3 Potential temperature perturbation at time $t = 900$ s for the density current test and resolutions $\Delta x = $ (a) 400 m, (b) 200 m, (c) 100 m, (d) 50 m, and (e) 25 m. Contours are plotted in the range $[-16, -1]$ K with a 1 K interval [Colour figure can be viewed at wileyonlinelibrary.com]

In particular, a hydrostatic flow is simulated by setting the height $h_m = 1$ m, half-width $a = 10$ km, and background wind speed $U = 20$m/s. The domain is a 240 km $\times$ 50 km isothermal atmosphere with background temperature $T = 250$ K. A damping layer is used in the topmost 20 km of the domain, with $\overline{\mu}\Delta t = 0.3$, and final time $t = 15000$ s.

A non-hydrostatic flow is simulated using the same height h_m, a half-width $a = 1$ km, background wind speed $U = 10$m/s, a 144 km $\times$ 35 km atmosphere with surface temperature $T_{\text{surf}} = 300$ K and constant buoyancy frequency $N = 0.01$ s^{-1}, a damping layer in the topmost 10 km with

TABLE 4 Minimum and maximum θ perturbation from the background state $\theta_{\text{surf}} = 300$ K and front location (rightmost intersection of -1 K contour with $z = 0$) in the density current test.

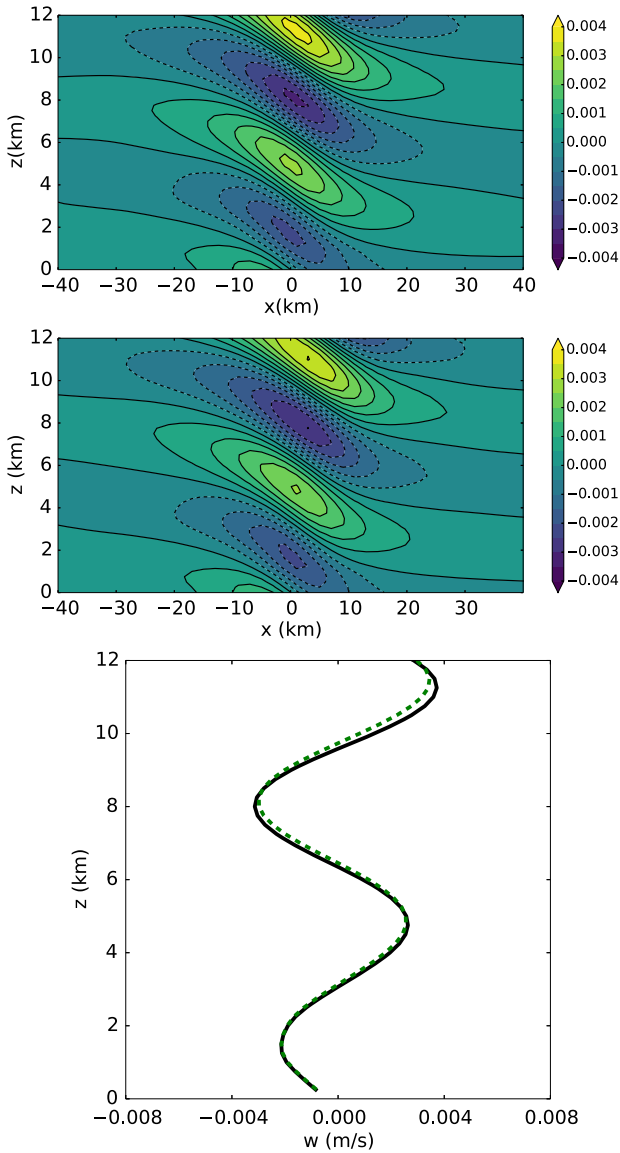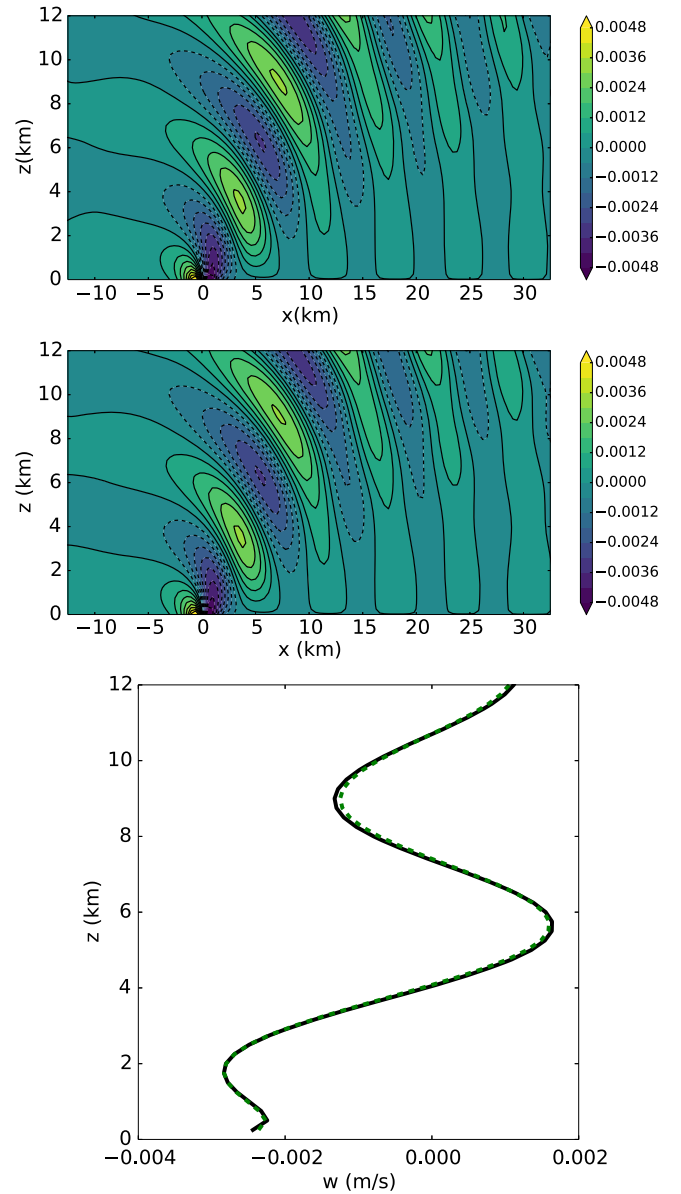| Grid size (m)	$\Delta\theta_{\min}$ (K)	$\Delta\theta_{\max}$ (K)	Front location (m)
400	−4.0704	0.5194	13939
200	−7.6091	0.1158	14941
100	−10.1768	0.1233	15313
50	−9.5342	0.0626	15384
25	−9.6589	0.0047	15402

FIGURE 4 Vertical velocity perturbation w for the hydrostatic mountain wave test. (a) Hydrostatic case result at time $t = 15,000$ s, with contours in the range $[-4, 4] \times 10^{-3}$ m/s with a 5×10^{-4} m/s interval, (b) ENDGame (Melvin *et al.*, 2010), and (c) profiles at $x = 0$ for the results from the current model (solid line) and from ENDGame (dashed line) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 5 Vertical velocity perturbation w for the non-hydrostatic mountain wave test. (a) Non-hydrostatic case result at time $t = 9,000$ s, contours in the range $[-4.8, 4.8] \times 10^{-3}$ m/s with a 6×10^{-4} m/s interval. (b) ENDGame (Melvin *et al.*, 2010), and (c) profiles at $x = 0$ for the results from the current model (solid line) and from ENDGame (dashed line) [Colour figure can be viewed at wileyonlinelibrary.com]

$\overline{\mu} \Delta t = 0.15$, and final time $t = 9000$ s. In both the hydrostatic (Figure 4) and non-hydrostatic (Figure 5) flow cases, the shape of the vertical velocity at final time of the mixed finite-element simulation compares favourably both with the results from the semi-implicit semi-Lagrangian model of Melvin *et al.* (2010) (referred to as ENDGame) and also the linear analytic solution presented there.

7.4 | Schär hill

In the third test (Schär *et al.* 2002) a mountain range is considered with bottom boundary profile given by the function

$$z_S = h_m e^{-(x/a)^2} \cos^2\left(\frac{\pi x}{\lambda}\right), \tag{96}$$

with $h_m = 250$ m, $\lambda = 4$ km, and $a = 5$ km. The domain is a $100\,\text{km} \times 30\,\text{km}$ atmosphere with surface temperature $T_{\text{surf}} = 288$ K, constant buoyancy frequency $N = 0.01\,\text{s}^{-1}$, and a background wind $U = 10$ m/s. A damping layer is used in the top 10 km, with $\overline{\mu} \Delta t = 1.2$. At final time $t = 2250$ s, the vertical velocity distribution matches the ENDGame result reasonably well (Figure 6), and the amplitude of the waves above the mountain is again similar to both ENDGame and the results of Klemp *et al.* (2003). As with ENDGame (Melvin *et al.*, 2010), in order to obtain solutions that compare well with the linear solution, a consistent discretization of certain metric terms is needed, as described in section 5.3.3. Without this correction, the same distortion as seen in Klemp
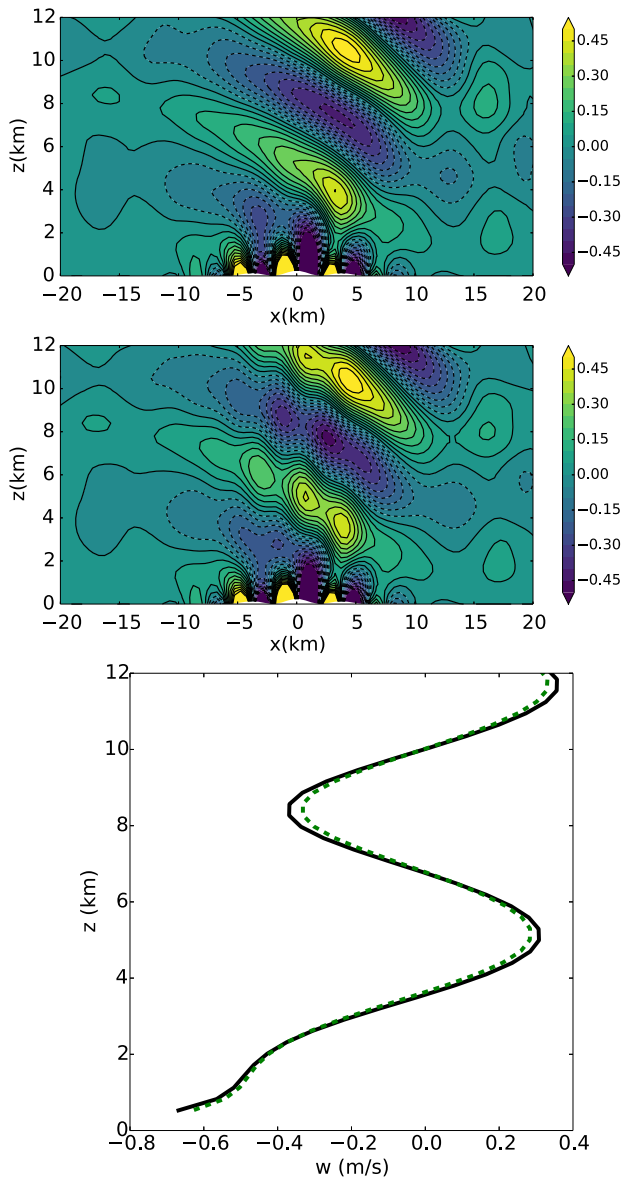
FIGURE 6 Vertical velocity perturbation w after 5 hr for the 2D Schär hill test. Contours are plotted in the range $[-0.5, 0.5]$ m/s with a 5×10^{-2} m/s interval. (a) $\Delta t = 8$ s, (b) $\Delta t = 8$ s with inconsistent metric terms, and (c) profiles at $x = 0$ for the results with consistent metrics from the current model (solid line) and from ENDGame (dashed line) from Melvin *et al.* (2010) [Colour figure can be viewed at wileyonlinelibrary.com]

et al. (2003) and Melvin *et al.* (2010) of the waves above the mountain is present (Figure 6b).

7.5 | 3D medium-steep bell-shaped hill

Next, the three-dimensional flow over a bell-shaped hill of Lock *et al.* (2012) and Yamazaki *et al.* (2016) is considered. The bottom boundary profile is:

$$ z_S = \frac{h_m}{\left[1 + (x/a)^2 + (y/a)^2\right]^{3/2}}, \quad (97) $$

with $h_m = 400$ m, $a = 1$ km. The maximum value of the derivative of Equation 97 corresponds to an approximate slope of 20°. The domain is a $60\,\text{km} \times 40\,\text{km} \times 16\,\text{km}$
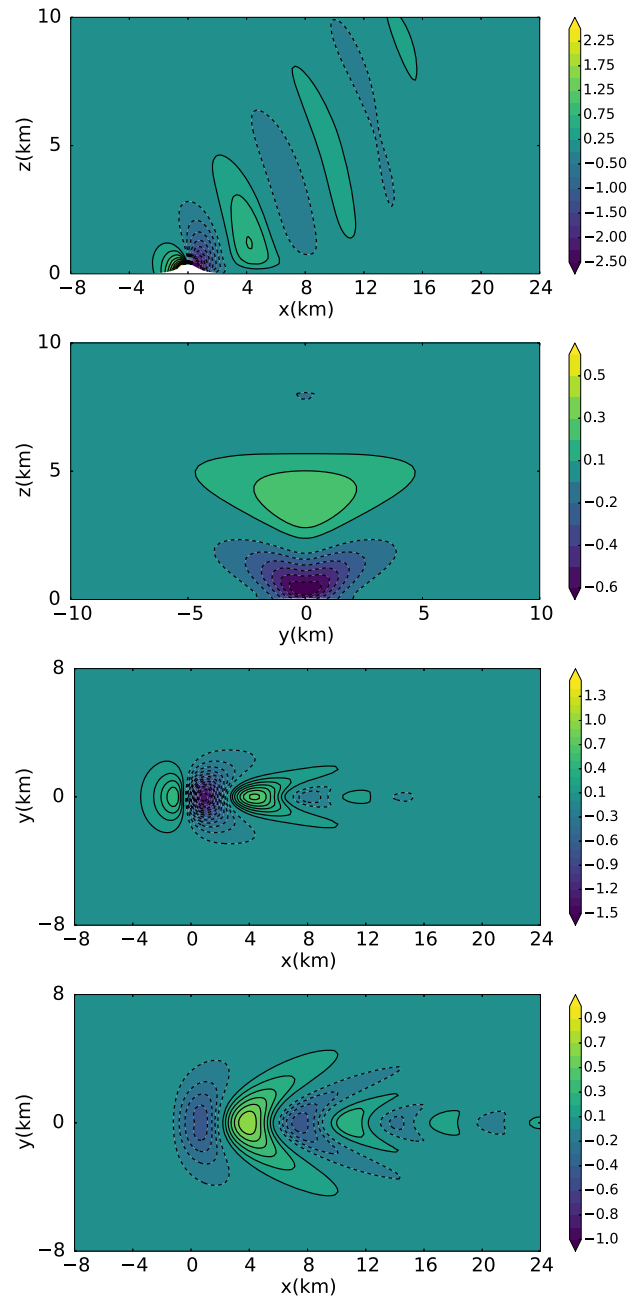
FIGURE 7 Vertical velocity perturbation after 1 hr (900 time steps) for the bell-shaped hill test. (a) x–z slice at $y = 0$, (b) y–z slice at $x = 2,000$ m, (c) x–y slice at $z = 800$ m, (d) x–y slice at $z = 2,000$ m,. The contour interval is 0.25 m/s for (a), and 0.1 m/s otherwise [Colour figure can be viewed at wileyonlinelibrary.com]

atmosphere with surface temperature $T_{\text{surf}} = 293.15$ K, constant buoyancy frequency $N = 0.01\,\text{s}^{-1}$, and a background wind $U = 10$m/s. A damping layer is used in the top 6 km, with $\overline{\mu}\Delta t = 1.2$. The vertical velocity distribution at final time $t = 3600$ s is in line with the literature (Figure 7; cf. figure 7 in Lock *et al.*, 2012 and figure 10 in Yamazaki *et al.*, 2016).

7.6 | 3D rising bubble

The 3D rising bubble test of Kelly and Giraldo (2012) is used. This simulates a buoyant thermal bubble on a neutrally
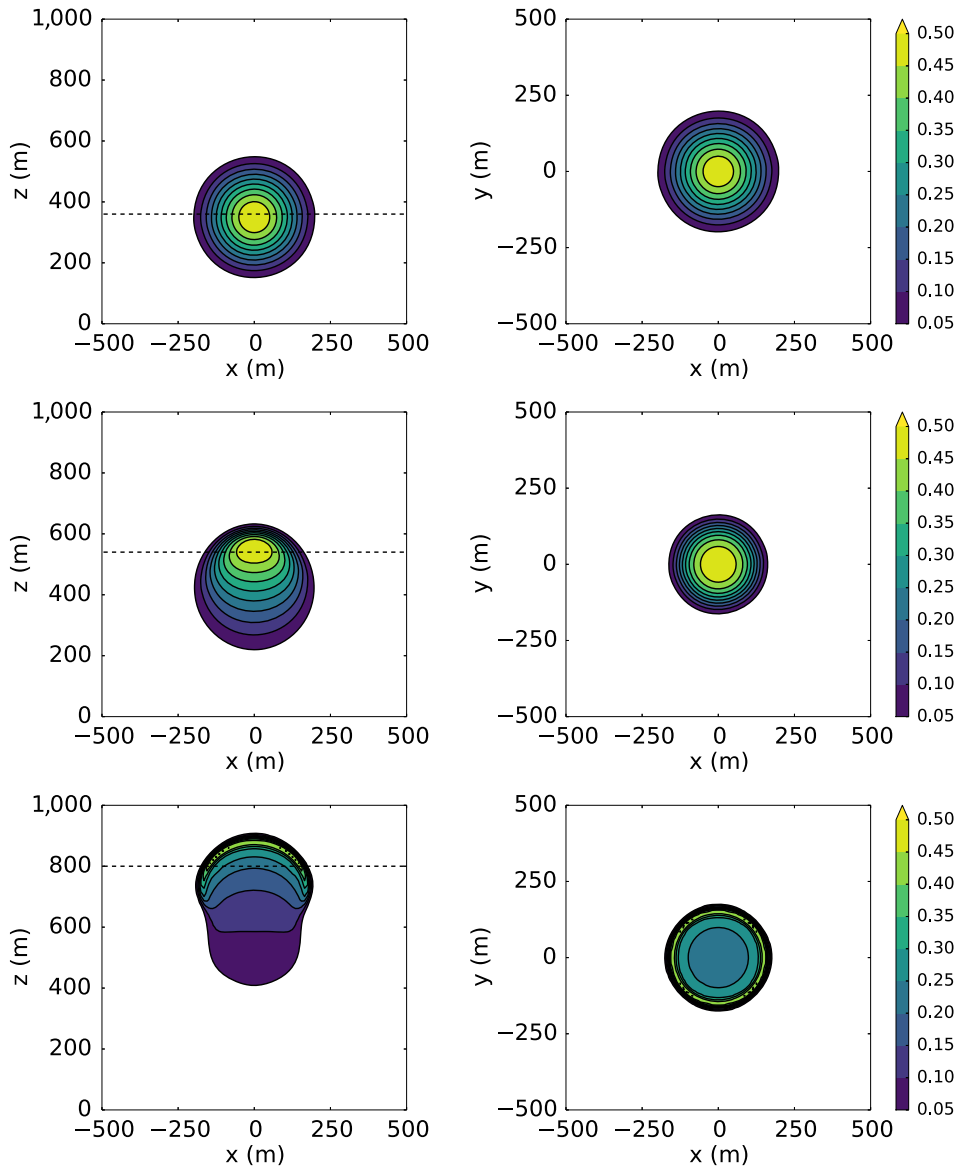
FIGURE 8 Potential temperature perturbation from background state of $\theta = 300$ K for the 3D rising bubble test. (a, c, e) x–z slices at $y = 0$. (b, d, f) x–y slices taken at the level indicated by the dashed line in the x–z slices. (a, b) show initial data, (c, d) at $t = 200$ s and (e, f) at $t = 400$ s, (all with contours in the range [0.05, 0.5] K with a 0.05 K interval) [Colour figure can be viewed at wileyonlinelibrary.com]

stratified isentropic background state with $\theta = T_{surf} = 300$ K in a domain 1 km $\times$ 1 km $\times$ 1.5 km. A spherical perturbation of radius $r_0 = 250$ m, located at $(x_0, y_0, z_0) = (0, 0, 350)$ m is added to the background state. The perturbation is defined by

$$\theta' = \begin{cases} A\left[1 + \cos\left(\frac{\pi r}{r_0}\right)\right] & r \le r_0, \\ 0 & r > r_0, \end{cases} \quad (98)$$

with $r = \sqrt{\left[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right]}$ and $A = 0.25$ K as in Abdi and Giraldo (2016). Snapshots of the bubble at $t = 0, 200$ and 400 s are shown in Figure 8 and a one-dimensional cross-section at $x = y = 0$ in Figure 9. These results compare well with those of Kelly and Giraldo (2012), notably maintaining a maximum value of the perturbation close to 0.5 K and only exhibiting a small undershoot above the bubble, comparable to the discontinuous Galerkin (DG) method of Kelly and Giraldo (2012).

8 | SUMMARY

A method for coupling a mixed finite-element method alongside a finite-volume transport scheme and an iterative semi-implicit time scheme has been presented. This method seeks to combine the benefits of all three schemes: the numerical consistency (independent of the mesh) and accurate wave dispersion properties of the mixed finite-element scheme, the flexibility and accuracy of a high-order upwind treatment of scalar advection from the finite-volume transport scheme, and the stable treatment of physically insignificant fast waves by the semi-implicit scheme.

The resulting model has been applied to a standard set of two-dimensional and three-dimensional test cases in Cartesian domains from the literature. This model has a similar level of accuracy to other models on these tests, including the semi-implicit semi-Lagrangian ENDGame dynamical core
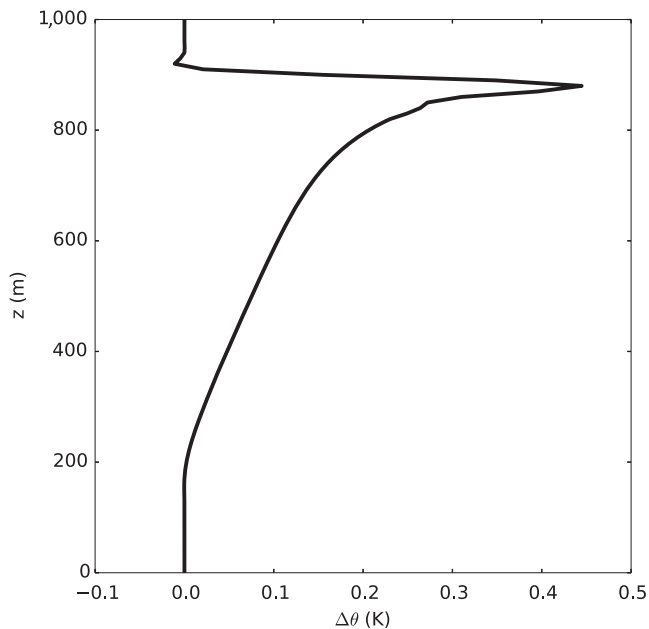
FIGURE 9 Potential temperature perturbation from background state of $\theta = 300$ K at $x = y = 0$ after 400 s for the 3D rising bubble

currently used at the Met Office. Nevertheless, a range of improvements and extensions to the model presented here are currently being developed:

1. The discretization presented here is valid for arbitrary quadrilateral-based meshes and future work will report on the extension of this model to spherical three-dimensional domains using quasi-uniform grids.
2. The computation of the vorticity in the $\mathbb{W}_1$ space (19) used here can introduce spurious oscillations. Although this has minimal effect on the results shown here, upwind-based schemes such as that used by Natale *et al.* (2016) are being investigated to improve on this aspect.
3. The method of lines scheme used to transport scalars imposes a time-step restriction on the model due to the CFL number constraint. Future work will investigate removing this constraint through using a flux-form semi-Lagrangian scheme for the scalars and possibly also for the velocity components.

ORCID

Thomas Melvin http://orcid.org/0000-0003-3933-0851

REFERENCES

Abdi, D.S. and Giraldo, F.X. (2016) Efficient construction of unified continuous and discontinuous Galerkin formulations for the 3D Euler equations. *Journal of Computational Physics*, 320, 46–68. https://doi.org/10.1016/j.jcp.2016.05.033.

Bochev, P.B. and Ridzal, D. (2010) Rehabilitation of the lowest-order Raviart–Thomas element on quadrilateral grids. *SIAM Journal on Numerical Analysis*, 47(1), 487–507.

Boffi, D., Brezzi, F. and Fortin, M. (2013) *Mixed Finite-Element Methods and Applications*. Springer, Berlin. https://doi.org/10.1007/978-3-642-36519-5.

Bott, R. and Raoul, L. (1982) *Differential Forms in Algebraic Topology*. Springer, New York, NY.

Brezzi, F. and Fortin, M. (1991) *Mixed and Hybrid Finite-Element Methods*. Springer, Berlin. https://doi.org/10.1007/978-1-4612-3172-1.

Cotter, C.J. and Shipton, J. (2012) Mixed finite elements for numerical weather prediction. *Journal of Computational Physics*, 231, 7076–7091.

Cotter, C.J. and Thuburn, J. (2014) A finite-element exterior calculus framework for the rotating shallow-water equations. *Journal of Computational Physics*, 257, 1506–1526.

Gottlieb, S. (2005) On high order strong stability preserving Runge–Kutta and multi step time discretizations. *Journal of Scientific Computing*, 25, 105–128.

Guerra, J.E. and Ullrich, P.A. (2016) A high-order staggered finite-element vertical discretization for non-hydrostatic atmospheric models. *Geoscientific Model Development*, 9(5), 2007–2029.

Kelly, J.F. and Giraldo, F.X. (2012) Continuous and discontinuous Galerkin methods for a scalable three-dimensional non-hydrostatic atmospheric model: limited-area mode. *Journal of Computational Physics*, 231(24), 7988–8008. https://doi.org/10.1016/j.jcp.2012.04.042.

Klemp, J.B., Skamarock, W.C. and Fuhrer, O. (2003) Numerical consistency of metric terms in terrain-following coordinates. *Monthly Weather Review*, 131, 1229–1239.

Lauritzen, P., Jablonowski, C., Taylor, M. and Nair, R. (Eds.) (2011) *Numerical Techniques for Global Atmospheric Models*. Lecture Notes in Computational Science and Engineering Vol. 80, Springer, Berlin.

Lawrence, B.N., Rezny, M., Budich, R., Bauer, P., Behrens, J., Carter, M., Deconinck, W., Ford, R., Maynard, C., Mullerworth, S., Osuna, C., Porter, A., Serradell, K., Valcke, S., Wedi, N. and Wilson, S. (2017) Crossing the chasm: how to develop weather and climate models for next generation computers? *Geoscientific Model Development Discussions*, 11, 1799–1821. https://doi.org/10.5194/gmd-2017-186.

Lock, S.J., Bitzer, H.W., Coals, A., Gadian, A. and Mobbs, S.D. (2012) Demonstration of a cut-cell representation of 3D orography for studies of atmospheric flows over very steep hills. *Monthly Weather Review*, 140(2), 411–424. https://doi.org/10.1175/MWR-D-11-00069.1.

Melvin, T., Benacchio, T., Thuburn, J. and Cotter, C.J. (2018) Choice of function spaces for thermodynamic variables in mixed finite-element methods. *Quarterly Journal of the Royal Meteorological Society*, 144, 900–916. https://doi.org/10.1002/qj.3268.

Melvin, T., Dubal, M., Wood, N., Staniforth, A. and Zerroukat, M. (2010) An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the non-hydrostatic vertical slice equations. *Quarterly Journal of the Royal Meteorological Society*, 136, 799–814.

Monk, P. (2003) *Finite-Element Methods for Maxwell's Equations*. Oxford University Press, Oxford, UK.

Natale, A., Shipton, J. and Cotter, C.J. (2016) Compatible finite-element spaces for geophysical fluid dynamics. *Dynamics and Statistics of the Climate System*, 1(1). https://doi.org/10.1093/climsys/dzw005.

Rognes, M.E., Kirby, R.C. and Logg, A. (2009) Efficient assembly of H(div) and H(curl) conforming finite elements. *SIAM Journal on Scientific Computing*, 31(6), 4130–4151.

Schär, C., Leuenberger, D., Fuhrer, O., Lüthi, D. and Girard, C. (2002) A new terrain-following vertical coordinate formulation for atmospheric prediction models. *Monthly Weather Review*, 130, 2459–2480. [Correction added on 21 June 2019, after first online publication. The reference "Schär *et al.* 2002" has been added in this version.]

Skamarock, W.C. and Klemp, J.B. (1994) Efficiency and accuracy of the Klemp–Wilhelmson time-splitting technique. *Monthly Weather Review*, 122, 2623–2630.

Staniforth, A. and Côté, J. (1991) Semi-Lagrangian integration schemes for atmospheric models – a review. *Monthly Weather Review*, 119, 2206–2223.

Staniforth, A. and Thuburn, J. (2012) Horizontal grids for global weather prediction and climate models: a review. *Quarterly Journal of the Royal Meteorological Society*, 138, 1–26.

Straka, J.M., Wilhelmson, R.B., Wicker, L.J., Anderson, J.R. and Droegemeier, K.K. (1993) Numerical solutions of a non-linear density current: a benchmark solution and comparisons. *International Journal for Numerical Methods in Fluids*, 17, 1–22.

Thuburn, J. and Cotter, C.J. (2015) A primal-dual mimetic finite-element scheme for the rotating shallow-water equations on polygonal spherical meshes. *Journal of Computational Physics*, 290(suppl C), 274–297. https://doi.org/10.1016/j.jcp.2015.02.045.

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J., Willett, M., Woolnough, S. and Xavier, P. (2017) The Met Office Unified Model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development*, 10, 1487–1520. https://doi.org/10.5194/gmd-10-1487-2017.

Williams, K.D., Harris, C.M., Bodas-Salcedo, A., Camp, J., Comer, R.E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S.F., Roberts, M.J., Rowell, D.P., Sanchez, C., Shelly, A., Sinha, B., Walters, D.N., West, A., Woollings, T. and Xavier, P.K. (2015) The Met Office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, 8(5), 1509–1524. https://doi.org/10.5194/gmd-8-1509-2015.

Wood, N., Staniforth, A., White, A.A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S.B., Zerroukat, M. and Thuburn, J. (2014) An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, 140, 1505–1520. https://doi.org/10.1002/qj.2235.

Yamazaki, H., Satomura, T. and Nikiforakis, N. (2016) Three-dimensional cut-cell modelling for high-resolution atmospheric simulations. *Quarterly Journal of the Royal Meteorological Society*, 142, 1335–1350.

APPENDICES

A. FINITE-ELEMENT SPACES

Consider a domain D partitioned into a mesh consisting of D_C cells, D_F faces, D_E edges and D_V vertices. Denoting the coordinates by (χ_1, χ_2, χ_3), where χ_1 and χ_2 are horizontal coordinates and χ_3 a vertical coordinate, the finite-element function spaces used here at order l for hexahedral elements correspond to:

$\mathbb{W}_0$ The Q_{l+1} space of scalar functions built from the tensor product of $P^{l+1}(\chi_1) P^{l+1}(\chi_2) P^{l+1}(\chi_3)$ polynomials of order $l + 1$ with full continuity between cells. The dimension of this space is $\dim(\mathbb{W}_0) = D_V + l D_E + l^2 D_F + l^3 D_C$.

$\mathbb{W}_1$ The Nédélec space of vector functions built from the tensor product of two P^{l+1} polynomials and one P^l polynomial with continuity between cells only in the tangential direction. The dimension of this space is $\dim(\mathbb{W}_1) = (l + 1) D_E + 2l(l + 1) D_F + 3l^2(l + 1) D_C$.

$\mathbb{W}_2$ The Raviart–Thomas space of vector functions built from the tensor product of one P^{l+1} polynomial and two P^l polynomials with continuity between cells only in the normal direction. The dimension of this space is $\dim(\mathbb{W}_2) = (l + 1)^2 D_F + 3l(l + 1)^2 D_C$.

$\mathbb{W}_3$ The Q_l^{DG} space of scalar functions built from the tensor product of $P^l(\chi_1) P^l(\chi_2) P^l(\chi_3)$ polynomials with no continuity between cells. The dimension of this space is $\dim(\mathbb{W}_3) = l^3 D_C$.

$\mathbb{W}_\theta$ The space of scalar functions $P^l(\chi_1) P^l(\chi_2) P^{l+1}(\chi_3)$ based on the vertical part of $\mathbb{W}_2$. These are discontinuous between cells in the horizontal directions but continuous between cells in the vertical direction. The dimension of this space is $\dim(\mathbb{W}_\theta) = (l + 1)^2 D_F^* + l(l + 1)^2 D_C$, where D_F^* is a subset of D_F containing only the faces in the χ_3 direction.

$\mathbb{W}_\chi$ The Q_m^{DG} space of scalar functions built from the tensor product of $P^m(\chi_1) P^m(\chi_2) P^m(\chi_3)$ polynomials with no continuity, where m may be different from l. The dimension of this space is $\dim(\mathbb{W}_\chi) = m^3 D_C$

B. BASIS FUNCTIONS

For hexahedral elements each basis function can be decomposed into the tensor product of three orthogonal polynomials, multiplied by a unit vector for the basis functions in the vector spaces $\mathbb{W}_1$ and $\mathbb{W}_2$. Two orders of polynomial functions are required in order to fully specify the basis functions; if the functions in $\mathbb{W}_3$ are order l, the two sets of polynomials can be denoted by

$$F_i(\eta) \equiv \prod_{\substack{j=0 \\ j \neq i}}^{l+1} \frac{\eta - \eta_j}{\eta_i - \eta_j}, \quad i = 0, \dots, l + 1, \qquad \text{(B1)}$$

and

$$G_i(\eta) \equiv \prod_{\substack{j=0 \\ j \neq i}}^{l} \frac{\eta - \eta_j}{\eta_i - \eta_j}, \quad i = 0, \dots, l. \qquad \text{(B2)}$$

These are the Lagrange interpolating polynomials that take the value 0 at $\eta = \eta_j, j \neq i$, and the value 1 at $\eta = \eta_i$. Here η denotes a generic coordinate. The polynomials F_i are of order $l + 1$ and those for G_i are of order l. For $l = 0$ equation B2

is the empty product giving $G_0(\eta) \equiv 1$. The locations of the basis nodal points η_j are evenly spaced in the computational coordinate η and include the endpoints, $\eta = 0, 1$. For constant functions, the centre point $\eta = 1/2$ is nominally used as the nodal point. Using Equations B1 and B2 then the basis functions, in the reference coordinates $\hat{\chi} \equiv (\hat{\chi}_1, \hat{\chi}_2, \hat{\chi}_3)$, for space $\mathbb{W}_0$ are given by

$$\gamma_{ijk}(\hat{\chi}) \equiv F_i(\hat{\chi}_1) F_j(\hat{\chi}_2) F_k(\hat{\chi}_3). \tag{B3}$$

Those in $\mathbb{W}_1$ are given by

$$\mathbf{c}_{ijk}(\hat{\chi}) \equiv \begin{cases} G_i(\hat{\chi}_1) F_j(\hat{\chi}_2) F_k(\hat{\chi}_3)\, \mathbf{i}, \\ F_i(\hat{\chi}_1) G_j(\hat{\chi}_2) F_k(\hat{\chi}_3)\, \mathbf{j}, \\ F_i(\hat{\chi}_1) F_j(\hat{\chi}_2) G_k(\hat{\chi}_3)\, \mathbf{k}. \end{cases} \tag{B4}$$

Those in $\mathbb{W}_2$ are given by

$$\mathbf{v}_{ijk}(\hat{\chi}) \equiv \begin{cases} F_i(\hat{\chi}_1) G_j(\hat{\chi}_2) G_k(\hat{\chi}_3)\, \mathbf{i}, \\ G_i(\hat{\chi}_1) F_j(\hat{\chi}_2) G_k(\hat{\chi}_3)\, \mathbf{j}, \\ G_i(\hat{\chi}_1) G_j(\hat{\chi}_2) F_k(\hat{\chi}_3)\, \mathbf{k}. \end{cases} \tag{B5}$$

Those in $\mathbb{W}_\theta$ are given by

$$w_{ijk}(\hat{\chi}) \equiv G_i(\hat{\chi}_1) G_j(\hat{\chi}_2) F_k(\hat{\chi}_3). \tag{B6}$$

Finally those in $\mathbb{W}_3$ are given by

$$\sigma_{ijk}(\hat{\chi}) \equiv G_i(\hat{\chi}_1) G_j(\hat{\chi}_2) G_k(\hat{\chi}_3). \tag{B7}$$

For the results presented here only the lowest-order elements, $l = 0$, are used. In this case, and assuming the basis nodal points for the linear functions F_i are at the extremities of the reference cell so that $\eta_0 = 0$ and $\eta_1 = 1$, Equations B1 and B2 give:

$$F_0(\eta) = 1 - \eta, \tag{B8}$$

$$F_1(\eta) = \eta, \tag{B9}$$

and

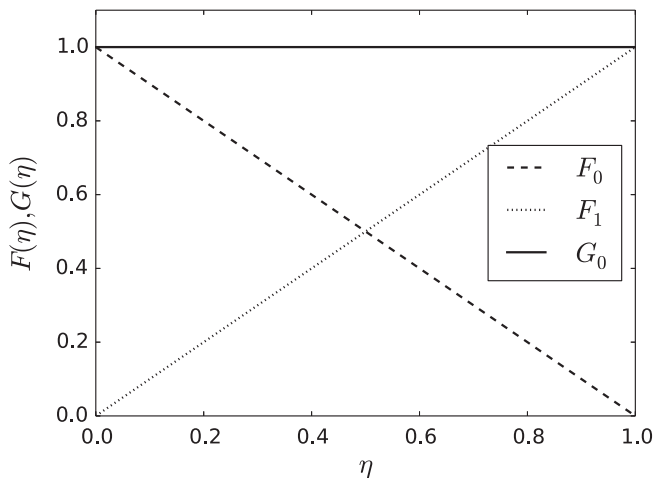$$G_0(\eta) = 1, \tag{B10}$$

and these are shown in Figure B1.

FIGURE B1 Basis functions needed for $l = 0$ order elements

C. EVALUATION OF DISCONTINUOUS FIELDS AT AND ACROSS CELL FACES

Consider a function φ and an interior face F shared by cells C^- and C^+. Let $\mathbf{n}_\mathsf{F}^-$ be the vector normal to the face F that is outward pointing for cell C^- and let $\mathbf{n}_\mathsf{F}^+$ be the corresponding outward-pointing vector normal to face F for cell C^+. The inner and outer traces of φ on the face F are denoted by φ_F^- and φ_F^+, respectively. These are defined to be:

$$\varphi_\mathsf{F}^\pm(\mathbf{x}, t) \equiv \lim_{\varepsilon \to 0} \varphi(\mathbf{x} \pm \varepsilon \mathbf{n}_\mathsf{F}^-, t). \tag{C1}$$

Then the face value of φ and the jump in φ across that face are defined, respectively, to be:

$$\{\varphi\}_\mathsf{F} \equiv \frac{1}{2}\left(\varphi_\mathsf{F}^+ + \varphi_\mathsf{F}^-\right), \tag{C2}$$

and

$$[\![\varphi]\!]_\mathsf{F} = \varphi_\mathsf{F}^- \mathbf{n}_\mathsf{F}^- + \varphi_\mathsf{F}^+ \mathbf{n}_\mathsf{F}^+. \tag{C3}$$

If the face F of cell C^- is an exterior face then define:

$$\{\varphi\}_\mathsf{F} = \varphi, \qquad [\![\varphi]\!]_\mathsf{F} \equiv \varphi \mathbf{n}_\mathsf{F}^-. \tag{C4}$$

If φ is a vector, then the multiplications in the definition of $[\![\varphi]\!]$ are dot products. Therefore, if φ is a scalar then $\{\varphi\}$ is also a scalar but $[\![\varphi]\!]$ is a vector, whilst if φ is a vector then $\{\varphi\}$ is also a vector but $[\![\varphi]\!]$ is a scalar.

D. EVALUATION OF THE JACOBIAN

With a linear coordinate space the coordinates within a cell C are given by

$$\begin{aligned} \chi_i = {}& \chi_i^{(1)} F_0(\hat{\chi}_1) F_0(\hat{\chi}_2) F_0(\hat{\chi}_3) \\ & + \chi_i^{(2)} F_1(\hat{\chi}_1) F_0(\hat{\chi}_2) F_0(\hat{\chi}_3) \\ & + \chi_i^{(3)} F_0(\hat{\chi}_1) F_1(\hat{\chi}_2) F_0(\hat{\chi}_3) \\ & + \chi_i^{(4)} F_1(\hat{\chi}_1) F_1(\hat{\chi}_2) F_0(\hat{\chi}_3) \\ & + \chi_i^{(5)} F_0(\hat{\chi}_1) F_0(\hat{\chi}_2) F_1(\hat{\chi}_3) \\ & + \chi_i^{(6)} F_1(\hat{\chi}_1) F_0(\hat{\chi}_2) F_1(\hat{\chi}_3) \\ & + \chi_i^{(7)} F_0(\hat{\chi}_1) F_1(\hat{\chi}_2) F_1(\hat{\chi}_3) \\ & + \chi_i^{(8)} F_1(\hat{\chi}_1) F_1(\hat{\chi}_2) F_1(\hat{\chi}_3), \end{aligned} \tag{D1}$$

for $i = 1, \dots, 3$. Here the mesh is further assumed to be uniform in the horizontal such that

$$\chi_1^{(1)} = \chi_1^{(3)} = \chi_1^{(5)} = \chi_1^{(7)} = \chi_1^-, \tag{D2}$$

$$\chi_1^{(2)} = \chi_1^{(4)} = \chi_1^{(6)} = \chi_1^{(8)} = \chi_1^+, \tag{D3}$$

$$\chi_2^{(1)} = \chi_2^{(2)} = \chi_2^{(5)} = \chi_2^{(6)} = \chi_2^-, \tag{D4}$$

$$\chi_2^{(3)} = \chi_2^{(4)} = \chi_2^{(7)} = \chi_2^{(8)} = \chi_2^+, \tag{D5}$$

and $\Delta\chi_1 \equiv \chi_1^+ - \chi_1^-$ and $\Delta\chi_2 \equiv \chi_2^+ - \chi_2^-$. Therefore the Jacobian can be simplified to

$$\mathbf{J} = \begin{bmatrix} \Delta\chi_1 & 0 & 0 \\ 0 & \Delta\chi_2 & 0 \\ \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \end{bmatrix}, \tag{D6}$$

where ε_i denotes a linear interpolating function in the $\widehat{\chi}_i$ direction of the change in the χ_3 coordinate, i.e.

$$
\begin{aligned}
\varepsilon_1 \equiv & \left[\left(\chi_3^{(2)}-\chi_3^{(1)}\right)\left(1-\widehat{\chi_2}\right)+\left(\chi_3^{(4)}-\chi_3^{(3)}\right)\widehat{\chi_2}\right]\left(1-\widehat{\chi_3}\right) \\
& + \left[\left(\chi_3^{(6)}-\chi_3^{(5)}\right)\left(1-\widehat{\chi_2}\right)+\left(\chi_3^{(8)}-\chi_3^{(7)}\right)\widehat{\chi_2}\right]\widehat{\chi_3}, \quad (D7)
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_2 \equiv & \left[\left(\chi_3^{(3)}-\chi_3^{(1)}\right)\left(1-\widehat{\chi_1}\right)+\left(\chi_3^{(4)}-\chi_3^{(2)}\right)\widehat{\chi_1}\right]\left(1-\widehat{\chi_3}\right) \\
& + \left[\left(\chi_3^{(7)}-\chi_3^{(5)}\right)\left(1-\widehat{\chi_1}\right)+\left(\chi_3^{(8)}-\chi_3^{(6)}\right)\widehat{\chi_1}\right]\widehat{\chi_3}, \quad (D8)
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_3 \equiv & \left[\left(\chi_3^{(5)}-\chi_3^{(1)}\right)\left(1-\widehat{\chi_1}\right)+\left(\chi_3^{(6)}-\chi_3^{(2)}\right)\widehat{\chi_1}\right]\left(1-\widehat{\chi_2}\right) \\
& + \left[\left(\chi_3^{(7)}-\chi_3^{(3)}\right)\left(1-\widehat{\chi_1}\right)+\left(\chi_3^{(8)}-\chi_3^{(4)}\right)\widehat{\chi_1}\right]\widehat{\chi_2}. \quad (D9)
\end{aligned}
$$

E. FINITE-ELEMENT ADVECTION OF THE REFERENCE POTENTIAL TEMPERATURE

The linear approximation $\mathcal{L}$ to the Jacobian $\mathcal{J}$ requires the evaluation of a finite-element estimate of the advection of the reference potential temperature field θ^* by the increment to the wind, $\mathbf{u}'$, specifically $\langle w, \mathbf{u}' \cdot \nabla \theta^* \rangle$. However, since $\theta^* \in \mathbb{W}_\theta$ is horizontally discontinuous, this term has to be integrated by parts. Following the same procedure as used to derive equation 18 together with the definitions of Appendix C, the result is:

$$
\begin{aligned}
\langle w, \mathbf{u}' \cdot \nabla \theta^* \rangle &= \sum_C \left\langle w\mathbf{u}', \nabla \theta^* \right\rangle_C \\
&= \sum_F \left\langle\!\!\left\langle [\![w\mathbf{u}']\!]_F, \{\theta^*\}_F \right\rangle\!\!\right\rangle_F \\
&\quad - \sum_C \left\langle \nabla_C \cdot (w\mathbf{u}'), \theta^* \right\rangle_C \\
&= \left\langle\!\!\left\langle [\![w\mathbf{u}']\!], \{\theta^*\} \right\rangle\!\!\right\rangle - \left\langle \nabla_C \cdot (w\mathbf{u}'), \theta^* \right\rangle \quad (E1)
\end{aligned}
$$

Transforming this expression to use the reference cell gives

$$
\langle w, \mathbf{u}' \cdot \nabla \theta^* \rangle = \left\langle\!\!\left\langle [\![\widehat{w}\widehat{\mathbf{u}}']\!], \left\{\widehat{\theta}^*\right\} \right\rangle\!\!\right\rangle - \left\langle \widehat{\nabla}_C \cdot (\widehat{w}\widehat{\mathbf{u}}'), \widehat{\theta}^* \right\rangle, \quad (E2)
$$

where $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ now denotes the surface integrals over the collection of all cell faces evaluated using the reference cell and $[\![\cdot]\!]$ is defined in terms of normal vectors defined for the reference cell. The form for $P_{\theta 2}^{\theta^*}$ given by Equation 87 follows from this expression.