Received: 08 August 2018 | Revised: 29 January 2019 | Accepted: 12 March 2019

DOI: 10.1002/qj.3532

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

Comparison of the Moist Parcel-in-Cell (MPIC) model with large-eddy simulation for an idealized cloud

Steven J. Böing[1] | David G. Dritschel[2] | Douglas J. Parker[1] | Alan M. Blyth[1,3]

[1]School of Earth and Environment, University of Leeds, Leeds, UK

[2]Mathematical Institute, University of St Andrews, St Andrews, UK

[3]National Centre for Atmospheric Science, University of Leeds, Leeds, UK

Correspondence
Steven J. Böing, School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK.
Email: s.boeing@leeds.ac.uk

Funding information
EPSRC (grant EP/M008525/1), NERC (grant NE/N013840/1), Met Office Academic Partnership, Royal Society, NSF (grant PHY17-48958)

Abstract

The ascent of a moist thermal is used to test a recently developed essentially Lagrangian model for simulating moist convection. In this Moist-Parcel-In-Cell (MPIC) model, a number of parcels are used to represent the flow in each grid cell. This has the advantage that the parcels provide an efficient and explicit representation of subgrid-scale flow. The model is compared against Eulerian large-eddy simulations with a version of the Met Office NERC Cloud model (MONC) which solves the same equations in a more traditional Eulerian scheme. Both models perform the same idealized simulation of the effects of latent heat release and evaporation, rather than a specific atmospheric regime.

Dynamical features evolve similarly throughout the development of the thermal using the two approaches. Subgrid-scale properties of small-scale eddies captured by the MPIC model can be explicitly reconstructed on a finer grid. MPIC simulations thus resolve smaller features when using the same grid spacing as MONC, which is useful for detailed studies of turbulence in clouds.

The convergence of bulk properties is also used to compare the two models. Most of these properties converge rapidly, though the probability distribution function of liquid water converges only slowly with grid resolution in MPIC. This may imply that the current implementation of the parcel mixing mechanism underestimates small-scale mixing.

Finally, it is shown how Lagrangian parcels can be used to study the origin of cloud air in a consistent manner in MPIC.

KEYWORDS

clouds, convection, numerical method, thermals

1 | INTRODUCTION

Detailed studies of moist convection continue to play an important role in the development of weather and climate models. Such studies were initially performed using two-dimensional slab-symmetric or axisymmetric models of clouds (e.g. Ogura, 1963). However, over the past 40 years it has become possible to run three-dimensional large-eddy simulation (LES) models, which resolve the internal dynamics of clouds and the boundary layer, on increasingly large domains.

LES studies form a bedrock of atmospheric research: they have been used to investigate the fine-scale dynamics

of mixing between clouds and their environment (e.g. Grabowski and Clark, 1993, Heus *et al.*, 2008), but also to aid the development of convective parametrizations (e.g. Siebesma and Cuijpers 1995) and investigate shortcomings of convection-permitting models (Hohenegger *et al.*, 2015; Panosetti *et al.*, 2016). Nevertheless, questions around fine-scale mixing remain hard to tackle because of uncertainty about its representation even in LES. Moreover, some research questions ask for a Lagrangian perspective: Lagrangian particles can be used in an LES model, but the representation of mixing in the model core is not fully consistent with the way in which particles are treated.

Existing LES models and convection-permitting models use a grid-based dynamical core, in which advection can be performed using either Eulerian or semi-Lagrangian methods. In Eulerian dynamical cores, the governing equations are discretized onto an underlying mesh, and advection is performed using the discretized equations on this mesh. Semi-Lagrangian methods, on the other hand, use so-called departure points to perform advection. Within each time step, the flow from these departure points is tracked, and interpolation is used to construct the field after advection. These methods permit the use of larger time steps at the same grid resolution, but tend to violate conservation and poorly preserve correlations between tracers (Lauritzen and Thuburn, 2012). Recent work on semi-Lagrangian methods has focussed on addressing conservation issues and maintaining tracer correlations (e.g. Zerroukat *et al.*, 2002; Kaas, 2008; Aranami *et al.*, 2015).

Both types of grid-based approach introduce a degree of numerical mixing: Eulerian methods because of truncation of the equations and because of the finite size of each grid box, and semi-Lagrangian methods because of the truncation and the interpolation that is performed each time step. Numerical mixing is particularly relevant for moist convection for two reasons:

1. The equation of state of moist air contains a discontinuity between the saturated and the unsaturated regime, which makes the problem highly nonlinear. Surprisingly, a mixture of a dry and a moist parcel can become denser than either of the two constituent parcels after its pressure adjusts to the environment (Paluch, 1979). The reason for this is the occurrence of evaporation. Numerical mixing and poor correlation between tracers are a cause for concern when representing such phase transitions.
2. Formation of precipitation is also highly dependent on the highest values of liquid water that occur, again in a nonlinear way. These values in turn depend on the degree of mixing in the interior of the cloud (Twomey, 1966; Blyth *et al.*, 2005; Cooper *et al.*, 2013).

LESs of both cumulus (Matheou, 2011; Pressel *et al.*, 2015) and stratocumulus (Stevens and Bretherton, 1999; Stevens *et al.*, 2005; Pressel *et al.*, 2015) clouds show substantial sensitivities to resolution and numerical method. Both types of moist convection represent a particular challenge: for stratocumulus convection, the sharp interface at the top of the cloud layer needs to be well represented. For cumulus convection, it is important to capture regions with high liquid water content where rain formation takes place, which can have a scale well below 100 m (Blyth *et al.*, 2005), as well as the dynamics of turbulent eddies and entrainment.

The aim of the current work is to test the performance of a recently developed Lagrangian model for detailed studies of moist convection. This Moist Parcel-In-Cell (MPIC) model is designed to provide a higher effective resolution in studies where convection is marginally resolved, as well as to provide a more reliable, physically based approach to mixing. It differs from approaches commonly used in atmospheric modelling in that it represents the atmosphere through a set of parcels, rather than through a mesh-based approach. A companion paper (Dritschel *et al.* 2018, hereafter D18) introduced MPIC and tested its internal consistency. Parcel-based models like MPIC have previously been used for various fluid dynamical problems (see the introduction of D18). MPIC also differs from most of these parcel-based approaches as it accounts for parcel buoyancy and has an explicit description of stretching of parcels at small scales.

As we expect the MPIC approach to have major benefits both in terms of computational cost and in accurately representing the physics of moist convection, we perform a more detailed comparison with an LES model here. As a reference model, we use the Met Office NERC Cloud (MONC) model (Brown *et al.*, 2015). For both models, a range of simulations with different resolutions is considered, including those where convection is marginally resolved, as well as simulations which include fine-scale turbulence deep into the inertial range.

For moist convection, a reliable description of stratification and the mixing of heat and moisture is important. D18 describe the way in which buoyancy and mixing are represented in the MPIC framework. We argue that this approach has advantages for modelling cumulus convection, although further improvements may be possible. Both the development of mean properties of the cloud as well as those of turbulence and dilution on small scales are considered. The convergence of these properties across a range of resolutions is also tested.

The paper is outlined as follows. Section 2 describes our methodology and includes a short description of both MPIC and the LES model used, MONC. Section 3 contains our main results. In this section, both the volume-integrated (bulk) properties of the flow, as well as its detailed turbulent evolution, are compared to MONC. Section 4 discusses the advantage of Lagrangian diagnostics, and Section 5 summarizes our findings.

2 | METHODOLOGY

The test case consists of an idealized single cloud with a simplified approach to evaporation and condensation. D18 describe this case in detail. Below, we mention some key properties of the code and the test problem set-up considered.

2.1 | Governing equations

We use a simplified set of governing equations introduced in D18. The governing equations for velocity $\mathbf{u}$, liquid water buoyancy b_l and specific humidity q are

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\frac{\boldsymbol{\nabla} p}{\rho_0} + b\widehat{\mathbf{e}}_z, \tag{1}$$

$$\frac{\mathrm{D}b_l}{\mathrm{D}t} = 0, \tag{2}$$

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \tag{3}$$

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0, \tag{4}$$

where $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \mathbf{u} \cdot \boldsymbol{\nabla}$ represents a material derivative. In Equation 1 the total buoyancy b (including the effects of latent heating) is approximated by

$$b = b_l + \frac{gL}{c_p \theta_{l0}} q_l \,, \tag{5}$$

where

$$q_l = \max\left(0, q - q_0 \mathrm{e}^{-\lambda z}\right) \tag{6}$$

is the liquid water content. The pressure p in Equation 1 excludes the part due to the hydrostatic background state of constant density ρ_0. The other symbols appearing in Equations 1–6 are the vertical unit vector $\widehat{\mathbf{e}}_z$, the gravitational acceleration g, the latent heat of condensation L, the specific heat at constant pressure c_p, the surface saturation humidity q_0, and the inverse condensation scale height λ. The liquid water buoyancy is defined by $b_l = g(\theta_l - \theta_{l0})/\theta_{l0}$, where θ_l is the liquid water potential temperature and θ_{l0} is a constant reference value.

2.2 | Parcel-based model

As discussed in D18, the parcels in the MPIC model carry liquid water buoyancy b_l, specific humidity q, parcel volume V and vorticity $\boldsymbol{\omega}$. The prognostic equation for vorticity reads

$$\frac{\mathrm{D}\boldsymbol{\omega}}{\mathrm{D}t} = \boldsymbol{\omega} \cdot \boldsymbol{\nabla}\mathbf{u} + \left(b_y, -b_x, 0\right), \tag{7}$$

where subscripts on b denote spatial partial differentiation. The velocity field is needed both to evolve the positions of the parcels as well as to calculate the right-hand side of the vorticity equation. D18 discuss the use of tri-linear interpolation on a grid to obtain a Poisson equation for the gridded velocity field given the gridded vorticity field. The different scale of the advecting velocity field compared to the particle size resembles a similar effective distinction made in layer-wise two-dimensional contour advection methods (Dritschel and Ambaum, 1997; Fontane and Dritschel, 2009).

Particle-mesh methods typically treat fluid elements as indivisible. This is not a suitable approach for the simulation of clouds, as mixing is needed on longer time-scales in order to represent the eventual evaporation of the cloud, which occurs at the end of the process of turbulent mixing. MPIC contains an explicit representation of the stretching and eventual mixing of parcels, detailed in D18. The amount of stretching that a parcel has experienced is prognostically calculated, and when this exceeds a critical threshold, a parcel is split. The choice of threshold value is also related to the assumed distance between parcels after splitting (D18). When this splitting leads to the parcel becoming smaller than a critical volume (here $1/216$ of the grid box volume), the parcel properties are merged with those of surrounding parcels through operations on the grid, using the conservative method developed in D18. D18 considered the sensitivity to the threshold; they found limited sensitivity over a wide range of values.

Another key numerical advantage of our approach is that buoyancy in Equation 5 is calculated at the level of each individual parcel, before it is summed to grid-based values. The order of these operations implies that nonlinearities in the thermodynamics are retained (Tsang and Vallis, 2018).

2.3 | Non-dimensionalization

As discussed in D18, the equations are non-dimensionalized by setting the length-scale $1/\lambda = 1$ in Equation 5. The characteristic squared buoyancy frequency $g\lambda\Delta\theta_{l0}/\theta_{l0} = 1$ is used to non-dimensionalize time. Here, $\Delta\theta_{l0}/\theta_{l0} = 0.01$ is a characteristic fractional variation of the liquid water potential temperature. This gives a dimensionless gravity of $g = 100$. The specific humidity q is scaled by its saturation value q_0 at ground level (i.e. we use $\widetilde{q} = q/q_0$ in what follows). We obtain the following dimensionless expression for the buoyancy b:

$$b = b_l + b_{\mathrm{m}} \max\left(0, \widetilde{q} - \mathrm{e}^{-z}\right) \,, \tag{8}$$

where

$$b_{\mathrm{m}} = \frac{gLq_0}{c_p \theta_{l0}} \,. \tag{9}$$

We use $L/c_p = 2,500\,\mathrm{K}$, $q_0 = 0.015$ and $\theta_{l0} = 300\,\mathrm{K}$, which gives $b_{\mathrm{m}} = 12.5$. Both MONC and MPIC are run using the same dimensionless set of equations, which make the incompressible Boussinesq approximation. MONC can also be used with a reference density which varies with height,

that is, an anelastic formulation; this option is not used in the present work. A rough estimate of the length-scale λ in the atmosphere is that it is of the order 2 km, which would make the domain 12 km high and the grid spacing in the 256^3 simulation about 50 m. In this case, some of the other assumptions we have made such as a lack of precipitation and a Boussinesq approach are not valid. The boundary layer would also be much deeper than is typical for an atmospheric case.

Alternatively, we can think of our simulation as a more shallow cloud, but with stronger saturation specific humidity dependence on temperature than in the atmospheric case. This means convection will be more vigorous than in atmospheric shallow convection.

Possibly, scaling all heights that define the case by a factor of about ⅓ to ½ with respect to λ would have been a good choice, as this would make them characteristic of shallow convection over land. The main reason we chose the parameters in the way we have done is that these allow us to test the dynamics with a strong nonlinearity across a range of resolutions. In summary, the set-up can be thought of as a thermal which releases a substantial amount of latent heat, rather than a realistic cloud in a particular atmospheric regime.

2.4 | Case description

The different models are compared using a moist buoyant thermal which is initially at rest and located near the surface. The thermal first rises through a neutrally stable lower atmospheric layer before encountering a layer with a constant stratification aloft. In order to prevent the flow from being overly symmetric, the buoyancy field in the spherical interior of the thermal is initialized as:

$$b_l = b_{l\mathrm{th}} \left(1 + \frac{e_1 x' y' + e_2 x' z' + e_3 y' z'}{R^2} \right). \qquad (10)$$

Here, $b_{l\mathrm{th}}$ is the mean thermal liquid water buoyancy, and x', y' and z' denote the position with respect to the thermal centre and R the thermal radius, while e_1, e_2 and e_3 are dimensionless parameters. These parameters are chosen as $R = 0.8$, $e_1 = 0.3$, $e_2 = -0.4$ and $e_3 = 0.5$. The top of the neutrally stratified layer is located at $z \approx 2.38$ and the dimensionless Brunt–Väisälä frequency above is $N \approx 0.97$. The environment relative humidity is set to 80% at all levels in the stratified layer, and the specific humidity is constant in the neutrally stratified layer, with a value chosen such that there is no discontinuity at the top of this layer. This value is also 90% of the specific humidity in the thermal.

A full overview of the procedure and parameters used to construct the parcel properties is given in D18. In MONC, the same procedure is used, but the specific humidity and buoyancy are initialized at the relevant locations on the (staggered) grid, rather than on parcels.

The reason for using a warm and moist thermal is that the detailed flow evolution can be compared across models and resolutions not only on a statistical basis, but also directly. Recently, we have made a number of changes to the solver which enable us to use a non-zero buoyancy gradient at the boundary. This will help us to implement boundary conditions that are suitable for real-case atmospheric simulations in future work.

The moist thermal has a *sharp* edge, that is, there is a discontinuity of buoyancy at the edge. This has the drawback that it is less realistic and leads to vorticity values which depend on resolution. In general, our results are likely to be more sensitive to resolution than simulations with a smoother initial thermal edge. However, the advantages of this approach are that it makes it easier to track the parcels that are defined as the initial thermal, and it accelerates the transition from rest to a fully turbulent flow.

2.5 | MONC simulations

MONC is a LES model that has been developed for research on atmospheric boundary layers and clouds. The model formulation of MONC follows that of its predecessor, the Met Office Large-Eddy Model, which has been used for many years for atmospheric process studies and LES intercomparison studies (e.g. Petch and Gray, 2001; Abel and Shipway, 2007). However, the code has been completely rewritten for use on modern parallel computing architectures (Brown *et al.*, 2015).

In the MONC simulations, the same idealized thermodynamics are used as in the MPIC simulations and we have adapted the model formulation and set-up (domain size and thermodynamics) to match that of MPIC. The model integrates prognostic equations for the different components of the momentum equation and the scalars b_l and q on a grid with Arakawa C-staggering. For scalars, it uses the positivity-preserving ULTIMATE advection scheme (Leonard *et al.*, 1993). For the advection of momentum and the subfilter-scale fluxes, we consider two different approaches in this work.

1. MONC-Smagorinsky: Velocity components are advected using a second-order kinetic energy conserving scheme (Piacsek and Williams, 1970). Hence there is a need for subgrid-scale dissipation: subfilter-scale fluxes of scalar quantities and momentum are determined using the Smagorinsky approach (Smagorinsky, 1963). Effects of stratification are taken into account in the determination of the eddy viscosity, as they are in the default version of MONC.

The Richardson number is used in calculating the eddy viscosity, as described in Mason and Brown (1999). For

unsaturated air, we define the Richardson number as $Ri = (\partial b_l/\partial z)/S$. Here, S is the modulus of the rate of strain tensor. For saturated air, a correction is made for latent heat release: $Ri = (\partial b_l/\partial z + b_m e^{-z})/S$.

2. MONC-Implicit: velocity components are advected using the ULTIMATE advection scheme, and no explicit subgrid-scale model is used. This approach is referred to as implicit LES.

The pressure is determined diagnostically using a Poisson equation, which is solved using Fast Fourier Transforms in the horizontal and a tridiagonal solver in the vertical. MONC was run with a constant time step Δ_t, proportional to the grid spacing Δ_x, namely $\Delta_t = 0.2\Delta_x/\pi$. We have used a constant time step in MONC as this was the easiest way to obtain output at (or near) the exact time when it was required with MONC's IO-server.

We take the grid spacing to be the same in x, y and z. Horizontal boundary conditions are doubly periodic. Free-slip boundary conditions are used at both domain bottom and domain top for horizontal momentum, whereas for temperature and moisture, the flux at the boundaries is set to zero.

2.6 | Use of subgrid-scale diagnostics for MPIC

For MPIC, we mostly show results that have been obtained by post-processing the particle information using a finer grid. MPIC uses a large number of parcels per grid box (typically 10–200), each of which occupies only a small fraction of the grid box volume. This means that gridded quantities fail to show the actual variability below grid scale that is present in the model. In order to visualize this subgrid variability, a projection algorithm is employed that works with a representative radius for each parcel. A full description of the projection algorithm can be found in Appendix A.

For computing statistics of the liquid water field in MPIC, we make use of the direct statistics of parcels. Appendix B discusses how the probability distribution function of liquid water depends on the method used to compute it.

2.7 | Computational cost

In the following, the number of grid points is mentioned in comparisons between MPIC and MONC. However, a like-for-like comparison is not easy to infer, since MPIC exploits the existence of many parcels within each grid box. Although the computational cost of the solver is largely determined by the number of grid points, the total computational cost is harder to compare.

MPIC was run using shared-memory parallelism (OpenMP) for this study. This is the reason that the maximum number of grid points used was limited to 384^3, whereas for MONC, which uses MPI parallelism, 1024^3 points could be used. In the meantime, a project to produce a hybrid parallel (OpenMP+MPI) version of MPIC in collaboration with the Edinburgh Parallel Computing Centre, which uses MONC's infrastructure, is close to being finished.

The number of parcels plays an important role in the cost of MPIC simulations, even though parcel operations are relatively straightforward. Parcel operations (including interpolation of values to and from the grid) are responsible for the bulk of the cost of the MPIC simulations (79% for parcel operations and 9% for parcel time stepping in a 128^3 grid points simulation). The FFT operations are responsible for about 10% of the computational cost.

The simulation with 256^3 grid points took approximately 99 core hours (the number of hours spent multiplied by the number of computational units involved) on ARCHER for MONC and 751 core hours on a local machine for MPIC using 24 cores. The OpenMP-only parallelism does not scale perfectly: as an example, for a 128^3 grid point simulation on eight cores the speed-up was a factor 5.35 compared to a single core, whereas using 24 cores only resulted in a speed-up of a factor 8.17. Early results from the hybrid parallel version on ARCHER show a total runtime of 453 core hours on 128 cores, with 17% of this time spent on FFT operations. The latter is a greater fraction of computational cost, but the total time spent on FFTs is similar to that in the version with shared-memory parallelism only. We aim to further optimize the relative cost of parcel and grid operations in future work.

The number of time steps for the computation with 256^3 grid points is 1667 for MPIC and 6400 for MONC. However, MPIC uses an RK4 time-integration scheme, whereas MONC uses a leap-frog approach with an Asselin filter. In MPIC, the solver is called four times per time step, so effectively the number of calls to the solver is similar. We have followed a relatively conservative approach in choosing the time step for both models. The time step length in MPIC mainly changes during the initial phase, and the smallest time step of 0.0039 time units occurs at $t = 5.92$.

3 | RESULTS

3.1 | Evolution of the flow at high resolution

Figure 1 shows the evolution of the liquid water specific humidity field q_l in the MONC and MPIC simulations at the highest resolutions at which we have run the models. The number of grid points used in the simulations is 1024^3 for MONC and 384^3 for MPIC (due to the limitations of shared-memory parallelism). The number of grid points is different as these simulations serve as both a benchmark for the lower-resolution simulations and as a test that both models give similar results as resolution increases. However, it is

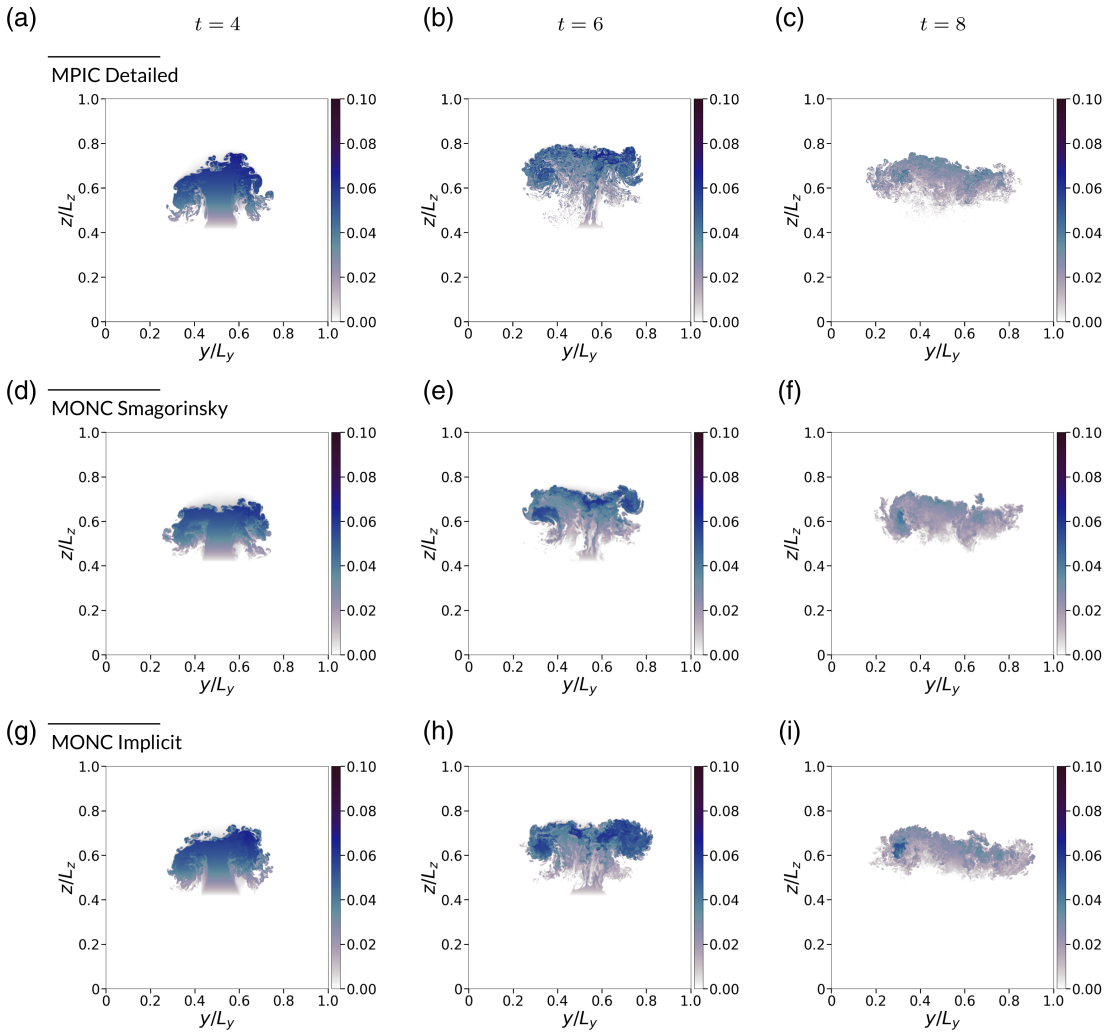

FIGURE 1 Cross-sections of dimensionless liquid water specific humidity q_l, for (a, b, c) MPIC, and for MONC with (d, e, f) a Smagorinsky and (g, h, i) an implicit subgrid formulation after (a, d, g) 4, (b, e, h) 6 and (c, f, i) 8 time units

important to keep in mind that, for producing Figure 1 and the detailed cross-sections shown in Figure 2, the MONC simulations had a much higher total computational cost. The number of parcels in the MPIC simulation is larger than the number of grid points by a factor of 8 at the start of the simulation, which increases to a factor of 16 at the end. The MPIC results are rendered using the algorithm described in Appendix A.

The large-scale evolution of the thermal is similar between the models, and even some of the smaller-scale features that arise due to the initial asymmetry are comparable at $t = 8$. Closer inspection also shows that, at $t = 6$, the thermal in MPIC has ascended further than the one in MONC.

MPIC shows a very large amount of fine-scale structure at the reference resolution, whereas the liquid water field looks relatively smooth in MONC, particularly for the Smagorinsky model. This is even more evident in the detailed cross-sections shown in Figure 2. In this figure, the MPIC results are also shown averaged to the coarse grid used in the dynamics. Some of the detail in the liquid water field (left column) is lost when only the grid-scale properties are used for visualization.

Larger differences between the reference simulations occur in the vorticity and velocity fields. The magnitude of the vorticity vector is shown in Figure 2b,e,h,k. MPIC shows a sharply confined region of high vorticity with a large amount of structure (in particular, it includes long swirling features), whereas in MONC the filaments are broken into multiple segments and are less confined. We will address the extent to which this depends on resolution below.

In MPIC, vorticity is a dynamical variable carried by the parcels, and we can therefore make use of the detailed projection algorithm for visualization. The magnitude of vorticity is larger at the subgrid scale, as partial cancellation between parcels with different orientations of the vorticity vector occurs on the grid scale. For MONC, the vorticity is calculated from the gradients in the grid-scale velocity fields. Conversely, in MONC the velocity fields evolve prognostically at the grid scale, whereas in MPIC these only exist at the scale of the interpolation grid. In principle, more detailed velocity fields can be constructed from the parcel vorticity by using a finer interpolation grid in the solver as compared

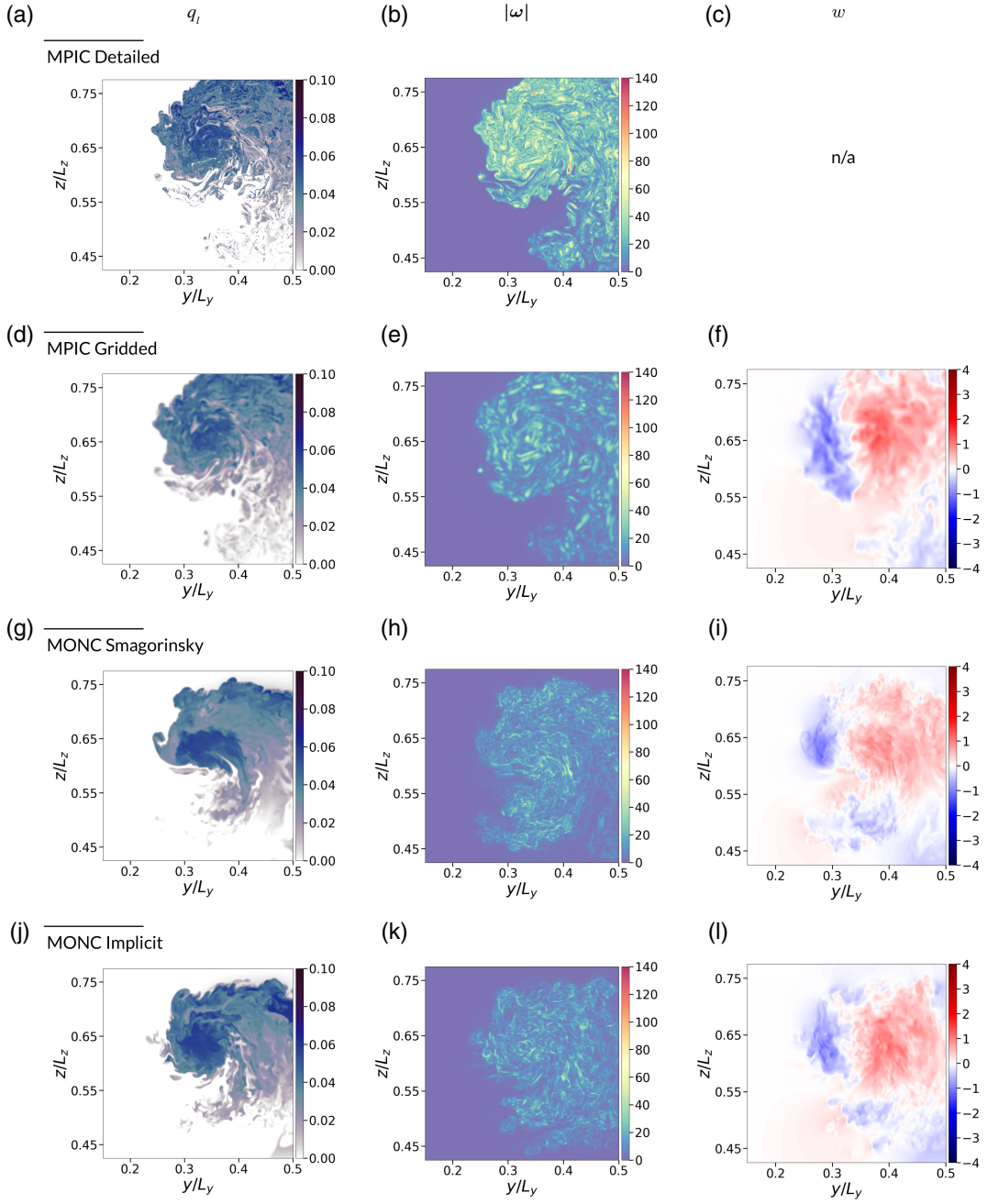

FIGURE 2 Detailed zoom into (a, d, g, j) the liquid water specific humidity q_l, (b, e, h, k) vorticity and (c, f, i, l) vertical velocity fields in the most active part of convection at $t = 6$. For MPIC, both results using (a, b, c) the method in Appendix A and (d, e, f) at the scale of the grid are shown. MONC uses either (g, h, i) the Smagorinsky subgrid scheme or (j, k, l) an implicit LES formulation

to the vorticity tendency equations, but this is currently not done (as parcels are advected with the gridded velocity fields in any case).

The vertical velocity shown in Figure 2f,i,l shows less structure in the MPIC fields than in the MONC ones. This is to be expected as this variable only exists on the grid, which is three times coarser for MPIC than for MONC.

In conclusion, the large-scale evolution of the thermal is similar in the two reference simulations. MPIC contains more small-scale features in the liquid water and vorticity fields, though the velocity field at the grid scale shows more fine-scale structure in MONC, which uses a larger number of grid points.

3.2 | Resolution sensitivity of the flow evolution

The flow evolution in MONC and MPIC across a range of resolutions is compared next. The stage when convection is at its most active ($t = 6$) is considered first. The liquid water and buoyancy fields at this stage are shown in Figures 3 and 4, respectively. These figures show results

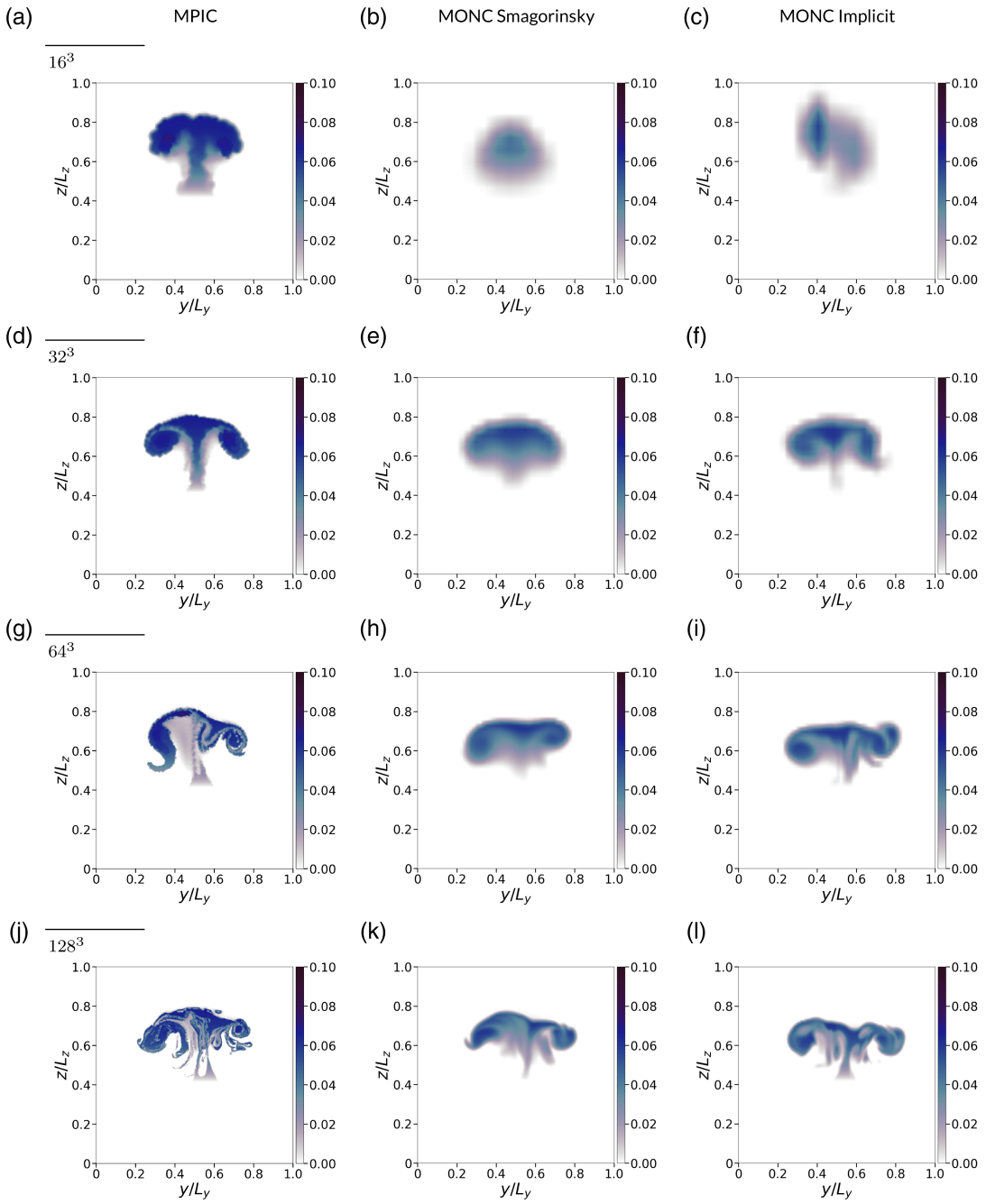

FIGURE 3 Cross-sections of liquid water specific humidity at $t = 6$ using (a, b, c) 16^3 to (j, k, l) 128^3 grid points, for (a, d, g, j) MPIC, and for MONC with (b, e, h, k) Smagorinsky and (c, f, i, l) implicit subgrid formulation

with the same grid spacing; admittedly, the computational cost at this grid spacing is currently higher for MPIC than for MONC.

Two differences between the MONC and MPIC simulations are clear. First, the coarsest MONC simulations are overly diffusive as compared to the high-resolution reference simulations. Intermediate MONC simulations preserve the maxima in the liquid water field better. MPIC, on the other hand, may be underdiffusive when the flow is marginally resolved, as can be seen by comparing the buoyancy fields in a low- and an intermediate-resolution simulation (Figure 4); a significant local maximum in buoyancy occurs in the vortex ring in the low-resolution MPIC simulations, whereas this maximum is much less pronounced at higher resolutions.

The differences between MPIC and MONC are more readily apparent at later times, when diffusive effects accumulate in MONC. This is shown in Figure 5, for the condensed portion of the specific humidity (comparing times $t = 4$ with $t = 8$). The MPIC model exhibits many more fine-scale features reminiscent of actual cumulus clouds than MONC does. By $t = 8$, the cloud spreads laterally as it has lost buoyancy

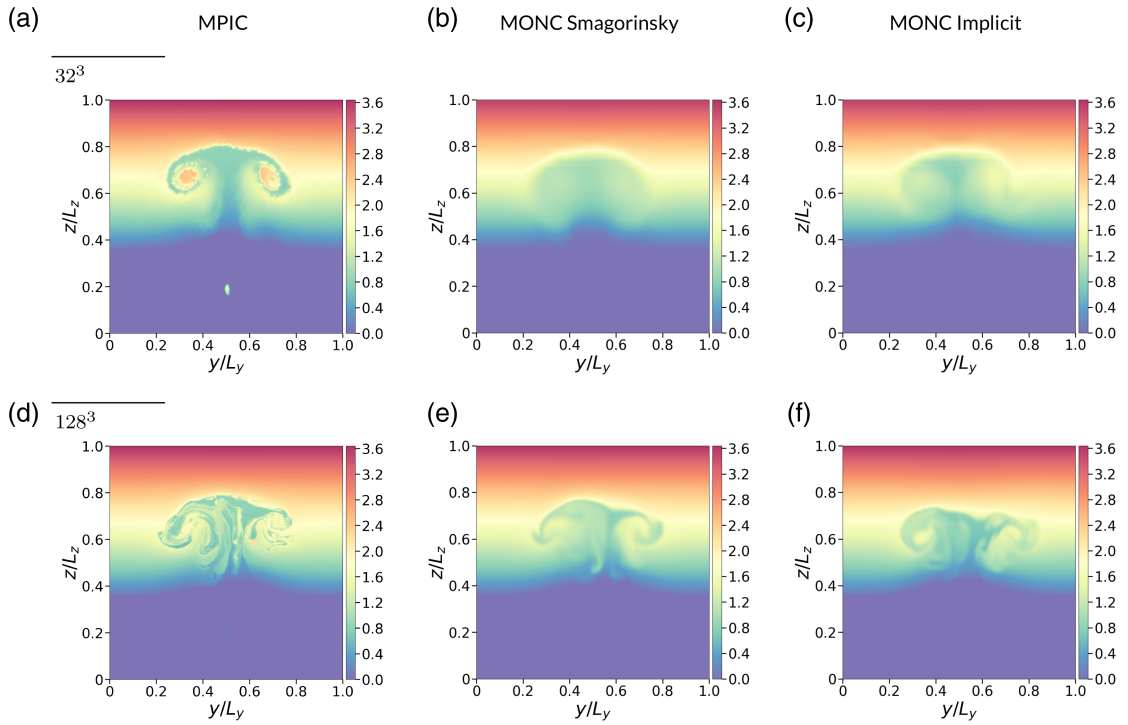

FIGURE 4 Cross-sections of buoyancy after 6 time units for simulations using (a, b, c) 32^3 and (d, e, f) 128^3 grid points. Columns are as in Figure 3

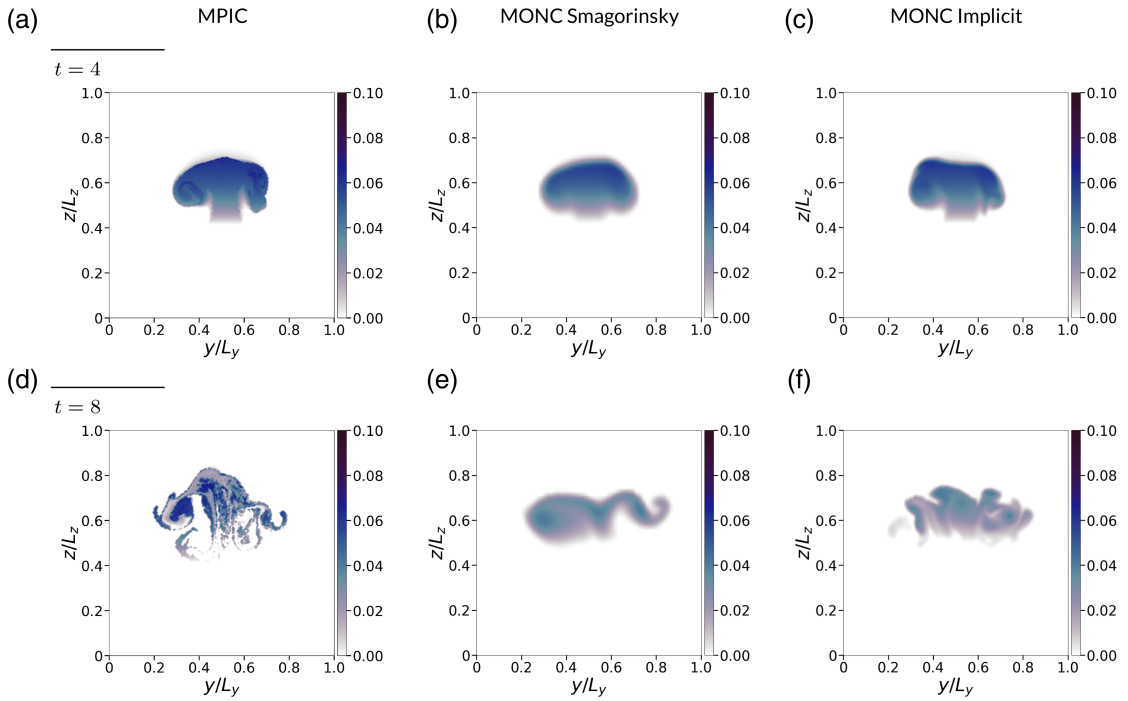

FIGURE 5 Time evolution of liquid water specific humidity in the simulations using 64^3 grid points after (a, b, c) 4 time units and (d, e, f) 8 time units. Columns are as in Figure 3

and gradually evaporates. Turbulent motions are better represented in the MPIC model than in MONC, and are sustained in time.

Both models capture the same intermediate and large-scale features when used across a range of resolutions. However, the small-scale features differ, and vary strongly with resolution. This is expected since the vorticity is generated by gradients in buoyancy, and so the finer the resolution used, the higher the gradients become and the more intense and localized the vorticity generation becomes.

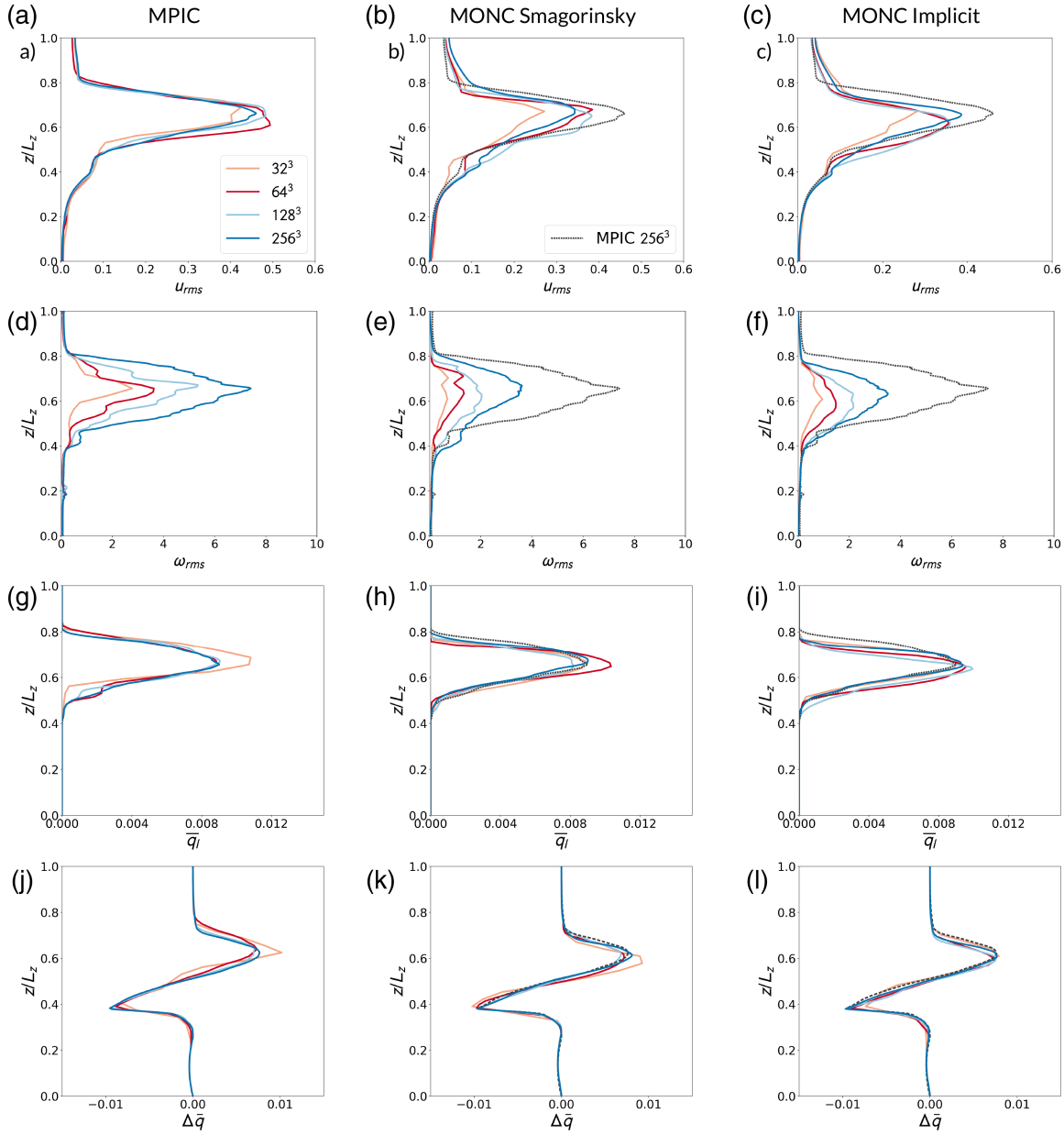

FIGURE 6 Resolution sensitivity of bulk flow properties at $t = 6$: (a–c) root-mean-square velocity, (d–f) root-mean-square vorticity and (g–i) liquid water specific humidity. (j–l) show the difference in total specific humidity between the start and end of the simulations. Columns are as in Figure 3. Colours represent the number of grid points (32^3 pink, 64^3 red, 128^3 light blue and 256^3 dark blue). The 256^3 MPIC simulation is represented as a black dotted line in the MONC plots

As expected, small-scale features can be resolved by MPIC simulations with a grid spacing that is about double that of the corresponding MONC simulations, but it is good to be cautious about the implications this has for mixing, as we shall see below.

In conclusion, the new MPIC model can simulate convection *purely in terms of parcels*. MPIC and MONC agree on large-scale flow features, although MONC tends to diffuse smaller-scale features at a low resolution. A major result is that the MPIC simulations show substantially more detail in the dynamical fields comparison. This is further quantified in Section 3.4.

3.3 | Resolution sensitivity of bulk flow properties

In order to quantitatively test convergence, the vertical profiles of a number of bulk measures of the flow properties are considered. Such bulk measures have previously been used in meteorology, where specific realisations of the flow are chaotic (Langhans *et al.* 2012).

A number of such bulk measures are plotted in Figure 6. In all the panels, the highest resolution MPIC results are shown for reference (dotted black curve in the MONC results). Here, the root-mean-square velocity profile (u_{rms}; Figure 6a–c)

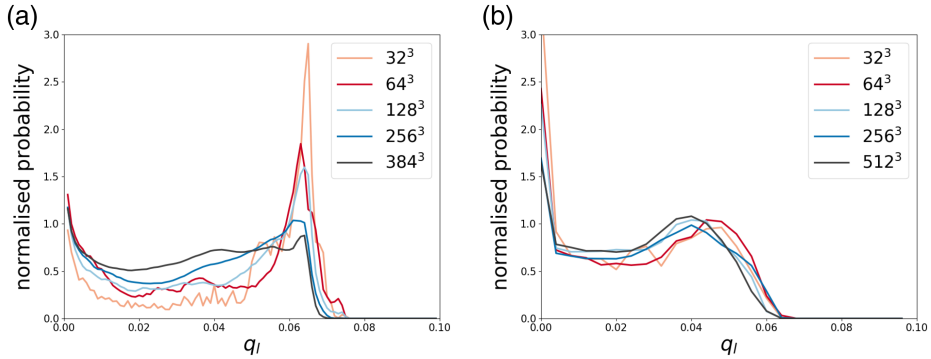

FIGURE 7 Liquid water specific humidity probability distribution function at $t = 6$ in (a) MPIC and (b) MONC using the Smagorinsky scheme, for various resolutions (colours are as in Figure 6, with an additional black curve indicating higher-resolution results)

gives a measure of the total amount of turbulent energy in the simulations, rather than a measure of error, and is therefore expected to converge across resolutions. This quantity is shown at $t = 6$, which is during the most active phase of the updraught. For MPIC, results based on gridded velocities are shown. As compared to MONC, the maximum in u_{rms} is higher and more localized in the MPIC simulations with the same grid resolution. The MPIC results also converge more rapidly.

For the root-mean-square vorticity (ω_{rms}; Figure 6d–f), values on the grid used for advection are used here rather than values on the finer grid. In both models, this quantity increases significantly each time the number of grid points is doubled. The rms vorticity measures gradients of velocity, which is much more sensitive to resolution as ever more small-scale gradients are resolved. (These small-scale gradients do not substantially change the magnitude of the velocity field, only its derivative at small scales.) D18 found ω_{rms} at early times in the simulation to roughly double each time the resolution was doubled. Its value is higher for MPIC simulations with the same number of grid points, which indicates that more small-scale structures are present in MPIC, as expected.

The vertical profile of horizontal mean liquid water specific humidity at $t = 6$ is shown in Figure 6g–i. The overall agreement between MPIC and MONC is good for this quantity, in particular at high resolution, although the cloud reaches somewhat higher levels in MPIC. In some regions, the results are sensitive to resolution; for MPIC this sensitivity is strongest near cloud base. This is because the inflow in the high-resolution simulations forms narrow structures that do not contribute a large amount of liquid water to the mean profile (Figure 3).

Figure 6j–l shows the total specific humidity change at $t = 10$ with respect to the initial profile. This serves as a bulk measure for the moisture transport by convection over time. Again, the overall agreement is good, though MPIC produces slightly deeper convection.

Figure 7 shows the probability density functions of the liquid water content. Regions of high liquid water content in MPIC (Figure 7a) are retained throughout the simulation as compared to MONC with the Smagorinsky scheme (Figure 7b). At all resolutions, higher values of liquid water specific humidity are retained in MPIC (also Figure B1). However, the coarser MPIC simulations underestimate the amount of dilution: the PDF of liquid water broadens significantly for MPIC simulations as the resolution increases. A larger number of weakly diluted parcels lead to reduced evaporation, and hence a higher cloud buoyancy. This is consistent with the higher centre of mass of thermals in MPIC.

The sharp peak in the liquid water probability density function in MPIC appears clearly only when probability density functions are determined directly from the parcel diagnostics. Appendix B describes how the use of parcel diagnostics allows us to accurately evaluate the rate of dilution in MPIC.

Figure 8 shows the probability distribution function of absolute vorticity in MPIC. The amount of vorticity in the simulation increases rapidly with resolution, consistent with our previous analysis of the root-mean-square vorticity. The amount of vorticity and therefore the stretching rate is very sensitive to resolution. This explains why the mixing rate is also sensitive to resolution.

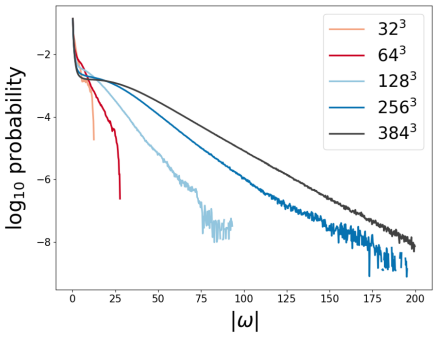

FIGURE 8 Probability distribution function of the absolute value of the vorticity in MPIC, for various resolutions at $t = 6$ (colours are as in Figure 7a)

3.4 | Measures of the representation of small-scale motions

A further measure of the representation of small-scale motions is given by the kinetic energy spectrum $\mathcal{K}(k)$, which is described in more detail in D18. These are calculated using gridded fields from both models. Results for $t = 6$ are shown in Figure 9a,b. The MPIC spectra generally retain more energy at small scales than MONC at the same resolution, but are also characterized by a sharp cut-off, which is due to a spectral filter used to prevent aliasing. We have performed sensitivity experiments that show a similar spectrum, but with a longer tail, can be obtained by applying this filter only for vorticity tendencies, and not in the solver (not shown). Both the faster convergence of the rms velocity and the spectra indicate that the velocity field is more detailed in MPIC when the same number of grid points is used, which suggests that small-scale eddies at the cloud edge can be better resolved. This may seem to contradict the slower reduction in liquid water content in low-resolution MPIC simulations as compared to MONC simulations with the same grid spacing shown in Figure 7. In MPIC, we explicitly represent the dilution process beyond the scale of resolved motion, whereas in MONC it is handled by the numerics or subgrid parametrization. However, we are currently only representing the effect of resolved dynamics on both entrainment and further mixing in MPIC. The mixing by turbulence on scales that are represented by neither model is therefore underestimated in MPIC.

Implicit MONC simulations also retain more energy on small scales than MONC simulations using the Smagorinsky scheme.

A second measure of small-scale motions can be obtained from the spectrum of the variance of total specific humidity $\mathcal{H}(k)$. This is calculated in the same manner as the kinetic energy spectrum. However, to isolate the cloud dynamics from the overall stratification, the mean vertical profile of total humidity has been subtracted before processing the fields.

For MPIC, we use the projection algorithm described in Appendix A to first construct results on a fine grid. Figure 9c shows results for simulations with 64^3, 128^3 and 256^3 grid points. Figure 9d shows spectra for the reference simulations, which have 1024^3 grid points for MONC and 384^3 for MPIC.

The humidity spectra of the MPIC simulations again show more variance at the intermediate and small scales, and in this case there is no clear cut-off. Compared to MONC, the MPIC spectra show a slope that is constant over a range that extends out to much smaller scales. This shows how additional detail below the grid scale is captured in MPIC simulations as compared to MONC due to the subgrid parcel representation.

4 | THE ORIGIN OF IN-CLOUD AIR

In this section we illustrate how MPIC can be used to tackle Lagrangian questions which are challenging to address in a Eulerian model. A tracer was added to track the level of origin of parcels. This tracer is treated in the same way as liquid water buoyancy and humidity when splitting and redistribution occurs.[1]

Figure 10 shows the vertical displacement of parcels from their altitude at $t = 0$ for an MPIC simulation with 256^3 grid points. This exhibits various noteworthy features. At $t = 4$, a region of in-cloud air originating from lower layers is shown in red between $z = 0.5$ and $z = 0.7$. This region has a sharp edge, beyond which there is an environment where subsidence has occurred. The effect of subsidence appears strongest at a level below the rising updraught. The cloud has entrained air at the centre of the updraught, which has ascended less than the air in the vortex ring. A column of rising air is found here, which is consistent with observational and modelling studies of cumulus convection (Blyth *et al.*, 2005). The region above the updraught has also ascended.

At $t = 6$, there is additional entrainment of air that has come down at the side of the cloud and is rising again. Pockets of this entrained air appear to have mixed into the cloud at $t = 8$. Though these pockets have descended, they have done so by a relatively small amount, which is consistent with earlier work on convection in moist environments (Blyth *et al.*, 1988; Heus *et al.*, 2008; Böing *et al.*, 2014). At this time, the air above the updraught has experienced a net descent and gravity waves become evident. A region of net descent appears below the cloud at $z/L_z \approx 0.3$. This region separates the cloud from another region of net ascent in the neutally stratified boundary layer below.

5 | CONCLUSIONS

The development of a warm, moist thermal was simulated in a novel essentially Lagrangian model, MPIC, as well as in the Met Office NERC Cloud model, MONC. The simulation was designed to capture the effects of latent heat release and evaporation, but is not aiming to address a particular type of atmospheric situation.

The overall development of the cloud is in good agreement between the MPIC and MONC simulations, as shown both in visual comparisons of the flow structure as well as in quantitative bulk measures of flow properties. The

[1]For studies that trace the origin of parcels over a longer duration, it could be argued that when redistribution occurs, it would be fairer to randomly choose a value of level of origin from the constituent parcels with appropriate weights, so that two parcels from a low and a high level do not produce a parcel that seems to originate from a level in between. Such an approach would not be possible in a conventional model.

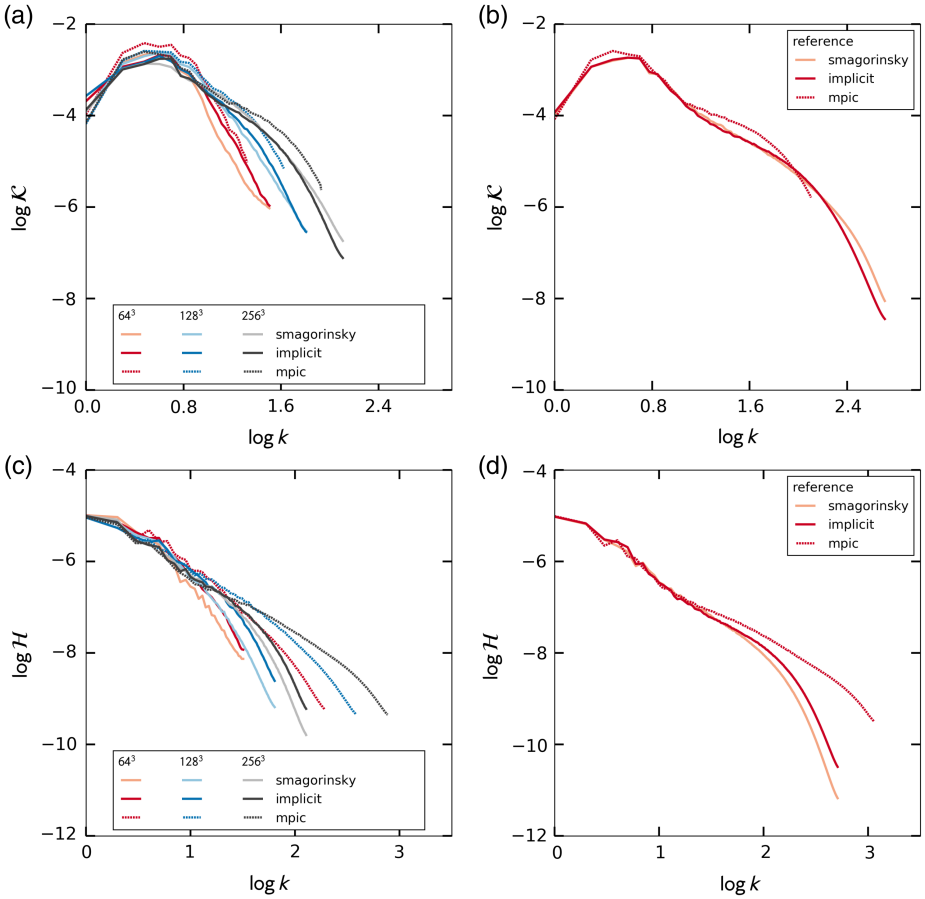

FIGURE 9 Kinetic energy spectra at $t = 6$ for (a) simulations with 64^3 (red), 128^3 (blue) and 256^3 (black/grey) grid points, and (b) for the reference simulations with 384^3 (MPIC) and 1024^3 (MONC) grid points. (c, d) show the corresponding total specific humidity variance spectra. Here, the MPIC spectra are based on detailed fields derived using the algorithm in Appendix A

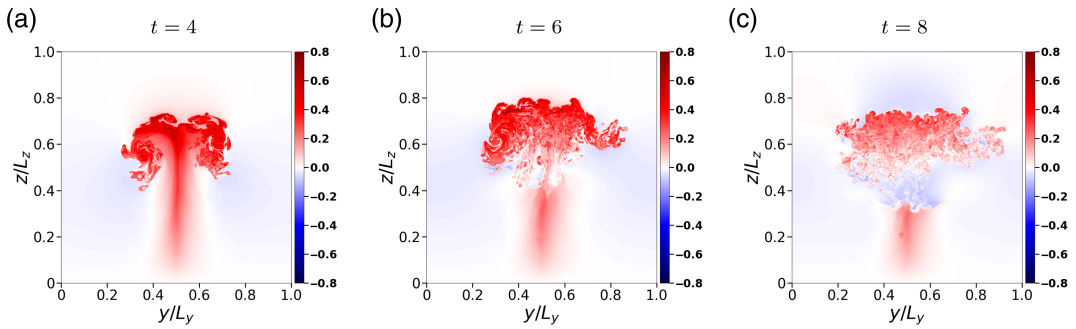

FIGURE 10 Detailed visualization of the vertical displacement of parcels since the beginning of the simulation normalized by the domain height for an MPIC simulation with 256^3 grid points after (a) 4 , (b) 6 and (c) 8 time units

MPIC simulations capture a large amount of detail in the subgrid-scale flow when compared to MONC at the same grid resolution due to the presence of multiple parcels in each grid box. This has been shown in comparisons of cross-sections and in corresponding spectra across model resolutions.

Subgrid-scale mixing is weaker in all the MPIC simulations than in MONC. Whereas low-resolution MONC simulations tend not to represent the cloud structure as well as low-resolution MPIC simulations, the latter have insufficient small-scale mixing compared to the high-resolution reference studies. The slow dilution of thermodynamic properties when MPIC is used at low resolution may be due to the absence of a representation of the velocity field on the parcel scale, that is, at subgrid scales. The consequence of weak mixing is that regions of high liquid water content remain present that are absent in higher-resolution simulations. This behaviour at low resolution sharply contrasts that of MONC, where regions with high liquid water content are eroded (diffused) quickly at the lowest resolution.

Even the highest resolution MONC and MPIC simulations are not fully in agreement when it comes to the probability density function of liquid water. At these resolutions, the MPIC results appear to converge, so this may indicate that the models disagree to some extent. However, the MPIC results also change if they are not calculated on parcels but by projection onto a grid (the projection smooths the field). In this case, they resemble MONC results to a greater extent (Appendix B).

A relatively slow convergence of the properties of the liquid water field is not unique to our study. It has also been found in previous work that tried to represent subgrid-scale mixing explicitly, e.g. Jarecka *et al.* (2009). Apart from this single issue, the models show a remarkable extent of agreement.

When used at sufficiently high resolution, MPIC may provide new perspectives on the role of entrainment and in particular vorticity dynamics in cloud formation (e.g. the role of vorticity extrema in mixing and the presence and dynamics of vortex rings). Previous work has shown the usefulness of Lagrangian analysis for understanding cloud processes and turbulence (e.g. Lasher-Trapp *et al.*, 2005; Cooper *et al.*, 2013; Böing *et al.*, 2014). MPIC takes this approach a step further by allowing for a Lagrangian analysis which is fully consistent with the model dynamics. A first example of such an analysis was given in Section 4. Further benefits in terms of computational cost can be expected in situations where a large number of tracers is carried by the parcels.

Developing a version of MPIC which uses the same thermodynamical formulation as the standard version of MONC, and testing this on more representative scenarios driven by surface fluxes, is one of our priorities for further development of MPIC. Furthermore, we want to improve the convergence of the liquid water probability density function with resolution in MPIC. The aim is first to attempt to implement methods that improve resolved and explicitly represented dynamics, for example the use of a finer grid in parts of the dynamical core (in particular for the inversion procedure) and a more physical model of parcel stretch (McKiver and Dritschel, 2003) and splitting. Kaas *et al.* (2013) also present an approach that looks at explicit deformation of parcels; in terms of mixing, a grid is used here as well, and it shares some properties with the approach currently used in MPIC. A second strategy would be to exploit subgrid information carried by parcels. Particle–particle particle-mesh (PPPM; e.g. Walther and Morgenthal 2002) methods use this strategy, but these are very expensive as many parcel–parcel interactions need to be represented. A stochastic approach based on methods used in dispersion modelling (Thomson, 1987; Weil *et al.*, 2004) could be more suitable, or alternatively a fractal representation of unresolved turbulence (e.g. Basu *et al.* 2004) could be explored. We believe a mostly parcel-based approach to mixing will be valuable for studies that use Lagrangian diagnostics. As a last resort, it would be possible to use a gridded approach only to subgrid-scale mixing.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support for this research from the EPSRC Maths Foresees Network. The numerical method development was carried out under the grant "A prototype vortex-in-cell algorithm for modelling moist convection" from March to October 2016. SJB, DJP and AMB are partially funded through the NERC/Met Office Joint Programme on Understanding and Representing Atmospheric Convection across Scales (grant NE/N013840/1). DJP is supported by a Royal Society Wolfson Research Merit Award and by the Met Office Academic Partnership. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk; accessed 3 April 2019). This research was supported in part by the National Science Foundation under Grant No. NSF PHY17-48958 (in particular, SB benefited from participation in the KITP Program on Planetary Boundary Layers in Atmospheres, Oceans, and Ice on Earth and Moons).

REFERENCES

Abel, S.J. and Shipway, B. (2007) A comparison of cloud-resolving model simulations of trade wind cumulus with aircraft observations taken during RICO. *Quarterly Journal of the Royal Meteorological Society*, 133, 781–794.

Aranami, K., Davies, T. and Wood, N. (2015) A mass restoration scheme for limited-area models with semi-Lagrangian advection. *Quarterly Journal of the Royal Meteorological Society*, 141, 1795–1803.

Basu, S., Foufoula-Georgiou, E. and Porté-Agel, F. (2004) Synthetic turbulence, fractal interpolation, and large-eddy simulation. *Physical Review E*, 70, 026310.

Blyth, A.M., Lasher-Trapp, S. and Cooper, W. (2005) A study of thermals in cumulus clouds. *Quarterly Journal of the Royal Meteorological Society*, 131, 1171–1190.

Blyth, A.M., Cooper, W.A. and Jensen, J.B. (1988) A study of the source of entrained air in Montana cumuli. *Journal of the Atmospheric Sciences*, 45, 3944–3964.

Böing, S.J., Jonker, H.J., Nawara, W.A. and Siebesma, A.P. (2014) On the deceiving aspects of mixing diagrams of deep cumulus convection. *Journal of the Atmospheric Sciences*, 71, 56–68.

Brown, N., Weiland, M., Hill, A., Shipway, B., Maynard, C., Allen, T. and Rezny, M. (2015) *A highly scalable Met Office NERC Cloud model In: Proceedings of the 3rd International Conference on Exascale Applications and Software*. pp. 132–137. Edinburgh: University of Edinburgh.

Cooper, W., Lasher-Trapp, S.G. and Blyth, A.M. (2013) The influence of entrainment and mixing on the initial formation of rain in a warm cumulus cloud. *Journal of the Atmospheric Sciences*, 70, 1727–1743.

Dritschel, D.G. and Ambaum, M. (1997) A contour-advective semi-Lagrangian numerical algorithm for simulating fine-scale

conservative dynamical fields. *Quarterly Journal of the Royal Meteorological Society*, 123, 1097–1130.

Dritschel, D.G., Böing, S.J., Parker, D.J. and Blyth, A.M. (2018) The moist parcel-in-cell method for modelling moist convection. *Quarterly Journal of the Royal Meteorological Society*, 144, 1695–1718.

Fontane, J. and Dritschel, D.G. (2009) The HyperCASL Algorithm: a new approach to the numerical simulation of geophysical flows. *Journal of Computational Physics*, 228, 6411–6425.

Grabowski, W.W. and Clark, T. (1993) Cloud-environment interface instability: Part II: Extension to three spatial dimensions. *Journal of the Atmospheric Sciences*, 50, 555–573.

Heus, T., Van Dijk, G., Jonker, H.J. and Van den Akker, H.E. (2008) Mixing in shallow cumulus clouds studied by Lagrangian particle tracking. *Journal of the Atmospheric Sciences*, 65, 2581–2597.

Hohenegger, C., Schlemmer, L. and Silvers, L. (2015) Coupling of convection and circulation at various resolutions. *Tellus A*, 67. 26678.

Jarecka, D., Grabowski, W.W. and Pawlowska, H. (2009) Modeling of subgrid-scale mixing in large-eddy simulation of shallow convection. *Journal of the Atmospheric Sciences*, 66, 2125–2133.

Kaas, E. (2008) A simple and efficient locally mass conserving semi-Lagrangian transport scheme. *Tellus A*, 60, 305–320.

Kaas, E., Sørensen, B., Lauritzen, P.H. and Hansen, A.B. (2013) A hybrid Eulerian–Lagrangian numerical scheme for solving prognostic equations in fluid dynamics. *Geoscientific Model Development*, 6, 2023–2047.

Langhans, W., Schmidli, J. and Schär, C. (2012) Bulk convergence of cloud-resolving simulations of moist convection over complex terrain. *Journal of the Atmospheric Sciences*, 69, 2207–2228.

Lasher-Trapp, S.G., Cooper, W.A. and Blyth, A.M. (2005) Broadening of droplet size distributions from entrainment and mixing in a cumulus cloud. *Quarterly Journal of the Royal Meteorological Society*, 131, 195–220.

Lauritzen, P.H. and Thuburn, J. (2012) Evaluating advection/transport schemes using interrelated tracers, scatter plots and numerical mixing diagnostics. *Quarterly Journal of the Royal Meteorological Society*, 138, 906–918.

Leonard, B., MacVean, M. and Lock, A.P. (1993) *Positivity-preserving numerical schemes for multidimensional advection*. NASA Lewis Research Center. Tech. Rep.

Mason, P.J. and Brown, A.R. (1999) On subgrid models and filter operations in large eddy simulations. *Journal of the Atmospheric Sciences*, 56, 2101–2114.

Matheou, G. (2011) On the fidelity of large-eddy simulation of shallow precipitating cumulus convection. *Monthly Weather Review*, 139, 2918–2939.

McKiver, W. and Dritschel, D.G. (2003) The motion of a fluid ellipsoid in a general uniform background flow. *Journal of Fluid Mechanics*, 474, 147–173.

Ogura, Y. (1963) The evolution of a moist convective element in a shallow, conditionally unstable atmosphere: a numerical calculation. *Journal of the Atmospheric Sciences*, 20, 407–424.

Paluch, I.R. (1979) The entrainment mechanism in Colorado cumuli. *Journal of the Atmospheric Sciences*, 36, 2467–2478.

Panosetti, D., Böing, S.J., Schlemmer, L. and Schmidli, J. (2016) Idealized large-eddy and convection-resolving simulations of moist convection over mountainous terrain. *Journal of the Atmospheric Sciences*, 73, 4021–4041.

Petch, J.C. and Gray, M.E.B. (2001) Sensitivity studies using a cloud-resolving model simulation of the tropical west Pacific. *Quarterly Journal of the Royal Meteorological Society*, 127, 2287–2306.

Piacsek, S.A. and Williams, G.P. (1970) Conservation properties of convection difference schemes. *Journal of Computational Physics*, 6, 392–405.

Pincus, R., Barker, H.W. and Morcrette, J.-J. (2003) A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *Journal of Geophysical Research, Atmospheres*, 108(D13). https://doi.org/10.1029/2002JD003322.

Pressel, K., Kaul, C., Schneider, T., Tan, Z. and Mishra, S. (2015) Large-eddy simulation in an anelastic framework with closed water and entropy balances. *Journal of Advances in Modeling Earth Systems*, 7, 1425–1456.

Siebesma, A.P. and Cuijpers, J. (1995) Evaluation of parametric assumptions for shallow cumulus convection. *Journal of the Atmospheric Sciences*, 52, 650–666.

Smagorinsky, J. (1963) General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, 91, 99–164.

Stevens, B., Moeng, C.-H., Ackerman, A.S., Bretherton, C.S., Chlond, A., de Roode, S., Edwards, J., Golaz, J.-C., Jiang, H., Khairoutdinov, M., Kirkpatrick, M.P., Lewellen, D.C., Lock, A.P., Müller, F., Stevens, D.E., Whelan, E. and Zhu, P. (2005) Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly Weather Review*, 133, 1443–1462.

Stevens, D.E. and Bretherton, C.S. (1999) Effects of resolution on the simulation of stratocumulus entrainment. *Quarterly Journal of the Royal Meteorological Society*, 125, 425–439.

Thomson, D.J. (1987) Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180, 529–556.

Tsang, Y.-K. and Vallis, G.K. (2018) A stochastic Lagrangian basis for a probabilistic parameterization of moisture condensation in Eulerian models. *Journal of Atmospheric Sciences*, 75, 3925–3941.

Twomey, S. (1966) Computations of rain formation by coalescence. *Journal of the Atmospheric Sciences*, 23, 405–411.

Walther, J.H. and Morgenthal, G. (2002) An immersed interface method for the vortex-in-cell algorithm. *Journal of Turbulence*, 3(N39). https://doi.org/10.1088/1468-5248/3/1/039.

Weil, J.C., Sullivan, P.P. and Moeng, C.-H. (2004) The use of large-eddy simulations in Lagrangian particle dispersion models. *Journal of the Atmospheric Sciences*, 61, 2877–2887.

Zerroukat, M., Wood, N. and Staniforth, A. (2002) SLICE: a Semi-Lagrangian Inherently Conserving and Efficient scheme for transport problems. *Quarterly Journal of the Royal Meteorological Society*, 128, 2801–2820.

How to cite this article: Böing SJ, Dritschel DG, Parker DJ, Blyth AM. Comparison of the Moist Parcel-in-Cell (MPIC) model with large-eddy simulation for an idealized cloud. *Q J R Meteorol Soc*. 2019;1–17. https://doi.org/10.1002/qj.3532

APPENDICES

A: AN ALGORITHM FOR RECONSTRUCTING PARCEL PROPERTIES ON THE FINE GRID

In some of our analysis, parcel properties are interpolated to a grid with a grid spacing that is six times finer than that of the grid used in the vorticity inversion. This approach is useful, for example, for producing two-dimensional cross-sections with a detailed representation of the flow field. While the tri-linear interpolation algorithm used in the vorticity solver aims to find mean parcel properties on a grid (where the spacing is coarser than the parcel radius), the fine grid reconstruction algorithm aims to represent the parcel properties at a scale similar to the smallest parcels. This requires a different approach, because parcel spacing and size vary.

The interpolation algorithm is used to calculate a representative value of a field at a location $\bar{\mathbf{x}}$ in the domain, where the overline denotes interpolated values. As an example, we consider the interpolated humidity field $\bar{q}(\bar{\mathbf{x}})$, which results from taking a weighted sum over the parcels in the domain:

$$\bar{q}(\bar{\mathbf{x}}) = \frac{\sum_{i\in\mathcal{P}} \mu(\bar{\mathbf{x}}, \mathbf{x}_i, V_i)\, q_i}{\sum_{i\in\mathcal{P}} \mu(\bar{\mathbf{x}}, \mathbf{x}_i, V_i)}. \tag{A1}$$

Here, V_i is the parcel volume and μ a weighting function. For each parcel, an equivalent radius is defined as

$$r_i = \left(\frac{3V_i}{4\pi}\right)^{1/3}. \tag{A2}$$

The weights are chosen to meet following conditions:

1. The weights are isotropic, that is, they depend only on the Euclidean distance $r \equiv \|\bar{\mathbf{x}} - \mathbf{x}_i\|$.
2. There is a maximum radius of influence r_m, which is the same for each parcel. Outside this radius of influence, $\mu(\bar{\mathbf{x}}, \mathbf{x}_i, V_i) = 0$. On the other hand, at each point in space, there must be at least one parcel which contributes to $\bar{q}(\bar{\mathbf{x}})$. As each grid box is guaranteed to contain at least one parcel, we can choose

$$r_\mathrm{m} = (1+\varepsilon)\sqrt{\Delta_x^2 + \Delta_y^2 + \Delta_z^2}. \tag{A3}$$

Here, ε is a small number (e.g. 0.02; this merely serves to avoid division by small numbers) and $\Delta_{x/y/z}$ denotes the grid spacing of the (vorticity) inversion grid, so that r_m is marginally bigger than the maximum distance within an inversion grid box.
3. The total contribution of a parcel to the three-dimensional projection domain (here the entire domain) is equal to its volume, that is,

$$\int \mu(\bar{\mathbf{x}}, \mathbf{x}_i, V_i)\, \mathrm{d}\boldsymbol{X} = V_i. \tag{A4}$$

This condition holds for parcels for which the projection does not extend beyond the domain top or bottom, otherwise the contribution of a parcel is equal to the contribution it gives on the interior of the domain using our method.
4. The contribution of a parcel decreases rapidly when $r > r_i$.

In order to meet these conditions, $\mu(r)$ is chosen as:

$$\mu(r) = V_i f(r)/N. \tag{A5}$$

Here, $f(r)$ specifies a radial dependence and N is a normalization factor, chosen such that

$$N = \int f(r)\, \mathrm{d}\boldsymbol{X}. \tag{A6}$$

We employ a Gaussian dependence on radius here:

$$f(r) = \begin{cases} \mathrm{e}^{-(r/r_i)^2} - \mathrm{e}^{-(r_\mathrm{m}/r_i)^2} & r < r_\mathrm{m}, \\ 0 & r \geq r_\mathrm{m}. \end{cases} \tag{A7}$$

This implies

$$\begin{aligned} N &= \int_{r=0}^{r_\mathrm{m}} 4\pi r^2 \left(\mathrm{e}^{-(r/r_i)^2} - \mathrm{e}^{-(r_\mathrm{m}/r_i)^2}\right) \mathrm{d}r \\ &= \pi r_i^2 \left(r_i\sqrt{\pi}\,\mathrm{erf}(r_\mathrm{m}/r_i) - 2r_\mathrm{m}\mathrm{e}^{(-r_\mathrm{m}^2/r_i^2)}\right) \\ &\quad - \frac{4\pi}{3} r_\mathrm{m}^3 \mathrm{e}^{-(r_\mathrm{m}/r_i)^2}. \end{aligned} \tag{A8}$$

Other choices of $f(r)$ can be made for $r < r_\mathrm{m}$ that also lend themselves to analytical integration; in particular the second power within the exponential function can easily be replaced by the third or fourth power. The choice of function impacts on the character of high-frequency modes in the projected field.

B: RECONSTRUCTING THE LIQUID WATER FIELD FROM PARCEL DATA

In Section 3.3, we used the parcel properties to derive the statistics of the liquid water field. In practice, we might need to reconstruct the liquid water field on the grid in order to couple the dynamics to other processes (e.g. a radiative transfer scheme).

Several approaches are considered here, and it is shown that the choice of approach significantly changes the probability density function of the liquid water field. As an example, we consider the impact of different approaches on the probability density function at $t = 6$ in Figure B1. The probability

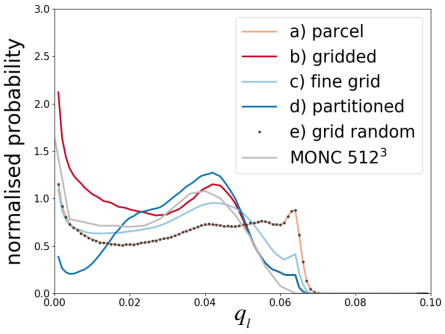

FIGURE B1 Liquid water specific humidity pdf at $t = 6$ in the 384^3 MPIC simulation, as diagnosed using different methods

distribution function from a 512^3 MONC simulation with the Smagorinsky closure is given as a reference.

a. If only the probability distribution function itself needs to be calculated, this can be done directly from the parcels. However, when the spatial distribution is also important, a different approach is needed.
b. One might simply derive the mean liquid water content of each grid box using the algorithm described in D18. This approach conserves the total liquid water content but has two drawbacks. First of all, some grid boxes are partially cloudy but are now simply seen as containing liquid water, so the cloud fraction increases. In addition, the extremes in the liquid water distribution are no longer present in the gridded field.
c. The detailed projection scheme described in Appendix A can be used. Here, q is calculated on the fine grid using the algorithm in the Appendix and q_l is calculated from the field of q. This approach is much more computationally costly, but it gives a better probability distribution than the coarse grid approaches and retains small-scale features. This approach is useful mainly for detailed analysis of a limited number of snapshots from MPIC.
d. A fourth approach would be to partition each grid box into a cloudy and a non-cloudy part (as is often done in radiative transfer), and to separately determine the mean value in the cloudy and the non-cloudy parts. This results in the correct cloud fraction and total liquid water content, but both high and low values of liquid water content are under-represented. This is an example of how MPIC's subgrid-scale information (the cloud fraction) can be used. It is possible to extend this, and for example to consider the variance of liquid water within the cloudy part of each grid box, fit an assumed probability distribution function, and use this in radiative transfer calculations.
e. For some purposes, it is possible to randomly select a parcel in each grid box, with odds proportional to parcel volume. This preserves the probability density function as well as some spatial information, and is similar to methods in radiative transfer where a randomly selected subset of model columns is used (Pincus *et al.*, 2003).