An Implicit–Explicit Vertical Transport Scheme for Convection-Allowing Models

LOUIS J. WICKER

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

WILLIAM C. SKAMAROCK

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 19 February 2020, in final form 6 July 2020)

ABSTRACT

An adaptive implicit–explicit vertical transport method is implemented in the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW), and improved integration efficiency is demonstrated for configurations employing convective-allowing horizontal and vertical resolutions. During the warm season over the continental United States, stable forecasts at convective-allowing resolutions are more challenging because localized regions of extreme thermodynamic instability generate large vertical velocities within thunderstorms that cause the integrations to become unstable because of violations of the Courant–Friedrichs–Lewy (CFL) condition for the explicit advection scheme used in WRF-ARW. The implicit–explicit vertical transport scheme removes the CFL instability but maintains accuracy for typical vertical velocities. Tests using this scheme show that the new scheme permits a time step that is 20%–25% percent larger, and it reduces the wall clock time by 10%–13% percent relative to a configuration similar to a current operational convection-allowing model while also producing more realistic updraft intensities within the most intense storms. Other approaches to maintain stability are either less efficient (e.g., reducing the time step) or significantly impact the solution accuracy (e.g., increasing the damping and/or reducing the latent heating, which severely limits the updraft magnitudes during the forecasts).

1. Introduction

Increases in computational speed, memory, and storage over the last decade have allowed many numerical weather prediction centers to increase horizontal and vertical grid resolutions to where convection can be explicitly represented. Though convective storms may be poorly resolved in convective-allowing models (CAMs), late spring and summertime continental environments often have very large thermodynamic instability that generate significant vertical velocities within simulated storms. Combined with strong vertical shear, long-lived convective storms can sustain updrafts exceeding $35\,\mathrm{m\,s^{-1}}$ even when the horizontal grid resolution is ~3 km (Potvin et al. 2019; Potvin and Flora 2015; Verrelle et al. 2015). These intense updrafts, combined with typical vertical grid spacing in the model's mid- and upper troposphere ($\Delta z \sim 500\,\mathrm{m}$), can severely limit a CAM's time step due to numerical stability considerations for the vertical transport. While small regions of intense convective activity may last only for a few hours, CAM forecasts of these potentially high-impact storms are often used by forecasters to help foresee when and where severe weather will occur. Therefore it is important that a CAM's numerical scheme capture the intensity of these storms and represent them with as much fidelity as possible.

For operational prediction, a CAM's time step needs to be as large as possible to minimize the use of computational resources and to generate forecasts as rapidly as possible. In simulated convective storms where updraft velocities larger than $35\,\mathrm{m\,s^{-1}}$ are present, the vertical Courant number $C_r^w = |W\Delta t/\Delta z|$ can greatly exceed 1 and violate the numerical stability limit for most explicit transport schemes. Several operational CAMs have opted to filter or reduce the most intense updrafts because this enables the use of a larger time step and reduces the computational cost. For example, the High-Resolution Rapid

Corresponding author: Louis J. Wicker, louis.wicker@noaa.gov

DOI: 10.1175/MWR-D-20-0055.1

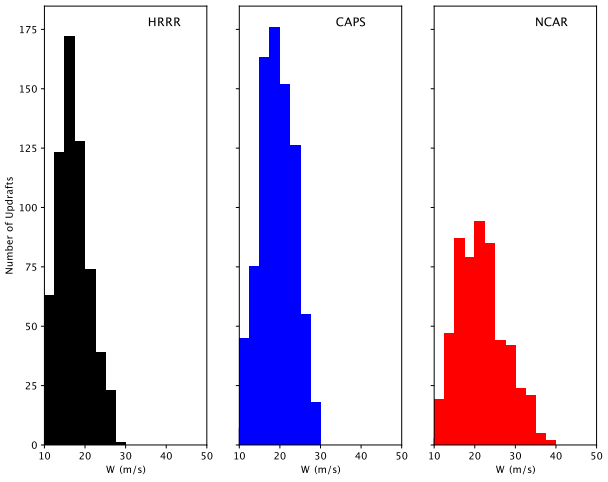

FIG. 1. Maximum vertical velocities at the 400-hPa level from individual storms obtained from three CAMs examined in Potvin et al. (2019) for the month of May 2017. The data are obtained from the 0000 UTC CAM forecasts during the diurnal precipitation maximum from 1800 to 0200 UTC. Shown are (left) the operational HRRRv3, (center) member 1 of the CAPS ensemble, and (right) member 1 of the NCAR ensemble. See Potvin et al. (2019) for more detail.

Refresh (HRRR) model (Benjamin et al. 2016), which is based on a configuration of the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW) community modeling system (Skamarock et al. 2008), uses two filters to control the magnitude of the vertical motion. The most significant filter is the reduction of latent heating by placing an upper limit on the heating rate produced by the physics parameterizations (mostly by phase changes within the cloud microphysics) at each time step. The second filter is a Rayleigh damping term that is activated whenever the C_r^w exceeds 1.2 in the model. The Rayleigh damping filter (hereinafter called the w damping) serves to reduce the updraft magnitude. While both of these filters are used in the HRRR, experiments show that the reduction in updraft intensity from the latent heating limiter has a much larger impact (S. G. Benjamin 2019, personal communication). Figure 1 shows the impacts of these filters on forecasts of convection using updraft climatologies from the Community Leveraged Unified Ensemble (CLUE; Clark et al. 2018) from May–June 2017. Using output from the CLUE experiment Potvin et al. (2019) presents detailed analyses of storm characteristics. The updraft data shown in Fig. 1 are adapted using data from Potvin et al. (2019). The Center for the Analysis and Prediction of Storms (CAPS; Snook et al. 2019) and the HRRR forecasts both use the latent heat limiting and w damping. The National Center for Atmospheric Research (NCAR, Schwartz et al. 2019) CAM forecast does not. The NCAR CAM forecast climatology has a much broader distribution, with nearly 100 updrafts greater than 30 m s^{-1}, whereas all of the updrafts in the CAPS and HRRR forecasts remain below 30 m s^{-1}. To keep the integration stable, the NCAR CAM forecast uses a 33% smaller time step than the HRRR model for integration (Δt = 15 s). Other operational CAM forecasts also have used vertical filters to limit their updraft intensities. The leapfrog time-integration scheme in the COSMO model (Baldauf et al. 2011) artificially limits the vertical Courant number to 1 for scalar transport, while using a semi-implicit vertical advection scheme for the dynamical variables. The Runge–Kutta version of COSMO also uses a semi-implicit scheme for the vertical advection of the dynamical variables.

Time-step limitations arising from isolated strong vertical motions as the horizontal grid resolution is increased is not only present in atmospheric models. Shchepetkin (2015, hereinafter S15) describes a similar problem occurring in the Regional Oceanic Modeling System (ROMS; Shchepetkin and McWilliams 2005). Similar to atmospheric CAMs, as an ocean model's horizontal grid resolution increases, the Courant number associated with updrafts along the sloped topography or along oceanic fronts becomes large enough to make the integration unstable when explicit Eulerian transport schemes are employed. The unstable regions are very localized and represent a small fraction of columns in the ROMS grid. To circumvent this time-step restriction S15 developed what we are calling the implicit–explicit vertical advection scheme (IEVA). The IEVA scheme is an adaptive vertical transport scheme combining explicit and implicit methods, and S15 implemented it for the predictor-corrector time integration scheme used in the ROMS model. The algorithm retains the explicit method wherever the vertical Courant number is small enough to be stable, and in locations where it is too large it combines the transport from the explicit scheme and an absolutely stable implicit method. S15 demonstrates that this adaptive hybrid method is very effective in permitting larger time steps with minimal loss of accuracy.

This paper shows how IEVA is adapted and implemented into the ARW's Runge–Kutta time integration scheme. The benefits from IEVA in the WRF-ARW Model (hereinafter the WRF-ARW Model will be referred to as the WRF Model) are similar to those presented in S15. IEVA can stably produce forecasts with large vertical motions using large timesteps without the need for vertical velocity filters, thus enabling a more accurate representation of the strongest storm updrafts. The adaptive scheme also permits the use of a larger time step than the current operational configuration of

the HRRR such that the integration is overall is faster despite the increase in computational cost from the IEVA scheme. Section 2 describes the formulation of IEVA within WRF's Runge–Kutta integration method as well as its formulation for positive-definite or monotonic fluxes. The IEVA scheme's linear stability and performance for several pure transport problems is shown in section 3 as well as a two-dimensional stability analysis for the time-split explicit/implicit system used to efficiently integrate the fully compressible equations. Results from using IEVA in the WRF for a 27 April 2011 "super outbreak" forecast are shown in section 4. Section 5 summarizes the results.

2. IEVA description

a. *IEVA formulation for the WRF's Runge–Kutta scheme*

Adapting IEVA into the WRF's third-order Runge–Kutta time-integration scheme for the compressible nonhydrostatic equations (hereinafter referred to as RK3; Wicker and Skamarock 2002) is relatively straightforward. The implementation is nearly identical to S15 and a brief summary is presented here with emphasis on the differences between S15 and our implementation of IEVA in WRF. An extensive presentation of the motivation and design behind IEVA is given in S15.

There are two major components of the S15 IEVA algorithm, the formulation of the explicit and implicit transport fluxes, and the method for splitting the vertical velocity into explicit and implicit contributions. We discuss these components first. The transport algorithms in WRF are consistent and, in the case of scalar transport, optionally positive definite and shape preserving (monotonic). The WRF IEVA implementation maintains these properties, as discussed in this section.

The transport equations in WRF can be written as

$$\frac{\partial \Phi}{\partial t} = -\nabla_\eta \cdot (\mathbf{V_h}\phi) - \frac{\partial(\Omega\phi)}{\partial \eta}, \tag{1}$$

where $\mathbf{V}_h = \mu_d(u, v, \omega) = (U, V, \Omega)$, $\phi = (u, v, w, \theta, q_m)$, and $\Phi = \mu_d\phi$. Following Klemp et al. (2007), one can write the temporally discretized version of RK3 in (1) as

$$\Phi^* = \Phi^n - \frac{\Delta t}{3}\nabla_\eta \cdot (\mathbf{V}_h\phi)^n - \frac{\Delta t}{3}\frac{\partial(\Omega\phi)^n}{\partial \eta}, \tag{2a}$$

$$\Phi^{**} = \Phi^n - \frac{\Delta t}{2}\nabla_\eta \cdot (\mathbf{V}_h\phi)^* - \frac{\Delta t}{2}\frac{\partial(\Omega\phi)^*}{\partial \eta}, \quad \text{and} \tag{2b}$$

$$\Phi^{n+1} = \Phi^n - \Delta t\nabla_\eta \cdot (\mathbf{V}_h\phi)^{**} - \Delta t\frac{\partial(\Omega\phi)^{**}}{\partial \eta}. \tag{2c}$$

WRF uses the Laprise (1992) vertical pressure coordinate η, where the dry-air column mass is μ_d and the covariant vertical velocity is $\omega = \dot{\eta}$. Scalar quantities such as water vapor, hydrometeors, or aerosol specific mixing ratios are represented by q_m. Potential temperature is represented by θ. The vertical mass flux for the transport in the dry dynamics (i.e., for u, v, w, and θ) is Ω, whereas for scalar transport Ω represents a time-averaged vertical flux that is consistent with the discrete mass-conservation equation in WRF. Since each substep of (2) is a forward Euler time step, it suffices to describe how IEVA is incorporated into a single forward step. While IEVA can be used on every RK3 substep, the WRF implementation only applies the scheme on the last RK3 step. This reduces the cost of IEVA significantly. No instabilities or significant solution differences are found when comparing simulations using IEVA on all three RK3 integration steps or just on the last step as long as the vertical Courant numbers remain less than 4.[1] The IEVA scheme only modifies the vertical transport term, and the discrete IEVA version of the transport equation given by (1), with the vertical transport separated into two pieces for the explicit and implicit transport as in S15, is

$$\Phi^{n+1} = \Phi^n - \Delta t\nabla_\eta \cdot (\mathbf{V}_h\phi)^n - \Delta t\delta_\eta[\Omega_e^n(\tilde{\phi})^n] - \Delta t\delta_\eta[\Omega_i^n(\overline{\phi})^{n+1}]. \tag{3}$$

Here Ω^e and Ω^i are the partitioned vertical mass fluxes at the top and bottom of the grid zone for the explicit and implicit transport and the operator δ_η represents a centered-flux difference between the upper and lower faces of the grid zone. The explicit fluxes use $\tilde{\phi}$, which represents the vertical interpolation of the variable ϕ to the faces of the grid cell. WRF is usually configured using third-, fourth-, or fifth-order polynomials, and fifth-order polynomials will be used for all the results shown in this paper. The implicit flux $\overline{\phi}$ can use either second-order centered or upwind interpolation. The implementation of the implicit fluxes in the test problems shown here and in the WRF Model implementation use upwind interpolation to maintain positivity. Either implicit interpolation method results in a tridiagonal system of equations similar to S15 and can be directly solved using standard techniques.

[1] Even with IEVA applied on every substep, WRF Model integrations can become unstable when the vertical Courant number continually exceeds 4 for a few tens of time steps. While infrequent, this situation usually occurs near the top of strong updrafts (near the tropopause) and generates large-amplitude gravity waves that break, have extreme three-dimensional velocity gradients on the grid scale, and appear to bring about integration instability.

b. Partitioning the vertical fluxes

The second component of the scheme involves splitting the vertical transport into explicit and implicit contributions by partitioning the vertical velocity. The formulas below are empirically designed by S15 and here adapted for the WRF Model. They are designed to provide a smooth transition between the explicit and implicit transport, and to accurately represent the vertical Courant number with a three-dimensional full physics model having terrain, uneven vertical grid layers, and within the WRF, three-dimensional mass divergence.

The splitting of the vertical velocity is accomplished by decoupling the vertical mass flux $\Omega_{k-1/2}$ ("k" indicates a scalar vertical level on the model grid) from the column mass μ_d and then partitioning the covariant vertical velocity $\omega_{k-1/2}$:

$$\omega^e_{k-1/2} = g\omega_{k-1/2} \quad \text{and}$$
$$\omega^i_{k-1/2} = (1-g)\omega_{k-1/2}, \tag{4}$$

where the function g is a transition function from pure explicit transport to a combination of explicit–implicit transport by partitioning the vertical velocity contributing to the explicit flux divergence. The formulation in (4) constrains the partitioning such that $\omega = \omega^e + \omega^i$. We will also require $0 \le g \le 1$ and that g be smooth, in this case at least C_1 continuous. This is achieved by using a formulation for g adapted from S15, but note that S15 uses a function f defined as the inverse of the partitioning function g used here (i.e., $f = g^{-1}$).

The partitioning function g is determined by three solution-dependent Courant numbers that we denote as α, $\alpha^*_{\max}$, and $\alpha^*_{\min}$; α is the vertical-velocity-based Courant number:

$$\alpha_{k-(1/2)} = \Delta t \left| \frac{\omega_{k-(1/2)}}{\Delta\eta_{k-(1/2)}} \right|, \tag{5}$$

and $\alpha^*_{\min}$ and $\alpha^*_{\max}$ are the Courant numbers used to define the blending zone for the function g. These Courant numbers necessarily take into account the *horizontal* Courant numbers because it is the combination of the Courant numbers in x, y and η that limit the explicit scheme stability. $\alpha^*_{\min}$ and $\alpha^*_{\max}$ are defined as

$$\alpha^\star_{\max} = \frac{\alpha_{\max} - \varepsilon\alpha^H}{\alpha_{\max}} \quad \text{and} \quad \alpha^\star_{\min} = \alpha_{\min}\frac{\alpha^\star_{\max}}{\alpha_{\max}}, \tag{6}$$

where $\alpha_{\min}$ and $\alpha_{\max}$ are user-specified parameters used to define the blending zone in the absence of horizontal flow and ε is also user specified and accounts for the difference in the stability limits due to the use of different horizontal and vertical interpolation formula for the fluxes. A 2D Courant number for the horizontal flow, α^H, accounts for the possibility of locally strong horizontal divergence in zone requiring the implicit scheme to turn on at a lower vertical Courant number to maintain numerical stability. This is written as

$$\alpha^H = \begin{cases} \Delta t\left\{\left[\dfrac{\max(u_{i+1,j,k-1},0) - \min(u_{i,j,k-1},0)}{\Delta x}\right] + \left[\dfrac{\max(v_{i,j+1,k-1},0) - \min(v_{i,j,k-1},0)}{\Delta y}\right]\right\} & \text{for} \quad \omega_{k-1/2} \le 0 \\ \Delta t\left\{\left[\dfrac{\max(u_{i+1,j,k},0) - \min(u_{i,j,k},0)}{\Delta x}\right] + \left[\dfrac{\max(v_{i,j+1,k},0) - \min(v_{i,j,k},0)}{\Delta y}\right]\right\} & \text{for} \quad \omega_{k-1/2} > 0 \end{cases}, \tag{7}$$

where (7) computes the horizontal Courant number from the grid level upwind of the vertical transport direction. As α^H increases, $\alpha^\star_{\max}$ and $\alpha^\star_{\min}$ decrease as indicated in (6). This stabilizes the scheme in locations where α^H is large enough to affect the vertical stability. Equations (6) and (7) are equivalent to S15's (3.11) and (3.7), respectively. Using the definitions given by (5) and (6), the partitioning function $g(\alpha, \alpha^*_{\max}, \alpha^*_{\min})$ is

$$g = \begin{cases} 1 & \text{if} \quad \alpha \le \alpha^\star_{\min} \\ \left[1 + \dfrac{(\alpha - \alpha^\star_{\min})^2}{4\alpha^\star_{\max}(\alpha^\star_{\max} - \alpha^\star_{\min})}\right]^{-1} & \text{if} \quad \alpha^\star_{\min} < \alpha \le 2\alpha^\star_{\max} - \alpha^\star_{\min}, \\ \dfrac{\alpha^\star_{\max}}{\alpha} & \text{if} \quad \alpha > 2\alpha^\star_{\max} - \alpha^\star_{\min} \end{cases} \tag{8}$$

where $\alpha = \alpha_{k-1/2}$. Equation (8) is equivalent to the inverse of S15's (3.13). Nondimensional vertical Courant numbers, partitioned on the basis of (8) with $\alpha_H = 0$, are given in Fig. 2. The blending region between $\alpha^*_{\min}$ and $(2\alpha^*_{\max} - \alpha_{\min})$ is designed to provide a smooth, rather than abrupt, transition between the contributions from explicit and implicit transports. When $\alpha_{k-1/2} < \alpha^{\star}_{\min}$, only the explicit transport is active. Beginning at $\alpha^{\star}_{\min}$, both explicit and implicit transport contributions are used. When the vertical Courant number exceeds $2\alpha^{\star}_{\max} - \alpha^{\star}_{\min}$ the transport tendency from the explicit scheme remains fixed while the remaining vertical transport contribution is obtained from the implicit scheme.

In the WRF implementation $\varepsilon = 0.9$ is used for multidimensional simulations, and this is the same value as used in S15. The RK3 transport scheme with a fifth-order interpolation is linearly stable for $\alpha \leq 1.43$ (Wicker and Skamarock 2002; Baldauf 2008). Experience suggests that IEVA requires a somewhat smaller $\alpha_{\max}$ to consistently generate stable integrations within full physics configurations. For simplicity, all linear stability analyses do not use the partitioning function g, while all of the numerical experiments use g with $\alpha_{\max} = 1.1$ and $\alpha_{\min} = 0.8$.

c. IEVA conservation and consistency

It is important that the IEVA transport scheme maintain conservation and produce solutions consistent with the mass field in WRF, especially for the scalar constituents. Consistency for scalar transport requires that a constant scalar mixing ratio ϕ remain constant for all time (i.e., the discrete scalar equation should collapse to the discrete mass conservation equation in this case). Scalar transport in WRF has this property, and to demonstrate that the IEVA scheme maintains that property, first consider the dry-air mass conservation equation in WRF:

$$\frac{\partial \mu_d}{\partial t} = -\nabla_\eta \cdot (\mu_d \mathbf{V}_h) - \frac{\partial(\mu_d \omega)}{\partial \eta} = -\nabla_\eta \cdot \mathbf{V}_h - \frac{\partial \Omega}{\partial \eta}. \tag{9}$$

The change in μ_d is diagnosed from the vertical integration of (9):

$$\mu_d^{t+\Delta t} = \mu_d^t - \Delta t \int_{\eta=1}^{0} \nabla_\eta \cdot \mathbf{V}_h \, d\eta, \tag{10}$$

where the term involving Ω in (9) drops out of the integral because of the boundary conditions $\Omega = 0$ at $\eta = (1, 0)$. The vertical mass fluxes Ω at intermediate levels are subsequently recovered by reintegrating (9). Because the covariant mass flux Ω is diagnosed from (9) after the integration of (10), no IEVA splitting of the vertical advection in the dry-air continuity equation is needed in the mass-coordinate system.

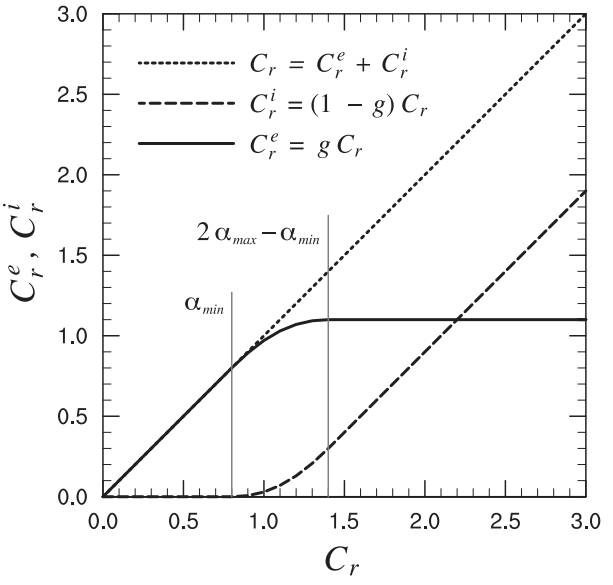

FIG. 2. Vertical Courant number partitioning based on the partitioning function $g(C_r, \alpha_{\min}, \alpha_{\max})$ and $\alpha_H = 0$ hence $\alpha = \alpha^*$. For all results presented in this paper and as depicted in the figure, $\alpha_{\min} = 0.8$ and $\alpha_{\max} = 1.1$.

We can recast (1) in terms of a scalar mixing ratio ϕ:

$$\frac{\partial(\mu_d \phi)}{\partial t} = -\nabla_\eta \cdot (\mathbf{V}_h \phi) - \frac{\partial(\Omega \phi)}{\partial \eta}. \tag{11}$$

As stated in section 2b, IEVA integration is only applied on the last stage of the Runge–Kutta integrator (from $n \rightarrow n + 1$) in (2c) where (11) are discretized as

$$(\mu_d \phi)^{n+1} = (\mu_d \phi)^n - \Delta t\{\delta_x (U\tilde{\phi})^n + \delta_y (V\tilde{\phi})^n + \delta_\eta (\Omega_e \tilde{\phi})^n + \delta_\eta [\Omega_i^n (\overline{\phi}^\eta)^{n+1}]\}. \tag{12}$$

One solves (12) for $(\mu_d \phi)^{n+1}$ using a tridiagonal solver. As previously discussed, the interpolation $\overline{\phi}^\eta$ could be either spatially centered or upwind biased, and both interpolations result in the tridiagonal system and both are conservative. Since the vertical mass transport has been separated into explicit and implicit pieces in (12), we require

$$\Omega^n = \Omega_e^n + \Omega_i^n. \tag{13}$$

The formulation (12) and the requirement (13) are sufficient to guarantee conservation and consistency in the scalar integration in the absence of shape-preserving (monotonic) or positive-definite (PD) flux limiting in WRF.

Conservation and consistency are achieved using the existing positive-definite and shape-preserving (monotonic) renormalizations currently available in WRF as follows. First, the scalar equation in (14) is advanced employing the PD or shape-preserving flux limiters using only the explicit mass fluxes:

$$(\mu_d\phi)^{**} = (\mu_d\phi)^n - \Delta t[\delta_x(U\tilde{\phi})^n + \delta_y(V\tilde{\phi})^n + \delta_\eta(\Omega_e\tilde{\phi})^n]. \tag{14}$$

Following Easter (1993), we use the mass fluxes employed in (14) to perform the corresponding intermediate update of the column mass:

$$\begin{aligned}\mu_d^{**} &= \mu_d^n - \Delta t(\delta_x U^n + \delta_y V^n + \delta_\eta \Omega_e^n)\\ &= \mu_d^{n+1} + \Delta t\delta_\eta \Omega_i^n.\end{aligned} \tag{15}$$

Note that, for any nonzero values of Ω_i^n, μ_d^{**} will be a 3D variable as opposed to the 2D variables μ^n and μ^{n+1} in (10). This intermediate mass μ_d^{**} must be used to decouple the predicted mass from the scalar mass to recover the mixing ratios, both in the intermediate upwind update that is part of the flux-limiting procedure and in the final explicit update. Thus, with the intermediate value of μ_d^{**} known, the final explicit updated values ϕ^{**} are computed as

$$\phi^{**} = (\mu_d\phi)^{**}/\mu_d^{**}. \tag{16}$$

The implicit (and final) update for the scalar mixing ratio is computed as

$$(\mu_d\phi)^{n+1} = (\mu_d\phi)^{**} - \Delta t\delta_\eta[\Omega_i^n(\overline{\phi}^\eta)^{n+1}], \tag{17}$$

and the final value of the mixing ratio is

$$\phi^{n+1} = (\mu_d\phi)^{n+1}/\mu_d^{n+1}. \tag{18}$$

If the values of $\overline{\phi}$ used in the flux divergence are the upwind values, then the PD and shape-preserving properties of the explicit update are preserved in the final implicit update and thus for the overall scheme.

d. Dynamical variable update

The scalar equation in (12) uses horizontal and vertical mass fluxes that are time-averaged over the acoustic steps in order to maintain a consistent solution between the density μ_d and the scalar mass. For the transport of the WRF dynamical variables (u, v, w, θ), these mass fluxes are not available. As in the implementation for scalar transport, we apply the IEVA scheme only to the final RK3 substep (2c), and to maintain consistency for the RK-step dynamics transport we use the following procedure. We begin by calculating a new density $\tilde{\mu}_d^{n+1}$ from the vertically integrated horizontal mass fluxes from the previous RK substep (2b):

$$\tilde{\mu}_d^{n+1} = \mu_d^n - \Delta t\int_{\eta=1}^{0} \nabla_\eta \cdot \mathbf{V}_h^{**}\, d\eta. \tag{19}$$

The implicit components of the vertical flux Ω_{i**}^n are reformulated using (19) and the requirement in (13):

$$\Omega_{i**}^n = \omega_i^n(\mu_d^{n+1}) = \tilde{\omega}_i^n(\tilde{\mu}_d^{n+1}), \tag{20}$$

and the IEVA transport for the dynamical variables is

$$\begin{aligned}(\tilde{\mu}_d^{n+1})\psi^{n+1} = (\mu_d^n)\psi^n - \Delta t\nabla_\eta \cdot (\mathbf{V}_h^{**}\psi^{**}) - \Delta t\delta_\eta[\Omega_e^n(\tilde{\psi})^{**}]\\ - \Delta t\delta_\eta[\tilde{\omega}_i^n(\tilde{\mu}_d^{n+1})(\tilde{\psi})^{n+1}],\end{aligned} \tag{21}$$

where $\psi = (u, v, w, \theta)$. If ψ is constant, (19)–(21) guarantee consistency of the large-time-step flux divergence (i.e., there are no spurious sources or sinks arising within the ψ update from the IEVA scheme). Note that solving (21) for the contravariant vertical velocity w requires incorporating the correct upper and lower boundary conditions into the tridiagonal solution.

3. Stability analyses and advection tests

Stability analyses for 1D and 2D advection, along with results from advection tests, are presented in this section. A stability analysis of the compressible time-split integration is also presented. A more detailed description of the stability analyses for the transport problem can be found in appendix A, and that for the compressible system can be found in appendix B. For the one- and two-dimensional numerical examples, the IEVA scheme is used for all three RK3 substeps.

a. 1D advection

The one-dimensional stability analysis sets the horizontal Courant number to zero in order to analyze only vertical transport using IEVA. As previously stated for the stability analyses, $\alpha_{\min} = \alpha_{\max} = 1.1$ (i.e., the implicit vertical Courant number C_r^i remains zero until the vertical Courant number exceeds 1.1, and then the explicit vertical Courant number C_r^e is held fixed at 1.1 while the implicit contribution increases). Figure 3 shows the amplification factors for the explicit RK3 scheme and RK3-IEVA. As previously noted RK3 is stable for $|C_r| < 1.43$. Figure 3b shows that RK3-IEVA is stable beyond the explicit stability limit, although with increased

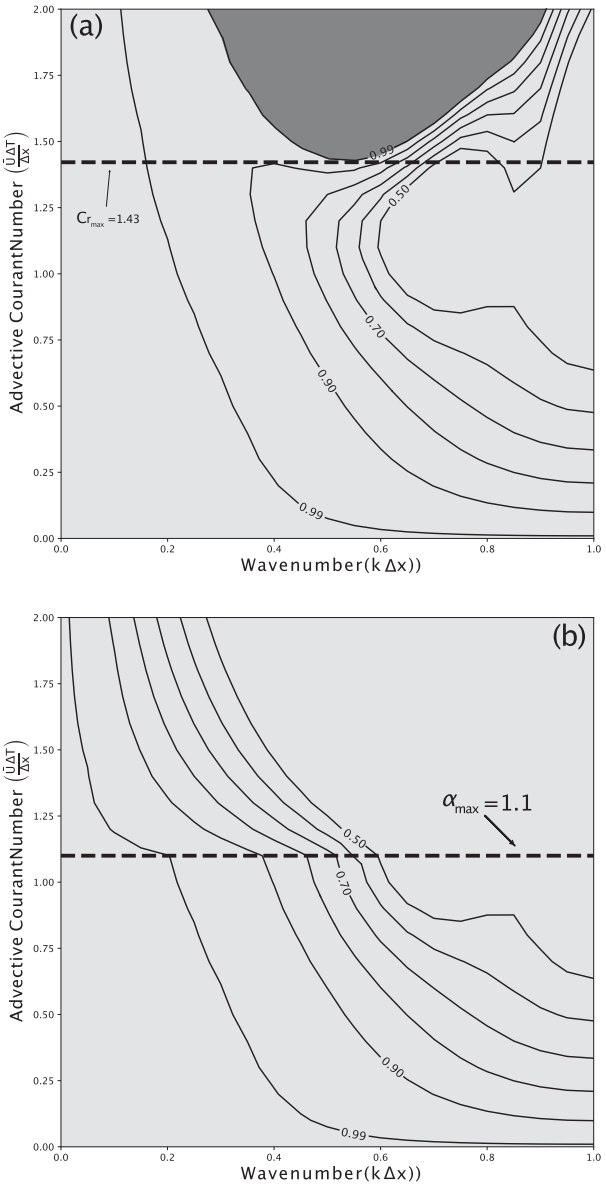

FIG. 3. Linear stability analysis for one-dimensional advection and the fifth-order flux. Light-gray-shaded regions with black contours are regions where the amplification matrix is less than 1.0, and regions that are darkly shaded indicate where the amplification matrix is > 1.0001 and the scheme is unstable. Contour values of amplification matrix values are labeled for values of 0.5, 0.7, 0.9, and 0.99. (a) Explicit RK3 time stepping. Maximum stability for RK3 is indicated by the horizontal dashed line. (b) Combined RK3-IEVA time stepping using first-order upwind differences for the implicit fluxes. The transition to implicit differencing is indicated by the horizontal dashed line.

damping as the C_r exceeds 1.1 due to the use of upwind interpolation within implicit time scheme. For linear problems RK3-IEVA is unconditionally stable as the vertical Courant number approaches (or exceeds) 2, but damping of the shortest wavelengths also increases monotonically. This damping is clearly seen in the linear advection solutions shown in Fig. 4. The solutions are from the advection of a smooth pulse (Wicker and Skamarock 2002) using several different Courant numbers. Grid resolution is moderate (50 points) such that the pulse is approximated across 20 Δx. The sharp gradients on the edges of the pulse are confined to 3–4 grid zones. When the Courant number is 0.8, the RK3 and RK3-IEVA solutions are identical (both solutions are plotted). As the Courant number increases beyond 1.1, the damping from the implicit upwind transport increases. However, even at Courant numbers >2, the numerical solution still represents the transport of the largest scales from the pulse while remaining absolutely numerically stable.

b. 2D advection

For the stability analysis of two-dimensional (x, z) flows, RK3-IEVA is only used for the vertical (z) transport. The maximum eigenvalues for each (x, z) pair of Courant numbers for the original WRF-RK3 scheme and the RK3-IEVA scheme are shown in Fig. 5. The WRF-RK3 results are identical to those reported in Baldauf (2008) where, for stability, the sum of both Courant numbers must be less than the maximum stability value for the one-dimensional case. For RK3 this means $|C_r^u| + |C_r^w| \leq 1.43$, and this is represented by the diagonal line (indicating the amplification factor $|\lambda| = 1$) between shaded and unshaded regions in Fig. 5 where the unshaded (white) region is the stable region for RK3. The region of instability for the RK3-IEVA is shown in the gray-hatched shaded region. As expected, the stability of the two-dimensional advection is now independent of the vertical Courant number and only depends on the horizontal Courant number.

Results from the two-dimensional Blossey–Durran test problem (Blossey and Durran 2008) using a time varying swirling flow is shown in Fig. 6 for both the RK3 and RK3-IEVA. Positive-definite fluxes are used. The 50×50 analytical solution (black contours) and the finite difference solution (filled contours) are shown in Fig. 6a after one full period. Maximum horizontal Courant numbers for the 50×50 grid are approximately 0.7. The numerical solution (filled contours) is similar to the Blossey and Durran (2008) results where the scalar field, initially a circular cone, is stretched out and contracted along the time-varying flow's principal axis of dilatation. The distortion of the final solution is reduced as the resolution is increased. Using RK3-IEVA on the 50×50 grid (solution not shown) yields an essentially identical solution, as the implicit transport is never activated. Increasing the vertical grid resolution by a factor

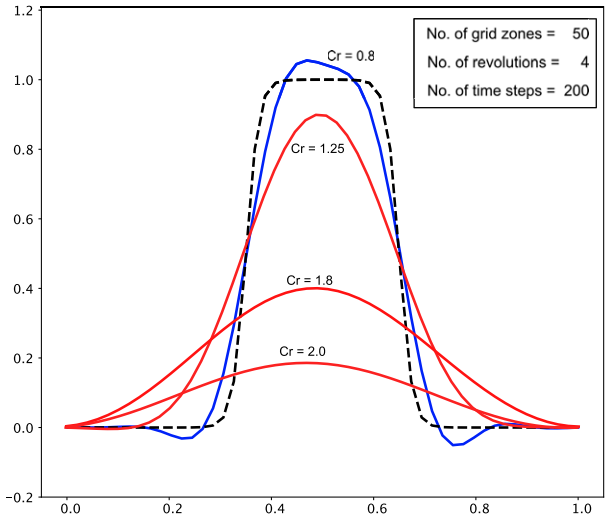

FIG. 4. Linear 1D advection of a Gaussian pulse. The dashed line is the analytical solution. The blue line shows both the RK3 and RK3-IEVA solutions for $C_r = 0.8$, which are identical to round-off. Red lines are the solutions of RK3–5 with IEVA for larger Courant numbers.

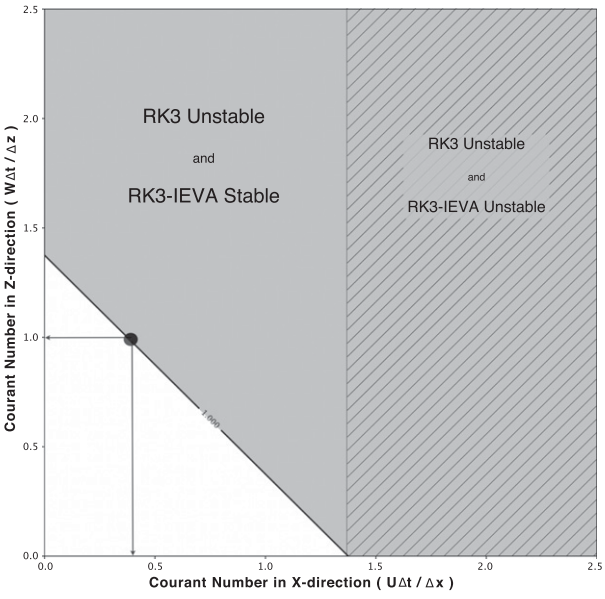

FIG. 5. Linear stability analysis for two-dimensional advection using fifth-order fluxes. The horizontal axis is the x Courant number, and the vertical axis is the z Courant number. Unshaded (white) regions are where both schemes (RK3 and RK3-IEVA) are stable. Shaded regions show where the RK3 time-split scheme is unstable and RK3-IEVA is stable. The shaded-plus-hatched region is where the both schemes are unstable.

of four (depicted by the density of the horizontal gray lines) is shown in Fig. 6b. The computational grid is now 50×200 with maximum vertical Courant numbers of ~2.8. The solution remains stable and accurate, with an slightly increased maximum peak value due to the finer resolution in along the vertical axis.

c. *Stability analysis of the time-split system*

We analyze the stability of the RK3 solver that employs the RK3-IEVA transport scheme and horizontally explicit vertically implicit splitting for the acoustic mode integration of the compressible nonhydrostatic equations. The analysis methodology follows that of Baldauf (2010, hereinafter B10) using the 2D linearized equations from Skamarock and Klemp (1992). The effects of buoyancy in the equations are excluded from this analysis. Further details are given in appendix B.

Figure 7 shows analysis results for the original RK3 integration scheme and for the scheme employing RK3-IEVA transport. The results are for a constant horizontal $C_r^u = 0.43$, which can be thought of as a two-dimensional slice from a three-dimensional cube where the three axes are the vertical, horizontal, and sound-wave Courant numbers. Each point in the cube represents the maximum eigenvalue for all possible spatial wavenumbers. As in the previous section, $\alpha_{\max} = 1.1$ and $\alpha_{\min} = 0.8$. As in Fig. 5, the region of white space (unshaded) below the horizontal line where $C_r^{we} \sim 1$ is the region where the RK3 scheme time-split scheme is stable. All gray and gray-hatched areas are unstable for the RK3 scheme. This stability boundary matches the values seen in Fig. 5a where for the chosen horizontal Courant number ($C_r^u = 0.43$) the maximum stable vertical Courant number is ~1 (see the horizontal and vertical arrows in Fig. 5). The region of instability associated with large sound-wave Courant numbers ($C_s > 0.9$) does not impact the choice of integration parameters and is very similar to the RK3 stability analyses shown in Wicker and Skamarock (1998, 2002) and B10. For RK3-IEVA, the region of stability is the region, white and gray, left of the gray-hatched region where $C_s > 0.9$ (Fig. 7). Similar to the pure 2D advection results, the stability is only constrained by the horizontal Courant number and the sound-wave Courant number. Other slices show the similar stability structure for various horizontal Courant numbers.

To help to increase computational efficiency, the 2D stability analysis is repeated where the implicit transport is only included on the last stage of the RK3 scheme. This would eliminate the tridiagonal solver during the first two stages. The explicit transport in the first two stages only uses the explicit piece of the vertical velocity. The stability analysis (not shown) is similar to that shown in Fig. 7 (i.e., the stability is not limited by the vertical Courant number). Therefore the RK3-IEVA scheme is stable when combined with the traditional

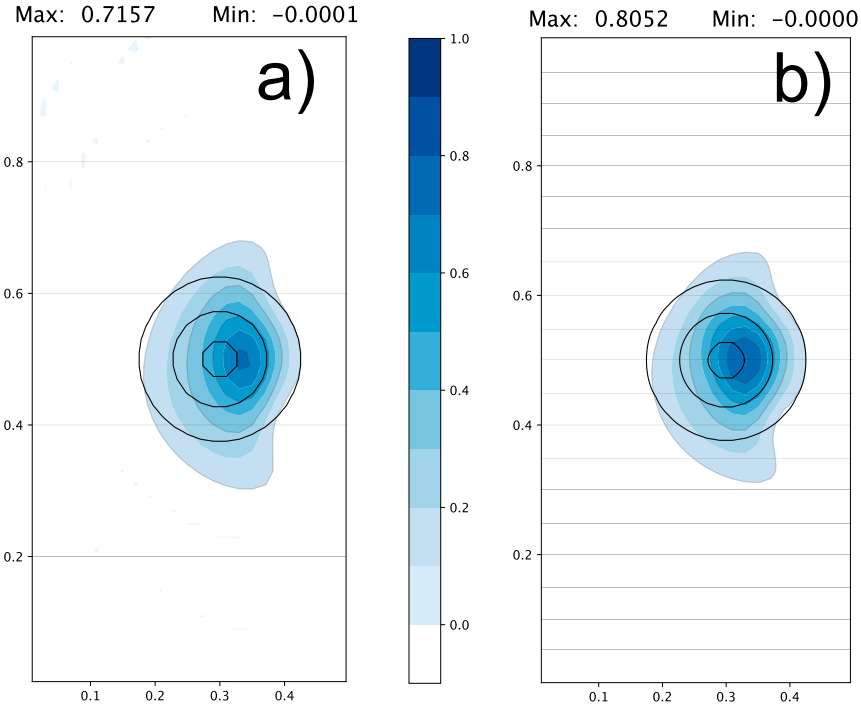

FIG. 6. Two-dimensional flow tests using the Blossy–Durran problem. The solutions are shown at $T = 4\pi$, and the analytical solution is depicted with solid black contours. The vertical grid resolution is shown using horizontal gray lines that are 10 y grid points apart. Positive-definite flux limiting is used in both solutions. Shown are the (a) RK3 scheme, with 50×50 grid and maximum $C_r = 0.7$, and (b) RK3-IEVA with a 50×200 grid and the maximum $C_r = 2.8$ in the y direction only.

time-splitting integration method used in WRF, and the numerical stability associated with large vertical Courant numbers can be achieved with minimal computational expense.

4. 27 April 2011 test case

WRF, version 4.1.1, has been modified to incorporate the IEVA algorithm.[2] To test the WRF IEVA implementation, forecasts of a severe weather event that took place over the south-central and southeastern United States on 26–27 April 2011 are performed. Over 400 tornadoes were reported during this very severe event, and over 300 deaths occurred. Several rounds of intense convection developed in eastern Texas, starting around 0000 UTC 27 April and moved eastward toward Georgia through 1200 UTC 28 April. During the evening of 26 April, strong convection developed in north-central into northeastern Texas. Convective available potential energy values in eastern Texas approached 4000 J kg^{-1} with 0–3-km storm-relative helicity approaching 400 m^2 s^{-2}. This environment supported numerous isolated supercells that produced tornadoes and other severe weather reports (Fig. 8a) from eastern Texas while moving into Louisiana and Mississippi overnight. The next day, two distinct rounds of convection with significant tornadoes occurred in Missisippi, Tennessee, and Alabama (Fig. 8b). This case is chosen to test the IEVA scheme when extremely intense supercells and strong updrafts are present in the forecasts. A much more detailed description of the outbreak can be found in Knupp et al. (2014).

Experimental CAM model forecasts for this case (Clark et al. 2013) generated long-lived supercell storms having intense rotation. Thus this case is ideal to test the IEVA scheme for its ability to represent the intensity of the convection while maintaining stability using a larger time step and none of the vertical velocity filters that have been used in, for example, the HRRR. Our discussion of the test results will focus on the temporal periods associated with strong convection within the parameters of the run [e.g., 0000–0300 UTC (the evening storms from 26 April) and 2000–2200 UTC during the second round of convection the afternoon of 27 April].

[2] The modified routines, initial conditions, and namelists used for these experiments can be found online (https://github.com/louiswicker/IEVA_Paper).

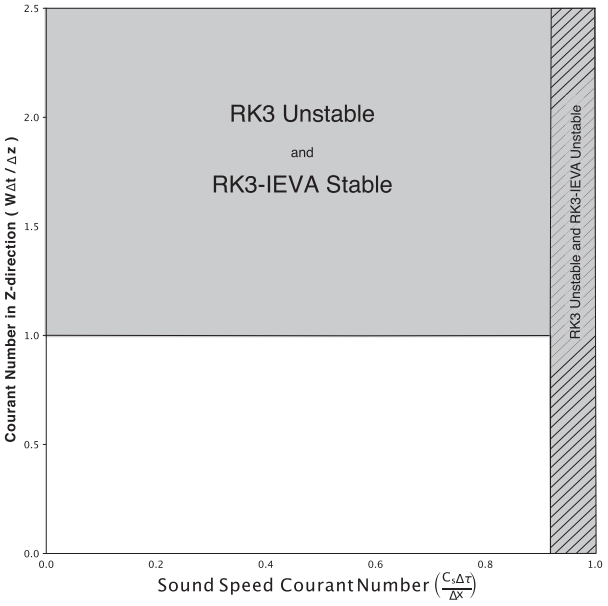

FIG. 7. Stability analysis for the two-dimensional time-split *U*–*W*–*P* system using fifth-order spatial interpolation for the advection and six acoustic time steps. The *W* Courant number is shown along the vertical axis. The analysis that is shown uses a horizontal Courant number of 0.43 to demonstrate time-split stability for 2D advection. Unshaded (white) regions are where both schemes (RK3 and RK3-IEVA) are stable. Shaded regions show where the RK3 time-split scheme is unstable and RK3-IEVA is stable. The shaded-plus-hatched region is where the both schemes are unstable. For reference, the horizontal Courant number (0.43) is depicted by the vertical arrow in Fig. 5.

The WRF Model is configured with a domain similar to the HRRR model, with 1601 × 1201 grid points (e.g., the contiguous U.S. domain) and 51 vertical levels. The horizontal grid spacing is 3 km, and the pressure at the top of the domain is 50 hPa. The initial and 3-hourly boundary conditions are from the North American Mesoscale (NAM) model and are obtained from NOAA's National Center for Environmental Information. Fluxes are computed using a fifth-order interpolant for all horizontal and vertical transport for all variables.[3] WRF is configured to use the NSSL two-moment microphysics scheme (Mansell et al. 2010), the Mellor–Yamada–Nakanishi–Niino (MYNN) planetary boundary layer scheme (Nakanishi and Niino 2004, 2006), and the Noah land surface model (Mitchell 2005). This WRF configuration is similar to the operational HRRR aside from the more complex microphysical parameterization and the use of the NOAH-mp land surface model. To increase computational efficiency, the RK3-IEVA is activated only on the last RK3 substep. The forecasts are essentially identical to those in which IEVA is run on each substep. Forecasts are initialized starting at 0000 UTC 27 April 2011 and are run for 24 h.

Table 1 lists the experiments performed for the case study, and the experiments are named according to the large time step used in the experiment (ranging from 12 to 25 s), and whether the operational HRRR model's two updraft filters or the IEVA algorithm are activated. The R12 forecast solution is stable for a time step of 12 s with no *w* damping or latent heat limiting. When the time step is increased to 15 s, an R15 forecast becomes numerically unstable after the first hour as updraft speeds approach 55 m s^{-1} and vertical Courant numbers >2 occur. The R20_WFILT run uses the latent-heating limiter and *w*-damping filter (as in the operational HRRR). These filters enable the large time step to be increased to 20 s while maintaining numerical stability. Increasing the time step any further makes the forecast unstable. With the latent-heating limiter and *w*-damping filter off and the IEVA scheme turned on, the large time step can be increased to as large as 25 s. Time steps larger than 25 s become unstable. Since results from the $\Delta t = 24$ s and $\Delta t = 25$ s are nearly identical (not shown), the analysis will show results from the $\Delta t = 24$ s experiment. This will be referred to as the R24_IEVA simulation.

The $z = 1$ km instantaneous reflectivity from the three forecasts after 3 h of integration from 0000 to 0300 UTC is shown in Fig. 9. Large-scale reflectivity patterns are essentially identical, with a squall line extending from central Arkansas southwestward into northern Louisiana. For example, all three forecasts suggest a mesoscale vortex is present in the northeast portion of the squall line east of Little Rock, Arkansas. Isolated supercells are also present in east central Texas consistent with the severe weather reports shown in Fig. 8a. Finer-scale details of individual supercells do vary across the experiments, but the timing and locations of the storms are very similar. Differences seen in the magnitude of the maximum reflectivity values are insignificant as they likely represent transient effects. The one-hour maximum updraft fields in Fig. 10 show considerable differences between the three simulations. The R20_WFILT has about one-half of the maximum relative to the reference (R12) and R24_IEVA experiments, and this is most apparent within the Texas supercells. The updrafts in the central Arkansas squall-line are similar in magnitude, with R20_WFILT only slightly weaker, for all the simulations. This suggests that the updraft speeds and the associated latent heating are small enough such that the w-filters are not activated in these storms very often. One-hour maximum updraft helicity (UH) is shown in Fig. 11. Since this field is strongly related to

[3] The HRRR configuration uses third-order interpolation for the vertical fluxes.

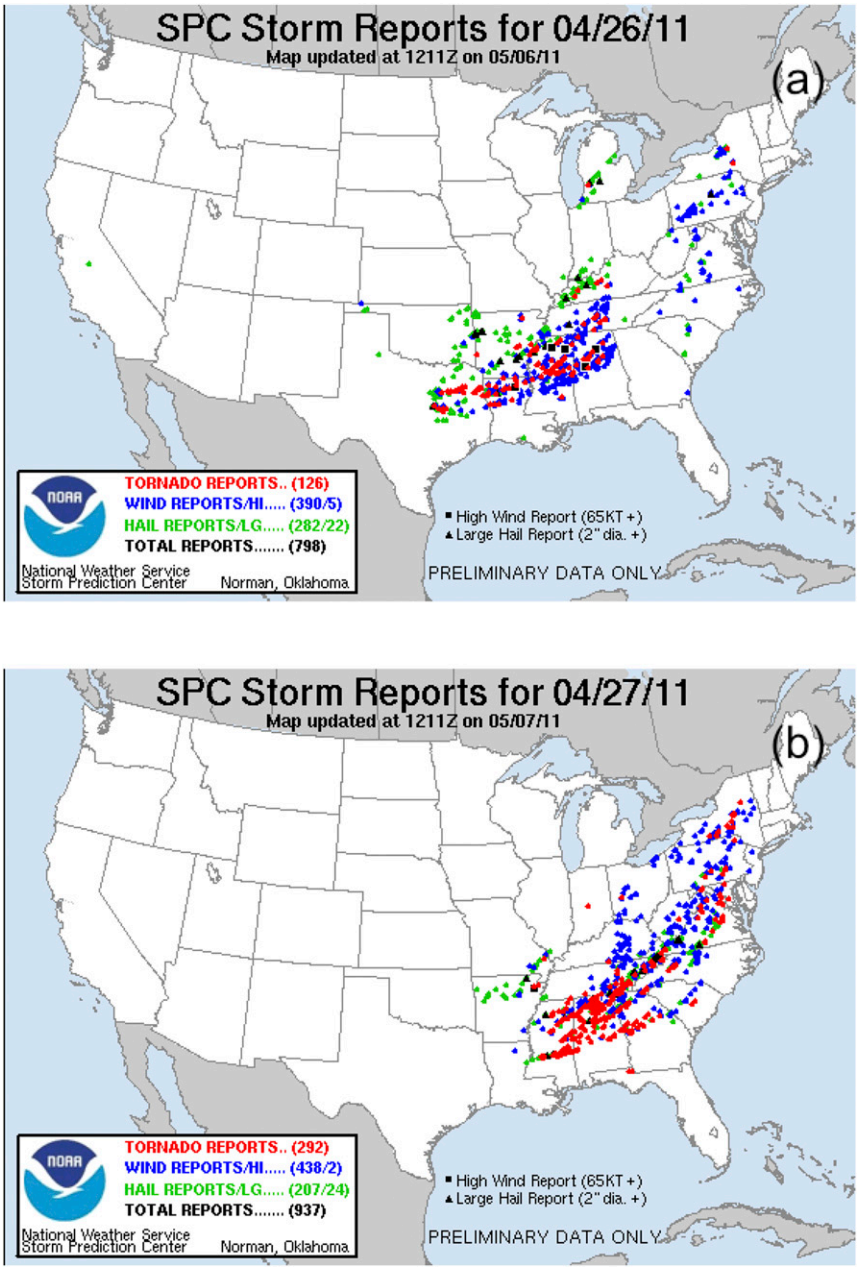

FIG. 8. Severe weather reports for 26–27 Apr 2011 from the Storm Prediction Center's database: reports for the 24-h period ending at 1200 UTC (a) 27 and (b) 28 Apr 2011.

the vertical motion field, significant differences between the R20_WFILT forecast and the other forecasts are expected. Maximum UH values exceed $500\,m^2\,s^{-2}$ in both the R12 and R24_IEVA while being only about $250\,m^2\,s^{-2}$ in the R20_WFILT forecast. The impacts from weaker updrafts within the R20_WFILT includes a reduction of the stretching of vertical vorticity and a reduction of the dynamical effects from updraft rotation that could impact the longevity of the individual storms. Similar updraft, UH, and reflectivity results are seen in the 2000–2300 UTC forecasts (not shown here).

Figure 12 shows the composite profile of vertical velocity for all storms where the column $W_{max} \geq 10\,m\,s^{-1}$ for two time periods, 0100–0300 and 2000–2200 UTC, from each of the experiments. The R20_WFILT updrafts are significantly weaker above 500 hPa for both time periods. Above 500 hPa (approximately 5–6 km) is the level where the maximum vertical velocities are

TABLE 1. WRF IEVA experiments and CPU timings.

Run name	Stable?	Updraft filters	CPU 3-h run	Remarks
R12	Yes	None	926 s	—
R15	No	None	—	Unstable after 1 h
R20_WFILT	Yes	mp_tend_limiter = 0.07; w_damp (cfl_thres = 1.2)	579 s	—
R20_IEVA	Yes	None	601 s	IEVA adds ~4% CPU cost
R24_IEVA	Yes	None	522 s	~10% speedup vs R20_WFILT
R25_IEVA	Yes	None	503 s	~13% speedup vs R20_WFILT

occurring in the convection in the simulations. Below this level, the profiles are quite similar. Since the differences are maximized at upper levels, this may explain why larger solution differences are not seen in the z = 1 km reflectivity fields. The R12 and R24_IEVA profiles are quite similar, with a tendency for the R24_IEVA to have slightly stronger updrafts than the reference solution in the layer where the scheme is likely to be activated. Since the implicit vertical advection scheme is more diffusive than the explicit scheme (Fig. 4) this may also impact the latent heating profile. Another impact is that most physical parameterizations, including cloud microphysical parameterizations, are sensitive to their time step. This can lead to systematic differences in

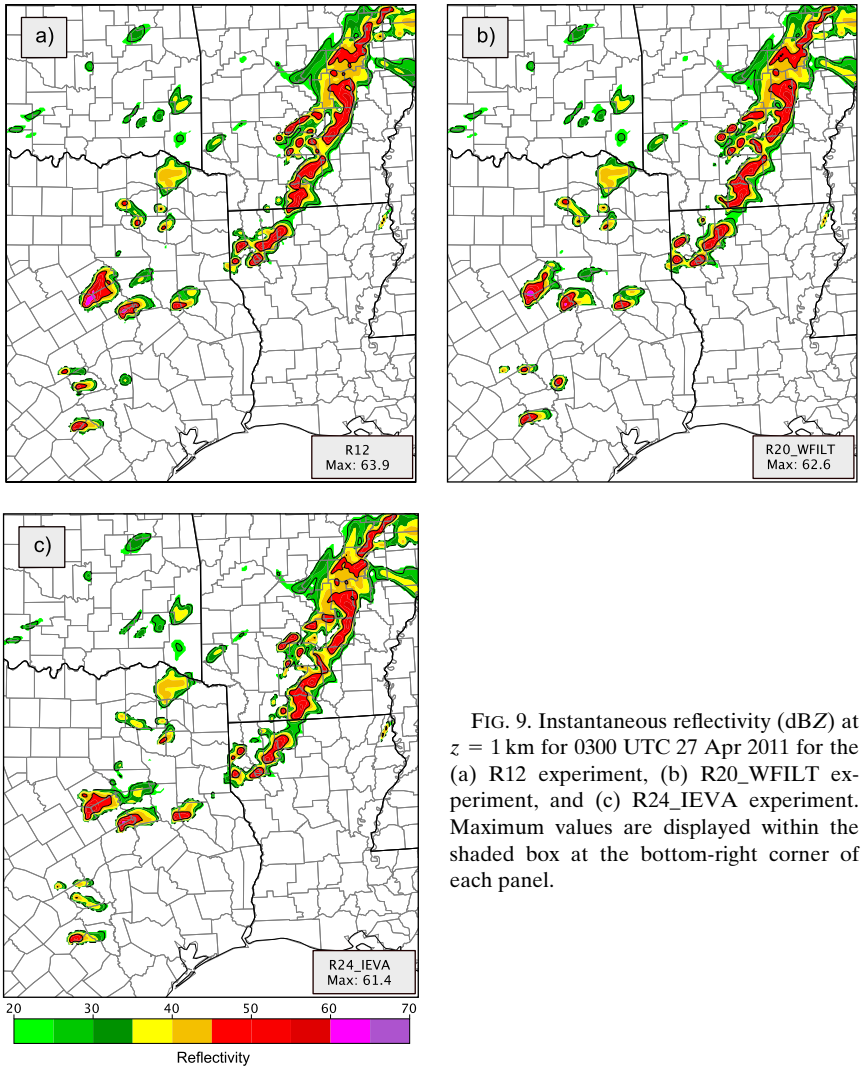

FIG. 9. Instantaneous reflectivity (dBZ) at z = 1 km for 0300 UTC 27 Apr 2011 for the (a) R12 experiment, (b) R20_WFILT experiment, and (c) R24_IEVA experiment. Maximum values are displayed within the shaded box at the bottom-right corner of each panel.

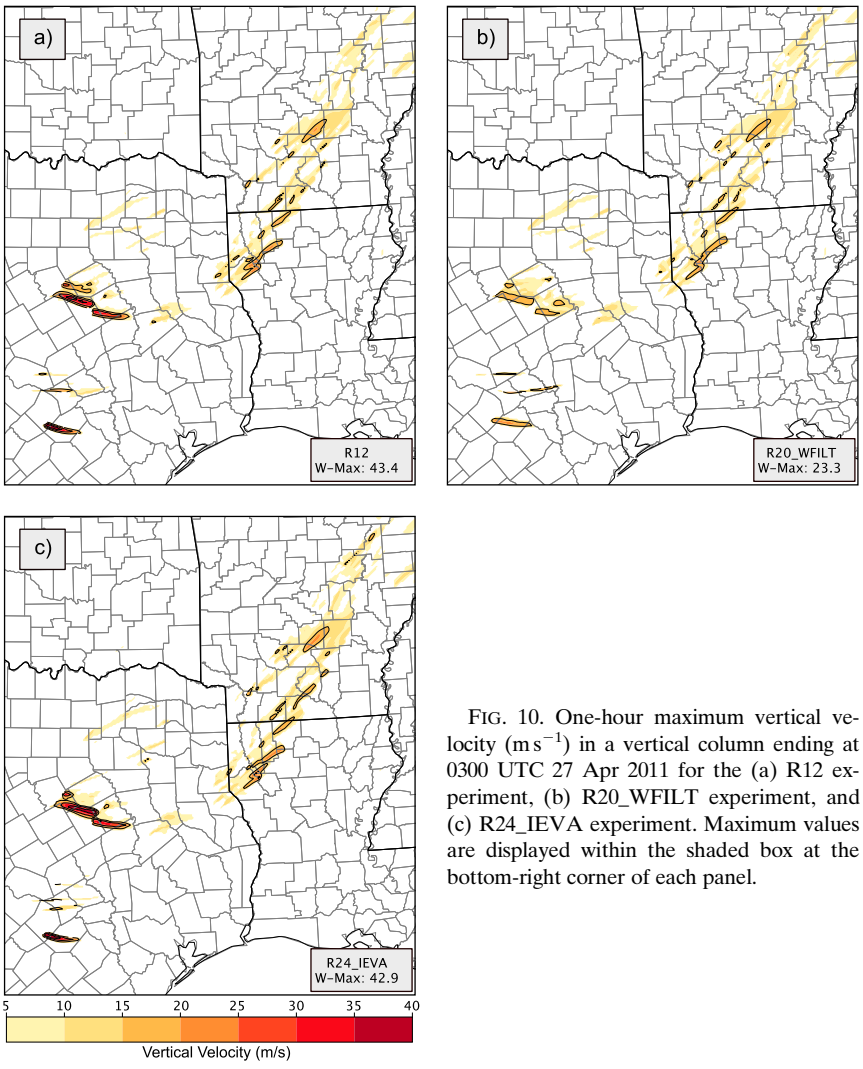

FIG. 10. One-hour maximum vertical velocity ($m\,s^{-1}$) in a vertical column ending at 0300 UTC 27 Apr 2011 for the (a) R12 experiment, (b) R20_WFILT experiment, and (c) R24_IEVA experiment. Maximum values are displayed within the shaded box at the bottom-right corner of each panel.

precipitation and hydrometeor distributions and presumably mean updraft magnitudes (Barrett et al. 2019). Figure 13 is a histogram generated from all three simulations for the two time periods that extracts the maximum updraft in any column that has $W_{max} \geq 10\,m\,s^{-1}$. The distributions for the R12 and R24_IEVA are similar, with a few updrafts exceeding $35\,m\,s^{-1}$ in both experiments. R20_WFILT has the most columns with $W_{max} < 12\,m\,s^{-1}$ and no updrafts larger than $25\,m\,s^{-1}$. The R20_WFILT updraft climatology is very similar to the HRRR and CAPS model forecasts from the CLUE experiment (see Fig. 1).

Table 1 compares 3-h forecast CPU timings from the experiments, plus another simulation using $\Delta t = 20\,s$ with RK3_IEVA to estimate the cost of IEVA algorithm relative to a R20_WFILT. The benchmarks are computed on a Cray XC30 system with Intel IvyBridge processors using 960 cores. The timings are an average of three simulations and do not include I/O time. The CPU time (i.e., wall-clock time) of the control run ($\Delta t = 12\,s$) without any filtering or limiting is 926 s. An unfiltered run using $\Delta t = 15\,s$ becomes unstable after 1 h of integration due to the development of strong updrafts in the Texas storms. Timings from experiments R20_WFILT and R20_IEVA using $\Delta t = 20\,s$ show that the added cost from IEVA is about 4%, mostly due to the cost of solving the tridiagonal system in each column. Using a $\Delta t = 24\,s$ with the R24_IEVA results in a 10% decrease in total wall-clock time relative to R20_WFILT. The R25_IEVA is approximately 13% faster that the R20_WFILT experiment. As the time step increases, is likely that the cost of the tridiagonal solver is increasing because more implicit calculations are needed when a larger time step is used. Nevertheless, an 10%–13% decrease in CPU time can be a significant improvement for real time

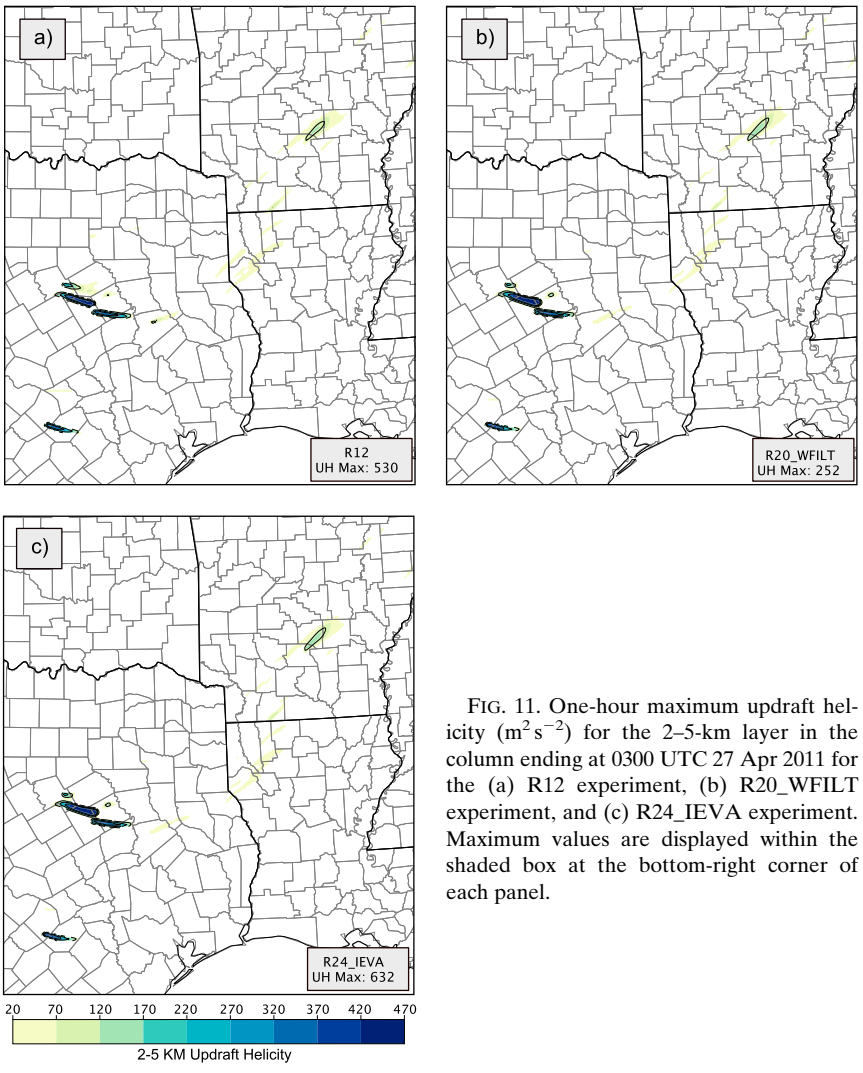

FIG. 11. One-hour maximum updraft helicity ($m^2 s^{-2}$) for the 2–5-km layer in the column ending at 0300 UTC 27 Apr 2011 for the (a) R12 experiment, (b) R20_WFILT experiment, and (c) R24_IEVA experiment. Maximum values are displayed within the shaded box at the bottom-right corner of each panel.

forecasts that must run within prescribed time windows at operational centers.

5. Conclusions

This paper presents an implicit–explicit vertical advection scheme for the WRF Model that enables the model to use a larger time step and allows for the reduction or elimination of two vertical velocity filters used in operational CAM forecasts—limiting the amount of diabatic heating per time step and a Rayleigh damping term applied to large vertical velocities. The IEVA scheme, originally from the UCLA ROMS ocean modeling system (S15), is modified for the Runge–Kutta time-stepping framework used in the WRF Model. For high-resolution atmospheric and oceanic models where strong vertical motions are observed at only a small fraction of points during the integration, IEVA combines the model's existing explicit vertical transport scheme with an absolutely stable implicit transport, and by smoothly blending the schemes stability is retained for vertical Courant numbers greater than that limiting the explicit scheme's stability. This approach greatly increases the numerical stability bounds associated with transport processes in the full model, permits a much larger time step, and the larger time step decreases the computational cost. The WRF-IEVA scheme also has minimal impact on the updraft velocities. In contrast, WRF's latent heat limiter and w-damping filter, used in a number of applications to enable larger time steps while keeping the integration stable, comes at the cost of significantly reducing the maximum updraft velocities that should occur in WRF configurations at CAM resolutions. The IEVA scheme does add cost by requiring

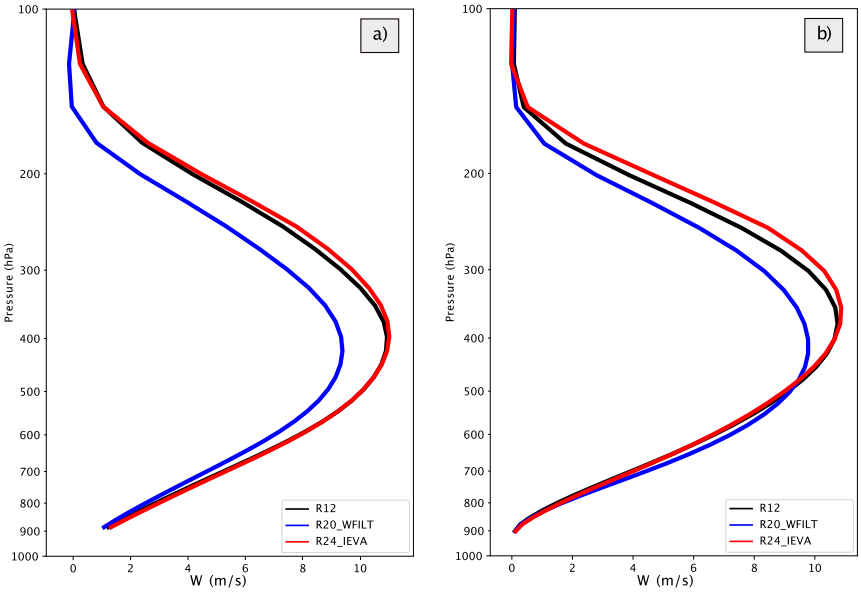

FIG. 12. Profiles of average vertical velocity from updrafts over a 3-h period where the maximum vertical velocity in the column is greater than 10 m s^{-1}. The reference solution is a black line, the WFILT solution is blue, and the IEVA solution is red. Shown are mean updraft profiles (a) from 0100 to 0300 UTC and (b) from 2000 to 2200 UTC.

the solution of a tridiagonal system for each transported variable within each vertical column. To reduce that cost, we show numerical experiments in which the IEVA scheme is activated only on the final substep of RK3 integration with no significant impact to solution stability or accuracy. A set of linear stability analyses demonstrate that IEVA is stable for both pure multidimensional transport as well as the split-explicit/implicit time-splitting scheme used in the WRF Model (W02). The IEVA method is tested in a full physics configuration of the WRF Model using the 27 April 2011 tornadic outbreak case by creating a benchmark simulation using a reduced large time step (Δt = 12 s) with no vertical velocity filters. Using the benchmark forecast as a proxy for the "correct" prediction, solutions from the IEVA and an operational configuration with the filters are compared using typical severe weather parameters as well as aggregated updraft statistics. The forecasts using RK3-IEVA were closer to the benchmark forecast. This was shown using horizontal cross-sections of reflectivity, vertical velocity, and updraft helicity, and profiles of vertical velocity from the RK3-IEVA forecast are closer to the benchmark solution than the filtered forecast. Average vertical velocity profiles and histograms of updraft magnitude are compared across the three experiments. The updraft and storm characteristics from the RK3-IEVA solution are much closer to R12 experiment than the R20_WFILT. These results are consistent from the analysis of Potvin et al. (2019) comparing CAM forecasts with and without the latent heating limiter and vertical velocity filtering (Fig. 1).

Even with a time step that is 2 times that of the benchmark run, convection in the RK3-IEVA forecast more closely reproduces storm characteristics seen in the R12 benchmark simulation than those from the

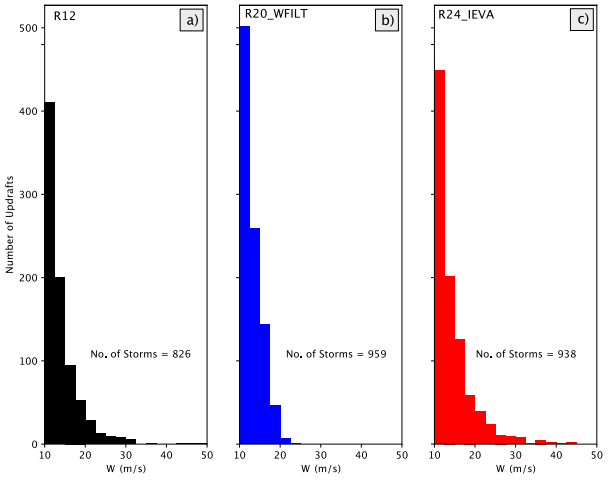

FIG. 13. Combined updraft histograms from 0100 to 0300 and from 2000 to 2200 UTC 27 Apr 2011 for any updraft for which the vertical velocity is greater than 10 m s^{-1} for the (a) R12 experiment, (b) R20_WFILT experiment, and (c) R24_IEVA experiment.

filtered operational configuration. Forecasts using the RK3-IEVA permit a time step that is 20%–25% larger than the filtered experiment, which results in a 10%–13% speedup relative to the filtered experiment's CPU time. The RK3-IEVA scheme is currently implemented in the newest version of the HRRR forecast system (i.e., the HRRRv4) currently scheduled for operational implementation in summer 2020 at NOAA's Environmental Prediction Center. Like the results shown here, RK3-IEVA enables a larger time step with more realistic updraft characteristics, while having no impact on the HRRRv4 system's computational reliability. In the near future, the RK3-IEVA scheme will be used in the ensemble data assimilation system used in the Warn-on-Forecast system (Skinner et al. 2018) and should enable larger time steps to be used despite the initial imbalances present in the analysis increments. This could significantly improve the efficiency in both the analysis cycling as well as the ensemble forecast system time and may permit larger ensembles or higher vertical resolutions to be used. Finally, almost any current modeling system that uses explicit finite difference or finite volume methods for transport should be able to adapt this method to help increase integration efficiency while maintaining reasonable solution accuracy if the transport time step is being limited by the Courant number along a single coordinate direction.

Acknowledgments. This research was funded primarily by the NOAA Warn-on-Forecast project. Additional funding was provided by NOAA/Office of Oceanic and Atmospheric Research under the NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce. The National Center for Atmospheric Research is supported by the National Science Foundation. The authors thank all of the reviewers for their careful review of the paper; Dr. Pedro Peixoto helped correct an error in our original procedure for when IEVA is used only on the last RK stage, and we are indebted to him. The first author thanks Drs. Joseph Klemp and Wei Wang of NCAR/MMM for their encouragement and support of this research over the last six years.

APPENDIX A

2D Stability Analysis for RK3-IEVA Advection

A brief description of the linear stability analyses is presented here for the RK3-IEVA. While the analysis is standard, since the formulation uses the two vertical Courant numbers some details are given. Since each stage of the Wicker and Skamarock (2002) scheme can be written as a series of forward steps, the amplification maxtrix can be written in a concise manner. Starting with a linearized two-dimensional version of (2) and using a two-dimensional wavenumber expansion in Fourier space (Durran 2013), the RK3-IEVA amplification factor for a single forward step is

$$\hat{\phi}^{n+1} = \left(\frac{1 - \lambda_u \frac{C_r^u}{N} - \lambda_{w_e} \frac{C_r^{w_e}}{N}}{1 + \lambda_{w_i} \frac{C_r^{w_i}}{N}} \right) \hat{\phi}^n . \tag{A1}$$

Since $N = 3, 2, 1$, one can combine all three steps into a general expression for the RK3-IEVA analysis system:

$$\begin{pmatrix} \hat{\phi}^{n+1} \\ \hat{\phi}^n \end{pmatrix} = \prod_{N=1}^{3} \begin{bmatrix} \dfrac{-\left(\lambda_u \frac{C_r^u}{N} + \lambda_{w_e} \frac{C_r^{w_e}}{N}\right)}{1 + \lambda_{w_i} \frac{C_r^{w_i}}{N}} & \dfrac{1}{\lambda_{w_i} \frac{C_r^{w_i}}{N}} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \hat{\phi}^n \\ \hat{\phi}^n \end{pmatrix}, \tag{A2}$$

where $\hat{\phi}$ are the solution amplitudes and C_r^u, $C_r^{w_e}$, and $C_r^{w_i}$ are the horizontal, explicit-vertical and implicit-vertical Courant numbers, respectively. The horizontal, explicit-vertical, and implicit-spatial Fourier expansions λ_u, λ_{w_e}, and λ_{w_i} use fifth-, fifth-,and first-order transport operators and are given in (Baldauf 2008). For the eigenvalue analysis, the vertical-splitting algorithm given by (8) is simplified by setting $\alpha_{\min} = \alpha_{\min} \equiv 1$; that is,

$$C_r^u = \frac{|\overline{U}|\Delta t}{\Delta x}$$

$$(C_r^{w_e}, C_r^{w_i}) = \begin{cases} C_r^{w_e} = \dfrac{|\overline{W}|\Delta t}{\Delta x} & \text{for} \quad \dfrac{|\overline{W}|\Delta t}{\Delta x} \le \alpha_{\max}^* \\ C_r^{w_i} = 0 & \\ C_r^{w_e} = \alpha_{\max}^* & \text{for} \quad \dfrac{|\overline{W}|\Delta t}{\Delta x} > \alpha_{\max}^* \\ C_r^{w_i} = \dfrac{|\overline{W}|\Delta t}{\Delta x} - \alpha_{\max}^* & \end{cases}, \tag{A3}$$

with $\alpha_{\max}^* = \alpha_{\max} - C_r^u$. As in (8), when the horizontal Courant number increases the implicit scheme ''kicks in'' at a lower value of the vertical Courant number to maintain the stability. In the extreme case in which

$C_r^u > \alpha_{\max}$, the entire vertical transport is computed implicitly. Eigenvalues are computed from the first term of (A2) using (A3) for each possible combination of horizontal and vertical Courant numbers and spatial wavenumbers ranging from 0 to π using the Python and SciPy software libraries (Virtanen et al. 2020). Eigenvalues are computed on a $31 \times 31 \times 31$ grid. The maximum eigenvalue along the spatial wavenumber dimension is computed for each pair of horizontal and vertical Courant numbers. Sampling the eigenvalues using a $61 \times 61 \times 61$ computational grid yields essentially identical results.

APPENDIX B

2D Stability Analysis for IEVA Time-Splitting Scheme

To examine the stability of RK3-IEVA within a solver for the fully compressible Navier–Stokes system such as that in WRF, the two-dimensional linear system of equations from Skamarock and Klemp (1992) is analyzed for the time-split horizontally explicit and vertically implicit forward–backward scheme used to integrate the system (Wicker and Skamarock 2002; Bryan and Fritsch 2002; Klemp et al. 2007). The RK3 analysis method from B10 is used, and the effects of buoyancy are excluded from the analysis. The IEVA scheme is again applied only to the vertical transport. Following B10, the RK3-IEVA integrator can be written as

$$Q^1_{\mathrm{RK3}} = Q_f^{ns/3} + \frac{1}{ns}\left[\sum_{i=0}^{ns/3-1} (Q_f)^i\right](Q_s - 1), \tag{B1}$$

$$Q^2_{\mathrm{RK3}} = Q_f^{ns/2} + \frac{1}{ns}\left[\sum_{i=0}^{ns/2-1} (Q_f)^i\right](Q_s - 1)Q^1_{\mathrm{RK3}}, \quad \text{and} \tag{B2}$$

$$Q^3_{\mathrm{RK3}} = Q_f^{ns} + \frac{1}{ns}\left[\sum_{i=0}^{ns-1} (Q_f)^i\right](Q_s - 1)Q^2_{\mathrm{RK3}}, \tag{B3}$$

with

$$Q_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & iC_s\lambda_{cz}\beta_1 \\ 0 & iC_s\lambda_{cz}\beta_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1+\gamma_d\lambda_{cx}^2 & \gamma_d\lambda_{cx}\lambda_{cz} & -iC_s\lambda_{cx} \\ \gamma_d\lambda_{cx}\lambda_{cz} & 1+\gamma_d\lambda_{cz}^2 & -iC_s\lambda_{cz}\beta_0 \\ -iC_s\lambda_{cx}(1+\gamma_d\lambda_{cx}^2) & -C_s\lambda_{cz}(\beta_0+\gamma_d\lambda_{cx}^2) & 1+C_s^2\lambda_{cx}^2 \end{pmatrix}$$

and

$$(Q_s - 1) = \begin{cases} -(C_r^u\lambda_u + C_r^{w_e}\lambda_{w_e})(\mathbf{Q}_I) & \text{for} \quad \alpha \le \alpha_{\max} \\ -\dfrac{(C_r^u\lambda_u + C_r^{w_e}\lambda_{w_e} + C_r^{w_i}\lambda_{w_i})}{(1 + C_r^{w_i}\lambda_{w_i})}(\mathbf{Q}_I) & \text{for} \quad \alpha > \alpha_{\max} \end{cases};$$

$\mathbf{Q}_I$ is the 3×3 identity matrix in this analysis. Here, C_s is the sound-wave Courant number (for each small time step), and C_r^u, $C_r^{w_e}$, and $C_r^{w_i}$ are the advective Courant numbers as previously defined in appendix A. The horizontal, vertical-explicit, and implicit-spatial advection Fourier components λ_u, λ_{w_e}, and λ_{w_i} use the fifth-, fifth-, and first-order Fourier representations as in appendix A. The remaining Fourier coefficients for the second-order divergence, pressure gradient, and the divergent damping (on the Arakawa C grid) are

$$\lambda_{cx} = 2i\sin\left(\frac{k\Delta x}{2}\right),$$

$$\lambda_{cz} = 2i\sin\left(\frac{l\Delta z}{2}\right),$$

$$\gamma_d = 0.05, \quad \text{and}$$

$$\beta_1 = 1 - \beta_0 = 0.55;$$

β_1 is the temporal off-centering for the vertically implicit divergence and pressure gradient forces. Eigenvalues are computed on a $(31)^4$ grid for each of the four degrees of freedom. The maximum eigenvalue along the spatial wavenumber dimension is computed for each triplet of horizontal, vertical, and sound-wave Courant numbers.

REFERENCES

Baldauf, M., 2008: Stability analysis for linear discretisations of the advection equation with Runge–Kutta time integration. *J. Comput. Phys.*, **227**, 6638–6659, https://doi.org/10.1016/j.jcp.2008.03.025.

——, 2010: Linear stability analysis of Runge–Kutta-based partial time-splitting schemes for the Euler equations. *Mon. Wea. Rev.*, **138**, 4475–4496, https://doi.org/10.1175/2010MWR3355.1.

——, A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1.

Barrett, A. I., C. Wellmann, A. Seifert, C. Hoose, B. Vogel, and M. Kunz, 2019: One step at a time: How model time step significantly affects convection-permitting simulations. *J. Adv. Model. Earth Syst.*, **11**, 641–658, https://doi.org/10.1029/2018MS001418.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Blossey, P. N., and D. R. Durran, 2008: Selective monotonicity preservation in scalar advection. *J. Comput. Phys.*, **227**, 5160–5183, https://doi.org/10.1016/j.jcp.2008.01.043.

Bryan, G. H., and J. M. Fritsch, 2002: A benchmark simulation for moist nonhydrostatic numerical models. *Mon. Wea. Rev.*, **130**, 2917–2928, https://doi.org/10.1175/1520-0493(2002)130<2917:ABSFMN>2.0.CO;2.

Clark, A. J., J. Gao, P. T. Marsh, T. Smith, J. S. Kain, J. Correia, M. Xue, and F. Kong, 2013: Tornado path length forecasts from 2010 to 2011 using ensemble updraft helicity. *Wea. Forecasting*, **28**, 387–407, https://doi.org/10.1175/WAF-D-12-00038.1.

——, and Coauthors, 2018: The Community Leveraged Unified Ensemble (CLUE) in the 2016 NOAA/Hazardous Weather Testbed spring forecasting experiment. *Bull. Amer. Meteor. Soc.*, **99**, 1433–1448, https://doi.org/10.1175/BAMS-D-16-0309.1.

Durran, D. R., 2013: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*. Springer Science and Business Media, 466 pp.

Easter, R. C., 1993: Two modified versions of Bott's positive-definite numerical advection scheme. *Mon. Wea. Rev.*, **121**, 297–304, https://doi.org/10.1175/1520-0493(1993)121<0297:TMVOBP>2.0.CO;2.

Klemp, J. B., W. C. Skamarock, and J. Dudhia, 2007: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. *Mon. Wea. Rev.*, **135**, 2897–2913, https://doi.org/10.1175/MWR3440.1.

Knupp, K. R., and Coauthors, 2014: Meteorological overview of the devastating 27 April 2011 tornado outbreak. *Bull. Amer. Meteor. Soc.*, **95**, 1041–1062, https://doi.org/10.1175/BAMS-D-11-00229.1.

Laprise, R., 1992: The Euler equations of motion with hydrostatic pressure as an independent variable. *Mon. Wea. Rev.*, **120**, 197–207, https://doi.org/10.1175/1520-0493(1992)120<0197:TEEOMW>2.0.CO;2.

Mansell, E. R., C. L. Ziegler, and E. C. Bruning, 2010: Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J. Atmos. Sci.*, **67**, 171–194, https://doi.org/10.1175/2009JAS2965.1.

Mitchell, K., 2005: The community Noah land-surface model (LSM): User's guide. NOAA Doc., 27 pp., ftp://ftp.emc.ncep.noaa.gov/mmb/gcp/ldas/noahlsm/ver_2.7.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31, https://doi.org/10.1023/B:BOUN.0000020164.04146.98.

——, and ——, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

Potvin, C. K., and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for Warn-on-Forecast. *Mon. Wea. Rev.*, **143**, 2998–3024, https://doi.org/10.1175/MWR-D-14-00416.1.

——, and Coauthors, 2019: Systematic comparison of convection-allowing models during the 2017 NOAA HWT spring forecasting experiment. *Wea. Forecasting*, **34**, 1395–1416, https://doi.org/10.1175/WAF-D-19-0056.1.

Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2019: NCAR's real-time convection-allowing ensemble project. *Bull. Amer. Meteor. Soc.*, **100**, 321–343, https://doi.org/10.1175/BAMS-D-17-0297.1.

Shchepetkin, A. F., 2015: An adaptive, Courant-number-dependent implicit scheme for vertical advection in oceanic modeling. *Ocean Modell.*, **91**, 38–69, https://doi.org/10.1016/j.ocemod.2015.03.006.

——, and J. C. McWilliams, 2005: The Regional Oceanic Modeling System (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modell.*, **9**, 347–404, https://doi.org/10.1016/j.ocemod.2004.08.002.

Skamarock, W. C., and J. B. Klemp, 1992: The stability of time-split numerical methods for the hydrostatic and the nonhydrostatic elastic equations. *Mon. Wea. Rev.*, **120**, 2109–2127, https://doi.org/10.1175/1520-0493(1992)120<2109:TSOTSN>2.0.CO;2.

——, and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast system. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Snook, N., F. Kong, K. A. Brewster, M. Xue, K. W. Thomas, T. A. Supinie, S. Perfater, and B. Albright, 2019: Evaluation of convection-permitting precipitation forecast products using WRF, NMMB, and FV3 for the 2016–17 NOAA hydrometeorology testbed flash flood and intense rainfall experiments. *Wea. Forecasting*, **34**, 781–804, https://doi.org/10.1175/WAF-D-18-0155.1.

Verrelle, A., D. Ricard, and C. Lac, 2015: Sensitivity of high-resolution idealized simulations of thunderstorms to horizontal resolution and turbulence parametrization. *Quart. J. Roy. Meteor. Soc.*, **141**, 433–448, https://doi.org/10.1002/qj.2363.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272, https://doi.org/10.1038/s41592-019-0686-2.

Wicker, L. J., and W. C. Skamarock, 1998: A time-splitting scheme for the elastic equations incorporating second-order Runge–Kutta time differencing. *Mon. Wea. Rev.*, **126**, 1992–1999, https://doi.org/10.1175/1520-0493(1998)126<1992:ATSSFT>2.0.CO;2.

——, and ——, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097, https://doi.org/10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2.