The Effects of Spatial Interpolation on a Novel, Dual-Doppler 3D Wind Retrieval Technique

JORDAN P. BROOK [ORCID],[a] ALAIN PROTAT,[b] COREY K. POTVIN,[c,d] JOSHUA S. SODERHOLM,[b] AND HAMISH MCGOWAN[a]

[a] *Atmospheric Observations Research Group, University of Queensland, Brisbane, Queensland, Australia*
[b] *Science and Innovation Group, Australian Bureau of Meteorology, Docklands, Victoria, Australia*
[c] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*
[d] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

ABSTRACT: Three-dimensional wind retrievals from ground-based Doppler radars have played an important role in meteorological research and nowcasting over the past four decades. However, in recent years, the proliferation of open-source software and increased demands from applications such as convective parameterizations in numerical weather prediction models has led to a renewed interest in these analyses. In this study, we analyze how a major, yet often-overlooked, error source effects the quality of retrieved 3D wind fields. Namely, we investigate the effects of spatial interpolation, and show how the common practice of pregridding radial velocity data can degrade the accuracy of the results. Alternatively, we show that assimilating radar data directly at their observation locations improves the retrieval of important dynamic features such as the rear flank downdraft and mesocyclone within supercells, while also reducing errors in vertical vorticity, horizontal divergence, and all three velocity components.

SIGNIFICANCE STATEMENT: We can attempt to estimate the wind speed and direction within a weather system when two weather radars measure it simultaneously. However, radars do not scan the whole atmosphere at once—instead, they measure along many cross sections, each at different heights. We show that a method commonly used to stitch the observations together degrades the accuracy of the winds. Additionally, we describe a way to feed the data directly into the analysis without stitching it together first, and show that this improves the wind retrievals considerably. We hope these improvements will help researchers better understand how various weather systems work, and help forecasters warn for dangerous weather such as tornadoes.

---

1. Introduction

Three-dimensional wind retrievals from Doppler radar measurements have become an indispensable tool for studying the kinematic properties of various weather systems, such as tropical cyclones, convective storms, and tornadoes (e.g., Kosiba and Wurman 2014; Betten et al. 2018; Markowski et al. 2018). These analyses are also set to play an important role in observationally verifying and improving numerical weather prediction model simulations and parameterizations of convection (e.g., Kumar et al. 2015; Nicol et al. 2015; Labbouz et al. 2018). Additionally, 3D wind retrievals herald exciting opportunities for severe weather forecasting, e.g., for damaging surface winds and updraft intensity identification, or other research applications such as real-time hail trajectory nowcasting and growth modeling (Kumjian et al. 2021; Brook et al. 2021). With these applications in mind, this study aims to investigate how the commonly overlooked effects of spatial interpolation impact current wind retrieval algorithms. Our study also particularly focuses on 3D wind retrievals for operational radar networks (e.g., Bousquet et al. 2007; Dolan and Rutledge 2007; Park and Lee 2009) where the undesirable effects imposed by spatial interpolation are most

severe.[1] We also expect our findings to be useful in retrievals using airborne, or rapid-scan mobile radar data.

It is now the literary consensus that variational 3D wind retrievals outperform those involving numerical integration of the mass continuity equation (Gao et al. 1999; Potvin et al. 2012a; North et al. 2017). These improvements are attributed to well-known methodological shortcomings of the latter, involving ill-posed boundary conditions and vertically compounding errors during numerical integration (Ray et al. 1980). Despite the improvements offered by variational analyses, the retrieval of accurate 3D winds from typical dual-Doppler observations remains a challenging problem. These challenges are greatest for the vertical velocity component, due to the focus on low elevations in typical operational scanning strategies. Under these conditions, the vertical component of the 3D velocity vector is poorly observed, meaning vertical velocity retrievals must rely heavily on the mass continuity constraint. Furthermore, the task is complicated by a number of secondary factors, including spatial interpolation errors, data gaps near the ground and near the storm top,

---

Corresponding author: Jordan Brook, j.brook@uq.edu.au

[1] Spatial interpolation errors are most severe in operational networks due to the large horizontal distances (>50 km) common between radars, and large elevation gaps (5°–6°) between constant elevation scans in operational scanning strategies.

measurement nonsimultaneity, discretization errors, and erroneous Doppler velocity data (e.g., due to incorrect dealiasing, dual-PRF artifacts, sidelobe contamination or ground clutter; Gao et al. 1999; Potvin et al. 2012b). All of these factors contribute to the considerable errors noted in variational 3D wind retrievals, especially in the vertical velocities (Potvin et al. 2012c; Oue et al. 2019).

Gao et al. (1999) summarized that the two most pronounced difficulties for vertical velocity retrievals are 1) data gaps at the vertical domain boundaries and 2) the spatial interpolation of radar information. In this study, we address the impacts of the latter, and aim to assess how errors introduced by spatial interpolation propagate into variational 3D wind retrievals. However, we first pause to note the significant progress made on the first of these issues. Perhaps the most notable contribution was made by Shapiro et al. (2009), who showed that including the vertical vorticity equation as a variational constraint can significantly improve retrievals with data voids near the surface in observing system simulation experiments (OSSEs) of simple analytic fields. Potvin et al. (2012b) consolidated this work by showing this constraint also aids vertical velocity retrievals in low-level data denial OSSEs of simulated thunderstorms, and this finding has been reproduced in more recent studies (Dahl et al. 2019; Gebauer et al. 2022). In practice, the provision of a vertical vorticity constraint requires data from multiple consecutive radar scans to estimate its temporal evolution (e.g., Protat and Zawadzki 2000). This can be achieved to varying degrees of accuracy by estimating a constant horizontal advection vector (Shapiro et al. 2009), or by using a combination of provisional retrievals at successive time steps and spatially varying advection correction techniques (Potvin et al. 2012b; Dahl et al. 2019; Gebauer et al. 2022). These studies, along with others such as Potvin et al. (2012c) and Oue et al. (2019), have shown that large errors can arise when the scanning time of a radar volume approaches that of operational weather radars (~5 min) due to erroneous temporal discretizations and temporal evolution of wind fields. In practice, the interpolation practices discussed here should be used in conjunction with these previous advances.

Spatial interpolation procedures for radial velocity data in 3D wind retrievals may be characterized into two types: 1) pregridded methods and 2) direct radar assimilation methods. The former requires that radar data are interpolated to a common, Cartesian analysis grid prior to the analysis. This simplifies the observational constraint, but introduces significant errors in radial velocities before the variational retrieval has begun (Gao et al. 1999; North et al. 2017; Oue et al. 2019). Gridding methods used in wind retrievals vary from simple nearest neighbor/linear interpolation methods such as that in Mohr and Vaughan (1979) (used in Sun and Crook 1998; Protat and Zawadzki 1999; Collis et al. 2010), weighted averages such as Cressman (1959) or Barnes (1964) (used in North et al. 2017; Oue et al. 2019; Gebauer et al. 2022), or a combination of the two in Dahl et al. (2019). The reader is referred to Brook et al. (2022; hereafter B22) for a comprehensive discussion on the types of errors introduced by the various radar gridding techniques. The propagation of gridding errors into

3D wind retrievals has been studied by Collis et al. (2010) for linear interpolation and Majcen et al. (2008) for multipass Barnes interpolation. However, it is difficult to draw reliable conclusions regarding the relative performance of gridding methods between these studies, due to the differences in study setups (analytical OSSEs vs NWP model OSSEs), weather types (simple horizontal flow vs supercell), and varying retrieval methodologies/software implementations. To the authors' knowledge, no effort has yet been made to provide an intercomparison of multiple gridding methods and their effects on 3D wind retrievals.

The second category of spatial interpolation procedures for 3D wind retrievals is the direct radar assimilation approach. In these methods, radar data are ingested into the observational constraint in its native spherical coordinates, forgoing the pregridding of radial velocities. A comparison of the radial velocities at their measurement locations is facilitated in the observation constraint by either a trilinear (e.g., Testud and Chong 1983; Gao et al. 1999, 2004) or a Cressman–Barnes weighted average (Shapiro et al. 2009; Potvin et al. 2012b) operator. Gao et al. (1999) notes this "reverse" interpolation from the regularly spaced analysis grid to the irregularly spaced radar locations is naturally well-defined, whereas interpolation from the radar locations to the analysis grid (as in standard radar gridding) is often ill-posed due to the large data voids between tilts in operational radar data. Encouragingly, Potvin et al. (2012c) found that the analysis was largely insensitive to the Cressman radius of influence used in their forward operator, an indication of the well-posed nature of this type of assimilation procedure. The radar assimilation method also benefits from performing the interpolation simultaneously with the other variational constraints, such that the mass continuity and smoothness penalties also assist in the interpolation procedure. Additionally, the exact radial contribution of each Doppler measurement is also preserved in the radar assimilation method, avoiding the spatial distortion involved with averaging many individual measurements as in the pregridded radial velocity data. These methodological advantages lead us to hypothesize that the radar assimilation method should result in smaller spatial interpolation errors relative to the pregridded retrieval methods (Gao et al. 1999; Rihan et al. 2005). We aim to verify this hypothesis by directly comparing pregridded 3D wind retrievals to those using direct radar assimilation.

We focus our analysis on supercells in this study for two reasons: 1) while they represent a small fraction of global convection, they are disproportionately responsible for severe weather hazards (e.g., Duda and Gallus 2010; Smith et al. 2012), and 2) they present the greatest challenge for current wind retrieval methodologies due to the presence of complex, mesoscale circulations and strong vertical motions (Potvin et al. 2012b; Shapiro et al. 2009). The remainder of the manuscript is organized as follows; first, we introduce the OSSE and wind retrieval methodologies used to test our hypothesis in section 2. Second, we present qualitative and quantitative results from our ensemble of supercell OSSE experiments in sections 3a and 3b, before assessing the applicability of these findings on

real supercell data in section 3c. Finally, we summarize our findings and discuss future research directions in section 4.

2. Methodology

a. OSSE setup

In this study, we examine the effects of spatial interpolation on 3D wind retrievals using OSSEs (e.g., Potvin et al. 2012a,b; Dahl et al. 2019). Under the OSSE framework, observations are simulated using a known source, meaning the accuracy of the resulting experiments may be quantified exactly (in contrast to experiments with real data, where the "truth" is unknown). The source of our OSSE data is an Advanced Regional Prediction System (ARPS; Xue et al. 2000) simulation of the 8 May 2003 Oklahoma City tornado, first described in Xue et al. (2014). The simulation contains a mature, nearly steady-state tornadic supercell shown in Fig. 1. The reader is referred to Xue et al. (2014) and Schenkman et al. (2014) for a detailed description of the supercell. The ARPS model contains a terrain-following grid with 53 vertical levels, increasing from ~20 m spacing at ground level to ~750 m spacing at the top of the domain, and a uniform 50 m spacing in the horizontal dimensions. The simulation begins at 2200 UTC and extends 20 km × 60 km × 80 km in the z, y, and x dimensions, respectively, having been successively downscaled through a series of three nested domains with 9 km, 1 km, and 100 m horizontal grid spacings. Mesonet, rawinsonde, and Weather Surveillance Radar-1988 Doppler (WSR-88D) data are all assimilated into the analysis, and the reader is referred to Xue et al. (2001, 2003) for additional details on the model physics.

Modeled wind fields from the twentieth minute of the 50-m-resolution ARPS simulation (2220 UTC) serve as the "true" atmospheric state,[2] upon which radar data are emulated at various positions around the domain. More specifically, our OSSE experiments contain an ensemble of 15 different dual-Doppler observation locations (Fig. 1). The azimuthal spacing between radar pairs is set to 51.7°, so that the model grid lies at the center of the respective dual-Doppler lobe in each experiment (defined where the cosine of the cross-beam angle < 0.9, shown as black dotted lines in Fig. 1), eliminating errors due to poor cross-beam angles. We implement the experimental ensemble to reflect the various sampling geometries possible in operational domains, thereby ensuring the generality and repeatability of the findings. We choose to extend our analysis grid from the lowest grid point at 500 m above mean sea level (MSL; note the terrain varies gradually between ~270 and 390 m MSL across the domain), up to the tropopause height (~15 km altitude) and over the entire ARPS domain (60 km meridionally and 80 km zonally). The chosen domain is significantly larger than previous rapid-scan OSSE experiments [e.g., the 5 km × 20 km × 20 km analysis grid in Dahl et al. (2019)]. Radars are positioned at an altitude of 350 m, either 60, 70, or 80 km from the center of the domain (illustrated by the varying line styles in Fig. 1), resulting in baseline
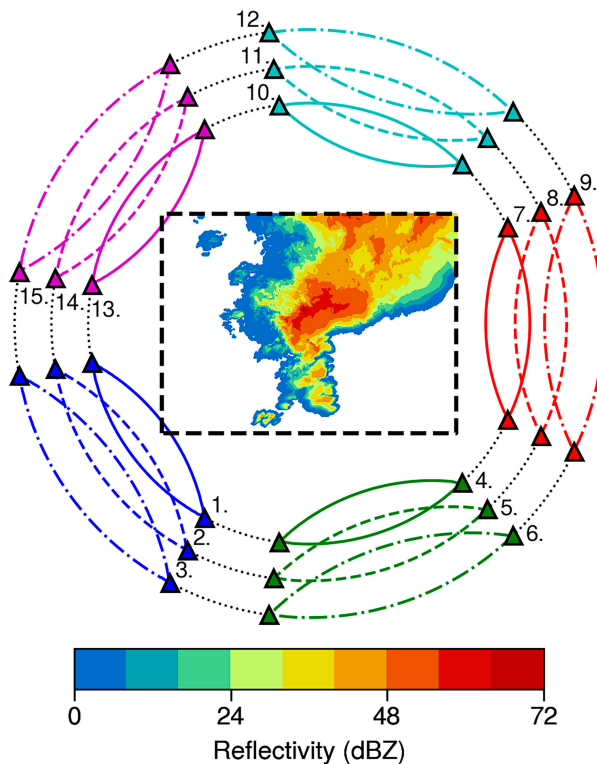

FIG. 1. The OSSE ensemble setup for this study. Dual-Doppler data are simulated for 15 pairs of locations around the model grid (pictured with reflectivity at the 20 m model level). Each simulated radar position is represented by a triangle, with colored lines indicating the collinear beam region between each radar pair (cosine of cross beam angle > 0.9). The ensemble is made up of five different viewing angles (216°, 288°, 0°, 74°, and 144°, clockwise from the positive x axis, shown by varying colors) and three different ranges (60, 70, and 80 km from the center of the model grid, shown by varying line styles).

distances for dual-Doppler pairs of ~50, ~60, and ~70 km, respectively. Furthermore, the farthest range for the most distant radars in our experiment is roughly 130 km. These settings reflect our focus on retrieval methodologies applicable to operational radar networks (as opposed to small-scale, rapid-scan analyses with mobile radars), where dual-Doppler baselines are commonly large,[3] and analysts wish to retrieve winds over large domains.

Radar data simulation is implemented identically to previous studies (e.g., Dahl et al. 2019; B22), and is described here for completeness. The radar beam propagates assuming standard atmospheric refraction, and accounts for the curvature of Earth using the 4/3 effective Earth radius model (Doviak and Zrnić 1993). Model points within each radar voxel are weighted according to their proximity to the center of the beam, effectively functioning as individual scatterers. Radar

---

[2] This analysis time is identical to that in Dahl et al. (2019), meaning retrieved winds may be qualitatively compared to their findings despite the smaller analysis domain used in their study.

---

[3] For example, the largest dual-Doppler baseline distance tested in this study corresponds to the distance between the two operational radars (70 km, Terrey Hills and Wollongong) in Sydney, Australia's most populous city.

observations are calculated as a weighted average of each scatterer as follows:

$$\varphi = \frac{\sum_{i=1}^{n} R_i B_i \varphi_i}{\sum_{i=1}^{n} R_i B_i}, \tag{1}$$

where φ is the radar field (either radial velocity or reflectivity in our case), R_i and B_i are radial and radar beam weights for the ith of n scatterers within each radar voxel. The resolution of the underlying model data ensures each radar measurement is informed by at least one "scatterer." Radial weights are implemented as

$$R_i = \begin{cases} 1, & |\Delta r_i| < 0.3d \\ \max\left(\dfrac{0.5d - |\Delta r_i|}{0.2d}, 0\right), & |\Delta r_i| \geq 0.3d \end{cases}, \tag{2}$$

where Δr is the scatterer's radial offset from the center of the radar voxel and $d = 250$ m is the radial data spacing. As in Potvin et al. (2009), the following beam weighting function simulates the angular sensitivity of radars within the U.S. operational radar network (WSR-88D):

$$B_i = \exp - 8\ln(2)\left[\left(\frac{\Delta \theta_i}{\theta_B}\right)^2 + \left(\frac{\Delta \phi_i}{\phi_B}\right)^2\right], \tag{3}$$

where $\Delta \phi_i$ and $\Delta \theta_i$ are the azimuthal and elevation offsets and $\phi_B = \theta_B = 1°$ are the half-power beamwidths (chosen to mimic standard operational S-band radars). We also implement the operational scanning strategy used in the Australian radar network, which contains 14 plan position indicator (PPI) sweeps at the following elevations: $\theta = (0.5, 0.8, 1.4, 2.4, 3.5, 4.7, 6.0, 7.8, 10.0, 13.0, 17.0, 23.0, 32.0, 45.0)°$. The realistic range, scan strategy, and dual-Doppler baselines used in these experiments ensure the applicability of these results to real, operational radar data. To this end, we follow B22 by contaminating all Doppler simulations with Gaussian observational noise (standard deviation of 1.0 m s^{-1}) and randomly masking 5% of all measurements to mimic measurement errors.

Radar reflectivity is calculated based on the modeled hydrometeor mixing ratios according to Tong and Xue (2005), and all Doppler measurements coincident with <5 dBZ reflectivity are masked to mimic data gaps in nonprecipitating regions (e.g., Potvin et al. 2012c). Finally, we assume all radar information is collected instantaneously to eliminate temporal discretization effects due to flow evolution and translation, and do not include any consideration on how hydrometeor terminal velocities affect the measured Doppler velocities (i.e., we calculate radial wind velocities directly from the modeled u, v, and w wind components assuming particle velocity is nil). While both error sources must be accounted for in retrievals from observations [e.g., using a vertical vorticity constraint and advection correction procedure (Shapiro et al. 2009; Potvin et al. 2012b; Shapiro et al. 2010), along with a reflectivity-based terminal velocity correction (Atlas et al. 1973)], we

control for these effects in our experiments to isolate the errors due to spatial interpolation alone. Wind retrieval outputs from all methods are masked using the 5 dBZ threshold from the true model reflectivity to permit the direct comparison of error statistics. These statistics include root-mean-squared errors of various quantities, including the wind components themselves, the total wind magnitude, and derived products such as vertical vorticity $\zeta \equiv \partial v/\partial x - \partial u/\partial y$ and horizontal divergence $\delta \equiv \partial u/\partial x + \partial v/\partial y$. Consistent with previous studies (e.g., Potvin et al. 2012c), centered finite derivatives with Neumann boundary conditions were used for these derived fields. Fractions skill scores (FSS) are also calculated on column-maximum, or horizontal slices of vertical velocity fields, to quantify how well each method retrieves updraft/downdraft regions (Roberts and Lean 2008). FSS scores are well-suited for assessing the ability to retrieve large vertical motions when slight spatial offsets are acceptable, owing to its original use as a neighborhood verification method for NWP rainfall accumulations.[4] Such cases may involve operational nowcasting or informing convection parameterizations, where identifying the size and strength of updrafts/downdrafts may be the priority. Finally, backward trajectory calculations for retrieved wind fields in section 3a were calculated similarly to Potvin et al. (2012c) by using a second-order Runge–Kutta scheme with an adaptive time step, along with trilinear interpolation for the surrounding velocity values.

b. Retrieval methodology

The variational 3D wind retrieval method proceeds with a cost function similar to those in previous literature, albeit with two important distinctions: an altered observational constraint (J_o) to facilitate the various interpolation methodologies, and a novel denoising constraint (J_d). The overall cost function is given below:

$$\min_{\mathbf{v}}\{J_o(\mathbf{v}) + J_m(\mathbf{v}) + J_s(\mathbf{v}) + J_d(\mathbf{v})\}, \tag{4}$$

where $\mathbf{v} = (u, v, w)$ is the 3D velocity vector, J_m is the mass continuity constraint, and J_s is the smoothing constraint, each of which are outlined below. First, in this study, the observational constraint in Eq. (4) varies depending on the structure of the input radial velocity data (V_r). We draw a distinction between methods that require pregridded radial velocity data and the proposed radar assimilation (RA) method, which ingests radial velocity data in their native, spherical coordinate system. We outline these two distinct observational constraints below in sections 2b(1) and 2b(2), before introducing the remaining three variational constraints that are common to all experiments in section 2b(3).

1) PREGRIDDED OBSERVATIONAL CONSTRAINT

When radar data are pregridded onto the analysis grid (the gridding techniques used here are outlined in section 2c), no

---

[4] We average all FSS calculations over five spatial scales in our study (namely, 1, 2, 3, 4, and 5 km), and found that our results were insensitive to shortening, lengthening, or adding additional length scales (not shown).

spatial interpolation operator is required in the optimization problem. Instead, the three Cartesian wind fields are iteratively compared to radial velocity observations on the $M = n_z \times n_y \times n_x$–sized analysis grid, through a gridded projection operator $\mathcal{P}_g : \mathbb{R}^{3M} \to \mathbb{R}^M$. The projection from Cartesian to radial velocities is achieved through the standard geometric relationship $V_r = \mathbf{v} \cdot \mathbf{p}$, where $\mathbf{p}$ is the normalized displacement vector between the radar measurement position and the location of the radar. Under these conditions, the observational constraint takes the following form:

$$J_o = \|V_{r,g} - \mathcal{P}_g\mathbf{v}\|_2^2, \tag{5}$$

where $V_{r,g}$ denotes the pregridded radial velocity measurements from both radars on a Cartesian analysis grid, and the common ℓ_2 (i.e., Euclidean) norm notation $\|x\|_2^2 = \sum_{i=1}^{N} x_i^2$ is also used throughout the text.

2) RADAR ASSIMILATION OBSERVATIONAL CONSTRAINT

The direct assimilation method ingests radial velocity observations (V_r) directly from the locations they were observed by the radar. The observational constraint in this case is as follows:

$$J_o = \|V_r - \mathcal{P}\mathcal{C}\mathbf{v}\|_2^2, \tag{6}$$

where $\mathcal{P}$ and $\mathcal{C}$ are radial velocity projection and Cressman interpolation operators, respectively. During forward passes, the Cressman interpolation operator is used to interpolate the three Cartesian velocity fields from the analysis grid to the exact radar observation locations $\mathcal{C} : \mathbb{R}^{3M} \to \mathbb{R}^{3N}$, where N is the cumulative number of observations from both radars. The reader is referred to appendix A for supplementary information on $\mathcal{C}$. Finally, the radial projection operator [as described in section 2b(1)] projects the three Cartesian velocity components at each radar data location into radial velocities $\mathcal{P} : \mathbb{R}^{3N} \to \mathbb{R}^N$. As in Gao et al. (2004), the projection and interpolation operators are combined to form a single operator in our software implementation.

3) REMAINING CONSTRAINTS

The second constraint in Eq. (4) is the anelastic form of the mass continuity constraint, commonly chosen for its suitability in deep convection (Batchelor 1953). This constraint is formalized identically to those in previous studies, using an idealized exponential density profile $\rho = \rho_0 \exp(-z/H)$, where ρ_0 is the reference density[5] and $H = 10$ km is the scale height of the atmosphere (e.g., Shapiro et al. 2009; Potvin et al. 2012b). We express it here in its simplest form:

$$J_m = \lambda_m \left\| u_x + v_y + w_z - \frac{w}{H} \right\|_2^2, \tag{7}$$

---

[5] Note there is no need to set a reference density value (ρ_0) as it is cancelled out in the derivation of Eq. (7).

where λ_m is a user-defined weighting constant used to control the relative importance of this constraint, and partial derivatives in each Cartesian dimension are denoted by subscripts (e.g., $u_x = \partial u/\partial x$).

B22 found that unmixed, second-order derivatives were a compelling regularization constraint for assimilating radar data into a Cartesian grid. The reader is referred to their appendix B for a discussion on the benefits and drawbacks of other common smoothing constraints. We follow B22, along with other recent wind retrieval studies (Potvin et al. 2012b; Dahl et al. 2019; Gebauer et al. 2022), in implementing second-order derivatives for this purpose. Theoretically, this formulation leads to a visually pleasing minimum curvature solution (Briggs 1974), while also spreading radar information into data voids due to the three-point, centered finite-difference stencil used for the numerical derivatives. The implementation used here is given below:

$$J_s = \sum_{\varphi \in \mathbf{v}} \lambda_v \|\varphi_{zz}\|_2^2 + \lambda_h(\|\varphi_{yy}\|_2^2 + \|\varphi_{xx}\|_2^2), \tag{8}$$

where λ_v and λ_h are weighting constants for the vertical and horizontal dimensions, respectively, and double subscripts denote second partial derivatives (e.g., $\varphi_{xx} = \partial^2\varphi/\partial x^2$).

Injudicious application of heavy second-order smoothing using Eq. (8) will eliminate observational radar noise at the expense of "oversmoothing" velocity fields, whereby meaningful information is filtered from the analysis (Testud and Chong 1983; Potvin et al. 2012c). B22 found that the inclusion of an anisotropic total variation denoising constraint, along with a conventional second-order smoothing constraint, is able to efficiently spread information in to data voids, eliminate observational noise, and preserve the sharp "edge" features that are common in radar observations of deep convection. In this study, we implement this constraint in an effort to better resolve the highly nonuniform, turbulent velocity information that is commonly underestimated in 3D wind retrieval analyses of strong thunderstorms (e.g., Potvin et al. 2012c; Oue et al. 2019; Evaristo et al. 2021). The denoising constraint is implemented as follows:

$$J_d = \lambda_d \sum_{\varphi \in \mathbf{v}} \|\varphi_z\|_1 + \|\varphi_y\|_1 + \|\varphi_x\|_1, \tag{9}$$

where λ_d is a tunable weighting parameter and the ℓ_1 norm notation used here is equivalent to $\|x\|_1 = \sum_{i=1}^{N} |x_i|$. Optimization in experiments containing the ℓ_1 denoising constraint is complicated due to the nonsmooth, nondifferentiable term in the cost function [arising from the first-order discontinuous absolute value operation in Eq. (9)]. As in B22, the split Bregman optimization method for mixed ℓ_1–ℓ_2 norms is used for optimizing the cost function in our experiments (Goldstein and Osher 2009; Ravasi and Vasconcelos 2020).

The provision of suitable tuning parameters (λ_m, λ_v, λ_h, and λ_d) is a challenging task in variational wind retrievals (Shapiro et al. 2009). Theoretically, parameters should reflect the intended strength of each constraint; however, such judgements are difficult to make a priori due to factors such as observational errors and data acquisition gaps. Recent OSSE studies

acquire suitable tuning parameters through experimentation, aided by the introduction of nondimensional tuning parameters that narrow the tuning parameter space (Shapiro et al. 2009; Potvin et al. 2012b; Dahl et al. 2019). We extend this experimental approach by programmatically retrieving optimal tuning parameters for each experiment using Bayesian optimization (e.g., Barth et al. 2022). Optimal tuning parameters are obtained by minimizing the RMS error for each OSSE retrieval, ensuring a fair comparison between each experiment. As in the aforementioned OSSE studies, these optimized parameters may then be employed in retrievals using observations. This approach is important when judging the potential performance for different variational formalisms, as optimal[6] parameters likely vary significantly for each experimental setup (e.g., pregridded data may benefit from a reduced smoothing constraint as some smoothing has already taken place during the gridding process). We use the Bayesian optimization routines within *scikit-optimize* for these purposes (Head et al. 2021).

c. Radar gridding

Most recent 3D wind retrieval studies (e.g., North et al. 2017; Dahl et al. 2019; Gebauer et al. 2022), along with open-source wind retrieval software (*PyDDA*; Jackson et al. 2019), require that dual-Doppler wind data must be pregridded prior to analysis. In this study, we aim to investigate how this preprocessing step may degrade the resulting wind retrievals relative to the direct assimilation approach described above, and provide guidance on how common gridding techniques perform relative to each other. We also include "control" dual-Doppler observations, which are upscaled directly from the original, high-resolution (50 m horizontal) model grid using Cressman (1959) weighted average interpolation with a radius of influence equal to the isotropic, analysis grid spacing (500 m). These control experiments will illustrate the baseline performance of our wind retrieval code for "perfect" observations, before they are degraded by the spatial interpolation errors in the "radar sampled" experiments.[7] Furthermore, we test two common radar gridding techniques involving Cressman weighted averages in two and three dimensions, both of which are detailed below.

1) 3D CRESSMAN

Perhaps the simplest and most commonly used radar gridding methods are the weighted average methods first proposed by Cressman (1959) and Barnes (1964). These methods are commonly used prior to variational wind retrieval algorithms

---

[6] We attempt to find the optimal parameters for each experiment (by experimentally limiting the parameter space in subsequent runs and running each optimization for 100 iterations); however, the reported parameters may not be truly "optimal" due to local, rather than global, convergence.

[7] Note that retrievals with "perfect" control observations do not result in wind fields with zero errors due to the effects of discretization errors, ill-posed boundary conditions and model assumptions (such as the anelastic mass continuity approximation) in the retrieval methodology.

(e.g., North et al. 2017; Oue et al. 2019; Gebauer et al. 2022), and operate by weighting all radar observations within a radius of influence (R) from each grid point in three dimensions. We test Cressman, rather than Barnes, weightings in this study due to their more prominent use in radar literature (Trapp and Doswell 2000).[8] The Cressman weighting function (W) is outlined below:

$$W(r) = \exp\left(\frac{R^2 - r^2}{R^2 + r^2}\right), \tag{10}$$

where r is the distance between an observation and the corresponding grid point. Despite being initially proposed as an iterative technique, radar analysts commonly implement Eq. (10) in a single pass [with some exceptions, e.g., Majcen et al. (2008)]. We follow this literary precedent by applying the 3D Cressman method in a single pass, using an open-source implementation (PyART; Helmus and Collis 2016). As in B22, we set R equal to the largest data spacing in the analysis domain ($R = d_{max} \approx 3$ km) and refer to the discussion in their section 2b regarding the provision of this parameter.

2) 2D CRESSMAN

The second gridding technique implemented here was first described by Dahl et al. (2019) for the purposes of dual-Doppler wind retrievals, and is subsequently referred to as the 2D Cressman method. Here, a single-pass weighted average with Cressman weights [Eq. (10)] is used to map radar data onto a horizontally regular, Cartesian grid along each conical, two-dimensional, PPI surface. These surfaces are then merged onto a 3D Cartesian grid by linearly interpolating along each vertical column according to

$$V_r = V_{r,1} + \frac{z - z_1}{z_2 - z_1}(V_{r,2} - V_{r,1}), \tag{11}$$

where subscripts 1 and 2 refer to data immediately below and above (respectively) the grid point at height z in each column.

Using Cressman interpolation solely along PPI surfaces permits the use of a considerably smaller R value (~1.7 km), set by the maximum azimuthal data spacing instead of the larger maximum elevation spacing (elevation gaps between PPI's commonly exceed 5° in operational scanning strategies). A decreased R value leads to the retention of small-scale details (on the order of d_{max}), which are commonly filtered or "oversmoothed" in the standard 3D Cressman gridding approach (Trapp and Doswell 2000; Zhang et al. 2005; B22). Resolution improvements gained through smaller R values are naturally offset by inaccuracies introduced by the vertical linear interpolation, such as first-order discontinuities in sparsely sampled regions and spectral distortion at poorly resolved wavelengths (Askelson et al. 2000; Trapp and Doswell 2000; B22). Three-dimensional wind retrievals are particularly sensitive to this type of spectral broadening due to their reliance on finite

---

[8] B22 found that this distinction is largely academic, as the Barnes and Cressman methods produce practically indistinguishable results for the purposes of domainwide radar gridding.

differences for numerical derivatives (Testud and Chong 1983), meaning it is not immediately clear whether the added resolution achieved by the 2D Cressman method is advantageous for our purposes. In this study, we aim to provide some experimental guidance regarding this question.

3. Results

a. Ensemble member case study

In this section, we analyze the performance of the first of 15 ensemble members pictured in Fig. 1 (i.e., experiment 1 of 15). We aim to derive a qualitative understanding of the performance of each retrieval method through the analysis of this single experiment, before providing more quantitative ensemble statistics in section 3b. First, Fig. 2 illustrates the various radial velocity inputs for the northwesternmost radar in experiment 1. The control field in Figs. 2a and 2c has not undergone any radar sampling; rather, it is a direct projection of the true Cartesian wind components in the radial direction. This setup renders it a ground truth description of the underlying radial velocities, limited spatially to regions with radar reflectivity ≥ 5 dBZ.

Perhaps the most noteworthy signature in the control field is the pronounced velocity couplet in region 1 of Fig. 2a, indicating the presence of a strong low-level mesocyclone. The shape and azimuthal shear values across the couplet are largely smoothed in the pregridded radial velocities in Figs. 2b and 2e. These effects are due to spatial aliasing in the radar gridding methods, which reliably filters observational noise, at the expense of high-frequency information required to accurately resolve such small-scale features (B22). We hypothesize that this type of smoothing will hamper the ability of pregridded methods to retrieve important dynamical features such as mesocyclone rotation in the resulting 3D wind retrievals. In contrast, the simulated radar data used in the RA method (Fig. 2f) accurately resolve the velocity couplet at ~60 km range in experiment one, albeit in the presence of notable speckle noise added during the radar simulation process. This level of observational noise is common throughout the simulated radar data (e.g., in region 2), and it is not clear a priori whether the benefits of finer data resolution will outweigh the added observational noise in the RA method.

Another important consideration for each wind retrieval method is the spatial continuity, or equivalently, the size and shape of data voids, in the radial velocity data. The "true" extent of the simulated radar echoes is shown in the control experiment (with the minimum detectable signal defined at 5 dBZ). The 3D Cressman method artificially extends valid data past their true extent (e.g., between regions 3 and 4 in Fig. 2), effectively "filling in" many true data voids. It is not obvious whether this data extension is advantageous in 3D wind retrievals, as the added observations may actually assist in retrieving valid winds in data voids, contingent on the quality of the extrapolated radial wind values. Velocity observations in the RA method contain large gaps aloft due to the spacing between constant elevation radar sweeps (e.g., region 3 in Fig. 2h), and this may have serious implications for the accuracy of retrieved winds if

the interpolation operator ($\mathcal{C}$) and the smoothing constraints (J_s) cannot effectively propagate radar information into these data voids. Similarly, the radial velocities in the 2D Cressman method are severely limited above the highest tilt, and below the lowest tilt (e.g., region 2 in Fig. 2e and region 4 in Fig. 2g). A notable consequence of these spatial continuity considerations may be observed in region 4, where a signature indicating easterly inflow in the low levels is evident in the control observations (Fig. 2c). This signature is largely absent in the pregridded methods, due to oversmoothing in the 3D Cressman observations, and data gaps below the lowest tilt in the 2D Cressman observations. Once again, we hypothesize that the pregridded retrievals will be adversely affected by the omission of this important dynamical information in the underlying radial velocity data.

Horizontal cross sections of retrieved vertical velocity for each method are given in Fig. 3. Visual comparisons of both shading colors and streamlines between Figs. 3a and 3c indicate that the control experiment is able to reproduce the true 3D wind field very accurately. One notable exception in the accuracy of the control experiment are small patches of spurious vertical velocity retrievals, such as those in region 2 of Fig. 3a. These are due to boundary effects in data-sparse regions, and are an example of how errors are still present in 3D wind retrievals with "perfect" radial velocities. For these reasons, the accuracy of any 3D wind retrieval method simulated using practical radar scanning geometries should also be judged against such a control experiment, and not merely the "true" model winds. The reader is referred to appendix B for further discussion on boundary effects in 3D wind retrievals.

All three retrievals that ingest radar simulated radial velocities (either pregridded in 3D/2D Cressman methods, or directly in the RA method) are able to reproduce the broadscale southwesterly horizontal flow, along with the approximate size and shape of the updraft within the inset region in Fig. 3 (hereafter referred to as the analysis region). The largest errors arise in the vertical velocity retrievals for the pregridded methods, which largely underestimate the true magnitude of the updraft velocities (by more than a factor of 2 in the 3D Cressman method), and are unable to retrieve the small downdraft features around the edges of the main updraft (blue shading, northwest and southeast of region 1 in Fig. 3c). We attribute these poor vertical velocity retrievals to the errors introduced by pregridding the underlying radial velocity fields. More specifically, filtering and oversmoothing of the highest radial velocity magnitudes, along with the poorly resolved low-level convergence discussed in Fig. 2, both contribute to the large vertical velocity errors noted in Fig. 3.

Unlike pregridded methods, the RA method does accurately resolve the amplitude and extent of the main updraft region (with maximum updraft velocities in Fig. 3e approaching the true values evident in Fig. 3c). Furthermore, the position of the aforementioned downdraft features are retrieved well in this method, while slightly underestimating maximum downdraft speeds. We attribute this success to the higher resolution of the input radial velocity data, which allows the technique to resolve the small-scale and highly dynamic circulation within the analysis region. Another example of the
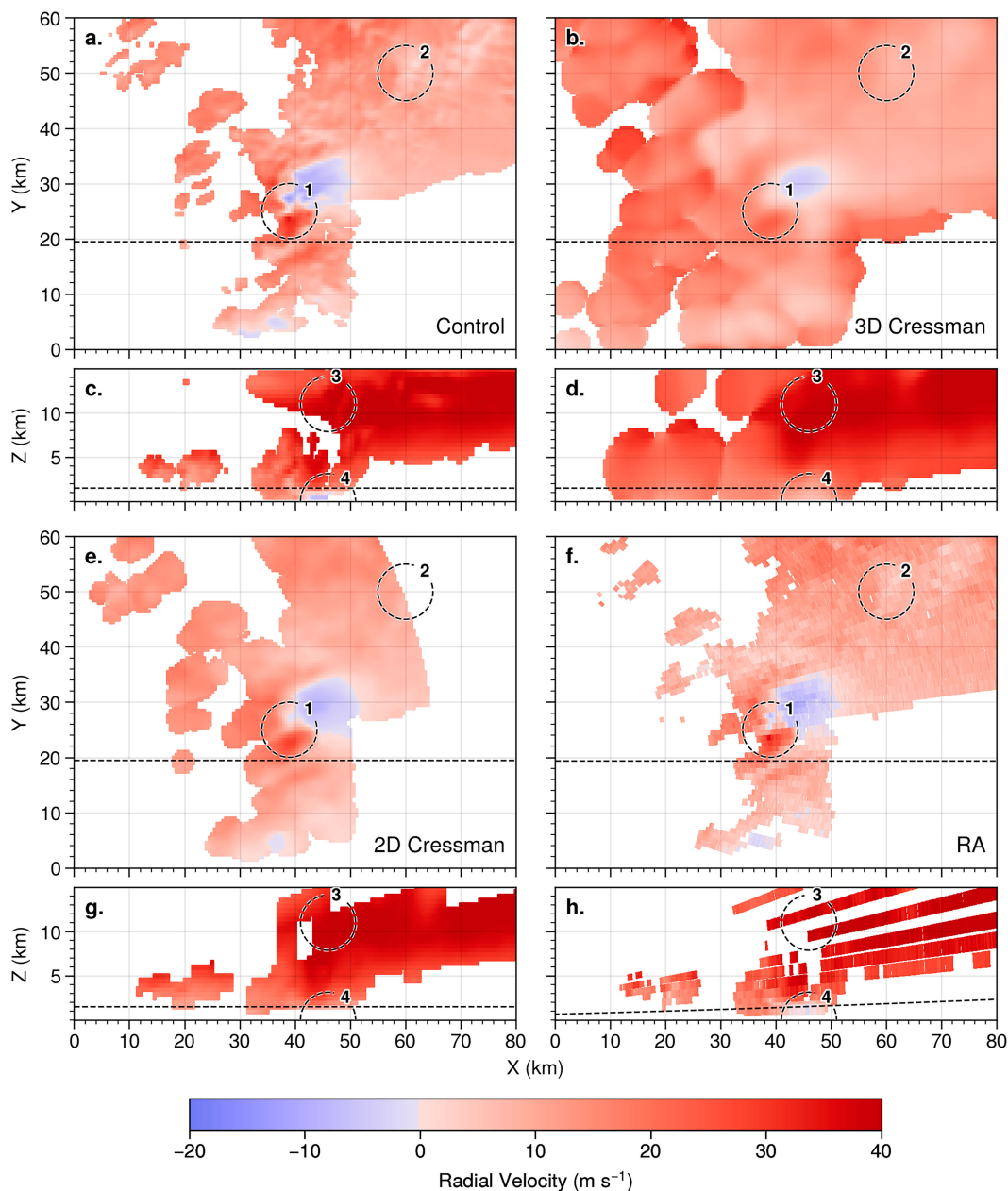
FIG. 2. Simulated radial velocities for radar position 1 for each of the 3D wind retrieval methods in this study. (a),(b),(e) Horizontal cross sections at $Z = 1.5$ km. (f) PPI from the second sweep ($\theta = 0.8°$). (c),(d),(g) Vertical cross sections at $Y = 19.5$ km. (h) "Pseudo" RHI from the 90° azimuth, where radar data are shaded inside the bounds of the 1° angular beamwidth of each ray. Straight dotted lines indicate the positions of the corresponding cross sections, and circular regions are highlighted for discussion in the text.

retrieval of important small-scale dynamic features is the cyclonic rotation within region 1 in Fig. 3c. Streamlines in Fig. 3e indicate this feature is retrieved well in the RA method, where the easterly component of the mesocyclone rotation is clearly visible against the prevailing westerly background flow. Both pregridded methods were unable to retrieve any easterly wind component within the mesocyclone region, which has important implications for identifying supercell thunderstorms,

which cause a disproportionate amount of large hail and significant tornadoes (e.g., Duda and Gallus 2010; Smith et al. 2012).

Vertical cross sections of vertical velocities from experiment 1 in Fig. 4 further illustrate the deficiencies of the pregridded retrieval techniques. We will highlight these deficiencies in the context of two important dynamic features present in the analysis inset in Fig. 4c: a rear flank downdraft (RFD) signature
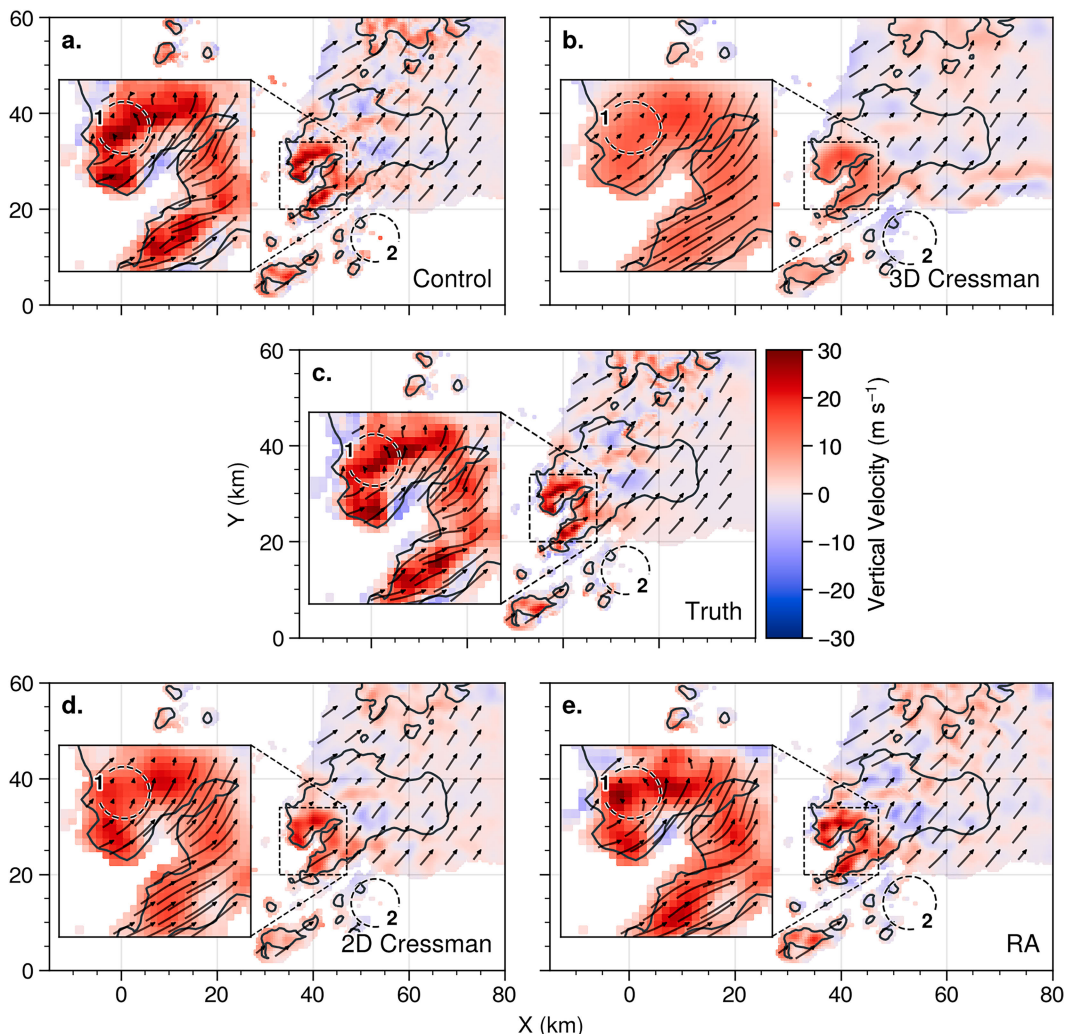
FIG. 3. Horizontal cross sections at $Z = 5$ km with a 50 dBZ reflectivity contour in gray, streamlines illustrating horizontal velocities (length proportional to horizontal wind magnitude), and shading indicating vertical velocities: (c) model "truth" winds, along with retrieved winds from the (a) control, (b) 3D Cressman, (d) 2D Cressman, and (e) RA methods. Insets in each panel show a zoomed view of the analysis region discussed in the text (spanning $X = 33$–47 km and $Y = 20$–34 km).

close to the surface within region 1, and an overturning circulation resulting from an updraft/downdraft couplet within region 2. Both pregridded methods were unable to retrieve any downdraft areas within the inset region, meaning both the RFD and overturning circulation are absent from the resulting wind fields. However, the RA was again able to retrieve these features accurately in terms of their shape and position, while slightly underestimating the true vertical velocity maximum values. The retrieval of these features has important implications for nowcasting meteorological hazards, such as identifying potentially strong straight line winds at the surface due to the RFD, and the existence of favorable hail growth trajectories due to the overturning circulation (e.g., Kumjian et al. 2021).

While the wind retrieval clearly benefits from the added resolution of radial velocity observations in the RA method,

this method still displays considerable limitations when compared to the control retrieval. First, consider the third highlighted region in Fig. 4c, which contains two spatially proximal, but distinct regions of very high vertical velocities (>30 m s^{-1}). The control method in Fig. 4c accurately retrieves these separated updraft regions, meaning this scale of motion (these features, and most convective towers, are approximately 1–2 km in diameter) is practically resolvable on the 500 m isotropic analysis grid used in this study. The RA method is once again able to retrieve the magnitude of this updraft feature in region 3 much more accurately than the pregridded methods; however, the two updraft features are aliased into one large updraft region in Fig. 4e. This spatial aliasing is present throughout the upper levels (>10 km), and is a result of the sparsity of observed radar information at these altitudes (refer to Fig. 2h). Clearly, the accuracy of the
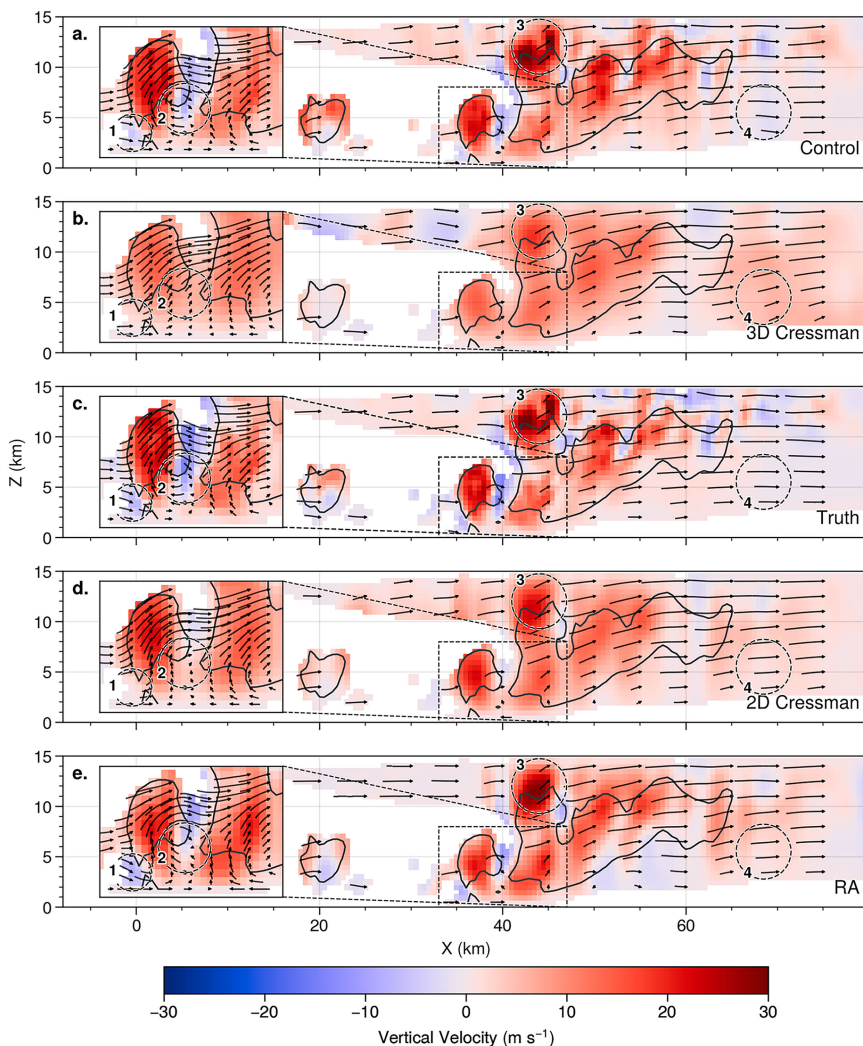
FIG. 4. As in Fig. 3, but for a vertical cross section at $Y = 30$ km. The inset panel shows a zoomed view of a region spanning $X = 33$–47 km and $Z = 0.5$–8 km.

RA method is still fundamentally limited by the scanning geometry of the underlying radar observations, albeit to a lesser extent than the pregridded methods. Finally, we also note the poor performance of all three radar-derived retrieval methods in the forward flank of the storm (e.g., region 4 in Fig. 4). Here, spurious vertical velocity signatures are present throughout the depth of the storm, and are coincident with significant data gaps due to clear air in the low levels. We posit these errors are caused by a well-known problem in the wind retrieval literature, whereby the impermeability boundary condition ($w = 0$ at $Z = 0$) in the mass continuity constraint is poorly imposed in regions with data voids near the surface. Studies have shown these issues may be alleviated through the inclusion of a vertical vorticity constraint (Shapiro et al. 2009; Potvin et al. 2012b), which is not implemented here due to our sole focus on the effects of spatial interpolation.

The final qualitative assessment of the wind fields from experiment 1 is a comparison of simulated air parcel trajectories in Fig. 5 (the control experiment is omitted due to its similarity with the model truth). Trajectories are initiated within the inflow of the storm at 600, 700, and 800 m altitudes, at 1 km spacing within the horizontal bounds of the analysis box shown in the inset in Fig. 3. Broadly speaking, trajectories calculated using the true wind fields in Fig. 5a follow two main paths: strong vertical ascent to the top of the domain (>8 km) within the updraft [mostly those initiated in the western (blue) and northern (orange) quadrants], or initial ascent followed by ejection into the forward flank of the storm in the midlevels [4–8 km, mostly the eastern (green) and southern (red) quadrants]. The RA winds in Fig. 5d produce trajectories that are qualitatively similar to the true trajectories for all four quadrants. Most notably, the RA method reproduces the strong split between trajectories from the southern (blue) and western (red) trajectories, which is also observed in Fig. 5a. Many trajectories in the pregridded fields also exhibit very similar trajectories to those in Fig. 5a; however, there are also
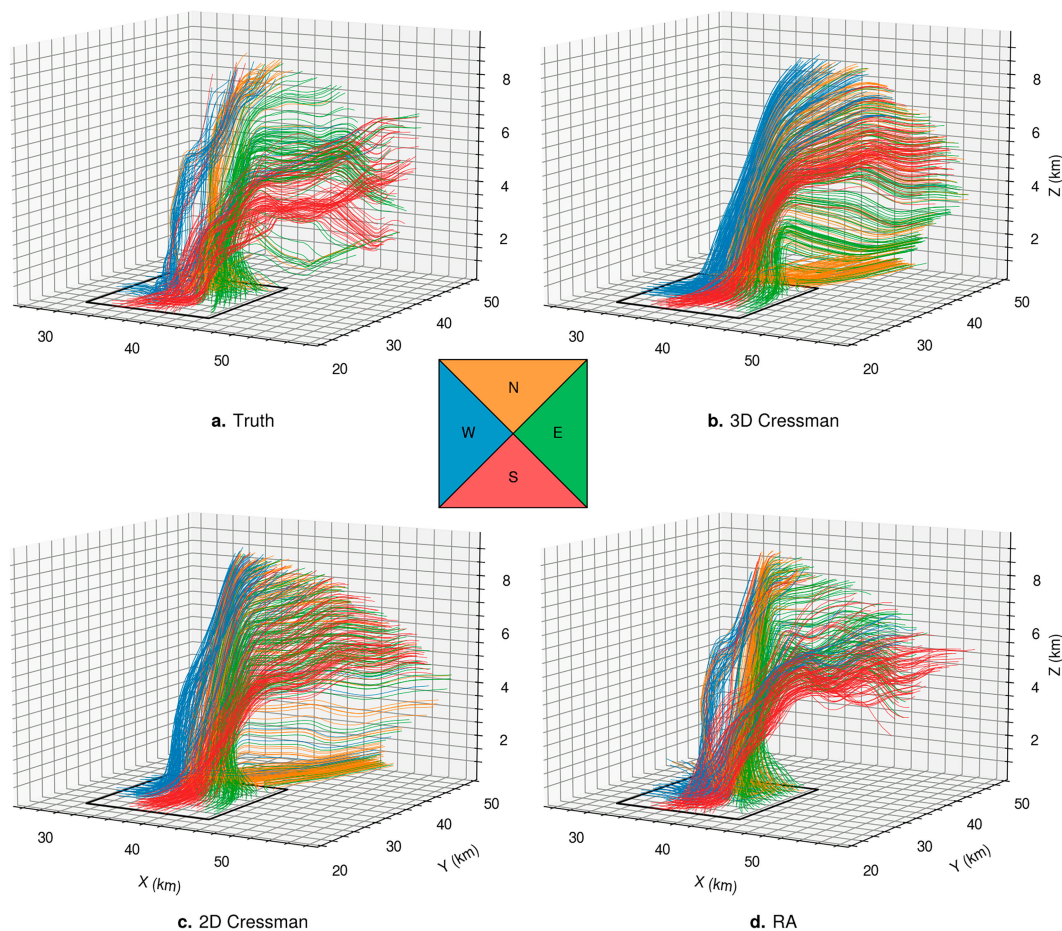
FIG. 5. Trajectory calculations for air parcels initiated once every 1 km in the horizontal dimensions within the analysis region (pictured in black), and 600, 700, and 800 m in altitude. Trajectories are modeled for (a) true model wind fields, and (b)–(d) retrieved wind fields, and colored according to their initial quadrant position (refer to the inset for details on quadrant coloring).

some notable differences. For example, the aforementioned split between the western and southern inflow regions is not observed in Figs. 5b and 5c. Rather, there exists a smooth transition between these pathways, resulting in a roughly continuous spatial distribution between the trajectories at the top of the storm and those ejected into the midlevels.

Another notable difference in the pregridded retrievals (vs the radar assimilation retrievals) is the comparative smoothness, or uniformity, of the trajectories, especially in the forward flank. This is once again a result of filtering of important information in the underlying gridded radial velocities, resulting in oversmoothed retrieved wind fields. Last, perhaps the most notable difference in the pregridded fields is the spurious collection of low-level trajectories which initiate from the northern quadrant of the initiation region. The presence of these trajectories indicates a lack of convergence within the main updraft, which ultimately results in the aforementioned underestimation of updraft strength for these methods. The lack of convergence in the low levels was hypothesized through the analysis of Fig. 2, where the inflow signature was

poorly resolved in the pregridded radial velocities. Figure 5 qualitatively confirms this assumption, and emphasizes the importance of ingesting accurate radial velocity information, especially in the low levels.

b. Ensemble analysis

Table 1 presents a summary of the experimental results for each retrieval method, including ensemble mean and standard deviation values. The control experiment achieves excellent results across all radar geometries tested in this study. Considering the average wind magnitude for the experiment is ~51 m s^{-1}, the ensemble average RMSE of 2.00 m s^{-1} results in an average percentage error of just ~4% in wind magnitudes across the domain. A standard deviation of 0.05 m s^{-1} for the 15 experiments also indicates the technique is not particularly sensitive to different radar positions or cross beam angles. As expected, the vast majority of error comes from the poorly observed vertical wind component, which has an average RMSE of 1.96 m s^{-1}, compared to the horizontal

TABLE 1. Summary error statistics for the 15-member ensemble for each retrieval method. The ensemble mean and standard deviation (in parentheses) are shown for the total RMSE, RMSE within the analysis region (an.), and RMSE for individual wind components (u, v, w). Fractions skill scores (FSS) are also listed for updrafts (up, 95th percentile of the column maximum vertical winds) and downdrafts (down, 5th percentile of the column minimum vertical winds).

	RMSE	RMSE an.	RMSE u	RMSE v	RMSE w	FFS up	FSS down
Control	2.00 (0.05)	2.68 (0.04)	0.23 (0.13)	0.22 (0.11)	1.96 (0.03)	0.95 (0.00)	0.80 (0.02)
3D Cressman	7.17 (0.16)	11.61 (0.22)	4.42 (0.28)	4.25 (0.36)	3.68 (0.16)	0.70 (0.04)	0.26 (0.04)
2D Cressman	7.43 (0.37)	8.90 (0.37)	4.96 (0.45)	4.28 (0.29)	3.49 (0.14)	0.76 (0.04)	0.43 (0.05)
RA	6.22 (0.27)	8.03 (0.61)	3.77 (0.57)	3.61 (0.52)	3.30 (0.13)	0.88 (0.01)	0.63 (0.06)

component RMSE's of 0.23 and 0.22 for u and v, respectively. We also note the control experiment is able to retrieve winds with an average RMSE of 2.68 m s^{-1} in the analysis region, which indicates that the "perfect" radial velocities on a 500 m analysis grid are capable of resolving the scales of motion very well in this highly dynamic region of the storm. Finally, we observe that FSSs for updrafts outperform those for downdrafts in the control experiment (0.95 vs 0.80, respectively), and note this finding is consistent across the three radar observed methods. This is likely a result of the proximity of downdraft regions to the top of the analysis domain ($z = 15$ km) at this time step, where boundary errors are prevalent (cf. downdrafts in Figs. 4a,c).

As expected from the qualitative analysis in section 3a, the quality of retrievals is considerably degraded when winds are retrieved with simulated radar radial velocities, resulting in a 300%–350% increase in error magnitudes. Interestingly, we did not observe reliable accuracy reductions for ensemble members with more distant radars (not shown), indicating the storm was adequately sampled by the volume coverage pattern at each of the three ranges used in this study (60, 70, and 80 km). First, both pregridded methods score similarly in terms of overall RMSE scores (7.17 m s^{-1} for 3D Cressman and 7.43 m s^{-1} for 2D Cressman); however, ensemble standard deviations indicate that error scores are much more consistent in the former. This is easily explained by the superior error filtering qualities in the 3D Cressman analysis, which limits the amount of observational error propagating into the retrieved wind fields. While the smoothing properties of the 3D Cressman analysis are beneficial in the consistency of error scores, they also hinder the retrieval of strong winds in the analysis region, where the RMSE score is considerably higher at 11.61 m s^{-1}. The inability of the 3D Cressman method to retrieve winds in highly dynamic regions is further underscored by the poor updraft and downdraft FSS values (defined by the 5th and 95th percentiles of column-maximum vertical velocities, $w \approx 8.5$ and -3.4, respectively) relative to the control experiment (0.70 vs 0.95 for updrafts, and 0.26 vs 0.80 for downdrafts). These results ultimately illustrate the trade-off between the 3D Cressman and 2D Cressman gridding methods: the former results in more consistent retrievals with slightly lower total error measures, but degrades the retrieval in the most dynamic regions of the storm, which are perhaps the most important in terms of understanding storm dynamics and nowcasting potential hazards.

The average RMSE score of 6.22 m s^{-1} in the RA method supports the aforementioned qualitative improvements over the pregridded methods. The use of this technique results in an average reduction in errors of 18% and 22% compared with the 3D Cressman and 2D Cressman methods, respectively, relative to the control experiment. These domainwide improvements relative to the pregridded methods are also reflected in RMSEs in the analysis region (8.01 m s^{-1}), and for individual wind components over the entire domain ($\sim$3.5 m s^{-1} errors for u, v, and w). Interestingly, while the average vertical velocity error is slightly lower in the RA method, the majority of RMSE improvements come from reducing errors in the horizontal components, which are much lower than in the pregridded retrievals. This belies the qualitative observations made in Figs. 3 and 4, which showed considerable improvements in RA vertical winds, especially within the analysis region. An explanation for this seeming discrepancy is that high-amplitude fields are penalized more heavily by pixel-to-pixel accuracy measures such as RMSE, as slight spatial offsets in wind fields result in higher error scores compared to smoother fields such as the 3D Cressman method. This effect also accounts for the lower RMSE values observed in the 3D Cressman method compared to the 2D Cressman method, despite the considerable oversmoothing noted in the former in section 3a. We further this hypothesis by noting that the FSSs for updraft and downdraft regions are considerably higher for the RA method. The updraft and downdraft FSSs in the RA method are only 7% and 21% lower, respectively, than the control experiment, while the 3D Cressman and 2D Cressman methods show larger a reduction of 20% and 46%, and 26% and 68%, respectively. These statistics indicate that while the RA method may not achieve a significantly lower RMSE for vertical velocities, it does retrieve regions of high vertical velocity much more reliably than the pregridded methods.

After presenting the mean RMSE and FSS values for the entire domain in Table 1, we proceed to examine the variation of these statistics with the magnitude of the underlying variables in Fig. 6. First, Fig. 6a confirms that horizontal velocities are retrieved very well with the "perfect" observations in the control experiment (<0.5 m s^{-1} errors across the entire range of magnitudes). For the other retrieval methods, the sampling effects of the simulated radar observations considerably degrade the accuracy of horizontal winds, with comparatively large RMSEs (>3 m s^{-1}). The lowest horizontal velocity errors actually arise in regions with large velocity magnitudes ($\sim$45 m s^{-1}), which correspond to the mean environmental flow in the upper levels of the storm (8–12 km, refer to Fig. 4). Horizontal velocity errors are lower here due to the uniformity
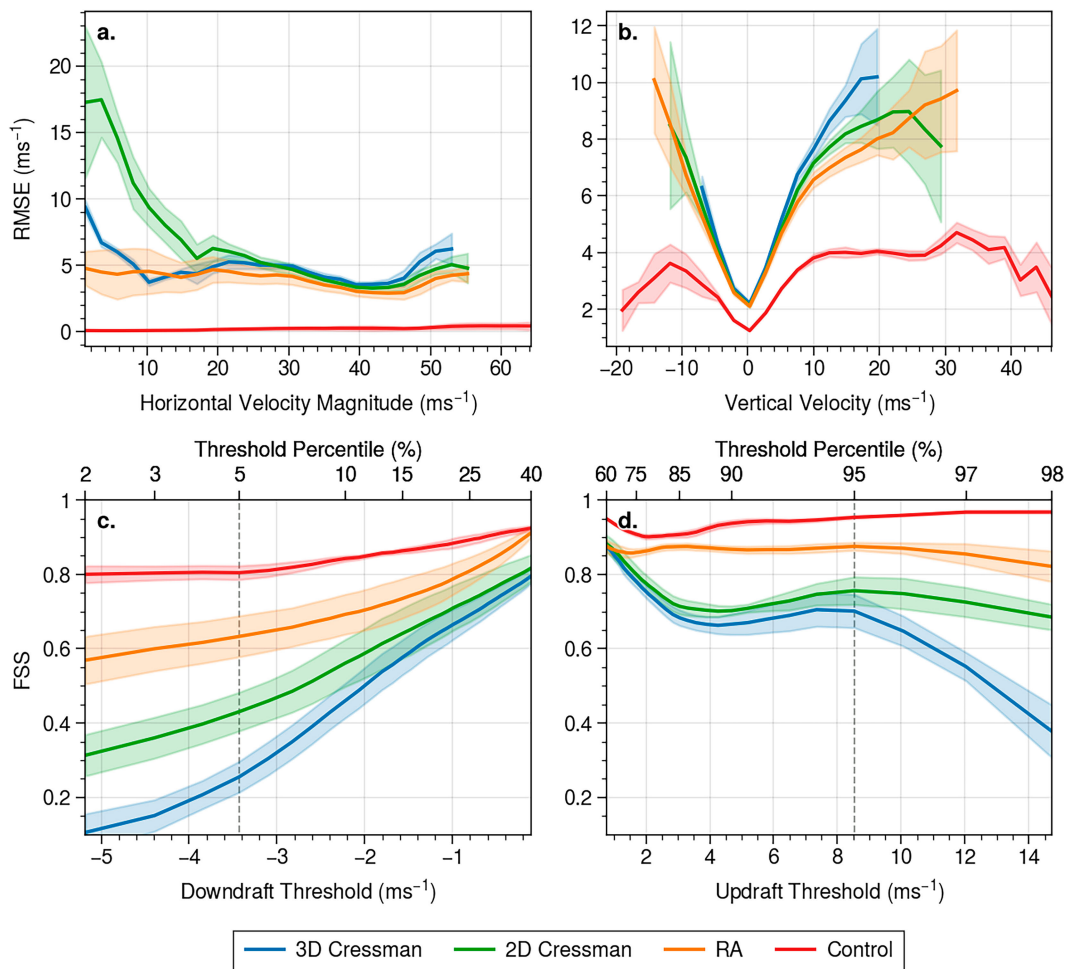
FIG. 6. Root-mean-square errors (RMSE) across a range of magnitudes for various quantities and retrieval methods. All values are classified into 30 bins according to their magnitude between the maximum and minimum of the true values for each field. The RMSE is calculated within each bin, and the ensemble mean error is plotted with a solid line. One standard deviation of the ensemble values is shown above and below the mean to indicate the ensemble spread. A line is plotted at each magnitude only if all 15 ensemble members have at least one valid value in the bin, and (c),(d) thresholds used to compute FSS values are indicated by gray dotted lines.

of the flow, with limited storm-induced perturbations and smaller vertical velocities. In contrast, regions with low horizontal velocities ($<$10 m s^{-1}) occur only very close to the surface, or in strongly dynamic regions of the storm, which oppose the broadscale environmental flow (such as within the mesocyclone, refer to Fig. 3). Both of these regions are poorly resolved in the pregridded retrievals, leading to large errors at low velocity magnitudes (especially for the 2D Cressman method, see Fig. 6a). The 3D Cressman method also predictably shows larger errors for the strongest horizontal velocities, as the oversmoothed radial velocities cannot reproduce the highest magnitude winds. Generally speaking, the RA method produces the lowest horizontal velocity errors and has the most consistent accuracy across the range of horizontal velocity magnitudes.

The difference in error magnitudes between the control experiment and the other retrieval methods is less pronounced for vertical velocities (cf. the red lines in Figs. 6a,b). This

reflects the comparative difficulty of retrieving accurate vertical velocities with low-elevation radial velocities, even with "perfect" radar observations. Errors are lowest for all methods in areas with small vertical velocities ($\sim$0 m s^{-1}), and generally increase at higher velocity magnitudes. As in the horizontal velocities, 3D Cressman retrievals most poorly resolve the highest wind magnitudes due to oversmoothing of important high-frequency information, and the RA method produces the lowest errors across most vertical velocity magnitudes. The RA method is also able to retrieve considerably larger vertical velocity magnitudes (both updrafts and downdrafts) relative to the pregridded techniques; however, these retrieved maximum values are still considerably underestimated. We also examine how FSS values vary with respect to the thresholds used to delineate the updraft and downdraft regions. We test thresholds between the 60th and 98th percentiles of the column-maximum vertical velocity field to compute
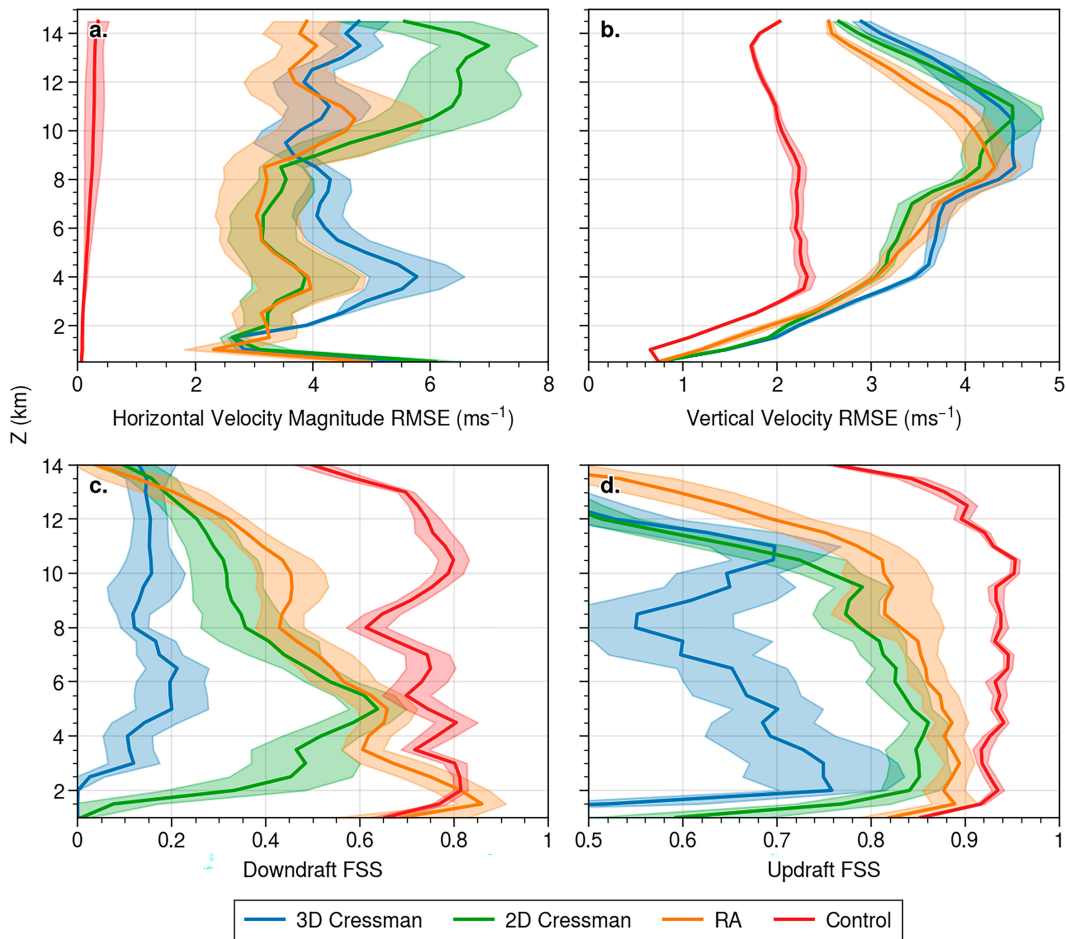
FIG. 7. Vertical profiles of various error scores for each retrieval method. The quantities are grouped and averaged at each altitude level, and the ensemble mean and standard deviation are then calculated and plotted, as described in Fig. 6.

updraft FSSs, and between the 2nd and 40th percentiles of the equivalent column-minimum field for downdrafts. Figures 6c and 6d show the qualitative FSS results presented in Table 1 (which used the 5th- and 95th-percentile values) are not sensitive to this threshold choice, as the RA method consistently outperforms the pregridded methods over the range of thresholds. This is particularly evident for the updraft and downdraft thresholds at the 98th and 2nd percentiles, confirming the RA method is better suited to retrieving regions with the most intense vertical velocities. These ensemble findings concur with the qualitative observations of vertical velocities made based on Figs. 3 and 4 in section 3a.

We also analyze how each of the diagnostics presented in Table 1 vary with altitude in Fig. 7. Consistent with Fig. 6, the control experiment exhibits very small horizontal velocity errors throughout the depth of the domain, but most notably at the lowest grid point (500 m altitude). This indicates that the impermeability condition imposed at the lowest grid point, as opposed to the ground surface, does not adversely affect the retrieval in this case (a more sophisticated implementation considering the surface elevation should be used in areas with

complex terrain; Chong and Cosma 2000; Liou et al. 2012). All three radar sampled methods exhibit large horizontal velocity errors (>5 m s^{-1}) at the surface due to the missing data below the lowest radar tilt elevation (0.5°), especially the 2D Cressman method (refer to Fig. 2g). Recall that Table 1 showed the majority of RMSE reductions in the RA method were attained through improvements in the horizontal velocity components. Figure 7a illustrates the origin of these improvements. The RA method produces roughly consistent RMSE scores throughout the depth of the storm (~ 4 m s^{-1}); however, both the pregridded methods show considerable deficiencies at different height levels. First, the 3D Cressman method produces large horizontal velocity errors in the midlevels of the storm (~ 4 km) due to a combination of not resolving the small-scale dynamics such as the mesocyclone within the core of the storm (e.g., region 1 in Fig. 3b), and boundary effects due to missing low-level data within the forward flank of the storm (e.g., region 4 in Fig. 4b, also refer to appendix B). Second, the 2D Cressman method exhibits large errors in the upper levels (>10 km), where the linear interpolation scheme used to fill large data gaps between PPI scans propagates

observational noise and creates first-order discontinuities in the underlying radial velocity data (Trapp and Doswell 2000; Askelson et al. 2000; B22). The ability of the RA method to mitigate these error sources results in overall lower RMSE scores.

Errors in vertical velocities for all four methods (including the control) increase steadily to a height of roughly 4 km, coinciding with the lower portion of the strong (>30 m s^{-1}) updraft in the midlevels (refer to Fig. 4c). Above 4 km, the errors in the control experiment stay roughly constant at around 2 m s^{-1}, whereas the radar sampled methods are significantly degraded above 8 km. We attribute this poor performance to the operational radar scanning pattern used in our experiments, which results in large elevation gaps that fundamentally limit the accuracy achievable in the upper levels. As noted in Table 1, the average improvement in vertical velocities gained by the RA method is modest throughout the depth of the storm, with 2D Cressman retrievals actually producing the lowest mean RMSE between 5 and 9 km in altitude. Again, we note that this result contrasts with the substantial qualitative improvements in vertical velocity observed for the RA method in Figs. 3 and 4, and speculate that the discrepancy may be due to the RMSE scores being biased in favor of pixel-to-pixel comparisons of smoother analysis fields. To verify this assertion, we recompute updraft and downdraft FSSs for the 5th- and 95th-percentile thresholds at each horizontal slice of the analysis grid. Figures 7c and 7d show that the RA method exhibits modest improvements in FSS compared to the 2D Cressman method between 5 and 9 km in altitude, despite having slightly higher RMSE scores over the same altitude range. This may indicate a slight spatial displacement in vertical velocities, or that errors from vertical velocities below the FSS thresholds contribute significantly to the RMSE in this region. Furthermore, Fig. 7d illustrates particularly low FSS values and high ensemble spread for the 3D Cressman method, further emphasizing its inability to accurately capture highly dynamic regions of the storm. Overall, Fig. 7 shows that the RA method produces lower error scores, and more reliably retrieves updraft and downdraft regions throughout the depth of the storm.

While the error statistics presented thus far are useful for understanding the error distribution over the entire domain (Table 1), for various magnitudes (Fig. 6), or over a range of altitudes (Fig. 7), some analysts may prioritize accurately identifying the maximum values of certain quantities, such as the maximum horizontal velocity in the low levels for forecasting severe wind hazards, or the maximum updraft velocity for diagnosing storm severity or verifying/informing convection parameterizations in numerical weather prediction models. These maximum value statistics are provided in Fig. 8. As expected, the pregridded retrieval methods considerably underestimate the true horizontal and vertical velocity maxima, due to oversmoothing of important radial velocity information. Encouragingly, the RA method is able to accurately reproduce maximum vertical velocity values below 8 km, before considerably underestimating those farther aloft (at least partly due to the aforementioned observational gaps at high altitude). However, the poor performance of the control

experiment at the domain top and bottom in Fig. 8b also indicates some of these errors may be partly attributed to the ill-posed vertical boundary conditions in the mass continuity equation.

Finally, the oversmoothing in pregridded methods is especially apparent in the derived maximum vorticity and divergence fields in Figs. 8c and 8d, and the qualitative effects of underestimating these dynamic quantities were observed in Figs. 5b and 5c, where simulated trajectories appeared unrealistically uniform or laminar. The RA method considerably outperforms these methods, while also underestimating the true vorticity and divergence maxima throughout the depth of the storm. The most notable improvements gained by the RA method occur in the lower levels in Fig. 4f, where the maximum convergence near the surface is retrieved accurately. This confirms the qualitative findings from the simulated radial velocity data in Fig. 2 and simulated trajectories in Fig. 5, which both suggested the superior characterization of low-level convergence and inflow in the RA method. The accurate characterization of low-level storm inflow is crucial for resolving storm dynamics (Coffer et al. 2023), and this factor likely contributes to the qualitative and quantitative improvements observed from the RA method.

c. Real data case study

To investigate the applicability of the OSSE findings to real cases, we now introduce a supercell case study from 27 November 2014 in Brisbane, Australia. This case was selected as an ideal candidate for verifying the ability to retrieve strong vertical motions, which are known to have occurred based on prior studies (e.g., Parackal et al. 2015; Soderholm et al. 2017). The storm tracked northward through the Brisbane metropolitan area between 0200 and 0700 UTC, resulting in over AUD 1.5 billion in insured losses due to giant hail (~70 mm), severe straight-line winds, and localized flooding (Insurance Council of Australia 2017). The storm was observed by two S-band, Doppler radars: the CP2 research radar and the operational Mount Stapylton radar (hereafter MS), located to the southwest and southeast of Brisbane, respectively. Figure 9 shows the analysis domain for our dual-Doppler 3D wind retrieval, centered on the leading edge of the storm at (27.4°S, 153.05°E), situated almost entirely within the dual-Doppler lobes (~40 km between radars). Quality controlled radar volumes are sourced from the Australian Unified Radar Archive (Soderholm et al. 2022), containing 9 and 14 constant elevation sweeps for CP2 and MS, respectively.[9]

To process real data, two minor methodological adjustments are made to the partially idealized OSSE retrieval methodology. First, we apply a reflectivity-based correction to account for the effects of hydrometeor terminal velocities in radial velocity measurements. We utilize the terminal velocity

---

[9] Exact elevations are $\theta = (0.9°, 1.7°, 2.4°, 3.2°, 4.7°, 6.5°, 9.1°, 12.8°, 17.8°)$ for CP2, and $\theta = (0.5°, 0.9°, 1.3°, 1.8°, 2.4°, 3.1°, 4.2°, 5.6°, 7.4°, 10.0°, 13.3°, 17.9°, 23.9°, 32.0°)$ for MS. Half-power beamwidths and range gates are 0.93° and 150 m for CP2 and 1.0° and 250 m for MS.
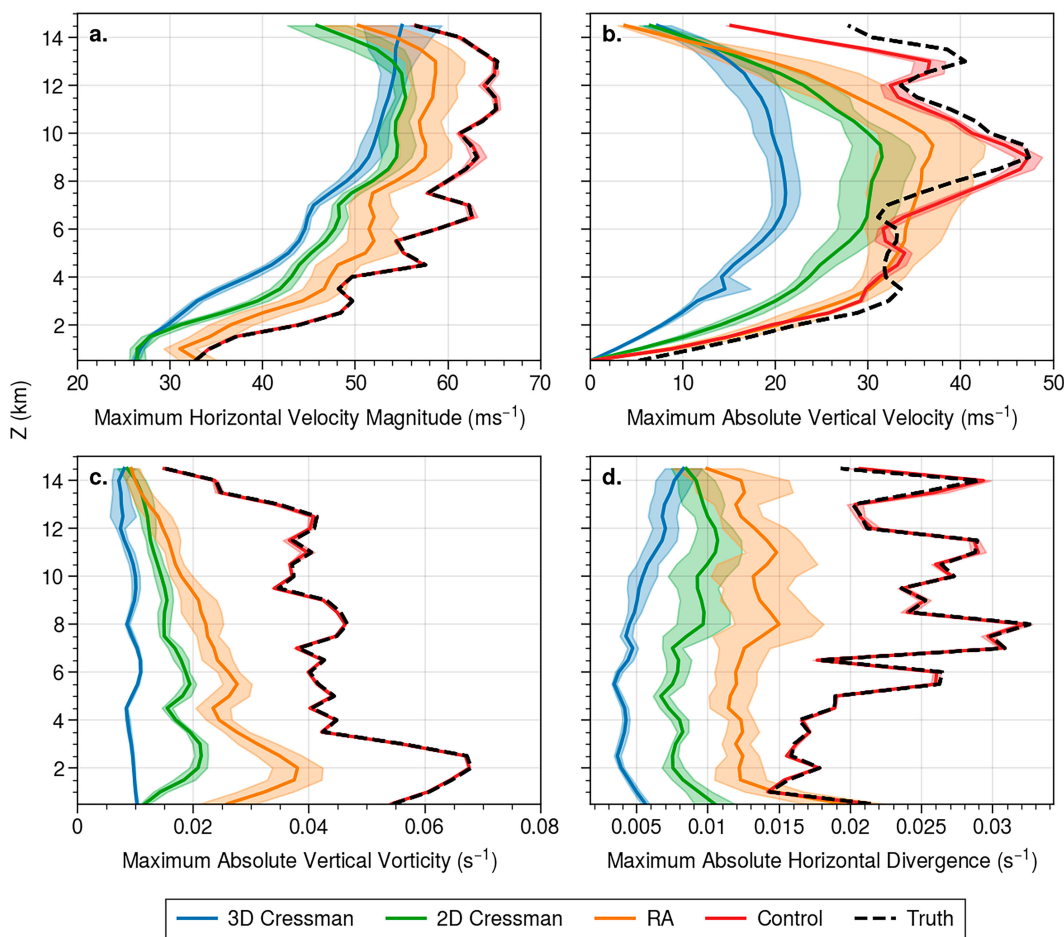
Fig. 8. As in Fig. 7, but for maximum values of various quantities. The "true" model maximum values for these quantities are also shown in black dashed lines.

correction from the PyDDA package (Jackson et al. 2020), which varies from pure rain (Joss and Waldvogel 1970) to hail (Conway and Zrnić 1993) based on reflectivity thresholds. Second, we employ a simple horizontal advection correction to account for storm motion during the finite sampling period of radar volumes (e.g., Shapiro et al. 2009). The average horizontal advection velocity was calculated by taking the mean optical flow velocity (PySTEPS Lucas–Kanade implementation; Pulkkinen et al. 2019) for constant-altitude reflectivity slices between 1 and 5 km. Reflectivity grids were interpolated using the variational method outlined recently by B22. Observations from both radars were then shifted to the midpoint of the radar scanning time (0644 UTC) according to this average advection velocity prior to gridding or ingestion into the RA retrieval method. The analysis grid spans 60 and 50 km in the x and y axes, respectively, with a 500 m grid spacing. The grid is also spaced at 500 m increments in the z axis, but extends from 200 to 15 200 m MSL.

Figure 10 presents retrieved wind fields for the 2014 Brisbane hailstorm for each of the retrieval strategies discussed in this study. Horizontal streamlines in Figs. 10a–c are broadly similar across each method, resolving the southeasterly change

to the rear of the storm, and the strong mesocyclone circulation around highlight region 1 (as in Soderholm et al. 2017). Similar to the OSSE experiments in section 3a, horizontal and vertical velocities appear visually smoother in the pregridded retrievals (Figs. 10b,c), whereas small-scale structures in the vertical velocity fields, especially in the southwest of the domain, are resolved in the RA method (Fig. 10a). The largest vertical velocity differences between methods occur to the north of region 1 in this case. All three methods capture an updraft at the leading edge of the storm, likely a result of dynamic lifting ahead of the advancing gust front. However, the leading-edge updraft appears spuriously large and intense for 2D Cressman method when compared to the main rotating updraft in region 1. This is likely a result of boundary effects caused by data scarcity at the storm's leading edge in the 2D Cressman method (refer to appendix B), which underrepresents the spatial extent of radar data, particularly in the low levels (e.g., region 4 in Fig. 2g).

The overall position and southwest–northeast orientation of the updraft at $Z = 4$ km in highlight region 1 is retrieved by all retrieval methods. The shape of the inflow notch (indicated by the reflectivity contour in region 1) aligns well with
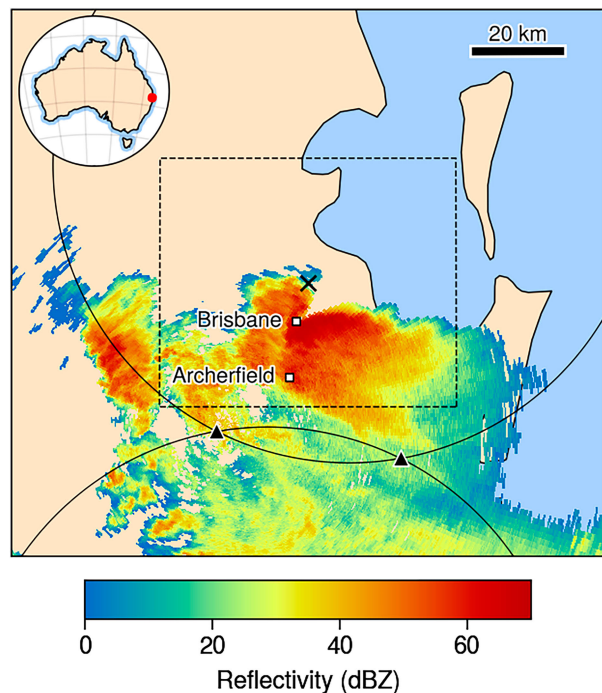
FIG. 9. Experimental setup for the Brisbane hailstorm case study, indicating the center and extent of the analysis domain with a cross and dotted lines, respectively. Radars used in the analysis (triangles) and dual-Doppler lobes (solid lines) are shown alongside radar reflectivity from the 3.7° tilt at 0642 UTC.

the mesocyclone position and updraft orientation. Maximum updraft velocities for the RA method in Fig. 10a are much larger than that in the pregridded retrievals. Quantitatively, the maximum updraft speed retrieved by the RA method was 52.5 m s^{-1}, compared to 30.8 and 30.3 m s^{-1} for the 3D and 2D Cressman methods, respectively. While the veracity of the increased updraft velocities in the RA method cannot be directly confirmed in this real data experiment, the retrieved maximum value of ~50 m s^{-1} is closer to the mean values found in supercell simulations (Peters et al. 2020), and the OSSE experiments suggest it is likely a more accurate estimate of the true updraft speed. Further qualitative evidence of considerable updraft strength for this case is given by the large, elevated hail core (>65 dBZ, ~0 Z_{DR}), and bounded weak echo region well-above the freezing level (Soderholm et al. 2017; Brook et al. 2021).

Figures 10d–f corroborate these findings for a vertical cross section through the main updraft. The pregridded retrievals adequately estimate the position and shape of the updraft (as evidenced by the position of the bounded weak echo region in region 2 for Figs. 10d–f), but underestimate vertical velocities relative to the RA method. Notably, the strong RFD signature shown in region 2 for the RA method is also entirely absent in the pregridded retrievals. This outflow feature is well documented for this event (Parackal et al. 2015; Soderholm et al. 2017; Brook et al. 2021), resulting in severe wind gusts and strong (>10 km) hail advection toward the west. The RA

method's ability to resolve the RFD also leads to stronger divergence at lowest grid level, resulting in 40.2 m s^{-1} maximum horizontal winds in the outflow region (west of region 2 in Figs. 10d–f), compared with 24.0 and 30.3 m s^{-1} for the 3D and 2D Cressman methods, respectively. By comparison, a 39.2 m s^{-1} wind gust was measured roughly 10 min prior to the analysis time at Archerfield Airport (refer to Fig. 9), along with widespread ~20 m s^{-1} gusts throughout Brisbane's western suburbs (Parackal et al. 2015; Soderholm et al. 2017). While the difficulties associated with comparing retrieved winds aloft to wind measurements at the surface preclude a more quantitative comparison,[10] the strength and position of these surface measurements provide qualitative support for the RFD signature retrieved by the RA method.

Overall, the results for the Brisbane supercell case study support the findings of the OSSEs, suggesting that the advances made in this study may be realized in practice. However, the practical implementation of the RA method also warrants a brief discussion on the relative computational efficiency of each retrieval method. The introduction of an additional interpolation step in the observational operator for the RA method (i.e., $\mathcal{PC}$ rather than $\mathcal{P}_g$, refer to section 2b) roughly doubles the computational cost of the operator relative to the pregridded methods. The observational constraint is only one of four otherwise identical constraints in the cost function [Eq. (4)]; however, it must be performed twice at each iteration of the optimization procedure (forward and adjoint). In our implementation, the overall execution time of the RA method is thus ~2.3 times slower than the pregridded methods for this case study (15.5 vs 6.8 s, respectively). However, when considering the additional costs involved in pregridding the observations (4.4 and 8.0 s for 3D and 2D Cressman, respectively), the RA method introduces a relatively modest computational overhead.

4. Summary and future directions

The use of 3D wind retrievals for applications such as storm-scale dynamics research, constraining convection parameterizations, and nowcasting wind hazards is set to increase in the future, with open-source implementations such as PyDDA (Jackson et al. 2019) increasing in popularity within the radar science community. Despite previous studies noting that spatial interpolation is a significant source of error in these analyses, there is currently no direct comparison or practical recommendations for spatial interpolation methodologies. In this study, we aimed to fill these knowledge gaps for severe convective storms by analyzing supercell 3D wind retrievals for both OSSEs based on high resolution simulations (50 m horizontal grid spacing) and a real data case study. In both experiments, radar data were either pregridded using two- or three-dimensional Cressman weighted average gridding

---

[10] Retrieved winds at the lowest grid level (~170 m AGL) are calculated using a free slip, impermeable, constant elevation boundary condition, and do not resolve the small-scale interactions (i.e., due to topography and surface roughness) that strongly influence wind gust measurements at the surface.
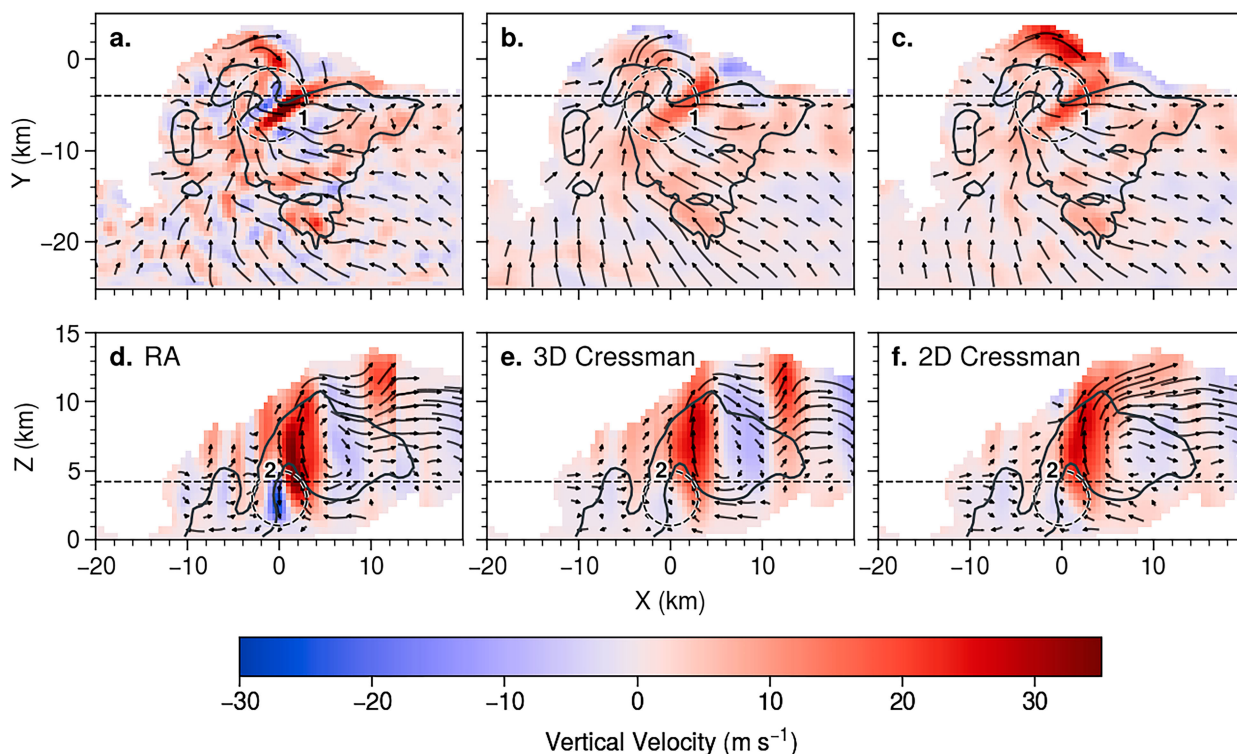
FIG. 10. Retrieved winds for the 2014 Brisbane hailstorm for each of the methods discussed in this study. The 50 dBZ reflectivity contour is shown in gray, vertical velocities are shaded, and streamlines illustrate along-plane velocities. (a)–(c) Horizontal cross sections at $Z = 4$ km and (d)–(f) vertical cross sections at $Y = -4$ km. Straight dashed lines indicate the positions of the corresponding cross sections, and circular regions are highlighted for discussion in the text.

techniques, or assimilated directly into the variational wind retrieval algorithm using the radar assimilation (RA) method. The outcomes of these experiments broadly confirmed the hypothesis that directly assimilating, rather than pregridding, radial velocity measurements results in more accurate 3D wind retrievals for the most challenging cases (i.e., the supercell thunderstorms investigated here). The main findings from our study are summarized below:

- Spatial interpolation has a very large effect on the quality of wind retrievals. Wind error magnitudes increased 300%–350% when using operational radar scanning geometries with realistic observational errors, compared to a control experiment with "perfect" observations defined everywhere within precipitating regions.
- Spatial interpolation errors are greater for methods using pregridded radial velocities, resulting in average total error magnitudes of ~7.3 m s^{-1}, compared with 6.2 m s^{-1} in the RA experiments.
- If restricted to pregridded retrieval strategies, the 2D Cressman gridding performs similarly to the 3D Cressman method in terms of error scores (albeit less consistently), but more accurately resolves highly dynamic regions of the simulated storm due to the increased detail in radial velocity inputs.
- The RA method qualitatively resolves important dynamic features that the pregridded methods do not, including the

mesocyclone, rear-flank downdraft, overturning circulation signatures, and all major updraft trajectory pathways present within the modeled storm.
- The RA method also reduces errors in vertical vorticity, horizontal divergence, and all three wind components (to an RMSE of ~3.5 m s^{-1} for u, v, and w), more accurately retrieves the maximum values of these quantities, and more reliably retrieves regions with intense updrafts/downdrafts as evidenced by greater fractions skill scores.
- An investigation of a supercell case study showed promising agreement with the OSSE findings above, suggesting that the improvements observed for the RA method are applicable to real data in practice.

Future work will be aimed at generalizing these results using further OSSEs and real data examples containing different weather phenomena (e.g., isolated convection, stratiform precipitation, and tropical cyclones). These studies should investigate whether the RA method shows comparable benefits to those described here for strong convection. We also aim to assess how the improvements achieved with the RA technique will function alongside other recent 3D wind retrieval developments, such as vertical vorticity constraints (Shapiro et al. 2009; Potvin et al. 2012b) and spatially variable advection corrections (Shapiro et al. 2010). This is particularly relevant for the OSSE portion of our study, which assumed an instantaneous radar volume coverage pattern to control for

the effects of nonstationarity. Finally, we aim to verify our findings against vertical velocity measurements from observed cases, as demonstrated recently in Gebauer et al. (2022).

Data availability statement. Radar data used in this study are available from the Australian Unified Radar Archive (https://www.openradar.io/). Model data are available upon request due to the considerable storage requirements required to house the high-resolution model outputs used in this study.

APPENDIX A

Projection/Interpolation Operators

The extensive use of pregridded radial velocities in recent wind retrieval publications (e.g., North et al. 2017; Dahl et al. 2019; Oue et al. 2019; Gebauer et al. 2022), and exclusive use in the open-source software packages MultiDop (Lang et al. 2017) and PyDDA (Jackson et al. 2019), has obscured the once obvious distinction between pregridded and direct radar assimilation techniques (e.g., Gao et al. 1999; Rihan et al. 2005). In this section, we aim to clarify this distinction by providing a technical explanation of the projection and interpolation operators used in the RA method.

In the pregridded observational constraint in Eq. (6), the difference between the analysis fields and radar observations is a straightforward calculation on the analysis grid. However, when the observations are provided in their native observation geometry in the RA method, the comparison requires a forward operator to interpolate the analysis fields to the observation locations. Perhaps the simplest interpolation operator is a trilinear interpolation of the eight surrounding grid points (e.g., Gao et al. 1999, 2004; B22). We tested this operator in our OSSE experiments (not shown) and found it performed poorly compared to the Cressman interpolation operator used throughout this study. We attribute these shortcomings to the propagation of observational noise and the introduction of first-order discontinuities into the analysis, which are characteristic of linear interpolation (Trapp and Doswell 2000; Askelson et al. 2000; B22). Wind retrievals are particularly sensitive to these effects due to their reliance on finite difference derivatives in the mass continuity equation (Testud and Chong 1983), meaning a Cressman interpolation operator (with the superior filtering qualities) is more suited to this application. The radius of influence in $\mathcal{C}$ was set to 1400 m for all of our experiments, having been experimentally optimized in a set of preliminary experiments (not shown), by setting it as a free variable in the Bayesian parameter optimization process described in section 3b(3). As in Potvin et al. (2012c), we find our results are not particularly sensitive to the provision of this parameter.
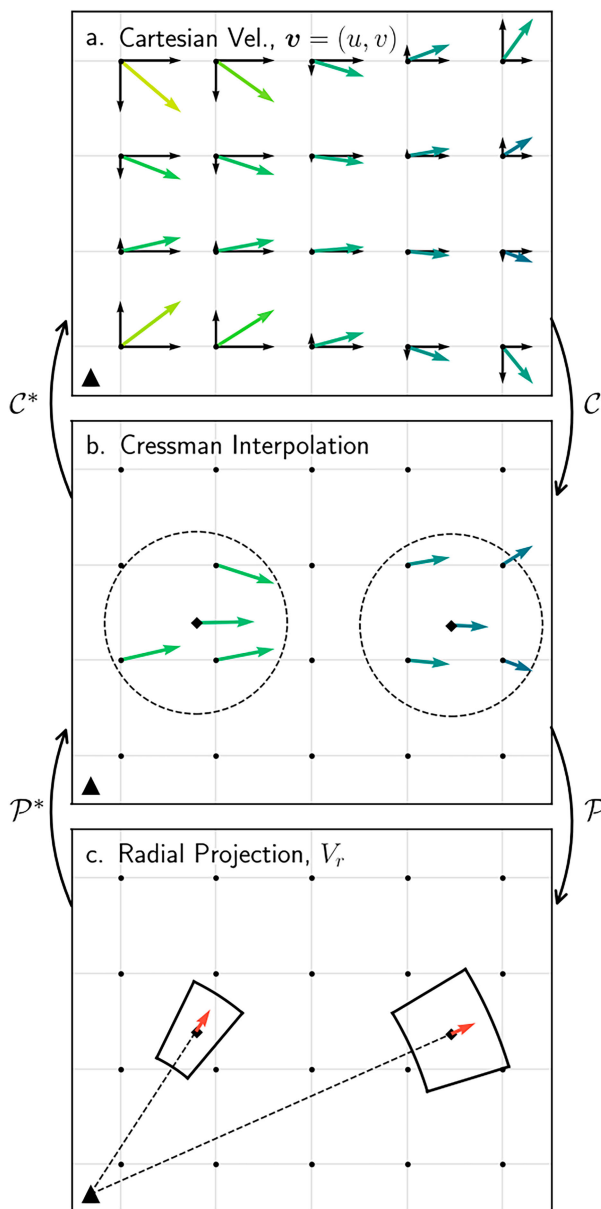

FIG. A1. An illustration of the Cressman interpolation ($\mathcal{C}$) and radial velocity projection ($\mathcal{P}$) operators and their adjoints ($\mathcal{C}^*$ and $\mathcal{P}^*$, respectively). (a) Colored arrows show a gridded, Cartesian velocity field (warmer colors represent larger magnitudes), with the individual velocity components shown in black. (b) Cressman interpolation from grid points within a radius of influence (dotted circles) of two radar observation points (diamonds). (c) Radial projection (red arrows) of the interpolated Cartesian velocities at the observation points. The black triangle indicates the radar position, and black circles and grid lines indicate the Cartesian analysis grid.

We illustrate the implementation of the Cressman interpolation operator with a 2D Cartesian velocity field in Fig. A1a. Computationally speaking, the operator is set up prior to the analysis by noting which grid points are within a constant radius of influence from each observation (as in Fig. A1b), and

the Cressman weights are precalculated according to Eq. (10). These weights are used to interpolate the analysis fields to the observation points, before the newly interpolated Cartesian velocities are projected into radial velocities by the radial projection operator in Fig. A1c. Importantly, the $\mathcal{C}$ and $\mathcal{P}$ operators are not commutative, meaning the order in which they are applied is significant. The order presented here is more accurate as the radial projection takes place directly at the observation location, instead of attempting to interpolate a strictly nonsmooth quantity in the reverse order. Also note that the Cressman interpolation in Fig. A1b is well defined, as more than one grid point will always be within a radius of influence from the observation points (due to the regularity of the Cartesian grid). Conversely, if one attempted to interpolate the radial velocity data in Fig. A1c to the grid points with a similar radius of influence (as is attempted in pregridding methods), many grid points would have no valid observations, producing data voids in the result. The radius of influence must then be inflated to mask the data acquisition gaps (which are very large between constant elevation sweeps), thereby filtering important information from the resulting data field.

APPENDIX B

Data Boundary Effects

Dual-Doppler 3D wind retrievals are severely hampered by data coverage limitations that arise from two main sources: 1) weather radars only return reliable information within regions with backscatterers (usually within clouds) and 2) radars only take measurements at set locations defined by their scanning strategy, meaning there are significant data gaps even within strong echo regions. We showed in our control OSSE experiments that eliminating the latter error source (by assigning valid data everywhere within echo regions) reduced total errors by more than 300% compared to a standard operational scanning pattern. Furthermore, if error source 1 is eliminated (by assigning valid radial velocity data everywhere, not shown), the retrieval errors drop to near zero. However, when both error sources are present (as is the case in real data), care must be taken to ensure boundary errors do not propagate into valid data regions. Here, we detail how imposing boundary conditions on the observational constraint can mitigate these effects.

First, the errors introduced at data boundaries may be best understood by considering an isolated, high-altitude data point with strong horizontal velocity, and no vertical velocity. If the radar is situated parallel to the strong horizontal velocities, the measured radial velocity will be large—such

that when it is projected back into Cartesian coordinates, a significant portion will be interpreted as vertical velocity (especially for higher elevation angles). In the absence of valid radial velocities in the surrounding grid points, the mass continuity and smoothing constraints will then construct a spurious updraft/downdraft region to match the isolated observation point, which is what we observed in isolated regions in Fig. 3a, and along data boundaries in other studies (e.g., Collis et al. 2010). We have found that setting the vertical gradient of the observational constraint equal to zero at these boundary grid points (boundaries are defined as neighbors to a data void), significantly reduces these errors, and we have designed a simple OSSE to demonstrate these effects.

The simple OSSE contains purely westerly flow ($V = W = 0$), which we have derived to mimic the thought experiment above. To this end, we fit a fourth-order polynomial to the average u profile from the ARPS simulation (containing an upper-level jet at ~10 km): $U = 0.0021z^4 - 0.067z^3 + 0.30z^2 + 5.14z + 0.46$, where z is in kilometers. The domain is the same as in our model OSSEs (80 km $\times$ 60 km $\times$ 15 km in the x, y, and z dimensions, with an isotropic 500 m grid spacing), and the radars are positioned roughly parallel to the westerly flow at position 7 in Fig. 1. We limit the radar echo regions to an elliptical shape spanning the entire domain horizontally, with a depth of roughly 6 km, centered at a height of 10 km. We then simulated radial velocities as in section 3a, and finally retrieved vertical winds using the radar assimilation retrieval method both without and with boundary masking (Figs. B1a,c and Figs. B1b,d, respectively).

While the westerly horizontal flow is retrieved well in the experiment with no boundary masking, significant spurious vertical motions are created within the valid data region (illustrated by the gray shading). These spurious features within the cloud are associated with the aforementioned data boundary effects, while the strongest updrafts/downdrafts at the horizontal edges of the cloud are induced by the mass continuity constraint at the domain edges (which also creates spurious return flow near the surface in Fig. B1c). In Figs. B1b and B1d, the w component of the observational adjoint is set to zero (i.e., $\mathcal{C}^* \mathcal{P}^* V_r = 0$) at the boundary points (plotted in black). Note that the spurious updraft features within the cloud are almost completely eliminated, leaving only those outside the valid data region associated with mass continuity. The total RMSE (within the valid data region) drops from 2.47 to 0.95 m s^{-1} as a result of this computational procedure. We have noted a similar elimination of spurious updrafts in the model OSSEs throughout this study.
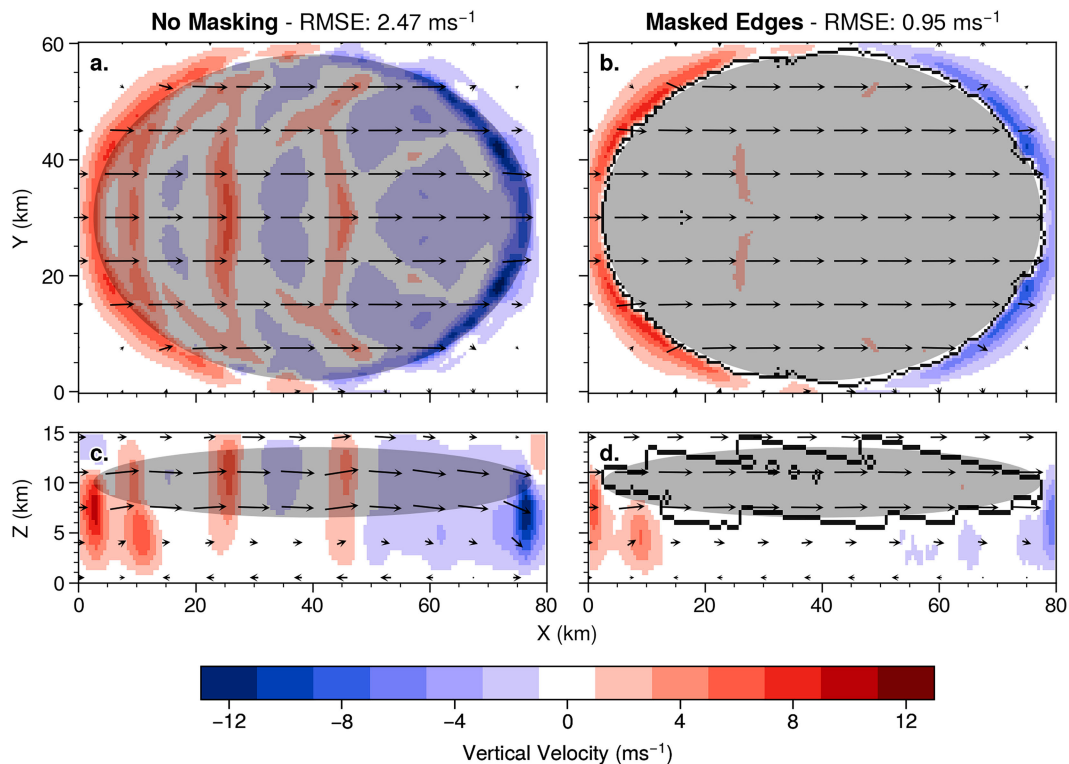
FIG. B1. Wind retrievals from the simple observing system simulation experiment described in appendix B. (a),(b) Horizontal cross sections at $Z = 10$ km, along with (c),(d) vertical cross sections at $Y = 30$ km. The radar echo region is highlighted in gray, while vertical winds are shaded and arrows show wind velocities along the cross-section plane. Pixels containing radar observation boundaries are colored black in (b) and (d).

REFERENCES

Askelson, M. A., J.-P. Aubagnac, and J. M. Straka, 2000: An adaptation of the Barnes filter applied to the objective analysis of radar data. *Mon. Wea. Rev.*, **128**, 3050–3082, https://doi.org/10.1175/1520-0493(2000)128<3050:AAOTBF>2.0.CO;2.

Atlas, D., R. C. Srivastava, and R. S. Sekhon, 1973: Doppler radar characteristics of precipitation at vertical incidence. *Rev. Geophys.*, **11**, 1–35, https://doi.org/10.1029/RG011i001p00001.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409, https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2.

Barth, A., A. Alvera-Azcárate, C. Troupin, and J.-M. Beckers, 2022: DINCAE 2.0: Multivariate convolutional neural network with error estimates to reconstruct sea surface temperature satellite and altimetry observations. *Geosci. Model Dev.*, **15**, 2183–2196, https://doi.org/10.5194/gmd-15-2183-2022.

Batchelor, G. K., 1953: The conditions for dynamical similarity of motions of a frictionless perfect-gas atmosphere. *Quart. J. Roy. Meteor. Soc.*, **79**, 224–235, https://doi.org/10.1002/qj.49707934004.

Betten, D. P., M. I. Biggerstaff, and C. L. Ziegler, 2018: Three-dimensional storm structure and low-level boundaries at different stages of cyclic mesocyclone evolution in a high-precipitation tornadic supercell. *Adv. Meteor.*, **2018**, 9432670, https://doi.org/10.1155/2018/9432670.

Bousquet, O., P. Tabary, and J. P. du Châtelet, 2007: On the value of operationally synthesized multiple-Doppler wind fields. *Geophys. Res. Lett.*, **34**, L22813, https://doi.org/10.1029/2007GL030464.

Briggs, I. C., 1974: Machine contouring using minimum curvature. *Geophysics*, **39**, 39–48, https://doi.org/10.1190/1.1440410.

Brook, J. P., A. Protat, J. Soderholm, J. T. Carlin, H. McGowan, and R. A. Warren, 2021: HailTrack—Improving radar-based hailfall estimates by modeling hail trajectories. *J. Appl. Meteor. Climatol.*, **60**, 237–254, https://doi.org/10.1175/JAMC-D-20-0087.1.

——, ——, ——, R. A. Warren, and H. McGowan, 2022: A variational interpolation method for gridding weather radar data. *J. Atmos. Oceanic Technol.*, **39**, 1633–1654, https://doi.org/10.1175/JTECH-D-22-0015.1.

Chong, M., and S. Cosma, 2000: A formulation of the continuity equation of MUSCAT for either flat or complex terrain. *J. Atmos. Oceanic Technol.*, **17**, 1556–1565, https://doi.org/10.1175/1520-0426(2000)017<1556:AFOTCE>2.0.CO;2.

Coffer, B. E., M. D. Parker, J. M. Peters, and A. R. Wade, 2023: Supercell low-level mesocyclones: Origins of inflow and vorticity. *Mon. Wea. Rev.*, **151**, 2205–2232, https://doi.org/10.1175/MWR-D-22-0269.1.

Collis, S., A. Protat, and K.-S. Chung, 2010: The effect of radial velocity gridding artifacts on variationally retrieved vertical velocities. *J. Atmos. Oceanic Technol.*, **27**, 1239–1246, https://doi.org/10.1175/2010JTECHA1402.1.

Conway, J. W., and D. S. Zrnić, 1993: A study of embryo production and hail growth using dual-Doppler and multiparameter

radars. *Mon. Wea. Rev.*, **121**, 2511–2528, https://doi.org/10.1175/1520-0493(1993)121<2511:ASOEPA>2.0.CO;2.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Dahl, N. A., A. Shapiro, C. K. Potvin, A. Theisen, J. G. Gebauer, A. D. Schenkman, and M. Xue, 2019: High-resolution, rapid-scan dual-Doppler retrievals of vertical velocity in a simulated supercell. *J. Atmos. Oceanic Technol.*, **36**, 1477–1500, https://doi.org/10.1175/JTECH-D-18-0211.1.

Dolan, B. A., and S. A. Rutledge, 2007: An integrated display and analysis methodology for multivariable radar data. *J. Appl. Meteor. Climatol.*, **46**, 1196–1213, https://doi.org/10.1175/JAM2524.1.

Doviak, R., and D. Zrnić, 1993: *Doppler Radar and Weather Observations*. 2nd ed. Dover, 562 pp.

Duda, J. D., and W. A. Gallus Jr., 2010: Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Wea. Forecasting*, **25**, 190–206, https://doi.org/10.1175/2009WAF2222338.1.

Evaristo, R., R. Reinoso-Rondinel, S. Tromel, and C. Simmer, 2021: Validation of wind fields retrieved by dual-Doppler techniques using a vertically pointing radar. *2021 21st Int. Radar Symp.*, Berlin, Germany, IEEE, https://doi.org/10.23919/IRS51887.2021.9466173.

Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev.*, **127**, 2128–2142, https://doi.org/10.1175/1520-0493(1999)127<2128:AVMFTA>2.0.CO;2.

——, ——, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic Technol.*, **21**, 457–469, https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2.

Gebauer, J. G., A. Shapiro, C. K. Potvin, N. A. Dahl, M. I. Biggerstaff, and A. A. Alford, 2022: Evaluating vertical velocity retrievals from vertical vorticity equation constrained dual-Doppler analysis of real, rapid-scan radar data. *J. Atmos. Oceanic Technol.*, **39**, 1591–1610, https://doi.org/10.1175/JTECH-D-21-0136.1.

Goldstein, T., and S. Osher, 2009: The split Bregman method for L1-regularized problems. *SIAM J. Imaging Sci.*, **2**, 323–343, https://doi.org/10.1137/080725891.

Head, T., M. Kumar, H. Nahrstaedt, G. Louppe, and I. Shcherbatyi, 2021: Scikit-optimize/scikit-optimize. Zenodo, https://doi.org/10.5281/zenodo.5565057.

Helmus, J. J., and S. M. Collis, 2016: The Python ARM Radar Toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *J. Open Res. Software*, **4**, e25, https://doi.org/10.5334/jors.119.

Insurance Council of Australia, 2017: ICA catastrophe database. Insurance Council of Australia, accessed 8 August 2023, https://insurancecouncil.com.au/industry-members/data-hub/.

Jackson, R., S. Collis, T. Lang, C. Potvin, and T. Munson, 2019: PyDDA: A new Pythonic wind retrieval package. *18th Annual Scientific Computing with Python Conf.*, Austin, TX, SciPy, 111–117, https://doi.org/10.25080/Majora-7ddc1dd1-010.

——, ——, ——, ——, and ——, 2020: PyDDA: A Pythonic direct data assimilation framework for wind retrievals. *J. Open Res. Software*, **8**, 20, https://doi.org/10.5334/jors.264.

Joss, J., and A. Waldvogel, 1970: Raindrop size distributions and Doppler velocities. Preprints, *14th Conf. on Radar Meteorology*, Tucson, AZ, Amer. Meteor. Soc., 153–156.

Kosiba, K. A., and J. Wurman, 2014: Finescale dual-Doppler analysis of hurricane boundary layer structures in Hurricane Frances (2004) at landfall. *Mon. Wea. Rev.*, **142**, 1874–1891, https://doi.org/10.1175/MWR-D-13-00178.1.

Kumar, V. V., C. Jakob, A. Protat, C. R. Williams, and P. T. May, 2015: Mass-flux characteristics of tropical cumulus clouds from wind profiler observations at Darwin, Australia. *J. Atmos. Sci.*, **72**, 1837–1855, https://doi.org/10.1175/JAS-D-14-0259.1.

Kumjian, M. R., K. Lombardo, and S. Loeffler, 2021: The evolution of hail production in simulated supercell storms. *J. Atmos. Sci.*, **78**, 3417–3440, https://doi.org/10.1175/JAS-D-21-0034.1.

Labbouz, L., Z. Kipling, P. Stier, and A. Protat, 2018: How well can we represent the spectrum of convective clouds in a climate model? Comparisons between internal parameterization variables and radar observations. *J. Atmos. Sci.*, **75**, 1509–1524, https://doi.org/10.1175/JAS-D-17-0191.1.

Lang, T., M. Souto, S. Khobahi, and R. Jackson, 2017: MultiDop. Zenodo, https://doi.org/10.5281/zenodo.1035904.

Liou, Y.-C., S.-F. Chang, and J. Sun, 2012: An application of the immersed boundary method for recovering the three-dimensional wind fields over complex terrain using multiple-Doppler radar data. *Mon. Wea. Rev.*, **140**, 1603–1619, https://doi.org/10.1175/MWR-D-11-00151.1.

Majcen, M., P. Markowski, Y. Richardson, D. Dowell, and J. Wurman, 2008: Multipass objective analyses of Doppler radar data. *J. Atmos. Oceanic Technol.*, **25**, 1845–1858, https://doi.org/10.1175/2008JTECHA1089.1.

Markowski, P. M., T. P. Hatlee, and Y. P. Richardson, 2018: Tornadogenesis in the 12 May 2010 supercell thunderstorm intercepted by VORTEX2 near Clinton, Oklahoma. *Mon. Wea. Rev.*, **146**, 3623–3650, https://doi.org/10.1175/MWR-D-18-0196.1.

Mohr, C. G., and R. L. Vaughan, 1979: An economical procedure for Cartesian interpolation and display of reflectivity factor data in three-dimensional space. *J. Appl. Meteor.*, **18**, 661–670, https://doi.org/10.1175/1520-0450(1979)018<0661:AEPFCI>2.0.CO;2.

Nicol, J. C., R. J. Hogan, T. H. M. Stein, K. E. Hanley, P. A. Clark, C. E. Halliwell, H. W. Lean, and R. S. Plant, 2015: Convective updraught evaluation in high-resolution NWP simulations using single-Doppler radar measurements. *Quart. J. Roy. Meteor. Soc.*, **141**, 3177–3189, https://doi.org/10.1002/qj.2602.

North, K. W., M. Oue, P. Kollias, S. E. Giangrande, S. M. Collis, and C. K. Potvin, 2017: Vertical air motion retrievals in deep convective clouds using the ARM scanning radar network in Oklahoma during MC3E. *Atmos. Meas. Tech.*, **10**, 2785–2806, https://doi.org/10.5194/amt-10-2785-2017.

Oue, M., P. Kollias, A. Shapiro, A. Tatarevic, and T. Matsui, 2019: Investigation of observational error sources in multi-Doppler-radar three-dimensional variational vertical air motion retrievals. *Atmos. Meas. Tech.*, **12**, 1999–2018, https://doi.org/10.5194/amt-12-1999-2019.

Parackal, K., and Coauthors, 2015: Investigation of damage: Brisbane, 27 November 2014 severe storm event. Cyclone Testing Station Tech. Rep. 60, 47 pp., https://www.jcu.edu.au/__data/assets/pdf_file/0003/321996/Technical-Report-60-Investigation-of-Damage-Brisbane-27-November-2014-Severe-Storm-Event.pdf/_noproxycache.

Park, S.-G., and D.-K. Lee, 2009: Retrieval of high-resolution wind fields over the southern Korean Peninsula using the Doppler weather radar network. *Wea. Forecasting*, **24**, 87–103, https://doi.org/10.1175/2008WAF2007084.1.

Peters, J. M., H. Morrison, C. J. Nowotarski, J. P. Mulholland, and R. L. Thompson, 2020: A formula for the maximum vertical velocity in supercell updrafts. *J. Atmos. Sci.*, **77**, 3747–3757, https://doi.org/10.1175/JAS-D-20-0103.1.

Potvin, C. K., A. Shapiro, T.-Y. Yu, J. Gao, and M. Xue, 2009: Using a low-order model to detect and characterize tornadoes in multiple-Doppler radar data. *Mon. Wea. Rev.*, **137**, 1230–1249, https://doi.org/10.1175/2008MWR2446.1.

——, D. Betten, L. J. Wicker, K. L. Elmore, and M. I. Biggerstaff, 2012a: 3DVAR versus traditional dual-Doppler wind retrievals of a simulated supercell thunderstorm. *Mon. Wea. Rev.*, **140**, 3487–3494, https://doi.org/10.1175/MWR-D-12-00063.1.

——, A. Shapiro, and M. Xue, 2012b: Impact of a vertical vorticity constraint in variational dual-Doppler wind analysis: Tests with real and simulated supercell data. *J. Atmos. Oceanic Technol.*, **29**, 32–49, https://doi.org/10.1175/JTECH-D-11-00019.1.

——, L. J. Wicker, and A. Shapiro, 2012c: Assessing errors in variational dual-Doppler wind syntheses of supercell thunderstorms observed by storm-scale mobile radars. *J. Atmos. Oceanic Technol.*, **29**, 1009–1025, https://doi.org/10.1175/JTECH-D-11-00177.1.

Protat, A., and I. Zawadzki, 1999: A variational method for real-time retrieval of three-dimensional wind field from multiple-Doppler bistatic radar network data. *J. Atmos. Oceanic Technol.*, **16**, 432–449, https://doi.org/10.1175/1520-0426(1999)016<0432:AVMFRT>2.0.CO;2.

——, and ——, 2000: Optimization of dynamic retrievals from a multiple-Doppler radar network. *J. Atmos. Oceanic Technol.*, **17**, 753–760, https://doi.org/10.1175/1520-0426(2000)017<0753:OODRFA>2.0.CO;2.

Pulkkinen, S., D. Nerini, A. A. Pérez Hortal, C. Velasco-Forero, A. Seed, U. Germann, and L. Foresti, 2019: Pysteps: An open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev.*, **12**, 4185–4219, https://doi.org/10.5194/gmd-12-4185-2019.

Ravasi, M., and I. Vasconcelos, 2020: PyLops—A linear-operator Python library for scalable algebra and optimization. *SoftwareX*, **11**, 100361, https://doi.org/10.1016/j.softx.2019.100361.

Ray, P. S., C. L. Ziegler, W. Bumgarner, and R. J. Serafin, 1980: Single- and multiple-Doppler radar observations of tornadic storms. *Mon. Wea. Rev.*, **108**, 1607–1625, https://doi.org/10.1175/1520-0493(1980)108<1607:SAMDRO>2.0.CO;2.

Rihan, F. A., C. G. Collier, and I. Roulstone, 2005: Four-dimensional variational data assimilation for Doppler radar wind data. *J. Comput. Appl. Math.*, **176**, 15–34, https://doi.org/10.1016/j.cam.2004.07.003.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Schenkman, A. D., M. Xue, and M. Hu, 2014: Tornadogenesis in a high-resolution simulation of the 8 May 2003 Oklahoma City supercell. *J. Atmos. Sci.*, **71**, 130–154, https://doi.org/10.1175/JAS-D-13-073.1.

Shapiro, A., C. K. Potvin, and J. Gao, 2009: Use of a vertical vorticity equation in variational dual-Doppler wind analysis. *J.*

Atmos. Oceanic Technol., **26**, 2089–2106, https://doi.org/10.1175/2009JTECHA1256.1.

——, K. M. Willingham, and C. K. Potvin, 2010: Spatially variable advection correction of radar data. Part I: Theoretical considerations. *J. Atmos. Sci.*, **67**, 3445–3456, https://doi.org/10.1175/2010JAS3465.1.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Soderholm, J. S., H. A. McGowan, H. Richter, K. Walsh, T. Wedd, and T. M. Weckwerth, 2017: Diurnal preconditioning of subtropical coastal convective storm environments. *Mon. Wea. Rev.*, **145**, 3839–3859, https://doi.org/10.1175/MWR-D-16-0330.1.

——, V. Louf, J. Brook, and A. Protat, 2022: Australian operational weather radar level 1b dataset. National Computing Infrastructure, accessed 8 August 2023, https://www.openradar.io/operational-network.

Sun, J., and N. Crook, 1998: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.*, **55**, 835–852, https://doi.org/10.1175/1520-0469(1998)055<0835:DAMRFD>2.0.CO;2.

Testud, J., and M. Chong, 1983: Three-dimensional wind field analysis from dual-Doppler radar data. Part I: Filtering, interpolating and differentiating the raw data. *J. Climate Appl. Meteor.*, **22**, 1204–1215, https://doi.org/10.1175/1520-0450(1983)022<1204:TDWFAF>2.0.CO;2.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

Trapp, R. J., and C. A. Doswell III, 2000: Radar data objective analysis. *J. Atmos. Oceanic Technol.*, **17**, 105–120, https://doi.org/10.1175/1520-0426(2000)017<0105:RDOA>2.0.CO;2.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161–193, https://doi.org/10.1007/s007030070003.

——, and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165, https://doi.org/10.1007/s007030170027.

——, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170.

——, M. Hu, and A. D. Schenkman, 2014: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data. *Wea. Forecasting*, **29**, 39–62, https://doi.org/10.1175/WAF-D-13-00029.1.

Zhang, J., K. Howard, and J. J. Gourley, 2005: Constructing three-dimensional multiple-radar reflectivity mosaics: Examples of convective storms and stratiform rain echoes. *J. Atmos. Oceanic Technol.*, **22**, 30–42, https://doi.org/10.1175/JTECH-1689.1.