A Variational Interpolation Method for Gridding Weather Radar Data

JORDAN P. BROOK, A ALAIN PROTAT, B JOSHUA S. SODERHOLM, B ROBERT A. WARREN, B AND HAMISH MCGOWAN A

A Atmospheric Observations Research Group, University of Queensland, Saint Lucia, Queensland, Australia
B Science and Innovation Group, Australian Bureau of Meteorology, Docklands, Victoria, Australia

(Manuscript received 18 February 2022, in final form 2 June 2022)

ABSTRACT: Observations made by weather radars play a central role in many aspects of meteorological research and forecasting. These applications commonly require that radar data be supplied on a Cartesian grid, necessitating a coordinate transformation and interpolation from the radar’s native spherical geometry using a process known as gridding. In this study, we introduce a variational gridding method and, through a series of theoretical and real data experiments, show that it outperforms existing methods in terms of data resolution, noise filtering, spatial continuity, and more. Known problems with existing gridding methods (Cressman weighted average and nearest neighbor/linear interpolation) are also underscored, suggesting the potential for substantial improvements in many applications involving gridded radar data, including operational forecasting, hydrological retrievals, and three-dimensional wind retrievals.

KEYWORDS: Atmosphere; Convective storms; Data processing/distribution; Radars/radar observations; Remote sensing; Variational analysis

1. Introduction

Gridding is a computational procedure involving the interpolation of data from their original coordinates to a regularly spaced set of points known as a “grid” or “mesh.” It is an important and often unavoidable requirement for many applications involving radar observations, including assimilation in numerical weather prediction models (e.g., Dowell et al. 2004; Xue et al. 2014), hydrological retrievals (e.g., Steiner et al. 1995), image processing algorithms (e.g., optical flow; Anagnostou and Krajewski 1999), and three-dimensional wind retrievals (e.g., Duhl et al. 2019). Gridded radar fields also enable the visualization of cross sections at various orientations (e.g., constant altitude slices) displayed as standard 2D images. These displays are better suited for visualization purposes, as no prior knowledge of the underlying radar scanning geometry is required.

Operational weather radars (i.e., stationary weather surveillance radars such as those in the U.S. WSR-88D network) typically operate in plan position indicator (PPI) mode, obtaining high-resolution measurements along conical PPI surfaces (250–1000 m in range and 0.5°–2° in azimuth). Data quality requirements and temporal constraints on the collection of radar volumes (~5 min operationally) naturally limit the number of elevation scans to around 10–15 per volume. As a result, gaps between successive elevation scans are typically much larger than between azimuths, especially at high elevations where elevation spacings commonly approach 5° (Langston et al. 2007). This coarse angular resolution means that the vertical spacing between data points increases considerably with range, leading to a highly nonuniform spatial data distribution. Furthermore, the quality of individual measurements is commonly affected by error sources such as nonuniform beam filling, velocity aliasing, sidelobe contamination, and spurious returns from nonmeteorological targets (Doviak and Zrnić 1993; Fabry 2015). Both the geometry and quality of operational observations make gridding radar data a challenging problem.

When consulting the radar gridding literature, a useful distinction can be drawn between “weighted average” and “nearest neighbor” approaches. The first category contains the iterative, inverse distance weighting approaches first proposed by Cressman (1959) and Barnes (1964). These methods were originally proposed for meso- and synoptic-scale analyses, where preservation of large-scale phenomena is the key objective. More recently, they have been adopted for use in formal weather radar literature (see Trapp and Doswell 2000; Askelson et al. 2000, and references therein), and are implemented in popular open-source radar software packages [e.g., Python ARM Radar Toolkit (PyART) and wradlib; Heistermann et al. 2015]. The second category contains comparatively simpler methods that either apply the nearest-neighbor interpolation method directly (e.g., Jorgensen et al. 1983), or use it as an intermediate step before subsequent processing such as linear interpolation (e.g., Mohr and Vaughan 1979). Nearest-neighbor methods were preferred in earlier studies because of their computational simplicity and have garnered renewed interest due to their implementation in producing multimodal products, such as the U.S. Multi-Radar/Multi-Sensor (MRMS) system (Lakshmanan et al. 2006).

Here, we pause to discuss the considerations involved in choosing a gridding method from either category, and argue that this decision is consequential for more than just computational efficiency. The justification for weighted-average approaches is best outlined using concepts from spectral analysis and classical sampling theory. Under this framework,

1 This process is commonly referred to as “objective analysis” in applications involving weather radars (e.g., Trapp and Doswell 2000); however, the more general term is used in this study.

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

DOI: 10.1175/JTECH-D-22-0015.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBLSReuseLicenses).
meteorological fields of interest are defined as a linear combination of many discrete waves, described individually by their wavelength \( \lambda \). When fields are measured at discrete locations, wavelengths smaller than the Nyquist wavelength (\( \lambda < 2\Delta \), where \( \Delta \) is the data spacing) are unresolvable, meaning their energy is aliased and mistaken for larger wavelengths (Shannon 1998). These theoretical guidelines are complicated by heterogeneous data spacings and varying smoothness requirements, meaning that the definition for a “resolved” wavelength typically varies between 2 and 10 \( \Delta \) across the literature (where \( \Delta \) is alternatively defined as the largest data point spacing in radar applications; Carbone et al. 1985; Trapp and Doswell 2000; Majcen et al. 2008). In general, the goal of these analyses is to eliminate wavelengths smaller than what is deemed unresolvable, while retaining those above. Observational noise is filtered from the result, producing a smooth analysis field that is well-suited for synoptic-scale analyses where the important features are well-described through large wavelengths alone (Given and Ray 1994).

Problems arise when weighted-average approaches are used to resolve convective-scale phenomena. Consider measurements from an operational weather radar with 1° azimuthal spacing and a maximum range of 100 km. Here, the maximum data spacing between azimuthally adjacent measurements is \( \sim 1.75 \text{ km} \), leading to a Nyquist wavelength of \( \sim 3.5 \text{ km} \). Carbone et al. (1985) noted that the horizontal extent of convective updrafts is typically on the order of 1–2 km, meaning the amplitude and extent of radar returns from these small-scale features are intentionally filtered from the analysis. This level of smoothing is undesirable for severe weather applications such as storm tracking and nowcasting that require accurate resolution of both the amplitude and extent of reflectivity signatures (Zhang et al. 2005). The manner in which these methods also stretch valid information past its true extent is particularly evident with altitude, where echo-top heights are artificially increased, and unfiltered ground clutter is propagated vertically to produce spurious features (Dowell et al. 2004; Zhang et al. 2005). Attempts to counter oversmoothing by decreasing the radius of influence below theoretical recommendations (e.g., Koch et al. 1983; Pauley and Wu 1990; Trapp and Doswell 2000) necessarily introduces other unwanted artifacts (see, e.g., appendix A in Dahl et al. 2019), meaning that these methods are fundamentally limited in their retention of small-scale features.  

Some combination of nearest neighbor and linear interpolation is commonly used when generating multidarad mosaic products (e.g., Zhang et al. 2005; Langston et al. 2007; Warren et al. 2020). Utilizing individual, spatially proximal data points in these methods prioritizes maximized detail, and results in analyses that closely mimic the underlying radar information in terms of amplitude and extent (Trapp and Doswell 2000). However, these lower-order methods also propagate observational noise and generate smaller-wavelength waves not present in the underlying data by creating first- and second-order discontinuities in the analysis fields (Trapp and Doswell 2000; Askelson et al. 2000). This reduction in spectral fidelity (i.e., spectral broadening) has adverse effects on applications such as dual-Doppler retrievals, which require smooth analysis fields because of their reliance on numerical derivatives (Testud and Chong 1983). The introduction of small-scale noise in mosaics is somewhat mitigated by spatiotemporally weighting additional measurements from multiple instruments (e.g., Langston et al. 2007). However, even for grid points with contributions from multiple data sources, these additional filtering steps will not completely eliminate the effects of spectral broadening. An analyst who requires smooth outputs may then consider the application of ad hoc filtering techniques (e.g., Dahl et al. 2019).

Unlike the aforementioned weighted-average or nearest-neighbor approaches that prioritize either spectral fidelity or fine-scale resolution, variational interpolation, the method used in this study, can potentially offer a compelling trade-off between the two. This method was introduced in a broad meteorological context by Sasaki (1970) and Wahba and Wendelberger (1980), and applied specifically to radar gridding by Testud and Chong (1983). The technique individually penalizes observational errors and analysis smoothness, allowing the analyst to weight either constraint in a manner that reflects their desired outcome. Despite this promise, the original Testud and Chong (1983) implementation is not directly suited for use in contemporary radar gridding applications. First, it was intended for use along two-dimensional, constant elevation surfaces (otherwise known as the COPLANE coordinate system; Miller and Strauch 1974). These 2D surfaces yielded problems with \( \sim 900 \) grid points, whereas the three-dimensional \((40 \times 600 \times 600)\), Cartesian grids used in this study\(^4\) contain more than 14 million grid points. The relatively small size of their 2D implementation permitted the use of a finite element approach, requiring the explicit specification and iterative solution of large banded matrices. This type of approach is computationally impractical for the problem sizes considered here. In this study, we build upon the original approach outlined by Testud and Chong (1983) and detail how advances in large-scale optimization in fields such as medical imaging and geophysics now permits the solution of three-dimensional variational interpolation for radar gridding.

In light of the preceding discussion on the strengths and limitations of various gridding methods, we describe five

\(^2\) The angular spacing between PPI elevations typically exceeds the azimuthal spacing for operational weather radars, meaning that the maximum data spacing considered here is actually an underestimate in practice.

\(^3\) Note that an iterative implementation of weighted-average methods (as they were originally intended; refer to section 2b) results in steeper response functions, partially mitigating the damping of amplitudes at resolved wavelengths. However, these implementations still filter those wavelengths below the Nyquist.

\(^4\) Here, we adhere to standard gridding nomenclature by referring to a grid defined at regular intervals with respect to the \( x, y \), and \( z \) axes as “Cartesian.” An azimuthal equidistant projection with its origin at the radar may then be used to project the Cartesian grid to geographic coordinates.
desired criteria for an ideal radar gridding procedure. These criteria are used to assess the performance of the various gridding methods tested in this study and are outlined as follows:

1) **Resolution**: Retrieved fields should closely resemble the underlying radar data in terms of both amplitude and spatial extent.

2) **Smoothness**: The methodology must filter observational noise, and ideally allow the user to control the degree of smoothness for their desired application.

3) **Interpretability**: The method should not introduce biases in analysis shape or amplitude, so that features in gridded fields, or spatiotemporal changes therein, may be correctly attributed to the underlying radar information.

4) **Continuity**: Retrieved gridded fields should be visually continuous, thereby masking data acquisition gaps between missing sweeps and rays.5

5) **Efficiency**: The retrieval method must be computationally efficient enough to permit real-time applications.

The remainder of this study is structured as follows. First, the variational interpolation gridding procedure is outlined in section 2a, before the methodology of two existing methods (Cressman weighted average and nearest neighbor/linear interpolation) is introduced for comparison purposes in sections 2b and 2c. Two observing system simulation experiments (OSSEs) are used to test each of these gridding methods in sections 3 and 4, before experiments using real data are analyzed in section 5. The findings from these experiments are summarized in section 6, along with suggestions for future research in this area.

2. **Methodology**

a. **Variational interpolation**

The interpolation technique introduced here is broadly categorized as a variational inverse method (e.g., Sasaki 1970; Walba and Wendelberger 1980), which aims to retrieve a continuous, 3D Cartesian grid of selected radar fields in model space: \( \varphi \in \mathbb{R}^M \) where \( M = n_x \times n_y \times n_z \) where \( n_i \) is the number of grid points in the \( i \)th dimension. The retrieval is informed by comparison with a discrete set of \( N \) radar observation points in data space \( \{d \in \mathbb{R}^N\} \). We apply the methodology to stationary, volumetric radar data in this study; however, in principle it may also be applied to alternate instrumentation such as airborne radar observations. Gridded fields must be close to the data in a least squares sense (without necessarily passing through the observations) to ensure retrievals are robust to noise and spurious radar returns. As in all variational problems, solutions in our implementation are achieved through minimizing a cost function. The formulation used in this study is as follows:

\[
\min_{\varphi} \{||d - \mathcal{R}\varphi||_2^2 + J_S(\varphi_x) + J_B(\varphi_a) + J_D(\varphi_r)\}, \tag{1}
\]

where \( \mathcal{R} \) is a trilinear interpolation operator used to facilitate a direct comparison between field values in model space and discrete observations in data space. Refer to appendix A for supplementary information on the theoretical aspects of \( \mathcal{R} \). The common \( L_2 \) norm (i.e., Euclidean norm) notation \( ||x||_2^2 = \sum_{i=1}^k x_i^2 \) is used in Eq. (1) and subsequently throughout the text. Minimizing Eq. (1) using the data fidelity term [first term in Eq. (1)] alone is equivalent to finding the least squares solution (Barth et al. 2021). However, such a problem is underdetermined in practice due to the incomplete data coverage of operational weather radars. The remaining regularization terms in Eq. (1) assist in resolving the system through the addition of a smoothness constraint \( J_S \), a background constraint \( J_B \), and a denoising constraint \( J_D \). Field values passed to the regularization constraints \( (\varphi_a = \varphi - \varphi_b) \) are normalized by subtracting a background estimate \( \varphi_b \) so that unconstrained regions naturally default to this value. For experiments presented here (involving the interpolation of radar reflectivity in dBZ units), it is suitable to set \( \varphi_b \) equal to zero, such that \( \varphi_a \) or \( \varphi \) can be used interchangeably.

A number of smoothing constraints were considered for implementation in this study, and interested readers are referred to appendix B for a justification of those used here. Unmixed, second-order smoothing terms (e.g., Potvin et al. 2012a) were found to efficiently spread radar information throughout the grid, while adequately filtering observational noise. The former property is of greater importance in this application, as noise filtering is predominantly handled by the denoising constraint \( J_D \). Furthermore, the relative computational simplicity of this formulation, in comparison with more complex smoothing norms (e.g., Barth et al. 2014), makes it an ideal candidate for radar interpolation on grids as large as those used in this study. The constraint is formalized as

\[
J_S(\varphi) = \lambda_v ||\varphi_{xx}||_2^2 + \lambda_H(\|W_x \varphi_y\|_2^2 + \|W_y \varphi_x\|_2^2), \tag{2}
\]

where \( \varphi_{ij} = \partial^2 \varphi / \partial i \partial j \) (computed using centered finite differences and Neumann boundary conditions) and \( \lambda_H \) and \( \lambda_V \) are tunable smoothing weights in the horizontal and vertical directions, respectively. Azimuthally varying weights \( W_x \) and \( W_y \) are strategically applied to horizontal smoothing terms to account for discrepancies in the maximum data spacing between range and azimuth (\( D_R \) and \( D_A \)). The maximum azimuthal spacing is typically much larger than the maximum range spacing, or equivalently \( f = D_R/D_A \ll 1 \). Horizontal smoothing constraints account for this discrepancy by applying the full horizontal smoothing weight \( (\lambda_H) \) for an axis when it is parallel with radar azimuth and a spacing-weighted fraction when it is perpendicular (\( f \lambda_H \)). This formulation ensures that the total horizontal smoothing weight is constant throughout

---

5 The term “performance” is defined relative to the five desirable criteria above and should not be confused with purely “computational performance.”

6 While continuity is an important criterion for gridding in general, some applications may require the preservation of data gaps. Postprocessing may be applied in such cases to either “flag” or “mask” grid points that are not sufficiently proximal to observation points.
the grid while spreading radar information along azimuths more readily than between radii. Weighting functions are defined as

$$W_y = C + A \cos(2\phi) \quad \text{and} \quad W_x = C - A \cos(2\phi),$$  \hspace{1cm} (3)

where $\phi$ represents azimuth (measured clockwise from north by convention), $A = |f - 1/2|$, and $C = (f + 1/2)$. The suitable provision of tunable constraint weights ($\lambda_H$ and $\lambda_V$) is a common problem in variational analysis that must be done with the analyst’s application in mind. Testud and Chong (1983) noted that higher smoothing weights act to increase the cutoff frequency of the low-pass filtering, in effect shrinking the pass band and producing smoother analysis fields. Higher smoothing weights should be used in applications where smoothness is prioritized, such as the preparation of gridded Doppler velocity fields for 3D wind retrievals. Conversely, relatively lower weights should be used in more common radar gridding applications such as the production of gridded reflectivity fields for weather monitoring, where the priority is to maintain fidelity to the original radar data in terms of amplitude and extent.

The background constraint penalizes deviations away from a background field value $\varphi_0$, ensuring retrievals are bounded in unconstrained regions. Unlike previous implementations (e.g., Gao et al. 1999), the constraint is not applied uniformly across the domain. Instead, penalties are weighted by their spatial proximity to the nearest valid observation. This implementation constrains solutions within data voids, while having a negligible effect within regions containing valid data. The weighting scheme used here is formalized below:

$$J_B(\varphi) = \lambda_B (\|\varphi - \varphi_0\|_1 + \|\varphi - \varphi_0\|_1 + \|\varphi - \varphi_0\|_1),$$  \hspace{1cm} (5)

where the $\ell_1$ norm notation implies $||x||_1 = \sum_{i=1}^{N} |x_i|$ and $\lambda_B = 0.2$ is a weighting parameter set invariably for all experiments through means of trial and error. Convention dictates that the first-order derivatives in Eq. (5) are taken using forward differences, with Neumann boundary conditions. The use of $\ell_1$ norms encourages sparsity in model derivatives, which reduces speckle noise without oversmoothing or causing extraneous inflections in regions with large gradients (such as reflectivity at the leading edge of a frontal system or Doppler velocity at the edge of a strong updraft; Rudin et al. 1992). The capability to reproduce these “sharp” edges with the $J_B(\varphi)$ constraint is demonstrated in appendix B.

The inclusion of $\ell_1$ norms in Eq. (1) (which are nonsmooth and therefore nondifferentiable), dictate that typical optimization methods based on gradient descent (such as the conjugate gradient method) cannot be used. Instead, we use an open-sourced implementation of the split-Bregman optimization method for mixed $\ell_1$ and $\ell_2$ norms (Ravasi and Vasconcelos 2020). This method proceeds by decoupling the $\ell_1$ and $\ell_2$ elements, and solving the resulting system by a series of nested optimizations (a maximum of 10 outer iterations and 5 inner iterations are used throughout this study; Goldstein and Osher 2009). All regularization constraints are implemented as hardware-agnostic linear operators, permitting matrix-free, real-time optimizations of large-scale 3D problems using GPU architecture (Ravasi and Vasconcelos 2020).

b. Comparison 1: Cressman interpolation

Comparison with a weighted-average gridding method requires one to choose between the popular Cressman (1959) and Barnes (1964) schemes. Our decision is guided by Trapp and Doswell (2000), who noted that the Cressman (1959) scheme is used more frequently in weather radar applications judging by citations in the literature. This may be due to the relative simplicity of the weighting function, which requires the provision of a single parameter [the radius of influence (ROI), or $R$] whereas Barnes (1964) requires an additional smoothing parameter. The Cressman scheme is selected on this basis. In general, inverse distance weighting schemes perform a weighted sum of all radar datum within a radius of influence from each analysis point. Cressman weights are calculated as follows:

$$W(r) = \exp\left(\frac{R^2 - r^2}{R^2 + r^2}\right)$$  \hspace{1cm} (6)

---

7 The anisotropic smoothing weights implemented in this study differ from those studied previously (e.g., Askelson et al. 2000; Trapp and Doswell 2000) because they do not impose a range dependence in the filtering properties, thereby avoiding the associated interpretability issues.

8 The ability to analytically derive response functions for the Barnes (1964) scheme has led to a comparatively richer catalog of theoretical studies (e.g., Barnes 1973; Koch et al. 1983; Achtemeier 1986; Pauley and Wu 1990) and more frequent use in the wider meteorological community. For the sake of completeness, we repeated the Cressman experiments in this study with the Barnes scheme and found that the results were notably insensitive to the choice of weighting scheme.
where \( r \) is the distance between an observation and the cor-
responding grid point. The original Cressman (1959) for-
mulation is an iterative method, which proceeds by making
successive corrections to an initial guess field. Radar applica-
tions historically employ a single pass (one iteration) with a
zero guess field (Given and Ray 1994; Trapp and Doswell 
2000), and implementations in open-source radar toolkits ex-
clusively permit these settings (Heistermann et al. 2015). One
such implementation is used in this study, namely, PyART
(Helmus and Collis 2016).

The principal consideration in a Cressman analysis is the ra-
dius of influence used in Eq. (6). In applications involving ran-
domly spaced data samples, \( R \) is usually set to some multiple
of the “average” data spacing (Stephens and Stitt 1970; Koch
et al. 1983). This treatment is ill-suited to radar applica-
tions, as data spacings systematically increase with range.
Instead, radii of influence are usually set in relation to the maxi-
mum data spacing \( d_{\text{max}} \) (the largest data spacing anywhere on
the grid). Using this methodology ensures the entire grid ex-
periences a uniform degree of smoothing, albeit at the expense of
adequately resolved wavelengths at smaller ranges. This treat-
ment also guarantees the interpretability of the results (see
section 1), as the minimum resolvable wavelength is constant
throughout the grid, prohibiting the introduction of any artifi-
cial bias with range or direction (Trapp and Doswell
2000; Potvin et al. 2012b). Due to these considerations, anisotropic
weighting schemes (e.g., Askelson et al. 2000) are not consid-
ered here.

In their two-dimensional experiments, Trapp and Doswell
(2000) justified the use of \( R = 1.83 d_{\text{max}} \) by numerically deriv-
ing the response function of the Cressman scheme, and ad-
justing \( R \) to ensure the near-total damping of “unresolved
wavelengths” \( \lambda < 2d_{\text{max}} \). When applied to operational radar
data, large elevation gaps between adjacent PPI scans natu-
rally lead to prohibitively large influence radii at significant
ranges (>100 km), which both filters desirable wavelengths
and extends valid data signatures beyond their true extent.
Dahl et al. (2019) explored decreasing \( R \) in an attempt to
mitigate oversmoothing and resolve small-scale convective in-
formation. They found that using a very small value of \( R \)
(slightly larger than 0.5\( d_{\text{max}} \)) the smallest value that en-
sures each grid point is calculated using at least one observa-
tion), introduces artifacts in analysis fields that necessitate ad hoc
postprocessing steps [these artifacts may be due to the discre-
tization effects of small influence radii discussed in Pauley
and Wu (1990)]. Setting \( R \) equal to the maximum data spacing
\( R = d_{\text{max}} \) did not exhibit these artifacts but did not resolve
field amplitudes to the same extent.

The analyst is then faced with a common decision in
weighted-average approaches: accept the oversmoothing limi-
tations imposed by large \( R \) values or reduce \( R \) beyond theo-
retical recommendations and potentially introduce spurious
artifacts in gridded fields. In practice, there is often no satis-
factory trade-off between these decisions. In light of these
considerations, we use \( R = d_{\text{max}} \) in these experiments, which
limits oversmoothing, permits some small-scale wavelengths
possibly due to noise, and does not suffer from the artifacts
described by Dahl et al. (2019). We do not consider a range of
\( R \) values to enhance the clarity of results, and we posit that
these settings will provide a sufficiently challenging compari-
sion with the variational method for small-scale convection
purposes.

c. Comparison 2: Linear vertical interpolation

The second comparison method tested here uses nearest
neighbors in the radial and azimuth direction, and linear
interpolation with elevation [named the “vertical inter-
polation” method by Zhang et al. (2005)]. To maintain the
nomenclature introduced in section 1, we refer to this
method as the nearest-neighbor method throughout the
text, despite the linear interpolation between elevations.
The algorithm proceeds by identifying the nearest PPI mea-
surement (in range and azimuth) on the adjacent tilt below
and above each Cartesian grid point \( (d_1 \) and \( d_2 \), re-
spectively). A value is then calculated for the corresponding grid
point \( (\varphi_i) \) by linearly interpolating these two measurements
according to

\[
\varphi_i = \frac{(w_1d_i + w_2d_2)(w_1 + w_2)}{w_1 + w_2},
\]

where \( w_1 \) and \( w_2 \) are interpolation weights specified as

\[
w_1 = (\theta_2 - \theta)/(\theta_2 - \theta_1)
\]

and

\[
w_2 = (\theta_1 - \theta)/(\theta_2 - \theta_1),
\]

and \( \theta \) values indicate elevation angles associated with \( \varphi_i \), \( d_1 \),
and \( d_2 \). The procedure described above can be implemented
effectively very efficiently by vectorizing the calculations in Eq. (7). This
is achieved by calculating \( w_1 \) and \( w_2 \) on the analysis grid, re-
quiring a coordinate transform into the radar’s native spheric-
ical coordinates \( (r, \phi, \theta) \), where \( r \) is the slant range along the
radar beam.

The spherical-to-Cartesian coordinate transformation is
complicated by the curvature of Earth and refraction of the
radar beam as it propagates through the atmosphere. Here,
we account for these effects by assuming the 4/3 effective
Earth radius model (Doviak and Zrnić 1993) and refer the
reader to a diagram outlining the relevant geometry in appen-
dix A of Warren et al. (2018). To the authors’ knowledge, the
transform from Cartesian to radar coordinates in this model is
not outlined in the radar gridding literature and will be pre-
presented here to aid computational reproducibility. First, the
distance along Earth’s surface \( s \) and the radar azimuth \( \phi \) can be
calculated directly from the horizontal grid coordinates:

\[
s = (x^2 + y^2)^{1/2} \quad \text{and} \quad \phi = \frac{\pi}{2} - \arctan2(y, x),
\]

where \arctan2 is a common computational function used to
specify the arc tangent angle with regard to the positive x axis.
The slant range and elevation of the radar beam are calcu-
lated by taking into account the height of the radar antenna
above sea level \( h \) and the effective radius of Earth \( a_e \) (4/3 times
the actual radius at the radar’s longitude):
The grid transformation outlined above greatly simplifies the provision of $d$ and $w$ arrays in Eq. (7), rendering this method suitably efficient for large multiradar mosaic grids (e.g., Zhang et al. 2005; Langston et al. 2007; Warren et al. 2020). Readers should consult these texts for information on how to project and assimilate radar information from multiple sources onto a common grid.

3. Continuous field experiment

a. Experimental setup

This experiment aims to characterize the performance of each gridding method with a simple harmonic control field. Radar observations are defined analytically at the center of each sample volume (i.e., voxel) as in previous observing system simulation experiment studies (e.g., Trapp and Doswell 2000; Shapiro et al. 2009). This approach eliminates any radar beam sampling artifacts, isolating the response due to interpolation alone (the inclusion of a more realistic radar simulator is discussed in section 4). The analytic function for the control field is outlined below:

$$f(x, y, z) = A \sin \left( \frac{\pi n_x x}{\lambda_x} \right) \sin \left( \frac{\pi n_y y}{\lambda_y} \right) \sin \left( \frac{\pi n_z z}{\lambda_z} \right),$$  \hspace{1cm} (13)

where $A = 10$ is the field amplitude, the fundamental wavelength $\lambda$ is set to the extent of the analysis domain in each dimension, and the constant $n$ controls the number of amplitude peaks in each dimension (subsequently referred to as meteorological “features”). Fields derived from Eq. (13) resemble a “checkerboard” pattern along horizontal slices (e.g., Fig. 1c). A range of horizontal wavelengths are tested to determine the effects of gridding meteorological features at varying spatial scales. These scales are discussed in terms of the number of features in each dimension ($n_x$ and $n_y$) throughout this section for the purposes of clarity and ease of visualization.

The Cartesian analysis grid is defined with the radar positioned at the origin, which is situated $\sim 28$ km away from the closest point on the analysis grid $[(x, y, z) = (20, 20, 0) \text{ km}]$. The grid extends 40 km in both horizontal dimensions, and 15 km vertically, with an isotropic grid spacing of 500 m. Range and azimuth spacing for the simulated radar observations is set to 250 m and $1^\circ$, respectively, and measurement error is simulated by adding random noise with a standard deviation equal to 1.0 (10% of the field amplitude). A series of constant elevation sweeps are taken at $1.5^\circ$ intervals between $0^\circ$ and $30^\circ$ inclusive to ensure radar coverage throughout the depth of the domain, and the 4/3 Earth radius assumption is once again employed to account for beam refraction and Earth’s curvature (Doviak and Zrnić 1993). Such a scanning pattern is unsuitable for operational purposes due to the effects of ground clutter in the lowest scan ($0^\circ$) and temporal constraints on the maximum number of sweeps within an operational volume. However, both the domain extent and the optimistic scanning pattern have been chosen to represent the “best case scenario”
in terms of observational coverage in this experiment. As outlined in section 2b, the Cressman radius of influence is set to the maximum data spacing \( (R \approx 2275 \text{ m}) \). The provision of variational weights \( k_V = 1.1; k_H = 0.4 \) was done similarly to previous variational interpolation studies (e.g., Barth et al. 2021), using a “black box” parameter optimization with respect to root-mean-squared error (Knysh and Korkolis 2016). These experiments are designed to provide an insight into the optimal expected performance for each method in practice.

b. Gridding performance

The first result shown in Fig. 1 exhibits gridded retrievals with \( n_x = n_y = 9 \). Under these conditions, each feature on the "checkerboard" grid has a diameter of \( \sim 4.5 \text{ km} \), which should ideally be well-resolved for radar applications involving convection. At this scale, the Cressman and nearest-neighbor methods perform similarly in terms of error measures, but the resulting fields highlight the relative strengths and weaknesses of each method particularly well. The variational output is visually very similar to the analytical field, which has been omitted for brevity. The visual likeness of the remaining Cressman and nearest-neighbor methods to the “truth” should be judged against the variational outputs in Figs. 1c and 1f.

The Cressman gridding method produces a visually pleasing field in terms of smoothness and symmetry in Figs. 1a and 1d. The regularity of the error field in Figs. 2a and 2d also suggests a near total extinction of short wavelengths introduced by measurement errors. An unavoidable compromise for this level of spectral fidelity is a partial damping of resolvable wavelengths, which results in a systematic underestimation of the extrema amplitudes. This effect is easily observed in the error fields as a smooth, persistent offset from the true field in Figs. 2a and 2d, resulting in relative errors of \( \sim 30\% \) at the center of each feature. These findings resemble those of previous radar gridding studies (e.g., Askelson et al. 2000; Trapp and Doswell 2000), and follow from the theory outlined in section 1. The Cressman error fields in Figs. 2a and 2d also show an important exception to these findings: a consistent overestimation of the true field along all grid edges. This effect is particularly evident in the vertical dimension (Fig. 2d), and is a consequence of data scarcity within the search radius at the edge of the domain. Boundary errors of this nature are mitigated in the interpolation of synoptic-scale meteorological applications by ensuring the grid domain is suitably contained within the extent of the observations (Koch et al. 1983; Pauley and Wu 1990). Such recommendations are not practical for operational weather radars, as we seek to grid observed fields onto a constant analysis domain regardless of the spatial distribution of hydrometeors. In section 5 we show that this effect has serious implications in real applications.

In contrast to the Cressman method, extrema are retrieved well by the nearest-neighbor method (mostly to within the simulated level of noise, refer to Fig. 2), and exhibit no systematic underestimation. Instead, absolute errors in Figs. 2b and 2e show a level of speckle noise similar to the simulated observation errors. Clearly, the majority of measurement noise has been retained in the results, confirming the poor smoothing and filtering characteristics of this method. Further inspection of Figs. 1 and 2 reveals azimuthal “striping” in regions most distant from the radar. These discontinuities arise due to the inherent limitations of nearest-neighbor interpolation methods.

Selecting a grid domain to avoid boundary errors would mean only analyzing grid points within the center of precipitating regions, which is undesirable for weather surveillance purposes and historical archiving.
in nearest-neighbor approaches in poorly sampled regions (Askelson et al. 2000), and manifest as a range-dependent bias in radar applications due to the reduction in data resolution with range. A lack of similar artifacts with elevation (despite a greater angular spacing of 1.5°, as compared with 1.0° in azimuth) indicates that first-order linear interpolation may be more suitable in the azimuthal dimension for radar scanning geometries (e.g., Warren et al. 2020). While linear interpolation between elevations does aid the interpretability of results by mitigating range-dependent biases between elevations, it does come at a cost to the continuity of the grid by introducing data gaps outside the observation bounds of each column. The result is a familiar occurrence in gridded radar data gaps or data acquisition artifacts) in this case.\footnote{Despite a greater angular spacing of 1.5°, as compared with 1.0° in azimuth.}

visually, the variational retrievals in Figs. 1c and 1f appear smooth and free of measurement errors, while accurately resolving the amplitude of extrema. This visual assessment is supported by a small root-mean-squared error (RMSE) of 0.32—a decrease of 71% and 73% relative to the Cressman and nearest-neighbor methods, respectively. The variational absolute error field in Figs. 2c and 2f also shows a significant reduction in error magnitudes throughout the grid. The variational errors do not exhibit the same harmonic structure as the underlying field, meaning no systematic underestimation has occurred, and resolvable frequencies have not been reliably attenuated. Furthermore, a lack of speckle noise indicates that high-frequency observational noise has been successfully filtered from the analysis. Instead, error fields take on an irregular, “patchy” appearance, implying some combination of aliasing from higher-frequency errors and retention of lower-frequency noise. Overall, small error magnitudes and the lack of speckle noise indicate that random Gaussian observation errors have been damped by the variational smoothing and denoising constraints. The retention of resolvable wavelengths has not come at the cost of interpretability (no clear spatial biases with range, azimuth, or elevation) or continuity (no data gaps or data acquisition artifacts) in this case.

We now turn our attention to determining whether the promising performance of the variational gridding method generalizes to a wider range of input fields. These experiments also aim to verify that variational smoothing weights ($\lambda_V = 1.1$ and $\lambda_H = 0.4$ in this experiment) do not need to be optimized to suit the structure of the underlying precipitation, ensuring one set of parameters may be used across a wide range of events (e.g., both convective and stratiform). Gridded retrievals with the same experimental setup as Figs. 1 and 2 were repeated for a range of horizontal frequencies by varying $n_x$ and $n_y$ between 1 and 25, corresponding to features with widths between 40 and 1.6 km, respectively. Varying numbers of features in each dimension result in various degrees of eccentricity, and the RMSE results for all experiments are shown in Fig. 3.

Gridding error measures are similar for all three methods for very low numbers of features ($n_x = n_y < 5$), meaning that the choice between gridding methods is less consequential for lower-frequency fields such as stratiform precipitation. In general, gridding accuracy drops off with increasing numbers of features for all gridding methods, highlighting the difficulty in resolving high-frequency, convective-scale phenomena. The largest, and most rapid drop off in accuracy is experienced by the Cressman method, resulting in RMSE values roughly double that of the other two methods for the highest numbers of features ($n_x = n_y < 25$). Accuracy reductions are less pronounced for the nearest-neighbor method due to its ability to retain small-scale information; however, the variational method clearly outperforms both the Cressman and nearest-neighbor methods for all experiments (Fig. 3). An interesting feature of Fig. 3 is that the variational retrievals are the only method affected by the degree of eccentricity of the input fields (as indicated by the comparatively noncircular appearance of constant-RMSE contours). Equivalently, the accuracy of the variational method is more strongly modulated by the most poorly resolved dimension of the input fields, perhaps indicating the variational method is approaching a practical accuracy “limit” given the input field and the sampling characteristics of this experiment.

![Fig. 3. RMSE errors for 625 experiments over a range of horizontal feature numbers ($n_x$ and $n_y$) for the continuous field experiment.](https://example.com/fig3)
d. Spectral analysis

Section 1 and references therein argued that depending on one’s application, error measures such as RMSE may not be the principal consideration when evaluating a radar gridding method. Namely, for applications such as 3D wind retrievals, the spectral fidelity of the resulting gridded fields may be of primary concern. Here, we aim to test the smoothing and noise-filtering properties of each gridding method by analyzing experimental spectral responses for various input fields with known spectral properties. One-dimensional spectral responses are preferred here for ease of presentation and were calculated following the methodology outlined in Trapp and Doswell (2000), except for the application of a 3D Hanning window to mitigate edge effects in the discrete Fourier transformation (Press et al. 2007). Spectral responses are provided in relation to their three-dimensional wavenumber $k = (k_x^2 + k_y^2 + k_z^2)^{1/2}$, where wavenumbers in each dimension are related to parameters already introduced in the text by $k_i = (m_i)\lambda_i$.

Spectra are presented for three symmetric input fields in Fig. 4, with 8, 18, and 26 features in each horizontal dimension. With these parameters, spectra should contain isolated peaks at $K \approx 0.9, 2.0, \text{and} 2.9$, respectively. The response of the analytic field shows significant spectral broadening due to discretization effects. This has been provided for comparison and demonstrates the “ideal” response for each experiment. Figure 4a represents an easier test of the gridding methods, with features 5 km in diameter. Noise reduction in the Cressman method is extremely effective at this scale, attenuating high wavenumber amplitudes by nearly five orders of magnitude. The Cressman responses in all three experiments exhibit a notable oscillating behavior. The authors offer no convincing explanation for this signature but suspect that it may have its origins in the aforementioned boundary effects at the edge of the domain. As expected from the visual inspection of Figs. 1 and 2, the spectral fidelity of the variational and nearest-neighbor methods is affected by the retention of high wavenumber signal from the observational noise. However, the level of noise retention is notably lower in the former. For higher-frequency input fields in Figs. 4b and 4c, some input signal is aliased to lower wavenumbers to the left of the peak wavelength. This is particularly severe for the Cressman method in Fig. 4c, where the peak is incorrectly attributed to a lower wavenumber. The variational method is least prone to aliasing in each of the experiments, meaning the amplitude and extent of the original radar information is most likely to be preserved using this method. The variational method is also the most performant in terms of noise filtering in the higher wavenumber experiments, and more generally exhibits the most consistent spectral performance across all three experiments.

4. Discrete storm experiment

a. Experimental setup

Having investigated the theoretical performance of each gridding method in an optimized set of OSSEs in section 3, we now aim to verify these findings using a more realistic set of experiments. Section 3 experimentally confirmed the difficulties associated with resolving small-scale features, and we further that inquiry in this section by extending a small 2D convective cell simulation first proposed by Trapp and Doswell (2000) into
three dimensions. The analytic field \( Z \) (henceforward storm, with field values analogous to radar reflectivity in a small convective storm cell) is temporally invariant but undergoes constant horizontal advection along the \( x \) axis away from the radar position at \( (x, y, z) = (-10, 15, 0) \) km. The addition of a temporal component in these experiments permits an investigation into the range dependence of each gridding method as the same storm is measured at varying distances out to 100 km. The analytical formulation of the simulated storm is

\[
Z(r') = \begin{cases} 
A \cos(2\pi r'), & r' < 1/4 \\
0, & r' \geq 1/4 
\end{cases}, 
\]

where \( A = 50 \) is the field amplitude and \( r' \) is a function controlling the extent of the storm, expressed as

\[
r'(t) = \left( \frac{m_1^2}{\lambda_x} \right)^2 \left| x - (\hat{x} + C_x t) \right|^2 + \left( \frac{m_3^2}{\lambda_z} \right)^2 \left| z - (\hat{z} + C_z t) \right|^2 
+ \left( \frac{m_2^2}{\lambda_z} \right)^2 \left| y - (\hat{y} + C_y t) \right|^2 \right)^{1/2},
\]

where \((C_x, C_y, C_z) = (20, 0, 0)\) is the constant advection velocity in kilometers per hour, \( t \) is time in hours, and \((\hat{x}, \hat{y}, \hat{z}) = (0, 15, 10/3)\) is the initial position of the storm cell in kilometers. Fundamental wavelengths are set to \( \lambda_x = \lambda_y = 30 \) km and \( \lambda_z = 10 \) km, along with shape parameters \((m_x, m_y, m_z) = (1, 1/2, 5/2)\), which determine the extent of the storm in each dimension. The result is an elliptical storm in the horizontal plane with a width of 6 km parallel to the direction of motion along the \( x \) axis, a length of 15 km on the \( y \) axis, and a depth of 5 km vertically (refer to Figs. 5e and 6e to visualize the analytic storm). The realistic amplitude and extent of the simulated reflectivity, along with a more sophisticated radar simulation methodology, ensure this OSSE closely resembles the common, yet challenging task of resolving a small convective cell in gridded radar data.

b. Radar simulation

Rather than simulating radar measurements analytically at the center of each radar voxel as in section 3, a method involving a weighted average of “scatterers” within each sample volume is employed in this section. The weighting methodology follows that of previous radars OSSEs (e.g., Potvin et al. 2009; Dahl et al. 2019), which define weighting functions according to how operational weather radars preferentially weight scatterers according to their proximity to the center of the sample volume. First, a collection of scatterers is defined on a finely spaced (20 m) isotropic grid, and field values are provided according to Eq. (14). A weighted average of all scatterers within each sample volume is subsequently calculated according to

\[
Z(r, \phi, \theta) = \frac{\sum_{i=1}^{n} R_i B_i Z_i}{\sum_{i=1}^{n} R_i B_i},
\]

where \( R_i \) and \( B_i \) are weights penalizing the radial and radar beam proximity for \( i \) of \( n \) scatterers within each sample volume. Radial weights are derived from a simple trapezoidal function defined as

\[
R_i = \begin{cases} 
1, & |\Delta r| < 0.3d \\
\max(0.5d - |\Delta r|, 0), & |\Delta r| \geq 0.3d 
\end{cases},
\]

where \( \Delta r \) is the scatterer’s radial offset from the center of the sample volume and \( d = 250 \) m is the range resolution of the radar. The beam weighting mimics the power distribution of an operational WSR-88D and is defined identically to that of Dahl et al. (2019) as

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

where \( \Delta \phi_i \) and \( \Delta \theta_i \) are the azimuthal and elevation offsets from the center of the sample volume and \( \sigma_B = \theta_B = 1^\circ \) denotes the half-power beamwidth of the radar in the azimuthal and elevation directions, respectively. Measurement error is once again simulated through the addition of Gaussian noise with a standard deviation equal to 1.0, with an additional removal (or masking) of 5% of all valid measurements. The minimum detectable signal is set to 5 dBZ for these experiments, with radar returns lower than this threshold masked from the analysis. The azimuthal spacing of radar measurements is set equal to the half power beamwidth of 1°, and the volume scanning strategy is set identically to the 10 lowest sweeps of the real data in section 5, at elevations \( \theta = (0.5, 0.9, 1.3, 1.8, 2.4, 3.1, 4.2, 5.6, 7.4, 10)\°. Volume scanning is performed instantaneously for all scans to ensure that no advection correction procedures are required. The maximum data spacing with this scanning strategy occurs between the sixth and seventh tiles (1.1° spacing) at the maximum radar range of 120 km, resulting in a Cressman radius of influence ~2500 m and experimentally determined variational smoothing weights \( \lambda_r = 0.01, \lambda_H = 0.35. \) Refer to Figs. 5a and 6a for horizontal and vertical cross sections of simulated radar observations.

c. Gridding performance

The Cressman gridded fields in Figs. 5b and 6b clearly suffer from the resolution limitations discussed in section 3. Maximum reflectivities are underestimated by roughly 30% for all time steps, and valid radar echoes are extended spatially by a distance equal to the radius of influence in all directions (particularly refer to Fig. 6b). Furthermore, the amplification of edge values due to boundary effects is evident at each time step, resulting in a local reflectivity maximum at the top of each cell. Cressman outputs are generally visually smooth despite the noise in the underlying radar data, with the exception of \( t = 5 \) in Fig. 5b. This grid exhibits some artificial striping at its center, indicating the radius of influence is too small to filter all high-frequency radar noise (Dahl et al. 2019). However, increasing the radius of influence naturally amplifies the aforementioned issues with amplitude
underestimates and boundary effects, leading one to conclude that single-pass weighted-average approaches may not be suited to this type of problem, no matter the choice of ROI.

Nearest-neighbor methods are better suited to observations containing small convective cells because of their enhanced resolution (Zhang et al. 2005) and do more accurately resolve the amplitude and extent of the cell as a result. However, Figs. 5c and 6c show once again that the added resolution comes at the cost of smoothness, with a clear visual retention of speckle noise and the propagation of missing radar data. Noise retention is more severe at later times due to the range dependence in data density, introducing a spatial bias in the resulting accuracy similar to Fig. 2b. The vertical extent of nearest-neighbor gridding is clearly limited by the sampling strategy of the radar, introducing data gaps below, and above the cell, and producing spurious echo tops in Fig. 6c. Overall, these experiments support the literary consensus that nearest-neighbor methods do more accurately resolve small-scale convective features, albeit at the expense of smoothness, interpretability, and continuity (Trapp and Doswell 2000; Zhang et al. 2005).

**Fig. 5.** Horizontal cross sections for all five time steps at $Z = 3.5$ km: (a) simulated radar observations, along with (e) the analytical “truth” field; also shown are gridded retrievals for the (b) Cressman, (c) nearest, and (d) variational methods.
As in section 3, the variational gridding method outperforms existing radar gridding methods in this experiment. Variational retrievals in Figs. 5d and 6d closely mimic the underlying radar information in terms of both amplitude and spatial extent, while successfully filtering radar noise to produce visually smooth outputs. These assessments are supported by temporally varying validation statistics presented in Fig. 7. Variational RMSEs vary between 0.25 and 0.30 throughout all five time steps, roughly a factor-of-3 and factor-of-6 less than those from the nearest and Cressman methods, respectively (RMSE calculations are limited to 10 km either side of the cell location along the x axis for each time step). The extent of variational retrievals, as measured by cell volume, is also very close to that of the analytical field throughout the experiment. In contrast, fields gridded using the Cressman method overestimate the storm extent due to boundary effects, and those from the nearest-neighbor method slightly underestimate storm extent due to data gaps introduced by the radar’s scanning strategy.

The final validation statistic in Fig. 7 is the maximum reflectivity value within the cell as it traverses the grid. This value should theoretically exhibit no temporal dependence [Eq. (14)] mandates a constant reflectivity maximum of 50 dBZ; however, we observed that the simulated radar reflectivities

![Figure 6](image-url)

Fig. 6. As in Fig. 5, but for a vertical cross section at Y = 15 km. The scanning geometry of simulated PPI radar sweeps is also shown with black dotted lines.
steadily decrease with range (red line, Fig. 7c). The familiar range dependence in radar sensitivity arises from low-pass filtering by the radar beam illumination functions in Eqs. (16) and (17) (i.e., nonuniform beam filling; Srivastava and Atlas 1974; Doviak and Zrnić 1993). Amplitudes of all gridded radar fields also incur this anisotropic range bias as a result. However, these effects are not of concern in this study as the success criteria for gridded fields (defined in section 1) are defined in relation to the radar data, not the underlying physical phenomena. Accordingly, the gridded maxima in Fig. 7 should be judged with radar maxima as “truth.” Variational maxima closely mimic the radar maxima through time, albeit with a near-constant negative offset of ~2–3 dBZ arising from the filtering of observational noise present in the radar measurements. Larger reflectivity offsets in the nearest-neighbor and, especially, Cressman techniques (10 dBZ in this example), have serious consequences for applications involving the direct use of radar reflectivity. Examples of such applications include rain rates $R$ (assuming standard $Z$–$R$ relationships), hail retrievals such as MESH (refer to appendix A in Warren et al. 2020), convective/stratiform delineation (Steiner et al. 1995) and storm identification/tracking (Dixon and Wiener 1993; Lakshmanan et al. 2009).

Fig. 7. Statistics for the discrete storm experiment across all five time steps: (a) the RMSE score for each time step, (b) the volume of each discrete storm cell, and (c) the maximum field value for each gridding method and the simulated radar data.

5. Real data case study
   a. Gridding performance

The last experiment presented here is a case study from the CPOL research radar, located in tropical Darwin, Australia (Louf and Protat 2021). The event took place on 18 February 2014, during a period when CPOL operated using identical sampling characteristics to those outlined in section 4b (half-power beamwidth and azimuthal spacing 1°, 250-m range gates, and same scanning pattern). The Cartesian analysis domain has an isotropic 500-m spacing, extending 300 km in the $x$ and $y$ dimensions and 20 km in the vertical direction. The extended radar range in this section warrants a change in gridding parameters to a 3500-m Cressman ROI and smoothing weights $k_V = 0.1$, $k_H = 0.5$ for the variational method. The case contains a period of severe multicellular convection, making it an ideal test case for radar gridding due to the presence of storm cells at varying stages of maturity. The following retrievals were performed on a volumetric scan measured between 1255 and 1300 UTC. No advection correction is employed to account for measurement nonsimultaneity so as to increase the clarity and reproducibility of the results.

Figures 8 and 9 show gridded radar reflectivity for each gridding method, along with raw radar reflectivity in polar coordinates for comparison. First, the theoretical issues involving Cressman oversmoothing are clearly present in practice. Take, for example, two small, newly initiating storm cells on the northern edge of the radar reflectivity field in Fig. 8.
These features are nearly completely aliased into one larger cell in the Cressman grid in Fig. 8c, with an estimated reflectivity ∼20 dB lower than measured by the instrument. Similar oversmoothing occurs throughout Figs. 8a–c, leading to a loss of important small-scale, high-reflectivity features. The expansion of radar reflectivity beyond its true spatial extent (discussed in section 4, Figs. 5 and 6), is also particularly evident across all gridded Cressman fields, but perhaps best observed in the vertical cross section between 0 and 50 km along the x axis in Fig. 9a. Spurious storm-top reflectivity maxima mentioned in OSSE experiments in sections 3 and 4 are also present in Figs. 9b and 9c. Without judicious consultation of the raw radar data in Figs. 9k and 9l, these artifacts could be falsely interpreted as meteorological features. Overall, the poor resolution and interpretability of Cressman interpolation method may render it unsuitable for many common radar applications.

The most notable feature of gridded nearest-neighbor fields may be the limited extent of the echoes above and below the lowest and highest elevation tilts. Retrievals in the z = 1.5 km horizontal cross section in Fig. 8d are limited to an ∼100-km range, and these effects are more severe at lower altitudes (note the missing data in the low levels in Fig. 9d). These discontinuities may pose challenges in generating valid hydrological retrievals such as rain rate from the lowest altitude horizontal cross section. Data extent limitations may also be noted in Figs. 9e and 9f, where echo tops spuriously mimic the measurement geometry of the observations in Figs. 9k and 9l. Some degree of observational noise is also clearly present in fields gridded using the nearest-neighbor method (e.g., Fig. 8e). Importantly, no user-defined parameters can manually increase the level of smoothness in the nearest-neighbor method, which may pose serious challenges to applications requiring high-level spectral fidelity. Despite the problems involving interpretability, continuity, and smoothness, the nearest-neighbor method does reproduce the raw radar information with very high resolution. Small features and high reflectivity values are notably well-resolved everywhere outside the aforementioned missing data regions.

Figures 8 and 9 confirm that the variational method demonstrates similarly promising performance as in sections 3 and 4.
All small-scale features (e.g., the two small, initiating cells on the northern edge of Fig. 8i) are retrieved with accurate amplitudes and extent. Comparisons between variational and radar insets (e.g., Figs. 8h,k) also reveal that observational noise is filtered from the variational analysis. The result is a visually smooth reflectivity grid, with no obvious compromise in data resolution. Data acquisition artifacts are masked by the variational methodology because of the spatial expansion of valid radar information into data voids. For example, note the difference between Figs. 9f and 9i in terms of the vertical continuity of grid retrievals. Despite this ability to mask data acquisition gaps, valid information is not “smudged” into weak echo regions, as evidenced by the retention of nonprecipitating regions between Figs. 8h and 8k.

The largest disparity in variational retrievals may be the introduction of a layer of low-reflectivity signatures (5–10 dBZ)

**FIG. 9.** As in Fig. 8, but for vertical cross sections at Y = 0. (j),(k),(l) “Pseudo” RHI data from the 90° and 270° azimuths, where raw radar data are shaded inside the bounds of the 1° angular beamwidth of each ray.
atop the available radar data (refer to Fig. 9g). There is a notable absence of low reflectivity echoes at storm top, due to the poor sampling extent and the aforementioned range dependence in radar sensitivity, meaning that the highest observed radar echoes may actually arise from well within precipitating regions (e.g., Fig. 9l). The variational method produces more visually realistic echo-top heights by vertically extending high reflectivity signatures at storm top via spatial continuity and smoothness constraints.12 Figure 10 shows how variationally derived echo tops are more structurally realistic than either comparison method. The first highlighted region is well-observed by the radar at ~75-km range; however, echo-top heights are still plagued by circular data geometry artifacts in the Cressman and nearest-neighbor methods. In contrast, the variational method produces a realistic depiction of a roughly circular, deep convective plume. Data geometry artifacts are still present in the variational method in the second highlighted region due to the radar’s cone of silence at smaller range; however, it still offers a considerable improvement over artifacts present in Figs. 10a and 10b. The third highlighted region in Fig. 10b contains significant speckle noise due to missing data in weak echo regions at storm top. In contrast, the variational method produces a realistic, physically continuous echo-top height through the vertical propagation of valid information. Verification of

12 Note that the current best practice for estimating echo-top heights (Lakshmanan et al. 2013) cannot perform these corrections because tops are taken from the height of the highest tilt in regions of missing data aloft.
actual height retrievals against an independent echo-top height source (such as radar instruments aboard precipitation monitoring satellites) should be performed in the future to ascertain which variational smoothing parameters produce the most accurate echo-top height fields in practice.

b. Computational efficiency

The final point of discussion in this work is about the computational efficiency of each gridding method. The efficiency criterion outlined in section 1 stated that a method was deemed acceptable if it was efficient enough for real-time applications. This is an intentionally vague definition, as it is contingent on application-specific variables such as input size (resolution, and number of radar inputs), output size (analysis grid resolution), and hardware choices (e.g., personal computer vs GPU cluster). The intention of this section is to provide some nonexhaustive guidance to analysts on which gridding methods may be computationally acceptable for their intended applications. To this end, we have chosen the CPOL radar data example as a representative, baseline example for the efficiency one may expect for each method in practice. We deem a method to be suitable for real-time applications if its wall time is less than the updating frequency of the data source (every ~5 min in this case). Standard CPU experiments (including the Cressman and nearest-neighbor implementations) are performed on a high-performance computing cluster consisting of 12 × Intel Xeon Platinum 8268 CPUs, and GPU experiments are performed on a single Nvidia V100-SXM2 GPU.

Figure 11 illustrates how the wall time of each gridding method varies as a function of the analysis grid size. For all methods, the wall time varies linearly with respect to the number of grid points in the analysis (resulting in a logarithmic appearance with the logarithmic y axis). Analysis grid sizes were varied from a coarse spacing of 5 km up to the high-resolution 500-m grids shown in Figs. 8–10. Note that the standard operational CPOL gridding resolution is 500 m vertically, and 1 km horizontally (resulting in 3.6 × 10^6 grid points; Louf and Prost 2021). The nearest-neighbor and Cressman methods are notably efficient over the entire range of analysis sizes, particularly the former, owing to the vectorized implementation outlined in section 2c. The variational method remains well below the ~5-min real-time efficiency threshold because of its implementation on GPU hardware, despite being slower by a factor of between 2 and 4 than the Cressman method across all analysis sizes. Variational methods are particularly suited to GPUs due to the large number of matrix operations performed during the recursive optimization process, resulting in a factor-of-20–40 speedup relative to the standard CPU implementation in Fig. 11. Using the aforementioned real-time definition, the variational method may be employed in real time on CPU hardware at 1000-m resolution (~3.5-min wall time) but not for the high-resolution grids used in this study (500-m spacing; ~30-min wall time).

6. Summary and future directions

A variational interpolation method for gridding radar data was introduced in this study. A series of OSSE experiments, and a real case study, showed that the variational method can offer substantial improvements over existing distance-weighting and nearest-neighbor methods. These improvements are summarized in Table 1, in which each method is assessed in terms of the gridding performance criteria outlined in section 1. Evidence for these classifications is provided in sections 3–5 and is briefly summarized below.

The main drawback in the Cressman gridding method is its poor resolution, whereby small features and high reflectivity regions are filtered from the analysis. This criterion alone renders the method unsuitable for many applications, including operational forecasting and hydrological retrievals. Perhaps the only practical applications for the Cressman method are those that are highly intolerant to noise, such as single and dual-Doppler wind retrievals. The large radius of influence used by the Cressman method masks data acquisition gaps, but also smudges valid radar data to an unacceptable extent. This, along with the introduction of spurious artifacts due to boundary effects, make the results difficult to interpret for untrained analysts.

As expected, the nearest-neighbor method scores poorest in the smoothness category, due to the propagation of measurement noise and the introduction of high frequencies due to discontinuities in the gridded fields. The interpretability and continuity of nearest-neighbor fields are inherently tied to the observational geometry of the radar, producing significant data gaps below and above the radar measurement extent and propagating missing data into the gridded fields. Despite these drawbacks, and especially in applications where spectral fidelity may be of secondary importance, the nearest-
The most significant finding in this study is the ability of the variational method to simultaneously achieve excellent resolution and smoothness characteristics in gridded fields. The results are visually pleasing, and mask data acquisition gaps by spreading valid information into data voids. The continuity and faithfulness of gridded variational fields to the underlying physical phenomena mean they may be more easily interpreted by a radar nonexpert, lowering the degree of specialist knowledge required to use radar data in practice. Furthermore, the ability to control the degree of spatial smoothness and noise filtering means the new method is suitable to a wider range of applications involving gridded radar data.

The most significant barrier to the method’s widespread use may be its computational efficiency in the absence of GPU computing hardware. To this end, future work will focus on implementing preconditioning techniques to reduce the number of iterations required during optimization, permitting efficient use on standard CPUs (e.g., Claerbout 2014). We also aim to validate the improvements achieved with the variational method in secondary applications such as echo-top heights, hail retrievals, and three-dimensional wind retrievals. Future work aims to show how this method is naturally suited to multiradar mosaic generation through the assimilation of radar data from multiple radars.

**Acknowledgments.** Author Brook acknowledges industry research funding provided by Guy Carpenter and Company, LLC. We also thank Valentin Louf, Corey Potvin, and Xavier Poncini for stimulating discussions related to this work.

**Data availability statement.** Download instructions for the CPOL research radar data used in this study are available online (https://www.openradar.io/cpol).

### APPENDIX A

**Linear Interpolation Operator**

The foundation of the variational interpolation approach used in this study is a linear interpolation operator \( \mathcal{R} : \mathbb{R}^M \rightarrow \mathbb{R}^N \), which maps gridded analysis fields in model space \( \varphi \in \mathbb{R}^M \) to radar information in data space \( d \in \mathbb{R}^N \). Refer to Fig. A1 for a two-dimensional illustration of this process, which is identical to those outlined in Testud and Chong (1983) and Barth et al. (2021). In our study, this process occurs in three dimensions, necessitating a trilinear interpolation from the 8 model points surrounding each radar data point. In forward mode, the interpolation step proceeds according to the standard trilinear equations:

\[
\varphi(x, y, z) = \varphi_{i,j,k}(1 - w_x)(1 - w_y)(1 - w_z)
+ \varphi_{i+1,j,k}w_x(1 - w_y)(1 - w_z)
+ \varphi_{i,j+1,k}(1 - w_x)w_y(1 - w_z)
+ \varphi_{i,j,k+1}(1 - w_x)(1 - w_y)w_z
+ \varphi_{i+1,j+1,k}w_xw_y(1 - w_z)
+ \varphi_{i+1,j,k+1}(1 - w_x)w_yw_z
+ \varphi_{i,j+1,k+1}w_xw_z(1 - w_y)
+ \varphi_{i+1,j+1,k+1}w_xw_yw_z
\]

(A1)

**FIG. A1.** A 2D schematic demonstrating the linear interpolation operator methodology. Black points indicate grid points, and red points indicate radar data points. The corresponding solid black and dotted red lines assist in visualizing the associated Cartesian and spherical coordinates, respectively. Black arrows illustrate the flow of information for operator forward projection.
where \( \varphi(x, y, z) \) denote field values at observation points, which are estimated via linear weighting along each axis \( [e.g., w_x = (x - x_i)/\Delta x] \), where \( \Delta x \) is the corresponding grid spacing. The trilinear adjoint \( (\mathcal{R}^*: \mathbb{R}^N \to \mathbb{R}^M) \) projects data residuals back to model space (often referred to as back projection; Claerbout 2014), inverting the flow of information indicated for the forward projection (black arrows in Fig. A1). The adjoint is implemented by adding back a weighted radar datum to model points according to the trilinear weights provided in Eq. (A1).

**APPENDIX B**

**Smoothing Regularization**

The addition of regularization constraints is crucial for underresolved inverse problems such as radar gridding. Model regularization through smoothing norms is common for variational problems in meteorology; however, there is currently no consensus on the optimal formulation of such constraints. Here, we discuss the benefits and drawbacks of several common approaches. Smoothing constraints generally
contain squared partial derivatives of model fields, which act as a low-pass filter by suppressing high frequencies that are likely due to measurement error and retaining resolvable frequencies containing valid information. Testud and Chong (1983) showed that in general, higher-order derivatives result in steeper gain curves around the cutoff frequency (theoretically resolving more small-scale features). Furthermore, methodologies for 3D oceanographic applications recommend more complex smoothing constraints containing derivatives up to third order, along with mixed partial derivatives of varying order (McIntosh 1990; Barth et al. 2014). By contrast, smoothing regularization terms in meteorological literature are generally simpler, involving unmixd partial derivatives of either first (e.g., Shapiro et al. 2009) or second order (e.g., Potvin et al. 2012a). Regularization with second-order (Laplacian) terms leads to the minimum curvature solution, which produces a visually pleasing result analogous to a thin elastic plate flexed to fit through data points (Briggs 1974). These constraints efficiently “spread” valid radar information throughout the model grid by propagating valid measurements into data voids (a highly desirable property of interpolation methods); however, they can also cause unwanted extraneous inflections (i.e., flexural bulges in the earlier elastic plate analogy, shown in Fig. B1b). Conversely, models constrained by first-order (gradient) penalties theoretically lead to solutions of the Laplace equation, which do not permit local extrema within unconstrained regions. This damps extraneous inflections, albeit with compromised smoothness properties due to a relatively flat gain curve and the permission of gradient discontinuities (Shapiro et al. 2009). Finally, a common alternate approach combines the desirable smoothing qualities of first- and second-order derivatives by using both as regularization terms (e.g., Smith and Wessel 1990). The addition of first-order derivatives constrains solutions in data voids, analogous to the application of “tension” at the extremities when fitting a thin elastic plate. Aptly, this method is known as splines under tension and has been successfully employed in a range of variational interpolation applications (e.g., Fomel et al. 2003).

All of the aforementioned smoothing constraints were tested by the authors in the preparation of this study, and the resulting choice of smoothing constraints was made on this experimental basis. Unmixed, second-order partial derivatives were found to offer the optimal regularization properties, and the result of optimization with this constraint alone is shown in Fig. B1b for the OSSE experiment introduced in section 4. Second-order derivatives alone produce a suitably smooth retrieved field that accurately resolves the maximum reflectivity within the cell. However, the aforementioned extraneous inflections characteristic of second-order derivatives are created along first-order discontinuities on the cell boundary (refer to the 1D cross section in Fig. B1b). These artifacts could cause significant issues for analyses such as 3D wind retrievals that rely on accurate divergence and vorticity fields and may be incorrectly interpreted as physical phenomena.

The introduction of a background constraint [Eq. (4)] to constrain regions outside radar coverage suppresses these reflections, while having a negligible effect on the model fit within the cell region. The RMSE fit of the cell is improved from 0.39 to 0.26 with this addition alone, and we observe similar performance increases in other OSSE experiments and real data. The one-dimensional cross section in Fig. B1c reveals that the background constraint alone cannot completely eliminate the presence of extraneous inflections, necessitating the introduction of the final, denoising constraint Jp. The anisotropic total variation denoising constraint suppresses speckle noise, permitting one to decrease the reliance on second-order smoothing constraints for denoising purposes. Reduced Jp weightings decrease the production of flexural bulges, and the ℓ2 penalty of the denoising constraint permits the retention of sharp “edges” at the cell boundary in Fig. B1d (Rudin et al. 1992). The addition of the denoising constraint means that the peak reflectivity value is not fully resolved (resulting in a slightly increased RMSE value relative to Fig. B1c); however, we argue that this is a desirable trade-off for the complete elimination of extraneous inflections.

REFERENCES


optimization of expensive black-box functions, arXiv, 1605.00998v1,
Barnes objective map analysis scheme for use with satellite
and conventional data. J. Climate Appl. Meteor., 22, 1487–1503,
Lakshmanan, V., T. Smith, K. Hondl, G. J. Stumpf, and A. Witt,
2006: A real-time, three-dimensional, rapidly updating, hetero-
geneous radar merger technique for reflectivity, velocity, and
doi.org/10.1175/WAF942.1.
Barnes objective map analysis scheme for use with satellite
and conventional data. J. Climate Appl. Meteor., 22, 1487–1503,


