

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Reducing Errors in Velocity–Azimuth Display (VAD) Wind and Deformation Retrievals from Airborne Doppler Radars in Convective Environments

CHARLES N. HELMS,^{a,b} MATTHEW L. WALKER MCLINDEN,^a GERALD M. HEYMSFIELD,^a AND STEPHEN R. GUIMOND^{c,a}

^a NASA Goddard Space Flight Center, Greenbelt, Maryland

^b Universities Space Research Association, Columbia, Maryland

^c University of Maryland, Baltimore County, Baltimore, Maryland

(Manuscript received 12 March 2020, in final form 30 September 2020)

ABSTRACT: The present study describes methods to reduce the uncertainty of velocity–azimuth display (VAD) wind and deformation retrievals from downward-pointing, conically scanning, airborne Doppler radars. These retrievals have important applications in data assimilation and real-time data processing. Several error sources for VAD retrievals are considered here, including violations to the underlying wind field assumptions, Doppler velocity noise, data gaps, temporal variability, and the spatial weighting function of the VAD retrieval. Specific to airborne VAD retrievals, we also consider errors produced due to the radar scans occurring while the instrument platform is in motion. While VAD retrievals are typically performed using data from a single antenna revolution, other strategies for selecting data can be used to reduce retrieval errors. Four such data selection strategies for airborne VAD retrievals are evaluated here with respect to their effects on the errors. These methods are evaluated using the second hurricane nature run numerical simulation, analytic wind fields, and observed Doppler radar radial velocities. The proposed methods are shown to reduce the median absolute error of the VAD wind retrievals, especially in the vicinity of deep convection embedded in stratiform precipitation. The median absolute error due to wind field assumption violations for the along-track and for the across-track wind is reduced from 0.36 to 0.08 m s^{−1} and from 0.35 to 0.24 m s^{−1}, respectively. Although the study focuses on Doppler radars, the results are equally applicable to conically scanning Doppler lidars as well.

KEYWORDS: Data processing; Radars/Radar observations; Wind profilers; Error analysis

1. Introduction

The velocity–azimuth display (VAD) technique has traditionally been used to retrieve vertical profiles of horizontal wind and deformation using ground-based Doppler radar or lidar. The technique, originally developed by Lhermitte and Atlas (1961) and Browning and Wexler (1968), uses Fourier analysis to fit sine and cosine waves to the Doppler velocity measurements as a function of azimuth angle and requires the radar to perform conical scans about the vertical axis. A traditional ground-based VAD analysis using a single scan with a low elevation angle produces five kinematic variables: zonal wind, meridional wind, shearing deformation, stretching deformation, and divergence (Browning and Wexler 1968). The divergence is computed by assuming that the Doppler vertical velocity is negligible. The Doppler vertical velocity W is defined here as

$$W = V_T - w, \quad (1)$$

where V_T is the fall speed of the scattering particle, w is the vertical air motion, and positive values indicate downward motion. By incorporating multiple elevation angles into the analysis, the Doppler vertical velocity can also be computed (Srivastava et al. 1986; Matejka and Srivastava 1991).

The traditional VAD technique makes a number of assumptions regarding the ingested observations and the underlying

wind field. Critical among these assumptions is that the underlying wind field, as represented by the three-dimensional motion of the scattering particles, is linear in the horizontal and constant in the vertical (Browning and Wexler 1968). The technique also assumes that any temporal variations in the wind field are negligible over the scan period (Browning and Wexler 1968). Furthermore, the retrievals are also typically treated as being oriented in the horizontal plane (e.g., the horizontal wind) and, as such, the scan traced out by a single range gate is assumed to occur entirely in the horizontal plane (Browning and Wexler 1968). Finally, because the retrievals are often treated as a point observation, there is also the inherent assumption that the radar instrument platform remains stationary throughout the scan (Tian et al. 2015).

VAD retrievals have been used to study a wide variety of phenomena including frontal systems (Browning and Harrold 1970; Hagen 1992; Neiman and Ralph 2004), tropical cyclones (Lorsolo et al. 2008; Kao et al. 2019), orographic flow (Reitebuch et al. 2003; Witschas et al. 2017), and even migratory bird behavior (Horton et al. 2016). Michelson and Seaman (2000) demonstrated that assimilating VAD wind retrievals improved the initial conditions of a numerical model. Studies have shown that assimilating observations with a higher data density than that of the model grid can degrade the accuracy of the initial conditions (e.g., Liu and Rabier 2002; Bick et al. 2016). The VAD technique effectively reduces the observation density by summarizing data over a relatively large area into a single profile of wind observations that is then assimilated into the model.

Corresponding author: Charles N. Helms, charles.n.helms@nasa.gov

Tian et al. (2015) demonstrated that the VAD technique could be applied to a downward-pointing, conically scanning, airborne Doppler radar if the aircraft-mounted radar is assumed to be stationary for the duration of the scan. In their error analysis, Tian et al. (2015) found that errors due to the assumption of a stationary instrument platform were acceptably small within stratiform regions where horizontal wind gradients are relatively small and the variation in vertical velocity within each scan is minimal. Additionally, the authors derived a set of equations to describe the errors introduced by the stationary instrument platform assumption.

In total, the present study will examine six sources of error in the VAD retrievals, enumerated below. The first error source is that which is introduced by assuming the radar platform is stationary, as discussed in Tian et al. (2015). The second error source of interest is associated with the noise inherent in the Doppler radial velocity retrievals themselves. The third source is associated with gaps in the Doppler velocity data coverage, either due to insufficient scattering particles (i.e., clear air) or due to technical issues during the data collection process. The fourth error source is the violation of the underlying assumptions that the VAD technique makes about the wind field: specifically that the horizontal winds vary linearly and the vertical particle motions are constant over the ingested data. The fifth error source of interest is produced by temporal variability during a scan and can be associated with variability in the aircraft orientation or in the underlying wind field. The final error source that will be discussed in this paper is associated with the spatial weighting function of the ingested data, where, traditionally, all of the information comes from the edge of the spatial footprint¹ and no information comes from the interior of the spatial footprint. While other sources of error may exist, the present study will focus on addressing these six sources. These error sources are discussed in further detail in section 2c.

The goal of the present study is to reduce the influence of the above VAD error sources when applied to airborne Doppler radar in stratiform regions with embedded deep convection (e.g., tropical cyclones). This error reduction will be accomplished by designing strategies for choosing which data are used in a VAD retrieval. These strategies, hereafter referred to as data selection strategies, are designed based on two sets of trait pairs. The first trait pair, sequential–synthetic, determines how the data are selected. A sequential strategy includes all data within a temporal window while a synthetic strategy includes all data within an along-track distance window. As will be discussed later, the synthetic strategies have a reduced spatial footprint from the sequential strategies at the expense of having an increased temporal footprint, defined here as the time between the earliest and latest observation incorporated into the VAD retrieval. The second trait, single scan–multiscan,

determines how much data are included in each VAD retrieval. A single-scan strategy includes an equal number of data points as is collected during a single revolution of the radar antenna while a multiscan strategy includes more data than would be collected by a single revolution. By this convention, a traditional VAD retrieval uses a sequential single-scan strategy. Examples of the four data selection strategies are depicted in Fig. 1.

Although VAD retrievals are typically applied to a single scan, there are no inherent assumptions that prevent the use of more than one scan of data to perform retrievals. Likewise, the VAD technique does not require Doppler velocity data to be fed into the retrieval algorithm as a continuous sequence of data collected at monotonically increasing azimuth angles. As such, retrievals can be performed using Doppler velocity data collected from multiple scans or from a subset of angles within a series of scans. The sequential multiscan, synthetic single-scan, and synthetic multiscan data selection strategies take advantage of this flexibility and provide benefits over the traditional sequential single-scan strategy. These benefits primarily take the form of error reduction (reducing the magnitude of the errors) and error isolation (reducing the number of retrievals contaminated by errors) and will be elucidated over the course of this paper using a combination of analytic wind fields, numerical simulation data, and observations.

Section 2 will introduce both the VAD method in general and the four data selection strategies in particular as well as provide a discussion of the six error sources noted above. Section 3 will discuss the data and methods used to evaluate the four data selection strategies. This will be followed by an error analysis of the four data selection strategies in section 4. Finally, an overview of the results will be provided in section 5.

2. VAD methods and error sources

a. VAD equations

The derivation of the VAD analysis equations for a moving instrument platform generally follows that of Tian et al. (2015). Assuming that the track-relative radar elevation angle Φ is constant, the radial wind V_r at a particular track-relative azimuth angle θ can be expressed as

$$V_r = u \cos\Phi \sin\theta + v \cos\Phi \cos\theta - W \sin\Phi, \quad (2)$$

where u is the across-track wind (positive to the right of the track), v is the along-track wind (positive in the direction of aircraft travel), and W is the Doppler vertical velocity (positive downward). Recall, the Doppler vertical velocity is the sum of the vertical air motion and the fall speed of the scattering particles, both of which are taken as the reflectivity-weighted volume averages. Furthermore, it is assumed that, prior to the calculations presented here, the aircraft motion has already been removed from the Doppler velocities.

The VAD technique assumes linear horizontal wind fields and constant Doppler vertical velocities along the scan (Browning and Wexler 1968). The linear decomposition of

¹ For our purposes, the spatial footprint of a scan will specifically refer to the region within the convex hull of the data points used for an individual VAD retrieval.

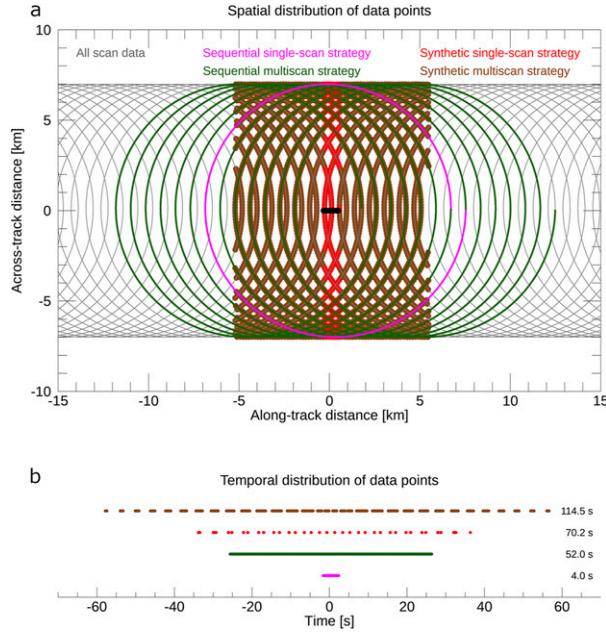


FIG. 1. Examples of the (a) spatial distribution and (b) the temporal distribution of data points in the four data selection strategies using the NASA ER-2 X-band radar (EXRAD) scan geometry. The gray line in (a) indicates the locations of all scan data. The data employed by a sequential single-scan retrieval are indicated in pink while the additional data employed by a sequential multiscan retrieval are indicated in green. The thick red points indicate data employed by a synthetic single-scan retrieval while the thick brown points indicate the additional data employed by a synthetic multiscan retrieval. The thick black bar in the center indicates the centers of curvature for the sequential single-scan strategy. Note, both the sequential multiscan and synthetic multiscan retrievals use the same data as their single-scan retrieval counterparts in addition to the data indicated by their respective colors. The temporal footprints of each strategy in seconds are indicated by the numbers on the right side of (b).

the horizontal winds, assuming constant Doppler vertical velocity, gives the following set of equations:

$$\begin{aligned} u &= u_0 + \frac{\partial u}{\partial x}(x - x_0) + \frac{\partial u}{\partial y}(y - y_0), \\ v &= v_0 + \frac{\partial v}{\partial x}(x - x_0) + \frac{\partial v}{\partial y}(y - y_0), \quad \text{and} \\ W &= W_0, \end{aligned} \quad (3)$$

where x indicates the across-track position, y indicates the along-track position, and the subscript 0 indicates the value at the platform stationary position. For a stationary platform, the scan will be a perfect circle; for a moving platform, the scan will be a spiral pattern. To replicate the spiral scan pattern produced by the moving platform, the definition of $(y - y_0)$ must be modified to include aircraft motion. Following Tian et al. (2015), $(x - x_0)$ and $(y - y_0)$ are defined as

$$\begin{aligned} (x - x_0) &= r \cos \Phi \sin \theta \quad \text{and} \\ (y - y_0) &= r \cos \Phi \cos \theta + \left(\frac{U_a \tau}{2\pi} \right) \theta, \end{aligned} \quad (4)$$

where r is the range from the radar, U_a is the aircraft ground speed, and τ is the time required to complete a single scan. Following the derivation of Tian et al. (2015), combining Eqs. (2)–(4) gives us

$$\begin{aligned} V_r &= C_0 + C_1 \cos \theta + C_2 \sin \theta + D_1 \cos(2\theta) \\ &\quad + D_2 \sin(2\theta) + E_1(\theta \cos \theta) + E_2(\theta \sin \theta), \end{aligned} \quad (5)$$

with

$$\begin{aligned} C_0 &= -W_0 \sin \Phi + \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \frac{r \cos^2 \Phi}{2}, \\ C_1 &= v_0 \cos \Phi, \quad C_2 = u_0 \cos \Phi, \\ D_1 &= \frac{r \cos^2 \Phi}{2} \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \right), \quad D_2 = \frac{r \cos^2 \Phi}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right), \\ E_1 &= \frac{U\tau}{2\pi} \cos \Phi \frac{\partial v}{\partial y}, \quad \text{and} \quad E_2 = \frac{U\tau}{2\pi} \cos \Phi \frac{\partial u}{\partial y}. \end{aligned} \quad (6)$$

The seven coefficients, defined in Eq. (6), are used to compute the kinematic variables that describe the wind field: C_0 is a linear combination of Doppler vertical velocity and horizontal divergence, the horizontal wind components are derived from C_1 and C_2 , the stretching and shearing deformations are found using D_1 and D_2 , and the along-track gradients of the along-track and across-track winds can be found using E_1 and E_2 .

Although it is possible to retrieve values for E_1 and E_2 by employing a multiple linear regression to retrieve the coefficients in Eq. (5), the independent variables associated with the E_1 and E_2 coefficients, $\theta \cos \theta$ and $\theta \sin \theta$, suffer from collinearity issues with the independent variables for the C_1 and C_2 coefficients, $\cos \theta$ and $\sin \theta$. Because of this collinearity, the regression has issues correctly assigning variability in the Doppler velocities to these four variables, resulting in unreliable retrievals of the horizontal wind and the along-track gradients. It should be noted, however, that the direct retrieval of the E_1 and E_2 coefficients would be highly desirable as including these coefficients in the VAD analysis would remove the need to assume a stationary platform.

b. Data selection strategies

As previously mentioned, the four data selection strategies are described by combinations of two trait pairs: sequential–synthetic and single scan–multiscan. Recall, sequential strategies employ all data within a temporal window, synthetic strategies employ all data from an along-track distance window, single-scan strategies employ an equivalent amount of data to that collected during a single revolution of the radar antenna, and multiscan strategies employ more data than are collected during a single revolution of the radar antenna.

It has already been established that traditional VAD retrievals, and those described in Tian et al. (2015), use the sequential single-scan strategy (pink points in Fig. 1), whereby the data from a single revolution of the radar antenna are

ingested into the VAD algorithm. In contrast, a synthetic single-scan retrieval (red points in Fig. 1) is produced by performing the VAD retrieval on all data that are located within a window whose along-track width is equal to the distance that the aircraft travels during a single revolution of the radar antenna (i.e., $U_a \tau$). As a result, a synthetic single-scan retrieval uses data collected over multiple scans.

For the synthetic single-scan strategy, the full 360° of azimuth angles subtended during a single revolution of the radar antenna are accounted for within the selection window. Visually, the equivalence of the synthetic single-scan and sequential single-scan arc lengths can be demonstrated by a careful inspection of the synthetic single-scan retrieval points depicted in red in Fig. 1a: the arc segments fit together without overlap.² Because the arc lengths are equal, the subtended azimuth angles must also be equal. It is worth noting that the synthetic single-scan strategy has the smallest scan spatial footprint that is possible without reducing the number of data points from what is included by a sequential single-scan strategy.

Kelberlau and Mann (2019) have previously demonstrated the application of a synthetic single-scan approach, referred to there as the “squeezing” method, for VAD retrievals of turbulence from a stationary scanning lidar. Their approach assumes that the turbulence is steady state and is advected over the lidar scan circle by the mean wind, which effectively replaces the aircraft ground speed U_a used in the present study. The mean wind, however, must be either measured by an external source or retrieved using a traditional sequential single-scan method in order to implement their method. Due to the similarity of the synthetic single-scan approach discussed here and the squeezing method of Kelberlau and Mann (2019), a portion of the error analysis discussed below will likely be relevant to both methods, although there are error sources that will be specific to each method.

The sequential multiscan (green points in Fig. 1) and synthetic multiscan (brown points in Fig. 1) strategies are produced by including data from a series of single-scan strategies. For the purpose of this paper, we have chosen a sequence of scans for the sequential multiscan retrieval by selecting all radar revolutions that have a scan centroid located within a 10-km window, whose size was chosen arbitrarily, centered on a given sequential single-scan centroid. In practice, however, a sequential multiscan retrieval can also be performed by selecting all data from within a predetermined temporal window. Regardless of how the multiple scans are selected, the ideal multiscan spatial footprint should balance the ability to resolve features of interest with the error reduction benefits discussed in section 4, some of which favor larger spatial footprints. Here, our selection criterion results in a sequence of 13 complete revolutions of the radar; as such, the multiscan retrieval examples in this paper use 13 times the number of data points as the single-scan retrievals. It should also be noted that the multiscan windows overlap in the analyses presented here. To

simplify the error analysis, we chose the data selection window for the synthetic multiscan strategy to be exactly 13 times the along-track distance window of the synthetic single-scan strategy. Thus, each synthetic multiscan retrieval includes all data collected within a 10 712-m window centered on each synthetic single-scan centroid and contains the same number of data points as the sequential multiscan strategy.

c. Error sources

As stated above, the present study focuses on six VAD retrieval error sources: the assumption that a moving instrument platform is stationary, Doppler velocity uncertainty, gaps in data coverage, violations of the basic VAD assumptions regarding the observed wind field, temporal variations, and representativeness issues introduced by the spatial weighting function inherent in the VAD technique. This section will briefly discuss each of these error sources.

As discussed in Tian et al. (2015), in order to perform a VAD retrieval on airborne radar data, it is necessary to assume that the instrument platform is stationary. A VAD retrieval using a stationary instrument platform returns the wind field valid at the center of curvature of the scan, assuming the wind field is linear and there is no variation in the vertical velocity over the scan. When the instrument platform is moving, the center of curvature of the scan traces a line rather than a point (Fig. 1a). As a result, there are terms in the VAD equation (see section 4a) that include the along-track gradient of the horizontal winds over the center-of-curvature line (Tian et al. 2015). These terms are colinear with the along-track and across-track wind terms in the VAD equation and, as such, the retrieval cannot properly assign the azimuthal variability in the Doppler velocity between these two sets of terms. As already mentioned, the solution to this is to assume the instrument platform is stationary at some point along the aircraft flight path; the error due to the stationary instrument platform assumption depends on where the platform is assumed to be stationary.

Doppler velocity data have an inherent level of uncertainty associated with them. This uncertainty in the Doppler velocity is introduced from a variety of sources (e.g., noise, partial beam filling, side-lobe contamination). Frequently, airborne radars have an engineering goal of achieving Doppler velocity uncertainty levels below $\sim 2 \text{ m s}^{-1}$, although, based on the authors' experience, the actual uncertainty is often an order of magnitude smaller than this. While the uncertainty in the raw data is often reduced via averaging during the quality control process, residual errors will remain in the final data. A key strength of VAD retrievals over multi-Doppler retrievals (e.g., Guimond et al. 2014; Didlake et al. 2015) is that the technique reduces the impact of Doppler velocity uncertainty due to the large number of data points that go into each retrieval. Data selection strategies that increase the number of data points used in an individual VAD retrieval should have the effect of further reducing the impact of Doppler velocity uncertainty.

As the radar beam traces out a scan, the beam often passes through regions devoid of scattering particles capable of producing a detectable signal (i.e., clear-air regions). Additionally, technical issues with the radar or data recorder

² Note that the width of the highlighting in Fig. 1a conceals a gap at the back of the scan directly below the aircraft into which the isolated red point, located at the center of the domain, fits.

can result in periods of missing data and excessive attenuation of the radar beam can result in regions of missing data despite the presence of scattering particles. These gaps in the radar data, regardless of cause, result in fewer data points being used to produce a VAD retrieval. Because each remaining point carries more weight in the retrieval when fewer data points are present, a data gap has the effect of increasing the influence of individual data points as well as any errors introduced by the remaining data points (e.g., Doppler velocity uncertainty). If the data coverage gaps are sufficiently extensive, noise can overwhelm the signal and prevent the technique from producing an accurate retrieval. Adding requirements regarding the number and distribution of data points that go into a retrieval is typically sufficient to reduce this error; however, certain data selection strategies can also isolate the impacts of errors due to data coverage gaps.

A major source of error in VAD retrievals is the error due to violating the underlying assumption that the horizontal wind field is linear and the Doppler vertical velocity is constant over the scan (hereafter referred to as the wind-field assumption). Small-scale features³ may violate the wind-field assumption, but they may only affect a small portion of the data points used in a retrieval. For these small features, the wind-field assumption violation errors can be reduced by including more data points in the retrieval, thus reducing the weight of the individual offending data points via spatial averaging. When the wind-field assumption violations are associated with larger features and contaminate a majority of the data points in a retrieval, however, a larger number of additional data points, located outside the offending feature, would be required to sufficiently reduce the associated error. In these large feature cases, a more practical solution is to limit the influence of the assumption violations to a minimal number of retrievals. This can be done by using a data selection strategy that has a small spatial footprint, thus limiting the number of retrievals that draw data from within the offending feature.

Both multi-Doppler (e.g., [Guimond et al. 2014](#); [Didlake et al. 2015](#)) and VAD retrievals assume that there are negligible temporal variations on time scales relevant to the retrievals. For either type of retrieval, the relevant time scale is determined by the difference between the times of the earliest and latest observation going into the retrieval (i.e., the temporal footprint). For VAD retrievals, the assumption of negligible temporal variation can be violated in two ways: either by variations in the aircraft orientation or by the evolution of the underlying wind field.

For high-altitude aircraft, on which most downward-pointing conically scanning radars are flown, the aircraft flight tracks typically include numerous long straight flight legs with sharp turns. Flying in the lower stratosphere, these aircraft do not typically encounter strong crosswinds or turbulence, so their

orientation during the straight flight legs is almost constant. For lower-altitude aircraft exposed to turbulence and crosswinds, the aircraft may need to change its orientation frequently as conditions change. These changes in aircraft orientation will introduce errors to the VAD either by changing the plane in which the winds are calculated⁴ or, in the case of a changing heading, the effective track-relative rotation rate of the antenna. While the present study will not attempt to characterize the errors due to changes in aircraft orientation, we will still consider the effects of the various scan strategies on minimizing the likelihood of errors due to changes in aircraft orientation. As aircraft orientation changes tend to be relatively isolated events, this study will assume that the likelihood of experiencing errors due to aircraft orientation changes is proportional to the length of the temporal footprint of each data selection strategy.

In addition to temporal variations in aircraft orientation, VAD retrieval errors can also be introduced by changes in the wind field over the duration of the temporal footprint (e.g., evolution of convective updrafts). Unlike the variations in aircraft orientation, these temporal variations impact data collected by aircraft at any altitude and are continuous. However, as the rate at which the temporal variations in wind field occur varies depending on the conditions, this study will not attempt to characterize these errors either. As with the aircraft orientation variations, the study instead will assume the likelihood of encountering nonnegligible temporal variations in the wind field to be proportional to the length of the temporal footprint.

The final VAD error source of interest to this paper is that of the spatial weighting function impacting the representativeness of the retrieval. The spatial weighting function of a VAD retrieval refers to the distribution of the input data points. For sequential single-scan retrievals, the data are distributed along the edge of the scan spatial footprint. This edge-heavy spatial weighting function can produce double images of small features, causing further issues in the presence of other error sources, particularly wind-field assumption violations. As will be demonstrated in [section 4](#), other data selection strategies contain data inside of the sequential single-scan footprint, producing a more representative spatial weighting function and reducing, or eliminating, the double image effect.

3. Evaluation data and error analysis methods

The four data selection strategies presented in this study will be evaluated using a combination of model data, analytic wind fields, and observed Doppler velocities. The model data and analytic wind fields will be sampled by either interpolating or calculating wind data to mimic the scan geometry of the NASA ER-2 X-band radar (EXRAD; [Li and Heymsfield 2009](#)) as flown on the NASA ER-2 aircraft during the NASA Integrated

³ What qualifies as a “small-scale feature” will depend on a number of factors, particularly the radar elevation angle and the range of the feature from the radar. Note, this means that a “small-scale feature” at a range of 15 km may not be considered to be small scale at a range of 5 km.

⁴ It is worth noting that aircraft typically have a slight upward pitch where the nose is slightly higher than the tail. As a result, the VAD-retrieved along-track winds will actually include a small contribution from the Doppler vertical velocity.

Precipitation and Hydrology Experiment (IPHEX) field campaign (Barros et al. 2014). For the model data analytic wind fields, the data will be sampled as though ER-2 is flying due east at a constant 206 m s^{-1} at 20-km altitude with a 2° upward pitch. The radar antenna will make one revolution every 4 s at an off-nadir angle of 25° and will make 70 observations per second (280 observations per scan) at 18.75-m range increments between the aircraft and the surface. This prescribed scan geometry results in the aircraft moving 824 m during each revolution of the radar antenna (i.e., $U_a \tau$).

Since this study will focus on retrievals of the horizontal winds and the total horizontal deformation, we designed an analytic wind field with nonzero quantities in both these fields. The linear analytic wind field represents a deformation zone within a southwesterly flow and is defined in units of meters per second by the following set of equations:

$$\begin{aligned} u_e &= (2.5 \times 10^{-5})x_e + (-2.5 \times 10^{-5})y_e + 10 \sin 45^\circ \quad \text{and} \\ v_e &= (-2.5 \times 10^{-5})x_e + (-2.5 \times 10^{-5})y_e + 10 \cos 45^\circ, \end{aligned} \quad (7)$$

where the subscript e indicates Earth-relative coordinates such that u_e and v_e are the zonal and meridional winds, respectively, and x_e and y_e are the zonal and meridional distances, respectively. These equations produce a wind field with $5 \times 10^{-5} \text{ s}^{-1}$ of both shearing and stretching deformation embedded in a 10 m s^{-1} southwesterly wind.

The model data used in this study are taken from the second hurricane nature run (HNR2; Nolan et al. 2013; Nolan and Mattocks 2014). HNR2 is a simulation of the life cycle of a tropical cyclone identified within a lower-resolution 13-month free-running global simulation. The data used in this study are available at 1-km horizontal grid spacing with 60 vertical levels between the surface and the model top at 20-km altitude. For the purpose of the present study, we only use a single time, 1930 UTC 21 August 2005, during which the tropical cyclone was still a disturbance with scattered deep convection embedded in a region of thick stratiform cloud.

To evaluate the performance of the various data selection strategies when using observed Doppler velocities, we will perform VAD retrievals on data collected during the NOAA Sensing Hazards with Operational Unmanned Technology (SHOUT; Dunion et al. 2018) field campaign. Specifically, the data were collected during a flight leg into Hurricane Matthew (2016) between 1306 and 1345 UTC 7 October 2016. Unlike the simulated retrievals that used the EXRAD instrument, the Doppler velocity observations were collected by the NASA High-Altitude Imaging Wind and Rain Airborne Profiler (HIWRAP; Li et al. 2008) radar, flown aboard the NASA Global Hawk. HIWRAP has two beams at 30° and 40° from nadir, each of which simultaneously operate at both Ku and Ka bands. The retrievals will be performed using the Ku-band inner (30°) beam masked to remove any data points with a reflectivity factor (hereafter reflectivity) less than 5 dBZ or aircraft roll greater than 3° . Furthermore, only those retrievals having valid data for at least 50% of the total number of data points for a given retrieval are considered in the error analysis.

The data selection strategies will be evaluated against one another by comparing VAD wind retrievals from each strategy to winds retrieved using the three-dimensional variational method of Guimond et al. (2014). This wind retrieval method works by minimizing a cost function to arrive at an estimated wind field that best matches the observed Doppler velocities with the added weak constraint of mass continuity. Including mass continuity as a weak constraint improves the retrieved wind field by reducing the effects of sampling and interpolation errors in the solution process. The accuracy of the retrievals is greatest in regions where the radar views a target volume from a large variety of angles. As a result, the total horizontal wind is most accurate just off nadir. Furthermore, the retrieved vertical and along-track velocities are most accurate at nadir while the across-track velocities are most accurate just off nadir. To minimize the impact of low angle diversity near nadir, the variational approach weights data within overlapping circles so as to draw information from data points just off nadir. Finally, near the edge of the swath, the view angles become nearly colinear and the retrieval accuracy can suffer from potential degradation. To mitigate issues with the variational retrieval accuracy and provide a consistent comparison point, the variational retrievals will be averaged over a disk with an off-nadir angular radius of 25° (5° less than the angular swath width). Figure 2 depicts the HIWRAP inner-beam Ku-band area-averaged reflectivity and area-averaged variational retrievals of horizontal and vertical wind speeds from the previously mentioned flight leg into Hurricane Matthew (2016).

4. Error analysis

a. Analysis of assumed platform position errors

As mentioned above, Tian et al. (2015) performed a detailed analysis of the errors incurred by assuming that the moving airborne instrument platform was stationary for the duration of the VAD retrieval data collection period (i.e., the temporal footprint). For their analysis, the authors of that study assumed the instrument platform remained stationary at the aircraft position at the start of the scan in order to simplify the error calculations. However, if the instrument platform is instead assumed to remain stationary at the aircraft position at the midpoint of the scan, the resulting error should be smaller as the distances between the assumed stationary point and the centers of curvature are shorter. This can be checked by rederiving the error equations of Tian et al. (2015) using different limits of integration. The Tian et al. (2015) limits of integration were 0 and 2π ; changing the limits of integration to $-\pi$ and π has the effect of changing the assumed stationary position to be located at the aircraft position at the midpoint of the scan.

Similar to Tian et al. (2015), the error due to assuming a stationary instrument platform can be computed by comparing the coefficients in Eq. (6) to the coefficients derived using a Fourier sine and cosine series. Following the process laid out in Tian et al. (2015), the Fourier-series-derived coefficients, under the assumption that the moving platform is stationary at

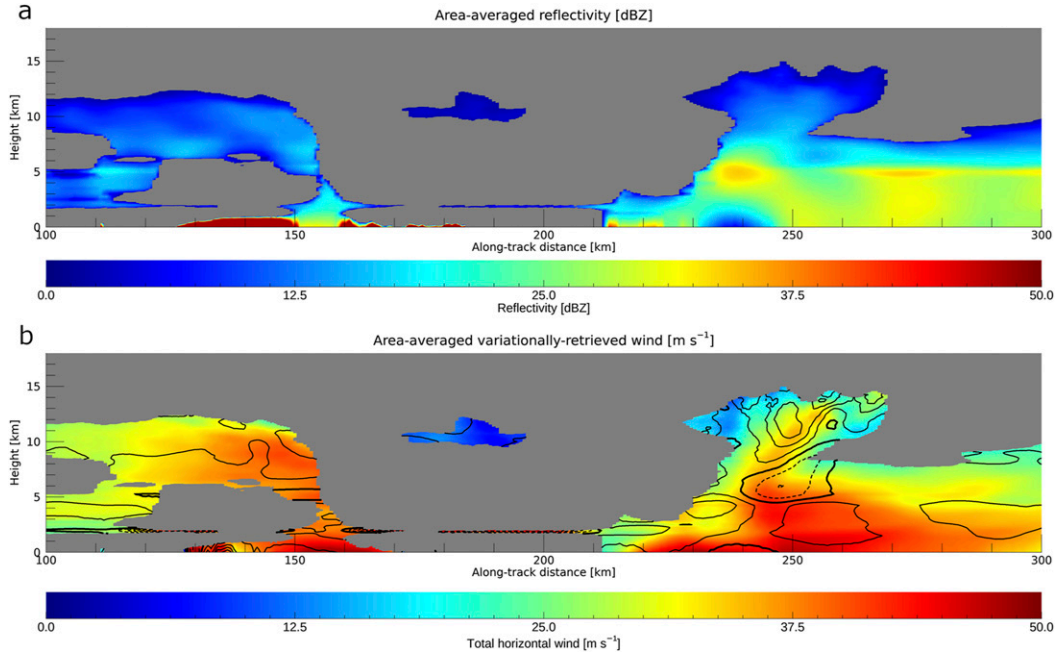


FIG. 2. Curtains of (a) area-averaged reflectivity and (b) area-averaged variably retrieved horizontal wind speed using HIWRAP data collected during the NOAA SHOUT field campaign flight leg into Hurricane Matthew (2016) between 1306 and 1345 UTC 7 Oct 2016. Area-averaged variably retrieved vertical wind speed is contoured in black in (b) every 2 m s^{-1} with solid contours indicating upward motion, dashed contours indicating downward motion, and the thick contour indicating zero vertical motion. The gray regions indicate areas where the area-averaged reflectivity was less than or equal to 0 dBZ. Averages are computed over a disk with an off-nadir angular radius of 25° centered on the aircraft position.

the scan midpoint, are derived by evaluating the following equations:

$$\begin{aligned} C_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} V_r d\theta, & C_1 &= \frac{1}{\pi} \int_{-\pi}^{\pi} V_r \cos\theta d\theta, \\ C_2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} V_r \sin\theta d\theta, & D_1 &= \frac{1}{\pi} \int_{-\pi}^{\pi} V_r \cos 2\theta d\theta, \quad \text{and} \\ D_2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} V_r \sin 2\theta d\theta. \end{aligned} \quad (8)$$

For the purposes of the present study, we will focus on the horizontal wind and deformation errors. Evaluating the relevant integrals, we find

$$\begin{aligned} C_1 &= v_0 \cos\Phi - \cos\Phi \frac{U_a \tau}{2\pi} \left(\frac{1}{2} u_y \right), \\ C_2 &= u_0 \cos\Phi - \cos\Phi \frac{U_a \tau}{2\pi} \left(\frac{1}{2} v_y \right), \\ D_1 &= \frac{r \cos^2\Phi}{2} S_1 - \cos\Phi \frac{U_a \tau}{3\pi} u_y, \quad \text{and} \\ D_2 &= -\frac{r \cos^2\Phi}{2} S_2 + \cos\Phi \frac{2U_a \tau}{3\pi} v_y, \end{aligned} \quad (9)$$

where S_1 and S_2 are the stretching and shearing deformations, respectively, and are defined in track-relative coordinates as

$$S_1 = \frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \quad \text{and} \quad S_2 = -\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}. \quad (10)$$

Equation (6) defines the coefficients for a VAD analysis using radar observations taken from a moving platform while the Fourier-derived coefficients in Eq. (9) are the coefficients that result from assuming the moving platform is stationary at the scan-midpoint aircraft position. The errors that result from assuming that the moving platform is stationary at the midpoint of the scan can be determined by subtracting Eq. (9) from Eq. (6). For this, the variables that originate from Eq. (6) will be denoted with an asterisk superscript. This gives us the following errors:

$$\begin{aligned} v_0^* - v_0 &= -\frac{U_a \tau}{2\pi} \left(\frac{1}{2} u_y \right), & u_0^* - u_0 &= -\frac{U_a \tau}{2\pi} \left(\frac{1}{2} v_y \right), \\ S_1^* - S_1 &= -\frac{U_a \tau}{2\pi} \left(\frac{4}{3} r \cos\Phi \right), & \text{and} \\ S_2^* - S_2 &= -\frac{U_a \tau}{2\pi} \left(\frac{8}{3} r \cos\Phi \right). \end{aligned} \quad (11)$$

Interpreting Eq. (11), we find that the errors introduced by the stationary platform assumption will be independent of the data selection strategy when the underlying wind field is linear and the Doppler vertical velocity is constant. When these wind field assumptions are not satisfied, the stationary platform

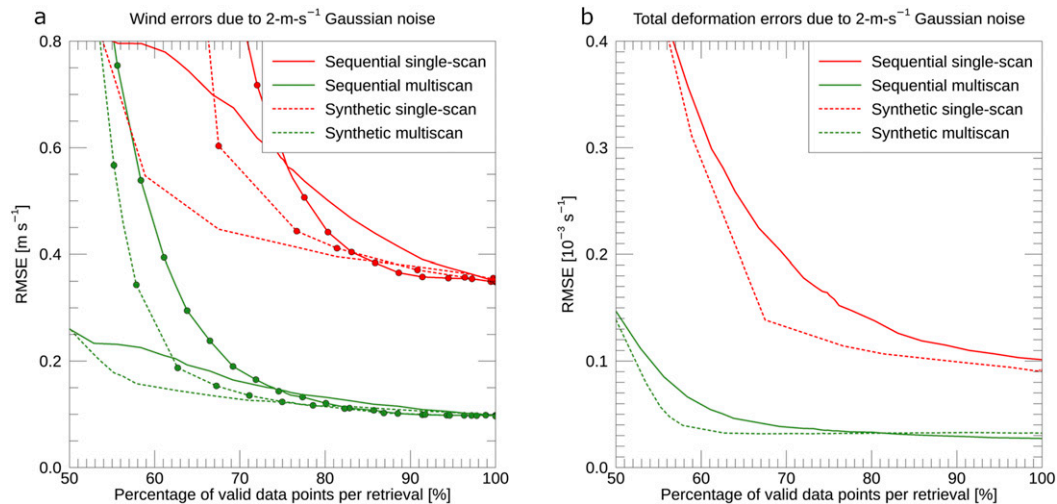


FIG. 3. Root-mean-square error (RMSE) of the VAD-retrieved (a) wind components and (b) total deformation for a range of data coverages for each data selection strategy. In (a), the across-track wind RMSEs are indicated by the lines with circle symbols and the along-track wind RMSEs are indicated by the lines without circle symbols. Doppler velocity uncertainty is generated by introducing 2 m s^{-1} of Gaussian noise into 10 000 retrievals of a linear deformation zone analytic wind field with $5 \times 10^{-5} \text{ s}^{-1}$ of both shearing and stretching deformation and 10 m s^{-1} southwesterly wind.

assumption errors will differ between data selection strategies based on the distribution of data points. Because the stationary platform assumption errors become dependent on the choice of data selection strategy only when the wind-field assumptions are violated, the resulting error differences between strategies will be handled as wind-field assumption violation errors rather than stationary platform assumption errors. As such, the stationary platform assumption error will be treated as being independent of the choice of data selection strategy.

Comparing Eq. (11) to the error equations of Tian et al. (2015), their Eqs. (9)–(12), reveals that moving the assumed stationary position from the start of the scan to the scan midpoint reduces the magnitude of the wind errors while the magnitude of the deformation errors remains the same. The magnitude of the v_0 and u_0 wind error differences⁵ are $U_a \tau (v_y/2)$ and $U_a \tau (u_y/2)$, respectively, with the midscan stationary position errors being smaller. While the differences will be relatively small for scanning geometries where $U_a \tau$ is small (i.e., a slow-moving aircraft or a fast scan speed), those scan geometries with a larger $U_a \tau$ could see considerable wind error reduction simply by assigning the retrieved wind values to the scan midpoint rather than the start of the scan.

Theoretically, if the along-track gradients of the horizontal wind components v_y and u_y are known, then the error due to the stationary platform assumption could be removed. Although these wind gradients can be computed using consecutive VAD profiles, noise within the gradient calculations introduces additional errors into the VAD wind retrievals in place of the errors that are being removed. As such, the error

due to the stationary platform assumption is best minimized by using the aircraft position at the scan midpoint as the stationary location. Therefore, for the remainder of this paper, all calculations will be performed with the assumed stationary platform location positioned at the scan midpoint.

b. Analysis of strategy-dependent errors

The remaining five error sources (Doppler velocity uncertainty, data gaps, wind field assumption violations, temporal variations, and spatial weighting function representativeness) are all dependent on the data selection strategy employed for the VAD retrieval. This section will discuss how these error sources are manifest in retrievals performed using each of the four data selection strategies.

The influence of Doppler velocity uncertainty on the VAD retrieval is largely dependent on the number of valid data points being ingested into the retrieval. The small number of data points ingested into the single-scan retrievals relative to the multiscan retrievals can cause minor issues in terms of Doppler velocity uncertainty reduction. These minor issues can become problematic when noisy data are combined with gaps in data coverage. Figure 3 depicts the root-mean-square error (RMSE) of noisy VAD retrievals using the various data selection strategies as a function of data coverage, expressed as the percentage of valid data points per retrieval. The uncertainty is added by introducing 2 m s^{-1} of Gaussian noise⁶ to each retrieval of the analytic wind field described by Eq. (7) and repeating the process 10 000 times; recall, 2 m s^{-1} of noise is equal to the engineering goal for accuracy on many radars,

⁵ Note that there appears to be a sign error in Eq. (10) of Tian et al. (2015).

⁶ The distribution of 2 m s^{-1} of Gaussian noise has a mean of $\sim 0 \text{ m s}^{-1}$ and a standard deviation of $\sim 2 \text{ m s}^{-1}$.

including EXRAD and HIWRAP, and is intended as a worst-case upper bound on the noisiness of the data. The data coverage is prescribed by setting a sector of data to missing values for each data selection strategy in order to mimic the loss of data coverage introduced by clear-air gaps or excessive attenuation.

For retrievals performed with full data coverage, the uncertainty errors for the sequential and synthetic single-scan strategies tend to be approximately equal to one another; likewise, the uncertainty errors for the sequential and synthetic multiscan strategies also tend to be approximately equal to one another (Fig. 3). As the data coverage decreases, the RMSE due to Doppler velocity uncertainty increases, albeit at different rates for each strategy. Not surprisingly, the multiscan strategies reliably outperform the single-scan strategies at all data coverage levels. Furthermore, the RMSEs for the multiscan strategies, particularly the synthetic multiscan strategy, remain relatively low for lower data coverages compared to the RMSEs for the single-scan strategies. It is important to note that the presence of data coverage gaps amplifies the effects of most error sources, not just errors due to Doppler velocity uncertainty. Likewise, the introduction of additional data points from outside of error-inducing features can reduce the influence of error sources, particularly when the offending feature is relatively small compared to the spatial footprint.

While the multiscan data selection strategies are effective at reducing the RMSE introduced by Doppler velocity uncertainty, even in the presence of data gaps, the synthetic strategies provide a benefit not captured by the RMSE analysis in Fig. 3. Specifically, the smaller spatial footprint of the synthetic strategies, particularly the synthetic single-scan strategy, acts to isolate errors to a minimal number of retrievals. This is schematically depicted in Fig. 4 for VAD retrievals in the presence of clear-air gaps using a sequential single-scan strategy (Fig. 4a) and a synthetic single-scan strategy (Fig. 4b). In this schematic, valid data are represented by the green curves and missing data are represented by the dashed red curves. The locations of retrievals using this data are indicated by the green and red circles: the green circles indicate retrievals where all data points were valid (hereafter “green retrievals”) while the red circles indicate retrievals that had missing data due to a clear-air gap (hereafter “red retrievals”). As a result of the smaller spatial footprint, four out of nine of the synthetic single-scan retrievals are uncontaminated by missing data while only two out of nine of the sequential single-scan retrievals escape the influence of missing data. However, there is a trade-off: the red synthetic retrievals contain fewer valid data points than the corresponding red sequential retrievals. As such, the red synthetic retrievals will likely produce larger errors than the matching sequential retrievals.

The error isolation effect of the synthetic strategies, depicted in Fig. 4, may be desirable for data assimilation purposes. It is expected that the error-prone red retrievals in this example would be removed before the data assimilation system ingests the data by screening the VAD retrievals for data coverage. In instances where such screening is not performed, however, the data assimilation system should be more likely to disregard error-prone red retrievals that disagree too strongly with the

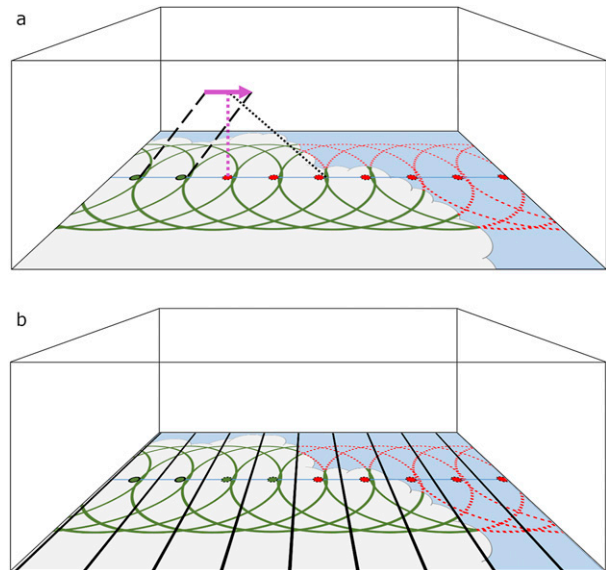


FIG. 4. Schematic example of data gaps due to clear-air regions impacting (a) sequential single-scan and (b) synthetic single-scan retrievals. The blue areas represent clear-air regions while the gray areas represent regions of scattering particles (e.g., cloud). Dotted red curves indicate locations of Doppler velocity data gaps due to clear air while the broken-bordered red dots indicate retrievals with missing data due to the clear-air data gaps. In contrast, green curves indicate locations of Doppler velocity data and green dots indicate retrievals that do not ingest any missing data points caused by clear-air data gaps. The magenta arrow in (a) indicates the aircraft flight path during a single scan, centered over the dot indicated by the dotted magenta line. The first and last radar beam positions of the scan are indicated by the black dashed lines while the radar beam position at the scan midpoint is indicated by the black dotted line. The solid black lines in (b) denote the divisions between subsequent synthetic single-scan data regions.

first guess model fields than the green retrievals that have smaller errors and, hopefully, greater agreement with the first guess fields. Furthermore, because the red synthetic retrievals are likely to have larger errors than the red sequential retrievals, the data assimilation system should, in theory, be more likely to discard the red synthetic retrievals than the red sequential retrievals. While Fig. 4 uses clear-air gaps to demonstrate the error isolation effect of the synthetic strategies, this effect is equally relevant to other error sources.

Although errors due to temporal variations in the underlying wind field will typically be negligibly small in regions where the wind-field assumptions are approximately satisfied, temporal variations in aircraft orientation can occur in any conditions and are potentially nonnegligible. As mentioned in section 2c, we consider the likelihood of a retrieval encountering a temporal variation error to be proportional to the temporal footprint of the retrieval strategy. Figure 1b depicts the temporal distributions of data points ingested into the VAD algorithm under each strategy for an EXRAD-like scan geometry. The sequential single-scan strategy has the shortest temporal footprint, being equal to the revolution period of the radar antenna, while the synthetic multiscan strategy has the

longest temporal footprint. As such, VAD retrievals using the sequential single-scan strategy (i.e., a tradition VAD retrieval) will be least likely to encounter temporal variation errors while those using a synthetic multiscan strategy will be most likely to encounter these errors. That said, data collected during changes in aircraft orientation can easily be removed via thresholding: Tian et al. (2015) removed data where the aircraft roll was greater than 3° . Furthermore, as previously mentioned, temporal variations in the underlying wind field are typically small unless the wind-field assumption is strongly violated.

Because each data selection strategy ingests data from a different region of the data swath, the representativeness of the retrievals will also vary. The pattern of data ingested under each strategy can be considered as a type of spatial weighting function. As the spatial weighting function is difficult to visualize for a two-dimensional pattern being applied to wind fields, we will instead use curtains of radar reflectivity averaged over the data points that would be ingested into a VAD retrieval. Figure 5 depicts the mean reflectivity produced when observing a 5-km-tall, 10-km-diameter cylindrical pillar of 35-dBZ reflectivity, ignoring the effects of beamwidth and attenuation.

While a nadir beam (Fig. 5a) sees the pillar exactly as it exists, the sequential single-scan strategy mean reflectivity (Fig. 5b) produces two weaker pillars and an area of zero reflectivity at the location of the actual pillar. This erroneous duplication is caused by the following sequence of events: first, the pillar is sensed by the forward portion of the scan, at which point the radar records the actual pillar as the left scan-mean pillar; then the pillar falls completely within the scan cone, at which point the beam never intersects the pillar and so zero reflectivity is recorded; and, finally, the actual pillar is sensed by the rearward portion of the scan, producing the right scan-mean pillar.

While the sequential multiscan strategy produces a single pillar of mean reflectivity (Fig. 5c), the resulting mean pillar is considerably wider than the actual pillar and there is still a local minimum in mean reflectivity at the center (Fig. 5f) due to the strategy ingesting fewer data points at the center of the pattern. Examination of Fig. 1a reveals that the 13 consecutive antenna revolutions that are ingested into a sequential multiscan retrieval do not include all data at the center of the spatial footprint. To include all scans that contain relevant data, the scan-selection window would need to be equal to the horizontal diameter of the radar scan circle (recall, the examples here use all sequential single scans within a 10-km window to produce the sequential multiscan strategy). For the scan geometry used in the examples here, the scan-selection window would need to be ~ 45 -km long to include all relevant data for the near-surface sequential multiscan retrieval and the resulting spatial footprint would be ~ 90 -km long in the along-track direction (the size of the scan-selection window plus twice the radius of the scan). The spatial footprint that results from using such a long scan-selection window is too large for many applications within the convective regions in which these methods are being evaluated, hence a more practically sized scan-selection window must be used at the cost of a local minimum in the spatial weighting function.

The issue of low data density at the center of the spatial footprints for the sequential strategies is absent in both the synthetic single-scan strategy (Fig. 5d) and the synthetic multiscan strategy (Fig. 5e). The spatial weighting function of both synthetic strategies peaks at the center of the spatial footprint (Fig. 5f). As such, the synthetic retrievals will be more representative of the underlying wind field, ignoring the effects of other error sources. This is particularly the case for the synthetic single-scan strategy, which almost exactly replicates the horizontal extent of the reflectivity pillar. For the synthetic strategies, the difference between the width of the averaged pillar and the actual pillar will be equal to the along-track length of the spatial footprint, with half the width added on each side of the pillar. Note, all strategies produce pillars whose mean reflectivity increase with altitude due to more data being drawn from outside the pillar at lower altitudes. This is due to the pillar having a smaller diameter than the across-track dimension of the spatial footprint, which grows wider as the range from the radar increases.

The final error source to be discussed in this paper, violations of the underlying wind-field assumptions, typically produces the largest VAD retrieval errors. The wind-field assumption violation errors will be evaluated using the HNR2 wind fields (Fig. 6). By averaging the HNR2 winds over each data selection strategy before computing the errors, the effects of the differing spatial weighting functions are removed from the error analysis. Furthermore, the wind fields are sampled as though the aircraft was flying in a straight line and using a single model time for all the data, eliminating errors due to temporal variation. The only error sources present in the analysis are the errors due to the wind-field assumption violations, the assumption of a stationary instrument platform, and a small error introduced by the 2° upward pitch of the aircraft.⁷ Since the wind-field assumption violation error is the only error source that is strategy dependent, differences in the errors produced under each data selection strategy can be fully attributed to violations of the wind-field assumptions.

Examination of Fig. 6 reveals several general patterns that have already been alluded to above. The multiscan strategies tend to smooth out errors while the synthetic strategies tend to isolate errors. The wind-field assumption violation errors for the synthetic multiscan strategy benefit from both the smoothing and isolating effects: the errors appear to be smoothed out in comparison to the sequential single-scan retrieval errors while still retaining a similar spatial extent as the sequential single-scan retrieval errors. Also evidenced in Fig. 6 is that the largest errors, regardless of strategy, are found in and near updrafts (contoured). While Fig. 6 only depicts the errors in horizontal wind speed, the errors in the horizontal deformation retrievals display similar patterns and, as such, are not shown.

Examining cumulative frequency histograms of the percent absolute errors of both the horizontal wind speed and

⁷ The pitch introduces an error when the VAD retrieval is assumed to be in the horizontal plane instead of in the plane of the aircraft.

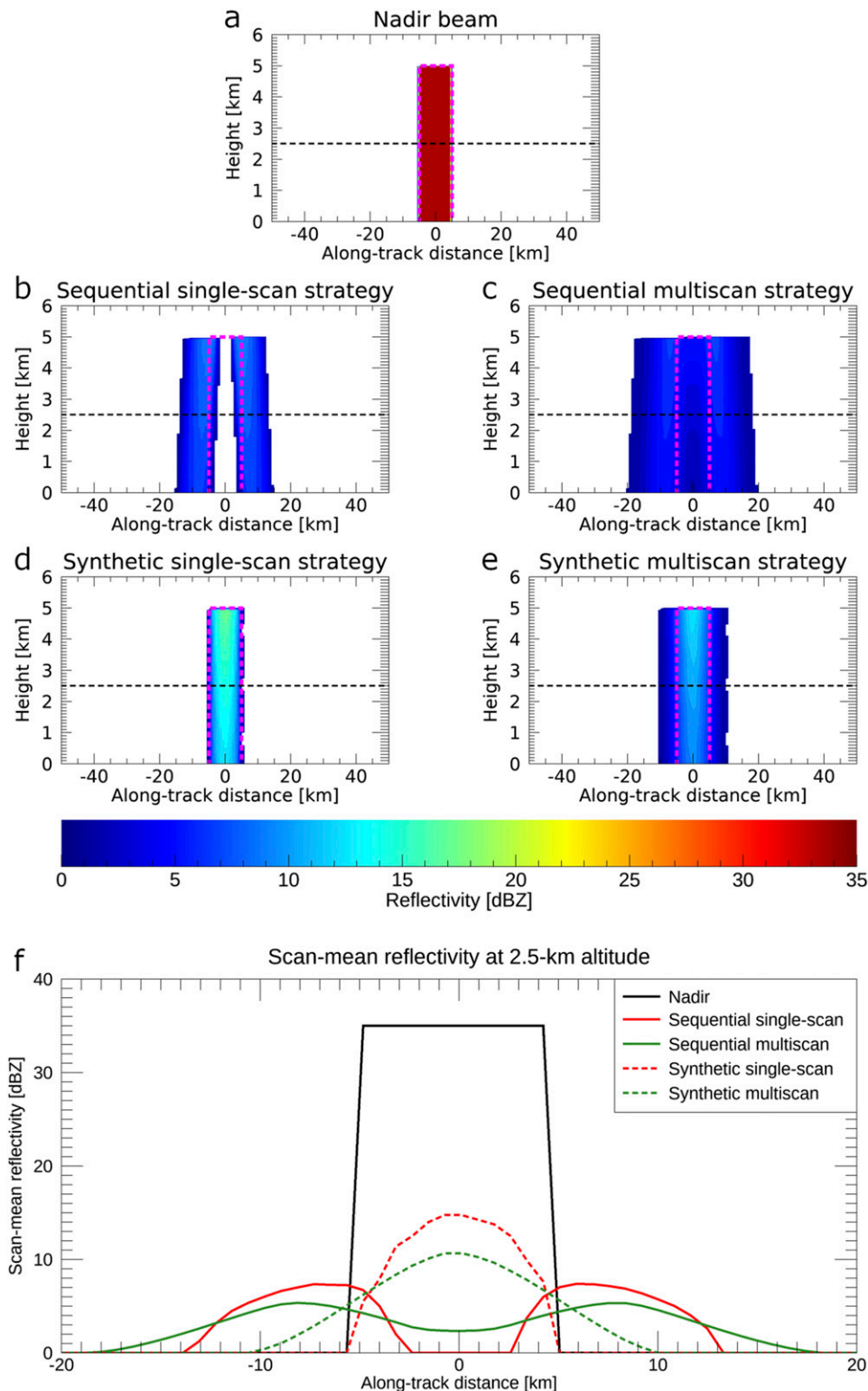


FIG. 5. Demonstration of the spatial weighting functions of the data selection strategies based on a 35-dBZ reflectivity pillar (pink dashed outline). (a) The reflectivity at nadir. (b)–(e) The scan-mean reflectivity resulting from the application of the (b) sequential single-scan, (c) sequential multiscan, (d) synthetic single-scan, and (e) synthetic multiscan data selection strategies. (f) The reflectivity at 2.5-km altitude for each of (a)–(e). The black dashed lines in (a)–(e) indicate the 2.5-km altitude level. For these plots, the reflectivity is taken as the value at the center of the pulse volume.

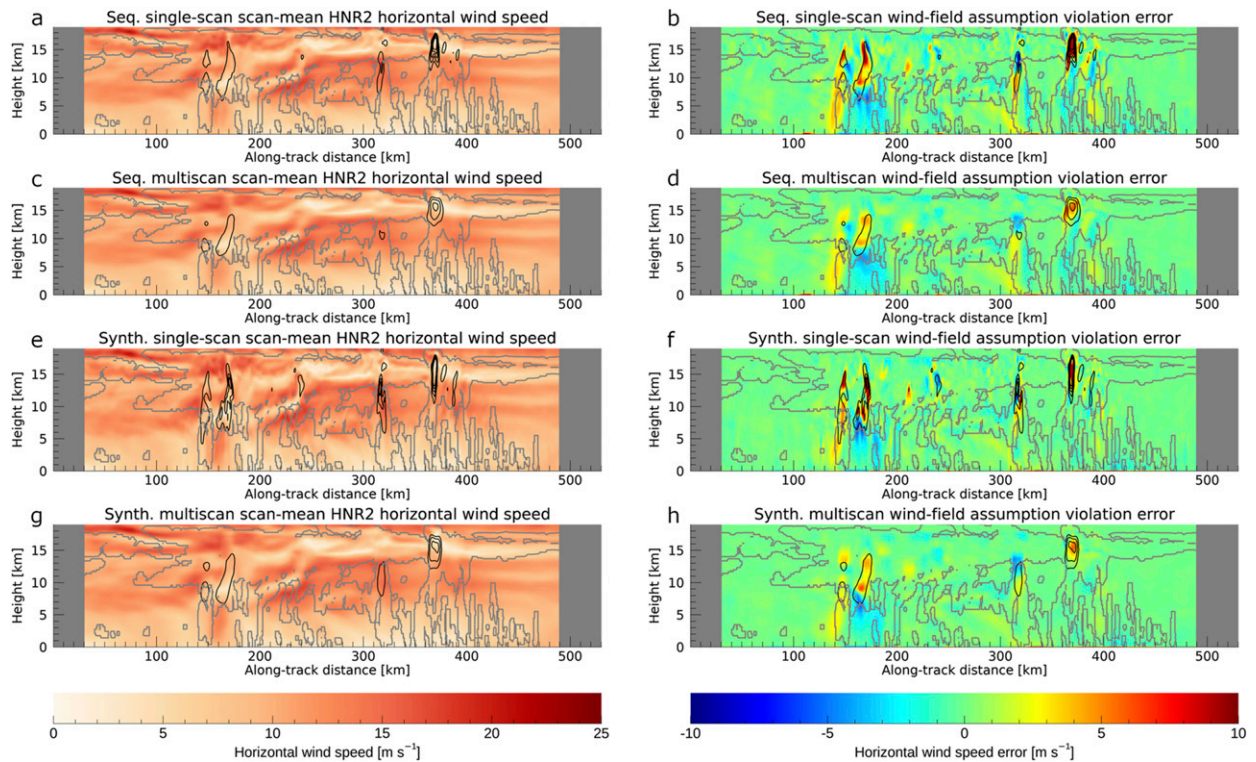


FIG. 6. Curtains of (a),(c),(e),(g) HNR2 total horizontal wind speed and (b),(d),(f),(h) VAD-retrieved total horizontal wind speed error with respect to the HNR2 winds computed using (a),(b) a sequential single-scan strategy, (c),(d) a sequential multiscan strategy, (e),(f) a synthetic single-scan strategy, and (g),(h) a synthetic multiscan strategy. HNR2 total horizontal wind speeds are computed by averaging over each data selection strategy so as to eliminate the influence of the spatial weighting function on the error analysis depicted in this figure. HNR2 updraft speeds are contoured in black at 2 m s^{-1} intervals starting at 2 m s^{-1} and the HNR2 cloud extent is contoured in gray. Note that HNR2-derived scan data are taken as the value at the center of the pulse volume.

horizontal deformation retrievals (Fig. 7) allows for a more direct intercomparison of the four data selection strategies. Based on the cumulative frequency of horizontal wind speed errors, the synthetic multiscan strategy performs best in terms of error reduction, followed by the synthetic single-scan, then sequential multiscan, and finally the sequential single-scan strategies. Overall, the horizontal wind retrieval errors due to wind-field assumption violations are relatively small: less than

~20% of data points experience horizontal wind speed errors greater than 10% of the actual horizontal wind speed. For synthetic multiscan retrieval, this improves to only ~10% of data points experiencing errors greater than 10%.

While the errors introduced to the horizontal deformation retrievals by violations of the wind-field assumptions may follow a similar pattern to that of the horizontal wind speed errors (Fig. 6), the horizontal deformation retrievals

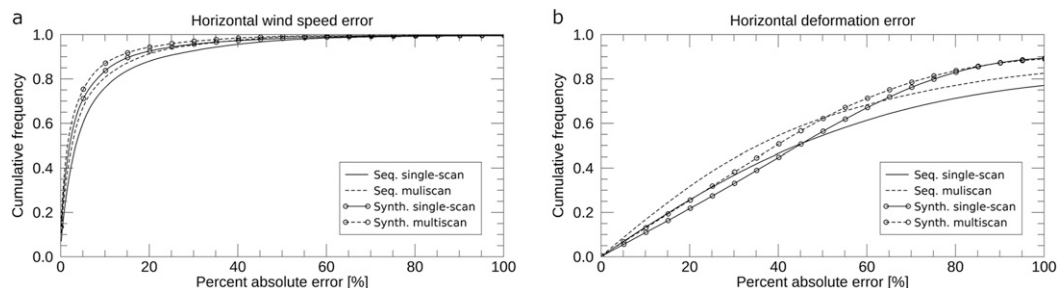


FIG. 7. Cumulative frequency diagrams for (a) horizontal wind speed and (b) horizontal deformation percent absolute errors due solely to violations of the wind field assumptions. Errors are computed as in Fig. 6, except that only errors associated with HNR2 horizontal wind speeds greater than 0.1 m s^{-1} and HNR2 horizontal deformations greater than 10^{-4} s^{-1} are considered.

TABLE 1. Median absolute errors (mean absolute error in parentheses) for the along-track wind component (AVEL), across-track wind component (XVEL), total horizontal wind (HVEL), and total horizontal deformation (DEF) for each data selection strategy for the HNR2 wind retrievals depicted in Fig. 6. Wind errors are in units of m s^{-1} and deformation errors are in units of 10^{-3} s^{-1} .

Data selection strategy	AVEL	XVEL	HVEL	DEF
Sequential single scan	0.36 (0.75)	0.35 (0.79)	0.37 (0.79)	0.17 (0.38)
Sequential multiscan	0.23 (0.43)	0.29 (0.65)	0.27 (0.55)	0.13 (0.23)
Synthetic single scan	0.12 (0.26)	0.34 (0.83)	0.22 (0.54)	0.16 (0.29)
Synthetic multiscan	0.08 (0.19)	0.24 (0.61)	0.16 (0.40)	0.10 (0.19)

performed much poorer (Fig. 7b). Over 80% of retrievals had errors greater than 10% of the actual horizontal deformation and over 50% of retrievals had an error of greater than 50% of the actual horizontal deformation. One potential reason that the error performance of the horizontal deformation retrievals is so much worse than that of the horizontal wind retrievals is that any appreciable changes in horizontal deformation will, by definition, violate the assumption that the underlying wind field is linear.

Further support for these insights regarding the retrieval errors due to wind-field violations can be found in the median and mean absolute errors (Table 1). In particular, the error performance for horizontal wind speed retrievals, discussed above, matches the median and mean absolute errors. The data selection strategies, ordered from smallest to largest median and mean absolute errors, are as follows: synthetic multiscan, synthetic single scan, sequential multiscan, and sequential single scan. This order of performance also holds when considering errors in the retrieval of along-track winds. Where the order of error performance does not hold is in the across-track wind and total horizontal deformation retrievals: for these retrievals, the multiscan strategies outperform the single-scan strategies, but the synthetic strategies do not clearly outperform the sequential strategies. This is likely due to the synthetic strategies only reducing the spatial footprint in the along-track direction while the across-track dimension of the spatial footprint remains unchanged, a conclusion also reached by Kelberlau and Mann (2020) in their analysis of VAD retrievals of turbulence using their “squeezing” method (Kelberlau and Mann 2019), which, as previously mentioned, is similar to the synthetic single-scan strategy discussed here.

Figure 8 depicts the difference between the VAD total horizontal wind retrievals, using each of the four data selection strategies, and the variationally retrieved total horizontal wind retrievals, which are averaged over a 25° -angular-radius circle, for the flight leg into Hurricane Matthew (2016) that is depicted in Fig. 2. Note, as the variational retrieval suffers from its own errors, the variationally retrieved winds are not technically a “ground truth” but instead serve as a common point of comparison that approximates the actual underlying wind field. As such, the differences should be considered in a relative sense rather than in an absolute sense. Note, all error sources discussed in this paper are present to some extent in the VAD retrievals used to produce Fig. 8.

The most striking feature depicted in the difference fields of Fig. 8 is the large difference associated with a convective burst embedded in the hurricane eyewall, near 235-km along-track

distance. The differences from the variational retrieval within the burst region tend to be an order of magnitude larger inside the burst region than outside the region (Table 2) for all data selection strategies. Interestingly, the sequential multiscan strategy has the lowest mean absolute difference within the burst region while the synthetic single-scan strategy has the lowest median difference within that region. The low median and large mean difference of the synthetic single-scan strategy is likely a result of the error isolation effect, where the error sources have a larger impact on a smaller number of retrievals.

Overall, there is some resemblance between the differences from the variational retrievals depicted in Fig. 8 and the HNR2-based error analysis depicted in Fig. 6. The multiscan strategies tend to smooth out the differences/errors and the sequential strategies tend to isolate the differences/errors. Interestingly, the synthetic single-scan strategy has the worst performance in terms of difference from the variational retrievals both overall and in the nonburst region according to Table 2. Examining the difference curtains (Fig. 8), the synthetic single-scan strategy contains peaks in difference in regions where the sequential single-scan strategy experiences minimal differences from the variational retrievals. It is possible that the larger spatial footprint of the sequential single-scan strategy is smoothing out some of errors despite its low data density. Regardless, the synthetic multiscan strategy provides the best performance overall and when limited to the nonburst region, although the sequential multiscan strategy performs almost as well in terms of median and mean absolute difference.

5. Conclusions

Four data selection strategies for selecting airborne Doppler radar data to be ingested into a VAD retrieval are presented here. The data selection strategies are characterized by two pairs of traits: sequential or synthetic and single scan or multiscan. The sequential–synthetic trait describes whether the data are selected using a temporal window (sequential) or an along-track spatial window (synthetic). The single-scan–multiscan trait describes the amount of data selected by the strategy: a single-scan strategy includes a number of data points equal to the number of data points collected during a single revolution of the instrument antenna/receiver while a multiscan strategy includes more data points than are collected during a single revolution of the instrument antenna/receiver. Pairing these traits results in the four data selection strategies: a

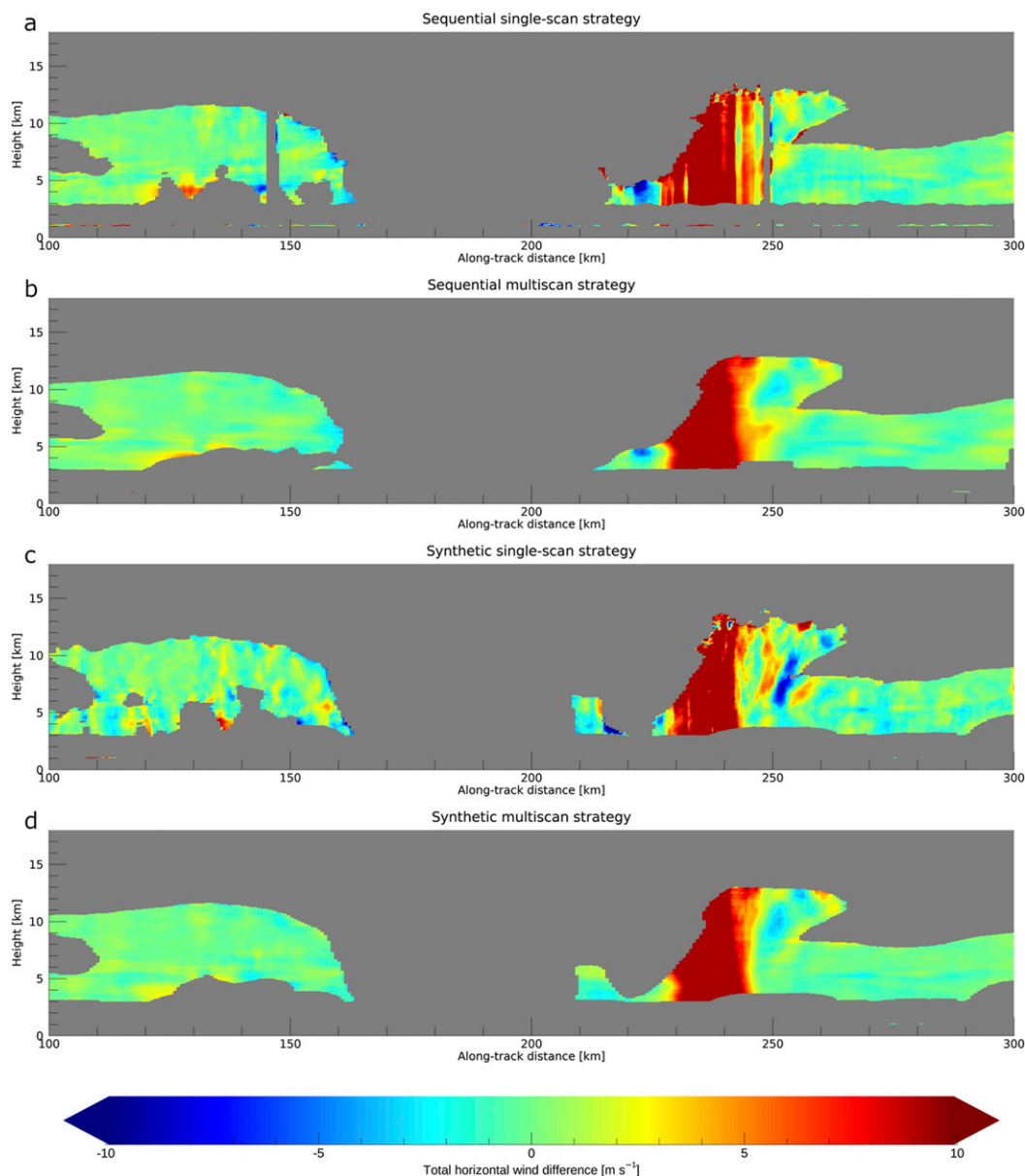


FIG. 8. Curtains of total horizontal wind difference between the VAD wind retrievals and area-averaged variational wind retrievals (Fig. 2b) using HIWRAP data for the (a) sequential single-scan, (b) sequential multiscan, (c) synthetic single-scan, and (d) synthetic multiscan data selection strategies. The HIWRAP data were collected during the NOAA SHOUT field campaign flight into Hurricane Matthew (2016) between 1306 and 1345 UTC 7 Oct 2016. Note that low-level data have been removed to eliminate contamination by the main-beam and sidelobe surface returns. Additionally, it should be noted that differences with a magnitude greater than 10 m s^{-1} exist; however, these larger differences have not been included in the color scale as these points almost certainly correspond to instances of extreme VAD error and are likely to be excluded from a careful analysis or during data assimilation.

sequential single-scan, sequential multiscan, synthetic single-scan, and synthetic multiscan strategy.

The four data selection strategies are evaluated with respect to their responses to errors introduced by a number of sources. The errors evaluated by this study are associated with six sources: the assumed stationary position of the moving

instrument platform, the Doppler velocity uncertainty, the gaps in the data coverage, the assumption violations regarding the underlying wind field required for the VAD retrievals, the temporal variations, and the representativeness of the spatial weighting function. The results of the error analysis are summarized as follows:

TABLE 2. Total horizontal wind median absolute differences (mean absolute differences in parentheses) between the VAD retrievals and the variational retrievals of the Hurricane Matthew (2016) flight leg depicted in Fig. 8. The burst differences include all differences within the region of the convective burst (220–250-km along-track distance) and nonburst differences include all differences outside the region of the convective burst.

Data selection strategy	All	Burst	Nonburst
Sequential single scan	0.74 (2.45)	9.14 (12.85)	0.63 (0.99)
Sequential multiscan	0.58 (1.86)	8.65 (10.12)	0.49 (0.67)
Synthetic single scan	0.99 (2.68)	7.81 (11.36)	0.85 (1.41)
Synthetic multiscan	0.48 (1.85)	9.39 (10.48)	0.41 (0.61)

- assuming the instrument platform is stationary at the mid-scan aircraft position minimizes related errors;
- sequential strategies tend to have edge-heavy spatial weighting functions and low data density near the center of their spatial footprints;
- synthetic strategies have high data density and have an error isolation effect, where the errors affect fewer retrievals but have a larger magnitude;
- single-scan strategies tend to be error prone due to having a small number of data points;
- multiscan strategies tend to smooth out errors due to their larger number of data points and larger spatial footprint;
- the sequential single-scan strategy has the smallest temporal footprint, but is error prone due to poor representativeness and a small number of data points;
- the synthetic single-scan strategy isolates error-contaminated data points to a minimal number of retrievals, but the remaining errors tend to be larger;
- the sequential multiscan strategy has smaller errors than the single-scan strategies, but has a considerably larger spatial footprint than the sequential single-scan strategy;
- the synthetic multiscan strategy slightly outperforms the sequential multiscan strategy in terms of minimizing errors and has a smaller spatial footprint than the sequential multiscan strategy;
- the deformation errors tend to be relatively large compared to typical deformation magnitudes; and
- the horizontal wind errors outside of updrafts tend to be relatively small compared to their typical magnitudes.

Based on the error analysis, a synthetic multiscan data selection strategy should be used when possible. Furthermore, the data selection window for the synthetic multiscan retrievals should be as large as possible without filtering out the features of interest or introducing error-inducing features (e.g., convective updrafts) into an excessive number of the retrievals. In cases where error-inducing features are closely spaced in the along-track direction or the features of interest are too small for a multiscan approach, a synthetic single-scan approach should be the preferred data selection strategy.

Although VAD retrievals are not able to capture the full variability of the observed wind field in three dimensions, the technique has several key applications. Performing VAD retrievals is exceedingly computationally efficient; hundreds of

thousands of VAD retrievals can be computed in the span of a few minutes. As such, the VAD technique is ideal for real-time applications such as situational awareness during airborne field campaign operations. The VAD technique also synthesizes large amounts of data into a single value, which has a number of important applications. Building on the previous example of real-time situational awareness, performing the VAD retrievals on board the aircraft means that scientists on the ground can monitor the environment using only a tiny fraction of the data bandwidth that would be required to transmit even a fraction of the radar data to the ground in real time.

Of more interest to scientific research though is the potential data assimilation applications for numerical weather prediction. Studies have shown that assimilating observations that have a higher density than the model grid spacing can degrade the simulation accuracy (e.g., Liu and Rabier 2002; Bick et al. 2016). The ability of the VAD technique to quickly summarize large quantities of high-resolution data is ideal for reducing the observation density, producing so-called superobservations. The multiscan data selection strategies presented here are ideal for data assimilation due to the feature smoothing effect reducing high-frequency variability that may cause issues for model analysis accuracy.

Acknowledgments. The authors would like to acknowledge the contributions of the late Lin Tian as her VAD retrieval code formed the basis of the sequential single-scan strategy discussed in this paper. Dr. Helms's work was supported by an appointment to the NASA Postdoctoral Program at NASA Goddard Space Flight Center, administered by Universities Space Research Association under contract with NASA. Dr. Guimond's contribution was supported by the NASA Weather and Atmospheric Dynamics program (Grant NNN16ZDA001N-WEATHER). The collection and processing of the underlying radar data used in this study was funded by the NOAA SHOUT program. Finally, the authors thank David Nolan and his colleagues for providing the HNR2 data set and the two anonymous reviewers for their efforts in improving this manuscript during the peer review process.

Data availability statement. All underlying radar data used in this study are available via the NASA High-Altitude Radar Group website at <https://har.gsfc.nasa.gov>. HNR2 data can be obtained by contacting David Nolan, whose contact information can be found in Nolan et al. (2013) and Nolan and Mattocks (2014). The processed VAD data can be obtained by contacting Dr. Helms and the processed 3DVAR radar wind data can be obtained by contacting Dr. Guimond.

REFERENCES

- Barros, A. P., and Coauthors, 2014: NASA GPM—Ground validation: Integrated Precipitation and Hydrology Experiment 2014. NASA Tech. Rep., 64 pp., <https://doi.org/10.7924/G8CC0XMR>.
- Bick, T., and Coauthors, 2016: Assimilation of 3D radar reflectivities with an ensemble Kalman filter on the convective scale. *Quart. J. Roy. Meteor. Soc.*, **142**, 1490–1504, <https://doi.org/10.1002/qj.2751>.
- Browning, K. A., and R. Wexler, 1968: The determination of kinematic properties of a wind field using Doppler radar.

- J. Appl. Meteor. Climatol.*, **7**, 105–113, [https://doi.org/10.1175/1520-0450\(1968\)007<0105:TDOCKPO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1968)007<0105:TDOCKPO>2.0.CO;2).
- , and T. W. Harrold, 1970: Air motion and precipitation growth at a cold front. *Quart. J. Roy. Meteor. Soc.*, **96**, 369–389, <https://doi.org/10.1002/qj.49709640903>.
- Didlake, A. C., G. M. Heymsfield, L. Tian, and S. R. Guimond, 2015: The coplane analysis technique for three-dimensional wind retrieval using the HIWRAP airborne Doppler radar. *J. Appl. Meteor. Climatol.*, **54**, 605–623, <https://doi.org/10.1175/JAMC-D-14-0203.1>.
- Dunion, J. P., G. A. Wick, P. G. Black, and J. Walker, 2018: Sensing Hazards with Operational Unmanned Technology: 2015–2016 campaign summary, final report. NOAA Tech. Memo. OAR-UAS-001, 49 pp., <https://doi.org/10.7289/V5/TM-OAR-UAS-001>.
- Guimond, S. R., L. Tian, G. M. Heymsfield, and S. J. Frasier, 2014: Wind retrieval algorithms for the IWRAP and HIWRAP airborne Doppler radars with applications to hurricanes. *J. Atmos. Oceanic Technol.*, **31**, 1189–1215, <https://doi.org/10.1175/JTECH-D-13-00140.1>.
- Hagen, M., 1992: On the appearance of a cold front with a narrow rainband in the vicinity of the Alps. *Meteor. Atmos. Phys.*, **48**, 231–248, <https://doi.org/10.1007/BF01029571>.
- Horton, K. G., B. M. V. Doren, P. M. Stepanian, W. M. Hochachka, A. Farnsworth, and J. F. Kelly, 2016: Nocturnally migrating songbirds drift when they can and compensate when they must. *Sci. Rep.*, **6**, 21249, <https://doi.org/10.1038/srep21249>.
- Kao, Y.-C., B. J.-D. Jou, J. C. L. Chan, and W.-C. Lee, 2019: An observational study of a coastal barrier jet induced by a landfalling typhoon. *Mon. Wea. Rev.*, **147**, 4589–4609, <https://doi.org/10.1175/MWR-D-19-0127.1>.
- Kelberlau, F., and J. Mann, 2019: Better turbulence spectra from velocity–azimuth display scanning wind lidar. *Atmos. Meas. Tech.*, **12**, 1871–1888, <https://doi.org/10.5194/amt-12-1871-2019>.
- , and —, 2020: Cross-contamination effect on turbulence spectra from Doppler beam swinging wind lidar. *Wind Energy Sci.*, **5**, 519–541, <https://doi.org/10.5194/wes-5-519-2020>.
- Lhermitte, R. M., and D. Atlas, 1961: Precipitation motion by pulse Doppler radar. *Proc. Ninth Weather Radar Conf.*, Kansas City, MO, Amer. Meteor. Soc., 218–223.
- Li, L., and G. Heymsfield, 2009: ER-2 X-band radar (EXRAD) for cloud, precipitation and wind measurements. NASA Airborne Instrument Technology Transition Final Rep., 31 pp.
- , —, J. Carswell, D. Schaubert, J. Creticos, and M. Vega, 2008: High-Altitude Imaging Wind and Rain Airborne Radar (HIWRAP). *Proc. IEEE Int. Geoscience and Remote Sensing Symp.*, Boston, MA, IEEE, 354–357, <https://doi.org/10.1109/TGRS.2015.2456501>.
- Liu, Z.-Q., and F. Rabier, 2002: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J. Roy. Meteor. Soc.*, **128**, 1367–1386, <https://doi.org/10.1256/003590002320373337>.
- Lorsolo, S., J. L. Schroeder, P. Dodge, and F. Marks, 2008: An observational study of hurricane boundary layer small-scale coherent structures. *Mon. Wea. Rev.*, **136**, 2871–2893, <https://doi.org/10.1175/2008MWR2273.1>.
- Matejka, T., and R. C. Srivastava, 1991: An improved version of the extended velocity–azimuth display analysis of single-Doppler radar data. *J. Atmos. Oceanic Technol.*, **8**, 453–466, [https://doi.org/10.1175/1520-0426\(1991\)008<0453:AIVOTE>2.0.CO;2](https://doi.org/10.1175/1520-0426(1991)008<0453:AIVOTE>2.0.CO;2).
- Michelson, S. A., and N. L. Seaman, 2000: Assimilation of NEXRAD-VAD winds in summertime meteorological simulations over the northeastern United States. *J. Appl. Meteor.*, **39**, 367–383, [https://doi.org/10.1175/1520-0450\(2000\)039<0367:AONVWI>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0367:AONVWI>2.0.CO;2).
- Neiman, P. J., and F. M. Ralph, 2004: Modification of fronts and precipitation by coastal blocking during an intense landfalling winter storm in Southern California: Observations during CALJET. *Mon. Wea. Rev.*, **132**, 242–273, [https://doi.org/10.1175/1520-0493\(2004\)132<0242:MOFAPB>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0242:MOFAPB>2.0.CO;2).
- Nolan, D. S., and C. A. Mattocks, 2014: Development and evaluation of the second hurricane nature run using the joint OSSE nature run and the WRF Model. *31st Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., P91, <https://ams.confex.com/ams/31Hur/webprogram/Paper244751.html>.
- , R. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a hurricane nature run using the joint OSSE nature run and the WRF Model. *J. Adv. Model. Earth Syst.*, **5**, 382–405, <https://doi.org/10.1002/jame.20031>.
- Reitebuch, O., H. Volkert, C. Werner, A. Dabas, P. Delville, P. Drobinski, P. H. Flamant, and E. Richard, 2003: Determination of airflow across the Alpine ridge by a combination of airborne Doppler lidar, routine radiosounding and numerical simulation. *Quart. J. Roy. Meteor. Soc.*, **129**, 715–727, <https://doi.org/10.1256/qj.02.42>.
- Srivastava, R. C., T. J. Matejka, and T. J. Lorello, 1986: Doppler radar study of the trailing anvil region associated with a squall line. *J. Atmos. Sci.*, **43**, 356–377, [https://doi.org/10.1175/1520-0469\(1986\)043<0356:DRSOTT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<0356:DRSOTT>2.0.CO;2).
- Tian, L., G. M. Heymsfield, A. C. Didlake, S. Guimond, and L. Li, 2015: Velocity–azimuth display analysis of Doppler velocity data for HIWRAP. *J. Appl. Meteor. Climatol.*, **54**, 1792–1808, <https://doi.org/10.1175/JAMC-D-14-0054.1>.
- Witschas, B., S. Rahm, A. Dörnbrack, J. Wagner, and M. Rapp, 2017: Airborne wind lidar measurements of vertical and horizontal winds for the investigation of orographically induced gravity waves. *J. Atmos. Oceanic Technol.*, **34**, 1371–1386, <https://doi.org/10.1175/JTECH-D-17-0021.1>.