Last updated: 8 June 2021

NASA IMPACTS 2020
VAD Wind Data Set

Contact: Charles N. Helms, USRA and NASA/GSFC
charles.n.helms@nasa.gov

Version:
Dataset v01r02
VAD algorithm code v03r03
See end of document for change log

This data set is produced from the NASA IMPACTS 2020 Level 1B EXRAD scanning-beam radar data. While the VAD retrievals are entirely based on the Doppler velocities, the radar reflectivity factor is used during the quality control (QC) process. This document provides an overview of the data set and the QC procedures. Note, the QC procedures and flags used to produce this data set are performed without human intervention and, as such, caution is needed when interpreting this data, particularly when using a bulk statistical, rather than a case study, approach.

1. VAD Technique Overview

The VAD technique used to produce this data set is described in Helms et al. (2020) and is an extension of the technique described in Tian et al. (2014). What follows here is a very general overview of this VAD technique.

The VAD technique used here computes track-relative horizontal wind components from aircraft-motion-corrected Doppler velocity measurements. To accomplish this, 5 coefficients are retrieved from a multivariable linear regression of Doppler velocity as a function of azimuth angle (and its various harmonics). The linear regression has the following form:

$$V_r = C_0 + C_1 cos\,\theta + C_2 sin\,\theta + D_1 cos\,2\theta + D_2 cos\,2\theta$$

where V_r is the Doppler velocity, C_0, C_1, C_2, D_1,and D_2 are the coefficients, and θ is the track-relative azimuth angle. These coefficients are then used to compute the retrievals via the following pairings: C_0corresponds to a linear combination of the vertical Doppler velocity and horizontal divergence, $C_1$and C_2correspond to the along- and across-track horizontal wind components, and $D_1$and D_2correspond to the track-relative stretching and shearing deformations. Earth-relative value can then be determined by rotating the vectors from a track-relative reference frame (i.e., 0° is in the direction of aircraft travel) to the Earth-relative reference frame (i.e., 0° is north).

A key assumption of all VAD techniques is that the underlying horizontal wind field varies only linearly and the vertical Doppler velocities (i.e., the vertical particle motions) are constant for all data points ingested into the retrieval. While this assumption is almost never exactly met in the real atmosphere, VAD retrievals usually give realistic values in horizontally homogeneous conditions, such as regions of stratiform precipitation. It is worth noting that any change in horizontal deformation constitutes a nonlinearity in the horizontal wind field. This assumption can be particularly problematic in convectively active regimes, such as in the vicinity of the tropical cyclone, or in regions of strong gradients in deformation, such as in the immediate vicinity of a front.

Typically, VAD retrievals use data from a single complete revolution of the radar antenna (i.e., a single scan) to produce a single wind retrieval. This results in the data being drawn from the edge of a fairly large footprint over which conditions can vary to a sufficient degree to introduce excessive errors to the retrieved winds. The Helms et al. (2020) approach minimizes these errors in two ways. First, by taking advantage of the movement of the radar platform, the technique can draw data from a smaller footprint in the along-track direction; all of the azimuth angles present in a single scan can be accounted for by taking all data within an along-track spatial window equal to the distance the aircraft moves during a single revolution of the radar antenna (referred to as a ‘synthetic’ scan; see Helms et al. 2020 for details). Second, by including additional data; the extra data points reduce the influence of any erroneous observations and should perform well when nonlinearities or vertical Doppler velocity fluctuations are on scales much smaller than the footprint of the retrieval. Combining the two methods can result in a reduction in error and an isolation of error-inducing regions such that the errors are limited to a small number of wind retrievals.

The VAD technique of Helms et al. (2020) results in four combinations of data selection (hereafter referred to as a ‘data selection strategy’). These are described in the following table:

Data Selection Strategy	**Description**
Sequential single scan	Traditional VAD using all data from a single revolution of the radar antenna
Sequential multiscan	Data points from a series of consecutive single scans are used to perform the retrieval
Synthetic single scan	Data is selected using an along-track spatial window of length equal to the distance traveled by the aircraft during a single revolution of the radar antenna
Synthetic multiscan	As per synthetic single scan, but the spatial window is larger than the distance traveled by the aircraft during a single revolution of the radar antenna

In general, the synthetic approaches tend to isolate errors while the multiscan approaches tend to smooth out errors. The synthetic multiscan approach combines both of these aspects.

2. VAD Internal Quality Control

This section covers the quality control (QC) process that is applied to the data before it is written to the output files. As such it does not include the QC flags that are provided in the data files and discussed in Section 3 (Output File Variables). Invalid data points are set to the missing value specified in the netcdf file metadata.

- Valid Doppler velocity data points are data points that meet the follow requirements:
  - Level 1B EXRAD scanning beam data indicates it's a valid data point
  - Aircraft roll at the time of observation is less than ±3 degrees
  - Passes the outlier test using an initial regression (see below)
- Doppler velocity outlier test:
  - Each VAD retrieval multiple linear regression is performed twice. The first regression is performed on all data that meets the first two bullet points under the definition of valid Doppler velocity data (see above).
  - The fitted sinusoid produced by this first regression is then subtracted from the observed Doppler velocities. Any Doppler velocities that lie farther from the fitted sinusoid than the amplitude of that fitted sinusoid are marked as invalid. This means that the observed Doppler velocity at the peak of the fitted sinusoid can be no greater than twice the amplitude above the mean Doppler velocity and no less than the mean Doppler velocity. Where the fitted sinusoid crosses the mean Doppler velocity, the observed Doppler velocity must lie between plus or minus the amplitude of the mean Doppler velocity.
  - The final VAD retrieval (as it appears in the output) is then computed using the remaining valid Doppler velocities.
- A VAD retrieval is considered valid if all of the following conditions are met during the second linear regression (i.e., after the Doppler velocity outlier test is applied):
  - At least 10 valid Doppler velocity data points are available for the regression
  - No more than 50% of the ingested data points are invalid
  - The multiple linear regression used to compute the VAD retrieval is successful

3. Output File Variables

This section covers the output variables and quality control (QC) flags. Mean, standard deviation, maximum, and minimum values will always be computed from all the data ingested into a VAD retrieval including data corresponding to missing/removed VAD wind or deformation data points. This distinction is important for any scan properties that are not based on wind or deformation (e.g., the scan-maximum reflectivity, *refl_max*). The dimensions of each variable are indicated in the parentheses.

- npoints_valid (range, time)
  - Number of valid (i.e., non-missing/removed) EXRAD scanning-beam Doppler radar data points used in the final VAD retrieval (i.e., during the second retrieval as described in Section 2). If there are no bad data points within the given data selection strategy scan, npoints_valid will equal npoints_total.
- npoints_total (range, time)
  - Total number of EXRAD scanning-beam Doppler radar data points (including missing/removed) available to be ingested into the VAD retrieval. If there are no bad data points within the given data selection strategy scan, npoints_valid will equal npoints_total.
- tilt (scalar)
  - Off-nadir (relative to the aircraft) tilt of the EXRAD scanning-beam radar antenna. For the IMPACTS-2020 deployment, this was 31.7 degrees.
- time (time)
  - Mean Julian timestamp of data points ingested in the VAD retrieval
- elapsed_time (time)
  - Elapsed seconds between the first mean Julian timestamp (i.e., *time* variable) in the current file for the current data selection strategy and the mean Julian timestamp of the current VAD retrieval. Note, this value can differ between data selection strategies.
- yt (time)
  - Mean along-track distance [meters] from the mean position of the first VAD retrieval in the current file for the current data selection strategy. Computed by integrating the aircraft ground speed from the start of each file.
- zt (range, time)
  - Mean distance from radar [meters]
- hght (range, time)
  - Mean height above mean sea level [meters]
- lat (time)
  - Mean latitude of the aircraft
- lon (time)
  - Mean longitude of the aircraft
- uvel (range, time)
  - VAD-retrieved zonal wind component [m/s] (converted from *avel* and *xvel* using *ac_track*)
- vvel (range, time)
  - VAD-retrieved meridional wind component [m/s] (converted from *avel* and *xvel* using *ac_track*)
- avel (range, time)
  - VAD-retrieved along-track wind component [m/s]. This is what is directly output by the VAD retrieval. Note, corrections for aircraft roll and pitch are not made to these values, so winds will always be retrieved in a plane perpendicular to aircraft nadir.

- xvel (range, time)
    - VAD-retrieved across-track wind component [m/s]. This is what is directly output by the VAD retrieval. Note, corrections for aircraft roll and pitch are not made to these values, so winds will always be retrieved in a plane perpendicular to aircraft nadir.
- dshr (range, time)
    - VAD-retrieved track-relative shearing deformation [1/s] as directly output by the VAD retrieval. The axis of dilitation is oriented 45 degrees counterclockwise of the along-track direction. Note, corrections for aircraft roll and pitch are not made to these values, so deformation will always be retrieved in a plane perpendicular to aircraft nadir. Also note that any changes in deformation within the footprint of the VAD retrieval are, by definition, a nonlinearity in the wind field, thus violating the VAD assumption of a linearly varying horizontal wind field.
- dstr (range, time)
    - VAD-retrieved track-relative stretching deformation [1/s] as directly output by the VAD retrieval. The axis of dilitation is oriented in the across-track direction. Note, corrections for aircraft roll and pitch are not made to these values, so deformation will always be retrieved in a plane perpendicular to aircraft nadir. Also note that any changes in deformation within the footprint of the VAD retrieval are, by definition, a nonlinearity in the wind field, thus violating the VAD assumption of a linearly varying horizontal wind field.
- refl (range, time)
    - Mean reflectivity factor [dBZ] of points ingested into VAD retrieval
- refl_max (range, time)
    - Maximum reflectivity factor [dBZ] of points ingested into VAD retrieval
- refl_std (range, time)
    - Standard deviation of reflectivity factor [dBZ] of points ingested into VAD retrieval
- footprint_maxdim (3, range, time)
    - Maximum spatial dimensions (along-track, across-track, and vertical) of VAD footprint [meters]. This value includes the effects of the along-beam averaging during the Doppler velocity retrievals and the effects of beam width on the overall VAD footprint
- footprint_maxdim_center (3, range, time)
    - Maximum spatial dimensions (along-track, across-track, and vertical) of VAD footprint [meters]. This value assumes all Doppler velocities are point observations and does not account for the effects of along-beam averaging or beam width on the overall VAD footprint.
- footprint_time (range, time)
    - Difference in time [seconds] between the earliest and latest Doppler radar data point ingested into the current VAD retrieval.
- delta_time (time)
    - Mean time [seconds] between subsequent Doppler velocities that are ingested into the VAD retrieval.

- delta_time_std (time)
    - Standard deviation of time [seconds] between subsequent Doppler velocities that are ingested into the VAD retrieval.
- delta_azimuth (time)
    - Maximum change in azimuthal angle [degrees] between subsequent Doppler velocity observations. For synthetic or multiscan data selection strategies, the azimuth angles (0 - 360 degrees) are sorted in numerical order before the calculation is performed.
- delta_azimuth_std (time)
    - Standard deviation of azimuth angle change [degrees] between subsequent Doppler velocity observations. For synthetic or multiscan data selection strategies, the azimuth angles (0 - 360 degrees) are sorted in numerical order before the calculation is performed.
- ac_alt (time)
    - Mean aircraft altitude above mean sea level [meters]
- ac_alt_std (time)
    - Standard deviation of aircraft altitude above mean sea level [meters]
- ac_track (time)
    - Mean aircraft track angle [degrees]. This is the direction the aircraft is actually traveling. 0° indicates the aircraft track is directed towards the north with clockwise being positive (i.e., +90° is pointed right/starboard). Uses a circular mean calculation.
- ac_track_std (time)
    - Standard deviation of aircraft track angle [degrees]. Uses a circular standard deviation calculation.
- ac_roll (time)
    - Mean aircraft roll angle [degrees]. 0° indicates the aircraft is flying with no roll, positive values indicate that the right wing is lower than the center of gravity (at least for any realistic attitude for the aircraft flown).
- ac_roll_std (time)
    - Standard deviation of aircraft roll angle [degrees].
- ac_pitch (time)
    - Mean aircraft pitch [degrees]. 0 degrees indicates the aircraft is flying with no pitch, positive values indicate that the aircraft nose is higher than the center of gravity (at least for any realistic attitude for the aircraft flown).
- ac_pitch_std (time)
    - Standard deviation of aircraft pitch angle [degrees]
- ac_heading (time)
    - Mean aircraft heading angle [degrees]. This is the angle towards which the aircraft nose points. 0° means the aircraft nose is pointed due north with clockwise being positive (i.e., +90° is pointed right/starboard). Uses a circular mean calculation.
- ac_heading_std (time)

    - Standard deviation of aircraft heading angle [degrees]. Uses a circular standard deviation calculation.
- ac_gspd (time)
    - Mean aircraft ground speed [m/s].
- ac_gspd_std (time)
    - Standard deviation of aircraft ground speed [m/s]
- vertical_resolution (range)
    - Approximate vertical resolution of the VAD data [meters]. Note that this does not account for the aircraft pitch or roll, so the resolution listed here is actually in the direction of the aircraft nadir. Features with a vertical dimension smaller than this value are unlikely to be properly resolved by this data set.
- antenna_rotdir (time)
    - Direction of antenna rotation flag. This value should not change during an individual flight. If this value does change and it is not noted in the latest version of the "Known Issues" section below, please let me know as the azimuth angles are probably erroneous and need to be fixed in the base EXRAD scanning data.
    - Values: 0 = unknown, 1 = clockwise, 2 = counterclockwise
- cor (range, time)
    - Multiple linear regression correlation coefficient from the VAD retrieval linear regression.
- qc1 (range, time)
    - QC flag indicating that an issue occurred in the multilinear regression used to perform the VAD retrieval that may result in a poor retrieval but was insufficient to cause the regression to completely fail. Users are encouraged to only use data where this flag is zero.
    - Values: 0 = quality_good, 1 = quality_bad
- qc2 (range, time)
    - QC flag indicating that the retrieved data point is probably contaminated by the radar side-lobe's surface return, which can produce incorrect winds, especially in low-reflectivity conditions. This flag is determined by flagging all data points with a mean reflectivity factor (*refl*) less than 0 dBZ that are located between +150 m and -1000 m of the expected peak nadir sidelobe altitude (based on antenna pointing angle). Users are encouraged to only use data where this flag is zero.
    - Values: 0 = quality_good (uncontaminated), 1 = quality_bad (contaminated)
- qc3 (range, time)
    - QC flag indicating that an azimuthal gap of more than 20 degrees exists in the data ingested in the VAD retrieval. This flag is computed using the *delta_azimuth* variable and a custom threshold can be applied by the user using that variable. As with the *delta_azimuth* variable, the azimuth angles (0 - 360 degrees) are sorted in numerical order before the calculation is performed. Users are encouraged to only use data where this flag is zero.
    - Values: 0 = quality_good (no large gaps), 1 = quality_bad (large gaps exist)
- qc4 (range, time)

- ○ QC flag indicating that the data ingested into the retrieval likely includes data contaminated by the main-beam surface return. This flag is set when the *refl_max* variable exceeds 45 dBZ. Note that there are no checks on altitude of the data, so high reflectivity thunderstorm cores may also trigger this flag. That said, a high reflectivity thunderstorm core is likely to produce a non-negligible violation of the VAD assumption that vertical Doppler velocity is constant across the ingested data points and should probably be ignored anyway. Custom thresholds can be applied using the *refl_max* variable. Users are encouraged to only use data where this flag is zero.
  - ○ Values: 0 = quality_good (no contamination), 1 = quality_bad (contamination)
- qc5 (range, time)
  - ○ QC flag set indicating that a certain percentage of the ingested Doppler velocity data was missing or removed. This flag is determined by dividing the *npoints_valid* variable by the *npoints_total* variable and, as such, a custom threshold can be applied using those variables. Users are encouraged to apply this flag as a “less than or equal to” test (e.g., qc5 LE 1). A good starting point for this flag is to only include retrievals with at least 90% valid data (i.e., flag values of 0 and 1).
  - ○ Values: 0 = all ingested points are valid, 1 = greater than or equal to 90% and less than 100% of points are valid, 2 = greater than or equal to 75% and less than 90% of points are valid, 3 = greater than or equal to 50% and less than 75% of points are valid, 4 = less than 50% of point are valid

4. Known Issues

This section contains a list of known issues with the data sets.

- Data from 1529 UTC through 1709 UTC on 7 February 2020 is missing due to bad azimuth data in the underlying EXRAD scanning-beam data. Some of this bad data seems to have gotten through the QC checks in the sequential single scan and sequential multiscan data sets and should be ignored.
- Occasionally, unreasonably large winds will be retrieved in a thin layer between 2 and 4 km altitude and are not flagged or removed by the automated QC algorithms. These erroneous winds do not appear to contaminate legitimate retrievals and can probably be safely screened out with an appropriate wind threshold.
- Sometimes the last few profiles are listed as having the antenna rotating the opposite direction from the rest of the file. This is being looked into.
- Several of the aircraft-related 1D variables should probably be 2D as they technically vary in both time and height for the synthetic scans (although they should not vary by much). These include: *delta_time, delta_time_std, delta_azimuth, delta_azimuth_std, ac_alt, ac_alt_std, ac_track, ac_track_std, ac_roll, ac_roll_std, ac_pitch, ac_pitch_std, ac_heading, ac_heading_std, ac_gspd,* and *ac_gspd_std*. At this moment, we don’t plan on expanding these unless specifically requested.

5. References

Helms, C. N., M. L. Walker McLinden, G. M. Heymsfield, and S. R. Guimond, 2020: Reducing errors in velocity–azimuth display (VAD) wind and deformation retrievals from airborne Doppler radars in convective environments, *J. Atmos. Oceanic Tech.*, **37**, 2251–2266. doi:https://doi.org/10.1175/JTECH-D-20-0034.1.

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, doi:https://doi.org/10.1175/JAMC-D-14-0054.1.

6. List of Acronyms

EXRAD - ER-2 X-band Radar
GSFC - NASA Goddard Space Flight Center
IMPACTS - NASA Investigation of Microphysics and Precipitation for Atlantic Coast-Threatening Snowstorms
QC - Quality Control
USRA - Universities Space Research Association
VAD - Velocity-azimuth Display

7. Data set change log

-- v01--

v01r02 - Fixed issue with across-track wind having the incorrect sign; VAD algorithm version v03r03

v01r01 - Fixed issue with zonal and meridional wind components (*uvel* and *vvel*); VAD algorithm version v03r02

v01r00 - Initial release; VAD algorithm version v03r01