

¹ **A Solo-based automated quality control algorithm for airborne
² tail Doppler radar data**

³ MICHAEL M. BELL ^{*}

University of Hawai'i at Manoa, Honolulu, Hawaii

⁴ WEN-CHAU LEE

CORY A. WOLFF

HUAQING CAI

National Center for Atmospheric Research, Boulder, Colorado [†]

^{*}Corresponding author address: Dept. of Meteorology, University of Hawaii at Manoa, 2525 Correa Rd., Honolulu, HI 96822

E-mail: mmbell@hawaii.edu

[†]The National Center for Atmospheric Research is sponsored by the National Science Foundation.

ABSTRACT

An automated quality control pre-processing algorithm for removing non-weather radar echoes in airborne Doppler radar data has been developed. This algorithm can significantly reduce the time and experience level required for interactive radar data editing prior to dual-Doppler wind synthesis or data assimilation. The algorithm uses the editing functions in the Solo software developed by the National Center for Atmospheric Research to remove noise, earth surface, side-lobe, second-trip, and other artifacts. The characteristics of these non-weather radar returns, the algorithm to identify and remove them, and the impacts of applying different threshold levels on wind retrievals are presented.

Verification was performed by comparison with published ELDORA datasets that were interactively edited by different experienced radar meteorologists. Four cases consisting primarily of convective echoes from the VORTEX, BAMEX, RAINEX, and TPARC/TCS08 field experiments were used to test the algorithm using three threshold levels for data removal. The algorithm removes 80, 90, or 95% of the non-weather and retains 95, 90, or 85% of the weather on average at the low, medium, and high threshold levels. Increasing the threshold level removes more non-weather echoes at the expense of also removing more weather echoes. The low threshold is recommended when weather retention is the highest priority, and the high threshold is recommended when non-weather removal is the highest priority. The medium threshold is a good compromise between these two priorities, and is recommended for general use. Dual-Doppler wind retrievals using the automatically edited data compare well with retrievals from interactively edited data.

²⁶ 1. Introduction

²⁷ Airborne Doppler radars are currently some of the best tools for deducing mesoscale
²⁸ three-dimensional wind and precipitation structure within weather systems at long range
²⁹ from an operations center. They have been primary tools for deducing the internal structure
³⁰ of mesoscale weather phenomena such as bow echoes, convective boundary layers, fronts,
³¹ hurricanes, mesoscale convective systems, tornadoes and winter storms (Lee et al. 2003 and
³² references therein). Raw airborne Doppler data contain both weather and non-weather
³³ echoes that require editing and quality control (QC) prior to wind synthesis, but interactive
³⁴ QC has been a hindrance for researchers due to the time and training required to properly
³⁵ identify non-weather radar echoes. To date, this interactive editing process has not been
³⁶ systematically documented. The purpose of the current study was: i) to document the
³⁷ characteristics of non-weather echoes in airborne Doppler radar, ii) to design an algorithm
³⁸ to pre-process airborne Doppler QC automatically to reduce both the time and necessary
³⁹ experience level for data users, and iii) to validate the algorithm and examine its impact on
⁴⁰ derived products such as dual-Doppler synthesis and radar data assimilation.

⁴¹ The ELDORA (ELectra DOppler RAdar)/ASTRAIA (Analyse Stereoscopique Par Radar
⁴² Aeroporte Sur Electra) system (Hildebrand et al. 1996), operated by the National Center for
⁴³ Atmospheric Research (NCAR), has the best spatial resolution of any currently operating
⁴⁴ airborne Doppler precipitation radar. ELDORA has two separate antennas that can scan
⁴⁵ rapidly at \sim 24 rpm, with a complete revolution of each radar available every 3 seconds.
⁴⁶ The two antennas are oriented \sim 16 degrees fore and aft of the cross-fuselage axis (i.e.,
⁴⁷ aircraft heading) so that the radar beams overlap in space as the aircraft moves forward.
⁴⁸ The ELDORA scanning strategy allows for dual-Doppler retrievals of the three-dimensional
⁴⁹ wind with 300 - 500 m spatial resolution, providing a tremendous advantage that can only
⁵⁰ be realized after this prodigious amount of data are quality controlled.

⁵¹ The QC process consists of airborne navigation and pointing angle corrections (Testud
⁵² et al. 1995, Bosart et al. 2002), followed by interactive data editing. The interactive QC

53 process of airborne Doppler radar data using the NCAR Solo software requires a significant
54 amount of time even for experienced radar meteorologists, with rough estimates of up to 30
55 minutes to edit a single radar scan. As a result, interactive editing of 10 minutes worth of
56 ELDORA radar data takes up to 240 hours by this estimate. With the algorithm described
57 herein the editing time for the 10 minutes of ELDORA data can be decreased *significantly*,
58 from several weeks down to a few days or hours, depending on the intended use of the radar
59 data.

60 A variety of algorithms to identify and remove non-weather radar echoes have been
61 developed for ground-based radars (Steiner and Smith 2002; Lakshmanan et al. 2007), but
62 many of the components from ground-based QC are not directly applicable to airborne
63 radar data (Bousquet and Smull 2003). For example, the surface echo characteristics are
64 quite different with airborne Doppler radars than with ground-based radars. Airborne tail
65 radars scan in an axis approximately perpendicular to the aircraft track, such that the radar
66 scans trace out a helix in space. A “clutter map” can be made for a ground-based radar
67 that indicates where stationary targets such as trees, terrain, or manmade structures are
68 present, but a static clutter map is not possible for a moving platform. Bousquet and Smull
69 (2003) alternatively used a digital terrain map to identify echoes from the earth’s surface
70 in their semi-automated QC algorithm for airborne radar data. QC algorithms based on
71 spatial patterns and numerical characteristics of the radar data may be generally applicable,
72 however. Bousquet and Smull (2003) also used a running-mean outlier detection algorithm
73 based on Bargen and Brown (1980) similar to that used in ground-based applications, with
74 further QC performed interactively.

75 Gamache (2005) developed an algorithm for automatic QC of data collected in hurricanes
76 by the National Oceanic and Atmospheric Association (NOAA) Tail-Doppler Radar (TDR).
77 The algorithm uses rule-based procedures similar to algorithm presented herein, but the
78 procedures and technical implementation are different. The Gamache (2005) algorithm was
79 developed for the specific context of hurricane reconnaissance, with a primary focus of pro-

80 viding real-time dual-Doppler winds for the NOAA National Hurricane Center and National
81 Center for Environmental Prediction. Hurricane specific features, such as the approximate
82 circular symmetry of the wind field and fixed operational flight patterns, can be exploited to
83 improve the QC. Since ELDORA is a research radar operated in the context of NSF funded
84 field experiments, both the observed phenomena and the data collection strategy change
85 from project to project, requiring a more general algorithm. One of the recorded data fields
86 from ELDORA that is an important part of the current QC algorithm is the normalized
87 coherent power (NCP), but this field is not currently computed by the TDR or used in the
88 Gamache (2005) algorithm.

89 Velocity aliasing can be of concern for the TDR when using a single pulse repetition
90 frequency with a relatively small Nyquist velocity (often 12.9 m s^{-1}). Although the technical
91 implementation is different, both NOAA TDR and ELDORA both have the ability to use a
92 dual pulse repetition frequency scheme to extend the Nyquist velocity to $60 - 80 \text{ m s}^{-1}$ and
93 essentially eliminate aliasing (Hildebrand et al. 1996, Jorgensen et al. 2000). The algorithm
94 presented here does not include dealiasing explicitly as in Gamache (2005), but can be
95 included if required (Bargen and Brown 1980).

96 The algorithm development was guided by two main design goals: i) to be general enough
97 to handle a wide variety of situations but specific enough to significantly reduce QC time,
98 and ii) to be implemented within the context of the ‘Solo’ interactive data editor. The first
99 design goal makes the algorithm generally applicable, but sacrifices advantages that could
100 be obtained from tuning for a specific weather regime. The current algorithm was therefore
101 designed in the context of convective weather given the preponderance of convective cases in
102 the ELDORA data archive. The trade-off is practical, but limits its applicability to a subset
103 of the ELDORA archive and may be sub-optimal for some weather regimes.

104 The second design goal was to develop the QC pre-processing algorithm using the ‘Solo’
105 interactive data editor (Oye et al. 1995). In contrast to a stand-alone, operational QC appli-
106 cation like that in Gamache (2005), a Solo script-based procedure can be easily distributed

107 to existing Solo users without requiring any additional software. The batch editing scripts
108 can also be used by researchers as a starting point for further interactive editing within the
109 Solo environment. The trade-off with the second design goal is that algorithm components
110 were limited to those available as existing Solo editing operations. A set of tunable param-
111 eters was developed with Solo components to meet different editing criteria using “low”,
112 “medium”, and “high” thresholds for data removal. It will be shown that increasing the
113 threshold level removes more non-weather echoes at the expense of removing more weather
114 echoes also. While good results are obtained with the current algorithm, one result of this
115 study is that rule-based procedures with tunable parameters such as those available within
116 Solo may be ultimately sub-optimal for weather discrimination. The current work helps pave
117 the way for a more complex algorithm with multi-dimensional classification metrics.

118 All radar QC algorithms suffer from a lack of an established “ground-truth” for vali-
119 dation. Non-weather echo characteristics are highly variable, and it is difficult to design a
120 synthetic verification dataset that captures this variability for different radars and weather
121 conditions. Determining whether a given radar volume consists of a weather or non-weather
122 echo ultimately depends on a radar meteorologist’s subjective appraisal of the scene and
123 experience level. While some volumes are easily classifiable, others are more ambiguous and
124 may be classified differently by different researchers. The approach taken in this study was
125 to validate the QC algorithm using datasets edited by different researchers from multiple
126 field projects. The selected datasets are taken from published papers by experienced EL-
127 DORA users, and provide a reasonable baseline for validation and statistical verification of
128 the algorithm using dichotomous classification metrics (Stanski et al. 1989; Woodcock 1976).
129 The verification approach contains some potential biases, but it is expected that differences
130 between datasets and researchers’ editing styles would remove any large systematic biases
131 in the verification statistics.

132 The desired product from a QCed airborne Doppler dataset is typically a wind retrieval
133 or improved initial conditions for a numerical simulation. While there are many studies that

134 address dual-Doppler and radar data assimilation techniques, considerably less attention
135 has been paid to the impacts of radar QC on these techniques. The impact of the different
136 thresholds of data removal on a dual-Doppler retrieval from a developing tropical cyclone
137 were examined, confirming the results of a previous study that wind retrievals using the
138 QC algorithm were robust compared to their interactively edited counterparts (Schmitz
139 2010). Impacts of the QC algorithm on radar data assimilation (Zhang et al. 2012) will also
140 be summarized here. These results suggest that the algorithm proves useful for preparing
141 airborne radar data for assimilation into numerical weather prediction models, real-time
142 dual-Doppler during a research flight to aid a mission scientist in making decisions about
143 where to next send the aircraft (Houze et al. 2006), or serve as a useful starting point for
144 further interactive editing for detailed case studies in post analysis mode.

145 An outline of the paper is as follows. Descriptions of the non-weather echoes in airborne
146 radar data and classification procedures are presented in section 2. The validation and
147 statistical verification of the QC algorithm are presented in Section 3. Impacts on dual-
148 Doppler retrievals and radar data assimilation are presented in Section 4. A summary and
149 concluding remarks are presented in Section 5. Technical details of the Solo software and
150 the automated QC algorithm are described in the appendices for interested readers.

151 **2. Classification of echoes in airborne tail Doppler radar** 152 **data**

153 *a. Characteristics of non-weather echoes*

154 A returned radar signal is quantified by the zero, first, and second radar moments that
155 correspond to the reflectivity (hereafter dBZ), Doppler velocity, and spectral width (SW),
156 respectively. The dBZ is directly proportional to the returned power and the strength of
157 the Doppler velocity signal. The Doppler velocity is a statistical estimate from multiple

158 pulses that can be affected by noise in the received signal and radar electronics. The SW
159 indicates the variance of the Doppler velocity within the radar pulse volume due to noise,
160 turbulence, wind shear, differential fall speed, and antenna motion. A less commonly used
161 radar parameter is the Normalized Coherent Power (NCP) that is defined by the ratio of the
162 power in the first moment (also called coherent power) to the total returned power. NCP
163 quantifies the relative power used to calculate the Doppler velocity with the total power used
164 to calculate dBZ, and is approximately inversely proportional to SW.

165 An unedited ELDORA radar sweep from the TPARC/TCS08 field campaign that illus-
166 trates these radar parameters is shown in Fig. 1. The radar image is from the aft-pointing
167 antenna at 23:51:31 UTC on 14 September 2008. The NRL P-3 aircraft is at the center of the
168 image, and a large convective echo and anvil are evident to the right of the aircraft. The four
169 panels show dBZ, Doppler velocity, SW, and NCP. Five main types of data can be identified
170 in Fig. 1: weather, noise, surface, radar side-lobes, and second-trip echoes. Gray shading in-
171 dicates values off the color scale, and the locations of the non-weather echoes have been high-
172 lighted. Noise regions are characterized by low dBZ with a logarithmic range-dependence,
173 and random Doppler velocities. The SW and NCP fields distinguish predominately signal
174 and noise regions with a relatively sharp demarcation.

175 The tail Doppler scanning strategy yields a prominent surface echo with a high dBZ,
176 near-zero Doppler velocity over land (or the corresponding ocean surface velocity), and a
177 low SW. The surface echo is a thin, multi-gate echo directly beneath the aircraft, but it
178 expands in width at longer ranges and becomes difficult to distinguish from the strong
179 convective echoes at low altitude to the right of the aircraft. For ELDORA, there is also
180 some cross-contamination from the multiple frequencies used in the complex “chip” pulse
181 (Hildebrand et al. 1996). The close proximity of these frequencies results in a broadening of

182 the surface echo in range compared with that of the single frequency NOAA TDR.¹ Reflected
183 atmospheric echoes appear beneath the ground.

184 While partial beam-filling by surface echo would occur even with a “perfect” Gaussian
185 radar beam, we define “side-lobe echoes” as those that are a result of distinct peaks in the
186 power in the tail of the radar beam by antenna diffraction. The side-lobe surface reflection
187 manifests as a ring of moderate dBZ and low velocity around the aircraft, with flared echoes
188 near the base of the ring that appear to emanate from the surface. The diameter of the side-
189 lobe surface ring is dependent on the altitude of the aircraft, with larger rings as the aircraft
190 flies higher. The strongest side-lobe echoes are located near the surface at approximately 20
191 - 30 degrees from nadir, and are characterized by a high NCP, moderate dBZ, and moderate
192 SW. Reflectivities decrease in the ‘flared’ portion of the echo further from the surface, but
193 the velocity signal is nearly constant throughout the echo. Side-lobe echoes occur primarily
194 in clear-air regions, but the boundary between the side-lobe echo and the weather echo is
195 often very subtle.

196 A “second-trip” echo is a radar echo from a previous pulse that returns from longer range
197 and is combined with the processing of the current pulse at shorter range. The second-trip
198 echo is evident as a large reflectivity wedge on the left side of the image, and is due to a
199 combination of returns from both the surface (below flight level) and convective weather
200 (above the flight level) at long range. On the right side of the aircraft, the stronger, first-
201 trip convective echo dominates the returned power and no second-trip echo is evident. The
202 second-trip echo appears as a spike in the reflectivity field with a random Doppler velocity,
203 a moderate NCP, and a high SW. Modern signal processing techniques have been identified
204 to separate multi-trip echoes, such as phase coding (Sachidananda and Zrnic 1999; Frush
205 et al. 2002) that may be implemented in future radar upgrades.

¹Multiple frequency contamination can be considered a “range side-lobe” similar to contamination by side-lobe energy found off the main beam axis. The range side-lobe can sometime also be found on the edge of strong weather echoes.

206 A close-up of reflectivity, Doppler velocity, and SW near the aircraft is shown in Fig.
207 2. The data has been partially edited for clarity by removing noise with a NCP threshold
208 value below 0.2. The second-trip echo has been partially, but not completely, removed by
209 the NCP threshold. Another feature of note that remains is a high dBZ, high velocity region
210 in the first few radar gates due to saturation of the receiver near the aircraft. The side-lobe
211 flare is still apparent on the left side of the aircraft, with reflectivity near 25 dBZ near the
212 surface and coherent velocity streaks around zero. The side-lobe echo is partially suppressed
213 in the shallow convection to the lower right of the aircraft, but can be seen in the clear-air
214 between the small convective cell and deeper convection further to the right.

215 The radar images shown in Fig. 2 illustrate the difficulty in discriminating weather and
216 non-weather echoes in some circumstances. It is hard to discern the side-lobe echo on the
217 right side in the dBZ field, and is only marginally easier in the velocity field even when a 1
218 m s^{-1} contour increment is used. Coherent velocity streaks around zero are evident in the
219 highlighted area upon close inspection, but the velocity magnitudes are very similar to those
220 found in the neighboring convection. There is a discernible increase in the SW field of a
221 few m s^{-1} in the side-lobe compared to the weather echo, however. The boundary between
222 the surface echo and deep convection is also difficult to discern in this example. The lack of
223 any ground-truth validation compounds the difficulty to definitively classify these ambiguous
224 echoes.

225 Every airborne radar scan contains varying combinations of these main echo types. Fig-
226 ure 3 shows histograms of radar parameters for weather (black) and non-weather (gray)
227 echoes derived from the entire verification dataset (details of the dataset are provided in
228 Section 3). The NCP histogram (Fig. 3a) shows the clearest separation of weather and
229 non-weather echoes, with high NCP values generally associated with weather and low NCP
230 values associated with non-weather. NCP can be regarded in some sense as a normalized
231 signal-to-noise ratio, with values near one indicating strong coherence of the velocity signal
232 and values near zero indicating noise. Intermediate values of NCP are ambiguous, such that

any threshold chosen to remove data invariably removes both weather and non-weather. The low, medium, and high thresholds used in the QC algorithm are indicated in Fig. 3a.

Figure 3b indicates that SW and NCP are generally negatively correlated, with similar but distinct information content. SW values above 8 are generally associated only with noise, but there is a large overlap region in which moderate SW values can be either weather or non-weather. The high frequency of non-weather with SW near zero is associated with surface echoes. Figure 3c indicates that dBZ by itself is a relatively poor discriminator of weather echoes. The mean dBZ of weather is higher than that of non-weather, but there is considerable overlap in the distributions. While SW and dBZ are not great discriminators by themselves, the combination of these parameters contains more information (Fig. 3d). In general, high SW associated with high dBZ would be more indicative of turbulent convection or wind shear, while high SW associated with low dBZ would be more indicative of noise. The ratio of SW to reflectivity in the QC algorithm is discussed further in the following section.

b. Automatic quality control algorithm

There have been relatively few automated QC algorithms developed for airborne weather radar, but one area that has received a thorough treatment is the removal of surface echo. For flat continental or oceanic returns, the primary radar gates affected by the surface can be identified by the intersection of the mainlobe of the radar beam with the surface (Lee et al. 1994a; Testud et al. 1995). The issue of surface removal in complex terrain was addressed in Georgis et al. (2000) and Bousquet and Smull (2003) through the use of a digital terrain map. Complex terrain removal is not currently part of the Solo software, highlighting one limitation of the existing QC framework used in this study.²

The above studies focused on removal of surface echoes within the main lobe (defined by the half-power beamwidth), but the strong backscatter from the surface can contaminate a

²This functionality has been implemented in a stand-alone version of the QC algorithm.

258 radar volume with only a small intersection of the “tail” of the beam and bias the Doppler
259 velocity towards zero. One advancement of the current QC algorithm is the use of a variable
260 “effective” beam width in the Testud et al. (1995) algorithm to remove data with partial
261 beam filling by surface echo. The use of a wider beam width removes partial surface echoes
262 left behind in the original form of the algorithm.

263 While the removal of all radar range gates even potentially affected by the surface may
264 be desired for some applications, excessive removal of near-surface echoes can be detrimental
265 for dual-Doppler synthesis. Over-estimating the amount of surface echo contamination can
266 significantly affect the magnitude and even the sign of the retrieved vertical velocity due to
267 the strong dependence on the measured low-level divergence. The variability of the partial
268 surface echo depends on the strength of the near-surface weather echo and beam broadening
269 with range, but without additional signal processing of the Doppler velocity time series it
270 is impossible to quantitatively determine how much a near-surface radar volume has been
271 affected by the surface.

272 A key advancement of the current study is the extension of automatic QC procedures
273 beyond the removal of surface echo. The basic editing procedure performed by the auto-
274 matic QC algorithm consists of nine steps shown in Table 1 and described in more detail
275 in Appendix B for interested readers. Some of the relevant details of the algorithm are
276 summarized here.

277 The Solo scripts used to execute the algorithm steps can be configured with user-defined
278 threshold values that compromise between the amount of weather and non-weather data
279 removed. Three basic threshold levels were defined as “low”, “medium”, and “high”, with
280 the corresponding script parameters shown in Table 2. The use of a general, procedural
281 approach necessitates the use of broad characteristics for data removal rather than specific
282 feature identification. The three threshold levels in the QC algorithm for NCP, SW, and
283 dBZ are indicated in Fig. 3a and 3d. Removing data with a low NCP is very effective at
284 removing noise, as evidenced by the histogram. At the low threshold almost no weather echo

285 is removed, but increasing the threshold to remove more undesirable non-weather echoes also
286 consequently removes more desirable weather echoes.

287 The ratio of SW to dBZ was found to be a good identifier of side-lobe echoes, and
288 Fig. 3d shows the logarithm of the ratio of SW to dBZ converted to its linear value in
289 $\text{mm}^{-6}\text{m}^{-3}$ for dBZ less than zero. A SW/Z ratio near zero is ambiguous, but increasing
290 values of the ratio are associated with higher probabilities of non-weather. It is noted that
291 the high threshold uses a 5 dBZ value, which increases the relative frequency of weather
292 slightly compared to Fig. 3d. While most ground-based algorithms include SW and dBZ
293 information, the SW/Z parameter is a novel combination to the authors' knowledge. The
294 ratio is not explicitly calculated in Solo, but a combination of SW and dBZ that mimics this
295 parameter is included in the algorithm.

296 It is also noted that the order the Solo components are executed plays an important
297 role in the algorithm, such that successively finer discriminations are made after coarser
298 data removal. For example, the "de-speckling" and "de-freckling" algorithms described in
299 Appendix B were found to be effective at removing isolated data, noise, and second-trip
300 echoes after the bulk of non-weather echo was removed.

301 An example of the QC results for the radar scan shown in Figs. 1 and 2 is shown in
302 Fig. 4. Figure 4a shows the "reference" dBZ, which was edited through a combination of
303 Solo operations described above and manual removal of remaining non-weather echoes. The
304 reference field represents the best estimate of the partition into weather and non-weather
305 echoes, and is regarded as "truth" for the purposes of validation. The effect of applying the
306 low threshold QC algorithm is shown in Fig. 4b. The algorithm removes the vast majority
307 of non-weather echoes, and the resulting dBZ field is generally comparable to the reference
308 field. A few distinct differences are apparent upon close inspection: i) the surface echo was
309 not removed completely to the lower-left of the aircraft, ii) small regions of side lobe echo
310 were not removed completely below the aircraft, and iii) low dBZ regions have been thinned,
311 especially in the anvil region to the upper-left of the aircraft.

312 The medium threshold algorithm (Fig. 4c) removes more of the surface echo and the
313 side lobe, but at the expense of additional reflectivity loss in the anvil and cloud edge above
314 the aircraft. An undesirable side effect of the additional surface removal is evident to the
315 right of the aircraft near the 30 km range ring where the base of the strong convective echo
316 has been removed. The high threshold algorithm (Fig. 4d) removes nearly all of the surface
317 echo and side lobe, as well as the anvil to the upper-left and even more of the convective
318 echo to the right of the aircraft, especially near the surface.

319 All three threshold levels of the QC algorithm are able to identify and remove the bulk
320 of the non-weather echo, but the details of the editing depend on the specific thresholds
321 used. The editing differences are most apparent along echo boundaries, in clear air echoes,
322 in side-lobe echoes near the surface, and in surface echoes at longer range from the aircraft.
323 There is an apparent trade-off between removing more non-weather echoes at the expense
324 of removing more weather echoes. The low algorithm is superior for maximum weather echo
325 retention, but leaves some non-weather echo that would require additional editing. The high
326 algorithm is superior for maximum non-weather echo removal, but removes more potentially
327 valuable weather echoes. The primary convective echo is well preserved at all three levels in
328 this example, suggesting that there would be no fundamental differences in the horizontal
329 winds retrieved by a dual-Doppler analysis. However, useful near-surface and echo-top radial
330 velocities are removed as the editing thresholds are increased, which may affect the derived
331 vertical velocity more significantly as will be shown in Section 4.

332 The example case illustrates that the choice of algorithm thresholds largely depends on
333 the user's application and amount of time they wish to spend on the editing. If vertical
334 velocity retrievals are critical in a research setting, the low algorithm provides an excellent
335 starting point for subsequent interactive editing. For data assimilation applications, the loss
336 of some good weather echoes may have a minimal impact compared to the insertion of radar
337 artifacts into the model, especially if the data are subsequently "super-obbed" after quality
338 control (Zhang et al. 2012). The trade-offs presented for the current example are generally

³³⁹ applicable for most ELDORA radar cases, but the details depend on the meteorological
³⁴⁰ situation and aircraft sampling strategy.

³⁴¹ 3. Algorithm verification

³⁴² a. Verification statistics

³⁴³ Verification of the QC algorithm was conducted using multiple radar datasets. The test
³⁴⁴ dataset was chosen from four field programs in which ELDORA was involved (Fig. 5). Each
³⁴⁵ case represents a different mesoscale regime so that the QC algorithm could be tested in a
³⁴⁶ variety of meteorological situations. The verification data are 1344 radar scans that were
³⁴⁷ interactively edited by experienced radar meteorologists and used in case studies published in
³⁴⁸ referred literature. The sample dataset contains 48 minutes of data and over 60 million radar
³⁴⁹ gates, which provides a reasonable sample for calculating statistics. It is recognized that data
³⁵⁰ editing is subjective and editing styles among these experienced radar meteorologists may
³⁵¹ vary, which could produce a bias in the statistical results presented. However, averaging
³⁵² over the variety of weather conditions and editing styles is expected to reduce the bias
³⁵³ from any one case. Table 3 provides information about each of the cases used for the
³⁵⁴ verification including the times of each of the flight legs. The field experiments are VORTEX:
³⁵⁵ Verification of the Origins of Rotation in Tornadoes Experiment, BAMEX: Bow Echo and
³⁵⁶ MCV Experiment, RAINEX: Hurricane Rainband and Intensity Change Experiment, and
³⁵⁷ T-PARC/TCS08: Thorpex Pacific Area Regional Campaign / Tropical Cyclone Structure
³⁵⁸ 2008. The VORTEX case (Wakimoto et al. 1998) and BAMEX case (Wakimoto et al. 2004)
³⁵⁹ represent different types of mid-latitude continental convection. The RAINEX (Houze et al.
³⁶⁰ 2006) and TPARC/TCS08 (Bell and Montgomery 2009) cases represent tropical oceanic
³⁶¹ systems in different stages of development.

³⁶² Measures of the skill of the QC algorithm can be derived using 2x2 contingency tables
³⁶³ where hits, misses, and false positives can be collected for each case. The edited data are

364 treated as dichotomous values where each gate in the sample is defined as either “positive”
365 weather or “negative” non-weather. A radar gate where both the QC algorithm and verifi-
366 cation dataset have identified a weather echo is considered a correct positive result. A gate
367 where both have identified a non-weather echo is considered a correct negative result. Correct
368 positive and negative results are considered hits. Radar gates falsely identified as weather
369 echoes by the QC algorithm are considered false positives. Radar gates falsely identified as
370 non-weather echoes are considered misses.

371 A typical sweep of data from ELDORA contains much more non-weather echo than
372 weather (cf. Fig. 1a to 4a). The algorithm skill scores would be dominated by data that are
373 obviously not weather without a baseline edit of the raw data. In order to get meaningful
374 statistics some very basic QC techniques were applied to the raw data before the fields were
375 run through the full QC algorithm. Data with NCP below 0.2, and radar echoes at and
376 below the surface identified with a 1.8 degree beam width and echoes above 25 km altitude
377 were removed prior to the verification. The baseline edited data allows evaluation of the
378 algorithm on data that are more difficult to classify as weather or non-weather.

379 Figure 6 shows a receiver operating characteristic (ROC; Joliffe and Stephenson 2003)
380 plot that displays the relationship between correctly identified events (hits) versus falsely
381 identified events (false positives) for the low, medium, and high thresholds. Anything to the
382 upper left of the 1:1 line across the ROC plot in Fig. 6a is considered to have positive skill.
383 A perfect algorithm with retention of all weather data and the removal of all non-weather
384 data would give results in the top left corner of the chart. The three data points for each test
385 case represent the average amount of data retained after application of the method using the
386 three different thresholds. The right-most points on each curve denote the low threshold and
387 the left-most points denote the high threshold. Thresholds in the algorithm could be tuned
388 by the user to yield different levels of editing between the high and low points illustrated
389 here. The curves are near the top left corner for the current algorithm, but trace counter-
390 clockwise as the QC thresholds are increased. The two primary goals of the algorithm to

391 retain all weather and remove all non-weather cannot be met simultaneously and involve
392 trade-offs as the thresholds are increased.

393 The ROC curve is the best for RAINEX (green line), with over 95% of the weather
394 retained and 88% of non-weather removed at all thresholds on average. The RAINEX case
395 was in the outer eyewall of hurricane Rita in 2005 and the data contain large regions of
396 stratiform and convective weather echoes. The other tropical case from TPARC/TCS08
397 (black line) performed the second best, with almost all non-weather data removed at the
398 high threshold. The continental cases from VORTEX (blue line) and BAMEX (red line)
399 still verify well, but with less skill than the oceanic cases. The differences in skill between
400 the oceanic and continental cases may be due to differences in aircraft altitude, the surface
401 echo characteristics between land and ocean, meteorological conditions, and editing style of
402 the researchers. The mean weather echo retained when averaged over all cases decreases
403 from 95% to 90% to 85% with increasing thresholds. The mean non-weather echo removed
404 increases from 80% to 90% to 95% with increasing thresholds.

405 Statistical distributions illustrating the algorithm performance over all the cases are
406 shown in Fig. 7. The top two panels show probability distribution functions (PDFs) for
407 each of the three threshold levels, and the bottom two panels show cumulative probability
408 distributions (CDFs). Figure 7a shows the probability that the algorithm met the first of
409 its primary goals by retaining weather echoes. There is considerable overlap in the PDFs,
410 but the distributions generally broaden for weather retention and sharpen for non-weather
411 removal with increasing thresholds. The low threshold PDF (green line) has the narrowest
412 distribution and peaks at 95% weather retained. The medium threshold PDF (blue line) has
413 a broader distribution and peaks closer to 85% weather retained. The high threshold PDF
414 (red line) has the broadest distribution with a nearly flat probability peak between 75 and
415 85% weather retained.

416 Figure 7b shows the probability that the algorithm met the second primary goal by
417 removing non-weather echoes. The PDFs are approximately the reverse of those shown in

418 Fig. 7a, with a sharpening of the PDFs as the thresholds are increased. The peaks of the
419 medium and high threshold PDFs are both at 95% non-weather removed, but the mean
420 probability that weather is removed at the medium threshold is lower due to the broader
421 distribution. The peak of the low threshold PDF is at 75% non-weather removed, but with
422 a very broad distribution. In general, increasing the thresholds increases the probability
423 that non-weather will be removed, but also decreases the probability that weather will be
424 retained. Applying different thresholds mostly affects the lower quality data that reside in
425 the gray area between weather and artifacts and has less impact on strong weather signals.

426 Figures 7c and 7d show the cumulative probability that the algorithm did not meet its
427 goals. The CDF of weather data that is removed is shown in Fig. 7c and the CDF that
428 non-weather data is retained is shown in Fig. 7d. Like the PDFs, the distributions are
429 approximately reversed with increasing thresholds. The CDFs can be used to estimate the
430 probability of error for the QC algorithm for a given threshold, with Fig. 7c corresponding to
431 misses and Fig. 7d corresponding to false positives. The low threshold CDFs indicate that
432 less than $\sim 12\%$ of the weather echo would be removed at the 95% confidence level, compared
433 to $\sim 26\%$ using the high threshold. However, no more than 16% of the non-weather would
434 be retained using the high threshold, compared to 30% using the low threshold. At the 99%
435 confidence level, the CDFs suggest that less than 20% of the weather echo would be removed
436 using the low threshold, and no more than 20% of the non-weather would be retained using
437 the high threshold.

438 Figures 5, 6, and 7 indicate a consistent depiction of the trade-off between weather reten-
439 tion and non-weather removal. It is clear that these two goals cannot be met simultaneously
440 with the current algorithm, but the thresholds provide a reasonable way of prioritizing one
441 goal over the other. The choice of a particular threshold then depends on the user's specific
442 application, experience level, and time commitment to further editing. The statistics sug-
443 gest that a user who wants to retain the maximum amount of weather echo should use the
444 low threshold. For example, an experienced user interested in dual-Doppler synthesis may

445 start the QC process with the low threshold script, followed by interactive editing to remove
446 the remainder of the non-weather data. The user can be confident that the vast majority
447 of weather echoes, including near-surface and echo-top data, are left intact. Likewise, the
448 statistics suggest that a user who wants to remove the maximum amount of non-weather
449 echo should use the high threshold. For example, radar data is typically thinned for data
450 assimilation into a numerical weather prediction model, such that only the highest quality
451 data may be desired. The statistics suggest that the medium threshold is a good compromise
452 that reduces editing time while still producing a high quality dataset.

453 *b. Skill metrics*

454 A variety of metrics are available to assess the skill of a dichotomous algorithm. The
455 threat score (TS; Schaefer 1990) quantifies how well weather identification by the algorithm
456 corresponds to weather identification by the experienced radar meteorologists. The TS
457 ranges from 0 (no skill) to 1 (perfect algorithm) and is calculated by dividing the number of
458 hits by the total number of hits, misses, and false positives. Figure 8a shows the TS for all
459 four of the cases tested along with their average score (cyan bar). The algorithm performed
460 well for these cases, with TS near or above 0.8 at all threshold levels. A slight decrease in
461 the TS is evident as the thresholds are increased, with average TS values of 0.89, 0.88, and
462 0.85 for the low, medium, and high thresholds respectively. The TS values indicate that 80-
463 90% of the algorithm attempts to identify weather and non-weather were successful. The
464 TS generally gives higher scores as the probability increases, which is consistent with the
465 highest scores for the RAINEX case that had the most weather echo.

466 The equitable threat score (ETS; Gandin and Murphy 1992) shown in Fig. 8b is calcu-
467 lated in a similar manner to the TS, but is adjusted for hits associated with random chance.
468 The ETS was developed so that scores could be more easily compared across different regimes
469 and account for climatology. For the current study, the ETS makes adjustments for higher
470 hits from random identification of weather when there is more weather echo in a radar scan.

471 The ETS is lower than the TS for all cases after accounting for random hits but shows simi-
472 lar relative magnitudes across the cases. The algorithm exhibits decreasing skill with higher
473 thresholds for the oceanic and VORTEX cases, and varying skill for the BAMEX case. The
474 average ETS values are 0.62, 0.63, and 0.57 for the three threshold levels, respectively. Thus,
475 the ETS indicates ~60% of the successful attempts to identify weather showed skill relative
476 to random chance.

477 Another metric for quantifying the performance of the algorithm is the True Skill Statistic
478 (TSS; Woodcock 1976), which quantifies how well the algorithm discriminates weather from
479 non-weather echoes within a range of -1 to 1, where 0 indicates no skill. The TSS shown in
480 Fig. 8c is calculated by subtracting the probability of a false positive for non-weather echoes
481 from the probability of a hit for weather echoes (Doswell et al. 1990). The best discrimination
482 between weather and non-weather according to this metric are equal at the medium and high
483 thresholds with an average TSS of 0.81. The TSS is only slightly lower at 0.75 at the low
484 threshold, indicating good discrimination skill for all threshold levels.

485 The metrics presented in Fig. 8 suggest that the QC algorithm has skill when validated
486 against echo classification performed by an experience radar meteorologist. The high TS
487 values correlate with the amount of weather echo present in the radar scans, but the ETS
488 and TSS values which corrected for this bias are considered skillful for all the test datasets.
489 The highest combined ETS and TSS skill is found at the medium threshold, suggesting the
490 medium threshold levels are nearly optimal for weather discrimination using the automated
491 Solo-based procedures. The algorithm is not perfect however, and perhaps additional skill
492 is achievable with a more sophisticated classification algorithm.

493 *c. Additional Tests*

494 An additional test of the method was also performed on ELDORA data from the Inter-
495 national H2O Project (IHOP) experiment that consisted primarily of clear-air echoes in the
496 boundary layer (Cai et al. 2006; Wakimoto and Murphrey 2010). The verification for the

497 IHOP case (not shown) was considerably poorer than the ELDORA convective cases, and
498 showed little to negative skill. The poor performance by the algorithm on the IHOP case can
499 be primarily attributed to removal of high SW and low dBZ echoes, or high SW/Z ratio (c.f.
500 Fig. 3d.) The IHOP counter-example emphasizes the fact that the current version of the
501 QC algorithm was designed for precipitating convection, and should be used with caution in
502 other meteorological situations, particularly in turbulent clear air returns.

503 A modified version of the algorithm was also tested using NOAA TDR data from the 2003
504 Coupled Boundary-Layer Air-Sea Transfer (CBLAST) experiment in Hurricane Isabel (Bell
505 and Montgomery 2008; Bell et al. 2012). The algorithm required modification since the TDR
506 data did not contain the NCP parameter and significant velocity dealiasing was present in the
507 test dataset. Replacing the NCP thresholding (step ii in Table 1) with a Bargen and Brown
508 (1980) dealiasing step provided a simple modification with which to test the algorithm using
509 a TDR dataset. The skill scores were generally consistent with the ELDORA convective
510 cases despite the algorithm modification and different radar characteristics. TS values were
511 near 0.9 for all three thresholds, and ETS values ranged from 0.51 at the high threshold to
512 0.57 at the low threshold. TSS values were the best for the high threshold at 0.68, and were
513 similar for the medium (0.60) and low (0.63) thresholds.

514 Though the NCP parameter is not recorded, the TDR data has significantly less data in
515 the clear air regions than ELDORA due to the reduced sensitivity of the radar receiver. The
516 reduced data coverage (and therefore reduced noise) in the clear air may help to minimize the
517 impact of the missing NCP thresholding step. However, failures in the velocity dealiasing can
518 have a large impact on the de-freckling step, and remove too much data. The CBLAST test
519 case indicates that the modified algorithm shows skill with NOAA TDR data, but further
520 verification and improvements to the velocity dealiasing, or use of newer dual-PRF data, are
521 needed.

522 4. Impacts of quality control on radar applications

523 Several tests of the QC algorithm's impacts on subsequent applications of airborne radar
524 data have been performed. Schmitz (2010) compared dual-Doppler wind retrievals using the
525 low, medium, and high automatic QC to interactively edited data from the BAMEX field
526 campaign (Wakimoto et al. 2006). She concluded that the wind fields from the automatically
527 edited data were virtually indistinguishable from those interactively edited. Further tests
528 of perturbation pressure retrievals (Gal-Chen 1978) indicated that the automatically edited
529 data had good, if not better, momentum cross-checks compared to the interactively edited
530 data (Schmitz 2010).

531 A similar test of the algorithm was performed for this study using the example dataset
532 presented in Sec. 2. A comparison of dual-Doppler retrievals using the methodology de-
533 scribed in Bell and Montgomery (2008) is shown in Fig. 9. The vertical velocity at 8 km
534 from the interactively edited data is shown in Fig. 9a in color, with 10 dBZ contours for
535 reference. A strong $\sim 20 \text{ m s}^{-1}$ updraft associated with the deep convective cell in Figs. 1,
536 2, and 5 is evident. The results using the QC algorithm in Fig. 9b-d indicate a vertical
537 velocity pattern that is nearly identical to the interactively edited data. Differences from
538 the original analysis larger than 1 m s^{-1} are contoured, revealing only minor differences in
539 the low and medium algorithms.

540 At the high thresholds, more significant differences exceeding 3 m s^{-1} are evident, with
541 a general reduction in the vertical velocity. This is believed to be due to the removal of
542 near-surface and echo top data that contains valuable divergence information. However, the
543 high threshold also removed the residual surface echo found on the left side of the aircraft
544 (see Fig. 5) that would negatively impact wind retrievals outside the chosen domain. The
545 maximum difference in the horizontal winds was found near the 3 km flight level, where some
546 low dBZ echoes present in the interactively edited data were removed at all three threshold
547 levels (not shown). Root mean square differences with the interactively edited data over the
548 entire domain of Fig. 9 are shown in Table 4. They are less than 1 m s^{-1} in all cases, but

549 do increase as the editing thresholds are increased. These results are consistent with those
550 from Schmitz (2010) suggesting that automatically edited wind fields are robust compared
551 to those from interactively edited data.

552 A test of the algorithm for data assimilation using the Weather Research and Forecasting
553 (WRF) model was performed by Zhang et al. (2012). They found positive impacts on both
554 track and intensity from assimilating automatically edited ELDORA data in a simulation
555 of Typhoon Jangmi (2008). The simulations suggested that the high threshold had a larger
556 impact on track, but the low threshold had a higher impact on intensity. It is unclear how
557 much of the difference is due to nonlinear processes in the model or specific differences in the
558 edited data. The positive impact on the simulation is encouraging however, and is consistent
559 with other recent studies assimilating airborne radar data into tropical cyclone simulations
560 (Aksoy et al. 2012, Weng and Zhang 2012).

561 5. Summary

562 The use of high resolution airborne Doppler datasets has typically been limited to ex-
563 perienced radar meteorologists and has required a significant amount of time for interactive
564 data editing. Non-weather echoes in airborne radar data vary from scan to scan, but exhibit
565 some consistent characteristics. The current study has documented and quantified some of
566 these characteristics in an attempt to move towards objective classification of weather and
567 non-weather echoes. An automated pre-processing quality control (QC) algorithm based on
568 interactive Solo editing procedures performed by experienced airborne Doppler data users
569 has been developed. The algorithm allows a user to QC a large number of radar scans in a
570 short period of time for subsequent dual-Doppler wind synthesis or data assimilation. An
571 estimated ten-fold time savings can be achieved by using the proposed pre-processing QC
572 algorithm compared with the traditional interactive editing using Solo.

573 The algorithm consists of nine steps based on the authors' experience editing data from

574 convective weather cases. Non-weather echoes resulting from noise, radar side-lobes, surface,
575 and second-trip returns were characterized and subject to a series of Solo radar editing
576 operations. Removing data with a low Normalized Coherent Power (NCP) was found to be
577 the most effective at removing noise. The Solo “de-speckling” and “de-freckling” algorithms
578 were found to be effective at removing isolated data, noise, and second-trip echoes. A
579 conditional threshold of high spectrum width and low reflectivity was found to be the best
580 for removing radar side lobes and preserve turbulent motions in the convective core, but was
581 detrimental to turbulent clear-air boundary layer returns.

582 The algorithm can be tuned at a combination of thresholds in different parameters for
583 data removal depending on the user’s application and priorities. The current study uses three
584 different combinations, termed low, medium and high thresholds, to illustrate and evaluate
585 the effectiveness of the proposed algorithm in removing non-weather radar returns. The
586 different threshold levels essentially tune the algorithm to classify echoes that are ambiguous
587 and not clearly weather or non-weather. Increasing the threshold level removes more non-
588 weather echoes at the expense of removing more weather echoes also. The low threshold is
589 recommended when weather echo retention is the highest priority, as it removes about 80%
590 of the non-weather echoes on average, and no more than ~20% of the weather echoes at the
591 99% confidence level in the verification dataset. The high threshold is recommended when
592 non-weather echo removal is the highest priority, as it removes about 95% of non-weather
593 echoes on average, and no more than ~30% of the weather echoes at the 99% confidence
594 level. The medium threshold is a good compromise between these two priorities, and is
595 recommended for general use.

596 Verification of the algorithm was performed using convective weather cases from the
597 ELDORA radar interactively edited by different radar meteorologists from the VORTEX,
598 BAMEX, RAINEX, and TPARC/TCS08 field experiments. The threat scores (TS) were
599 between 0.8 to 0.9 on average, suggesting that the algorithm has skill at identifying the
600 weather echoes corresponding to those identified as weather by a radar meteorologist. Lower

601 scores near 0.6 for the equitable threat score (ETS) suggest that the algorithm still has skill
602 after adjusting for hits associated with random chance. The true skill statistic (TSS), which
603 determines how well the method can discriminate between weather and non-weather echoes,
604 was between 0.6 and 0.8 on average. The TS decreased slightly with higher thresholds, while
605 the TSS increased slightly with higher thresholds. The medium threshold had the highest
606 combined ETS and TSS for the verification dataset, suggesting a near optimal combination
607 of threshold parameters. The metrics suggest that the algorithm performs similarly for the
608 different test cases in general, and it is applicable for a wide variety of convective datasets.
609 An additional test with NOAA Tail Doppler Radar data (TDR) from CBLAST indicates
610 a modified version of the algorithm shows skill for TDR convective cases as well as EL-
611 DORA data. The algorithm performed poorly on the IHOP dataset however, and is not
612 recommended for pre-processing primarily clear-air boundary layer echoes.

613 The application of different threshold levels is most apparent along echo boundaries, in
614 clear air echoes, in side-lobe echoes near the surface, and in surface echoes at longer range
615 from the aircraft. Large convective or stratiform echoes are largely unaffected at any of the
616 threshold levels, except near the surface and echo top. Application of the automatically
617 edited data in a dual-Doppler synthesis to retrieve the horizontal winds within convective
618 echoes was robust after applying the method and not strongly sensitive to the choice of
619 thresholds. The minimal impact on dual-Doppler wind retrieval is consistent with previous
620 studies that showed positive impacts of automatically edited data on both dual-Doppler
621 analyses (Schmitz 2010) and numerical simulations using radar data assimilation (Zhang
622 et al. 2012). However, accurate Doppler velocities near the surface and echo top are critical
623 for retrieving an accurate vertical velocity. Root mean square differences between wind
624 retrievals from interactively and automatically edited data were small, but point-wise values
625 differed by up to a few m s^{-1} . It is highly recommended to inspect and manually edit the
626 data after applying the QC algorithm for research quality wind synthesis.

627 Though the majority of the radar echo can be classified using a rule-based approach,

628 the real challenge for radar QC lies in the more ambiguous echoes. One of the primary
629 deficiencies of the current method is the use of hard thresholds for each of the discriminating
630 parameters. A more sophisticated algorithm must take into account the “fuzzy” nature of
631 the radar echoes in different scanning strategies and meteorological situations. For example,
632 it is evident that the probability of weather generally increases as the NCP increases, but use
633 of this single parameter does not contain enough information to define an optimal threshold
634 for all situations. The classification must ultimately be multi-dimensional, and requires more
635 complex logic that is not possible within the Solo framework. Furthermore, the classification
636 is currently subjective and could be improved with better objective characterization of non-
637 weather echoes using synthetic datasets. The medium threshold represents a near optimal set
638 of thresholds for meeting both requirements of removing non-weather and retaining weather
639 echoes identified by a radar meteorologist using Solo. However, further refinement appears
640 to be possible outside confines of the Solo software. Continued development of the method
641 to incorporate multi-dimensional fuzzy logic in a stand-alone software package is currently
642 underway.

643 *Acknowledgments.*

644 This research was funded by National Science Foundation through the American Recovery
645 and Reinvestment Act. The first author was also supported by the Office of Naval Research
646 Award N001408WR20129. The authors would also like to thank Cathy Kessinger and Andy
647 Penny for their comments on the manuscript.

APPENDIX A

Solo interactive radar editor

651 Solo was originally developed at NCAR in 1993 for perusing and editing Doppler radar
652 data (Oye et al. 1995), and was extensively revised and renamed as ‘SoloII’ in 2003. The title
653 of the software stems from the initial development location in the Solomon Islands as part of
654 the Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment
655 (TOGA-COARE) field program. The native data format for Solo is the Doppler Radar
656 Exchange (DORADE; Lee et al. 1994b) format developed for ELDORA, with individual
657 DORADE files colloquially known as “sweep files”. ELDORA data editing is typically done
658 by the principal investigator for a field project as part of their post-experiment research with
659 assistance from NCAR, where editing procedures can be tuned based on the specific project
660 requirements. Solo has been the primary editing software for many ELDORA and NOAA
661 tail Doppler radar users over the past years.

662 The Solo data viewer can display up to 12 color panels of different radar field variables
663 with range rings or Cartesian distance overlays. The numerical values of individual radar
664 gates can be inspected using the “Data” widget. A more detailed inspection of the data can
665 be performed using the “Examine” widget, which can display numerical values of multiple
666 data fields in a specified area, the metadata associated with the radar data header, or the
667 history of radar edits in the file. The Examine widget can also be used for point-and-click
668 deletion or velocity unfolding of individual radar gates or rays.

669 Both interactive and bulk radar data editing can be performed using the “Editor” wid-
670 get. In an interactive setting, the user can draw a boundary around an arbitrarily defined
671 patch of radar echo and apply a series of operators to that patch on a scan by scan basis.
672 In a batch mode, editing operations can be applied to multiple radar scans at one time.

Many editing operations are available for ground-based and airborne platforms, ranging from simple deletion to more complex series of logical operations. Solo also contains several diagnostic operators including histogram, rain rate, and radial shear calculations. The primary operators used for the QC algorithm presented in this study are the “remove surface”, “threshold”, “de-speckle”, and “de-freckle” operators. Ground removal operations use the aircraft altitude, aircraft attitude parameters, radar beam width, and the radar pointing angles to identify potential surface returns. At the present time, SoloII cannot handle complex terrain. Threshold operations remove data in one field based on the numerical values exceeding a specified parameter in another field. Data removal can be performed above, below, or between the specified threshold values. The de-speckle operator removes clusters of radar gates along a ray that are smaller than a specified “speckle” definition, defined by default as 3 gates. De-speckling therefore removes isolated radar gates that are not part of a larger weather echo. Isolated gates with large numerical values that are embedded within a weather echo can be identified by comparison with the average value of nearby radar gates. An outlier algorithm to identify these gates is implemented in the radial direction as the “de-freckle” algorithm, and in a circular patch as a “de-glitch” algorithm. Experiments for the current study with the de-glitch operator tended to remove more data compared to the de-freckle operator.

Other Solo editing operators include velocity dealiasing using the Bargen and Brown (1980) algorithm, and mathematical functions (i.e. add, subtract, multiply, and exponentiate). Solo can also be used to correct radar metadata such as aircraft inertial navigation system errors (Testud et al. 1995; Bosart et al. 2002). Earth-relative Doppler velocities and radar gate locations can then be recalculated using the corrected platform location and motion.

One deficiency of the “SoloII” software is that it depends on an older, deprecated 32-bit graphics library, and the code is not easily accessible for modification to add or improve the available editing steps. A stand-alone version of the batch editing functionality of the

700 software has been developed by the authors. The stand-alone version has an easily modified
701 Ruby script interface, but it currently lacks the full functionality of Solo as a data viewer and
702 interactive editor. Efforts to upgrade to “Solo3” including full 64-bit support are currently
703 underway at NCAR. The ‘SoloII’ software package, editing scripts described here, and stand-
704 alone editing software are available for interested users.

705 APPENDIX B

706

707 **Automated quality control procedures**

708 1) BACKUP ORIGINAL REFLECTIVITY AND VELOCITY FIELDS

709 The current version of Solo (called SoloII) does not have an ‘undo’ functionality. As a
710 result, it is important for any editing step using this software to store the results of unedited
711 or stepwise edited procedures that could result in unwanted loss of data. This step also
712 makes it easy to apply different levels of thresholding within the same file under different
713 data field names.

714 2) REMOVE NOISY DATA WITH LOW NCP

715 The NCP parameter is a normalized value between zero and one representing a spectrum
716 from pure noise to strong signal in the Doppler velocity estimate as described in Sec. 2. The
717 three thresholds selected for the automatic algorithm range from 0.2 to 0.4, above which
718 the removal of quality weather data becomes unacceptable for most applications. While the
719 NCP threshold is relatively simple, it removes the most data of any step in the algorithm.
720 Noisy gates that remain after this step are typically isolated and are further addressed in

721 step 6. The NCP parameter is not currently computed in NOAA TDR data, and therefore
722 this step is not applicable for TDR datasets.

723 3) REMOVE NOISY DATA NEAR THE EDGES OF RADAR RANGE

724 The data from the first few gates near the aircraft often have strong dBZ and noisy
725 Doppler velocities due to saturation of the receiver as evident in Fig. 2. Similarly, data in
726 the last few gates at the edge of the radar range are often unusable due to signal processing
727 requirements. While the amount of noise is dependent on the meteorological conditions and
728 the unambiguous range of the radar, the few gates retained by a careful examination of these
729 regions is usually not worth the effort. Approximately 5 gates at the beginning and end of
730 each radar beam are therefore removed in this algorithm step.

731 4) REMOVE THE DIRECT SURFACE RETURNS

732 Testud et al. (1995) derived a formula to identify the radar gates contaminated by the
733 surface echo based on the half-power beamwidth of the radar antenna. An increase in the
734 “effective” width of the radar beam is sufficient to identify and remove gates with partial
735 beam filling in the surface identification algorithm. The low threshold uses an effective
736 beamwidth parameter of 2 degrees that is slightly above the native resolution of ELDORA
737 and TDR (1.8 degrees). Higher effective beamwidths of 3 and 4 degrees at the medium
738 and high thresholds remove substantially more surface echo, especially away from the nadir
739 rotation angle. At longer ranges, the high threshold may remove too much near-surface echo
740 and should be used with caution when quality vertical velocity estimates are important.

741 5) REMOVE DATA WITH HIGH SW AND LOW dBZ

742 The ratio of the SW and dBZ fields is analogous to the NCP parameter, with the variance
743 of the Doppler signal used instead of the power in the velocity signal itself. In this step,

744 volumes with SW greater than 6 m s^{-1} and dBZ less than zero are removed at the low
745 threshold level. The SW threshold is decreased to 4 m s^{-1} at the medium level, and the dBZ
746 threshold is increased to 5 at the high level.

747 For radars that do not record NCP, this step can serve as a proxy for step 2. The
748 caveat with the SW / dBZ threshold is that regions of turbulence or wind shear with small
749 scatterers will be adversely affected. The thresholds were chosen to remove side-lobe echoes
750 and minimize the chance of convective turbulence being removed by the algorithm. However,
751 tests indicate that this step can remove significant echo from the clear-air boundary layer
752 and should be used with caution when preservation of boundary-layer echoes is important.

753 6) “DE-SPECKLE” THE DATA

754 The spatial scales of the weather observed by airborne Doppler radars are typically much
755 larger than individual radar range gates. After the previous five steps have been performed,
756 the majority of the non-weather echoes have been removed and remaining isolated gates
757 with few neighbors are unlikely to represent significant meteorological features. An efficient
758 algorithm for identifying “speckles”, or isolated gates along the radar beam, is part of the
759 Solo editor with a variable definition for the speckle size. The low threshold uses the default
760 speckle definition of 3 gates, effectively removing features with spatial scales less than 450
761 m (for a 150 m range gate). The medium and high thresholds increase the definition to 5
762 and 7 gates, or 750 m and 1050 m, respectively.

763 Unfortunately, the corresponding algorithm to identify isolated gates in the azimuthal
764 direction is not part of the Solo package. Thus, a larger speckle definition has the potential
765 of removing features which are relatively thin in the radial direction but have significant
766 horizontal extent. The most common examples of this are the anvil region of a thunderstorm
767 or shallow, near-surface structures such as boundary layer clouds or insect returns. Careful
768 consideration of the features of interest is important for determining the speckle definition.

769 7) “DE-FRECKLE” THE DATA

770 Complementary to radar “speckles” are radar “freckles”. Freckles are radar gates that
771 are outliers relative to their neighbors within a larger weather echo. A freckle can be defined
772 as a gate with a value that deviates significantly from a spatial mean in the neighborhood of
773 the gate. Though the freckles can be defined using any radar parameter, only the Doppler
774 velocity is used in the current algorithm. Since this step is primarily for outlier detection,
775 the freckle definitions are identical for the three threshold levels with a 5 gate running mean
776 and a 20 m s^{-1} outlier criterion.

777 In general, the de-freckling algorithm is effective at removing spikes in the velocity field
778 within weather echoes while retaining realistic velocity gradients. The de-freckle algorithm
779 is also effective at removing second-trip echoes that have moderate dBZ but random Doppler
780 velocities. A potential failure mode of the de-freckle algorithm occurs when the mean be-
781 comes ill-defined relative to the outliers. This appears to occur most often where a second-
782 trip echo occurs radially inward of a weather echo. The running mean is then defined by
783 noisy data and the entire beam can be removed. The algorithm can also fail where there is a
784 strong velocity gradient across a region of missing data such as that found between hurricane
785 rainbands, or where the velocity has aliasing errors.

786 8) SECOND “DE-SPECKLING” AND SYNCHRONIZATION OF REFLECTIVITY AND VELOC-
787 ITY

788 A final pass with the despeckling algorithm removes any gates that were isolated by the
789 defreckling algorithm, and the synchronization ensures that identical gates are removed from
790 both the dBZ and velocity fields. The synchronization is not ideal, since the reflectivity may
791 be representative of weather even when the velocity is noisy. Unfortunately the algorithm
792 cannot distinguish this situation since the primary criteria are based on velocity metrics.

REFERENCES

- 795 Aksoy, A., S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, and F. Zhang, 2012:
 796 The *hwrf* hurricane ensemble data assimilation system (*hedas*) for high-resolution data:
 797 The impact of airborne doppler radar observations in an *osses*. *Mon. Wea. Rev.*, **140**,
 798 1843–1862.
- 799 Bargen, D. W. and R. C. Brown, 1980: Interactive radar velocity unfolding. *Preprints, 19th*
 800 *Conf. on Radar Meteorology*, Miami Beach, FL, Amer. Meteor. Soc., 278–285.
- 801 Bell, M. M. and M. T. Montgomery, 2008: Observed structure, evolution and potential
 802 intensity of category five Hurricane Isabel (2003) from 12-14 september. *Mon. Wea. Rev.*,
 803 **136**, 2023–2045.
- 804 Bell, M. M. and M. T. Montgomery, 2009: Airborne radar observations of pre-depression
 805 Hagupit during TPARC/TCS-08. *Preprints, 34th Conf. on Radar Meteorology*, Williams-
 806 burg, VA, Amer. Meteor. Soc.
- 807 Bell, M. M., M. T. Montgomery, and K. E. Emanuel, 2012: Air-sea enthalpy and momentum
 808 exchange at major hurricane wind speeds observed during CBLAST. *J. Atmos. Sci.*, in
 809 press.
- 810 Bosart, B. L., W.-C. Lee, and R. M. Wakimoto, 2002: Procedures to improve the accuracy
 811 of airborne Doppler radar data. *J. Atmos. Oceanic Technol.*, **19**, 322–339.
- 812 Bousquet, O. and B. F. Smull, 2003: Airflow and precipitation fields within deep alpine
 813 valleys observed by airborne Doppler radar. *J. Appl. Meteor.*, **42**, 1497–1513.
- 814 Cai, H., W. C. Lee, T. M. Weckwerth, C. Flamant, and H. V. Murphrey, 2006: Observations

- 815 of the 11 June dryline during IHOP 2002 - A null case for convection initiation. *Mon.*
816 *Wea. Rev.*, **134**, 336–354.
- 817 Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in
818 rare event forecasting based on contingency tables. *Wea. and Fore.*, **5**, 576–585.
- 819 Frush, C., R. J. Doviak, M. Sachidananda, and D. S. Zrnic, 2002: Application of the SZ
820 phase code to mitigate range-velocity ambiguities in weather radars. *J. Atmos. Oceanic*
821 *Technol.*, **19**, 413–430.
- 822 Gal-Chen, T., 1978: A method for the initialization of the an elastic equations: Implications
823 for matching models with observations. *Mon. Wea. Rev.*, **106**, 587–606.
- 824 Gamache, J. F., 2005: Real-time dissemination of hurricane wind fields determined from
825 airborne Doppler radar data. Joint Hurricane Testbed final report, U.S. Weather Research
826 Program, 12 pp., Miami, FL. URL http://www.nhc.noaa.gov/jht/2003-2005reports/DOPLRgamache_JHTfinalreport.pdf.
- 827
- 828 Gandin, L. and A. H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon.*
829 *Wea. Rev.*, **120**, 361 – 370.
- 830 Georgis, J. F., F. Roux, and P. H. Hildebrand, 2000: Observation of precipitating systems
831 over complex orography with meteorological Doppler radars: A feasibility study. *Meteor.*
832 *Atmos. Phys.*, **72**, 185–202.
- 833 Hildebrand, P. H., et al., 1996: The ELDORA/ASTRAIA airborne Doppler weather radar:
834 High-resolution observations from TOGA COARE. *Bull. Amer. Meteor. Soc.*, **77**, 213–232.
- 835 Houze, R. A., et al., 2006: The hurricane rainband and intensity change experiment: Obser-
836 vations and modeling of hurricane Katrina, Ophelia, and Rita. *Bull. Amer. Meteor. Soc.*,
837 **87**, 1503–1521.

- 838 Joliffe, I. T. and D. B. Stephenson, 2003: *Forecast Verification: A Practitioner's Guide in*
839 *Atmospheric Science*. Wiley and Sons, Inc., New York, 254 pp.
- 840 Jorgensen, D. P., T. R. Shepherd, and . A. S. Goldstein, 2000: A dual-pulse repetition
841 frequency scheme for mitigating velocity ambiguities of the NOAA P-3 airborne doppler
842 radar. *J. Atmos. Oceanic Technol.*, **17**, 585–594.
- 843 Lakshmanan, V., A. Fritz, T. Smith, K. Hondl, and G. J. Stumpf, 2007: An automated
844 technique to quality control radar reflectivity data. *J. Appl. Meteor.*, **46**, 288–305.
- 845 Lee, W. C., P. Dodge, F. D. Marks, and P. H. Hildebrand, 1994a: Mapping of airborne
846 Doppler radar data. *J. Atmos. Oceanic Technol.*, **11**, 572–578.
- 847 Lee, W.-C., F. D. Marks, and C. Walther, 2003: Airborne Doppler radar data analysis
848 workshop. *Bull. Amer. Meteor. Soc.*, **84**, 1063–1075.
- 849 Lee, W. C., C. Walther, and R. Oye, 1994b: DOoppler RAdar Data Exchange format
850 DORADE. Tech. rep., National Center for Atmospheric Research, 18 pp., Boulder, CO.
- 851 Oye, R., C. Mueller, and S. Smith, 1995: Software for radar translation, visualization,
852 editing, and interpolation. *Preprints, 27th Conf. On Radar Meteor.*, Vail, CO, Amer.
853 Meteor. Soc., 359–361.
- 854 Sachidananda, M. and D. Zrnic, 1999: Systematic phase codes for resolving range overlaid
855 signals in a doppler weather radar. *J. Atmos. Oceanic Technol.*, **16**, 13511363.
- 856 Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea.*
857 *Forecasting*, **5**, 570–575.
- 858 Schmitz, A., 2010: An investigation to address the robustness of aircraft radar data. M.S.
859 thesis, Meteorology, St. Louis University, 188 pp.
- 860 Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verification
861 methods in meteorology. World Weather Watch Tech. Rept. No.8, WMO, 114 pp., Geneva.

- 862 Steiner, M. and J. A. Smith, 2002: Use of three-dimensional reflectivity structure for auto-
863 mated detection and removal of nonprecipitating echoes in radar data. *J. Atmos. Oceanic
864 Technol.*, **19**, 673–686.
- 865 Testud, J., P. H. Hildebrand, and W.-C. Lee, 1995: A procedure to correct airborne Doppler
866 radar data for navigation errors using the echo returned from the earth's surface. *J. Atmos.
867 Oceanic Technol.*, **12**, 800–820.
- 868 Wakimoto, R. M., H. Cai, and H. V. Murphrey, 2004: The Superior Nebraska, supercell
869 during BAMEX. *Bull. Amer. Meteor. Soc.*, **85**, 1095–1106.
- 870 Wakimoto, R. M., C. Liu, and H. Cai, 1998: The Garden City, Kansas, storm during
871 VORTEX 95. Part I: Overview of the storm's life cycle and mesocyclogenesis. *Mon. Wea.
872 Rev.*, **126**, 372–392.
- 873 Wakimoto, R. M. and H. V. Murphrey, 2010: Frontal and radar refractivity analyses of the
874 dryline on 11 June 2002 during IHOP. *Mon. Wea. Rev.*, **138**, 228–241.
- 875 Wakimoto, R. M., H. V. Murphrey, A. Nester, D. P. Jorgensen, and N. T. Atkins, 2006: High
876 winds generated by bow echoes. Part I: Overview of the Omaha bow echo 5 July 2003
877 storm during BAMEX. *Mon. Wea. Rev.*, **134**, 2793–2812.
- 878 Weng, Y. and F. Zhang, 2012: Assimilating airborne doppler radar observations with an
879 ensemble kalman filter for convection-permitting hurricane initialization and prediction:
880 katrina (2005. *Mon. Wea. Rev.*, **140**, 841859.
- 881 Woodcock, F., 1976: The evaluation of yes/no forecasts for scientific and administrative
882 purposes. *Mon. Wea. Rev.*, **104**, 1209–1214.
- 883 Zhang, L., Z. Pu, W.-C. Lee, and Q. Zhao, 2012: The influence of airborne doppler radar
884 data quality on numerical simulations of a tropical cyclone. *Wea. and Forecasting*, **27**,
885 231–239.

⁸⁸⁶ **List of Tables**

⁸⁸⁷ 1	Description of editing steps in QC algorithm.	37
⁸⁸⁸ 2	Parameters used at the different QC threshold levels.	38
⁸⁸⁹ 3	Date, time, field experiment, and brief description of the cases in the verifica-	
⁸⁹⁰	tion dataset. See text for acronym definitions.	39
⁸⁹¹ 4	Root mean square values of dual-Doppler winds from interactively edited data	
⁸⁹²	and the difference using automatic editing in m s ⁻¹ .	40

TABLE 1. Description of editing steps in QC algorithm.

- i. Backup original reflectivity and velocity fields
- ii. Remove data with NCP below a specified threshold
- iii. Remove noisy data at edges of radar range
- iv. Remove the direct surface returns
- v. Remove data with high SW and low DBZ
- vi. “De-speckle” the data
- vii. “De-freckle” the data
- viii. Second “de-speckle” pass
- ix. Synchronize the reflectivity and velocity fields

TABLE 2. Parameters used at the different QC threshold levels.

	<i>Low</i>	<i>Medium</i>	<i>High</i>
NCP Threshold	0.2	0.3	0.4
Effective beamwidth	2	3	4
SW / DBZ Threshold	6/0	4/0	4/5
Speckle definition	3	5	7
Freckle definition (Outlier / Gate Avg)	20/5	20/5	20/5

TABLE 3. Date, time, field experiment, and brief description of the cases in the verification dataset. See text for acronym definitions.

Date	Begin & End Time (UTC)	Field Experiment	Description
16 May 1995	22:19:44 - 22:25:59	VORTEX	Tornadic supercell
23 Jun 2003	02:51:00 - 03:01:59	BAMEX	Mesocyclone in MCS
22 Sep 2005	18:00:00 - 18:21:59	RAINEX	Mature hurricane eyewall
14 Sep 2008	23:47:02 - 23:55:59	T-PARC / TCS08	Tropical pre-depression convection

TABLE 4. Root mean square values of dual-Doppler winds from interactively edited data and the difference using automatic editing in m s^{-1} .

	RMS Value	Low Diff.	Med Diff.	High Diff.
u	4.80	0.34	0.52	0.92
v	3.94	0.31	0.45	0.72
w	3.12	0.27	0.44	0.90

893 **List of Figures**

- 894 1 Example of radar echoes from pre-depression Hagupit on 14 September 2008
895 at 23:51:30 UTC during TPARC/TCS08. Panels show (a) reflectivity in 4
896 dBZ increments , (b) Doppler velocity in 2 m s^{-1} increments, (c) spectrum
897 width in 1.5 m s^{-1} increments, and (d) normalized coherent power in 0.05
898 increments. Non-weather echoes are labeled in top panel. 43
- 899 2 Close-up of radar echo from Fig. 1 near the aircraft. A 0.2 NCP threshold has
900 been applied to the data. Panels show (a) reflectivity in 5 dBZ increments,
901 (b) Doppler velocity in 1 m s^{-1} increments, and (c) spectrum width in 0.5 m s^{-1}
902 increments. The side-lobe echo to the right of the aircraft has been
903 highlighted in panel (b) by the dashed ellipse. 44
- 904 3 Relative frequency of weather (black) and non-weather (gray) echoes with re-
905 spect to various radar parameters. Panels are (a) normalized coherent power,
906 (b) spectrum width, (c) reflectivity, and (d) the logarithm of the ratio of
907 spectrum width to linear reflectivity. 45
- 908 4 Example of QC algorithm for the radar echo from Figs. 1 and 2. “Reference”
909 interactive editing is shown in panel (a), with results from the low, medium,
910 and high threshold levels of the QC algorithm shown in panels (b - d). 46
- 911 5 Reference reflectivity examples from the verification dataset from (a) VOR-
912 TEX, (b) BAMEX, (c) RAINEX, and (d) TPARC/TCS08. See Table 3 for
913 further description of the dataset. 47

- 914 6 Receiver operating characteristic curves for the QC algorithm. Full axes from
 915 0 to 1 are shown in panel (a) with the diagonal line demarcating the skill
 916 (upper-left) and no-skill (lower right) regions. A close-up of the area marked
 917 by the gray dashed box is shown in panel (b). Symbols represent low, medium,
 918 and high QC thresholds moving counter-clockwise from upper-right to lower
 919 left. Blue, red, green, and black curves correspond to different field experi-
 920 ments, and the cyan curve represents the average over the test dataset. 48
- 921 7 Probability and cumulative probability distributions for the QC algorithm.
 922 Panel (a) indicates PDF of weather echoes retained and (b) indicates PDF
 923 of non-weather echoes removed for test dataset. Panel (c) indicates CDF of
 924 weather echoes removed and (d) indicates CDF of non-weather echoes re-
 925 tained. Dashed green, solid blue, and dash-dot red lines represent the low,
 926 medium, and high threshold levels, respectively. 49
- 927 8 Verification metrics for the QC algorithm. The threat score is shown in panel
 928 (a), the equitable threat score in panel (b), and true skill statistic in panel
 929 (c) for each of the three threshold levels. Blue, red, green, and black bars
 930 correspond to different field experiments, and the cyan curve represents the
 931 average over the test dataset. 50
- 932 9 Vertical velocity (color) at 8 km altitude from dual-Doppler retrieval of exam-
 933 ple in Fig. 1 using different editing techniques. Panel (a) shows results from
 934 interactively edited data, and panels (b)-(d) from automatically edited data.
 935 Black contours in panel (a) denote 10 dBZ radar reflectivity intervals above 5
 936 dBZ. Black contours in (b)-(d) indicate differences from (a) larger than 1 m
 937 s^{-1} , with dashed contours denoting negative differences. 51

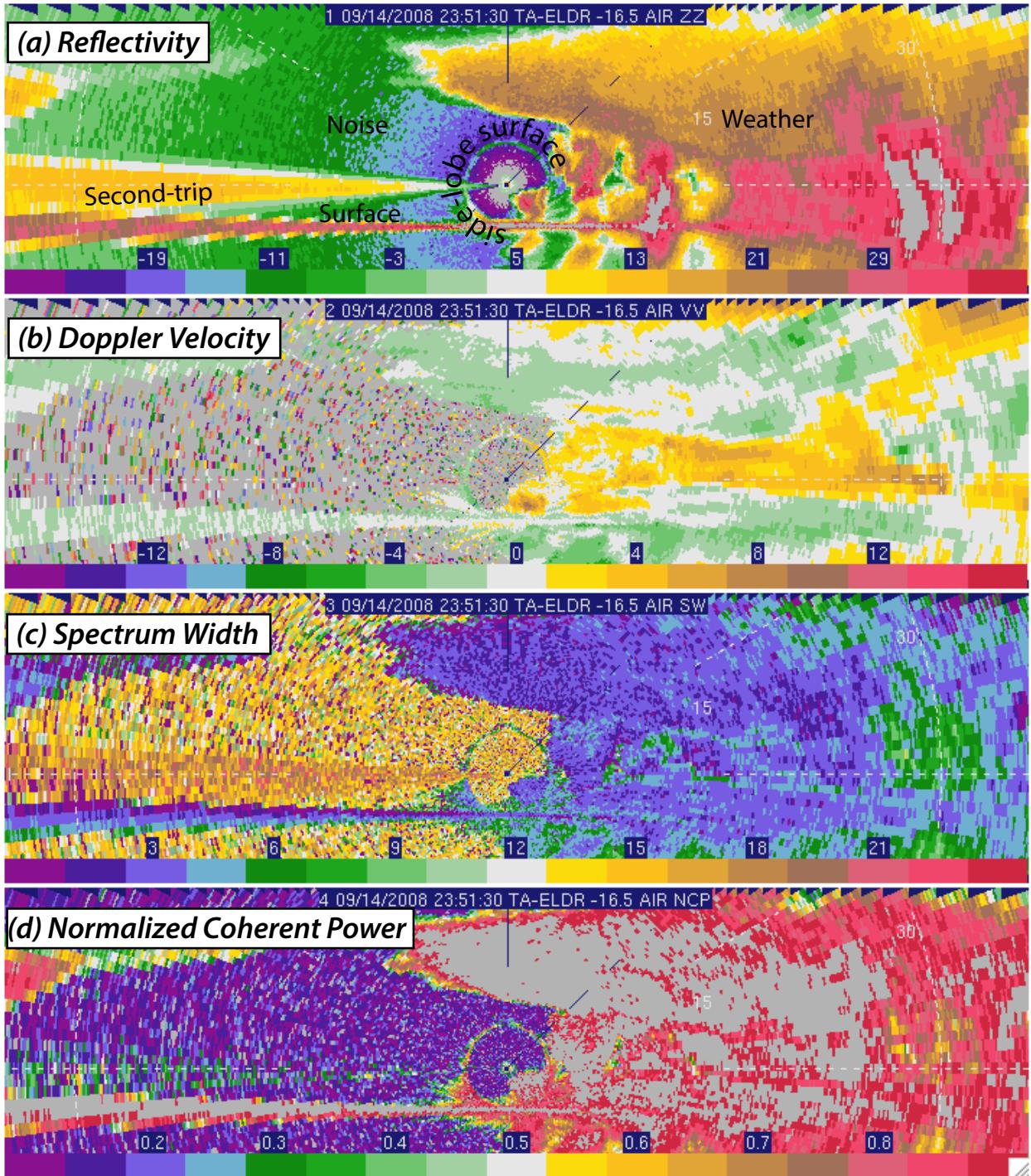


FIG. 1. Example of radar echoes from pre-depression Hagupit on 14 September 2008 at 23:51:30 UTC during TPARC/TCS08. Panels show (a) reflectivity in 4 dBZ increments, (b) Doppler velocity in 2 m s^{-1} increments, (c) spectrum width in 1.5 m s^{-1} increments, and (d) normalized coherent power in 0.05 increments. Non-weather echoes are labeled in top panel.

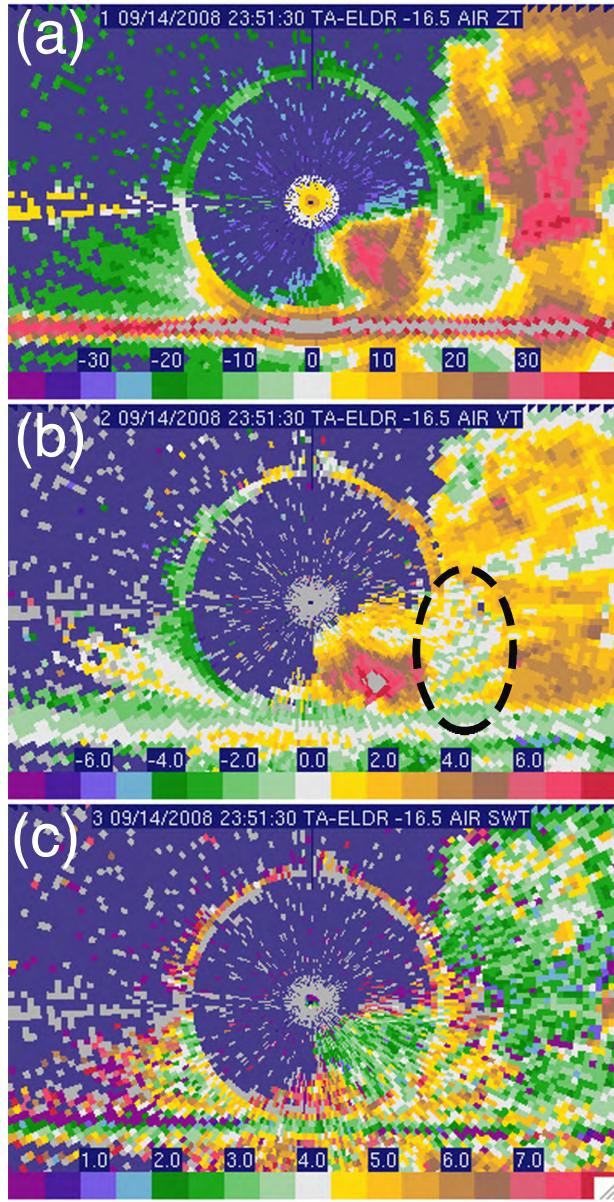


FIG. 2. Close-up of radar echo from Fig. 1 near the aircraft. A 0.2 NCP threshold has been applied to the data. Panels show (a) reflectivity in 5 dBZ increments, (b) Doppler velocity in 1 m s^{-1} increments, and (c) spectrum width in 0.5 m s^{-1} increments. The side-lobe echo to the right of the aircraft has been highlighted in panel (b) by the dashed ellipse.

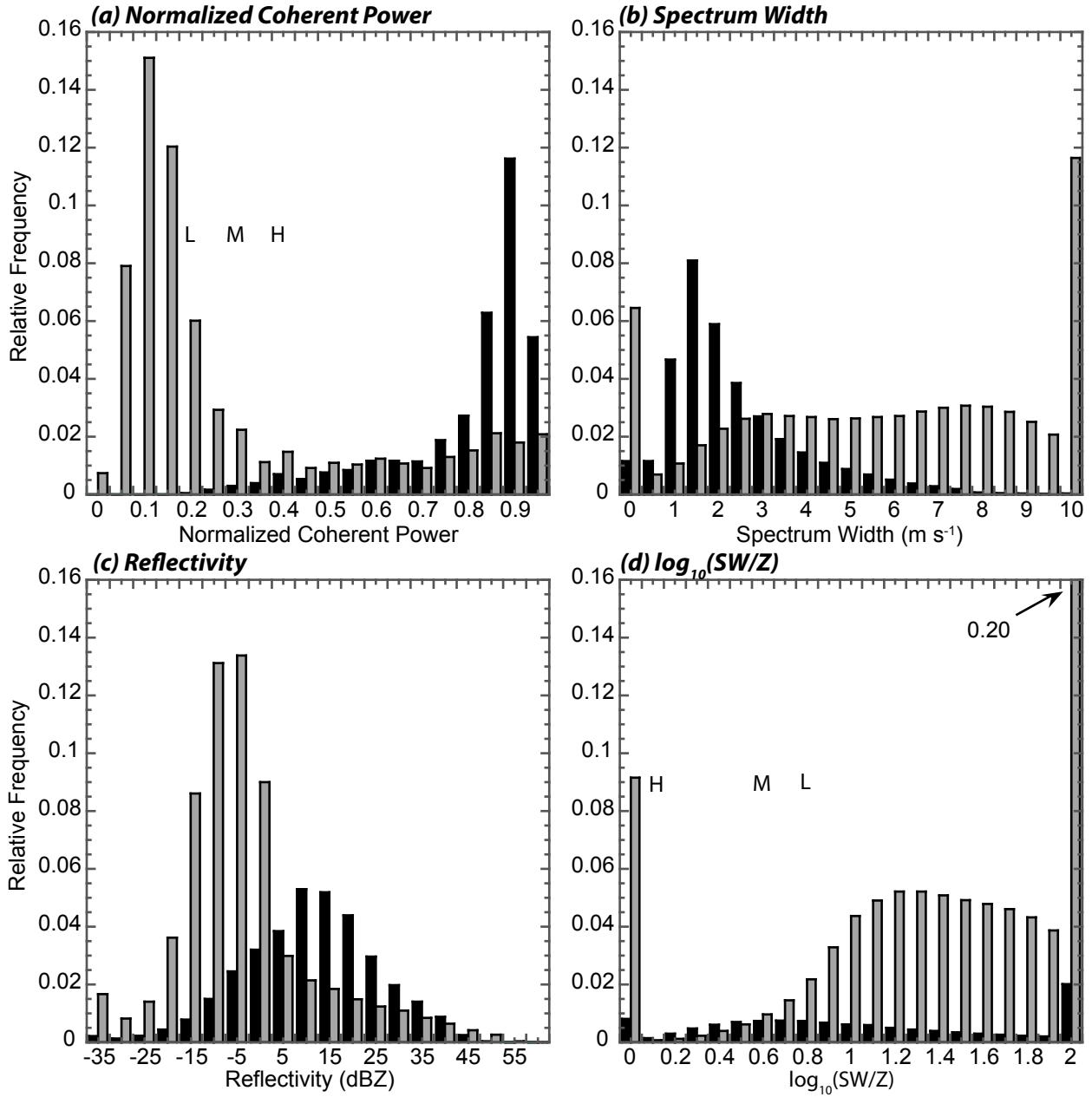


FIG. 3. Relative frequency of weather (black) and non-weather (gray) echoes with respect to various radar parameters. Panels are (a) normalized coherent power, (b) spectrum width, (c) reflectivity, and (d) the logarithm of the ratio of spectrum width to linear reflectivity.

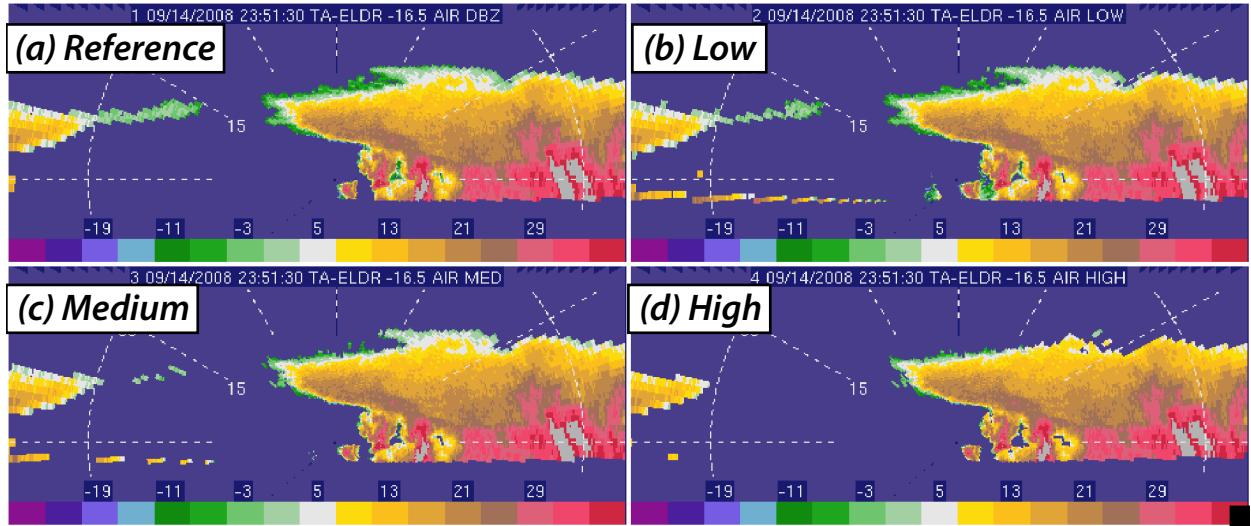


FIG. 4. Example of QC algorithm for the radar echo from Figs. 1 and 2. “Reference” interactive editing is shown in panel (a), with results from the low, medium, and high threshold levels of the QC algorithm shown in panels (b - d).

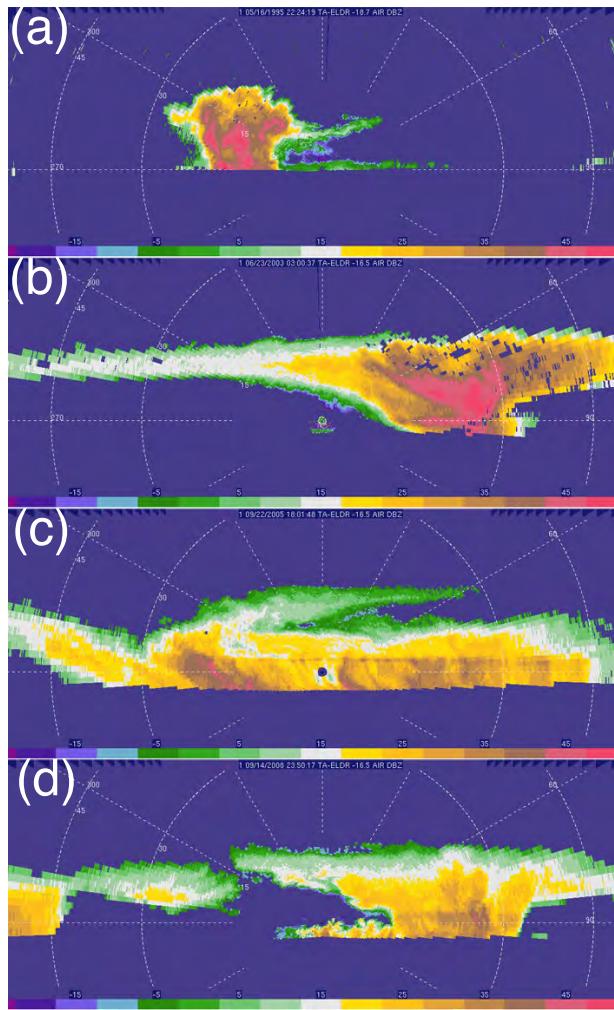


FIG. 5. Reference reflectivity examples from the verification dataset from (a) VORTEX, (b) BAMEX, (c) RAINEX, and (d) TPARC/TCS08. See Table 3 for further description of the dataset.

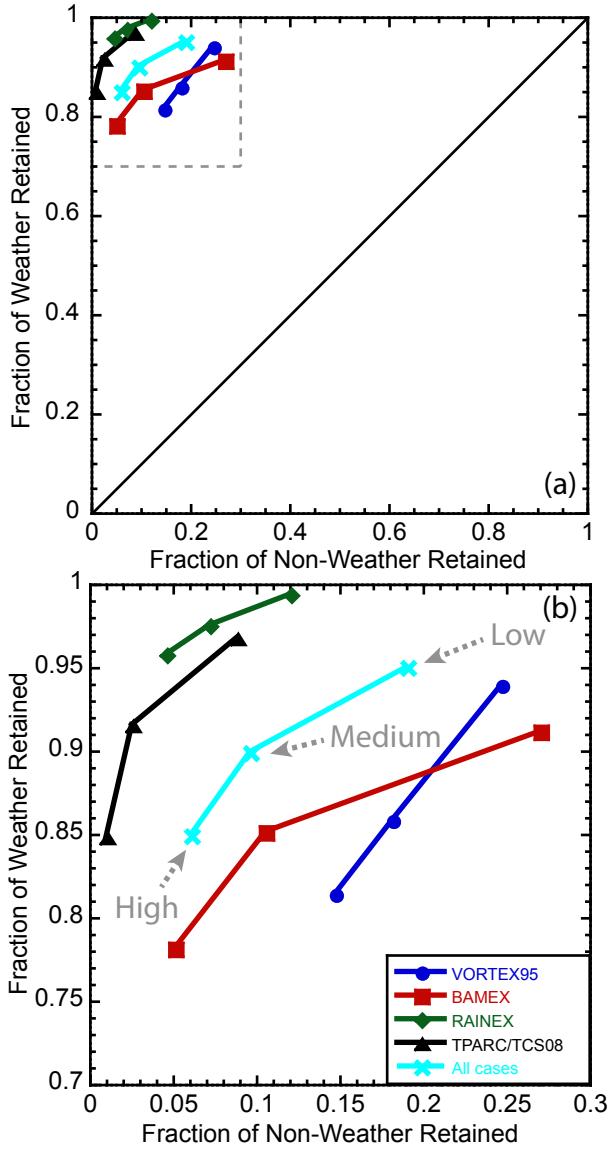


FIG. 6. Receiver operating characteristic curves for the QC algorithm. Full axes from 0 to 1 are shown in panel (a) with the diagonal line demarcating the skill (upper-left) and no-skill (lower right) regions. A close-up of the area marked by the gray dashed box is shown in panel (b). Symbols represent low, medium, and high QC thresholds moving counter-clockwise from upper-right to lower left. Blue, red, green, and black curves correspond to different field experiments, and the cyan curve represents the average over the test dataset.

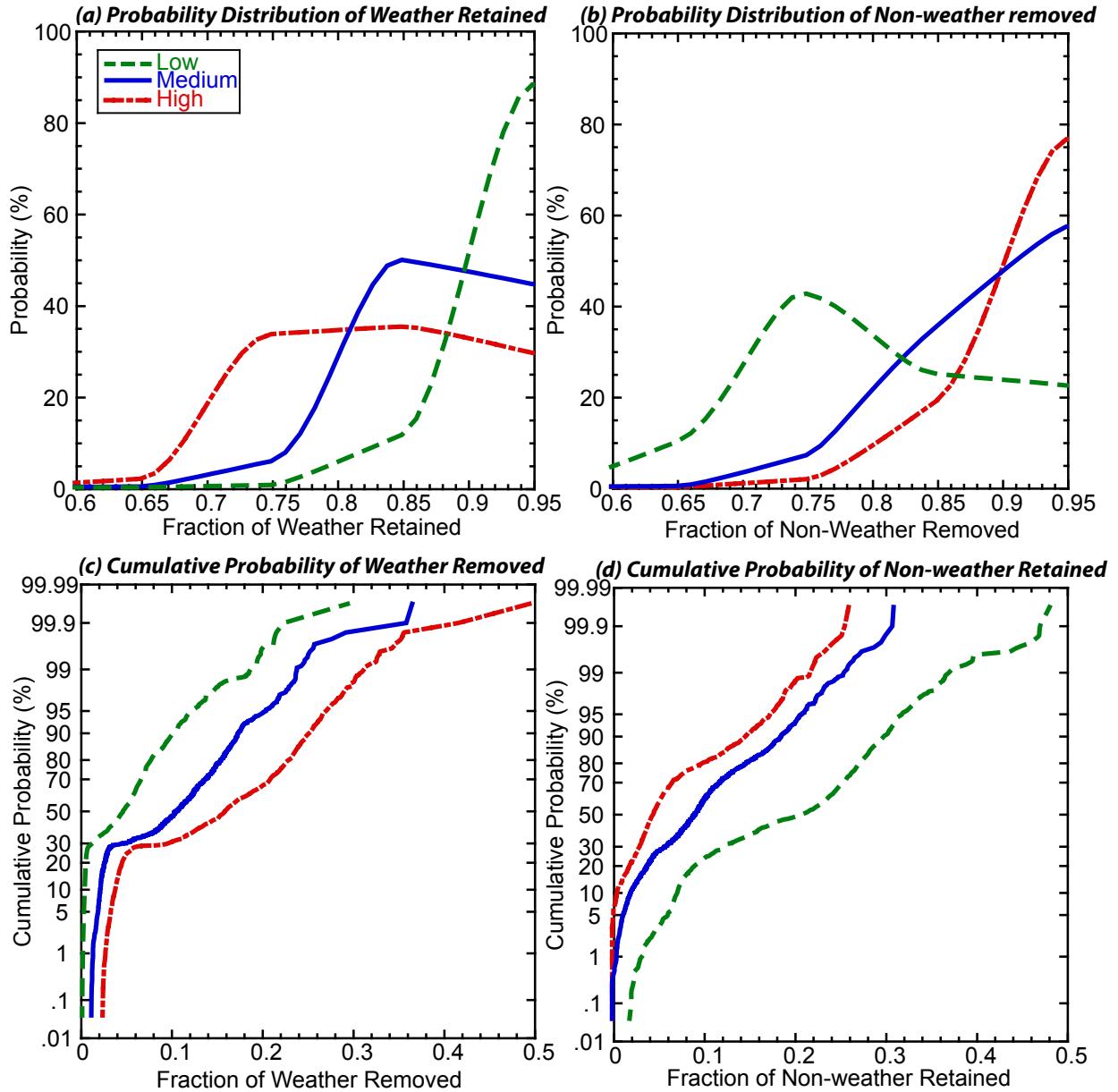


FIG. 7. Probability and cumulative probability distributions for the QC algorithm. Panel (a) indicates PDF of weather echoes retained and (b) indicates PDF of non-weather echoes removed for test dataset. Panel (c) indicates CDF of weather echoes removed and (d) indicates CDF of non-weather echoes retained. Dashed green, solid blue, and dash-dot red lines represent the low, medium, and high threshold levels, respectively.

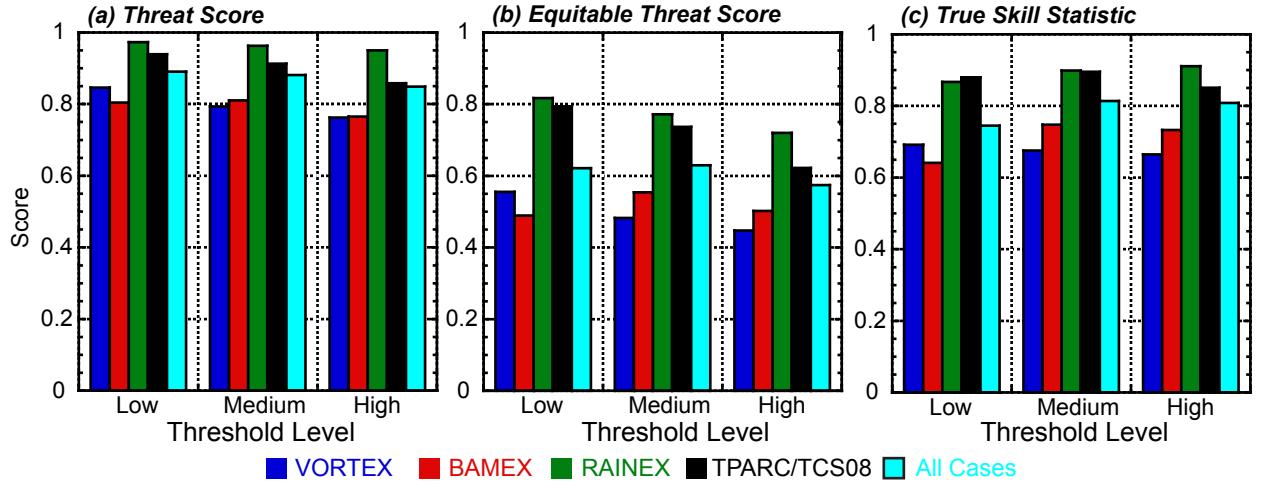


FIG. 8. Verification metrics for the QC algorithm. The threat score is shown in panel (a), the equitable threat score in panel (b), and true skill statistic in panel (c) for each of the three threshold levels. Blue, red, green, and black bars correspond to different field experiments, and the cyan curve represents the average over the test dataset.

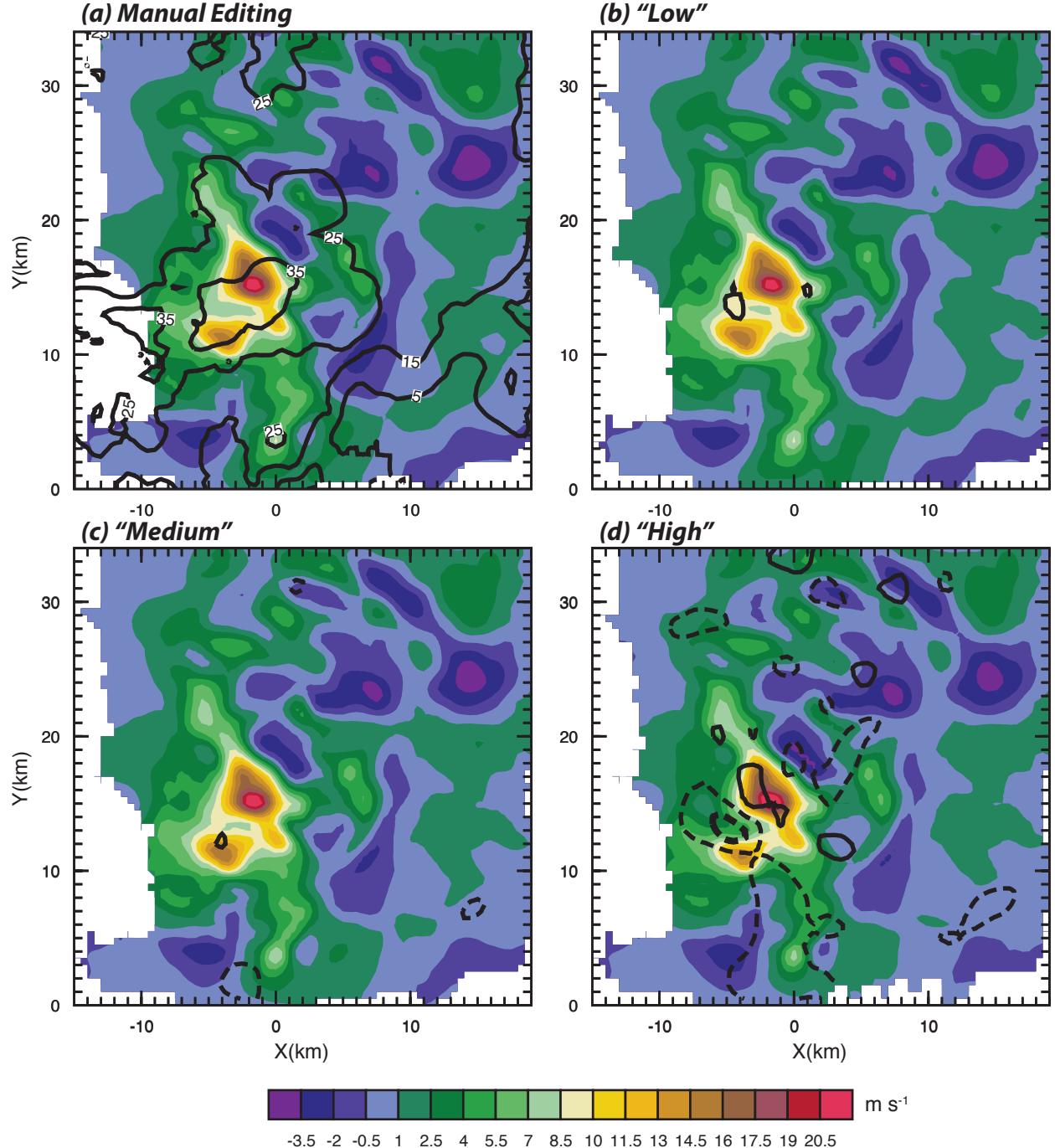


FIG. 9. Vertical velocity (color) at 8 km altitude from dual-Doppler retrieval of example in Fig. 1 using different editing techniques. Panel (a) shows results from interactively edited data, and panels (b)-(d) from automatically edited data. Black contours in panel (a) denote 10 dBZ radar reflectivity intervals above 5 dBZ. Black contours in (b)-(d) indicate differences from (a) larger than $1\ m\ s^{-1}$, with dashed contours denoting negative differences.