161

RECENT PROGRESS ON THE AUTOMATIC QUALITY CONTROL OF ELDORA DATA

Cory A. Wolff *, Michael M. Bell, Wen-Chau Lee, and Huaqing Cai
National Center for Atmospheric Research, Boulder, CO

1. INTRODUCTION

A method for automatic quality control of the data collected by the Electra Doppler Radar (ELDORA) was described in Wolff et al (2009). This method introduced fields that were thought to be good candidates for inclusion in an algorithm to automatically remove noise and non-weather returns. In this paper we discuss the performance of an early version of the algorithm when it is run on data collected by the ELDORA during five field programs.

2. ALGORITHM

The automatic quality control (QC) algorithm is part of an end-to-end process that is being developed to produce near real-time Dual-Doppler syntheses for use in the field, accelerate the post analysis data QC efficiency, and for quick assimilation into numerical weather prediction models. A technique for calculating the navigation corrections (Cai et al, 2011) is the first part of this process, and they are applied to the radar data to remove all aircraft motion. Once the algorithm has removed the noise and bad data the set of sweeps are ready to be input into an automatic Dual Doppler synthesis package, which is being developed separately. The resulting product will contain reflectivity and the wind field for the entire volume.

Currently, the QC algorithm uses four fields or calculations to remove non-weather data. This initial version of the algorithm uses hard thresholds for testing, which don't offer much flexibility. Plans call for the development of a fuzzy logic algorithm with interest maps instead of hard thresholds and the addition of a weather probability field that can be customized by the user.

The first field is the normalized coherent power (NCP), which is a ratio of the power calculated at lag one to the total received power, and is generally very efficient at removing noise. Higher NCP values are measured in valid radar echoes while lower values are associated with noise. In the QC algorithm a threshold of 0.3 is applied, meaning that any gates with NCP less than 0.3 are removed.

Next, the probability of a gate being contaminated by the surface is calculated. This determines the gate where the center of the beam impacts the surface based on its beamwidth and the elevation angle, taking into account the curvature of the earth and the aircraft altitude. Once this gate is known the ground probability for all succeeding gates is set to 1.0. For all preceding gates the probability of ground may be non-zero due to spreading of the beam, which increases with decreasing elevation angle. The ground probability for these gates is calculated based on a Gaussian beam shape and decreases the closer to the aircraft the gate is (Fig. 1). This field will obviously work best for data collected over a flat surface, but with a high-resolution terrain map the algorithm would be able to adapt to data collected in complex terrain as well. In the QC algorithm any gates with a ground probability above 70% are removed.

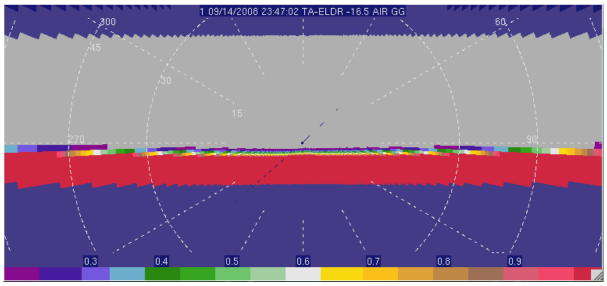

Figure 1. An example of the ground probability field. Areas in gray have no ground probability while areas in red are definitely below the surface.

After the likely ground has been removed the next field calculated is the ratio of spectral width to reflectivity. This field is based on the assumption that areas of high spectral width are associated with non-weather echo, unless the reflectivity is also high as in heavy precipitation. A more complete description is given in Wolff et al (2009) but the lower the ratio is the more likely that a given echo is weather. Gates with a ratio above 0.6 are removed in this version of the algorithm.

The final step in the QC algorithm is to remove any speckles that might be left after the thresholds have been applied. A speckle is a small area of

---

* *Corresponding author address:* Cory A. Wolff, NCAR/RAL, P.O. Box 3000, Boulder, CO 80307; e-mail: cwolff@ucar.edu

spurious echo away from other features and is usually leftover noise. A despeckling routine has been written that is applied along the beam and for a constant gate on adjacent rays. The number of gates that constitute a speckle can be defined, which means that any number of consecutive gates smaller than this threshold that contain echo are removed. The speckle threshold has been set to five for testing purposes.

An example of the results of the QC algorithm is shown in Fig. 2c along with the original, baseline, and manually edited reflectivity fields (Figures 2a, 2b, and 2d, respectively). The automatic QC keeps all of the main echo regions but removes some areas around the edges. The baseline field, which has some basic noise and ground removal applied (see Section 3), still has a large amount of bad data including second trip echo, sidelobes, a reflectivity “ring”, and many speckles. The automatically QCed data qualitatively resembles the manually edited data, and was quantitatively verified using dichotomous forecast metrics described in the next two sections.

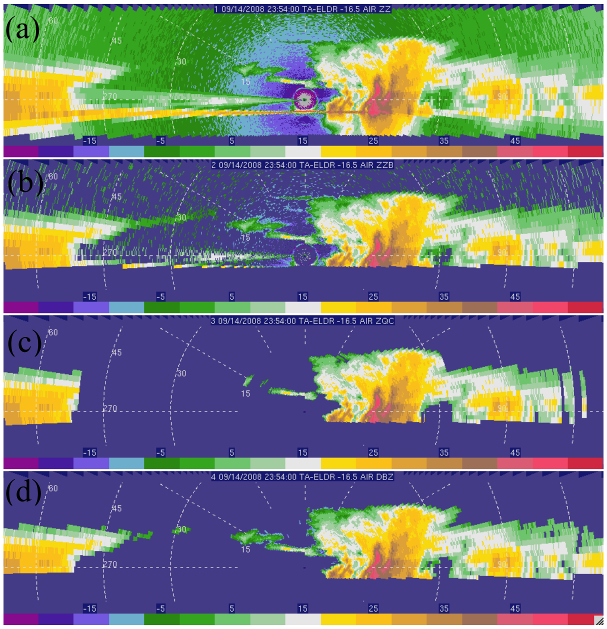

Figure 2. The original reflectivity field (a). The reflectivity field after basic noise removal (b). The reflectivity field after the automatic QC has been run on it (c). The manually edited (truth) reflectivity field (d).

3. VERIFICATION METHODOLOGY

The verification dataset was compiled from five different field programs that involved a variety of different weather conditions and is summarized in Table 1. The field program name is included as an aide to distinguish between cases in the plots and discussion in Section 4.

Field Program	Date	Description
VORTEX	16 May 1995	Tornadic supercell
IHOP	11 Jun 2002	Pre-convective boundary layer
BAMEX	23 Jun 2003	Mesocyclone in MCS
RAINEX	22 Sep 2005	Mature hurricane
T-PARC / TCS08	14 Sep 2008	Pre-depression tropical convection

Table 1. The five cases that the QC algorithm was tested on.

The data in each case have been manually edited and those fields are used as the truth to which the results of the automatic QC algorithm will be compared. We are not making changes to the reflectivity or velocity fields so the verification will be concerned only with comparing gates and determining the number of hits and misses produced by the automatic QC. The QC and manual fields are treated as dichotomous values where each gate in the sample is defined as weather or non-weather. A 2x2 contingency table can then be constructed from these values so that performance statistics can be calculated. A gate where both the QC and manually edited field have weather data are considered a correct yes result, whereas if both consider the gate to be bad it is a correct no result. If a gate in the field from the automatic QC has a weather echo and the manual field is bad then it is considered a false alarm while if it is the opposite then it is considered a miss.

A typical sweep of data from ELDORA contains much more bad echo than good (compare Fig. 2a to 2d). In order to get meaningful statistics some very basic noise removal techniques were applied to the raw data before the fields were run through the QC algorithm. Without this first step it would be difficult to see any differences in the verification as the signal would be dominated by gates that are obviously bad (e.g. those well below the ground or with very low NCP values). To gauge the effectiveness of the QC algorithm we want to evaluate its performance on those gates that are more difficult to distinguish between weather and noise. These baseline fields will also be useful for determining how well new fields perform in the removal of bad data as they provide a starting point for which to compare field distributions for weather and non-weather gates.

4. VERIFICATION RESULTS

From the contingency tables for each of the five cases some basic statistics can be calculated that show how well the current algorithm is performing when compared to hand edited sweeps. The following descriptions are taken from Murphy and Winkler (1987) and Doswell et al (1995). The most obvious measure is the probability of detection for yes and no events (PODy and PODn). PODy is also known as hit rate and is calculated by dividing the number of hits by the total hits and misses. PODn is calculated by dividing the number of correct non-weather gates by the total non-weather gates in the sample. The false alarm ratio (FAR) measures the fraction of time that the automatic QC kept an echo that should have been removed. The proportion of correct (PC) tells the fraction of all gates from the automatic QC algorithm that was correct, treating both weather and non-weather hits equally. Finally, the true skill score (TSS) is a simple measure of the success of an algorithm and is calculated by subtracting the probability of false detection (POFD), which is the probability of false alarms given that the event did not occur, from PODy. While all of the other measures are presented on a 0 – 1 scale the TSS is a value between -1 and 1 with 1 representing perfect skill, 0 representing no skill, and -1 representing negative skill (i.e. all yes events are classified as no events and vice versa).

Figure 3 shows PODy, PODn, and TSS for all of the cases while Figure 4 shows the PC and FAR. All of the cases had a high PODn, but this measure is likely skewed somewhat by the baseline fields, which still leave a large amount of obviously bad gates. The PODy for the tropical cases (RAINEX and T-PARC/TCS08) is very good, with values of 0.9 and 0.82, respectively. The widespread convection was fairly easy to distinguish from the non-weather data. A larger portion of their gates had data in them because the aircraft was flying inside of large-scale convective regions with good echo on both sides. The BAMEX and VORTEX cases (PODy of 0.69 and 0.71, respectively) were over land, albeit a relatively flat surface, but the ground gate calculation was still likely to include some areas of ground contamination. Also, these flights were sampling individual storms and kept them to one side of the aircraft. In the BAMEX case the aircraft was flying along the leading edge of an MCS, while for VORTEX it was flying around a supercell thunderstorm. This reduced the number of good gates in a given volume.

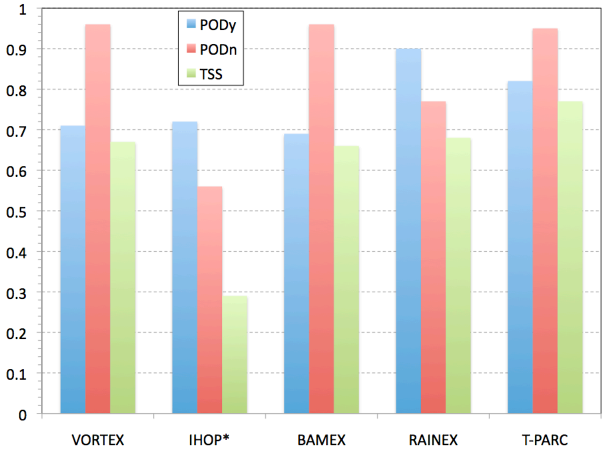

Figure 3. PODy (blue), PODn (red), and TSS (green) for the five cases tested.

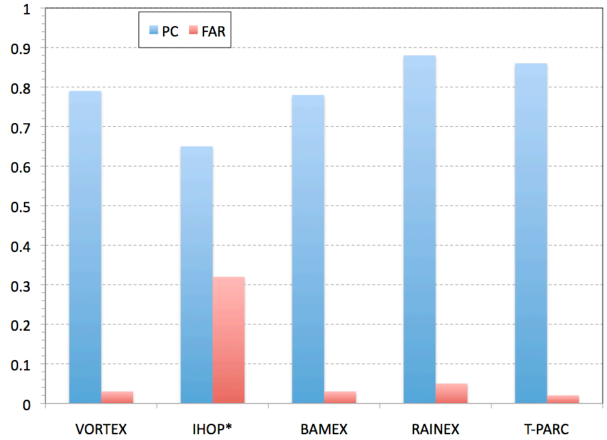

Figure 4. PC (blue) and FAR (red) for the five cases tested.

The convective boundary layer case (IHOP) has an asterisk next to it in both Figures to indicate that the QC algorithm was run for this case without the spectral width to reflectivity ratio. This case contains weak echoes with small values of spectral width, but the ratio was high for all gates and almost everything in the sweep was removed. This resulted in a very high PODn, but a PODy near zero and a negative skill score. Removing the ratio provided a similar PODy to the other continental cases but a much lower PODn, demonstrating the difficulty the algorithm has differentiating between good and bad echo at low reflectivity.

The TSS values generally follow PODy with the least skill found for the convective boundary layer. The low TSS for the boundary layer case is not surprising based on its high POFD (0.44). The only exception is the mature tropical system in the RAINEX case, which has the highest PODy but is similar to the continental thunderstorm cases (VORTEX and BAMEX) in TSS due to an unexpectedly high POFD (not shown). The RAINEX case had a POFD of 0.23 compared with 0.07 or lower for all of the other cases with active weather. The reasons for this are still unclear.

The PC for the all of the cases is 79%, and is even higher when only those cases with active weather are included, with 84% of gates correctly classified. The FAR is also low in every case except for the convective boundary layer. This means that the algorithm does not appear to be retaining too much bad data for the active weather cases, but that a higher proportion of non-weather gates are being misclassified in the boundary layer.

The active weather cases have a much larger percentage of gates that contain weather echo than the boundary layer one. Therefore, the statistics for the latter case are bound to be more sensitive to changes in the algorithm. The algorithm can remove good weak echo regions in the active weather cases without being penalized by the statistics or affecting the final synthesis much because of the overwhelmingly large numbers of strong echo gates. Statistics calculated for active weather cases that only include gates with weak echoes would likely be similar to the convective boundary layer case.

5. SUMMARY & FUTURE WORK

The algorithm as it stands does a good job of removing noise without eliminating too much weather echo. Improvements are still necessary. As the IHOP case showed, certain environments do not lend themselves well to the current algorithm. Customization of the algorithm by the user might allow for better results in those situations. Future plans for the algorithm involve replacing hard thresholds with interest maps and weighting functions. It will also include a new field called Probability of Weather that will assign a probability to each gate that it contains valid radar echo and not noise or spurious echo. The user will then be able to set his or her own threshold for the probability above which they would like to keep data, which may vary depending on the user, their interests, and the case being studied.

The current algorithm is already partially customizable. The core QC and input/output routines are written in C++, but a scripting interface has been added using Ruby that allow changes to be made to the thresholds and fields used without having to recompile or be familiar with the internal workings of the QC program. The code and instructions for installation are also freely available through an online repository called Github at https://github.com/mmbell/Airborne-Radar-QC.

6. REFERENCES

Cai, H., W. C. Lee, C. A. Wolff, and M. M. Bell, 2011: Study of airborne Doppler radar navigation correction using a variational technique. *35th Conf. on Radar Meteorology*, Pittsburgh, 25 – 29 Sep.

Doswell, C.A., III, R. Davies-Jones, and D.L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576-585.

Murphy, A.H. and R.L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.*, **115**, 1330-1338.

Wolff, C. A., M. M. Bell, and W. C. Lee, 2009: Automatic quality control of airborne Doppler radar data. *34th Conf. on Radar Meteorology*, Williamsburg, VA, 5 – 9 Oct.

Acknowledgments. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.