A VALIDATION OF RADAR REFLECTIVITY QUALITY CONTROL METHODS

Angela Fritz
NWC REU
Valparaiso University, Valparaiso, Indiana

V. Lakshmanan
CIMMS/NSSL, Norman, Oklahoma

Travis Smith
CIMMS/NSSL, Norman, Oklahoma

Beth Clarke
CIMMS/NSSL, Norman, Oklahoma

Eddie Forren
CIMMS/NSSL, Norman, Oklahoma

National Weather Center Research Experience for Undergraduates
Final Project

4 August 2005

Corresponding author address:
Angela Fritz
Valparaiso University Department of Meteorology
1809 Chapel Drive
Valparaiso, IN 46383
Email: angela.fritz@valpo.edu

Abstract

Quality control of radar reflectivity data is essential for accurate precipitation forecasts and products of algorithms that require clean data. Radar data is frequently contaminated with non-precipitation echoes. Quality control methods should be able to remove a majority of these non-precipitation echoes as well as retain all of the actual precipitation. In this validation study, three quality control methods are tested on sixteen independent radar cases. These cases included non-precipitation such as anomalous propagation, biological return, and electronic interference as well as actual precipitation including weak and strong convection and stratiform rain events. The data was analyzed and then hand-truthed to remove the contamination and create what we refer to as the target. The data was then run through the quality control methods and the results from each were scored against the target. Skill scores were calculated to determine which methods excel in the situations that were chosen.

I. Introduction

Many quality control methods for radar reflectivity data from the Weather Surveillance Radar 88 Doppler (WSR-88D) have been created and studied (e.g., Kessinger et al. 2003; Lakshmanan et al. 2003; Zhang et al. 2004). These quality control methods are important for automated applications that rely on clean radar data with only weather-related returns. The quality control methods should be designed to remove radar echoes corresponding to non-meteorological contaminants, including ground clutter, biological return (insects, birds, bats, etc.), anomalous propagation (AP), and electronic interference. Downstream algorithms are affected by this contamination, but results are improved when quality control methods are applied. For example, quality control of the radar reflectivity data can mitigate the problem of detecting mesocyclones in areas where there are no storms (Mazur et al. 2003), and can reduce error in rainfall estimates (Fulton et al. 1998, Kessinger et al. 2003).

This study focuses on the validation of some methods of quality control: the Radar Echo Classifier (REC) (Kessinger et al. 2002), a method created by the Worldwide Integrated Sensors for Hydrometeorology group (WISH QC) using horizontal and vertical reflectivity structure (Zhang et al. 2004) and the Quality Control Neural Network (QCNN) (Lakshmanan 2003). The REC is currently in operation at the National Weather Service, and the QCNN was designed as a part of the Warning Decision Support System - Integrated Information (WDSS-II) (Hondl, 2002).

The validation will be limited to sixteen cases of meteorological and nonmeteorological returns. These cases include biological returns, ground clutter, electronic interference, such as test patterns and interference from other radars in the area, as well as examples of good radar data such as convective and stratiform precipitation.

The rest of this report is organized as follows. Section 2 is a brief description of each QC method used in this study. In Section 3 we describe the data and method used to conduct the validation study. In Section 4, we discuss the results, and conclude in Section 5.

II. Quality Control Methods

a. REC

The REC (Kessinger et al. 2003) is currently in operation in the Open Radar Products Generator (ORPG). It was designed as a part of the AP Clutter Mitigation Scheme (e.g., Kessinger et al, 2001 and 2002; Ellis et al. 2003). The scheme's purpose is to improve radar-derived rainfall estimates by removing contaminants in the radar data, specifically AP. As a part of the scheme, the REC determines which echoes are precipitation, and removes those that are not. It was built and trained specifically for AP and ground clutter.

The REC uses reflectivity, radial velocity, and spectrum width to classify radar echoes by three algorithms: the AP detection algorithm (APDA), the precipitation detection algorithm (PDA), and the insect clear air detection algorithm (ICADA). The REC relies on a feature generator and a fuzzy logic engine to determine the types of echoes, and removes the echoes that are not precipitation. It uses this method to quality control the data out to a range of 230 km, and retains all the original data from 230 km to 460 km.

The REC Build 8 version was used in this validation study.

b. WISH QC

The WISH QC (Zhang et al. 2004a) is implemented in the National Radar Mosiac and QPE Project (Zhang et al. 2004b) and uses both horizontal and vertical reflectivity structure to perform the quality control. It operates under the assumption that precipitating and non-precipitating echoes have different vertical reflectivity structures. It uses this vertical reflectivity information, which is computed by height rather than by radar tilt, to determine which echoes are not precipitation. This method was created to mitigate precipitation uncertainty caused by beam spreading.

The WISH QC has four main steps. First, it runs the raw reflectivity data through a noise filter to remove minor speckle. Then the texture of reflectivity described in (Kessinger et al. 2003) is used to determine the horizontal reflectivity structure. In the next step, the upper reference tilt is used to examine vertical continuity by a parameter called the vertical difference of reflectivities. This difference is calculated for each range gate by subtracting an upper reference gate from the gate being quality controlled. A threshold value for this difference is applied and the echo is removed if larger than the threshold value. The highest tilt is quality controlled by reflectivity structure alone.

c. QCNN

The QCNN (Lakshmanan 2003) is implemented in the WDSS-II system (Lakshmanan 2004). It uses reflectivity, velocity, and spectrum width as well as horizontal reflectivity structure, SPIN (Kessinger 2003), SIGN (Kessinger 2003), and echo size to determine which echoes are and are not precipitation. The QCNN also considers maximum vertical reflectivity and the maximum reflectivity in the neighborhood of the gate in question in its analysis. It uses this information to calculate a precipitation confidence on a scale from 0 to 1, 0 being the least confident and 1 being the most confident. Any gate with less than 0.4 precipitation confidence is considered non-precipitation and is removed.

The QCNN version 20050620 was used in this validation study.

III. Data and Validation Method

To perform the quality control validation, 16 independent radar cases were chosen to include both good and bad data. The bad data that were selected included non-precipitating echoes that could be mistaken by an automated algorithm to be precipitation. The good data, while not immaculate, are lacking these types of returns.

Of the 16 cases, 8 included what we have defined to be bad data as well as precipitation in various forms. The data includes two AP cases, two biological cases (one of which includes bats), three electronic interference cases, and a general ground clutter case ("speckle clutter"). . All of these features are those that a quality control method would be expected to remove without affecting the actual precipitation in any meaningful way. The remaining 8 cases were chosen as good cases and lack significant bad data. Table 1 describes these cases in detail.

The cases were originally Level II data that were then converted to NetCDF format to be viewed and analyzed in the WDSS-II display. Using reflectivity and velocity information in the display, the data set was hand-truthed by drawing polygons around the areas of non-precipitation. Polygons in the good cases were drawn outside the areas of any reflectivity. The target was created by removing the polygons from the data, and is what these cases would be expected to look like after quality control.

The target was used to score each quality control method range gate by range gate. Thresholds for the scoring were set at 0, 10, 30, and 40 dBZ reflectivity, and 0 and 25 kg m^{-2} vertically integrated liquid (VIL). These thresholds were used to score how each method did on higher reflectivity and VIL values. Each method was scored individually.

In the scoring, each gate was a hit, miss, false alarm, or null. A hit was a gate in which there was precipitation and the method retained the gate. A miss was a gate that contained precipitation and the method removed it. A false alarm was a gate that contained non-precipitation but the method kept it, and a null was a gate in which there was non-precipitation and the method removed it. By defining a hit as retained precipitation, this method of scoring emphasizes the importance of retaining good data. Although a quality control method should be trained to remove bad data, it is just as important that it recognizes actual precipitation, as well.

From the hit, miss, false alarm and null information, the probability of detection, false alarm rate, critical skill indexes, and Heidke Skill scores were calculated. The probability of detection (POD) is the fraction of the "good echo" that is retained. Therefore, the POD measures how well the method recognizes actual precipitation. The false alarm rate (FAR) is the ratio of bad echo to good echo in the product generated by the quality control method. It measures how well the method removes non-precipitation. The critical skill index (CSI) is a combination of POD and FAR with CSI = 1 being a perfect score. The Heidke Skill Index (HSS) is another way to combine POD and FAR, as well as take into account the number of null cases – events where there was no precipitation and the method left that range gate alone. This null was calculated as the number of gates that had an echo in the range (-infinity, 0) dBZ in both the original and the quality-controlled reflectivity composite fields (or VIL = 0 in the case of the VIL fields). These skill scores were computed on all 16 cases using a "leave one out" statistic, also called jackknifing (Efron and Tibshirani 1997). This method of calculation was used to estimate the standard error of each score.

The methods were scored for all 16 cases as well as after breaking down the results into subcategories of AP, biological, and electronic interference.

IV. Results and Discussion

The mean skill scores achieved for each QC method as well as the 95% confidence interval assuming a normal distribution are presented in Table 2. Scores for the biological, electronic interference, and AP cases are located in Table 3. The measures of skill on the Reflectivity Composite product can serve as a proxy for visual quality, while the measures of skill on the VIL product can serve as a proxy for the effect that the quality control can have on warning decision algorithms.

QCNN outperforms the other methods at the 10 dBZ and greater thresholds (Table 2). Its lowest scores are in the 0 to 10 dBZ range. It struggles to remove all biological contaminants, but outperforms the WISH QC and the REC in retaining all actual precipitation. An example of this is shown in Figure 1. While QCNN does not remove all of the low reflectivity biological contamination close to the radar, it retains all of both storms despite the high gradient around the edges. Overall, QCNN is more likely to get higher reflectivity values correct, whether they are actual precipitation retained or non-precipitation removed.

The WISH QC outperforms QCNN in the 0 and 10 dBZ thresholds, but it received high FAR scores at those values, as well. It also does not do as well above 30 dBZ. An example of this is shown in Figure 1. While WISH successfully removes all of the contamination from the radar interference as well as the biological contamination around the radar, it removes large portions of precipitation from the high values within the storm. Another example of this is shown in Figure 3. Here, high reflectivity values have been removed from the outer edge of the storm core. Removal of such high values has a heavily weighted negative impact on the WISH QC's POD and CSI scores.

The REC struggled with correctly identifying higher reflectivity echoes. The exception to this is in the AP cases. Figure 3 shows that the REC was able to remove the high AP reflectivity but retained the low reflectivity non-precip. In most of the other cases the REC left the original data as-is. This result is to be expected because it was designed specifically for AP and ground clutter and not for other contaminants such as biological or electronic interference.

It is important to understand these results from the point of view of the automated applications that use these reflectivity data. Severe weather algorithms and precipitation estimate algorithms require that the input is “clean” – that none of the data contain information that is not weather-related. The scores for these algorithms give an idea of what the results would be should they be implemented in operation where all kinds of contaminants are present, not just AP and ground clutter. A quality control method used in operation needs to take into account all kinds of bad data.

IV. Conclusion

A quality control method for radar reflectivity data is necessary in operation where automated applications rely on clean data with only weather-related returns. A validation study was performed using the Radar Echo Classifier, the system used by the Worldwide Integrated Sensors for Hydrometeorology group, and the Quality Control Neural Network. 16 independent radar cases containing a variety weather related returns and non-weather contaminants were chosen. Each of the scans was hand-truthed by an expert and the target was created. The cases were then run through each of the quality control techniques and were scored against the target.

The methods were scored gate by gate. Each gate was a hit (method retained actual precipitation), miss (method removed precipitation), false alarm (method retained non-precipitation), or a null (there was no information and the method left it as-is). Using this, this probability of detection, false alarm rate, critical skill index, and Heidke Skill scores were calculated for each method.

In most cases, the REC left the reflectivity data as-is, with the exception of the AP case (KLBB 1995-10-05). The WISH QC performed well, but had some unfavorable results due to the removal of very high dBZ precipitation in a few of the cases. The QCNN outperformed the other methods overall, with its strength lying in the ability to correctly identify actual precipitation more often, while still being able to remove most non-weather contaminants.

VI. Acknowledgements

Funding for this research was provided by the National Weather Center Research Experience for Undergraduates (NWC REU), which is funded by the National Science Foundation grant 00997651.

VII. References

Efron, B. And R. Tibshirani, 1997: Improvements on cross-validation: The .632+ bootstrap method., *J. of the American Statistical Association,* **92**, 548-560.

Ellis, S., C. Kessinger, T.D. O'Bannon, and J. VanAndel, 2003: Mitigating ground clutter contamination in the WSR-88D. Preprints CD, *19th Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr.*, Amer. Meteor. Soc., Long Beach, CA.

Fulton, R.A., J.P. Breidenback, D.J. Seo, D.A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Weather and Forecasting*, **13**, 377-395.

Hondl, K., 2002: Current and planned activities for the warning decision support system-integrated information (WDSS-II). *21st Conference on Severe Local Storms*, Amer. Meteo. Soc., San Antonio, TX.

Kessinger, C., S. Ellis, and J. Van Andel, 2001: NEXRAD data quality: The AP clutter mitigation scheme. *30th Int'l Conf. on Radar Meteorology*, Amer. Meteor. Soc., Munich, Germany.

Kessinger, C., S. Ellis, and J. Van Andel, 2002: NEXRAD data quality: An update on the AP clutter mitigation scheme. *18th Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr.*, Amer. Meteor. Soc., Orlando, FL.

Kessinger, C., S. Ellis, and J. Van Andel: 2003, The radar echo classifier: A fuzzy logic algorithm for the WSR-88D. *19th Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr.*, Amer. Meteor. Soc., Long Beach, CA.

Lakshmanan, V., K. Hondl, G. Stumpf, and T. Smith, 2003: Quality control of WSR-88D data. *31st Radar Conference*, Amer. Meteor. Soc., 2003, Seattle, WA.

Mazur, Rebecca, 2003: Quality control of radar data to improve mesocyclone detection. *20th Int'l Conf. on Inter. Inf. Proc. Sys. (IIPS) for Meteor., Ocean., and Hydr.* Amer. Meteor. Soc., Seattle, WA, CD-ROM, P1.2.

Zhang, J., K. Howard, W. Xia, C. Langston, S. Wang, and Y. Qin, 2004: Three-dimensional high-resolution national radar mosaic. *11th Conference on Aviation, Range, and Aerospace Meteorology*, Amer. Meteor. Soc., Hyannis, MA.

Zhang, J., S. Wang, and B. Clarke, 2004: WSR-88D reflectivity quality control using horizontal and vertical reflectivity structure. *11th Conference on Aviation, Range, and Aerospace Meteorology*, Amer. Meteor. Soc., Hyannis, MA.

Radar	Scan Date	Scan Time	Type
KLBB	1995-10-05	01:44:29	AP
KTLX	1996-06-16	14:16:24	AP With Precip
KUEX	2002-06-13	02:31:26	Strong Convection
KICT	2003-04-19	20:32:04	Convection
KDVN	2003-05-01	04:36:06	Convection
KAMA	2003-05-03	21:50:10	Biological
KAMA	2003-05-04	01:05:34	Interference
KTLX	2004-04-30	22:31:56	Strong Convection
KFDR	2004-07-16	02:59:26	Biological
KINX	2004-08-17	11:21:55	Speckle Clutter
KCYS	2004-09-21	00:57:01	Sun Ray
KFFC	2004-09-27	16:15:49	Stratiform Precip
KICT	2004-10-11	08:01:37	Stratiform Precip
KILX	2004-10-26	23:56:51	Stratiform Precip
KHDX	2005-05-28	10:34:44	Test
KDGX	2005-06-07	22:15:02	Weak Convection

TABLE 1. The 16 independent volume scans used in this validation study.

Product	Data Range	Measure	No QC	REC	WISH QC	QCNN
Composite Refl.	> 0 dBZ	POD	1 ± 0	0.96 ± 0.0031	0.92 ± 0.0079	0.88 ± 0.088
		FAR	0.39 ± 0.06	0.4 ± 0.06	0.27 ± 0.084	0.02 ± 0.0072
		CSI	0.61 ± 0.06	0.59 ± 0.057	0.69 ± 0.074	0.86 ± 0.011
		HSS	0.89 ± 0.02	0.88 ± 0.019	0.92 ± 0.019	0.98 ± 0.0016
Composite Refl.	> 10 dBZ	POD	1 ± 0	0.94 ± 0.0023	0.96 ± 0.0031	0.92 ± 0.0039
		FAR	0.32 ± 0.071	0.32 ± 0.073	0.26 ± 0.086	0.02 ± 0.007
		CSI	0.68 ± 0.071	0.66 ± 0.069	0.72 ± 0.081	0.96 ± 0.0083
		HSS	0.93 ± 0.017	0.93 ± 0.016	0.94 ± 0.018	0.99 ± 0.0011
Composite Refl.	>30 dBZ	POD	1 ± 0	0.92 ± 0.0065	0.97 ± 0.005	1 ± 0.00029
		FAR	0.08 ± 0.02	0.09 ± 0.011	0.032 ± 0.0086	0 ± 0.00057
		CSI	0.92 ± 0.02	0.84 ± 0.014	0.76 ± 0.014	1 ± 0.00072
		HSS	1 ± 0.00064	0.99 ± 0.00052	1 ± 0.0008	1 ± 0
Composite Refl.	> 40 dBZ	POD	1 ± 0	0.88 ± 0.0088	0.93 ± 0.0084	1 ± 0
		FAR	0.09 ± 0.023	0.1 ± 0.0074	0.011 ± 0.00088	0 ± 0.00039
		CSI	0.91 ± 0.023	0.8 ± 0.013	0.92 ± 0.0082	1 ± 0.00072
		HSS	0.97 ± 0.0091	0 ± 0.00018	1 ± 0	1 ± 0
VIL	> 0 kg/m^3	POD	1 ± 0	0.9 ± 0.0078	0.96 ± 0.0048	1 ± 0.00084
		FAR	0.47 ± 0.16	0.49 ± 0.15	0.48 ± 0.17	0 ± 0.00053
		CSI	0.53 ± 0.16	0.48 ± 0.13	0.51 ± 0.16	1 ± 0.0011
		HSS	0.97 ± 0.0091	0.97 ± 0.0085	0.97 ± 0.01	1 ± 0
VIL	> 25 kg/m^3	POD	1 ± 0	0.76 ± 0.026	0.83 ± 0.05	1 ± 0.00075
		FAR	0 ± 0.0022	0.19 ± 0.025	0.013 ± 0.0037	0 ± 0.0022
		CSI	1 ± 0.0022	0.65 ± 0.033	0.82 ± 0.045	0.99 ± 0.0027
		HSS	1 ± 0	1 ± 0	1 ± 0.00033	1 ± 0

TABLE 2. The results of this validation study in terms of probability of detection (POD), false alarm rate (FAR), critical skill index (CSI), and Heidke Skill score (HSS). Scores achieved for each QC method as well as the 95% confidence interval assuming a normal distribution are shown.

Biological

Data Range	No QC	REC	WISH QC	QCNN
0 dBZ	0.854	0.814	0.896	0.852
10dBZ	0.935	0.891	0.940	0.923
30 dBZ	0.998	0.876	0.962	0.996
40 dBZ	0.999	0.818	0.921	0.997
0 kg/m² VIL	1.000	0.876	0.942	0.994
25 kg/m² VIL	1.000	0.701	0.813	0.981

AP

Data Range	No QC	REC	WISH QC	QCNN
0 dBZ	0.886	0.864	0.912	0.910
10dBZ	0.939	0.929	0.957	0.985
30 dBZ	0.945	0.871	0.964	0.998
40 dBZ	0.919	0.811	0.922	0.998
0 kg/m² VIL	0.990	0.866	0.944	0.994
25 kg/m² VIL	1.000	0.690	0.806	0.979

Electronic Interference

Data Range	No QC	REC	WISH QC	QCNN
0 dBZ	0.718	0.686	0.697	0.692
10dBZ	0.736	0.704	0.722	0.731
30 dBZ	0.972	0.855	0.940	0.973
40 dBZ	0.990	0.815	0.919	0.990
0 kg/m² VIL	0.519	0.863	0.487	0.516
25 kg/m² VIL	0.996	0.647	0.817	0.981

TABLE 3. Critical skill indexes for biological, AP, and electronic interference cases. No significant scores are indicated as there were not enough cases in each category to do so.

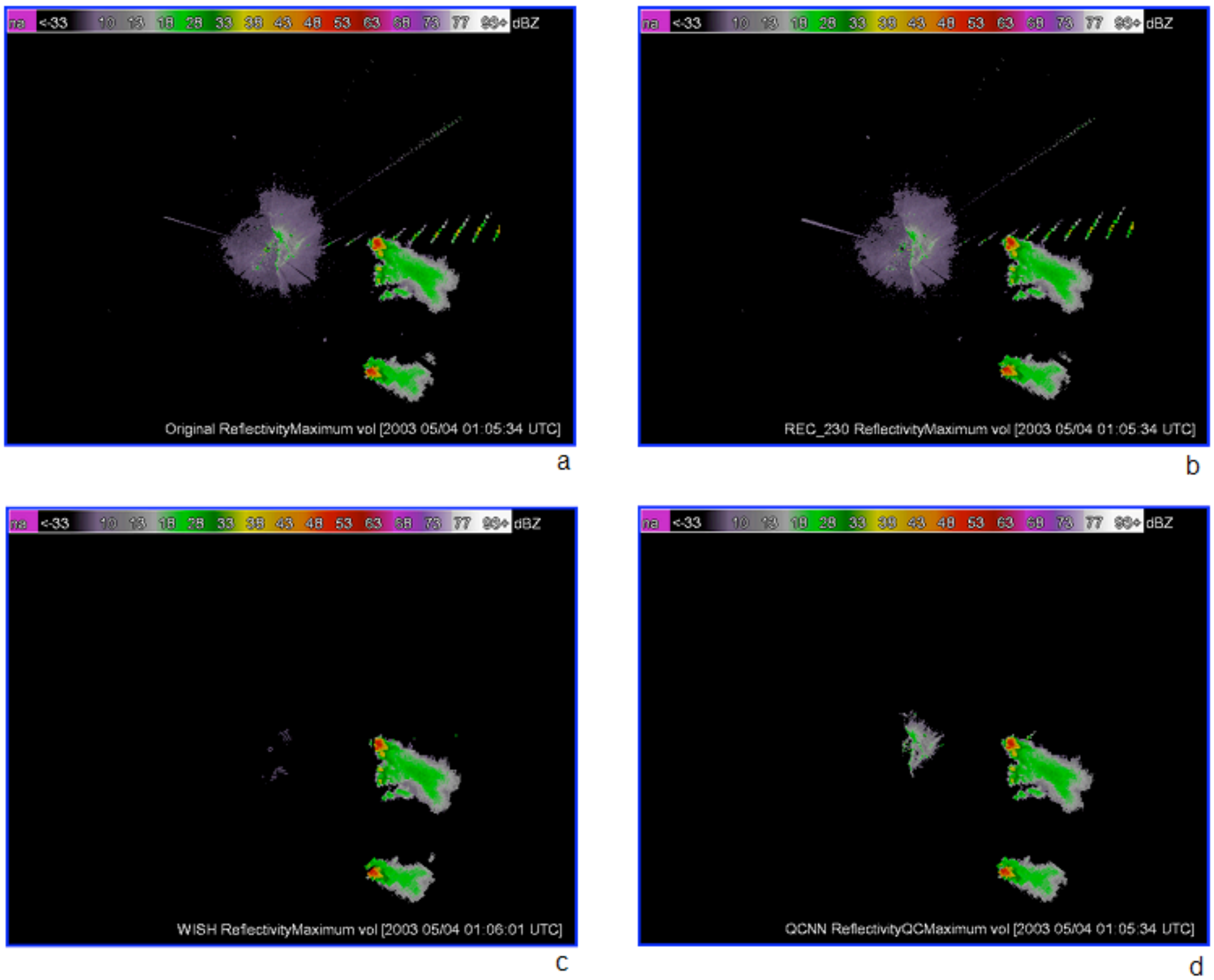

FIGURE 1. Performance of quality control methods in this validation study for KAMA 2003-05-04. (a) No quality control. (b) Result from REC. No visible quality control has been performed. (c) Result from WISH QC. All of the electronic interference has been removed, as well as most of the biological near the radar. WISH QC also removed large portions of high reflectivities within the storms. (d) Results from QCNN. The majority of the electronic and biological contaminants have been removed, and all the actual precipitation is retained.

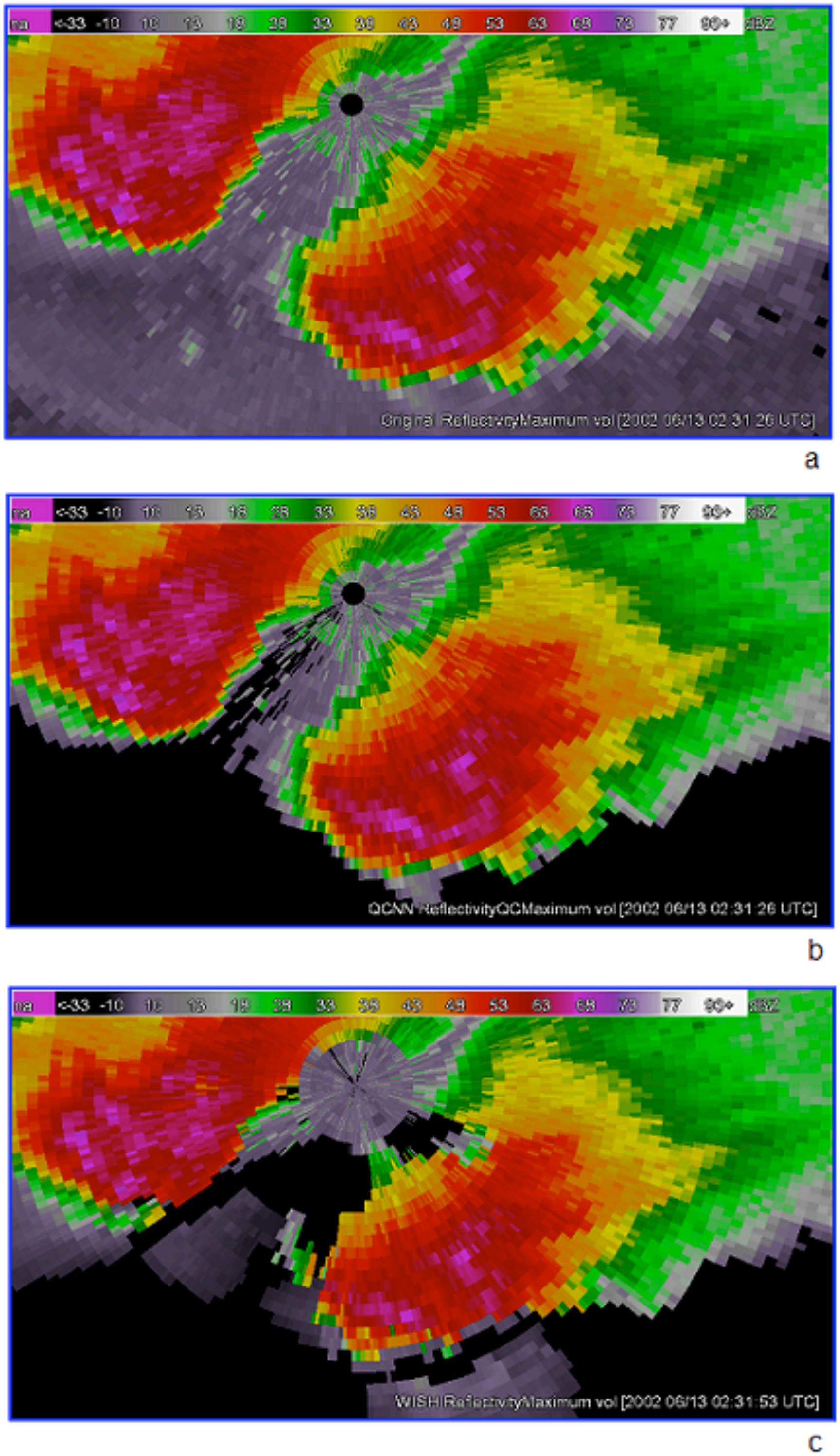

FIGURE 2. Performance of QCNN and WISH QC in this validation study for KAMA 2002-06-13. (a) No QC and REC results. REC did not remove any gates in this case. (b) Results from QCNN. Low reflectivities close to the radar were removed and all of the intense values of the storm were left. (c) Results from WISH QC. Low reflectivities close to the radar have been removed, as well as some higher values from the storm.

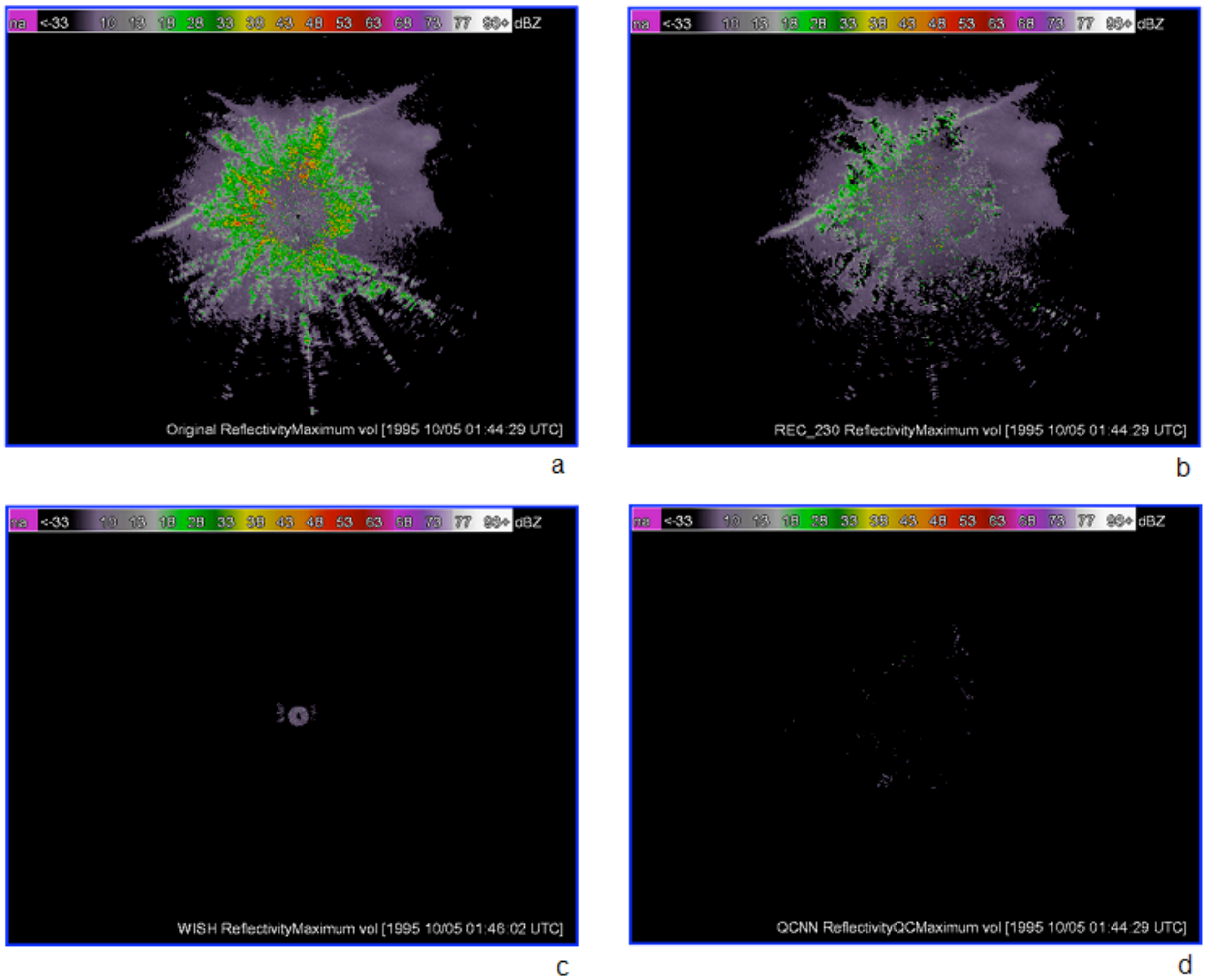

FIGURE 3. Performance of quality control methods in this validation study for KLBB 1995-10-05. (a) No quality control. (b) Results from the REC. Most of the high reflectivity AP has been removed, but the weak intensity AP is left. (c) Results from WISH QC. Most of the AP has been removed. (d) Results from QCNN. Most of the AP has been removed.