Evaluation of Anomalous Propagation Echo Detection in WSR-88D Data: A Large Sample Case Study

WITOLD F. KRAJEWSKI AND BERTRAND VIGNAL

Iowa Institute of Hydraulic Research, The University of Iowa, Iowa City, Iowa

26 June 2000 and 12 December 2000

ABSTRACT

A method of detecting anomalous propagation echo in volume scan radar reflectivity data is evaluated. The method is based on a neural network approach and is suitable for operational implementation. It performs a classification of the base scan data on a pixel-by-pixel basis into two classes: rain and no rain. The results of applying the method to a large sample of Weather Surveillance Radar-1988 Doppler (WSR-88D) level II archive data are described. The data consist of over 10 000 volume scans collected in 1994 and 1995 by the Tulsa, Oklahoma, WSR-88D. The evaluation includes analyses based on radar data only and on various comparisons of radar and rain gauge data. The rain gauge data are from the Oklahoma Mesonet. The results clearly show the effectiveness of the procedure as indicated by reduced bias in rainfall accumulation and improved behavior in other statistics.

1. Introduction

One can safely say that quality control (QC) of radar reflectivity data is the most important step in the overall process of radar rainfall estimation. The crucial aspect of such QC lies in the detection of ground clutter and echoes caused by anomalous propagation (AP) of radar waves (Moszkowicz et al. 1994). Recently, Grecu and Krajewski (2000) proposed a method of AP detection based on a neural network approach. The method classifies the base scan radar reflectivity data into rain or no-rain echo on a pixel-by-pixel basis. With this method, several characteristics of the reflectivity field are computed in the neighborhood of the pixel under investigation. They include pixel velocity and its local variance, height of the highest nonzero echo above the base scan pixel in question, height of the maximum echo above the pixel location, maximum horizontal gradient in an eight-pixel neighborhood, and others. A trained (i.e., calibrated) neural network uses these characteristics as inputs and performs the classification. A unique aspect of the method is the selection of the training dataset required by the neural network approach. The authors advocate selecting only the ''clear cut'' cases, both for rain and no-rain echo. This allows fast and efficient preparation of the training sample and, thus, rapid implementation of the methodology. The selected set is used for random drawing of the training and validation samples. Thus, the training and the validation do not include the challenging cases where AP and rain might be collocated.

The Grecu and Krajewski (2000) methodology includes self-evaluation through repeated resampling and cross-validation. Performance was monitored in terms of the number of misclassified pixels. In this short communication, we expand the evaluation methodology. We include several analyses based on radar data only, as well as various comparisons with rain gauge observations.

2. Summary of the data

We used the same, but somewhat expanded, database as Grecu and Krajewski (2000). The data were collected by the Tulsa, Oklahoma, Weather Surveillance Radar-1988 Doppler version (WSR-88D). In Oklahoma, the rainfall regime is dominated by midlatitude convective systems (Houze et al. 1990). The data cover mostly the warm season months of 1994 and 1995. Since the study of Grecu and Krajewski (2000), we filled several gaps in the radar data and, as a result, had available over 10 000 volume scans (see Fig. 1 for the histogram). These radar data were converted from the Archive level II format (Klazura and Imy 1993) to the efficient format ASCII-RLE (Kruger and Krajewski 1997), allowing the rapid access required for such a large sample study. The rain gauge data we used are from the Oklahoma Mesonet (Brock et al. 1995), with some 49 rain gauges located within the Tulsa radar domain.

Corresponding author address: Witold F. Krajewski, Iowa Institute of Hydraulic Research, The University of Iowa, Iowa City, IA 52242. E-mail: witold-krajewski@uiowa.edu

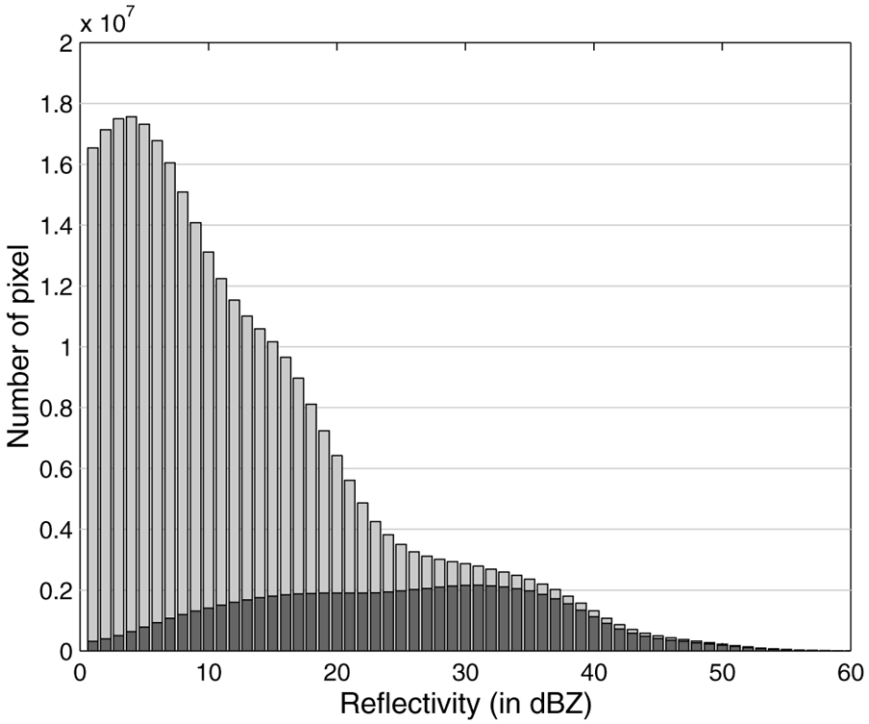


FIG. 1. Histogram of the radar reflectivity data used in the study. The light shaded region corresponds to the data before QC. The dark shaded region corresponds to the data after QC.

3. Results

We used the neural network trained by Grecu and Krajewski (2000), applying the same network configuration for the entire dataset. Our analyses are divided into two parts: that based on radar data only and that based on both radar and rain gauge observations.

a. Radar-only analyses

First, let us consider the probability of detection (POD) of an echo stronger than a certain threshold, calculated on a pixel-by-pixel basis. Clearly, the possible values range from 0, if no echo is ever detected at the given pixel (this may happen if the view of the pixel is completely blocked by an object such as a building or mountain), to 1, if an echo is always detected (as in the case of reflections off a mountain). If only rain-caused echoes were detected, the expected range of values would be around 0.06 (based on rain gauge data analysis), but the exact numbers are unknown. Furthermore, if the rainfall under the radar umbrella was climatologically and statistically homogeneous, and we had a sufficiently large sample, the range of values of the POD would be very narrow—a single spike in the limit.

How does the POD pattern look for the Tulsa WSR-88D? Figures 2 and 3 provide the answer. In Fig. 2 we show the effect of various thresholds (T) on the pattern of POD. Clearly, for $T = 0$ dBZ the pattern displays circular artifacts, which result from a combination of ground clutter effects (near the radar), and AP farther out. The slight shift of the pattern toward the southeast reflects the rainfall climatology of the region. As the threshold increases, the POD pattern becomes more uniform. This is because, by using thresholds, we eliminate much of the ground clutter and AP. However, we also eliminate some of the rainfall. For example, simple analysis of the histograms (Fig. 1) indicates that for $T = 20$ dBZ we may be removing rainfall while not eliminating all of the AP. In fact, the threshold eliminates 9% of the area-averaged rainfall accumulation, if we assume the original data are correct, and fails to remove 6% of rainfall-equivalent AP. Based on the corrected data (i.e., quality controlled) the threshold eliminates 3% of rainfall. Thus, while using thresholds may be considered the simplest QC method, it introduces the risk of distorting rainfall climatology.

Now let us compare the corresponding patterns of POD calculated from quality-controlled data (Fig. 3), that is, data obtained after eliminating the AP echo. Clearly, the POD is now consistently lower and the patterns more uniform. If, in addition, we consider the corresponding histograms for both sets of the POD patterns (Fig. 4), it becomes clear that the applied QC is more effective than simple thresholds. For the quality-controlled data, there is little effect of applying thresholds on the histogram of POD, which indicates that most false echoes were removed by the QC procedure.

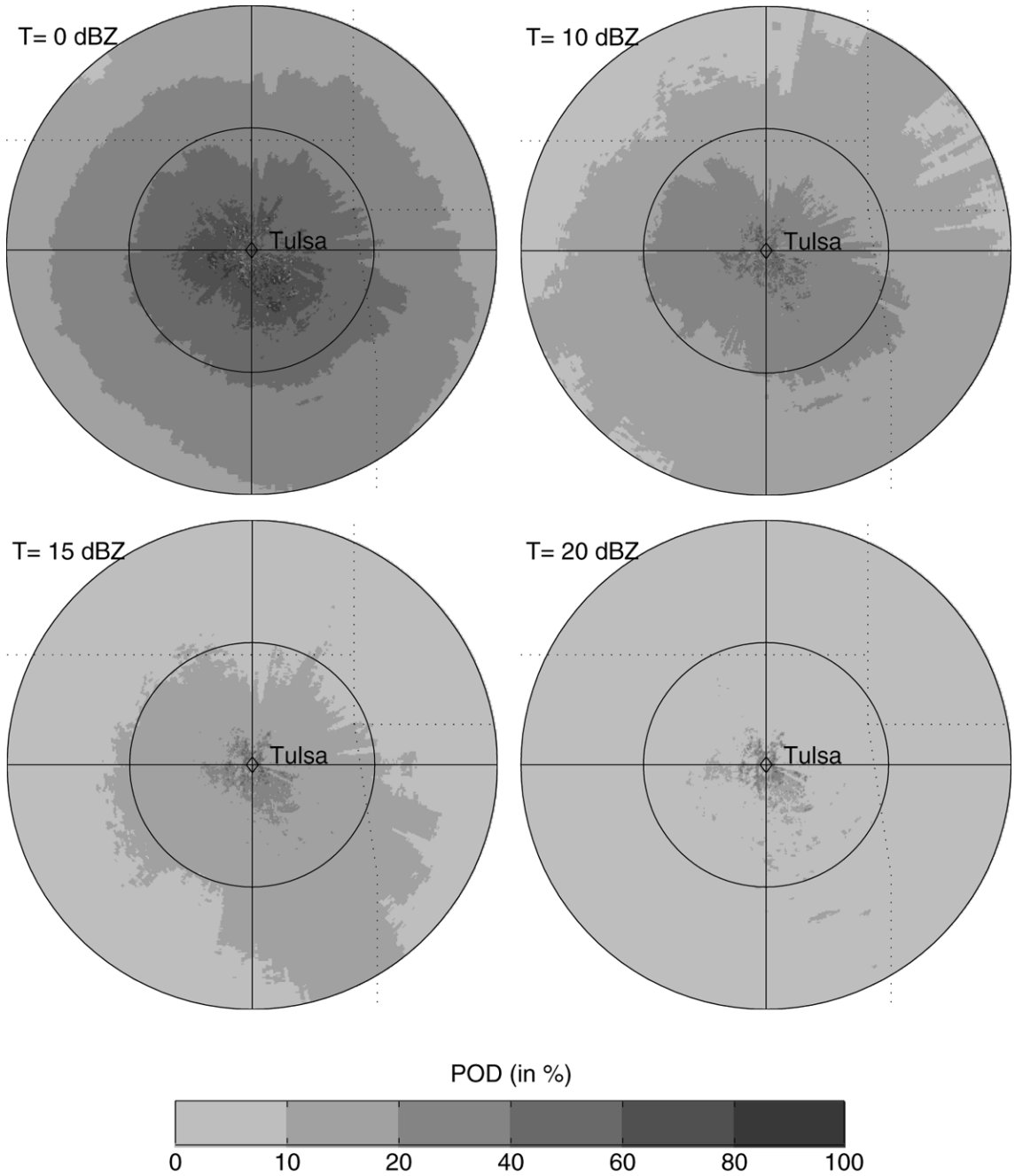


FIG. 2. POD before QC for different thresholds (from 0 to 20 dBZ) based on 10 000 volume scans collected by the Tulsa, Oklahoma, WSR-88D between 1994 and 1995. The rings, centered on the radar, denote 100- and 200-km ranges.

b. Radar and rain gauge analysis

Our analysis now includes hourly estimates of rainfall calculated by applying the ''standard'' Next Generation Weather Radar Z–R relationship $Z = 300R^{1.4}$ to the base scan data (antenna elevation angle of 0.47°). We also use hourly rain gauge data from the Oklahoma Mesonet.

First, consider the conditional probability that radar observes reflectivity (Z) greater than 10 dBZ given that a collocated rain gauge observes measurable rainfall ($R > 0.1$ mm). This statistic is useful, as the study by Grecu and Krajewski (2000) did not address the performance of AP detection in the presence of rain. Theoretically, the probability $P(Z > 10 \text{ dB}Z \mid R > 0.1 \text{ mm})$ should be high, but an AP procedure that eliminates too much rain would decrease it below the true (but unknown) level. As is evident from Fig. 5, the quality-controlled data results in the conditional probability more uniformly distributed along the distance from the radar, in line with data at far ranges where the QC procedure does little. (Our simplified calculations based on hourly rain gauge

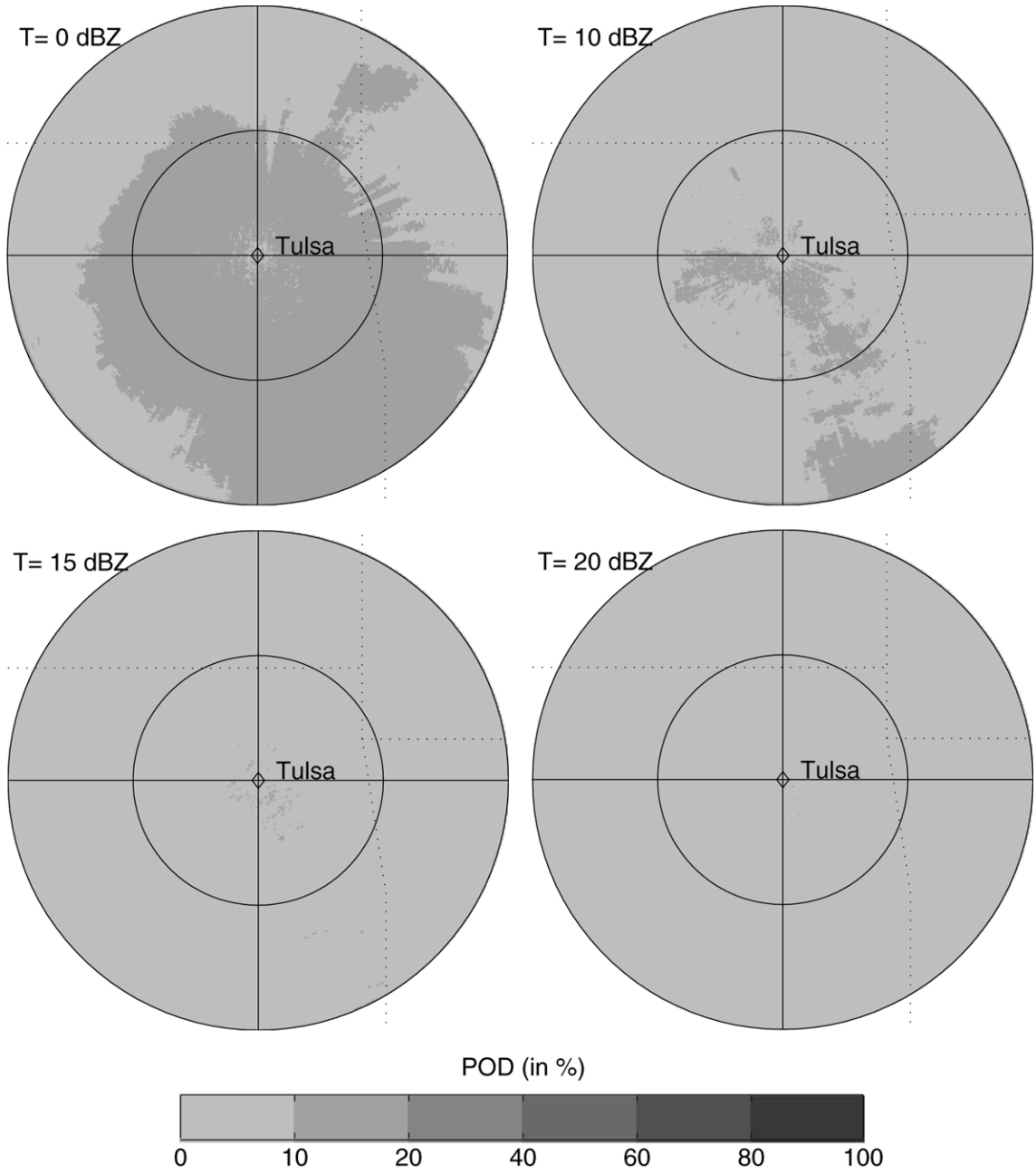


FIG. 3. Same as in Fig. 2, but after QC.

data assume that rainfall is uniformly spread over a given hour—a necessary assumption for comparing individual radar scans with time-integrated gauge observations.)

The effect of quality control is much more dramatic if we consider the conditional probability that radar observes significant echo, given that rain gauges observe no rainfall, that is, $P(Z > 10 \text{ dB}Z \mid R < 0.1 \text{ mm})$. Theoretically, this probability should be only slightly greater than zero, because radar ''sees'' larger areas and possibly weak echoes that produce no surface rainfall. Indeed, as we show in Fig. 6, this probability seems reasonable for the quality-controlled data, but too high for the original data. Most importantly, this probability is independent of the distance from radar, which is a strong indication of the effectiveness of the QC procedure.

Finally, consider the total rainfall accumulation from both the rain gauges and the radar. We calculated the field of rainfall accumulation for the entire duration of our dataset by interpolating between the locations of the gauges. The interpolation method is based on four quadrant near-neighbors and inverse distance weights, the same method used by the National Weather Service in

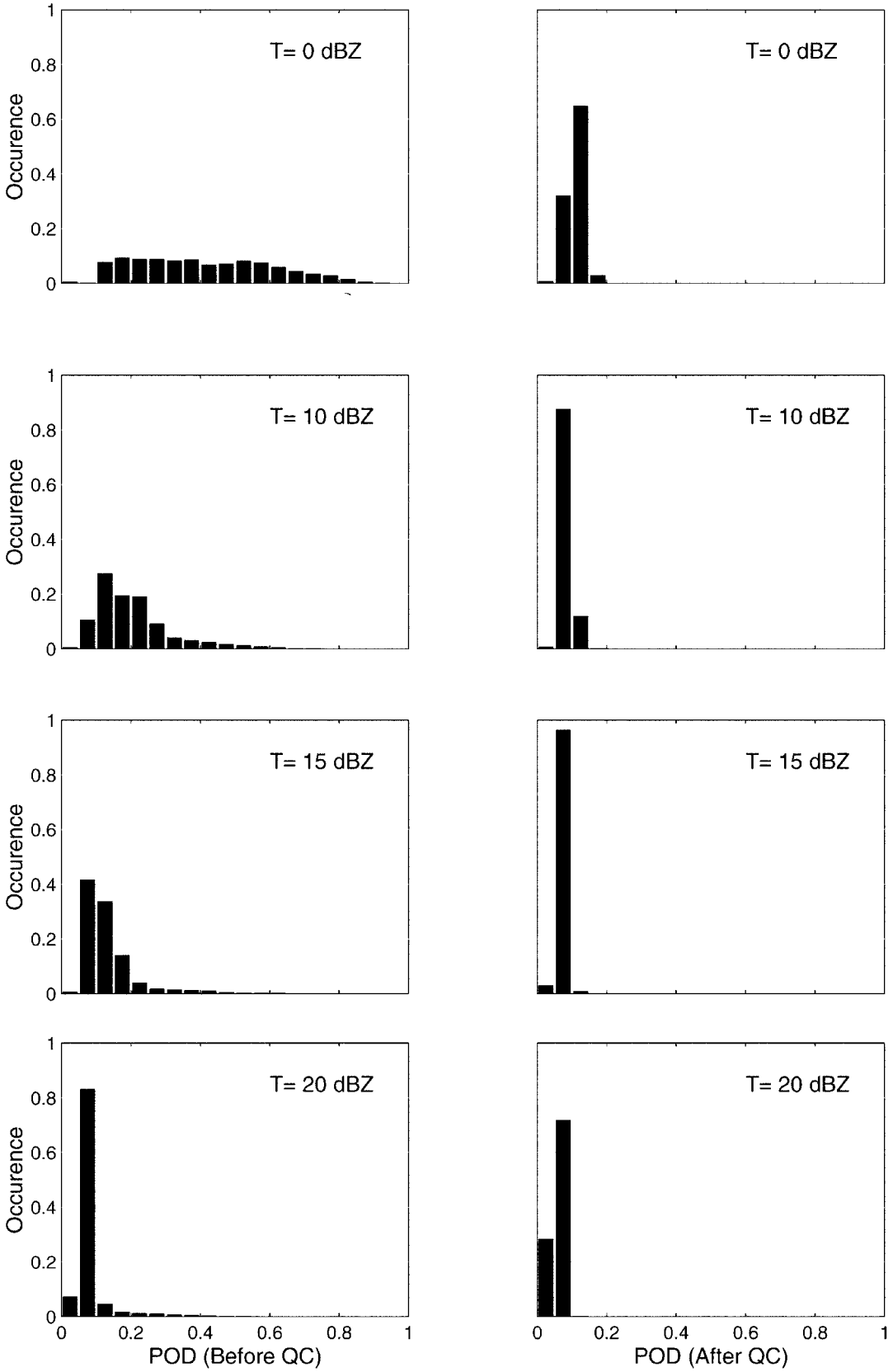


FIG. 4. Histogram of POD based on the 360 × 200 polar pixels of the radar domain.

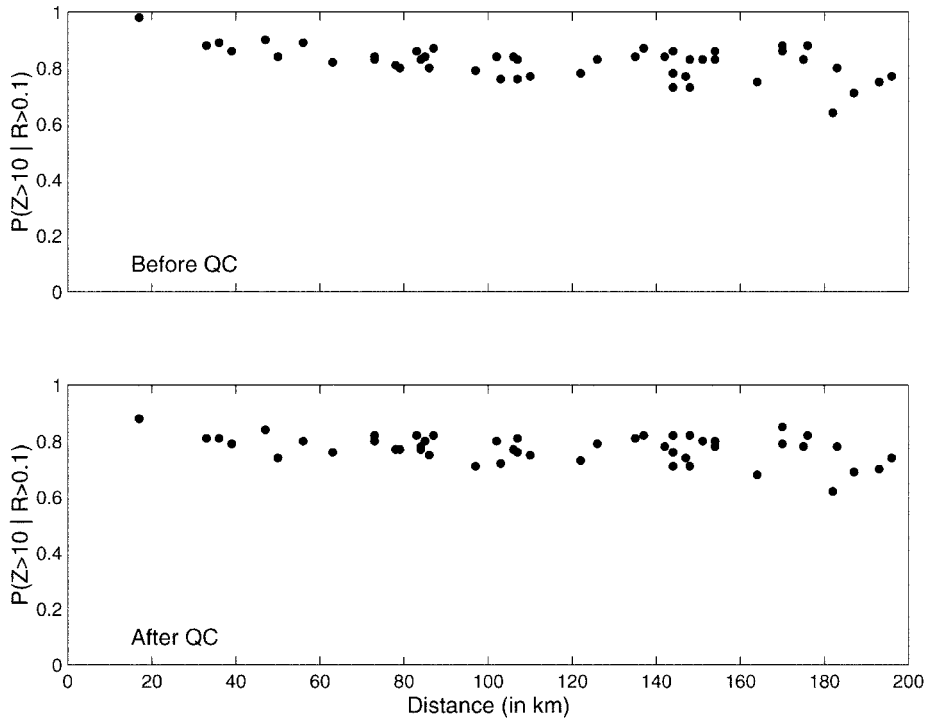


FIG. 5. POD conditional on the gauge measurements, $P(Z > 10 \text{ dBZ} \mid R > 0.1 \text{ mm})$.

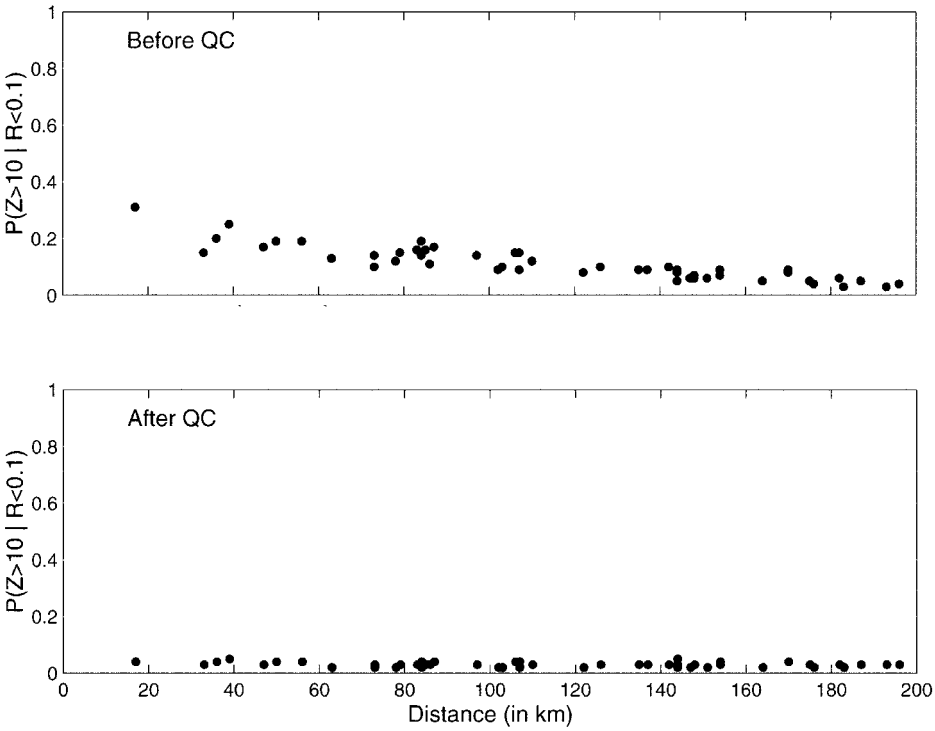
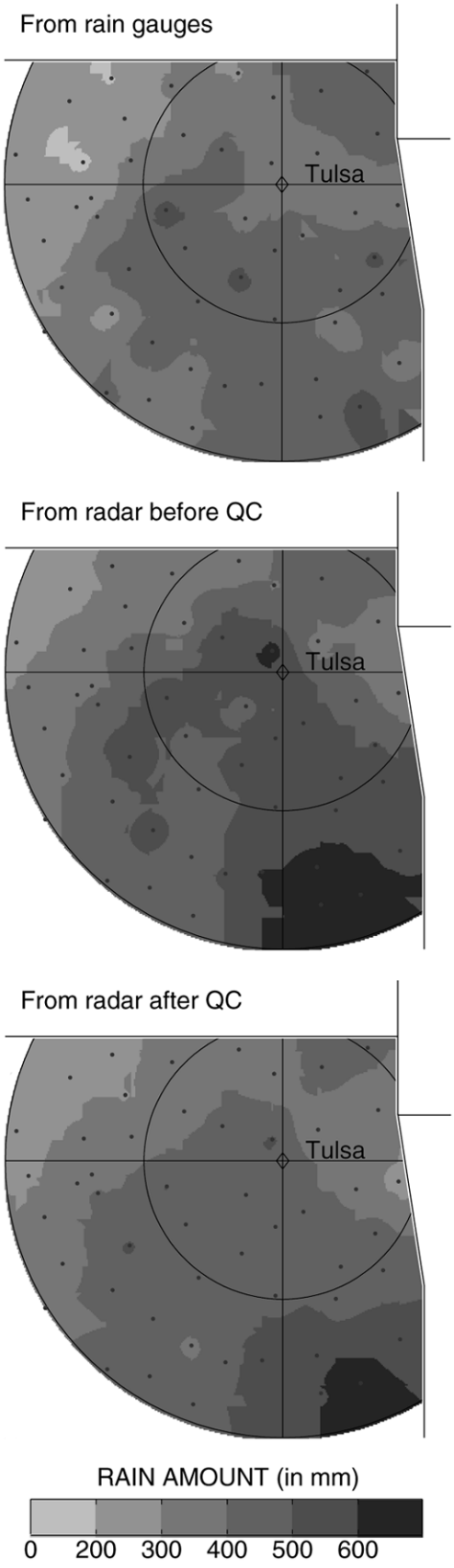


FIG. 6. Probability of false detection conditional on the gauge measurements, $P(Z > 10 \text{ dBZ} \mid R < 0.1 \text{ mm})$.



its hydrologic forecasting. We applied it to both the radar and the rain gauge fields (Fig. 7). We show the accumulated field within the boundaries of the state of Oklahoma only to avoid the possible inconsistency of using a rain gauge dataset outside the Oklahoma Mesonet. To make the comparison easier, we also show a map of the bias—defined as *difference between the two*—interpolated the same way (Fig. 8). It is clear that the accumulation field displays lower bias when calculated on the quality-controlled data. The ''hot spot'' on the figure (about 145 km west of the radar) is most likely due to an underestimation by the gauge. This gauge (Red Rock, Oklahoma; Mesonet code REDR) displays relatively long periods of reporting no rain during which radar observes rain. We summarize the quantitative results in Table 1. The table shows that the effect of the QC procedure is smallest at distances greater than 100–150 km from the radar. This is understandable, as the influence of AP diminishes farther out from the radar. It is the nonuniformity of the vertical profile of reflectivity that is mostly responsible for the discrepancy at those far ranges (Vignal et al. 1999).

4. Conclusions

The QC procedure developed by Grecu and Krajewski (2000) for detecting AP signals in WSR-88D data proved very effective when applied to the Tulsa, Oklahoma, radar. The agreement of long-term accumulation between the radar and rain gauge data is remarkable, especially considering that there is no fitting of the *Z–R* or any other parameters of the rainfall estimation algorithm. The algorithm we applied is the simplest it can be; it does not include rainfall classification, vertical integration, or pattern extrapolation—all elements that improve radar rainfall estimation, as shown in previous studies (e.g., Ciach et al. 1997; Anagnostou and Krajewski 1999). We used the standard *Z–R* relationship, which proved adequate.

Perhaps the most important aspect of our study is that we used a large sample of volume scan reflectivity data and mimicked an operational environment to the degree possible. We are convinced that, due to the highly variable nature of rainfall processes, large sample studies are critically important in evaluating new technologies of radar rainfall estimation, including the ''solve-it-all'' polarimetric methods (e.g., Zrnic and Ryzkov 1999). The traditional mode of testing radar methods using only a few selected events is simply misguided. Many statistics calculated from radar and rain gauge observations, for example, bias, are mean-

←

FIG. 7. Total rainfall accumulation interpolated from the Oklahoma Mesonet rain gauge data and from radar, both before and after QC. Dots denote the gauge locations; rings are every 100 km from radar.

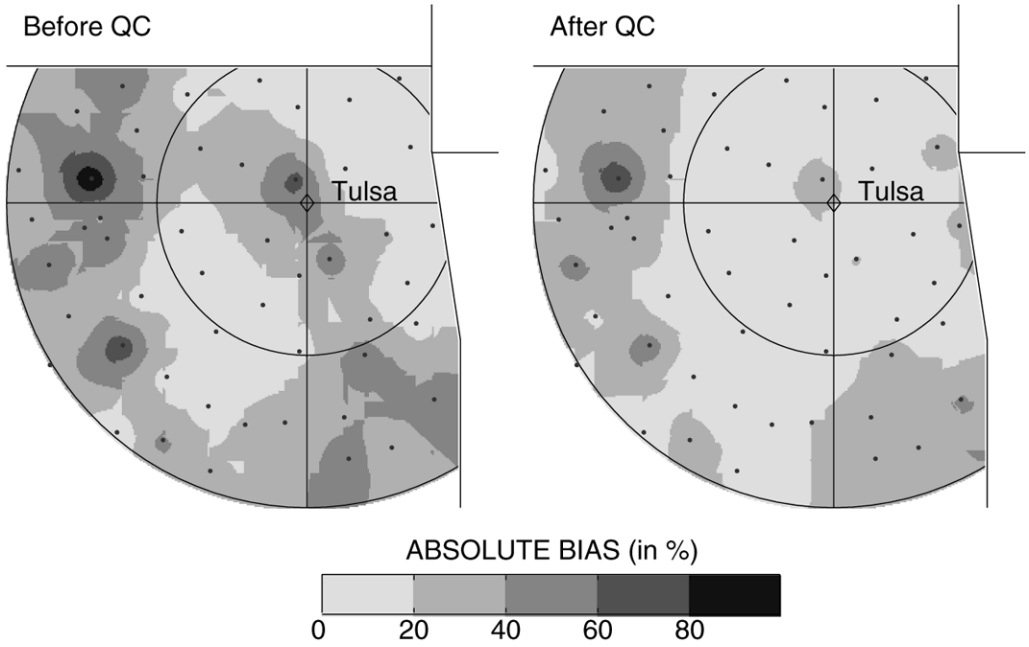


FIG. 8. Bias between radar and gauge total rainfall accumulation, considering radar data before and after QC. Dots denote the gauge locations; rings are every 100 km from the Tulsa WSR-88D radar.

ingful and/or statistically stable only for accumulations over a long period.

We have assembled large samples of radar reflectivity data for several WSR-88D locations, including Davenport, Iowa; Grand Forks, North Dakota; Memphis, Tennessee; and Melbourne, Florida. We are planning to use these datasets for a variety of comparative studies and analyses, and we invite collaboration.

Acknowledgments. This study was supported by the National Weather Service Office of Hydrology under cooperative agreement with the Iowa Institute of Hydraulic Research (NA47WH0495). We would like to thank Dr. Mircea Grecu for his assistance with the computer code and Dr. Grzegorz Ciach for many helpful discussions. We are grateful to the Oklahoma Climatological Survey and the Oklahoma EPSCoR program for providing the rain gauge data from the Oklahoma Mesonet.

TABLE 1. Evaluation of the QC procedure using total rainfall accumulations (averaged across the gauged). The criterion is the relative difference between accumulations from radar and rain gauge (i.e., bias in percent).

	Distance from radar (km)				
	0–200	0–50	50–100	100–150	150–200
Mean accumulation of gauges (mm)	375	391	427	356	336
Bias before QC (%)	26	34	9	32	35
Bias after QC (%)	14	12	−5	20	26

REFERENCES

Anagnostou, E. N., and W. F. Krajewski, 1999: Real-time radar rainfall estimation. Part 1: Algorithm formulation. *J. Atmos. Oceanic Technol.,* **16,** 189–197.

Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet—A technical overview. *J. Atmos. Oceanic Technol.,* **12,** 5–19.

Ciach, G. J., W. F. Krajewski, E. N. Anagnostou, J. R. McCollum, M. L. Baeck, J. A. Smith, and A. Kruger, 1997: Radar rainfall estimation for ground validation studies of the Tropical Rainfall Measuring Mission. *J. Appl. Meteor.,* **36,** 735–747.

Grecu, M., and W. F. Krajewski, 2000: An efficient methodology for detection of anomalous propagation echoes in radar reflectivity data using neural networks. *J. Atmos. Oceanic Technol.,* **17,** 121–129.

Houze, R. A., Jr., B. F. Smull, and P. Dodge, 1990: Mesoscale organization of springtime rainstorms in Oklahoma. *Mon. Wea. Rev.,* **118,** 613–654.

Klazura, G. E., and D. A. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. *Bull. Amer. Meteor. Soc.,* **74,** 1293–1311.

Kruger, A., and W. F. Krajewski, 1997: Efficient storage of weather radar data. *Software Practice Experience,* **27,** 623–635.

Moszkowicz, S., G. J. Ciach, and W. F. Krajewski, 1994: Statistical detection of anomalous propagation in radar reflectivity patterns. *J. Atmos. Oceanic Technol.,* **11,** 1026–1034.

Vignal, B., H. Andrieu, and J. D. Creutin, 1999: Identification of vertical profiles of reflectivity from voluminal radar data. *J. Appl. Meteor.,* **38,** 1214–1228.

Zrnic, D. S., and A. V. Ryzhkov, 1999: Polarimetry for weather surveillance radars. *Bull. Amer. Meteor. Soc.,* **80,** 389–406.