

Sensitivity of WSR-88D Rainfall Estimates to the Rain-Rate Threshold and Rain Gauge Adjustment: A Flash Flood Case Study

RICHARD A. FULTON

Hydrologic Research Laboratory, Office of Hydrology, NOAA/National Weather Service, Silver Spring, Maryland

(Manuscript received 27 May 1998, in final form 20 May 1999)

ABSTRACT

A strong thunderstorm produced a flash flood on the evening of 12 July 1996 in Buffalo Creek, Colorado, that caused two deaths and significant property damage. Most of the rain fell in a 1-h time period from 2000 to 2100 MDT. The performance of the WSR-88D rainfall algorithm, Precipitation Processing System, was examined in detail to determine how well it performed. In particular the sensitivity to the algorithm's rain-rate threshold (hail cap) parameter and the performance of the gauge-radar adjustment subalgorithm on the resulting radar rainfall estimates were examined by comparison with available rain gauge data.

It was determined that the WSR-88D rainfall algorithm overestimated the rainfall in general over the radar scanning domain for this event by about 60% relative to the rain gauges although the radar-derived rainfall for the flood-producing storm cell nearly matched the single gauge that sampled it. The derived rainfall over the radar scanning domain was not very sensitive to the setting of the rain-rate threshold parameter. Lowering it reduced the overestimation in general but did not bring the estimates satisfactorily close to unbiasedness. Other error sources were suggested, including use of an inappropriate $Z-R$ relationship and/or radar reflectivity miscalibration. Relative importance of these sources could not be determined.

The portion of the rainfall algorithm that adjusts the radar estimates using rain gauge data was tested to determine if it could have satisfactorily reduced the observed overestimation. It was found to have performed suboptimally due primarily to the methodology in the algorithm that forms the gauge-radar pairs. A simpler technique was proposed and tested, and the algorithm's performance was greatly enhanced as a result. Therefore the performance of the gauge-radar adjustment algorithm depends on the gauge-radar rainfall data that are passed to it, and that data are dependent on the method by which the pairs are formed.

1. Introduction

The performance of the Weather Surveillance Radar-1988 Doppler (WSR-88D) Precipitation Processing System (PPS) (Fulton et al. 1998) is examined for a flash flood event that occurred in the evening of 12 July 1996 in the small rural town of Buffalo Creek, Colorado, located about 50 km southwest of Denver in mountainous southern Jefferson County. Two drownings and extensive property damage occurred in the town and the immediate vicinity associated with the rapid rise of the Buffalo Creek that flows through town. In the appendix is a memorandum written by the National Weather Service (NWS) Forecast Office in Denver describing the impacts of the flood event. Brandes et al. (1997) describes a preliminary study of the Buffalo Creek event that compares rainfall estimates from the WSR-88D and the National Center for Atmospheric Research S-Pol dual-polarization radar.

Thunderstorms first developed in central Colorado around 1100 MDT 12 July (1700 UTC) and moved generally southeastward. The most intense rainfall over the Buffalo Creek watershed occurred primarily between 2000 and 2100 MDT 12 July (0200 and 0300 UTC 13 July). At that time east-southeasterly (upslope) surface winds in eastern Colorado were advecting moist air with dewpoints in the low 60°s F along the front range of the Rocky Mountains (Fig. 1a). A weak stationary front was draped northwest-to-southeast across the region, and an upper-level ridge over the West Coast produced mid- and upper-level west-northwesterly winds of 30–40 kt over Colorado. The atmosphere was convectively unstable with a lifted index of -4 from the modified 0000 UTC 13 July Denver sounding about 3 h before the flood event (Fig. 1b).

There were some unofficial reports of hail associated with the flood-producing storm. There were 11 official hail reports ranging from 0.75 to 1.75 in. in diameter and three F0 tornadoes throughout the same day associated with other severe thunderstorms in Colorado (NOAA 1996). In fact, near the time of the flash flood in southern Jefferson County, major hail damage and an F0 tornado were occurring in the extreme northern

Corresponding author address: Richard A. Fulton, Hydrologic Research Laboratory W/OH1, National Weather Service, 1325 East-West Highway, Silver Spring, MD 20910.
E-mail: fulton@skipper.nws.noaa.gov

portion of Jefferson and adjacent Boulder Counties from a different thunderstorm.

The storm-total rainfall over the Buffalo Creek watershed from the Denver WSR-88D radar (FTG) ending at 0800 UTC 13 July as estimated by the PPS algorithm is shown in Fig. 2. The location of the heaviest rainfall associated with the flash flood-producing storm is centered very close to the town of Buffalo Creek (star), located near the outlet of the Buffalo Creek watershed where it empties into the North Fork of the South Platte River. The terrain in this region is steep, and the Buffalo Creek drainage basin ranges in altitude from about 2070 m above mean sea level (6800 ft) at the outlet up to 3600 m (11 900 ft). It has an area of 120 km² and is located at a distance of 75–93 km and azimuth of 230°–237° from the Denver WSR-88D. The forest fire burn area, which is believed to have played a major role in the severity of the flash flood (see the appendix), is also outlined in Fig. 2.

2. WSR-88D reflectivity data

Archive level II radar reflectivity factor (hereafter called simply reflectivity) data (Crum et al. 1993) from the Denver WSR-88D for 141 volume scans were ingested into the PPS for the 14-h period from 1756 UTC 12 July to 0800 UTC 13 July 1996. Thunderstorms first developed in Colorado around 1700 UTC; however, Archive II data prior to 1756 UTC were not consistently available and therefore could not be used. No analyses have been performed prior to 1756 UTC.

Over the Buffalo Creek basin the PPS uses reflectivity data exclusively from the second elevation angle (nominally 1.5°) in constructing the reflectivity hybrid scan and resulting rainfall estimates. The reflectivity hybrid scan is the plan position indicator of reflectivity constructed each volume scan by choosing the reflectivity at the elevation angle specified in Fig. 3. The Buffalo Creek basin is outlined to the southwest of the Denver radar and is in the region labeled “elev 2 only” where the reflectivities from the second elevation angle are used. The WSR-88D off-line algorithm, described in O’Bannon (1997) and Fulton et al. (1998), that builds the site-unique, static, hybrid scan elevation angle data for each radar uses high-resolution digital terrain height data to find the most optimum elevation angle to use at each polar grid bin that is not likely to be contaminated by ground clutter or excessively blocked by terrain. Beam blockage by terrain at the lowest elevation angle of 0.5° ranges from 10% to 40% over the basin, but the second elevation angle is free of blockage (Fig. 4).

3. Rain gauge data

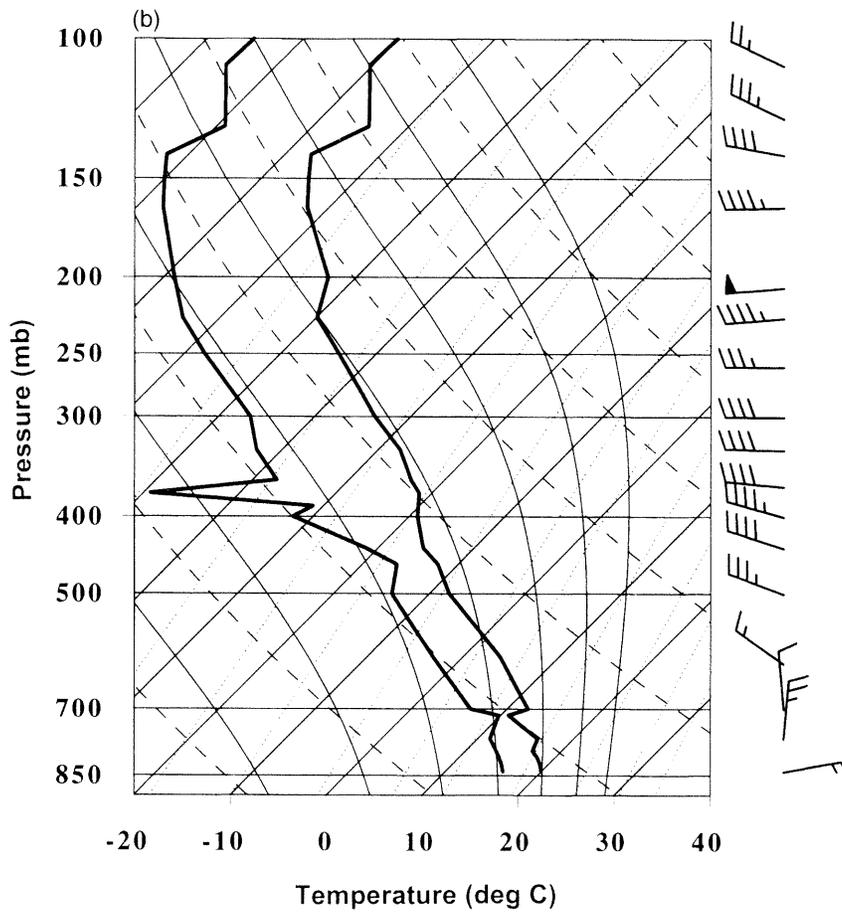
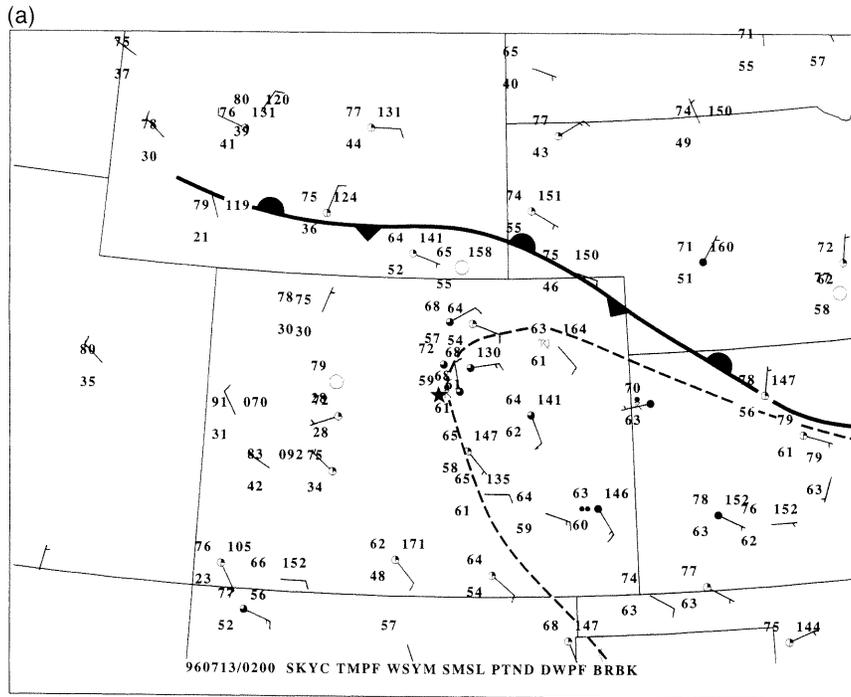
Automated digital data from 145 rain gauges within 230 km of FTG were collected from three sources: the National Weather Service (NWS) Arkansas—Red Basin River Forecast Center in Tulsa, Oklahoma; the NWS

Missouri Basin River Forecast Center in Pleasant Hill, Missouri; and the Denver Urban Drainage and Flood Control District (UDFCD), which maintains a gauge network primarily in the Denver metropolitan area. The 56 gauges available from the first two sources include automated NWS gauges as well as automated gauges from other federal agencies such as the U.S. Army Corps of Engineers and the U.S. Geological Survey. They represent typical gauge data available nationally for operational use by the NWS. The 88 UDFCD gauges are part of a local network whose data are received by the NWS forecast office in Denver. Manual quality control of the data was performed to remove any obviously bad reports.

Most of the rain gauges within 230 km of FTG are located in the mountains in the western sector of the radar scanning domain (Fig. 5). The two conglomerations of gauges just to the west and northwest of the radar are UDFCD gauges in the Denver metropolitan area, and the remainder are the federally owned gauges. There were no automated rain gauges in the Buffalo Creek basin; however, there was one storm-total gauge report from a resident of Buffalo Creek of 2.68 in. of rain. During its lifetime, the flood-producing storm did not pass over any rain gauge other than this privately owned gauge in Buffalo Creek. The closest automated gauge, SPTC2, was located just 9 km to the northeast of Buffalo Creek; however, it was missed by the flood-producing storm and only received 0.07 in. Of the 145 gauges used in this study, 61% received rain during the period 1600 UTC 12 July to 0800 UTC 13 July.

The gauge data used in this study serve both as independent verification of the PPS radar estimates as well as direct input into the PPS adjustment algorithm to compute hourly gauge–radar biases to automatically correct the operational radar estimates. Fulton et al. (1998) describes the PPS adjustment algorithm.

Since the current PPS gauge–radar adjustment algorithm will allow at most 50 rain gauges in its internal database for rainfall calibration purposes due to software limitations, the 145 available gauges were reduced to just 50 for use in algorithm performance testing. In order to reduce the number of gauges to 50 for testing, they were first ranked in descending order of storm-total rainfall over the period examined. Gauges that were spatially very close to other gauges were excluded because they may not have provided truly independent information. This was a common occurrence with the UDFCD gauges that were concentrated in a small region near metropolitan Denver and thus sampled only a very limited area of the total rainfall system. Any gauges that did not report consistently at frequencies better than or equal to about once per hour were also excluded from use in automated gauge–radar adjustment. The PPS adjustment algorithm would not be able to determine hourly incremental rainfall accumulations for these gauges, thus making them useless for real-time calibration purposes. The top 50 gauges from the resulting list were



chosen for use in automated gauge–radar adjustment. All 145 available gauges were retained for use in scatterplot comparisons and verification of gauge and radar rainfall, however.

4. Performance of the gauge–radar adjustment algorithm

This section describes how the gauge–radar adjustment procedure performed in attempting to produce automated, unbiased radar rainfall estimates. The FTG WSR-88D radar was not being supplied with real-time rain gauge data at the time of the flash flood; however, the PPS has been rerun using Archive Level II radar data and the archived rain gauge data in a simulated real-time mode. The gauge reports are assumed to have been received at the WSR-88D within 50 min of their observation time. In the real operational world, there can be longer delays in transmission of the gauge data from the gauge platform to the WSR-88D.

The build 9 software version of the WSR-88D PPS was run for this event. All PPS adaptable parameters [details are described in Fulton et al. (1998)] were set to the values in use at the FTG WSR-88D at the time of the flood. Specifically, a few of the more important ones include the Z – R equation, $Z = 300R^{1.4}$, and the rain-rate threshold (sometimes called the *hail cap*) of 74.7 mm h^{-1} , which corresponds to 51 dBZ assuming the Z – R equation above. Any rain rates that exceed this value are capped at this value to prevent hail contamination. The gauge accumulation time was set to $H + 00 \text{ min}$ (i.e., hourly radar and gauge accumulations ending at the top of every hour are used in the adjustment algorithm), and the time each hour when the bias estimation was performed was $H + 50 \text{ min}$ based on hourly accumulation data ending at $H + 00 \text{ min}$. An arbitrary 50-min time delay is built into the operational software to permit the gauge data enough time to be collected and transmitted to the radar prior to their use. Once a bias is computed near the end of the hour, that value remains in affect and is applied to the subsequent scan-to-scan radar estimates until a new bias is computed the following hour. Biscan maximization, the procedure in the algorithm that sets a grid bin's reflectivity to the highest value from the two lowest elevation angles, was disabled.

Figure 6a shows the total reflectivity echo area time series over the 230-km range domain of FTG computed from the hybrid scan of reflectivity each volume scan (every 6 min). The period of archive level II data examined covered the entire life cycle of the larger-scale rainfall event over Colorado from formation to dissi-

pation. The flood-producing storm passed over the Buffalo Creek basin during the period 0200–0300 UTC (hours 26–27 in Fig. 6), shortly after the time of most widespread rainfall coverage. The time series of mean reflectivity over the entire radar domain, weighted by the area of the polar bins and conditional on the existence of reflectivity exceeding 7 dBZ, increases to a maximum value near 24 dBZ several hours into the rainfall event, remains there for several hours, and then decreases as the storm system dissipates (Fig. 6b).

Figure 7 shows the output from the PPS adjustment algorithm over the course of the rainfall event. These hourly time series plots include (a) computed gauge–radar multiplicative bias estimate (G/R), (b) mean square error of the bias estimate, (c) total number of available gauge–radar pairs, and (d) the number of pairs used by the adjustment algorithm after automated quality control. Even though as many as 31 gauge–radar pairs (out of a maximum possible 50) were available for 1 h during the period, at most only 17 pairs passed the two quality control steps of the PPS (Figs. 7c and 7d). In order for a gauge–radar pair for a given hour to pass quality control, the hourly gauge and radar rainfall must both exceed 0.6 mm (i.e., it must have rained), and the difference between the gauge and radar amounts must not exceed 2.0 standard deviations of the mean difference for all gauge–radar pairs for that hour (Fulton et al. 1998). A minimum of six pairs must pass the quality control steps in order for a bias to be estimated for that hour.

Due to an insufficient number of pairs, an hourly bias estimate was not computed until 0150 UTC 13 July (hour 25.8), which is valid for the hourly period 0000–0100 UTC. Before this time, the bias estimate is simply the default value of 1.0, implying no bias. However, after that, bias estimates are computed for each of the next six consecutive hours until 0750 UTC (hour 31.8) when the number of pairs falls again below the minimum required number of six. The computed biases were exactly 1.0 during the first three hours (26–29) when enough gauge–radar pairs were available. The computed bias estimates from the adjustment algorithm range between 1.00 and 0.96 (Fig. 7a) based on 7–17 input gauge–radar pairs (Fig. 7d) and indicate that the radar estimates are apparently in very good agreement with the available gauge measurements.

Hourly scatterplots of gauge–radar accumulations are shown in Fig. 8. Filled (open) circles represent gauge–radar pairs that passed (did not pass) the quality control steps. If less than six pairs were available for any given hour, all of those pairs will be indicated as open circles for that hour, implying that a bias could not be com-

←

FIG. 1. (a) Surface observations over Colorado and surrounding states at 0200 UTC 13 Jul 1996. The 60°F dewpoint dashed contour is shown. Buffalo Creek, CO, is located at the star. (b) Denver sounding profile of temperature and dewpoint for 0000 UTC 13 Jul 1996 modified in the lowest 0.5 km for Denver surface observations at 0200 UTC. Wind barbs are in kt.

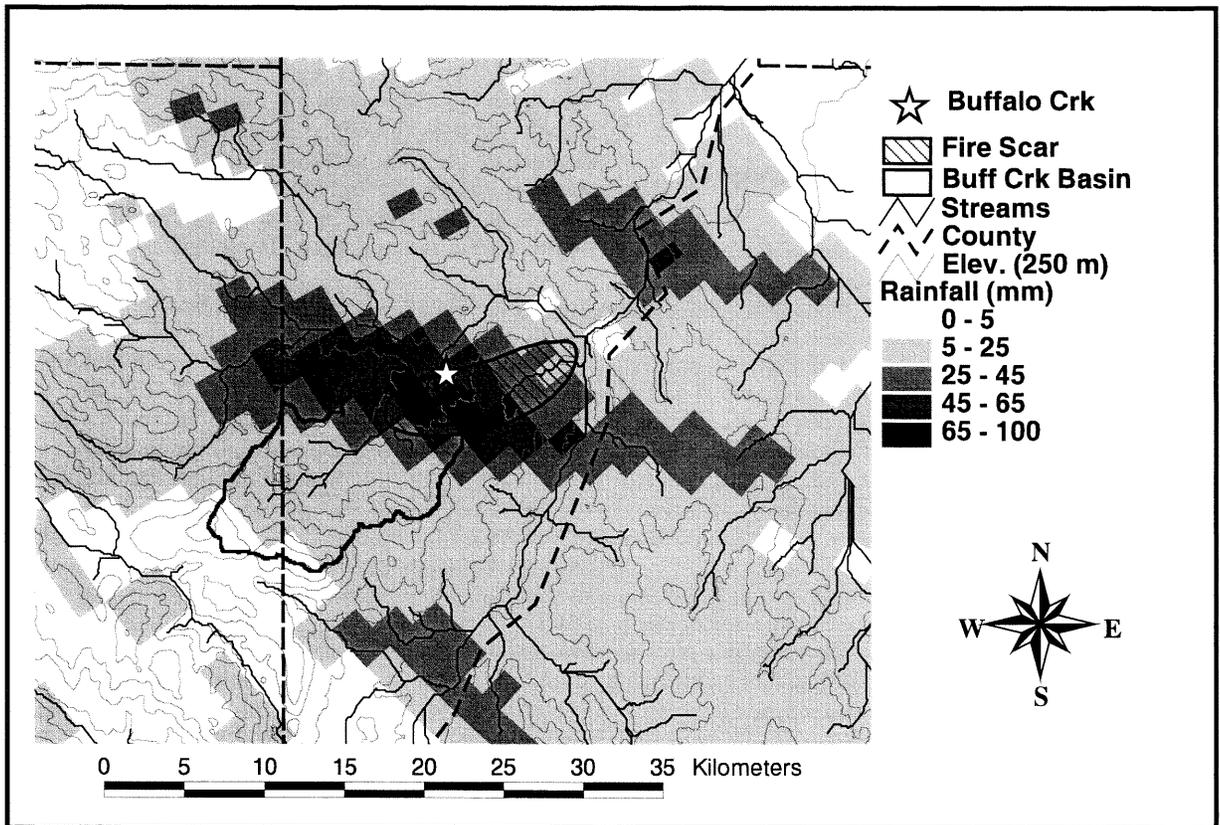


FIG. 2. Storm-total rainfall (mm) ending at 0800 UTC 13 July 1996 from the Denver WSR-88D radar. The outline of the Buffalo Creek basin is shown as a thick solid line, and the town of Buffalo Creek is indicated by the white star near the basin outlet. The forest fire burn area is hatched, and terrain heights are contoured every 250 m.

puted. Such was the case until hour 25. The time of bias estimation (in decimal form relative to 0000 UTC 12 July) is indicated at the top of each panel and is normally at or shortly after $H + 50$ min. For the 6-h period from 25.886 to 30.850 h when gauge-radar bias estimates were successfully computed by the algorithm, the gauge-radar pairs fall along or very close to the diagonal line indicating near-perfect agreement. The multiplicative bias estimates deviate most from 1.0 at hour 28.902 with a bias estimate of 0.96, indicating slight radar overestimation relative to the gauges for that hour. Thus the PPS seems to be yielding accurate estimates of the accumulation compared to the available hourly rain gauges.

5. Impact of the method of forming gauge-radar pairs

Because the neat alignment of points in Fig. 8 along the unbiased diagonal line seemed anomalously good, the method for forming the gauge-radar pairs within the algorithm was examined in greater detail for this case. It has been found that, in fact, this neat alignment is due to the method in the algorithm that pairs the radar

rainfall to a corresponding gauge value, and this method has an adverse effect on the computed gauge-radar biases such that they tend to be closer to unity (unbiasedness) than they really should be. The net effect of this is a degradation of the ability of the adjustment algorithm to properly identify and thus correct for biased radar estimates. This section describes these results.

There are two fundamental steps to estimating a gauge-radar bias in the PPS adjustment algorithm as described in Fulton et al. (1998). The first step is the pairing together of hourly gauge and radar rainfall estimates. The algorithm examines the radar estimates for the nine (3×3) polar grid bins surrounding each gauge location. If the gauge observation for a given hour falls within the range of radar estimates from the nine surrounding polar grid bins for that same hour, it assigns a radar rainfall that is exactly equal to the gauge observation; that is, it forms a perfect-match (unbiased) gauge-radar pair regardless of whether or not a radar bin actually has that exact amount. If the gauge observation falls outside of the nine-bin radar rainfall range, then the closest (in value) radar estimate of the nine bins to the gauge observation is chosen in forming the gauge-radar pair. The original motivation for this design

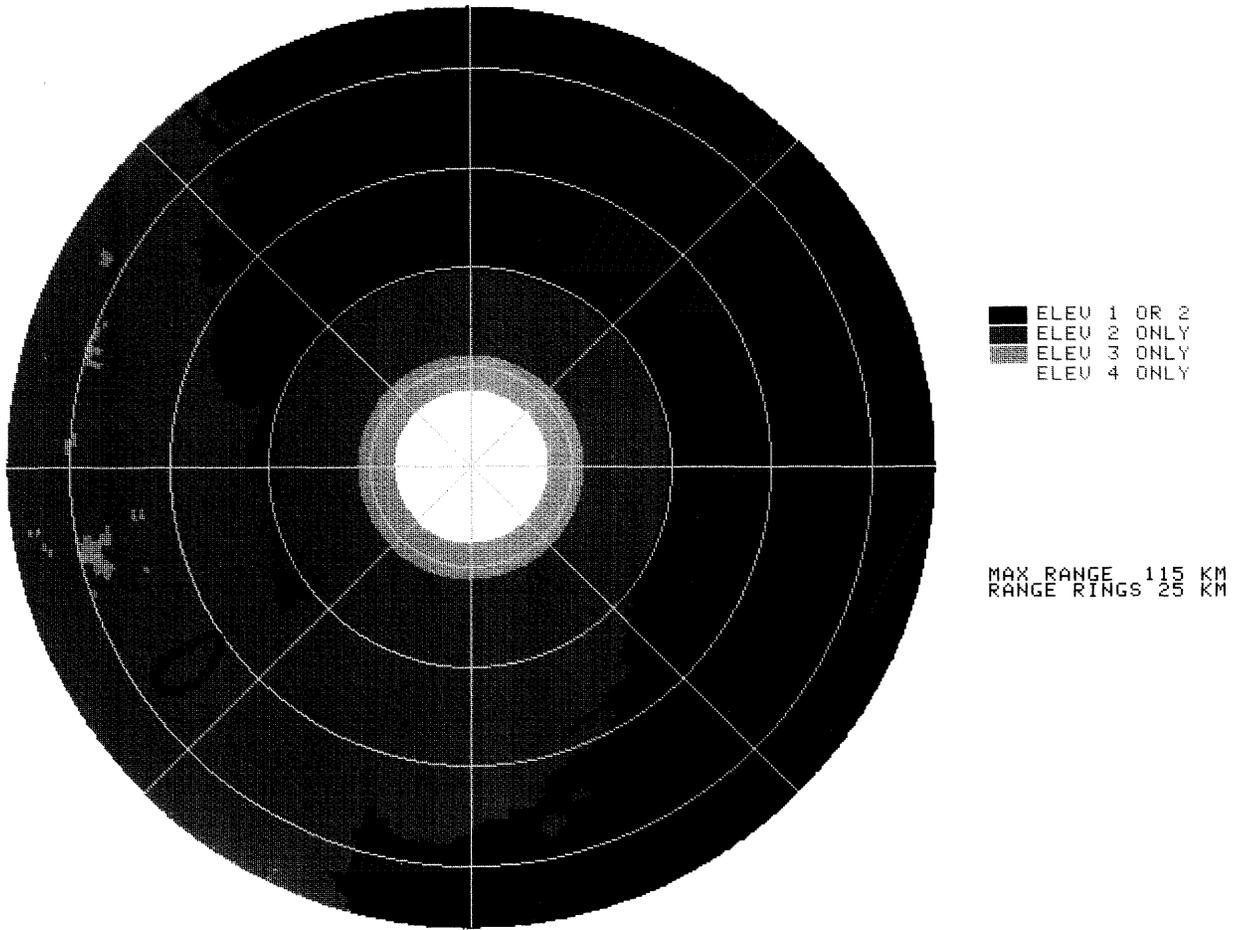


FIG. 3. Hybrid scan elevation angles used by the PPS at the Denver WSR-88D, located in the center, shown to a maximum range of 115 km (rain estimates are produced out to 230 km, however). The Buffalo Creek basin is outlined to the southwest of the radar. The gray shades identified in the legend indicate from which elevation angle for each particular range and azimuth grid bin the reflectivity will be extracted to produce rain estimates.

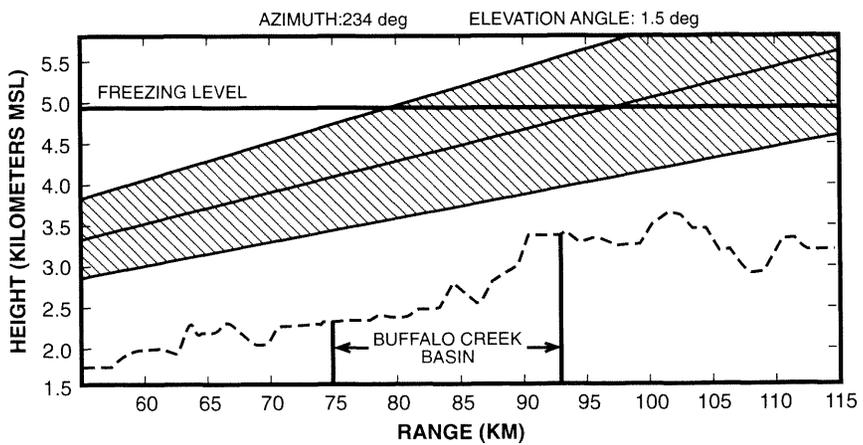


FIG. 4. Height of the 1.5° elevation angle half-power beam (hatched) relative to the terrain (dashed line) at the 234° azimuth angle that is centered over the basin. The two thick vertical lines delineate the extent of the basin. The height of the freezing level is shown at 4.9 km.

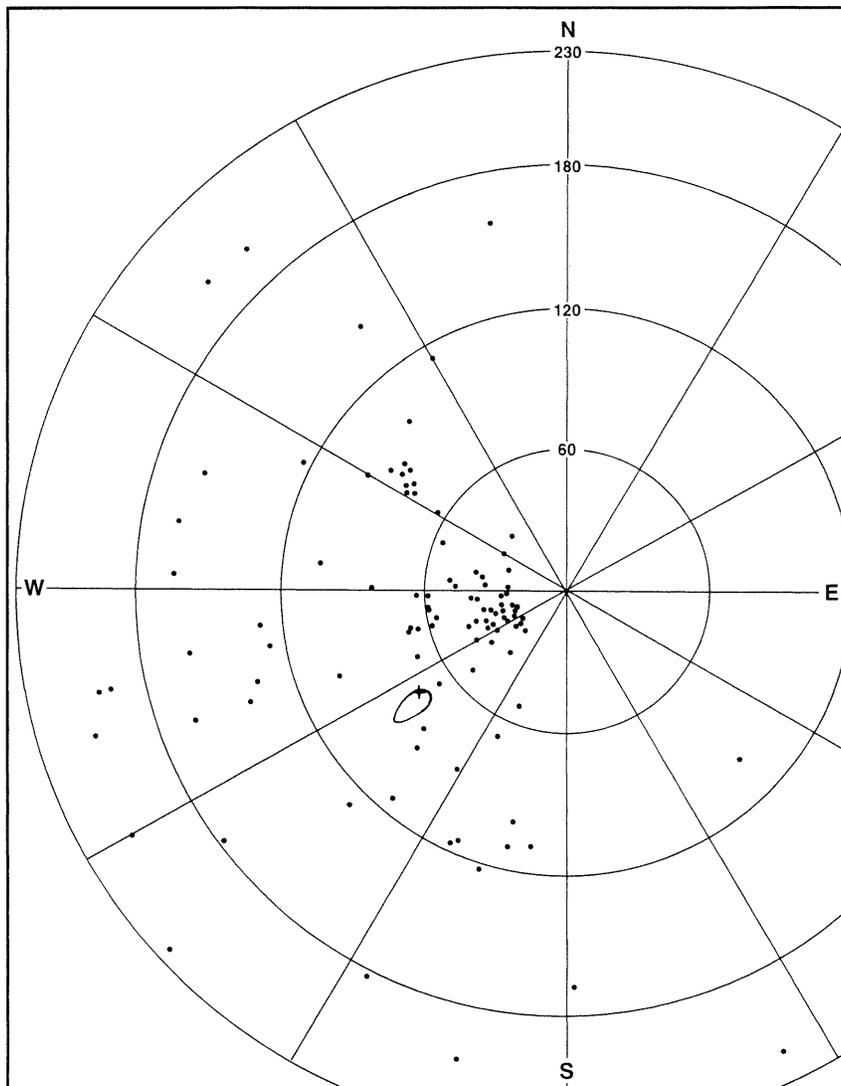


FIG. 5. Rain gauge locations surrounding the FTG radar (center of polar grid) used in this study. Range rings are at 60-km intervals. The Buffalo Creek basin is outlined, and the town of Buffalo Creek is indicated by the "+" symbol.

was to account for the uncertainties in precise knowledge of the gauge location and uncertainty in knowing the exact radar bin that contributes rain to the gauge due to horizontal advection of rain below the beam.

The second step is the actual estimation of the bias using these gauge-radar pairs in the algorithm's Kalman filter. The Kalman filtering procedure is a mathematical temporal-filtering technique that combines the current hour's gauge-radar observations with observations from previous hours using statistical measures of their estimated reliability. In its most fundamental form, the Kalman filter weighs more heavily the most recent gauge-radar observations compared to earlier observations, and the memory and smoothing characteristics of the

filter are controlled by adaptable parameters. These steps are performed once every hour at about H + 50 min using hourly accumulation ending at H + 00 min.

Figure 9 is a scatterplot of gauge-radar pairs of storm-total rainfall for all of the 85 gauges within 230 km of FTG that reported rain during this event. The radar estimates have not been adjusted by any computed biases in this plot. Although the PPS adjustment algorithm does not form or use such *storm-total* gauge-radar pairs, this plot illustrates several points about the general process by which gauge-radar pairs are formed for individual hours during the hourly bias estimation procedure.

For this Buffalo Creek case, there is a large range of radar rainfall estimates in the nine polar bins surround-

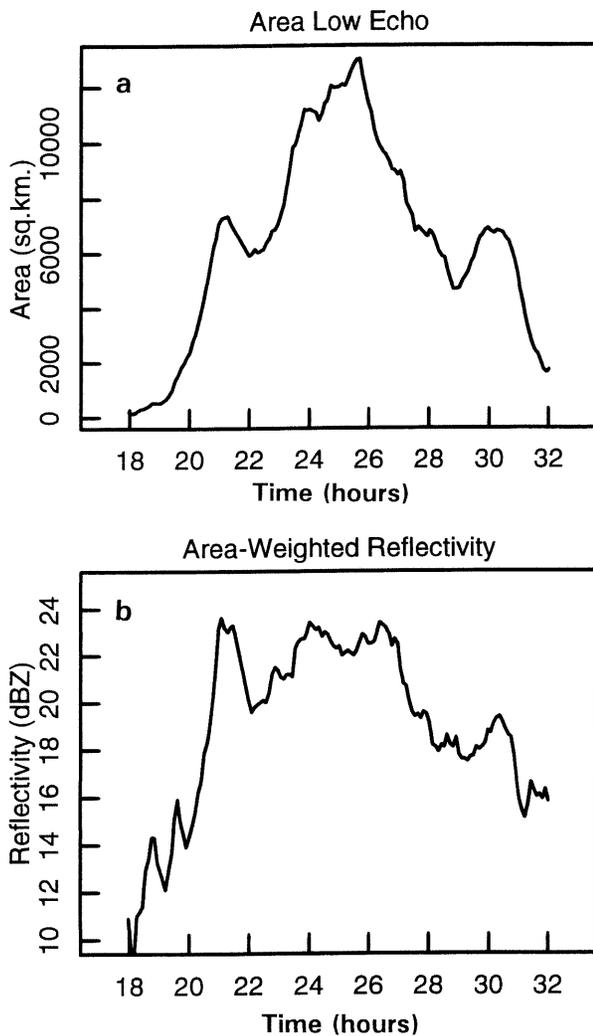


FIG. 6. Time series plots of (a) echo area at the lowest tilt and (b) area-weighted mean reflectivity. Times are relative to 0000 UTC 12 Jul.

ing most gauges due to the large spatial gradients associated with the convective nature of the rainfall on this day. Stratiform rainfall events with typically smaller horizontal gradients of rainfall may not exhibit such large 3×3 bin variability of rainfall. The net effect of this is seen by comparing the filled circles (the “best” radar value used in the algorithm) and the plus symbols (the radar value that is collocated with the gauge) for a given gauge value. The gauge–radar pairs used by the adjustment algorithm lie consistently closer to the diagonal line than the collocated pairs. This by itself is not necessarily bad unless it produces a gauge–radar bias that is itself biased. This appears to be the case based on the computed biases shown in Table 1.

Table 1 illustrates the sensitivity on the computed biases of the method by which the gauge–radar pairs

are formed. Shown are the storm-total sample biases¹ computed according to

$$B = \frac{\sum_{i=1}^n G_i}{\sum_{i=1}^n R_i},$$

where the summations are over all n collocated gauge–radar rainfall pairs for various methods of choosing the unadjusted radar estimates from the nine surrounding polar bins to pair with the corresponding gauge amount. Column 2 in the table is the bias computed using the current algorithm’s method discussed previously in this section. Column 3 is the bias using the radar bin whose rainfall most closely matches the gauge rainfall amount within the 3×3 array. Column 4 is the bias based on use of the radar bin collocated with the gauge (the center bin of the 3×3 array). Column 5 is the bias computed by averaging the radar rainfall for all nine bins surrounding the gauge, while the last two columns make use of the minimum and maximum radar amount in the 3×3 array. Up until now, the discussion has assumed a rain-rate threshold in the PPS algorithm corresponding to 51 dBZ, the value in use at the Denver WSR-88D on the evening of the flood, but three other thresholds ranging from 49 to 55 dBZ have also been examined and are included in the table for comparison and later discussion in section 7.

Focusing on the default case using a 51-dBZ rain-rate cap (second row in Table 1), it is clear that the choice of which of the nine radar estimates to pair with the gauge observation can have a potentially large impact on the sample bias and therefore, by deduction, on the computed bias estimate from the Kalman filter. Use of the radar estimate centered on the gauge (column 4) or the nine-bin average (column 5) results in bias estimates around 0.6 (significant radar overestimation). However, using the algorithm-chosen gauge–radar pairs (column 2), the sample bias is about 0.9, implying apparently less severe radar overestimation. Application of these larger (0.9) bias estimates to the raw radar estimates would have reduced the radar overestimation problem, but the adjusted radar rainfall would still be significantly biased even after adjustment. This was indeed the case as shown previously in the algorithm-computed bias estimates of Fig. 7a. The algorithm had computed bias estimates not much below 1.0 and thus made very little bias adjustment to the raw radar estimates that are shown from Table 1 to be significant overestimates. Because

¹ A distinction is made here between “sample biases” and “bias estimates.” The former are computed according to the above equation using the radar and gauge observations, while bias estimates refer to the computed biases from the adjustment algorithm’s Kalman filter using a time-weighted combination of sample biases from the current hour and past hours (Fulton et al. 1998).

FTG 16 UTC 12 July - 8 UTC 13 July 1996 51 dBZ

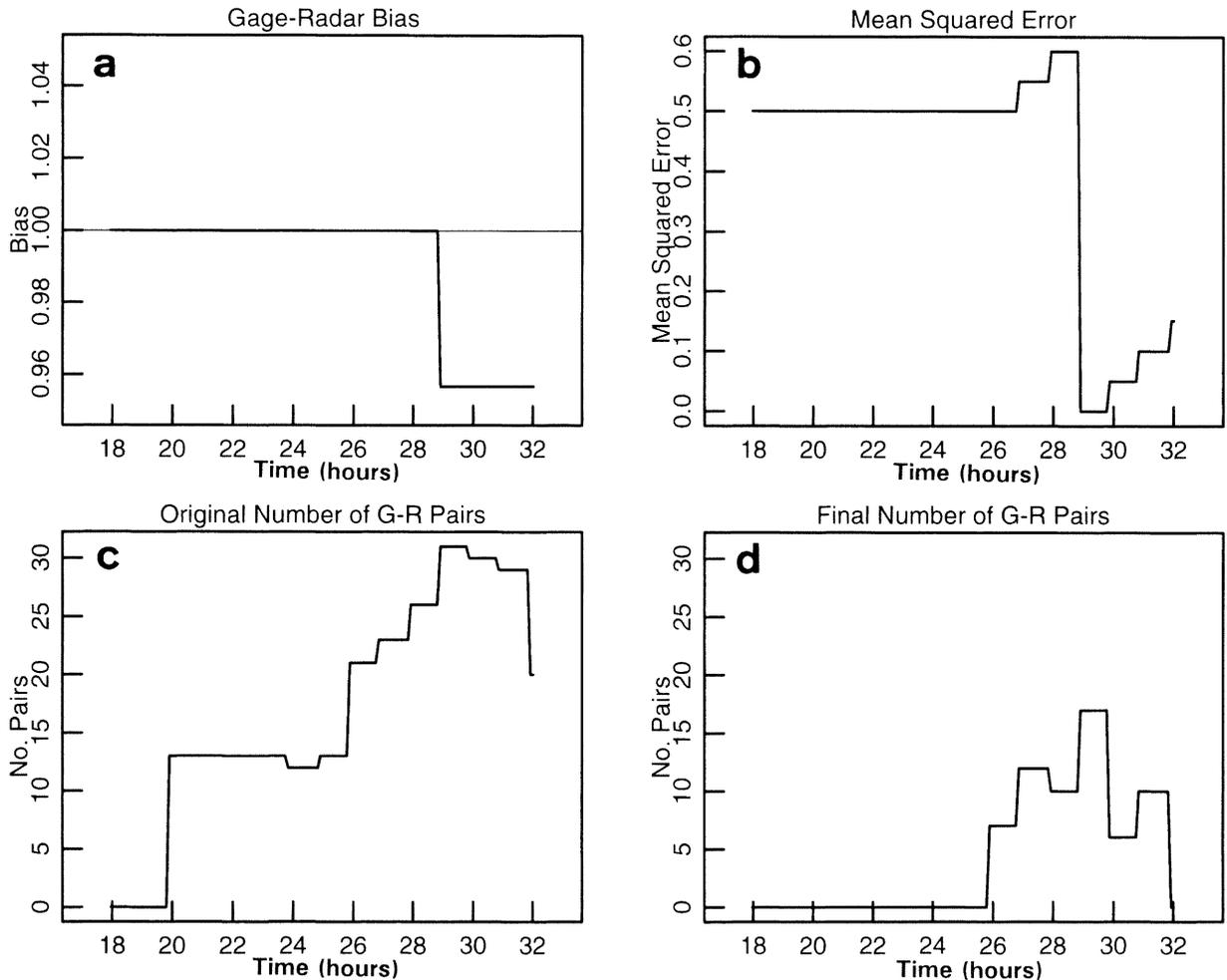


FIG. 7. Hourly time series of (a) gauge-radar bias estimate (G/R), (b) mean square error of the bias estimate, (c) total number of gauge-radar pairs available, and (d) the number of pairs that passed the quality control steps and contributed to the computed bias estimate. Times are relative to 0000 UTC 12 Jul.

the adjusted radar rainfall estimates were almost as significantly biased as the unadjusted ones, the adjustment algorithm did not perform satisfactorily.

The reason the adjustment algorithm did not perform well is due primarily to the method of forming the gauge-radar pairs. Clearly the sample bias shown in Table 1, column 2, for 51 dBZ (0.903), using the current default method, is significantly different than those in columns 4 or 5 (0.614 or 0.637) using either the collocated radar bin or the 3×3 average, and their use may have a significant impact on the total radar-estimated water volume over the FTG radar umbrella had these other methods been used to form the gauge-radar pairs.

In order to determine the impact on the computed bias estimates of using the center radar bin collocated with the gauge instead of the current logic in forming

gauge-radar pairs, the module that forms the pairs was altered. A change was made to force the use of the radar estimate that is collocated with the gauge. By comparing the resulting time series of bias estimates (Fig. 10a) using this new simpler methodology with the corresponding Fig. 7a using the existing methodology, it is clear from this case that the method by which the gauge-radar pairs are determined can have a large impact on the computed biases from the algorithm. The newly computed hourly bias estimates are much lower as desired and more in line with the storm-total sample bias of 0.614 shown in column four of Table 1. In this case the adjustment algorithm is successfully detecting the radar overestimation on an hourly timescale that is evident in the storm-total gauge-radar comparisons in Table 1. The performance of the adjustment algorithm is thus critically dependent upon the gauge-radar rainfall

Hourly Gage-Radar Pairs FTG 12-13 July 1996 51 dBZ

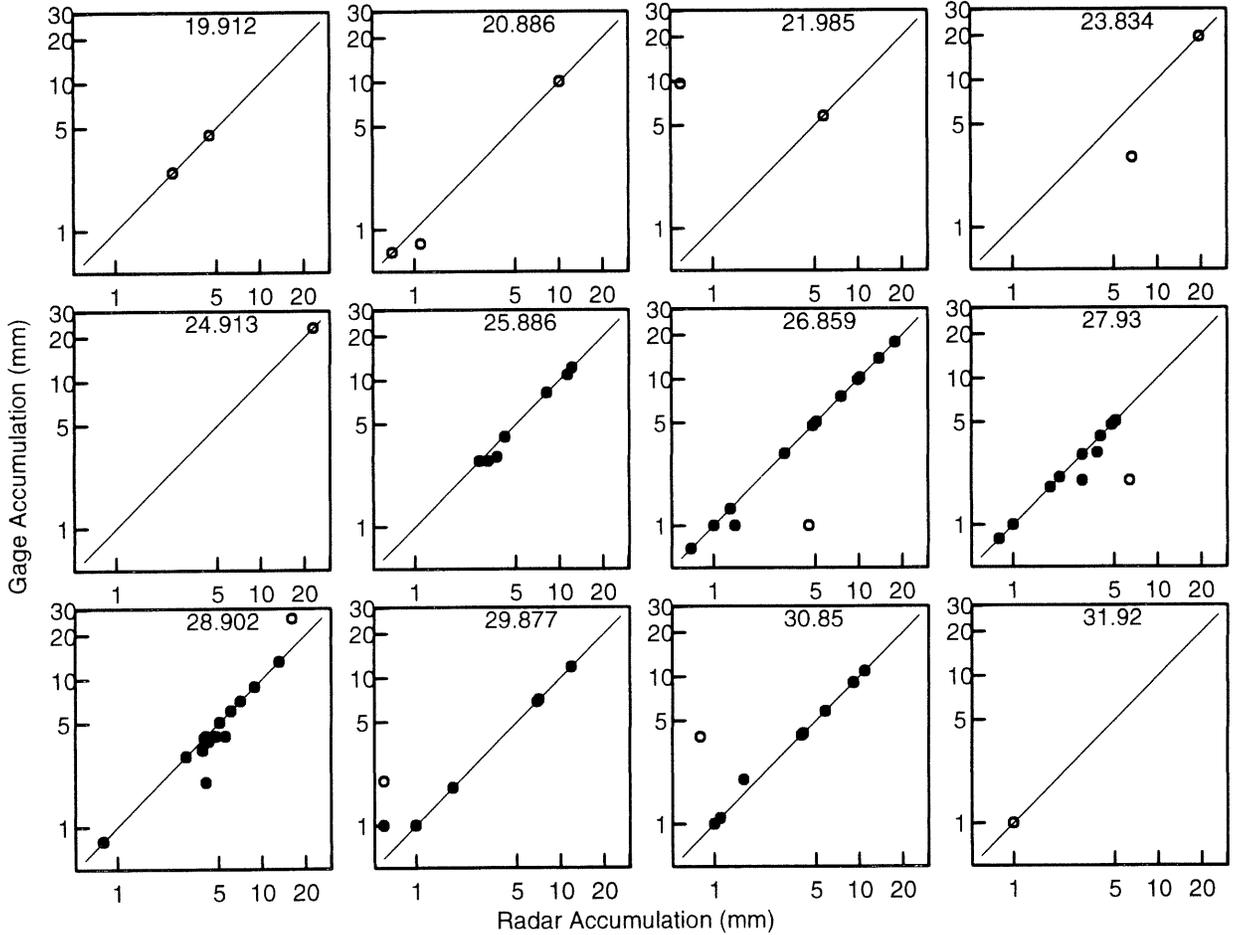


FIG. 8. Log-log scatterplots of the hourly gauge and radar rainfall amounts for all available pairs. Filled (open) circles represent pairs that were (were not) included in the computed bias estimate for that hour because they passed (did not pass) the two quality control steps. The time of bias estimation in decimal hours relative to 0000 UTC 12 July is shown at the top of each panel. The hour for which the pairs are valid ends at the time corresponding to the whole decimal hour that results after that time is truncated (e.g., the first panel corresponds to the hour period from 1800 to 1900 UTC, and the bias estimation was performed at hour 19.912, or 1955 UTC).

data that are passed into it, and those data are dependent on the method by which the pairs are formed.

By comparing the scatterplots of hourly gauge-radar pairs where the radar estimate chosen is the center bin over top of the gauge (Fig. 11) with Fig. 8, the previous neat alignment of many of the gauge-radar pairs along the 1:1 diagonal line is replaced with more widely scattered points more typical of what would normally be expected. Also, it is easy to see that the radar is overestimating in general compared to the rain gauges.

Figure 12 summarizes the results from this section by showing the net impact on the gauge-adjusted storm-total rainfall of changing the method by which the gauge-radar pairs are formed. The adjusted radar estimates associated with each pair of points in this figure reflect the multiplication of the computed hourly bias estimates from the adjustment algorithm as shown in

Figs. 7a and 10a with the radar rainfall estimates exactly as would have been done if the adjustment algorithm had been executing on an hourly basis over the course of the storm event. The open circles are associated with the adjusted radar estimates using the existing method to form the pairs, while filled circles are associated with the adjusted radar estimates using the center bin of the 3×3 array.

Focusing first on the results from the existing algorithm (open circles), the storm-total sample bias after adjustment is 0.615, equivalent to a 63% radar overestimation, which is not very close to the ideal postadjustment bias of 1.0 that the adjustment algorithm strives to achieve. This bias is slightly larger than the corresponding unadjusted value shown in Table 1 (0.614), reflecting a slight reduction in the radar overestimation. Although the adjustment algorithm has successfully

Storm Total Gage vs. Radar

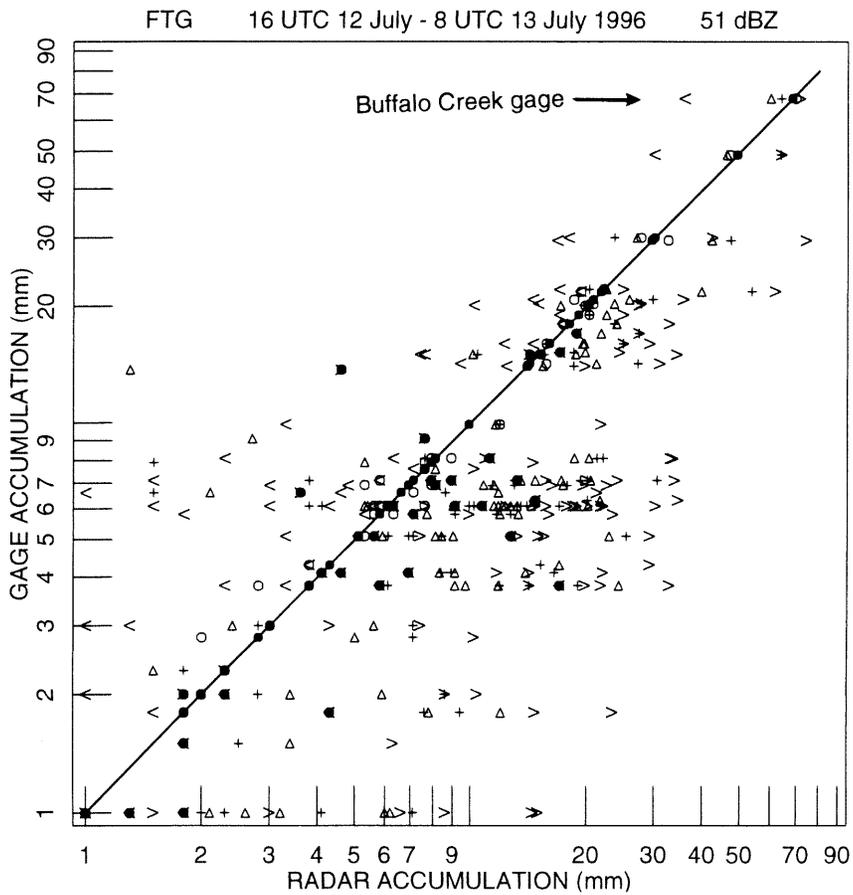


FIG. 9. Log-log scatterplots of storm-total gauge-radar pairs for the case with the rain-rate threshold set to 74.7 mm h^{-1} (51 dBZ). For each available storm-total gauge observation, a set of six symbols are plotted corresponding to the various possible radar estimates at that gauge location. The “<” and “>” symbols represent the minimum and maximum storm-total radar estimates from the nine surrounding polar 1° by 2 km bins at each gauge location. The “+” symbol represents the radar estimate in the center bin collocated with the gauge. The triangle represents the average radar estimate from the nine surrounding polar grid bins. The open circle represents the radar estimate closest in rainfall amount to the gauge observation from the nine surrounding bins. The filled circle represents the radar estimate that the adjustment algorithm would choose to pair with the gauge observation according to the rules described in the text. The top row of points associated with the largest gauge observation (68 mm = 2.68 in.) are those pairs associated with the public storm-total gauge report in the town of Buffalo Creek.

TABLE 1. Storm-total sample biases as a function of the rain-rate threshold, converted to reflectivity (leftmost column), and various radar estimates from the nine polar bins surrounding each gauge location (top row). All available gauges that reported rain were included. In column three, “closest to gauge rain” means closest in quantitative rainfall amount, not spatially closest (as is represented in column four). Sample biases less (greater) than 1.0 imply radar overestimation (underestimation) relative to the gauges.

Rain-rate threshold (dBZ)	Radar (chosen by algorithm)	Radar (3×3 closest to gauge rain)	Radar (3×3 center)	Radar (3×3 average)	Radar (3×3 min)	Radar (3×3 max)
49	0.938	0.951	0.700	0.729	1.349	0.482
51	0.903	0.906	0.614	0.637	1.225	0.409
53	0.890	0.894	0.544	0.564	1.147	0.350
55	0.874	0.860	0.491	0.508	1.096	0.306

FTG 16 UTC 12 July - 8 UTC 13 July 1996 51 dBZ Center Bin

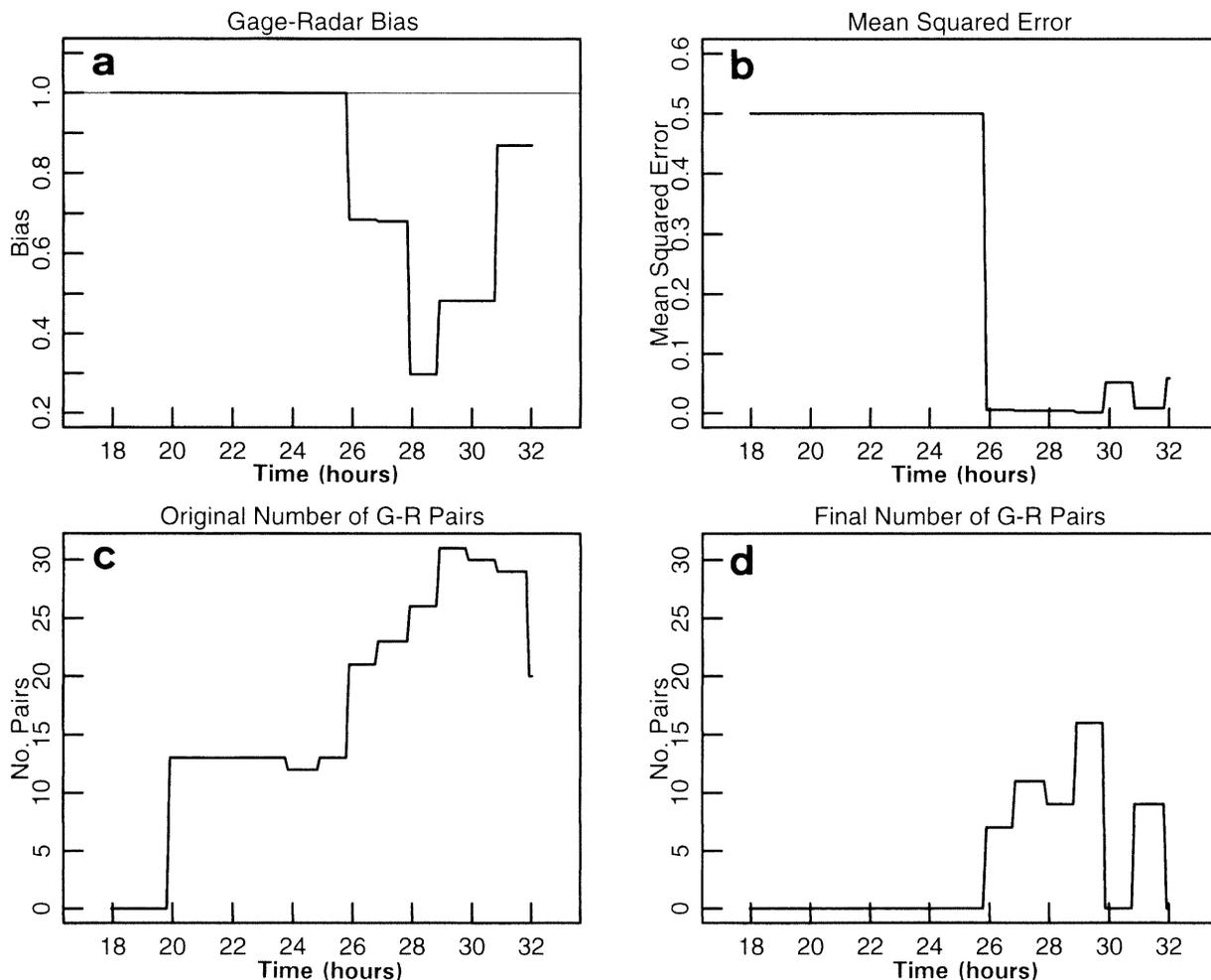


FIG. 10. Same as Fig. 7 except that the radar estimate collocated with the gauge location (the center bin of the 3×3 bins) was chosen in forming the gauge-radar pairs used in the adjustment algorithm.

moved the radar estimates in the right direction, that is, decreasing them in better accordance with the gauge observations, it has not moved them far enough as there is still a large overestimation even after automated bias adjustment has been applied.

On the other hand, the resulting adjusted storm-total sample bias associated with the new methodology (filled circles) is 0.797, about 30% larger than the previous 0.615, and closer to the ultimate goal of 1.0. A sample bias of 0.797 corresponds to a radar overestimation of 25%, which, though not totally unbiased, is much better than the previous radar overestimation of 63%. In this case, the new methodology has moved the radar estimates much farther in the right direction by reducing the overestimation and bringing the postadjustment storm-total sample bias closer to 1.0. However, since the postadjustment bias is still less than unity, there is

still some residual radar overestimation that the adjustment algorithm could not totally correct.

6. Local versus mean field gauge-radar biases

There is one last important point worthy of discussion from Fig. 12. Despite the new adjustment algorithm's reduction of the radarwide rainfall overestimation, it has inadvertently reduced the storm-total radar estimate over the town of Buffalo Creek that had previously been a relatively good estimate when no automated bias correction had been performed at all. The radar estimate in the polar bin over the town of Buffalo Creek was reduced from 63.8 mm to 43.2 mm using the new formulation, a 32% decrease. Since the rain gauge observation at this location was 68.1 mm, the net result is that for the gauge-radar pair over Buffalo Creek the

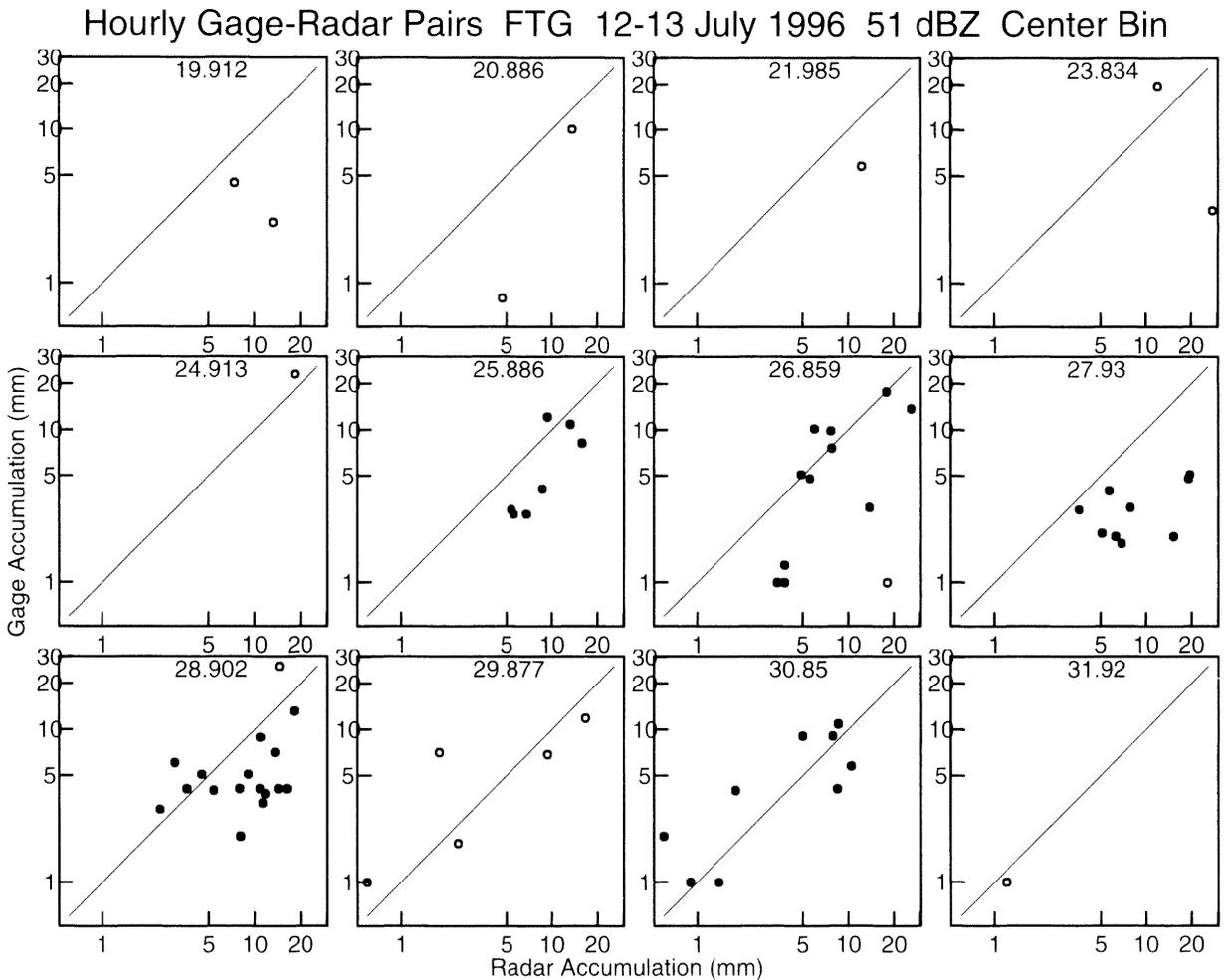


FIG. 11. Same as Fig. 8 except that the radar estimate collocated with the gauge location (the center bin of the 3×3 bins) was chosen in forming the gauge-radar pairs used in the adjustment algorithm.

radar estimates would have been degraded by the inclusion of real-time gauge data in the PPS algorithm.

This will not always be the case, but it is an unavoidable dilemma that may occur when estimating spatially uniform ("mean field") biases as in the current design of the adjustment algorithm. If the PPS performance for this case is judged based only on *radarwide* gauge-radar comparisons (e.g., sample biases computed from all available gauge data over the radar domain) as opposed to *point-specific* gauge-radar comparisons, then the adjustment algorithm would have improved the rainfall estimates relative to the case if the adjustment algorithm had not been run at all. In this case, the total water volume over the entire radar umbrella has been better estimated relative to the rain gauges (i.e., the radar overestimation has been decreased).

On the other hand, if one judges the performance of the adjustment algorithm based only on individual point-by-point comparisons of gauge and radar, for example,

at the town of Buffalo Creek, then it is unavoidable that there will be some locations where the adjusted radar estimates will be locally degraded relative to the corresponding unadjusted estimates. This is evident in Fig. 12. In general the majority of filled circles are closer to the diagonal line than the corresponding open circles (the latter of which nearly match the case if no adjustment had been done), consistent with storm-total sample biases closer to 1.0 as described previously. However, some filled circles have moved farther away from the diagonal line, implying a local degradation of the adjusted estimate. The net result is that the integrity of radar estimates are improved over the majority of the radar domain at the expense of reducing their integrity at a few locations. This is rooted quite simply in the fact that we place greater statistical faith in the gauge-radar bias information content of a large group of gauge-radar pairs than we do in just a few. However, the desire for a larger set of pairs necessarily implies

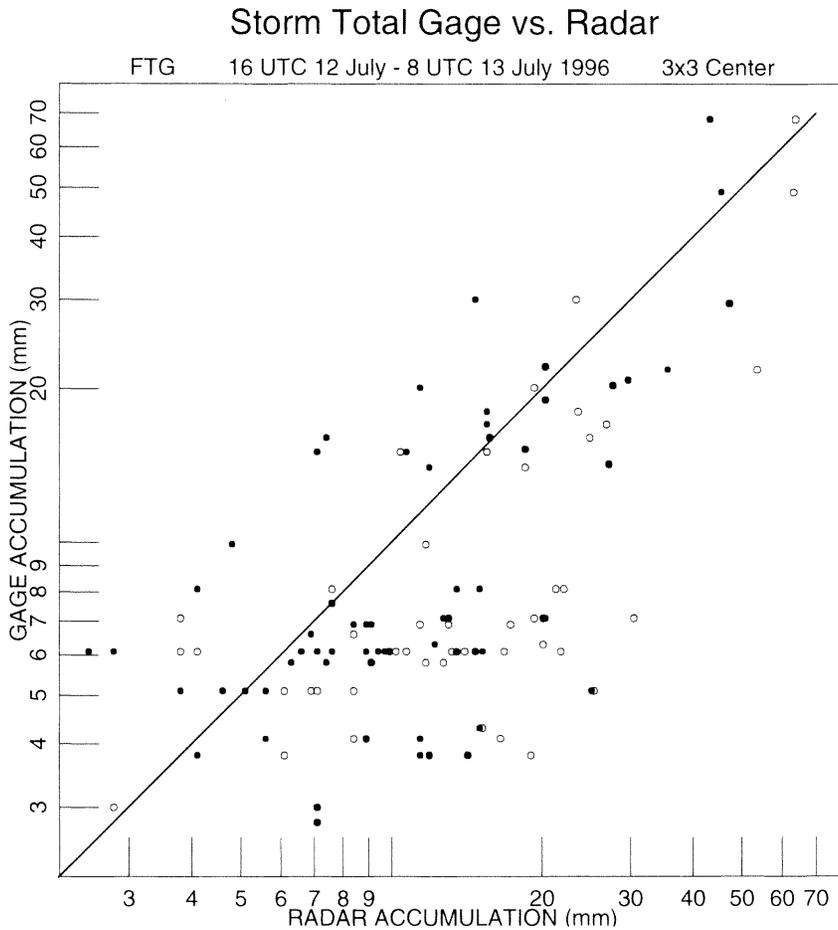


FIG. 12. Log-log scatterplots of adjusted storm-total gauge-radar pairs for the rain-rate threshold of 51 dBZ for all available gauge data. Open circles are associated with the PPS run in which the existing method for forming pairs was used. Filled circles are associated with the PPS run in which the gauge-radar pairs were formed using the 3×3 center bin. The topmost pair of points corresponds to the observation from the town of Buffalo Creek.

that the resulting bias will be valid over a correspondingly larger region, and any possible local gauge-radar bias variations are thus smoothed out.

The current adjustment algorithm is designed to perform hourly radarwide bias correction by combining all gauge-radar pairs together into a single mean field bias regardless of their relative spatial distribution. It is therefore expected to account for spatially uniform radar estimation errors such as reflectivity miscalibration, wet radome attenuation, and inappropriate $Z-R$ coefficients. However, it does not account for any possible spatial variability of gauge-radar biases that may exist across the radar domain due to 1) storm-to-storm variability in raindrop size distributions (and therefore $Z-R$ relations), 2) range-dependent variability associated with bright bands or range degradation at far ranges (this should be corrected first), and/or 3) azimuth-dependent variability due to terrain blockages not previously accounted for.

If the magnitude of these spatially nonuniform errors is small relative to the spatially uniform errors, then the mean field bias corrections should by themselves improve estimates at both small and large scales.

It is not always necessary that one has to sacrifice radar rainfall accuracy at large scales to obtain accuracy at small scales, or vice versa, as that will depend on the relative magnitude of all relevant radar error sources for each particular rain event and their spatial variability across the radar domain. Spatially varying bias adjustment techniques are being examined for possible future implementation at radar sites where abundance of real-time gauge data justify their use. It is also not necessarily true that mean field bias adjustment techniques will always degrade rainfall estimates where the flash floods occur as in this event. This may be the case, however, if the flash flood-producing storm cell differs significantly in its microphysical nature from the other rain

cells in the radar domain such that the mean field bias computed over the entire domain is not representative for that particular storm cell.

More work needs to be done in this area to determine if flash flood-producing storms are different microphysically and dynamically from their surrounding storms. This is supported by Bauer-Messmer et al. (1997), the results of this case, and the Fort Collins, Colorado, flash flood of 28 July 1997. In the Buffalo Creek case, the radar rainfall estimate associated with the flash flood-producing storm was quite good, while in general the estimates in other parts of Colorado were poor. This contrasts to the Fort Collins flood in which the radar estimates in Fort Collins were locally underestimated while the estimates in other parts of Colorado were quite good (Weaver et al. 1998). Both events suggest that the flood-producing storms are not always similar microphysically or dynamically to storms occurring simultaneously in other parts of the radar domain, and that would present difficulties to radar rainfall algorithms like the PPS that compute mean field bias adjustment factors. Perhaps the use of neural networks for radar rainfall estimation might add value in this area (Bringi et al. 1998).

7. Sensitivity of PPS estimates to the rain-rate threshold

As stated previously in section 1, there were hail reports as large as 1.75 in. in Colorado on this day. The PPS contains a rain-rate threshold parameter that caps the rate (or equivalently the reflectivity factor through the Z - R relation) that is considered to be the maximum likely associated with rain. Since reflectivity associated with hail has been observed to exceed 65 dBZ in intense storms and because the associated rain rate becomes unrealistically inflated, all radar bins with rates above the chosen threshold are lowered to it. The default threshold in use at the Denver WSR-88D on the day of this case study was 75 mm h^{-1} (51 dBZ). The maximum observed reflectivity in the lowest unblocked elevation angle (1.5°) associated with the Buffalo Creek storm was 62.5 dBZ (see section 9). The degree to which this rain-rate threshold impacts the derived rain estimates in Buffalo Creek and over all of Colorado is examined in this section.

Rutledge et al. (1998) have examined the sensitivity of rainfall accumulation to the rain-rate threshold for two heavy rain events in June and July in Colorado and found two different optimum values corresponding to 53 and 55 dBZ using the same default WSR-88D Z - R relationship as in this study. Petersen et al. (1999) found little impact on radar estimates when a 53-dBZ maximum threshold was applied for the Fort Collins, Colorado, flash flood event of July 1997. Baeck and Smith (1998) discuss rainfall errors associated with the rain-rate threshold for several intense thunderstorms across

the United States, and Glitto and Choy (1997) do the same for two tropical storms in Florida.

A gauge-radar scatterplot of unadjusted storm-total rainfall for the Buffalo Creek event for all gauges within 230 km of FTG (Fig. 13) illustrates the impact of various rain-rate thresholds on the storm-total rainfall. Table 2 defines the conversion between reflectivity and rain rate. In agreement with previous conclusions, Fig. 13 shows that the radar has overestimated the rainfall relative to the gauges as evidenced by the majority of points to the right and below the 1:1 diagonal line. This overestimation is more pronounced as the rain-rate threshold increases from 49 to 55 dBZ, which is expected if hail is contaminating the radar rainfall estimates. Referring back to the sample biases shown in column four of Table 1 (corresponding to Fig. 13), the storm-total sample biases decrease from about 0.7 to 0.5 as the rain-rate threshold increases from 49 to 55 dBZ, indicating increasing radar overestimation. Even with the lowest setting of 49 dBZ the radar estimates for this case are still too high relative to the gauges by at least 40% (0.7^{-1}) on the average. Thus one is not able to explain the radar overestimation for this case based only on a hail cap threshold that should have been lower. Other factors must be important.

As previously mentioned, it turns out that the radar estimate over the town of Buffalo Creek (the top row of symbols in Fig. 13) was very close to the actual rain gauge measurement for the case when the default value of 51 dBZ was used as a rain-rate threshold. The rain gauge measured 2.68 in. while the PPS estimated 2.51 in., which is only a 6% underestimate. The range of PPS estimates for the nine surrounding polar bins ranges from 1.41 to 2.81 in. (see Table 3). The largest storm-total PPS rainfall estimate at any location affected by the flood-producing storm was 2.85 in., only two polar grid boxes away from the town of Buffalo Creek, at an azimuth of 232.5° and range of 77 km.

8. Other possible error sources

The integrity of PPS rainfall estimates may be affected by other meteorological and/or algorithmic factors besides the rain-rate threshold. These include Z - R parameters, radar reflectivity calibration, evaporation below the beam, brightband contamination, or some contributions from all of these.

The radar overestimation bias for this case could be fully explained using a Z - R relation different from the default $Z = 300R^{1.4}$. Using the observed storm-total gauge-radar mean field bias, one can easily estimate an "effective" or "true" value of the multiplicative Z - R coefficient, A , in the standard power law relation $Z = AR^b$ that would produce unbiased radar rainfall estimates by assuming the power coefficient, b , is fixed. Doelling et al. (1998) and Smith and Joss (1997) recommend that b be fixed somewhere in the range 1.4 to 1.6 and A be allowed to vary to account for natural variations in the

Storm Total Gage vs. Radar

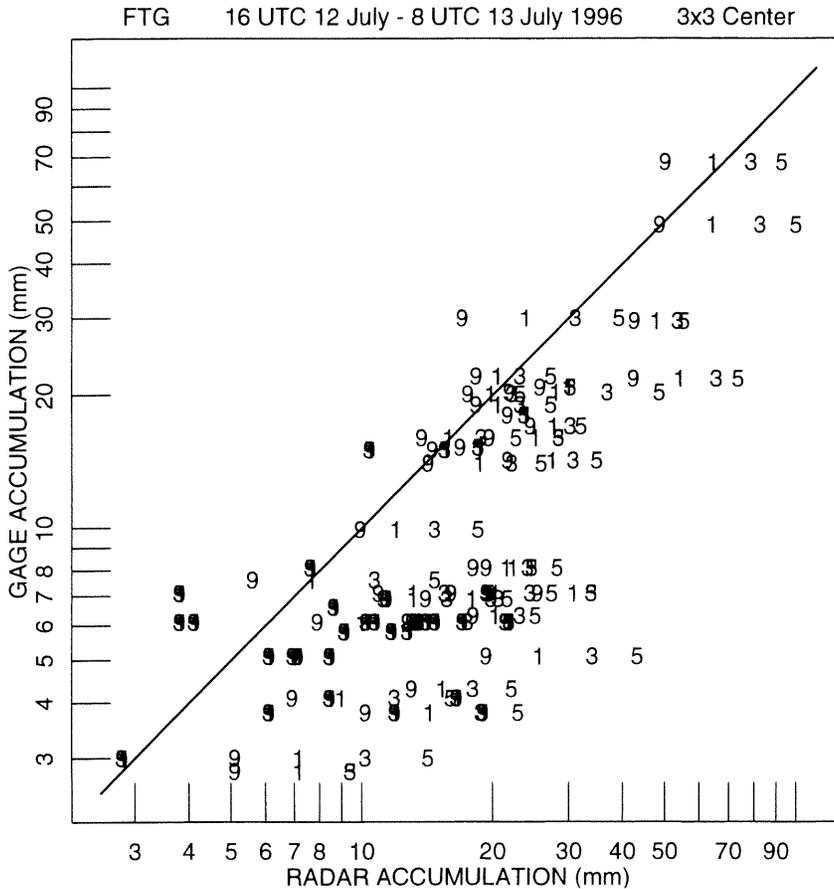


FIG. 13. Log-log scatterplot of storm-total gauge-radar pairs for four different rain-rate thresholds corresponding to 49, 51, 53, and 55 dBZ, denoted respectively as "9," "1," "3," and "5." The radar estimates used for each pair are the center of the 3 × 3 bins surrounding the gauge locations. The radar estimates have not been adjusted by any bias estimates.

drop size distribution of rainfall because the variations in *b* are typically small compared to *A*:

$$Z_{88D} = AR_{88D}^b \tag{1}$$

where Z_{88D} is the reflectivity factor ($\text{mm}^6 \text{m}^{-3}$) measured by the WSR-88D, R_{88D} is the derived rain rate (mm h^{-1}), and *A* and *b* are the default values as used in the PPS. If we define the observed gauge-radar mean field bias, $B = \Sigma G / \Sigma R$, by summing over all gauge-radar pairs for some chosen time period, then

$$R_{\text{true}} = R_{88D}B, \tag{2}$$

where R_{true} is the true, unknown, unbiased radar rain rate that we really want to know. Solving (2) for R_{88D} and substituting into (1) produces

TABLE 2. Conversions between reflectivity and rain rate assuming $Z = 300R^{1.4}$.

Reflectivity threshold (dBZ)	Rain-rate threshold (mm h^{-1} ; in h^{-1})
49	53.8; 2.12
51	74.7; 2.94
53	103.8; 4.09
55	144.4; 5.68

TABLE 3. Unadjusted PPS storm-total rainfall estimates (in.) for the 5 × 5 polar grid bins surrounding the town of Buffalo Creek, CO (located at the central bin with the bold italic "2.51"), using the 51-dBZ rain-rate threshold, shown at the raw spatial resolution of the PPS algorithm. Columns are the center azimuths in °, and rows are the center ranges in km of each polar 1° × 2 km rainfall bin. The period of rainfall is from 1758 UTC 12 Jul to 0800 UTC 13 Jul.

Range (km)	Azimuth (°)				
	232.5	233.5	234.5	235.5	236.5
81	1.13	1.51	1.89	2.35	2.43
79	2.11	2.76	2.81	2.74	2.49
77	2.85	2.79	2.51	2.17	1.80
75	2.44	2.19	1.86	1.41	1.05
73	1.47	1.63	1.20	0.86	0.48

$$Z_{88D} = AB^{-b}R_{\text{true}}^b = A'R_{\text{true}}^b, \quad (3)$$

where the gauge-adjusted multiplicative coefficient is $A' = AB^{-b}$. It represents the fixed coefficient A multiplied by a time-dependent correction factor based on observed gauge data. Thus the application of an observed, real-time multiplicative gauge–radar bias to radar rainfall estimates computed using the default, fixed Z – R relation as in (2) is analogous to adjusting the A coefficient in $Z = AR^b$ in real time and would produce equivalent results.

The observed storm-total mean field bias $B = 0.614$ from Table 1 would translate into an effective multiplicative Z – R coefficient of $A' = 300(0.614)^{-1.4} = 594$. Thus if the equation $Z = 594R^{1.4}$ had been used, the PPS would have produced unbiased radar estimates in the mean relative to the gauges for this rain event. However, such a large coefficient is somewhat outside of the range of historically estimated coefficients as reported in Battan (1973); therefore, other error sources may also be contributing.

While the PPS estimated the rainfall for the Buffalo Creek storm relatively well and overestimated it in general for all storms in the radar domain on that day, the PPS underestimated the rainfall for the Fort Collins, Colorado, flash flood that occurred almost exactly one year later (Petersen et al. 1999; NWS 1997). When a tropical Z – R relationship ($Z = 250R^{1.2}$) was tested for that event, the estimates were improved. Use of this tropical relation for the Buffalo Creek flood event would have resulted in even greater overestimation errors than observed here with the WSR-88D default relation. This suggests that there is no one single relationship that will consistently produce good results for all flash flood events in Colorado or elsewhere. One relation that performs well for one event may badly reproduce the observed rainfall for another event even though they may occur during the same season in the same geographic region.

Another possible error source for the Buffalo Creek event is the reflectivity factor calibration of the WSR-88D. The observed gauge–radar sample bias of 0.614, for example, could be explained by a radar in which the reflectivities were too high by 3.0 dB assuming use of the default WSR-88D Z – R relation. Unfortunately this magnitude of error is not out of the question for the WSR-88Ds. Calibration information for the FTG (or any WSR-88D) radar was not archived prior to software build 10 and was therefore not available on this day. Thus it is not possible to separate the relative contributions of possible miscalibration and improper Z – R coefficients on the radar overestimation errors for this event, though it appears that both are contributing factors for this case. Ulbrich et al. (1996) examines the relative impact of changes in the Z – R coefficients and reflectivity calibration in explaining the tendency for WSR-88D underestimation for tropical rainfall. They conclude that a combination of the two error sources

can explain the underestimation when the sign of the error is identical and thus additive.

Another possible explanation for the radar overestimation in this case is evaporation of the rain observed above the ground as it falls down to the gauges below (Rosenfeld and Mintz 1988). This is common in drier climates such as Colorado often dominated by “inverted V” sounding profiles. However, the ambient environment was much moister than usual in Colorado on this day with surface dewpoint temperatures above 60°F and relatively high dewpoints below cloud base (Fig. 1). Therefore below-beam evaporation would not be expected to play as large a role as it usually does in Colorado in explaining radar overestimation relative to the gauges on this day, though it cannot be totally ruled out in contributing to the observed radar overestimation, particularly in the far western portions of the radar scanning domain where the lower atmosphere was drier (Fig. 1a). This argument could be generalized to suggest that because flash flood-producing storms in Colorado or other similarly semiarid regions are usually accompanied by anomalously moist lower-atmospheric sounding profiles (Petersen et al. 1999; Bauer-Messmer et al. 1997; Caracena and Fritsch 1983; Maddox et al. 1978), evaporation would not be expected to be a primary contributor to radar rainfall overestimation errors on these days.

Finally, it is possible that some overestimation of the rainfall estimates could perhaps be explained by brightband contamination of the reflectivity. Enhanced reflectivity often occurs in a narrow vertical zone just below the melting level associated with increased backscatter from water-coated, aggregated snowflakes or graupel. However, brightband signatures are often non-existent in convective clouds compared to stratiform clouds due to the larger magnitudes of vertical velocity associated with deep convection typical of the Buffalo Creek storm (Steiner et al. 1995), and in fact there was no observed annulus of enhanced radar rainfall in the PPS storm-total rainfall accumulations in the limited range interval of the radar domain where the radar beam intersects the freezing level to lend support to such a proposition. Additionally, general radarwide overestimation as observed in this case could not be explained by brightband enhancement that would occur only over limited regions of the radar scanning domain. The vertical radar beam cross section in Fig. 4 previously showed that at the 77-km range of the rainfall centroid of the Buffalo Creek storm near the outlet of the basin (see Fig. 2), the reflectivity measurements were collected primarily below the freezing level. And in fact, the PPS estimates were quite good there.

Both the rain-rate threshold and Z – R parameters likely change over relatively small spatial and temporal scales in ways not fully understood (Rutledge et al. 1998), and they may be a function of atmospheric temperature and/or moisture sounding variables on any given day (Kelsch 1992). Optimizing these parameters as

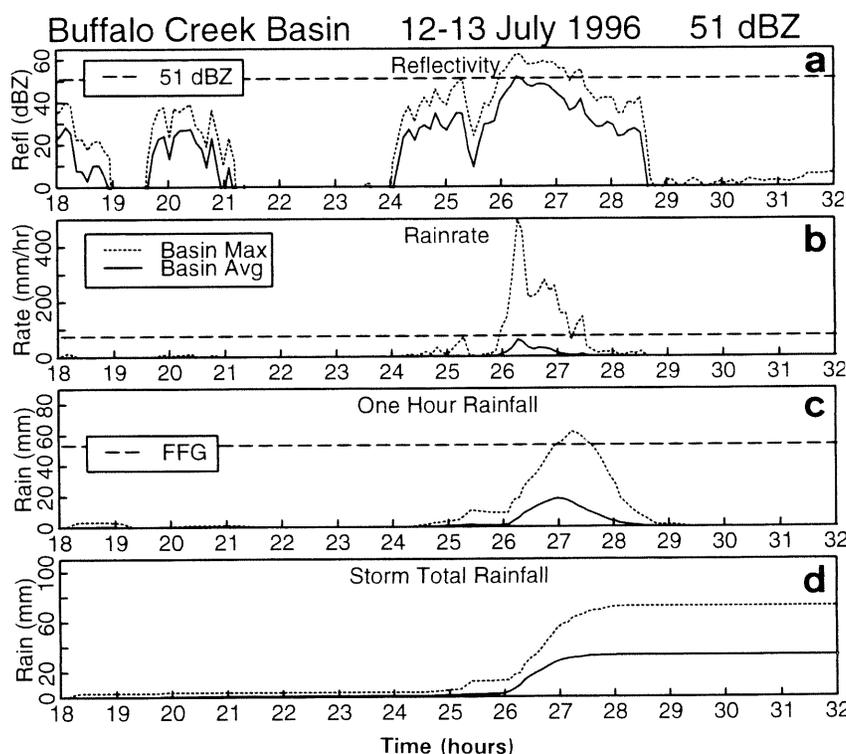


FIG. 14. Time series plots of basin-maximum (dotted lines) and basin-averaged (solid lines) values of (a) reflectivity, (b) rain rate, (c) 1-h rainfall, and (d) storm-total rainfall. Rainfall in this figure has not been adjusted by any bias estimates. The dashed lines in (a) and (b) represent the equivalent maximum rain-rate thresholds of 51 dBZ and 74.7 mm h^{-1} , respectively. The dashed line in (c) represents the 1-h Flash Flood Guidance rainfall. Times are in hours relative to 0000 UTC 12 Jul.

a function of storm type, geographic location, season, etc., across the entire United States is necessary but may be unattainable. And unfortunately the transportability of locally optimized parameter values from one time or location to another is questionable.

This is one area where dual-polarization radar technology has potential due to its increased ability to distinguish hydrometeor sizes and types such as rain versus hail and its lesser sensitivity to temporal and spatial variations in drop size distributions in rain clouds (Zrnić and Ryzhkov 1999). As a result, the dilemma of determining appropriate rain-rate thresholds and/or $Z-R$ parameters in real time associated with the WSR-88D PPS single polarization rainfall estimation technology may be partially sidestepped when using dual-polarization information such as specific differential phase and differential reflectivity measurements (Fulton et al. 1999; Ryzhkov and Zrnić 1996a,b).

9. Basin-averaged rainfall

Point measurements of radar and gauge rainfall have been compared thus far in this analysis. This section will briefly examine the basin-averaged rainfall as derived from the PPS. Time series plots of basin-average

and basin-maximum reflectivity, rain rate, 1-h rainfall and storm-total rainfall over the Buffalo Creek basin for the 14-h time period encompassing the flash flood event are shown in Fig. 14. Basin-average and basin-maximum values are computed by considering all $1^\circ \times 2 \text{ km}$ PPS grid cells whose centers fall within the basin. The heaviest rain occurred during the 1-h period from 0200 to 0300 UTC (hours 26–27 in Fig. 14), and the flash flood occurred around 0300–0400 UTC. According to public reports in the town, the Buffalo Creek had begun to rise around 0230 UTC and was out of its banks by 0300 UTC. The river had risen about 15 ft in the town and had returned to near-normal levels by midnight (0600 UTC). Light rain had passed over the basin earlier in the day but had contributed only about 4 mm (0.16 in.) to the basin maximum rainfall.

Had the rainfall rates not been capped at 51 dBZ ($=75 \text{ mm h}^{-1}$) in the PPS (dashed horizontal lines in Figs. 14a and 14b), the radar estimates of rainfall would have been significantly higher than plotted as evidenced by the large area of maximum rain rate that lies above the threshold in Fig. 14b (see also Fig. 13). As an example, if the rain-rate threshold had been set at 144 mm h^{-1} or 5.7 in h^{-1} (equivalent to 55 dBZ), the basin-maximum storm-total rainfall would have increased from 72 mm

TABLE 4. Storm-total radar-derived rainfall (in.) from 1758 UTC 12 Jul to 0801 UTC 13 Jul 1996 for all $1^\circ \times 2$ km polar bins whose centers lie within the Buffalo Creek basin. The rain-rate threshold is 51 dBZ, and no bias adjustment is performed. Columns are the center range (km) and rows are the center azimuths ($^\circ$) of each polar bin. Water flows to the right in general, and the basin outlet is near the underlined bin, which is also the same bin where the town of Buffalo Creek is located. The storm moved in general from the top to the bottom of the page. The fire burn area is located approximately where the rainfall totals are in bold type.

Azi- muth ($^\circ$)	Range (km)									
	93	91	89	87	85	83	81	79	77	75
236.5			0.76	0.77						
235.5		0.35	0.59	0.69	0.95	1.57	2.35			
234.5	0.18	0.33	0.39	0.62	0.82	1.20	1.89	2.81	<u>2.51</u>	
233.5	0.12	0.27	0.39	0.48	0.58	0.77	1.51	2.76	2.79	2.19
232.5		0.18	0.43	0.39	0.42	0.39	1.13	2.11	2.85	2.44
231.5		0.15	0.34	0.41	0.65	0.22	0.66	1.59	2.65	
230.5						0.53				

(2.85 in.) to 104 mm (4.09 in.). The maximum reflectivity at any bin within the basin exceeded 51 dBZ for an hour and a half, from 0159 to 0327 UTC, and the largest observed hybrid scan reflectivity of 62.5 dBZ occurred at 0217 UTC. This maximum reflectivity would correspond to an unreasonably large instantaneous rain rate of 19.5 in. h^{-1} using the default $Z-R$ relationship and would likely be resulting from hail contamination of the reflectivity.

Basin-maximum 1-h rainfall peaked at 0315 UTC at 62 mm (2.44 in.), while basin-averaged 1-h rainfall never exceeded 18 mm (0.72 in.) (Fig. 14c). This large discrepancy is a result of the fact that the heaviest rainfall occurred only in the lowest portion of the basin where the town of Buffalo Creek is located, which, coincidentally, is also the portion of the basin that had been ravaged by a forest fire two months earlier (see Table 4 and Fig. 2). The lack of vegetation in the burn area was suggested to be an important reason for the severity of the flash flood (see the appendix).

The 1-h Flash Flood Guidance (FFG) rainfall valid for Jefferson County as issued by the NWS Missouri Basin River Forecast Center earlier in the day was 53.3 mm (2.1 in.; see dashed line in Fig. 14c). FFG rainfall represents the approximate amount of basin-averaged rainfall over a given duration that is needed to bring small streams to bankfull (Sweeney 1992). It is estimated daily for rainfall durations of 1, 3, and 6 h by the operational hydrologic forecast models at the RFCs. NWS forecasters often compare the model-estimated FFG rainfall with observed gauge- or radar-derived rainfall in real time to assess flash flood risk. Since this rainfall event was of approximately 1-h duration, the 1-h FFG is the appropriate one to examine here.

The basin-maximum 1-h rainfall exceeded the 1-h FFG from 0257 to 0332 UTC; however, the basin-averaged rainfall at its peak at 0303 UTC was about three times smaller than the 1-h FFG. As a result, a comparison of FFG with basin-maximum radar rainfall instead of basin-averaged rainfall would have been more appropriate in alerting forecasters to the potential flood threat for this case. The basin-maximum 1-h rainfall first

exceeded the FFG value beginning at 0257 UTC, which corresponds closely to the time when the public reported that the Buffalo Creek came out of its banks. The intensity and small size of this rainfall event likely made the model-estimated FFG rainfall inappropriate for this event even if the unmodeled effect of the forest fire on the infiltration characteristics and resulting hydrologic runoff response of the basin had been considered.

10. Conclusions

The quantitative performance of the WSR-88D rainfall algorithm, PPS, is examined for a flash flood event along the Buffalo Creek in the town of Buffalo Creek, Colorado, in July 1996. A total of 145 rain gauges were available to serve as a comparison to the radar estimates derived from the Denver WSR-88D radar, though only one actually sampled the flood-producing storm. Over the 230-km range domain of the WSR-88D, the PPS overestimated the rainfall relative to the rain gauges by about 60%; however, the radar estimate over the town of Buffalo Creek where the flood deaths and property damage occurred was within 6% of the gauge observation of 2.68 in. collected by a citizen of the town living on the banks of the river.

A default rain-rate threshold of 75 mm h^{-1} (corresponding to 51 dBZ) was in use in the PPS adaptation data at the Denver WSR-88D radar on this day. This parameter is used in the algorithm to cap the rain rates to prevent typically large hail reflectivities from producing unrealistically large rain rates. The sensitivity of the derived radar rainfall to the setting of this parameter was examined to determine the extent to which overestimation could be explained. The radar estimates would still have been overestimated on the average even if it were lowered to 54 mm h^{-1} (49 dBZ) or even lower. Thus radar overestimation cannot be totally corrected by lowering the rain-rate threshold for this case.

Other potential causes for the radarwide overestimation are suggested to be a combination of several error sources including the use of a $Z-R$ relationship that may not have been optimal for the environment and

storms on that day and improper radar hardware calibration of the reflectivity measurements. Below-beam evaporation and brightband contamination may have also played minor roles. Quantitatively estimating the relative contribution of these error sources is difficult if not impossible to do without calibration information and is thus not attempted in this study.

The PPS adjustment algorithm was evaluated using the available gauge data to determine if it could properly correct (i.e., reduce) the radar overestimates to bring them more in line with the gauge estimates as it is designed to do. The algorithm performed suboptimally, not due to any deficiency in the Kalman filter formulation, but instead due to the particular methodology employed in the algorithm when it forms the hourly gauge-radar pairs that are input to the filter. The algorithm did not make large enough bias corrections to the radar estimates to fully account for the observed biases because the gauge-radar pairs it generated as input were themselves biased by this procedure. The algorithm computed bias estimates that moved the radar estimates in the right direction (i.e., decreasing them) as it should have, but it did not move them far enough to bring them into satisfactory agreement with the gauges.

An alternate and simpler methodology for forming the gauge-radar pairs by using the radar bin directly over the gauge was tested and found to significantly improve the algorithm's performance even though it did not totally remove all of the gauge-radar bias. The storm-total radar overestimation was reduced from 63% to 25% by running the adjustment algorithm with input gauge-radar pairs in which the radar estimate collocated with the gauge was chosen.

The 1-h basin-average radar rainfall never reaches more than one-third of the 1-h Flash Flood Guidance rainfall for the Buffalo Creek basin, yet a flash flood occurred. If the computed gauge-radar bias estimates had been applied to the radar estimates for this case, the basin average rainfall over the Buffalo Creek basin would have been even smaller. Comparisons of FFG and basin-maximum radar rainfall would have been more appropriate to evaluating the flash flood threat for this event due to the small space scales and timescales of the flood-producing thunderstorm relative to the larger scales at which FFG is estimated. In addition, the forest fire in the lower portion of the basin two months earlier likely altered the runoff characteristics of the basin and contributed to the severity of the flood.

Acknowledgments. The Denver Urban Drainage and Flood Control District rain gauge data were graciously supplied by Edward Brandes of the National Center for Atmospheric Research. Comments of the reviewers greatly improved the manuscript and are appreciated. This work was supported by the triagency NEXRAD Program through the WSR-88D Operational Support Facility.

APPENDIX

National Weather Service Summary of Flood Event

July 15, 1996

MEMORANDUM FOR: The Record

FROM: WSFO Denver

SUBJECT: Flood Deaths

EVENT: Flash Flood

DATE OCCURRED: Friday July 12th between 900 P.M. and midnight.

OFFICE: NWSFO Denver, CO

DEATHS: Two (43 year old male was swept away in his car; 73 year old male was swept away in a trailer he lived in).

INJURIES: None

DAMAGE: Rural area, but damage was extensive including: County Road 128 washed away, bridge washed away, firehouse destroyed, ambulance totaled, Recreation Center lifted from its foundation, several cars/tracks overturned, hundreds of mature trees uprooted, loss of electricity, telephone and water services.

WATCHES: Tornado watch in effect, but no flash flood watch.

WARNINGS: Numerous tornado warnings in effect east of the area, but no warnings (severe, tornado, or flash flood) for Jefferson County at the start of the flooding. A flash flood warning was issued by the WSFO at 1013 P.M. that was effective until 215 A.M. on July 13th.

SUMMARY OF OUR SERVICE

RADAR: The WSR-88D operated without problems during the event. Numerous thunderstorms covered northeast Colorado during the evening hours with the stronger storms all producing heavy rains, hail and tornadoes. A thunderstorm with heavy rain moved eastward out of Park County into southwest Jefferson County around 800 P.M. Radar indicated 50+ dBZ over the town of Buffalo Creek from 815 P.M.–915 P.M. Lighter showers occurred in the area through 1100 P.M. but probably did not aggravate the situation greatly. The Storm Total product indicated that 2–3 inches of rain fell, with most of it likely falling between 815 and 915 P.M. This storm had been monitored closely for severe potential. The storm was moving east at 20 mph, but slowed slightly as it moved across the area. During the heavier rains, numerous severe thunderstorm and tornado warnings were issued, with several in the Denver metropolitan area. Some locations in the Denver area received almost 3 inches of rain from 2 different storms, but only minor flooding resulted.

SATELLITE: Not Available (2½ days after event)

COMMENTS: A total of 18 severe thunderstorm and 5 tornado warnings were issued by forecasters between

600 P.M. and midnight. Two forecasters were handling the severe weather warnings while another forecaster did the routine duties and assisted in the severe weather operations, when time permitted. The storms were in the metro area between 800 P.M. and midnight, which greatly added to the workload. It should be noted that no other significant flooding occurred with the storm that struck Buffalo Creek, even though the storm was just as intense to the west and east of this location. The major fire that burned up 12 thousand acres in southwest Jefferson County 2 months ago was a big culprit in this event. Residents in the area have seen flash floods in this area before, but never to this magnitude. Unfortunately, some residents saw the Buffalo Creek raging, but never notified the local authorities or the NWS. Jefferson County Sheriff's office is on NAWAS [National Warning System] and never called us with any reports of flooding. Forecasters were aware that heavy rains would increase the chances of flash flooding in the Buffalo Creek area because of the lack of vegetation, but no one thought it would be of the magnitude we saw with as little as 2 or 3 inches of rain. We have already addressed the issue locally. Flash flood watches will be issued for Southern Jefferson Co. the rest of this convective season when thunderstorms with heavy rain are forecast. Also, when rainfall estimates within an hour reach 1 inch or more, flash flood warnings will be issued immediately. A slow moving thunderstorm with heavy rains will be enough to trigger a warning, even before the heavy rains occur.

REFERENCES

- Baeck, M. L., and J. A. Smith, 1998: Rainfall estimation by the WSR-88D for heavy rainfall events. *Wea. Forecasting*, **13**, 416–436.
- Battan, L. J., 1973: *Radar Observation of the Atmosphere*. University of Chicago Press, 323 pp.
- Bauer-Messmer, B., J. Smith, M. Baeck, and W. Zhao, 1997: Heavy rainfall: Contrasting two concurrent Great Plains thunderstorms. *Wea. Forecasting*, **12**, 785–798.
- Brandes, E., J. Vivekanandan, and J. Wilson, 1997: Radar rainfall estimates of the Buffalo Creek flash flood using WSR-88D and polarimetric radar data. Preprints, *28th Conf. on Radar Meteorology*, Austin, TX, Amer. Meteor. Soc., 123–124.
- Bringi, V. N., J. Ran, and V. Chandrasekar, 1998: Polarimetric radar and surface observations of a flash flood. *Proc. IGARSS '98*, Seattle, WA, 144–146.
- Caracena, F., and J. Fritsch, 1983: Focusing mechanisms in the Texas Hill Country flash floods of 1978. *Mon. Wea. Rev.*, **111**, 2319–2332.
- Crum, T., R. Alberty, and D. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.*, **74**, 645–653.
- Doelling, I., J. Joss, and J. Riedl, 1998: Systematic variations in Z-R relationships from drop size distributions measured in northern Germany during seven years. *Atmos. Res.*, **47–48**, 635–649.
- Fulton, R., J. Breidenbach, D.-J. Seo, D. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting*, **13**, 377–395.
- , A. Ryzhkov, and D. Zrnić, 1999: Areal rainfall estimation using conventional and polarimetric radar methods. Preprints, *29th Conf. on Radar Meteorology*, Montreal, PQ, Canada, Amer. Meteor. Soc., 293–296.
- Glitto, P., and B. Choy, 1997: A comparison of WSR-88D storm total precipitation performance during two tropical systems following changes to the multiplicative bias and upper reflectivity threshold. *Wea. Forecasting*, **12**, 459–471.
- Kelsch, M., 1992: Estimating maximum convective rainfall rates for radar-derived accumulations. Preprints, *Fourth AES/CMOS Workshop on Operational Meteorology*, Whistler, BC, Canada, Amer. Meteor. Soc., 161–168.
- Maddox, R., L. Hoxit, C. Chappell, and F. Caracena, 1978: Comparison of meteorological aspects of the Big Thompson and Rapid City flash floods. *Mon. Wea. Rev.*, **106**, 375–389.
- NOAA, 1996: *Storm Data*. Vol. 38, No. 7.
- NWS, 1997: The Fort Collins flash flood of July 28, 1997: Service assessment initial report. NOAA National Weather Service, 27 pp. [Available from NWS Central Region Headquarters, Kansas City, MO 64106.]
- O'Bannon, T., 1997: Using a "terrain-based" hybrid scan to improve WSR-88D precipitation estimates. Preprints, *28th Conf. on Radar Meteorology*, Austin, TX, Amer. Meteor. Soc., 506–507.
- Petersen, W., and Coauthors, 1999: Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull. Amer. Meteor. Soc.*, **80**, 191–216.
- Rosenfeld, D., and Y. Mintz, 1988: Evaporation of rain falling from convective clouds as derived from radar measurements. *J. Appl. Meteor.*, **27**, 209–215.
- Rutledge, S., V. N. Bringi, V. Chandrasekar, J. Hubbert, S. Bolen, P. Kennedy, C. Butler, and T. Saxen, 1998: Improving estimates of precipitation rates and hydrometeor classification using polarimetric radar: Implications for NEXRAD. Preprints, *Symp. on the Research Foci of U.S. Weather Research Program*, Phoenix, AZ, Amer. Meteor. Soc., 532–534.
- Ryzhkov, A., and D. Zrnić, 1996a: Rain in shallow and deep convection measured with a polarimetric radar. *J. Atmos. Sci.*, **53**, 2990–2995.
- , and —, 1996b: Assessment of rainfall measurement that uses specific differential phase. *J. Appl. Meteor.*, **35**, 2080–2090.
- Smith, P., and J. Joss, 1997: Use of a fixed exponent in "adjustable" Z-R relationships. Preprints, *28th Conf. on Radar Meteorology*, Austin, TX, Amer. Meteor. Soc., 254–255.
- Steiner, M., R. Houze, and S. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteor.*, **34**, 1978–2007.
- Sweeney, T., 1992: Modernized areal flash flood guidance. NOAA Tech. Memo. NWS HYDRO 44, National Weather Service/Office of Hydrology, 21 pp. [Available from National Weather Service, Office of Hydrology, Silver Springs, MD 20910.]
- Ulbrich, C., J. Pelissier, and L. Lee, 1996: Effects of variations in Z-R law parameters and the radar constant on rainfall rates measured by WSR-88D radars. Preprints, *15th Conf. on Weather Analysis and Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 316–319.
- Weaver, J., W. Petersen, and N. Doesken, 1998: Some unusual aspects of the Fort Collins, Colorado flash flood of 28 July 1997. Preprints, *Eighth Conf. on Mountain Meteorology*, Flagstaff, AZ, Amer. Meteor. Soc., 310–316.
- Zrnić, D., and A. Ryzhkov, 1999: Polarimetry for weather surveillance radars. *Bull. Amer. Meteor. Soc.*, **80**, 389–406.