Archival precipitation data set for the Mississippi River Basin: Evaluation

Brian R. Nelson,[1,2] Witold F. Krajewski,[1] James A. Smith,[3] Emad Habib,[4] and Gerrit Hoogenboom[5]

Received 2 May 2005; revised 29 July 2005; accepted 18 August 2005; published 21 September 2005.

[1] The goals of the Global Energy and Water Cycle Experiment Continental-Scale International Project (GCIP) point to the need for high resolution data sets on all elements of the land surface and atmospheric hydrologic cycle. A high resolution precipitation data set has been derived from radar reflectivity observations taken from the National Weather Service WSR-88D radars in the continental U.S. To evaluate the product the authors provide several case studies of radar-rain gauge comparisons at locations throughout the Mississippi River Basin. They present bias, root mean square difference, fractional standard difference, and correlation coefficient statistics for radar-rain gauge comparisons for the hourly, daily, monthly, yearly and warm season temporal scale. These point (gauge) and pixel (radar) comparisons show large discrepancies at the hourly scale, on the order of 600–800%. An evaluation of the differences associated with temporally integrated estimates shows marked reduction in these discrepancies. At the long-term (warm season), these reduce to about 10%. An estimate of the difference in the comparison of the long term accumulation of gridded gauge based estimates and radar estimates at 0.25° × 0.25° shows values in the range of 20% but decrease to about 15% after applying filtering techniques in the basin-wide comparisons.

Citation: Nelson, B. R., W. F. Krajewski, J. A. Smith, E. Habib, and G. Hoogenboom (2005), Archival precipitation data set for the Mississippi River Basin: Evaluation, *Geophys. Res. Lett.*, *32*, L18403, doi:10.1029/2005GL023334.

1. Introduction

[2] As part of the Global Water and Energy Cycle (GEWEX) Continental-Scale International Project (GCIP), we have developed an archival radar-based precipitation data set for use in a wide range of hydroclimatological analyses. The overall goal of GCIP, which is based in the Mississippi River Basin (MRB), is to demonstrate skill in predicting changes in water resources on seasonal and annual time scales [*Coughlan and Avissar*, 1996]. To aid in achieving this goal, we have developed this radar-based precipitation data set at a high resolution, 4×4 km^2 spatial and hourly temporal, for a 5-year period (1996–2000). We presented the issues related to the data management, organization, and format by *Nelson et al.* [2003a], and we also presented the algorithm used to develop the precipitation data set by *Nelson et al.* [2003b]. The result of that study was the radar-based precipitation data set which we evaluate in this study using rain gauge measurements. The radar-only rainfall estimates are produced based on the algorithm detailed by *Nelson et al.* [2003b] which used high quality rain gauge networks for Z-R calibration, a reflectivity enhancement algorithm, and visual quality control. We present an evaluation of the warm season (April–October) only as precipitation estimates during the cold season are subject to many more sources of error than in the warm season.

[3] As argued by *Krajewski and Smith* [2002] in the strict sense, and with respect to radar hydrology, validation is the determination of the space-time statistical structure of errors of the radar-rainfall product. Although the determination of the joint probability distribution of the radar-rainfall error is an ongoing research agenda, several studies address the statistical structure of radar-rainfall errors by investigating more issues than just simple comparisons between radar and rain gauge estimates of rainfall [*Zawadzki*, 1975; *Kitchen and Blackall*, 1992; *Ciach and Krajewski*, 1999; *Ciach et al.*, 2003]. These studies suggest that validation of an areal estimate is difficult when only one or few in-situ measurements are available, especially at the spatio-temporal scale of our product. Given the vast area of the MRB, we do not present a true validation of the product because there are few if any high density long-term rain gauge networks that could serve to represent the entire basin for the study period. We present an evaluation of the precipitation product including rain gauge data from high quality networks and gridded estimates of precipitation that cover almost the entire basin.

[4] The paper is organized as follows: In section 2, we present the case studies we used for evaluation. In section 3, we compare the product at the basin-wide scale, and we present conclusions in section 4.

2. Evaluation Case Studies

2.1. Small-Scale Networks

[5] We selected five small-scale networks in the MRB based on their data availability, data quality, and study period. The networks are the following: The Iowa City Airport Piconet in Iowa for 1999 and 2000 [*Krajewski et al.*, 1998]; The Atmospheric Boundary Layer Experiments

[1]IIHR-Hydrosciences and Engineering, and Department of Civil and Environmental Engineering, University of Iowa, Iowa City, Iowa, USA.
[2]Now at the National Climatic Data Center, Asheville, North Carolina, USA.
[3]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[4]Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, Louisiana, USA.
[5]Department of Biological and Agricultural Engineering, University of Georgia, Griffin, Georgia, USA.

0094-8276/05/2005GL023334$05.00

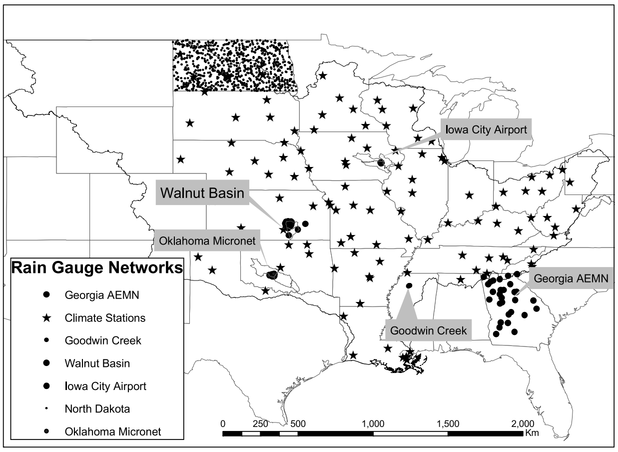

Figure 1. Rain gauge networks used in evaluation of radar rainfall precipitation product.

(ABLE) of the Walnut watershed in Kansas for 1998–2000 [*LeMone et al.*, 2000]; The U.S.D.A. Agricultural Research Service Oklahoma Micronet for 1998–1999; The Goodwin Creek experimental watershed in Mississippi for 1996–2000 [*Steiner et al.*, 1999]; The Georgia Automated Environmental Monitoring Network (AEMN) for 1996–2000 [*Hoogenboom et al.*, 2003].

[6] We compared radar rainfall pixel estimates (4×4 km^2) with rain gauge point estimates for the five small-scale networks (Figure 1). Comparisons are made for hourly, daily, monthly, yearly and total warm season analysis for these specialized networks. We performed quality control of the data or we obtained data that have been quality controlled. We then used four statistical measures for the quantitative analysis of the comparisons. They are the multiplicative bias between the radar-rainfall (R_r) and rain gauge (R_g) pairs, the root mean square difference between the two estimates, the fractional standard difference, and the sample correlation coefficient between the two estimates. We define them as follow: The multiplicative bias, $B = \sum R_r / \sum R_g$; The root mean square difference, $RMSD = \sqrt{\sum (R_r - R_g)^2}$; The fractional standard difference, $FSD = RMSD / \sum R_g$. The sample correlation coefficient, $r(R_r, R_g)$, is determined from the Pearson formula given in any common statistics text book.

[7] We integrated radar rainfall and rain gauge estimates from hourly time scale to daily, monthly, yearly warm season, and five-year warm season accumulations The hourly, daily, monthly, and yearly warm season results are presented in Table 1. Yearly accumulations are for the warm season only and thus are also the seasonal accumulations for the given year. Some of the main results of the small-scale network comparisons are as follows:

[8] 1. At the hourly scale, the fractional standard difference is large at 600–800%. The fractional standard difference decreases for the increasing temporal scale. At the daily scale, the range of the fractional standard difference is 100–200%. At the monthly scale, the range of the fractional standard difference is 30–40%. At the yearly scale, the range of the fractional standard difference is about 25%. At the total for all warm seasons, the fractional standard difference is on the order of 20%.

[9] 2. The correlation at the hourly scale ranges from 0.57–0.69; at the daily scale it ranges from 0.82–0.86; at the monthly scale, it ranges from 0.74–0.92; at the yearly scale, it ranges from 0.70–0.98. The warm season correlation is about 0.96 for all networks.

[10] 3. Based on the correlation estimates at the differing scales, there is a decrease in the correlation of yearly comparison versus the correlation of the monthly comparisons for certain networks due to the smaller sample size in the yearly radar, rain gauge pairs and the fact that we use the population coefficient which has the limitation that it is influenced by outliers and skewed distributions.

[11] 4. The multiplicative bias in the comparisons ranges from 1.01 to 1.2. The Walnut basin network has a bias of 1.2. We believe this is due to the quality of the rain gauge network as indicated in meta-data records and maintenance logs. The Goodwin Creek network has a bias of 1.15. The Georgia network has a bias of 1.09. The Iowa City network and the Oklahoma Micronet have a bias of 1.01.

2.2. North Dakota Network

[12] The North Dakota Atmospheric Resource Board Cooperative Observer Network (ARBCON) network provides for a relatively high density of gauges (~600) over a larger spatial scale. The network is located in the northern region of the MRB and this presents different problems as compared to other networks that we have analyzed. We used information in the input rain gauge data set to determine bad gauge data, and we found that during certain months there were many gauges not reporting precipitation. We found that the number of gauges not reporting for April (50%) and May (7%) was high as compared with the other months. Therefore we excluded the months of April and May from our analysis. We believe that in the northern part of the MRB, precipitation type such as freezing rain, graupel, snow, and ice is an issue which results in highly biased comparisons of gauge and radar estimates.

[13] Comparisons of the resulting gauges for June–September for monthly, yearly accumulations and warm season accumulations are presented in Table 2. The multiplicative bias for this network is similar to the other small-

Table 1. Range of Values for Bias, Root Mean Square Difference, Fractional Standard Difference, and Correlation for the Small-Scale Networks in MRB

	Bias	RMSD, mm	FSD, %	Correlation
Hourly	1.01–1.20	0.71–0.96	618.0–989.0	0.57–0.69
Daily	1.01–1.20	3.84–4.87	140.0–212.0	0.82–0.86
Monthly	1.01–1.20	22.53–38.7	30.0–46.0	0.73–0.92
Yearly	1.01–1.20	37.1–150.0	10.0–26.0	0.70–0.98

Table 2. Bias, Root Mean Square Difference, Fractional Standard Difference, and Correlation for Monthly, Yearly, and Total Warm Season Comparisons of North Dakota ARBCON Network

	Bias	RMSD, mm	FSD (%)	Correlation
Monthly	1.05	35.16	54.0	0.63
Yearly	1.05	60.62	25.0	0.65
Total w.s	1.05	128.07	14.0	0.77

Table 3. Bias, Root Mean Square Difference, Fractional Standard Difference, and Correlation for Monthly, Yearly, and Warm Season Comparisons of NCDC First Order Rain Gauges

	Bias	RMSD, mm	FSD, %	Correlation
Monthly	0.994	30.6	33.0	0.86
Yearly	0.994	97.2	15.0	0.84
Total w.s	1.02	250.7	8.0	0.94

scale networks. Interestingly, this suggests that over the long term, the simple plastic gauges perform almost as well as the complex tipping bucket gauges. There is considerable difference in the correlation of the gauge and radar comparisons as compared with the small-scale networks. For the North Dakota network, the correlation for the monthly scale is about 0.63 and for the small-scale networks, the correlation ranges from 0.74–0.92. The smaller correlation for the North Dakota network is due to the fact that as compared to the small-scale networks, the radar gauge comparisons are taken over the entire state and there are many more gauge radar pairs of estimates. Gauge radar comparisons evaluated over a larger geographic area (North Dakota) are likely to be more variable as compared to gauge radar comparisons for a specific gauge network that spans only a few kilometers.

2.3. NCDC First Order Stations

[14] At the large spatial scale, we compare the radar based precipitation product with specific rain gauges throughout the MRB. We used rain gauge data from first order stations from the Surface Land Daily Cooperative Summary of the Day Data set TD3200 [*National Climatic Data Center*, 2000]. None of these first order stations have been used for the radar-only product development. Table 3 presents the statistics for the radar rainfall and rain gauge estimates for the first order NCDC gauge locations in Figure 1. We found several locations in the MRB that were affected by problems that exist in the radar rainfall estimate or the rain gauge estimate, such as beam blockage, range dependency, or other gauge siting problems [see *Nelson et al.*, 2003]. We then removed these radar and gauge pairs from the analysis and the resulting 73 locations showed good agreement (Table 3). The fractional standard difference decreases from 33% at the monthly accumulation scale to 15% at the yearly accumulation scale to 8% at the five year warm season accumulation scale.

3. Basin-Wide Analysis

[15] At the basin scale, we compare the radar based precipitation product with a gridded rain gauge product (URD) [*Higgins et al.*, 2000] for the period 1996–1998 at daily temporal resolution and 0.25° × 0.25° resolution. The URD is estimated from point estimates using a Cressman scheme with modifications. We integrated the radar based precipitation product to the same spatial scale as the URD product and compared different temporal resolutions. This provides for comparisons of the MRB product over the entire basin.

[16] Figure 2 shows the accumulation of the warm season months for the years 1996–1998 for each pixel in the MRB. Here we present the joint probability of the MRB product and the URD product as there are some 5000 pixels for comparisons. The joint distribution in Figure 2 shows that most of the pixels in the basin compare well. However, there are locations where the radar based product severely underestimates precipitation as compared to the rain gauge product. We believe the underestimation in the radar product corresponds to areas in the basin that are affected by beam blockage. Even at the 0.25° × 0.25° scale, there are large spatial extents in the western part of the basin where the radars are partially or completely blocked by mountains. Figure 3a shows the spatial coverage of the MRB product. The underestimation of the product in the western mountains can be seen, but is not entirely evident. The URD product (Figure 3b) also shows drier regions in the west and this could be a climatological effect. Further, siting of gauges in the mountains is also a difficult task, and this could cause an underestimation of precipitation in the URD product. Figure 3c shows the absolute difference (%) between the integrated radar pixel

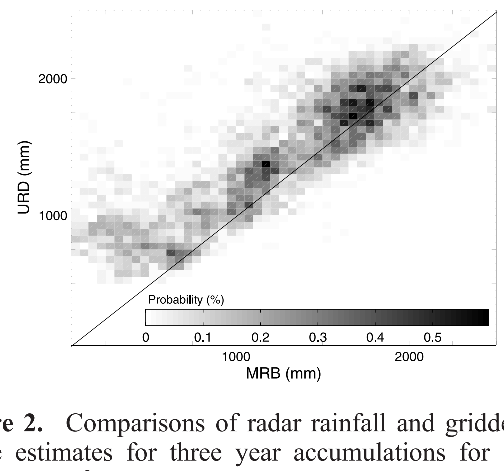

Figure 2. Comparisons of radar rainfall and gridded rain gauge estimates for three year accumulations for ~5000 0.25° × 0.25° pixels in the MRB. The joint probability of five year radar rainfall and rain gauge accumulations are presented.

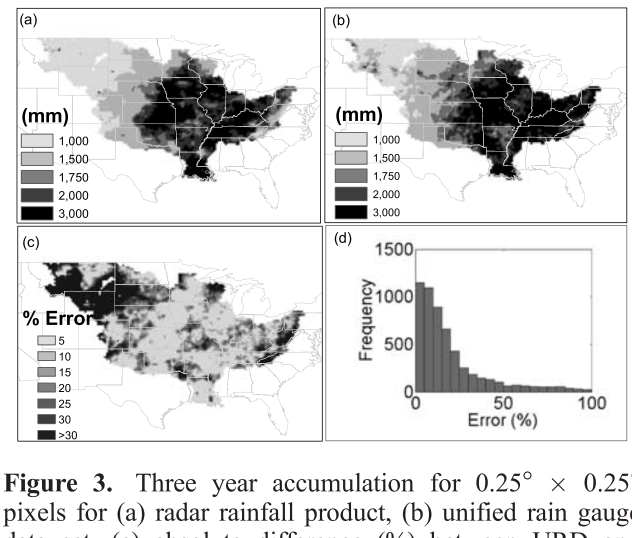

Figure 3. Three year accumulation for 0.25° × 0.25° pixels for (a) radar rainfall product, (b) unified rain gauge data set, (c) absolute difference (%) between URD and MRB products. (d) Histogram of absolute differences.

Table 4. Bias, Root Mean Square Difference, Fractional Standard Difference and Correlation for Three-Year Warm Season Comparisons for Masked Areas (a) of Discontinuous Probability of Detection Due to Beam Blockage and Close to the Radar Problems, (b) 200 km Radar Range, (c) 150 km Radar Range, and (d) 100 km Radar Range

	Bias	RMSD, mm	FSD, %	Correlation
Mask (a)	1.08	316.84	17.8	0.82
Mask (b)	1.08	382.49	22.8	0.82
Mask (c)	1.07	295.20	16.4	0.83
Mask (d)	1.06	289.57	16.0	0.84

and the interpolated rain gauge estimate with a percent difference that is less than about 20%. The histogram of these differences can be seen in Figure 3d with a mean of about 19%. There are, however, areas in the mountain regions where the difference is large.

[17] We then compared the integrated radar rainfall and interpolated rain gauge estimates based on filtering both sets of data temporally and spatially. Our assumption is that by filtering data due to assumed physical problems that could affect the estimate we can arrive at a more realistic comparison of precipitation due to rain gauge and radar. We filtered both sets of precipitation estimates spatially in two manners. We first filtered the estimates in areas we identified due to discontinuous probability of detection (POD). The POD is defined as the fraction of hours at each pixel reporting rainfall during the study period. The estimate of the POD over a long period can be used to identify areas that are blocked due to terrain or are affected by other factors close to the radar. Estimates of precipitation close to the radar are biased due to range dependent tilt selection, the cone of silence, and clutter suppression algorithms. We then filtered the estimates of precipitation based on the distance of the pixel estimates from the radar. We created radar coverage areas for all the radars in the MRB at 200, 150, and 100 km range. Several studies have shown that radar estimates, which are not adjusted by a vertical profile correction algorithm, and are located closer to the radar are quantitatively better than those that are far from the radar as compared with rain gauge estimates [e.g., *Vignal and Krajewski*, 2001].

[18] After masking out these areas, we compared the estimates of the URD precipitation estimate and the radar based precipitation estimate (Table 4). The correlation between the radar estimate and the rain gauge estimate increases slightly as the masks are applied from 0.82 to 0.84. The mask at the 200 km radar range does not filter many pixel locations. The POD mask provides an improvement in the comparisons as the decrease in the fractional standard difference is evident as compared to the 200 km radar range. The fractional standard difference improves still when the 150 km radar range filter is applied. The correlation in the radar and gauge estimates increases and the overall bias decreases with the POD mask and also with the 150 km radar range mask. We performed the same analysis after applying a temporal filter of the data. We extracted only the months of June, July, and August for the three year comparison of the radar and gauge estimates. There was no improvement in the statistics when applying the temporal filter except that the overall bias reduces from about 1.08 to 1.0.

4. Conclusions

[19] In this study, we evaluated a radar-based precipitation product that was previously developed as part of the GEWEX Continental-Scale International Project. This evaluation is intended to be a guide for users of the data set. Although the differences associated with the hourly radar rainfall estimates are large, these differences diminish with increasing temporal scales. Evaluation of the differences at many gauge locations throughout the basin shows values of about 10% for the five-year warm season scale. Further, the biases associated with the comparisons of the radar and rain gauge estimates are small. Case studies presented in this paper provide some guidelines for users of the radar based data set. Although quantitative comparisons of radar rainfall and rain gauge estimates should be taken with caution because of the representativeness error that is present in the comparison of a 4 $\times$ 4 km^2 pixel with a point estimate of rainfall, especially at the fine temporal scales. Filtering of this representativeness error was not possible in this study because we did not have any high density rain gauge networks for comparison of many radar rainfall pixels that also spanned the long period of record. As this is an important part of validation of radar rainfall products, the hydrologic and meteorological communities can be served well by implementing and maintaining high density quality rain gauge networks throughout the U.S.

[20] **Acknowledgments.** We gratefully acknowledge the support of NOAA Office of Global Programs through grant NA96GP0417. Portions of this work were performed while the first author held a National Research Council Research Associateship Award at the National Climatic Data Center, Asheville, NC.

References

Ciach, G. J., and W. F. Krajewski (1999), On the estimation of radar rainfall error variance, *Adv. Water Resour.*, *22*, 585–595.
Ciach, G. J., E. Habib, and W. F. Krajewski (2003), Zero-covariance hypothesis in the error variance separation method of radar rainfall verification, *Adv. Water Resour.*, *26*, 673–680.
Coughlan, M., and R. Avissar (1996), The Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP): An overview, *J. Geophys. Res.*, *101*, 7139–7147.
Higgins, R. W., W. Shi, and E. Yarosh (2000), Improved United States precipitation quality control system and analysis, *NCEP/Clim. Predict Cent. ATLAS 7*, 40 pp., U.S. Dep. of Comm., Washington, D. C.
Hoogenboom, G., D. D. Coker, J. M. Edenfield, D. M. Evans, and C. Fang (2003), The Georgia Automated Environmental Monitoring Network: 10 years of weather information for water resources management, in *Proceedings of the 2003 Georgia Water Resources Conference*, edited by K. J. Hatcher, pp. 896–900, Inst. of Ecol., Univ. of Georgia, Athens, Ga.
Kitchen, M., and R. M. Blackall (1992), Representativeness errors in comparisons between radar and gauge measurements of rainfall, *J. Hydrol.*, *134*, 13–33.
Krajewski, W. F., and J. A. Smith (2002), Radar hydrology: Rainfall estimation, *Adv. Water Resour.*, *25*, 1387–1394.
Krajewski, W. F., A. Kruger, and V. Nespor (1998), Experimental and numerical studies of small-scale rainfall measurements and variability, *Water Sci. Technol.*, *37*, 131–138.
LeMone, M. A., et al. (2000), Land-atmosphere interaction research, early results, and opportunities in the Walnut River Watershed in southeast Kansas: CASES and ABLE, *Bull. Am. Meteorol. Soc.*, *81*, 757–779.
National Climatic Data Center (2000), Surface land daily cooperative: Summary of the day, *Tech. Doc. TD-3200*, U.S. Dep. of Comm., Washington, D. C.

Nelson, B. R., W. F. Krajewski, A. Kruger, J. A. Smith, and M. L. Baeck (2003a), Archival precipitation data set for the Mississippi River Basin: Development of a GIS-based data browser, *Comput Geosci.*, *29*, 595–604.

Nelson, B. R., W. F. Krajewski, A. Kruger, J. A. Smith, and M. L. Baeck (2003b), Archival precipitation data set for the Mississippi River Basin: Algorithm development, *J. Geophys. Res.*, *108*(D22), 8857, doi:10.1029/2002JD003158.

Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden (1999), Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation, *Water Resour. Res.*, *35*, 2487–2503.

Vignal, B., and W. F. Krajewski (2001), Large-sample evaluation of two methods to correct range-dependent error for WSR-88D rainfall estimates, *J. Hydrometeorol.*, *2*, 490–504.

Zawadzki, I. I. (1975), On radar-raingage comparisons, *J. Appl. Meteorol.*, *14*, 1430–1436.

E. Habib, Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, USA.

G. Hoogenboom, Department of Biological and Agricultural Engineering, University of Georgia, Griffin, GA 30223, USA.

W. Krajewski, IIHR-Hydrosciences and Engineering, University of Iowa, Iowa City, IA 52242, USA. (witold-krajewski@uiowa.edu)

B. R. Nelson, National Climatic Data Center, Asheville, NC 28801, USA.

J. A. Smith, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA.