Hongjie Xie, Eric Small, Jan Hendrickx, Matt Richmond, and Xiaobing Zhou

GIS Based NEXRAD Precipitation (Stage III) Database

NEXRADs have been providing extremely useful information on precipitation, real time hazard weather monitoring and prediction. Long term precipitation data is a critical input to hydrologic and climate modeling, and drought monitoring. The difficulties to use NEXRAD products in conjunction with other geospatial products are that they are (1) in Hydrologic Rainfall Analysis Project; (2) extremely large and in multiple-compressed binary format. The goal of this study is to build a GIS based NEXRAD precipitation database with a commonly used coordinate system. This database will allow online access, querying, statistic analysis, and downloading. Techniques and approaches in this paper will cover automatic data conversion, projection, and reprojection. Spatial-temporal precipitation distribution in upper Rio Grande basin in the monsoon period of 1998 is presented as an application example.

INTRODUCTION

Radar has assisted weather predictions for over forty years but its operational use in hydrologic applications spans only a decade or so (Krajewski and Smith, 2002). The National Weather Service (NWS) began the installation across the U.S. of its NEXRAD in 1991. The resulting 160 radars, called WSR-88D (Weather Surveillance Radar-1988 Doppler), have revolutionized the NWS forecast and warning program through improved detection of severe wind, hail, and tornadoes and also for improved hydrologic forecast operations and services (Fulton, 2002). NEXRAD precipitation products (Stages I, II, III, and IV) have also been used to analyze the statistical characterization of extreme rainfall frequency, and validation of satellite remote sensing algorithms (Krajewski and Smith, 2002; Habib and Krajewski, 2002). In addition, long term precipitation data is a critical input to hydrologic modeling, weather and climate modeling, and drought monitoring.

The most useful NEXRAD product for most applications is the NEXRAD stage III data, in which, stage II estimates are combined from multiple radars into a single product covering an entire River Forecast Center (RFC) region (NWS/OH, 1997; Fulton et al., 1998) so that basin-wide stream flow study can be made. The stage III product is difficult to use in conjunction with other geospatial products. First, the data is in the Hydrologic Rainfall Analysis Project (HRAP), or secant polar stereographic projection (earth-centered datum) coordinate system. In contrast, many geospatial data are in geodetic datum (ellipsoidal earth) coordinate system, such as WGS 84 and NAD83, which is easily applied for multiple sources data integration and comprehensive analysis. Second, the dataset is extremely large and in multiple compressed binary format (XMRG). For example, one year of data for the area of West Gulf River Forecast Center (WGRFC) region is about 100 MB in compressed monthly files, and 15.33 GB in uncompressed ASCII files (hourly). Once the data is uncompressed and reprojected, retrieval of time series of rainfall for any region or cell of interest is possible.

Reed and Maidment (1995 and 1999) describe the HRAP coordinate system, and its transformations to other geodetic (ellipsoid datum) coordinate systems. The NWS website (NWS/OH, 1999) provides detailed information on how to display and transfer the HRAP/XMRG file format to ArcView or ArcInfo GIS format. These materials provided a starting point for our coordinate and data format transformations. In an effort to build a GIS database of stage III precipitation for the State of New Mexico, we found that it is extremely time consuming to transform the data format and coordinate system one file at a time since there are roughly 10,000 hourly files per year. In addition, for various validation studies and applications, it is necessary to retrieve a time series of rainfall for a region or cell of interest. Given these demands, we developed batch process approaches to (1) uncompress the multiple compressed binary files, and transfer the XMRG format to ASCII; (2) transfer ASCII to GIS grid format, define the polar stereographic projection, and reproject to chosen projection; and (3) clip and retrieve time series of rainfall information in any region or cell of interest, and create standard format. Figure 1 is the work flow chart includes all batch processes. The third process of time series precipitation clippling and retrieval will be covered in our other paper (in preparation).

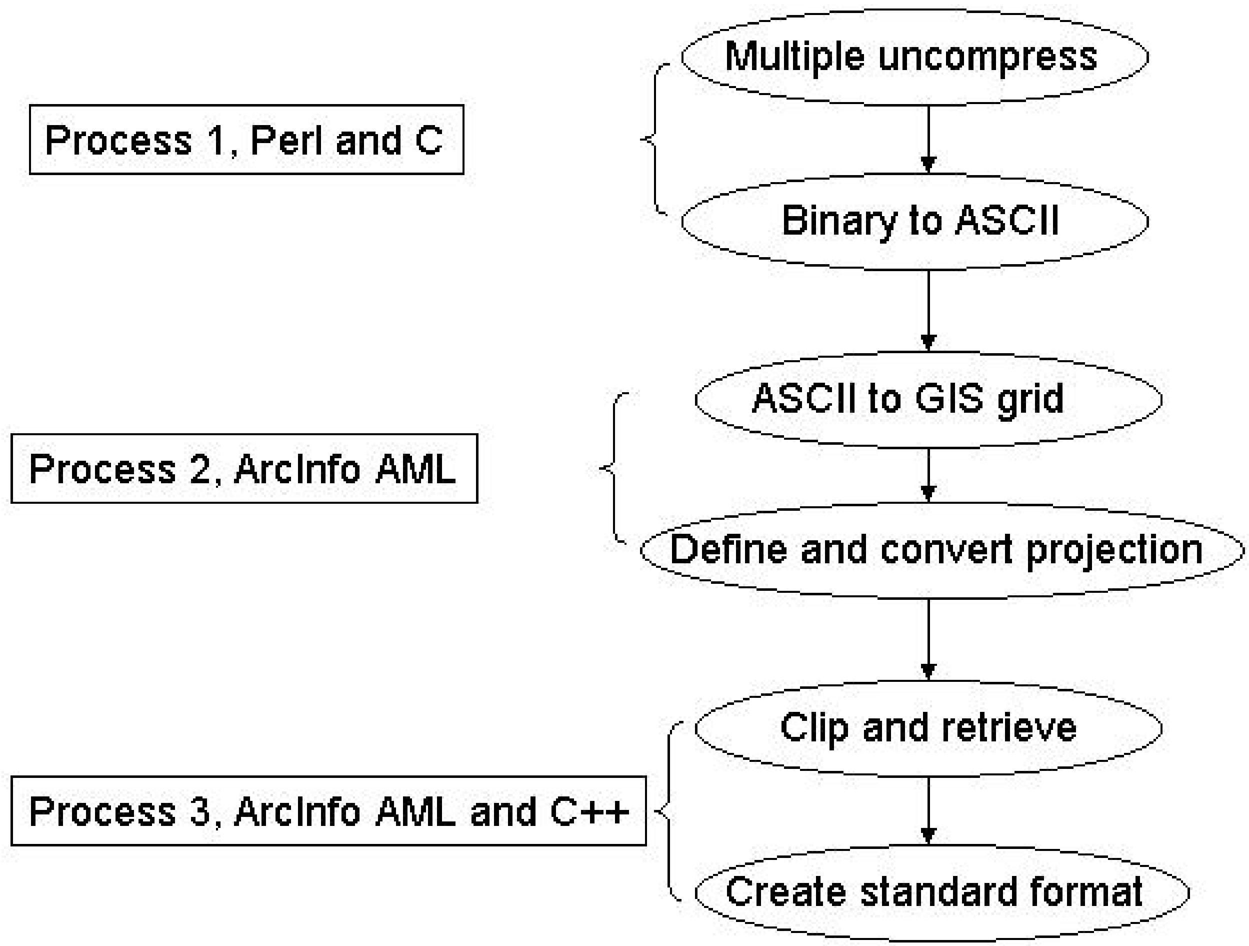


Figure 1. Work flow chart includes all batch processes

In addition to developing batch processes, we also describe improvement to the polar stereographic projection defined using ArcInfo from the NWS website (NWS/OH, 1999), In the NWS method, the ArcInfo's polar stereographic projection was used to define the HRAP's polar projection with radius of 6371200 m, and true latitude of 60 N, in contrast to ArcInfo's parameters of 6370997 m, 60.40884 N for radius and latitude, respectively. ArcInfo (or ArcInfo workstation) does not support this sort of customization. In this paper, we describe how to use ArcGIS's ArcToolBox to change the parameters from the default to match the HRAP's projection.

BASIC INFORMATION OF STAGE III DOWNLOADED FROM RFCs

Here we use stage III data downloaded from the West Gulf River Forecast Center (WGRFC), found at http://dipper.nws.noaa.gov/hdsb/data/nexrad/wgrfc_stageiii.html . This data is in a monthly tar format. In our database, we keep these monthly files under the particular years subdirectory, under stage3-original directory (see detail in Figure 4). Based on an initial exploration of the data, we found that there are three different data types depending on time period (table 1). Each requires slight differences in processing.

Table 1. Data parameters for different periods of time

	type A (Before Sept. 1997)	type B (Sept. 1997- Dec. 1998)	type C (After 1998)
xllcorner- polar ster. (HRAP)	-671512.5 (260)	-671512.5 (260)	-528637.5 (290)
yllcorner- polar ster. (HRAP)	-7620000.0 (1)	-7620000.0 (1)	-7577137.50 (10)
cell size (m)	4762.5	4762.5	4762.5
nclos	455	455	425
nrows	500	500	390
Notes: xllcorner- x coordinate at low left corner; yllcorner- y coordinate at low left corner; polar ster.- polar stereographic projection; HRAP- Hydrologic Rainfall Analysis Projection; nclos- number of columns; nrows- number of rows.			

BATCH PROCESS TO GIS GRID FORMAT

In this section, we describe our method to automatically transfer the original multiple- compressed binary files into GIS grid files, and define and convert the projection (processes 1 and 2 in Figure 1). To extract hourly Stage III products (in XMRG format), we first untar the monthly file to get the daily files, untar the daily files to get the compressed hourly files, and then uncompress the hourly files (XMRG format). The hourly XMRG file are then read and transferred to an ASCII file, which is transferred to an ArcGIS grid file. From this grid file, we can define its projection (polar stereographic) and then reproject the data to any selected projection. Here, we use UTM zone 13 and WGS 84 datum system as the final projection.

Perl and C scripts Run in UNIX System

We now describe scripts designed to uncompress monthly tar files to hourly XMRG files, and then transfer them to ASCII files. By understanding the 3 different types of data format for different time periods, Perl scripts have been written slightly different, while C scripts are exactly the same except one additional C script for reducing dimensions of types A and B (455, 500) to as the same as type C (425, 390).

In total, 6 scripts are required: munch-A.pl, munch-B.pl, munch-C.pl, reduce.c, aggregate.c, and xmrgtoasc.c. The first 5 scripts are developed here, while the last script can be downloaded from NWS website (NWS/OH, 1999). For our applications, we did modify the xmrgtoasc.c script to convert the units of precipitation into mm. The original unit in stage III is in hundredths of mm.

The following is the pseudo code for script munch-A.pl:

```
1 looking for files of Siii[mm][yy]WG.tar
   2 for each month
          untar each file to Siii[mm][dd][yy]WG.tar
      3 for each day
           untar each Siii[mm][dd][yy]WG.tar to xmrg[mm][dd][yy][hh]z.Z
         4 for each hour xmrg[mm][dd][yy][hh]z.Z
              4.1 uncompress it to xmrg[mm][dd][yy][hh]z
              4.2 call the xmrgtoasc to transfer xmrg[mm][dd][yy][hh]z into xmrg[mm][dd][yy][hh]z.asc. (selecting ster as the third argument: the polar
stereographic projection)
              4.3 call the reduce to cut the size from 455x500 to 425x390 as the ones type C
          end 4
            delete each hour *.Z file
            call aggregate (a c program) to add each hour *.asc to one day .asc
            delete the daily tar file
      end 3
   end 2
end 1
```

After this step, each monthly tar file is uncompressed to hourly files (*.asc), and hourly files in one day are aggregated into daily files. In total, there are roughly 750 (30 x 24 +30) files for each month and 9125 (365 x 24 +365) files for each year. The size of each file is about 1.68 MB. So for one month will about 1.23 GB, and for one year will be about 15.33 GB. For the WGRFC, there are data over 8 years from December 1994 to present, so the total size for those hourly files is over 122.64 GB (15.33 x 8).

AML Script Runs on ArcInfo Workstation

The purpose of this script (asc2grd.aml) is to batch transfer the ASCII files to ArcInfo grid files, define the polar stereographic projection for the grid files, reproject them to geographic (sphere) coordinate, and then to UTM 13, WGS 84 (ellipsoidal earth datum) coordinate, which is the standard coordinate system for our final database. In a GIS system, we need to define the coordinate system for the grid files, even though the polar stereographic coordinate has been assigned for ASCII files in the previous scripts. ArcInfo (or ArcInfo workstation) does not support customizing its polar stereographic projection to match the HRAP polar stereographic projection, but ArcGIS's ArcToolBox does support this sort of customization. So we first define the HRAP projection using the ArcToolBox, and then use the ArcInfo AML to complete the batch process.

Step 1. Select any one of the ASCII files, and then transfer the file to grid using ASCII to Grid tool in ArcToolBox, with float as grid type. Then define the projection of the grid file using Define Projection (coverage, grids, TINs) tool in ArcToolBox, and select polar as projection. Figures 2 and 3 illustrate the procedures for customizing the HRAP polar stereographic projection. Polar stereographic projection is a spherical, earth-centered datum of radius 6371.2 km, with a standard (true) latitude of 60° North and a standard longitude (longitude of the projection center) of 105° West.

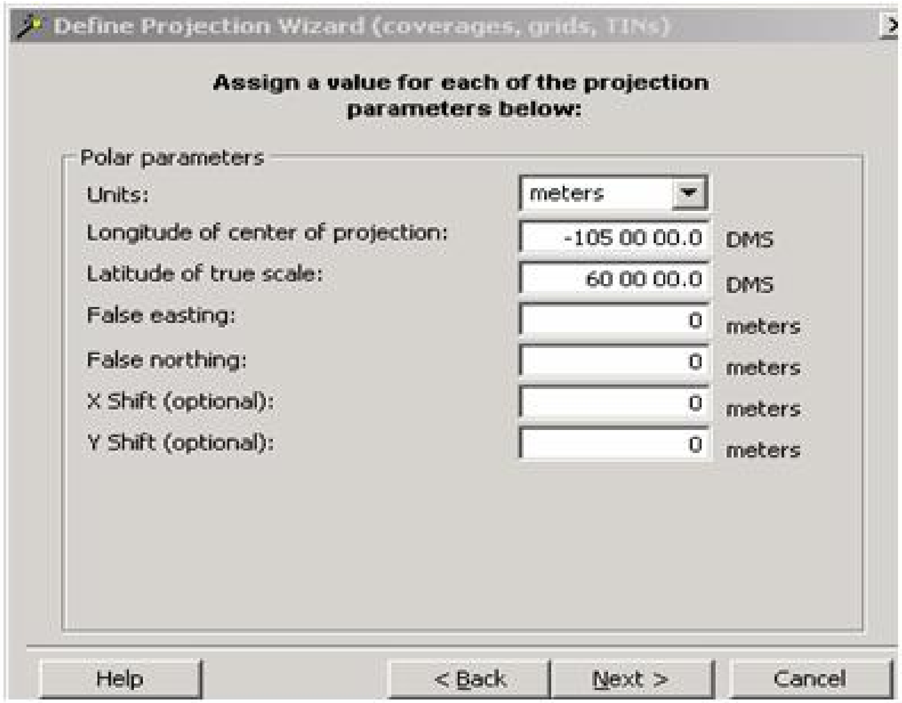


Figure 2. Define the polar stereographic projection in ArcToolBox

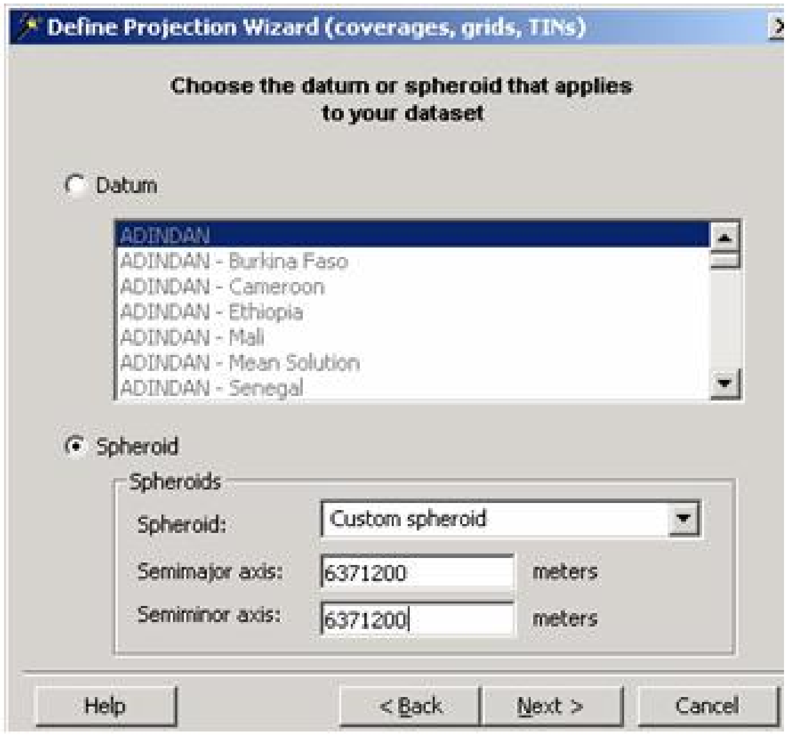


Figure 3. Customize the earth-centered datum of radius as 6371200 m

Step 2. Use the AML script (developed in this paper running on ArcInfo or ArcInfo workstation). This script first changes the workspace to where the data are located, and then creates a list of files to hold all .asc files. For each file within this list, the following steps are completed:
a. get the name of the file, create 4 new names to hold three files with different projections, and one as temporary file name;
b. transfer ASCII to grid, float type;
c. define the projection of the grid file by copy projection from step 1;
d. convert the polar stereographic projection to geographic (sphere) projection;
e. convert the geographic projection to UTM, WGS84 system, resampling method: nearest, grid size: 4000 m;
f. delete the temp files.
g. delete the filelist file.

After these steps, all hourly precipitation data are in ArcGIS grid files. Each hourly record has three different files, each with a different projection: polar stereographic projection, geographic projection, and UTM 13, and WGS 84 coordinate system. Here we achieve the exact same result by using two-step projection conversions instead of the three steps described by Reed and Maidment (1995 and 1999). We use the UTM coordinate system for our database, but we retained the data in other coordinate systems for special requests. The unit of precipitation is mm/hour. The size of each file is reduced by more than half from 1.68 MB (.asc file) to .8 MB (GIS grid file). Grid files can be input to any GIS system for analysis and/or further processing. The total file size for one month is about 600 MB, for one year is about 7.3 GB, and for 8 years is about 58.4 GB.

Strategies for Running the Scripts and Organizing the Data

Given the datasets are relatively large, we describe the strategies we employed to organize and store the data. In order to keep the original data and final rainfall data in store in our database, two different directories stage3-original and stage3-rainfall have been created (Figure 4). A subdirectory of each year is created under both those directories, such as year1994, year1995, and year2002. Then, for the stage3-rainfall data, each month of data is stored in a subdirectory under each year, for example, Jan, Feb, and Dec. This speeds up the data access from ArcGIS. Two other subdirectories, called polar-geo and sevi-rainfall, with tree structure of year and months are also created under the stage3-rainfall. The polar-geo is used to hold the hourly polar and geographic grid files, and sevi-rainfall holds the clipped subgrid files as an example.

All Perl and C scripts described above are located within the stage3-original directory. In the UNIX environment, we run the Perl script under the year subdirectory such as year1995 and year2000, by typing '../munch-A.pl' or '../munch-C.pl'. This script needs about 17 processing hours for each year of data of each year.

Then we copy each month of data (hourly ASCII files) into the related month subdirectory under the stage3-rainfall directory / polar-geo subdirectory. The AML script is executed in the stage3-rainfall directory, run on the ArcInfo workstation environment. This script needs about 35 processing hours for each year of data.

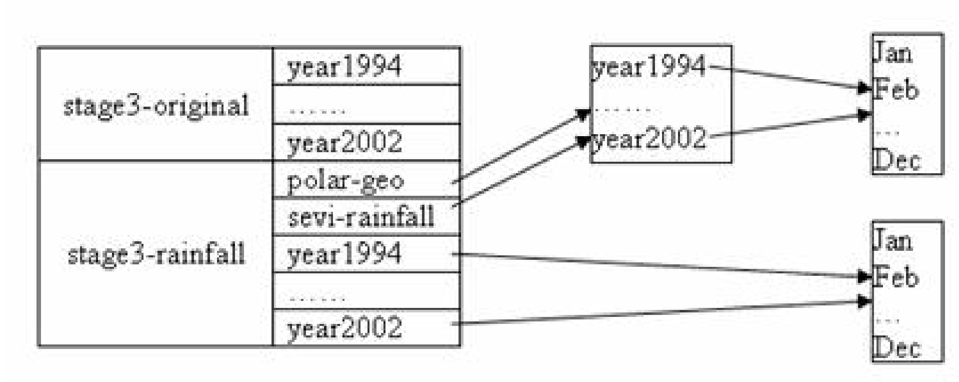


Figure 4. Data structure in the database

TIME SERIES OF RAINFALL CLIPPING AND RETRIEVING

For validation of the NEXRAD product and various applications, we need to automatically clip any region or cell of interest, extract time series of rainfall data for selected area, and create standard formats for comparisons with rainfall from rain gauge (process 3 in Figure 1). Some scripts (AML and C++) are developed to serve this purpose. In the future, some of these scripts will be incorporated into our ArcIMS interface, which will be accessible and executable through remote users to clip regions or cells of interest.

AN APPLICATION EXAMPLE IN THE UPPER RIO GRANDE BASIN

Once we have all data in GIS grid format (UTM zone 13, WGS 84) in our database, we can easily clip any region or cell of interest to get time series of precipitation for validation and application purposes. We can easily integrate different spatial information for comprehensive analyses. Here we use upper Rio Grande basin as an example to show how the rainfall datasets easily be used for various purposes. Figure 5 is an example showing two polygon coverages (upper Rio Grande basin and New Mexico state) overalying a daily NEXRAD precipitation map (grid format) of July 4, 1998. The upper Rio Grande basin coverage has been used to automatically clip precipitation from the grid, and then sum all daily or hourly precipitations into monthly, or even monthly to seasonally or yearly. The total area of the upper Rio Grande basin is about 101,272.8 km^2. As we can see that some of the area are not fully covered by the NEXRAD umbrella (no data available, white color), such as the north part (left) of the Rio Grande basin inside Colorado close to New Mexico.

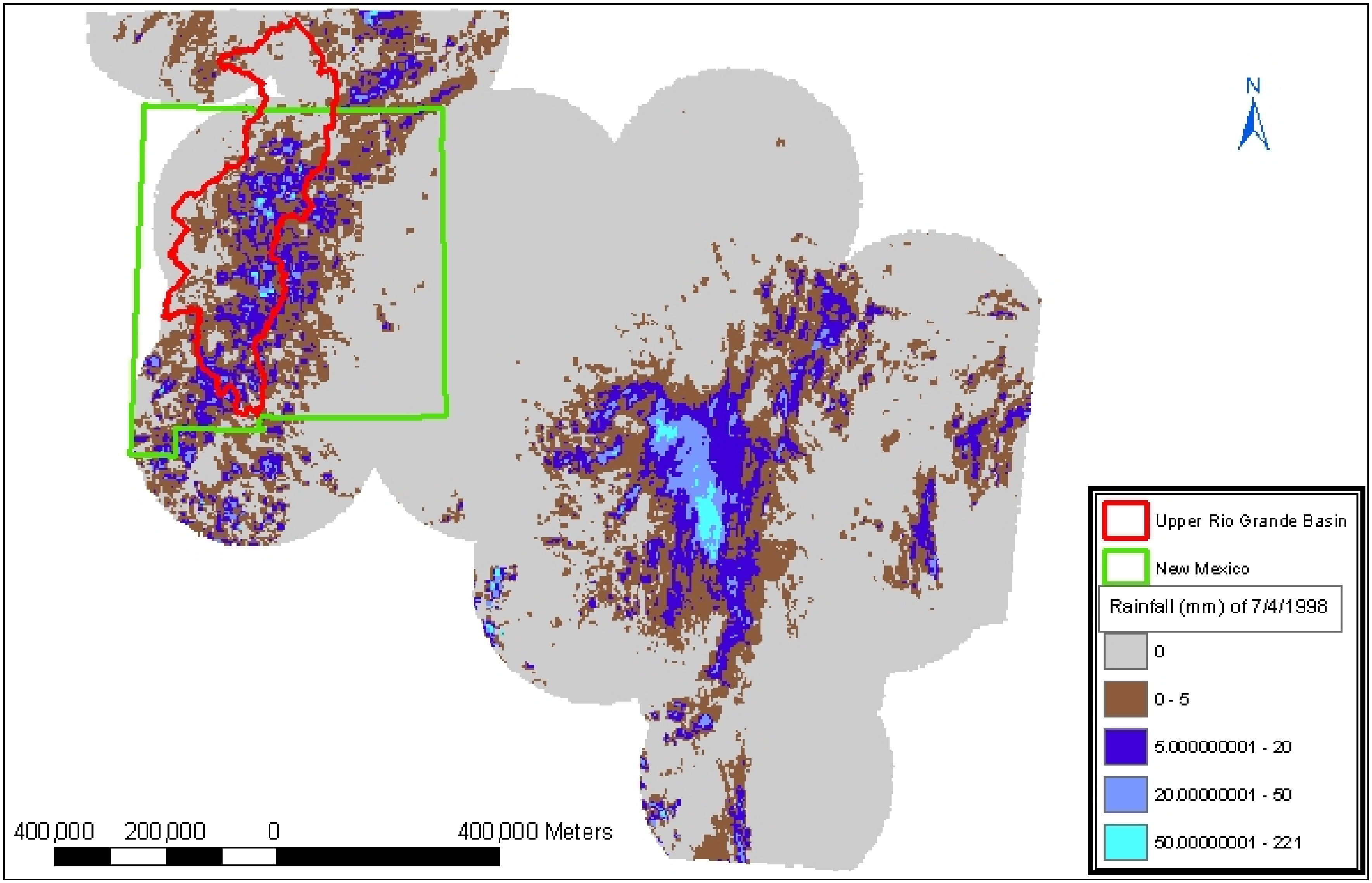
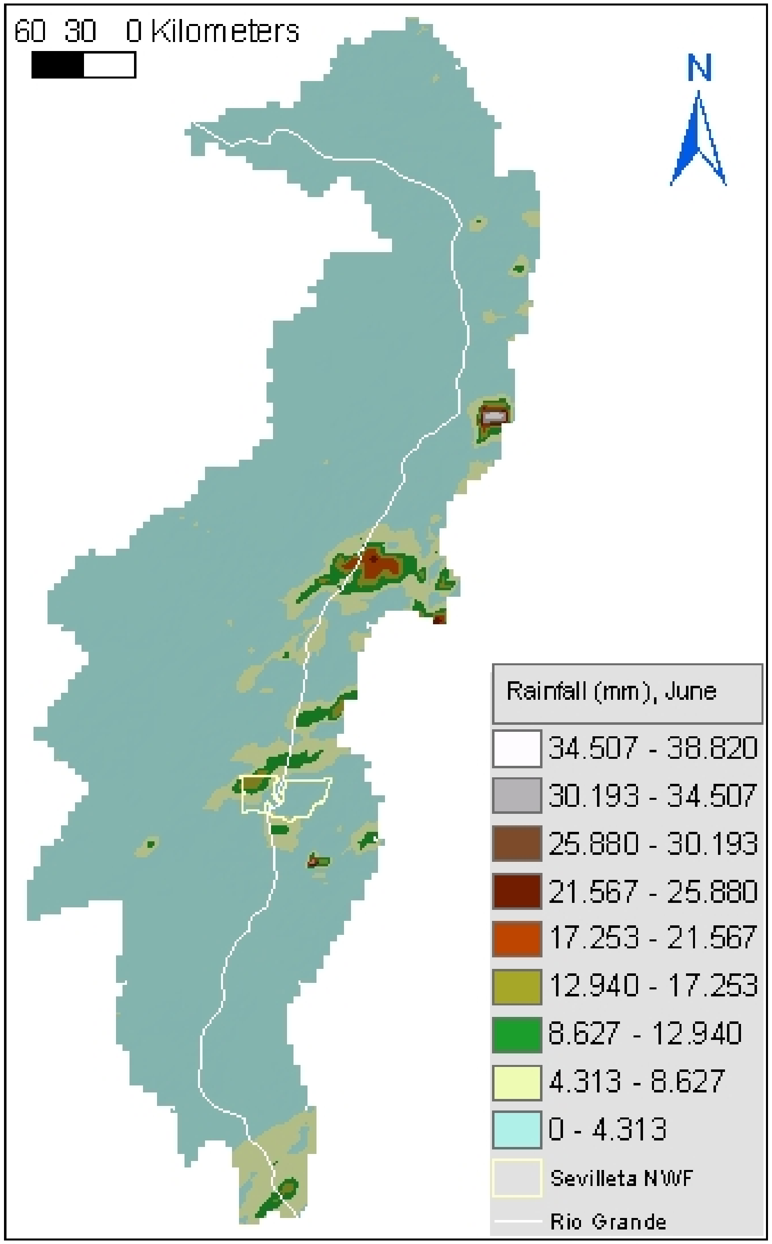
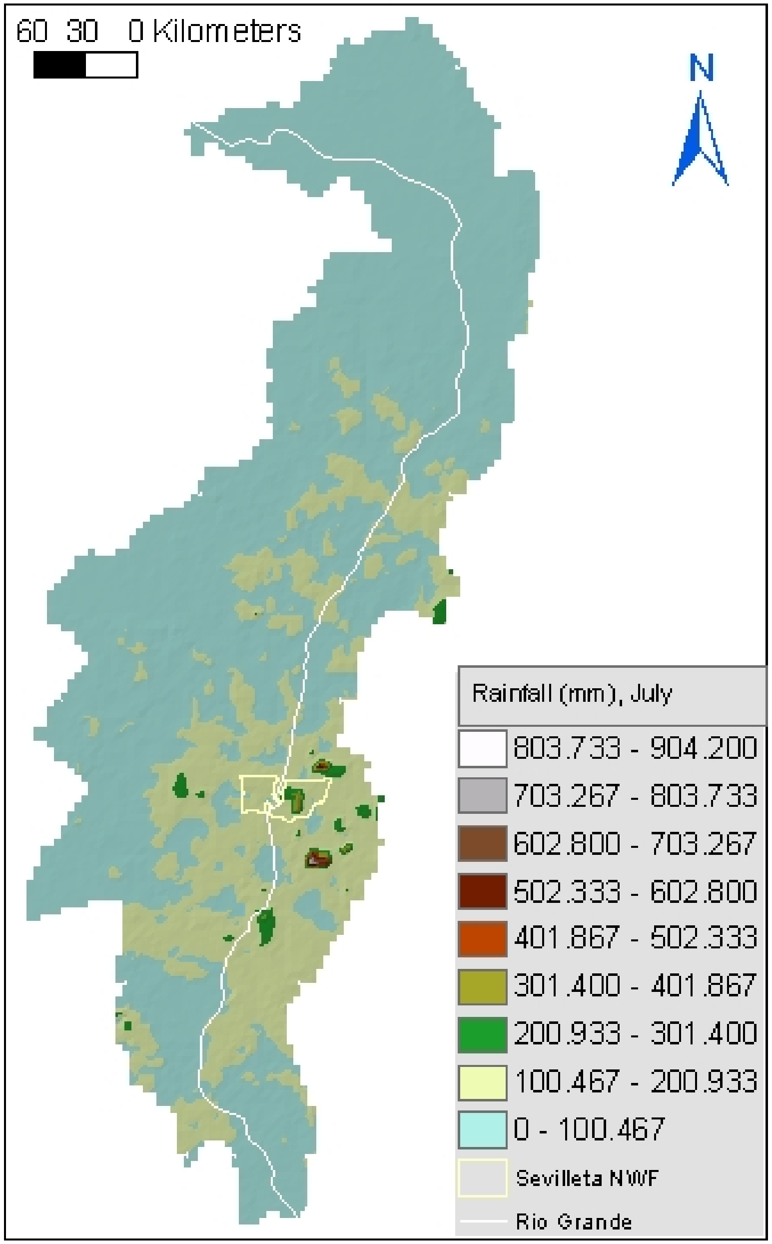


Figure 5. An example of hourly NEXRAD precipitation in the West Gulf River Forecast Center area

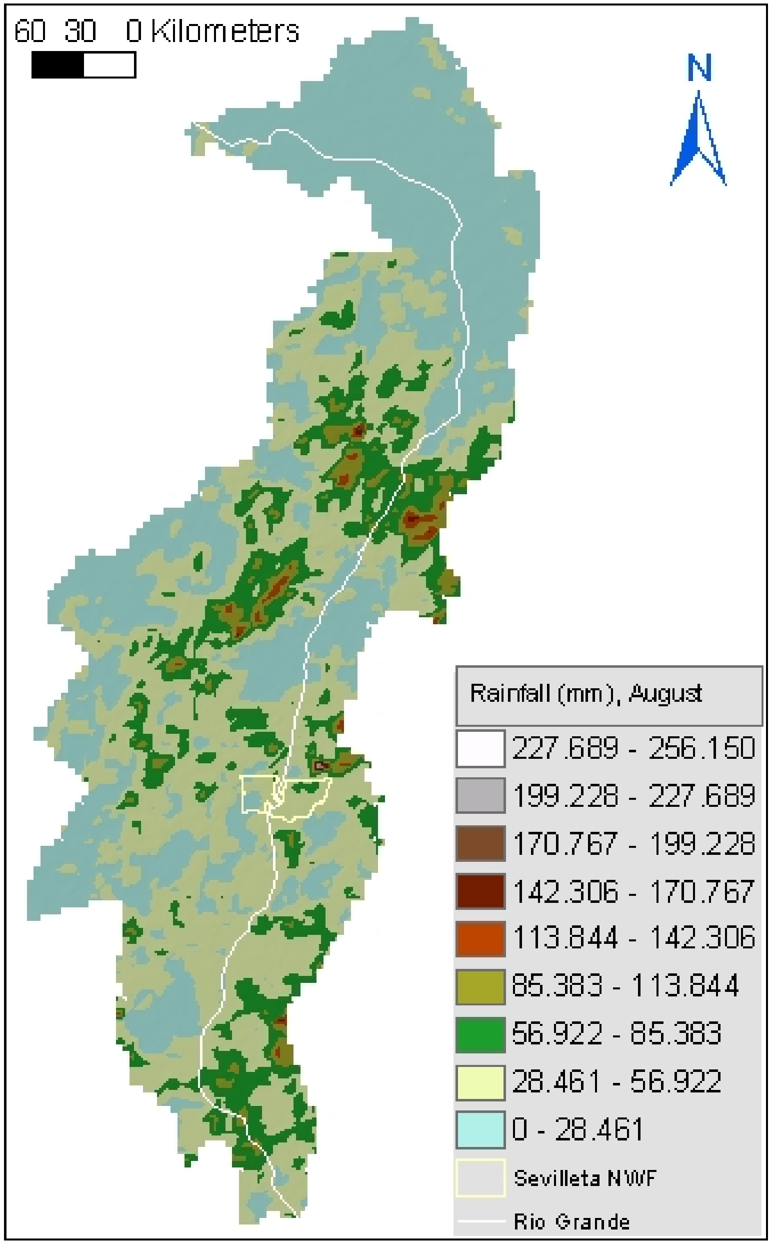
Figures 6-11 and table 2 show monthly precipitation distribution and statistic characteristics in the monsoon period of 1998 in the upper Rio Grande basin. Some results can be observed and summarized as follows:

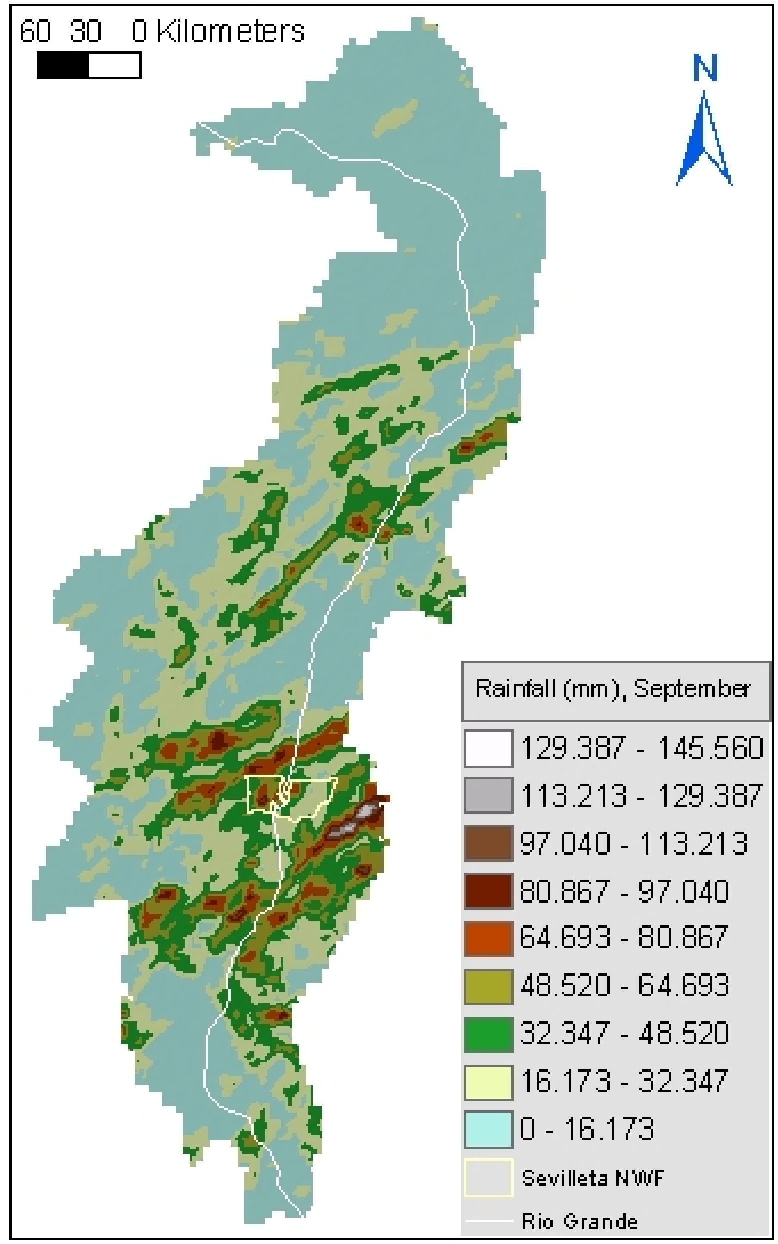
1. Total monsoon precipitation in the basin was 778,288.8 mm, or volume of 7.784 x 10^14 m^3. In fact, the NEXRAD fully covered area of 100,176 km^2 (Figure 11) is a little less than actual basin area of 101,272.8 km^2. So the actual amount of NEXRAD precipitation for the basin should be a little larger than the above one.
2. Precipitation in July was the biggest one (about 55%) in the monsoon period of 1998. Following July, August was about 28%, September 16%, while the June only 1% of the total monsoon precipitation.
3. From the rainfall maps (Figures 6-10), we can easily find the spatial distribution of precipitation changing through months. For example, in June, the heaviest precipitation was received in the south-east part of Taos county, second was the area between Santa Fe and Albuquerque, while northwest corner of Sevilleta received more precipitation than the east part of Sevilleta; In July, five point and its south area within the east part of the Sevilleta received 300-400 mm of precipitation, while other areas within the Sevilleta only obtained less than 200 mm, while the heaviest precipitation (800-900 mm) was in an area located in the southeast of Sevilleta, east of Socorro City and Rio Grande river, and north of U.S. highway 380; In August and September, they have totally different patterns of precipitation distribution.
4. From the comparison of radar and gauge precipitation in table 2, we find the estimated mean precipitation (124.3 mm) from NEXRAD for the upper Rio Grande basin matches very well to both the mean precipitation from 7 Sevilleta weather stations (136.2 mm) and mean precipitation (124.0 mm) of 80 years in Socorro. This means that NEXRAD precipitation in overall has very high accuracy compared to the rain gauge precipitation. But we find big difference when we try to compare precipitation (256.5 mm) from NEXRAD cells and precipitation (136.2 mm) from gauges within the related cells. This difference is understandable mostly because of the area-point error, and gauge representativeness error. A good deal of research effort on radar-gauge difference has been performed by Zawadzki (1975), Krajewski (1987), Kitchen and Blackall (1992), Ciach and Krajewski (1999), Anagnostou, et al. (1999), and many more.





Figures 6 and 7. TIN images of NEXRAD precipitation in June and July, 1998





Figures 8 and 9. TIN images of NEXRAD precipitation in August and September, 1998

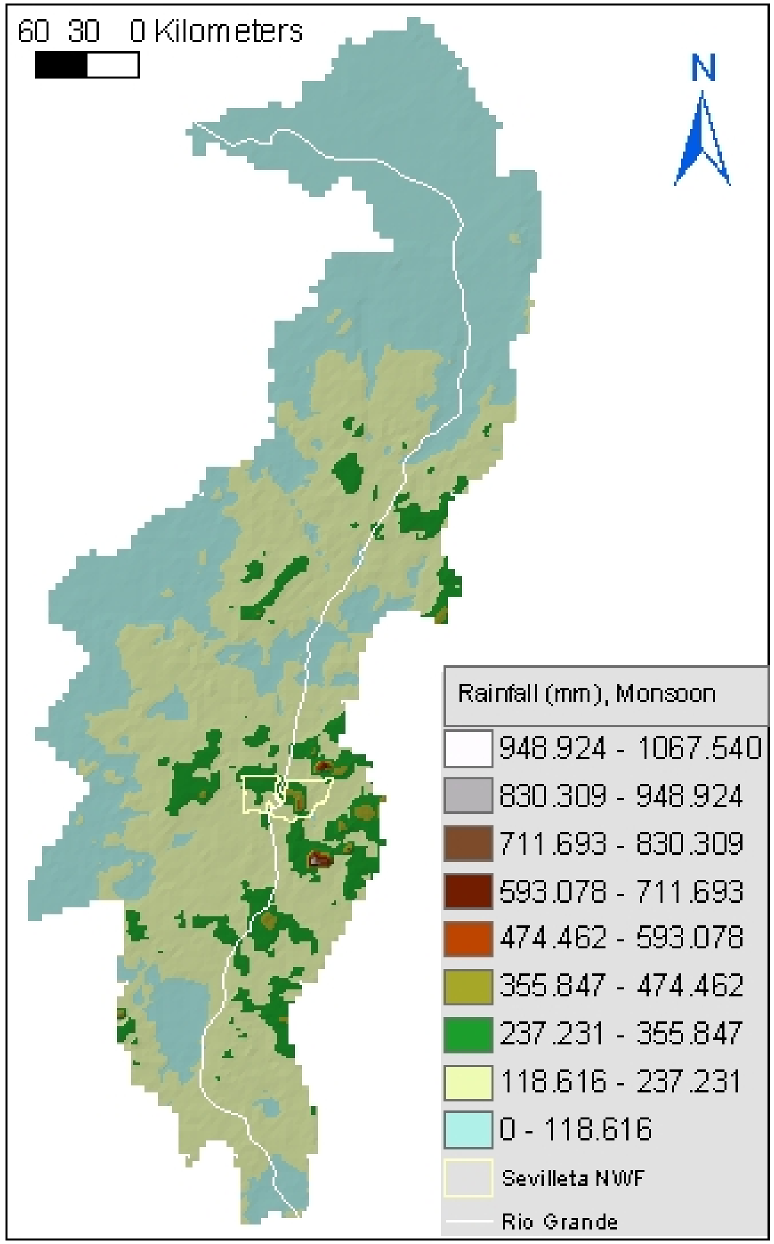


Figure 10. TIN images of NEXRAD precipitation, summer monsoon period of 1998 (June - September)

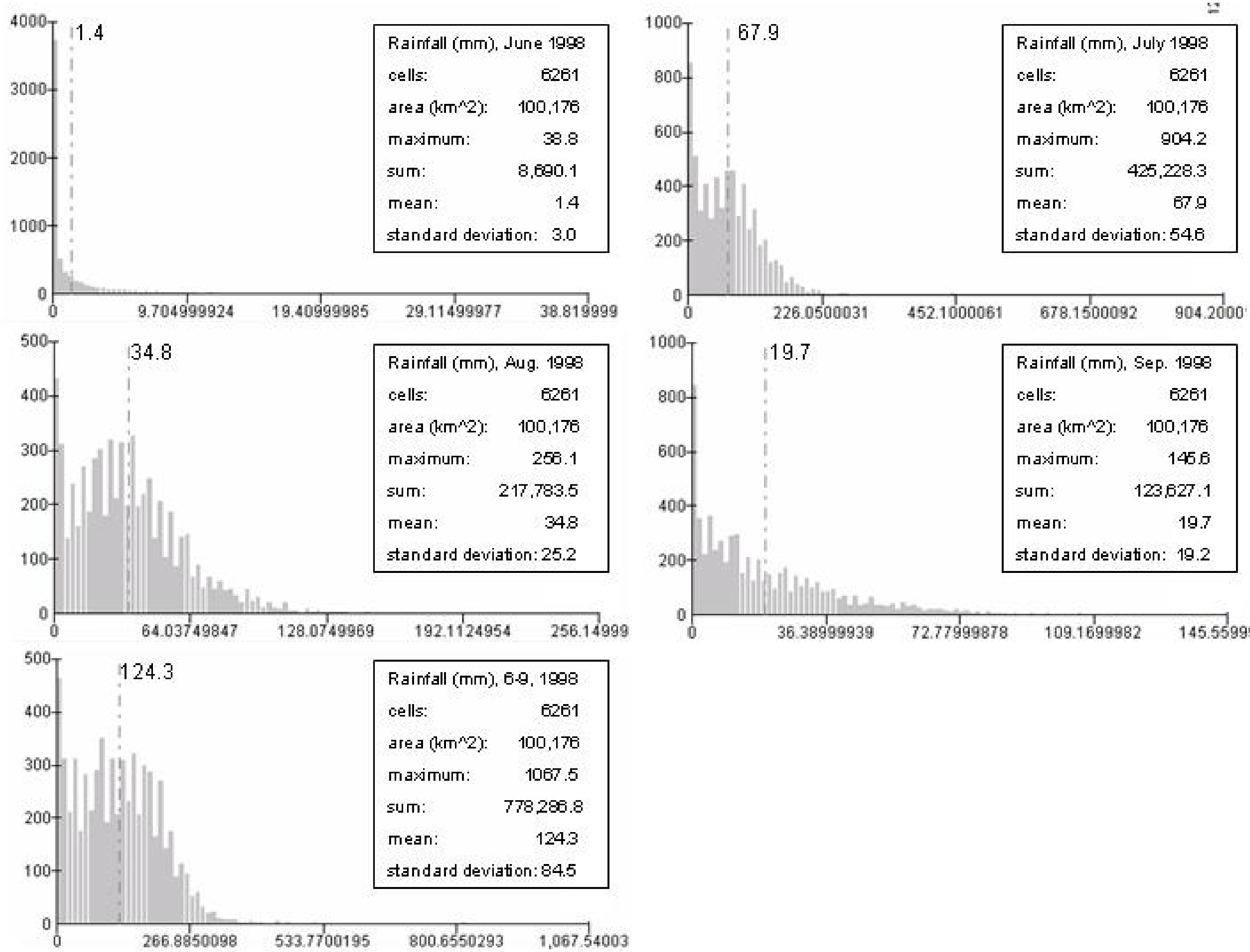


Figure 11. Statistic characteristics of NEXRAD precipitation in different months of 1998 monsoon period in upper Rio Grande basin

Table 2. Comparison of NEXRADA and rain gauge precipitation (mm)

	Monsoon period of 1998	Monsoon period of 1914-1993
NEXRAD for the upper Rio Grande basin	124.3	
Sevilleta weather stations (7)	136.2	

NEXRAD cells (7) related to Sevilleta stations (7)	256.5	
Socorro weather station (20 miles south of Sevilleta)		124.0 (Dahm and Moore, 1994)

CONCLUSION

In order to apply the NEXRAD products, it is important to have the data well prepared. The basic idea is to build a NEXRAD precipitation database in a GIS grid format, which allows fast display, requires less storage capacity, provides the best way to use ArcGIS's Geostatistical Analyst, Spatial Analyst, and 3D Analyst, and enables easier integration with other geo-spatial information. We developed a number of batch process approaches to automatically uncompress and transfer original stage III data to GIS grid format, define projection, convert projection, clip time series of region or cell of interest, and retrieve time series of rainfall in any region or cell of interest for purposes of validation and various applications. From an example of applying the NEXRAD precipitation for upper Rio Grande basin, we could easily find the spatial-temporal distribution of precipitation in the basin, which could provide good information for agriculture, forestry, water districts, and so on. This information is also a best input for various models and analyses such as rainfall-runoff model and water balance analysis. The NEXRAD estimation for the monsoon period of 1998 matches very well with the rain gauge observations, showing that radar provides very good accuracy in overall. In the same time, radar provides very high spatial resolution (4 x 4 km^2), which is very difficult, if it is not impossible, for sparse rain gauge networks to provide such a high spatial resolution of rainfall distribution. But the estimation difference between area (a radar cell) and point (a gauge within the related radar cell) is obvious, which has been addressed by a great deal of studies. One significant improvement, as shown in this paper, is to use ArcGIS's ArcToolBox to customize the parameters for the HRAP's polar stereographic projection definition, rather than using the ArcInfo (or ArcInfo workstation) that does not support the customization.

ACKNOWLEDGMENTS

This work was supported by the NSF EPSCoR grant through INRAM (Institute of Natural Resources Analysis and Management) in New Mexico. We acknowledge and appriciate Sevilleta Long-Term Ecological Research Program at University of New Mexico for sharing GIS layers and weather data.

REFERENCES

Anagnostou, E.N, Krajewski, W.F., and Smith, J., 1999, Uncertainty quantification of mean_areal radar-rainfall estimates, Journal of atmospheric and oceanic technology, V.16: 206-215.

Ciach, G.J. and Krajewski, W.F., 1999, Radar-rain gauge comparisons under observational uncertainties, Journal of Applied Meteorology, V.38: 1519-1525.

Dahm, C. N., Moore, D. I., 1994. The El Niño/Southern Oscillation Phenomenon & The Sevilleta Long- term Ecological Research Site Pages 12-20 in LTER Report. LTER Climate Committee, Edited by David Greenland. LTER Publication No. 18

Fulton, R. A., 2002, Activities to improve WSR-8D radar rainfall estimation in the National Weather Service, 2nd Federal Interagency Hydrologic Modeling Conference

Las Vegas, Nevada, July 28-August 1, 2002

Fulton R. A., Breidenbach, J. P., Seo, D., Miller, D. A., 1998, and O'Bannon, T., The WSR-D rainfall algorithm, Weather and forecasting, V. 13, June, pp377-395

Habib, E. and Krajewski, W. F., 2002, Uncertainty analysis of the TRMM ground-validation radar-rainfall products: application to the TEFLUN-B field campaign, Journal of applied meteorology, V. 41, pp. 558-572

Kitchen, M., and Blackall, R.M., 1992, representativeness errors in comparisons between radar and gauge measurements of rainfall. Journal of Hydrology, V134, pp13-33.

Krajeski, W. F. and Smith, J. A., 2002, Radar hydrology: rainfall estimation, Advances in Water Resources 25: 1387-1394

Krajewski, W.F., 1987, Co-kriging of radar-rainfall and rain gauge data, J. Geophysical Research, V.92: 9571-9580.

Molles, M.C., Jr., C.N. Dahm, and M.T. Crocker. 1992. Climatic variability and streams and rivers in semi-arid regions. p. 197-202. In R.D. Robarts and M.L. Bothwell (eds.), Aquatic ecosystems in semi-arid regions: implications for resource management. Environment Canada, Saskatoon.

Mitchell, V.L. 1976. The regionalization of climate in the western United States. Journal of Applied Meteorology 15:920-927.

Neilson, R.P. 1986. High resolution climatic analysis and southwest biogeography. Science 232:27-34.

NWS Office of Hydrology, 1997, Stage III precipitation processing syste, system guide, Hydrologic Research Laboratory.

NWS Office of Hydrology, 1999, Displaying and using NWS XMRG/HRAP files within ArcView or ArcInfo GIS, available at http://www.nws.noaa.gov/oh/hrl/distmodel/hrap.htm.

Redmond, K., 1998, El Nino, La Nina, and the Western U.S., Alaska and Hawaii, available at: http://www.wrcc.dri.edu/enso/ensofaq.html

Reed, S.M. and D.R. Maidment, 1999, Coordinate Transformations for Using NEXRAD Data in GIS-based Hydrologic Modeling, Journal of Hydrologic Engineering, 4(2): 174-182

Reed, S.M. and D.R. Maidment, 1995, A GIS procedure for merging NEXRAD precipitation data and digital elevation models to determine rainfall-runoff modeling parameters, Center for Research in Water Resources (CRWR) online report 95-3, 119pp. University of Texas at Austin, available at http://www.ce.utexas.edu/prof/maidment/GISHydro/docs/reports/seann/rep95_3.htm.

Zawadzki, I., 1975, On radar-raingauge comparison, Journal of Applied Meteorology, V.14: 1430-1436.

Hongjie Xie

Postdoctoral Research Associate, Earth and Environmental Sciences, New Mexico Tech, Socorro, NM 87801, 505-835-6448, hjxie@nmt.edu

Eric Small
Assistant Professor, Geological Sciences, University of Colorado - Boulder, CO 80309, 303-735-5033, eric.small@colorado.edu

Jan M.H. Hendrickx
Professor, Earth and Environmental Sciences, New Mexico Tech, Socorro, NM 87801, 505-835-5892, hendrick@nmt.edu

Matt Richmond
Computer Engineer, Earth and Environmental Sciences, New Mexico Tech, Socorro, NM 87801, segfault@ees.nmt.edu

Xiaobing Zhou
Research Assistant Professor, Earth and Environmental Sciences, New Mexico Tech, Socorro, NM 87801, 505-835-5068, xzhou@nmt.edu