ELSEVIER

Computers & Geosciences 31 (2005) 65–76

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

GIS-based NEXRAD Stage III precipitation database: automated approaches for data processing and visualization ☆

Hongjie Xie[a,*], Xiaobing Zhou[b], Enrique R. Vivoni[b], Jan M.H. Hendrickx[b], Eric E. Small[c]

[a]*Department of Earth and Environmental Science, University of Texas at San Antonio, 6900 N. Loop 1604 W., San Antonio, TX 78249, USA*

[b]*Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM 87801, USA*

[c]*Department of Geological Sciences, University of Colorado, Boulder, CO 80309, USA*

Received 14 November 2003; received in revised form 28 September 2004; accepted 28 September 2004

Abstract

This study develops a geographical information system (GIS) approach for automated processing of the Next Generation Weather Radar (NEXRAD) Stage III precipitation data. The automated processing system, implemented by using commercial GIS and a number of Perl scripts and C/C++ programs, allows for rapid data display, requires less storage capacity, and provides the analytical and data visualization tools inherent in GIS as compared to traditional methods. In this paper, we illustrate the development of automatic techniques to preprocess raw NEXRAD Stage III data, transform the data to a GIS format, select regions of interest, and retrieve statistical rainfall analysis over user-defined spatial and temporal scales. Computational expense is reduced significantly using the GIS-based automated techniques. For example, 1-year Stage III data processing (~9000 files) for the West Gulf River Forecast Center takes about 3 days of computation time instead of months of manual work. To illustrate the radar precipitation database and its visualization capabilities, we present three application examples: (1) GIS-based data visualization and integration, and ArcIMS-based web visualization and publication system, (2) a spatial–temporal analysis of monsoon rainfall patterns over the Rio Grande River Basin, and (3) the potential of GIS-based radar data for distributed watershed models. We conclude by discussing the potential applications of automated techniques for radar rainfall processing and its integration with GIS-based hydrologic information systems.

Keywords: NEXRAD; GIS; Precipitation retrieval; Visualization; Automated techniques

1. Introduction

The National Weather Service's (NWS) Next Generation Weather Radar WSR-88D (NEXRAD) precipitation products are widely used in hydrometeorology and climatology for rainfall estimation (e.g., Seo et al., 1999; Krajewski and Smith, 2002), precipitation and weather forecasting (e.g., Johnson et al., 1998; Grecu

☆Code available from server at http://www.iamg.org/CGEditor/index.htm

*Corresponding author. Tel.: +1 210 458 5445; fax: +1 210 458 4469.

E-mail address: hongjie.xie@utsa.edu (H. Xie).

0098-3004/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2004.09.009

and Krajewski, 2000) and flood forecasting (e.g., Johnson et al., 1999; Young et al., 2000). Currently, the NEXRAD precipitation products are categorized into four product levels according to the amount of preprocessing, calibration, and quality control performed. The Stage I product is an hourly digital precipitation (HDP) directly derived from reflectivity measurements using a Z–R (reflectivity–precipitation) relationship after the application of several quality control algorithms (e.g., Fulton et al., 1998). Stage II is the single radar site HDP product merged with surface rain gauge observations, with a mean field bias correction applied using a Kalman filter algorithm (Smith and Krajewski, 1991; Anagnostou et al., 1999). In Stage III, the Stage II rainfall data from multiple weather radars covering an entire River Forecast Center (RFC) is combined using the average of all available estimates for each Hydrologic Rainfall Analysis Project (HRAP) cell NWS/OH (1997) (Fulton et al., 1998; Reed and Maidment, 1999). Finally, Stage IV is the mosaicked Stage III rainfall product covering the entire Continental United States (CONUS).

The most commonly used NEXRAD product in hydrometeorological applications is the NEXRAD Stage III data (e.g., Young et al., 2000) since it involves the correction of radar rainfall rates with multiple surface rain gauges and has a significant degree of meteorological quality control by trained personnel at individual RFCs (Fulton et al., 1998). However, the Stage III product is difficult to use in conjunction with other geospatial products due to its HRAP projection (Reed and Maidment, 1995, 1999), the large size of storm or multi-year data sets, and the multi-tarred (i.e., compressed hourly rainfall binary files in 1 day are firstly tarred into daily file, daily tarred files in 1 month are then tarred into monthly file) and compressed binary format in which it is distributed by the NWS. These factors hinder researchers from the rapid acquisition and retrieval of high-resolution spatial–temporal precipitation data over large regions. Nelson et al. (2003) have recently recognized the need for easy-access to archival NEXRAD data sets and have implemented a series of procedures for retrieval and acquisition. Although useful, the system implemented by Nelson et al. (2003) does not address the automated processing of readily available NWS NEXRAD Stage III rainfall data nor its distribution and visualization through a geographical information system (GIS)-based web browser.

The motivation of this study is to automatically process the commonly distributed NEXRAD Stage III rainfall data set into a GIS that includes additional geospatial data coverages typically utilized within hydrologic applications, including a digital elevation model (DEM), watershed delineation, land use maps, and satellite imagery. Current GIS technology provides the gateway for the advancement of engineering hydrology into a more exact geoscience relying on accurate model representations of hydrological processes (e.g., Ogden et al., 2001). The data-handling capabilities and improved processing offered by a GIS allow the creation of hydrologic models that enhance the simulation of spatially distributed hydrology (Ogden et al., 2001). A similar need was recently identified by Nelson et al. (2003), who developed an ArcView-based browser to efficiently display NEXRAD reflectivity data for the Mississippi River Basin. In their system, the hourly reflectivity data (in dBZ at 4 km resolution) is in a compressed run-length encoding (RLE) format which can be directly converted to an image format for ArcView display. In contrast, the methods we present in this study build upon a rainfall database, where hourly NEXRAD Stage III precipitation data (in mm/h at 4 km resolution) is used as an input to an automated set of processing algorithms. The output precipitation data are in a GIS grid format which can be directly imported into GIS software, image processing systems, terrain and watershed analysis programs, and hydrological models. As a result, the automated GIS processing leads to an enhanced integration between radar rainfall data and hydrological models, as discussed recently by Garbrecht et al. (2001) and Ogden et al. (2001), who provide an overview of the integrated use of spatial data, GIS, and distributed watershed models. Despite the many advantages that GIS systems have provided to hydrological engineering (e.g., rapid data input, efficient data storage, and easy manipulation and visualization of geospatial data), current technology does not facilitate the input, storage, and manipulation of time-varying, spatially distributed data, such as radar rainfall, in a straightforward way. Here we attempt to ameliorate this problem by creating a structured radar rainfall data set that can be manipulated and queried in a GIS-based hydrologic application.

In order to establish a long-term, GIS-based rainfall database using NEXRAD Stage III precipitation for New Mexico, we use the archival data in the West Gulf River Forecast Center (WGRFC) which are currently available since December 1994. There are roughly 9000 files of hourly precipitation data per year for a large region including the states of Texas and New Mexico. These data files are compressed and multi-tarred to save storage space and enhance fast distribution through an interactive web directory. In a typical hydrological application of the radar rainfall product, the data sets have to be untarred, uncompressed, transformed into a GIS format, and finally projected into a common coordinate system so that it can be integrated with other geospatial data sets. In addition, to conduct watershed-based rainfall analyses, it is necessary to clip the data of a river basin and retrieve a rainfall time series for the subregion or cell of interest. Carrying out these processing steps manually would be extremely time-consuming and error-prone given the number of necessary rainfall fields for a particular storm event, seasonal or climatological analysis. As a result, we believe that automated GIS-based techniques made available to the hydrologic community would enhance the usability of the NEXRAD Stage III rainfall products and provide a means for rapid visualization and manipulation in hydrologic applications.

2. NEXRAD Stage III rainfall data

To establish a long-term GIS-based rainfall database for New Mexico, we utilize the hourly NEXRAD Stage III in a multi-tarred and compressed format directly distributed by the NWS. Compressed hourly rainfall files for a region covering an entire RFC area are tarred into daily files which are subsequently tarred into monthly files for a web distribution. Since most of New Mexico is under the umbrella of the WGRFC, we use the Stage III products archived at the WGRFC data server, available at http://dipper.nws.noaa.gov/hdsb/data/nexrad/wgrfc_stageiii.html. We categorize the data files from the WGRFC into three different types based on file naming conventions and time periods (see Table 1). Types A, B, and C refer to the data files archived before September 1997, between September 1997 and December 1998, and after December 1998, respectively.

An important characteristic of the available NEXRAD Stage III rainfall data is the HRAP geographic projection utilized to distribute the merged RFC product. The HRAP or secant polar stereographic projection is an earth-centered datum coordinate system, while most geospatial data are in a geodetic datum (ellipsoidal earth) coordinate system, such as WGS84 or NAD83. Reed and Maidment (1995, 1999) describe the HRAP projection and its transformations to other geodetic coordinate systems. Correct and accurate data projections are critical for the estimation of radar rainfall data and its subsequent use with topographic, vegetation, and soils data in hydrological models. The NWS (NWS/OH, 1999) provides detailed information on the properties, format, and projection of the original Stage III radar rainfall data distributed through the Internet directory. Preliminary instructions are provided in order to properly display and transfer the HRAP binary file format into a GIS-based product. These materials provided a starting point for constructing the automated processing algorithms discussed in the following section, in particular as related to the coordinate transformation of HRAP radar data over the WGRFC region.

3. Implementation of automated approaches for data processing and visualization

To extract hourly Stage III rainfall products from the distributed tarred bundles, a multi-step procedure for untarring and uncompressing files was implemented utilizing standard UNIX commands. The resulting hourly files are in a binary XMRG format that is converted to an ASCII grid format, which is then converted to a GIS grid format. The projection of the GIS data is defined and subsequently transferred into a common coordinate system using Arc Macro Language (AML) scripts. Multiple-period accumulative rainfalls are then created based on the retrieved hourly GIS files. For a subregion of interest, a set of algorithms are implemented to automatically select, retrieve, and analyze the spatial and temporal properties of the radar precipitation.

Based on the above procedures, batch processing procedures are developed to automatically (1) multi-untar and uncompress the binary (XMRG) files, and transfer the XMRG format to ASCII; (2) transfer the ASCII data to GIS grid format, define the polar stereographic projection, and reproject it to a chosen projection; and (3) clip, retrieve, and sum spatial–temporal rainfall for a region of interest. Fig. 1 shows the flow chart of these batch processes. A brief description of the approaches first appeared in Xie et al. (2003).

Table 1
Data parameters for different periods of time

	Type A (before September 1997)	Type B (September 1997–December 1998)	Type C (after 1998)
xllcorner-polar ster. (HRAP)	−671,512.5 (260)	−671,512.5 (260)	−528,637.5 (290)
yllcorner-polar ster. (HRAP)	−7,620,000.0 (1)	−7,620,000.0 (1)	−7,577,137.50 (10)
Cell size (m)	4762.5	4762.5	4762.5
nclos	455	455	425
nrows	500	500	390

Note: xllcorner, *x*-coordinate at low left corner; yllcorner, *y*-coordinate at low left corner; polar ster., polar stereographic projection; HRAP, Hydrologic Rainfall Analysis Project; nclos, number of columns; nrows, number of rows.

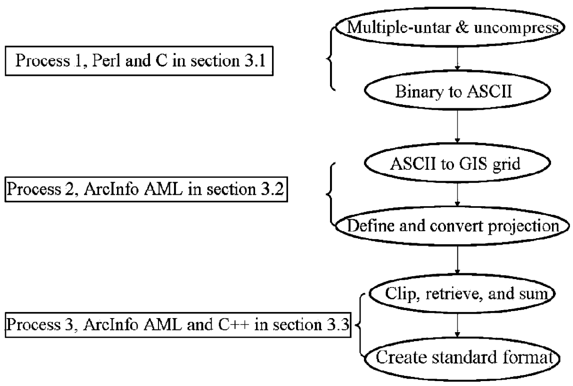

Fig. 1. Work flow chart showing entire automated processes by using commercial GIS and a number of Perl scripts and C/C+ + programs.

3.1. Process 1: from monthly tarred files to hourly ASCII files

To untar and uncompress the monthly bundled files to the hourly XMRG binary format and convert them to ASCII files, three Perl scripts are developed (*munch-A.pl*, *munch-B.pl*, and *munch-C.pl*), each with slight differences that account for the three types of data files described in Table 1. The conversion from the XMRG binary format to an ASCII grid format is carried out by a C language program (*xmrgtoasc.c*) distributed by the NWS Office of Hydrology NWS/OH (1999). For our purposes, we modified the *xmrgtoasc.c* program to convert the unit of precipitation from one hundredth of a millimeter per hour to millimeter per hour. A C program (*reduce.c*) is developed to reduce the dimensions of file types A and B from 455 × 500 pixels to match the domain size of type C files (425 × 390 pixels), the most recent format. A daily accumulation of rainfall over the domain is computed by aggregating the hourly files using another developed C program (*aggregate.c*).

The preprocessing steps outlined above are entirely automated and run in batch mode over radar rainfall data spanning multiple years. Care is taken to ensure the changes in the data format throughout the NEXRAD record are accounted for in the Perl scripts. In total, there are roughly 750 (30 days/month × 24 h/day + 30 daily accumulative files/month) files for each month and about 9000 files for each year. The size of each file is about 1.68 MB, which implies that 1.26 GB is required per month or 15.12 GB/year. In the WGRFC, the NEXRAD Stage III products span from 1995 to 2002, leading to a total database size of 120.96 GB. An example of the preprocessing steps for uncompressing and transforming the distributed binary Stage III radar data is presented below as pseudo-code (Perl 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 455 × 500 to 425 × 390 as the ones in 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
```

3.2. Process 2: from hourly ASCII files to GIS grid files

We developed an AML script (*asc2grd.aml*) to transfer the ASCII files retrieved from the Stage III binary distribution to GIS grid files (ESRI, 1992). As with most GIS systems, the raster grid files contain a matrix of the radar rainfall data values referenced to a particular geographic location that is defined via a set of projection properties (e.g., spheroid, datum, units). The Stage III binary files are referenced to a polar stereographic projection defined over the CONUS. We reproject the data to a more convenient geographic (sphere) coordinate system and then transfer it to a local UTM zone 13, WGS 84 (ellipsoidal earth datum) system for New Mexico. We selected the local UTM system as the coordinate projection for the radar rainfall GIS database for a variety of reasons, including (1) higher local accuracy measured in ground units, (2) square grid cells of equal dimensions in north–south and east–west directions, and (3) ease of conversion of other ancillary data sets (e.g., topography, vegetation, soils).

Coordinate projection is a critical issue in the processing of the NEXRAD Stage III radar rainfall data. Any misalignment with the actual ground coordinates can lead to misrepresentation of rainfall values over specific regions, thus introducing error into radar-rain gauge comparison and radar-based hydrologic modeling. In the distributed data from the NWS, the ArcInfo polar stereographic projection was used to define the HRAP polar projection NWS/OH (1999) (Reed and Maidment, 1999), characterized by a spherical, earth-centered datum of radius 6371.2 km, with a standard (true) latitude of 60°N and a standard longitude (longitude of the projection center) of 105°W. However, the ArcInfo polar stereographic projection actually has default parameters of 6,370,997 m for radius and 60.40884°N for latitude which are slightly different from the HRAP. Since ArcInfo, itself, does not support customization of the default coordinate parameters, we utilized ArcToolBox, a component of ArcGIS, to define the polar stereographic projection and then used the developed AML script to complete the batch process.

Two important substeps are taken to transfer the ASCII files to GIS grid files, to define the HRAP, and to convert the coordinate projection into the local UTM system, as outlined below:

Process 2a: Select one ASCII file of the Stage III rainfall data, and transfer it to grid using ASCII to Grid tool in ArcToolBox, with float as the data type. Then define the projection of the grid file using the *Define Projection (coverage, grids, TINs)* tool in ArcToolBox, selecting polar as the projection. Assign meters for the units, −105° for the longitude of center of projection, and 60° for the latitude of true scale. Finally, by selecting the sphere of radius of 6,370,997 m for the spheroid, we can customize both the semimajor and semiminor axis to 6,371,200 m.

Process 2b: Use the AML script (*asc2grd.aml*) on ArcInfo to first change the workspace to where the rainfall data are located and then to create a file to hold all ASCII files. For each file, the following steps are completed:

1. Get the name of the file, create four new names to hold three files with different projections (polar, geo, and UTM), and one as temporary file name.
2. Transfer ASCII to grid, float type.
3. Define the projection of the grid file by copying polar projection from step 1.
4. Convert the polar stereographic projection to geographic (sphere) projection; resampling method: nearest neighbor.
5. Convert the geographic projection to UTM zone 13, WGS84 system (used in our database); resampling method: nearest neighbor; grid size: 4000 m.
6. Copy the UTM grid file to assigned month directory under year directory.
7. Delete the temp files.
8. Delete the filelist file.

After these steps, all hourly precipitation data are in a GIS grid format. Grid files can be input to any GIS system for analysis and/or further processing. Each hourly record has three different files, each with a different projection: polar stereographic, geographic, and UTM 13 coordinate system. We implement the UTM coordinate system for the rainfall database, but retain the data in other systems for special requests. In this study, we achieved the same result by using a two-step projection instead of the three-step conversion presented by Reed and Maidment (1995, 1999). Their three steps are from (1) polar stereographic to latitude/longitude geocentric, (2) geocentric latitudes to geodetic latitudes, and (3) to other geodetic projections. By omitting the second step, we saved computational time in the batch process. In addition, the data conversion described here reduced the file size over the WGRFC from 1.68 MB (ASCII file) to 0.87 MB (GIS grid file). The total file size is about 652.5 MB/month, 7.83 GB/year, and 62.6 GB for 8 years in the WGRFC. For New Mexico alone, the size of one grid file is 0.17 MB, and the total size is 127.5 MB for 1 month, 1.53 GB for 1 year, and 12.2 GB for 8 years.

3.3. Process 3: clipping, retrieving, and summing of a time series of rainfall

For a regional or watershed application of the NEXRAD Stage III precipitation data, we developed algorithms to automatically clip the rainfall data for a subregion or cell of interest, to extract time series of radar rainfall for the selected area, to accumulate or time-average the gridded rainfall within the chosen time periods and selected area of interest, and to create standard formats that facilitate the comparison between radar and rain gauge measurements. A series of scripts in AML and C++ were developed for these purposes, as detailed below.

3.3.1. Clip a subregion or cell of interest

A script (*gridclip.aml*) is developed to clip a subregion or cell of interest from radar rainfall data in GIS grid format for each individual month for the WGRFC region (~750 files) located within a single subdirectory. The script performs a variety of functions, including (1) create an output text file to store rainfall information

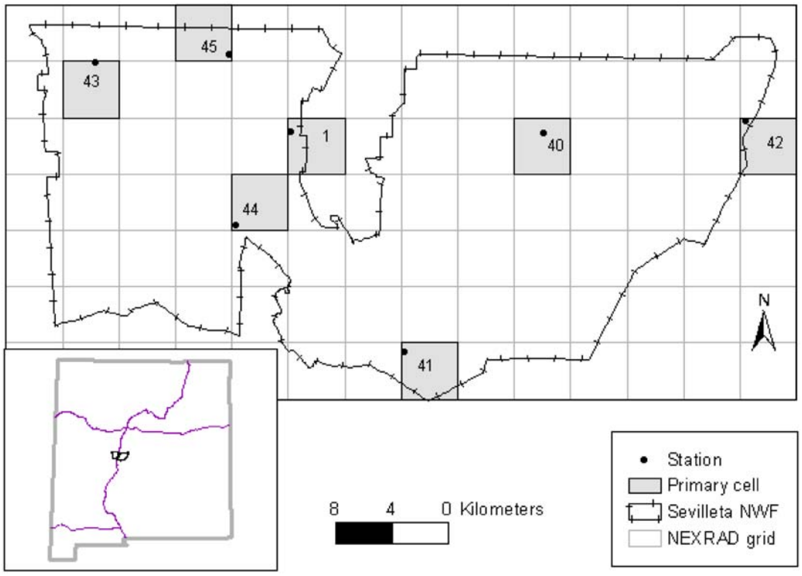

Fig. 2. Seven permanent weather stations with station numbers and NEXRAD grid (in UTM zone 13 and WGS 84 datum) with 4 km × 4 km cell size in the SNWR, central New Mexico, USA.

retrieved from the grid files (e.g., event time, cell position, and rainfall rate), (2) count the number of grid files within the monthly subdirectory, (3) utilize the ArcInfo Grid function *GRIDCLIP* to clip each grid to a user-defined box (x-min, y-min, x-max, and y-max) or polygon coverage and create subgrid files, and (4) load the retrieved rainfall information into the defined output text file. For example, radar rainfall data for the Sevilleta National Wildlife Refuge (SNWR) (Fig. 2) can be extracted as a user-defined region. Within the SNWR, there are 98 cells at 4 km × 4 km resolution. The following illustrates the format of the text output, event time in UTC (MDDYYYYHH or MMDDYYYYHH), x and y cell location (UTM meters), and rainfall value (mm/h):

```
717199923
312926.93068,3811827.52628,0.000
316926.93068,3811827.52628,5.000
320926.93068,3811827.52628,0.000
……
718199915
328926.93068,3811827.52628,0.000
332926.93068,3811827.52628,0.000
336926.93068,3811827.52628,12.000
……
```

We have also integrated the regional clipping within the batch preprocessing steps (Section 3.2) through the script *asc2grd2clip.aml* to facilitate the transfer of ASCII to GIS grid data and extraction of a subregion in a single step.

3.3.2. Sum grids within chosen time periods

To carry out radar rainfall analysis over a particular region within a selected time frame, we developed an AML script (*gridsum.aml*) to process the clipped spatial–temporal grid files (e.g., hourly, daily). This script is capable of accumulating radar rainfall data for a particular storm event, or over any user-defined period, including daily, monthly, seasonal, and multi-year. The script counts the number of grid files whose rainfall data are to be summed in a directory. Subsequently, the algorithm accumulates the radar rainfall rates and stores the cumulative fields in a new GIS grid file for subsequent analysis or query.

3.3.3. Retrieve and export data for cells of interest

The hourly or accumulative rainfall data in a region obtained using the previously described algorithms can be used for numerous applications including input to various hydrologic models and water balance analyses at various scales. However, for validating of radar rainfall data against rain gauge measurements, we need to retrieve rainfall data for each individual cell collocated with a surface gauge or gauges. This direct comparison is facilitated by converting the gridded or cell data into a standard text format that matches available rain gauge data. We developed two C++ programs for this purpose, as detailed below:

(1) The C++ program, *retrievecells.cpp*, generates a file that includes: the event time, weather station ID (in a primary radar cell), and rainfall amount of the cell. The program first reads the output text file (created in Section 3.3.1) storing 1 month of rainfall data and then stores the event time as a 1D array, and *x*, *y*, and rainfall as a 3D array. The program extracts the cells of interest with rainfall values greater than zero within each month. For example, the extracted rainfall data would appear as

```
EventTime StationID Rainfall
717199923 1 0.48
718199900 40 4.07
719199910 42 0.48
721199904 44 1.09
……
```

(2) The C++ program, *standard.cpp*, then produces a file with a data format that matches the rain gauge observations from the SNWR, except the first column used to store the radar ID (i.e., the event time). We keep this ID for referencing to the radar precipitation database. The only difference is that the file in the database has a prefix 'r' and a suffix 'u', saying 717199923 here is r717199923u in the database. The program transfers the event time from UTC to Mountain Standard Time (MST) in YYYY DD HH format.

```
RadarID StaionID Year JulianDay
Hour(MST) Rainfall
717199923 1 1999 198 16 0.48
718199900 40 1999 198 17 4.07
719199910 42 1999 200 3 0.48
721199904 44 1999 201 21 1.09
……
```

We illustrate this capability within the collocated radar-rain gauge measurements in the SNWR (Fig. 2). Within the SNWR, there are seven permanent weather stations with automated rain gauges (labeled 1, 40–45) that have continuous rainfall records from 1995 to the present, thus overlapping the NEXRAD Stage III product archived in the WGRFC. Each station is located within one primary radar cell of 4 km resolution. After retrieving the radar rainfall information for the primary cell, we can compare the rainfall rates from the Stage III product and the rain gauge. Since the Stage III data does not incorporate the SNRW rain gauges during the multi-sensor data fusion, the radar-gauge comparison is a means for validating the NEXRAD product. We can further analyze the rainfall characteristics, such as the correlation coefficient, bias, and error variance, between the radar estimation and gauge observation.

3.4. Geodatabase and data visualization

The processed rainfall data are first stored in a file-based data management system. In order to better store and manage the multi-year, regional rainfall data, we developed an ArcSDE-based enterprise geodatabase in which the spatial data are stored in the Oracle-based relational database management system (RDBMS). The ArcSDE database provides not only the capability for storing all spatial data such as vector, raster, CAD, and annotation, but also leads to better performance for multiple user access, management and query of extremely large raster data at multiple security levels (ESRI, 2003). ArcSDE is a gateway to store and manage spatial data with other business data in an RDBMS and provides the communication between the GIS client and the RDBMS. Especially, ArcSDE offers high-performance in terms of raster storage, query, and display by utilizing raster blocks, pyramids, and compression (Gaskill and Brooks, 2002; McAbee, 2002). We use the 3-tier ArcSDE client/server architecture with both the ArcSDE and Oracle RDBMS running on the same server, which minimizes network traffic and client load while increasing the server load compared to 2-tier system, in which the clients directly connect to the RDBMS (West, 2001; Gaskill and Brooks, 2002).

GIS software packages such as ArcGIS, ArcView, ArcInfo can directly access and display the rainfall data through the file-based database system (grid format) and ArcSDE geodatabase (ArcSDE raster format). In addition, the analysis capabilities inherent in a GIS software, such as geostatistical, spatial, and 3D analysis, can be used to characterize the spatial–temporal distribution and statistical characteristics of rainfall.

Finally, we develop an ArcIMS-based data visualization and publication system to display and publish our rainfall data and additional geo-spatial data coverages typically utilized within hydrologic applications, including a DEM, watershed delineation, state or county boundaries, land use maps, satellite imagery, etc. ArcIMS provides the foundation for distributing high-end GISs and mapping services via the Internet. ArcIMS software enables users to integrate local data sources with Internet data sources for display, query, and analysis in an easy-to-use web browser. ArcIMS revolutionizes the way users can access and interact with Internet mapping and GIS data (http://www.esri.com/software/arcims/).

4. Application examples

The automated approaches for data preprocessing, conversion, and visualization of the GIS-based NEXRAD rainfall database are best illustrated via application examples. For the WGRFC Stage III data, we can complete processes 1 and 2 for 1-year Stage III (~9000 files) within 3 days (based on window 2000 with Pentium IV CPU and 1GB RAM) and 8 years within 1 month of computational time. The automated procedure can be repeated if revisions to the code are required, including code errors and projection changes. Performing these repetitive tasks manually is difficult and time-consuming, especially if changes are required to the entire process. Using the approaches developed in process 3, we can clip any subregion to obtain spatial–temporal rainfall for application and validation purposes. In the following, we present the visualization and integration of various GIS data layers (e.g., topography, hydrography, vegetation) with the Stage III radar product over the SNWR and the Rio Grande Basin. This data integration is valuable for assessing the quality of the radar precipitation product relative to surface gauge measurements and for generating rainfall inputs to hydrologic models.

4.1. Data visualization and integration

The hydrologic application of NEXRAD radar rainfall estimates typically requires data visualization over various regions and time scales (Nelson et al., 2003) as well as the integration to geospatial information such as DEMs, watershed coverages, land use and soil maps, and satellite imagery. The precipitation data in the GIS database is readily displayed, analyzed, and integrated with available geospatial data. For example, Fig. 3 illustrates five GIS layers, describing the Rio Grande, SNWR, New Mexico, Texas, and upper Rio Grande basin, overlaid by a daily NEXRAD precipitation map for July 4, 1998 for the WGRFC region. The spatial distribution of rainfall over the WGRFC shows two major storm systems in the region: (1) a north–south trending system over the Rio Grande basin and (2) a large synoptic system over Texas. Some portions of the Rio Grande Basin are not fully covered by the NEXRAD radar umbrella (e.g., no data available) due to terrain effects, such as the northwestern region within Colorado. Other regions are well covered by multiple overlapping radars (e.g., Maddox et al., 2002). As a result, rainfall estimation from NEXRAD for each basin will depend on the radar coverage and can have artifacts

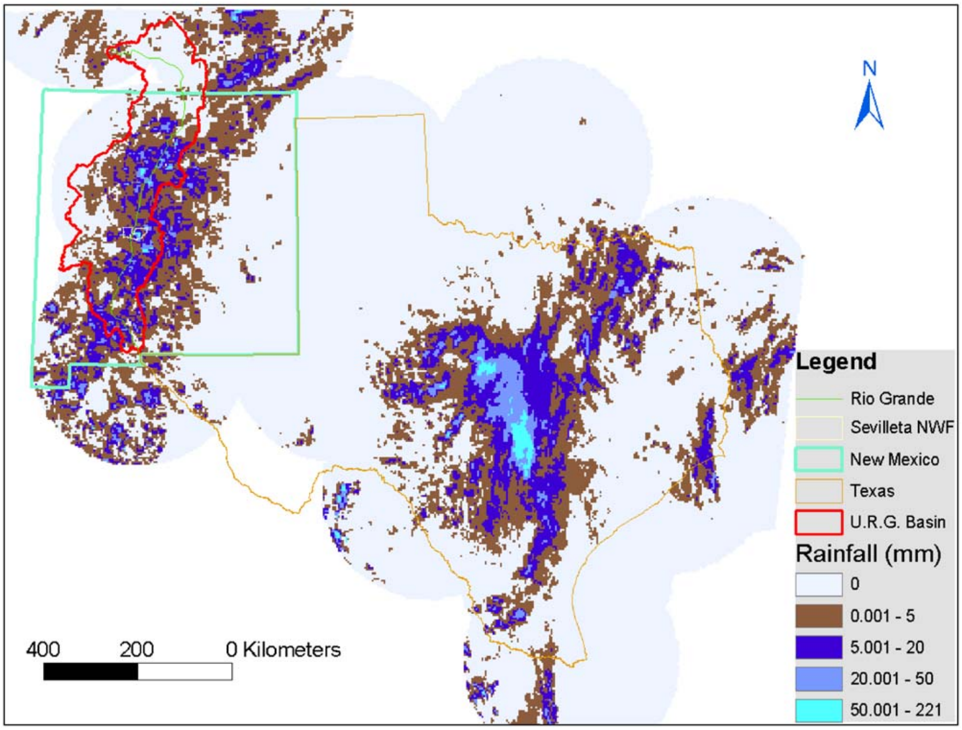

Fig. 3. GIS layers (Rio Grande, SNWR, Upper Rio Grande Basin, and New Mexico and Texas State boundaries) overlaying a daily precipitation (7/4/1998) map (grid) of the WGRFC area, showing two major storm systems: Rio Grande Basin region and central Texas region.

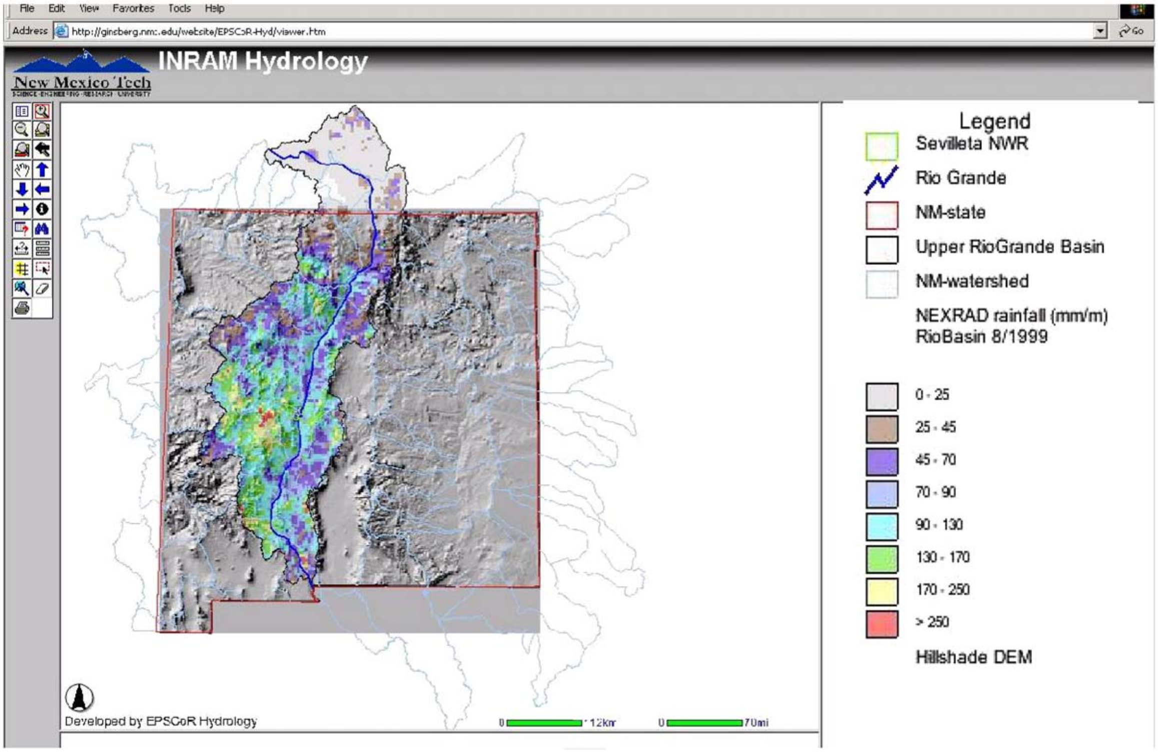

Fig. 4. ArcIMS-based web visualization and publication system (http://ginsberg.nmt.edu/website/EPSCoR-Hyd/viewer.htm).

that require corrections via comparisons to surface rain gauges. Thus, GIS-based radar rainfall visualization overlaid on topographic, hydrographic, and watershed features can improve the interpretation and analysis of storm events over large regions, including possible errors and radar artifacts.

Fig. 4 shows an example from our ArcIMS-based web visualization and publication system (http://ginsberg.nmt.edu/website/EPSCoR-Hyd/viewer.htm). Through this system, a client from a remote location can access, display, zoom in, zoom out, move, identify, and query our data sets, and even print and save maps. For example, in this figure, seven layers are displayed including five feature layers (SNWR, Rio Grande River, New Mexico state, Upper Rio Grande Basin, and watersheds) and two raster layers (rainfall and shaded relief DEM). The NEXRAD monthly rainfall accumulation of August 1998 in the Upper Rio Grande Basin is draped over the shaded relief DEM. We are working on further development for distributing our data sets via this website.

4.2. Spatial–temporal precipitation distribution

To illustrate the capability of the GIS-based automated approaches for analyzing spatio-temporal rainfall data, we extracted and accumulated the precipitation during the monsoon period of 1998 for the Upper Rio Grande Basin. The monsoon period in New Mexico spans from June to September and provides the majority of the annual rainfall through isolated, yet intense thunderstorm activity (Dahm and Moore, 1994). Table 2 summarizes the statistical rainfall characteristics in the 1998 monsoon period computed with algorithms in ESRI ArcGIS. The total monsoon precipitation volume for the season was $7.784 \times 10^{14}\,m^3$ over the $100{,}176\,km^2$ area covered by the NEXRAD radars (Fig. 3). July had the highest amount of precipitation (about 55%) in the 1998 monsoon period, followed by August (28%),

Table 2
Statistics of NEXRAD precipitation in different months of the 1998 monsoon period in Upper Rio Grande Basin

	Mean (mm)	SD (mm)	CV
June	1.4	3.9	2.8
July	67.9	54.6	0.8
August	34.8	25.2	0.7
September	19.7	19.2	1.0
Monsoon	124.3	84.5	0.7

Note: Mean, areal mean rainfall of a month in the Basin; SD, standard deviation of rainfall; and CV, coefficient of variation (SD/mean).

Table 3
Comparison of NEXRAD and rain gauge rainfall (mm)

	Rainfall in 1998 monsoon period	Mean rainfall in monsoon period (1914–1993)
NEXRAD for upper Rio Grande basin	124	
Sevilleta weather stations (7)	136	
NEXRAD cells (7) related to Sevilleta stations (7)	257	
Socorro weather station (24 km south of Sevilleta)		124 (Dahm and Moore, 1994)

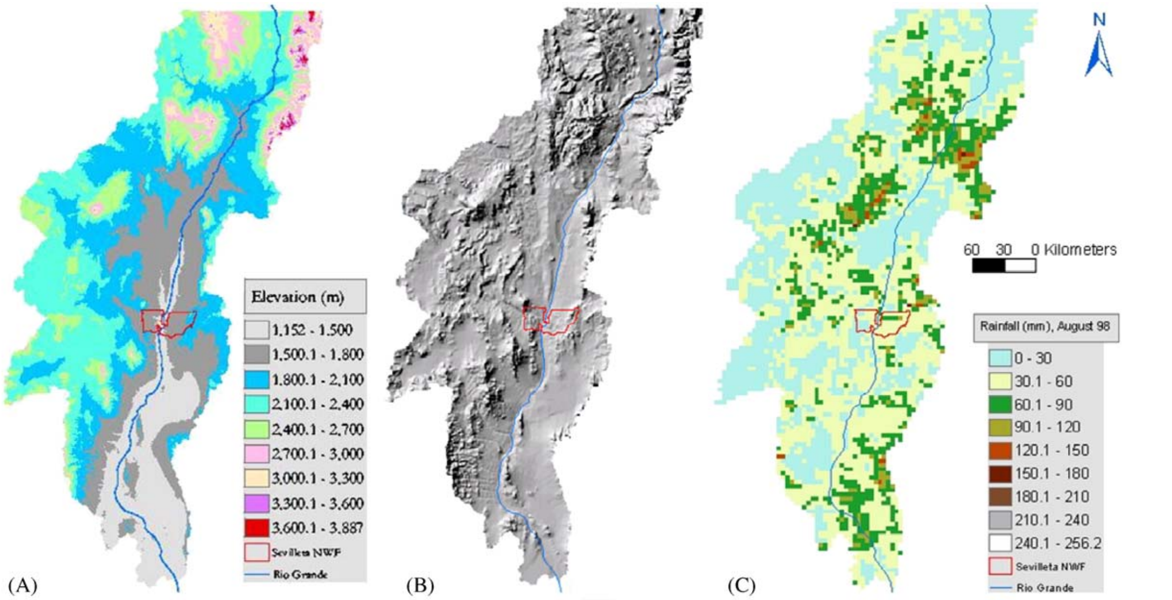

Fig. 5. DEM (A), shaded relief DEM (B), and NEXRAD precipitation (C) of August 1998 in the Upper Rio Grande Basin within New Mexico are readily for various GIS spatial and statistical analyses or input for hydrologic modeling.

September (16%), and June (1%). June has the lowest amount of rainfall accumulation, but highest coefficient of variation (2.8) compared with an average of 0.7 in the monsoon period. This means rainfall in this June was more isolated and scattered than in other months.

Table 3 presents a radar and rain gauge precipitation comparison for the 1998 monsoon period derived directly from the GIS-based Stage III database, as well as the historical rainfall values (1914–1993). We found that the areal mean precipitation (124.3 mm) from NEXRAD for the upper Rio Grande basin matches very well to both the mean precipitation from the seven SNWR weather stations (136.2 mm) and mean rainfall (124.0 mm) over the 80-year record in Socorro, New Mexico (24 km south of SNRW). There is a substantial difference, however, between the average precipitation (256.5 mm) from the seven NEXRAD cells and the collocated gauge measurements (136.2 mm) in the SNWR. This difference is probably due to area (radar) and point (gauge) error. Similar radar-gauge differences have been documented by Zawadzki (1975), Krajewski (1987), Kitchen and Blackall (1992), Smith et al. (1996), Ciach and Krajewski (1999), and Anagnostou et al. (1999).

4.3. Applications for distributed watershed models

The GIS-based Stage III radar rainfall database is a valuable tool for use within distributed hydrologic models that require spatial–temporal variations in precipitation. Typically, a hydrologic modeling operating over a watershed will utilize the topographic and land-surface characteristics to describe a catchment. Subsequently, observed or modeled rainfall is used to force the model and simulate the hydrologic response in the form of streamflow, soil moisture or groundwater table position. The topographic representation in the form of a DEM provides information such as the terrain slope, aspect, flow length, contributing area, drainage divides, and channel network that can be rapidly and reliably extracted using a GIS (Garbrecht et al., 2001). The NEXRAD precipitation data in a GIS grid format can then be merged with the DEM elevation and its by-products within a distributed hydrologic simulation. Fig. 5 illustrates three different GIS layers of Upper Rio Grande Basin in New Mexico: an elevation map (DEM), a hill-shaded DEM, and the NEXRAD precipitation for August 1998. For this month, the mean (40 mm), standard deviation (24 mm), and coefficient of variation (0.6) confirm the large spatial variations in rainfall. This spatial variability is essential for a distributed watershed model operating over the Upper Rio Grande. In addition, the relation between the GIS-based NEXRAD rainfall and the topographic or land-surface features (soils, vegetation, geology) in various subbasins can be investigated readily within this framework.

5. Summary and conclusions

In this study, we have described a set of automated approaches for preprocessing, integrating, visualizing, and analyzing radar rainfall data available from the National Weather Service utilizing GIS techniques. In order to effectively utilize NEXRAD Stage III radar rainfall data, it is important to establish methods for facilitating its use and storage within a GIS database. We developed a number of approaches to automatically uncompress and transfer Stage III data to GIS grid format, define and convert projections, clip and sum the spatial–temporal rainfall for a subregion of interest, and retrieve the data for various applications. We have illustrated our approaches with Stage III precipitation data in New Mexico, in particular for the Sevilleta National Wildlife Refuge and the Upper Rio Grande Basin. The data integration and visualization provided by the automated GIS algorithms allow the analysis of spatial–temporal precipitation distribution and its relation to other GIS data sets for input to hydrologic models. The ArcIMS-based web visualization and publication of the rainfall data allows for open access from the research community to the GIS-based radar rainfall data set.

The performance of the automated techniques has been shown to be superior. Compared to the ASCII format, the GIS grid format can save about 50% of disk space. With the automated approaches, only 3 days of computational time are required to complete the processing for 1 year of NEXRAD Stage III data (about 9000 files) covering the WGRFC region, instead of extensive manual work that can be error-prone, time-consuming, and tedious. Two improvements have been made over the previous methods: (1) an implementation of an ArcGIS ArcToolBox customization for projection definition, (2) a reduced two-step coordinate transformation procedure, saving additional computational effort. The reader is referred to Appendix A for more information on the processing algorithms and their distribution.

Acknowledgments

This work was supported by the NSF EPSCoR grant EPS-0132632 through Institute of Natural Resources Analysis and Management (INRAM) in New Mexico. We acknowledge and appreciate Sevilleta Long-Term Ecological Research Program (LTER) at University of New Mexico for sharing GIS resources and weather data. We would thank Matt Richmond for his help in computer programming. We would also thank Janet Greenlee at the New Mexico State University and Dr. Brian Nelson at the University of Iowa for their constructive comments to improve this manuscript. We appreciate the effort of the reviewers whose helpful comments improved this paper.

Appendix A

Following is a list of source code developed in this project. The source code *xmrgtoasc.c* has been downloaded from NWS website NWS/OH (1999); it has been slightly modified. All these codes and test data sets are available via the website "http://www.iamg.org/CGEditor/index.htm":

1. *munch-A.pl*, *munch-B.pl*, and *munch-C.pl*: batch process to uncompress monthly tar files to hourly XMRG files, and then transfer them to ASCII files;
2. *xmrgtoasc.c*: transfer the XMRG binary format to ASCII format;
3. *reduce.c*: reducing dimensions of types A and B (455, 500) to as the same as type C (425, 390);
4. *aggregate.c*: hourly files in 1 day to be aggregated into daily file;
5. *asc2grd.aml*: batch process to transfer ASCII files to ArcInfo grid files, define the polar stereographic projection for the grid files, reproject it to geographic (sphere) coordinate, and then to UTM 13, WGS 84 (ellipsoidal earth datum) coordinate;
6. *gridclip.aml*: batch process to clip any region of interest, and retrieve spatial–temporal rainfall for the region of interest;
7. *asc2grd2clip.aml*: combined asc2grd.aml and gridclip.aml;
8. *gridsum.aml*: batch process to sum hourly, daily, or monthly grids;
9. *retrievecells.cpp*: retrieve event time, station ID, and rainfall of cells of interest in a region;
10. *standard.cpp*: create a standard format table.

References

Anagnostou, E.N., Krajewski, W.F., Smith, J., 1999. Uncertainty quantification of mean_areal radar-rainfall estimates. Journal of Atmospheric and Oceanic Technology 16, 206–215.

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

Dahm, C.N., Moore, D.I., 1994. The El Niño/southern oscillation phenomenon & the Sevilleta long-term ecological research site. In: Greenland, D. (Ed.), LTER Report, LTER Climate Committee. LTER Publication 18, pp. 12–20.

ESRI, 1992. Understanding GIS: the ArcInfo Method. ESRI Press, Redlands, CA (400pp).

ESRI, 2003. Raster Data in ArcSDE 8.3, an ESRI White Paper. ESRI Press, Redlands, CA (32pp).

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

Garbrecht, J., Ogden, F.L., Debarry, P.A., Maidment, A.R., 2001. GIS and distributed watershed models I: data coverages and sources. Journal of Hydrologic Engineering 6, 506–514.

Gaskill, J., Brooks, D., 2002. Understanding ArcSDE. In: Proceedings of the 20th Annual ESRI International User Conference, San Diego, CA.

Grecu, M., Krajewski, W.F., 2000. A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. Journal of Hydrology 239, 69–84.

Kitchen, M., Blackall, R.M., 1992. Representativeness errors in comparisons between radar and gauge measurements of rainfall. Journal of Hydrology 134, 13–33.

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

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

Johnson, D., Smith, M., Koren, V., Finnerty, B., 1999. Comparing mean areal precipitation estimates from NEXRAD and rain gauge networks. Journal of Hydrologic Engineering 4 (2), 117–124.

Johnson, J.T., MacKeen, P.L., Witt, A., Mitchell, E.D., Stumpf, G.J., Eilts, M.D., Thomas, K.W., 1998. The storm cell identification and tracking algorithm: an enhanced WSR-88D algorithm. Weather and Forecasting 13, 263–276.

Maddox, R.A., Zhang, J., Gourley, J.J., Howard, K.W., 2002. Weather radar coverage over the contiguous United Sates. Weather and Forecasting 17, 927–934.

McAbee, J., 2002. The GeoDatabase and ArcSDE. In: Proceedings of the 17th Annual Northeast Arc Users Group Conference.

Nelson, B.R., Krajewski, W.F., Kruger, A., Smith, J.A., Baeck, M.L., 2003. Archival precipitation data set for the Mississippi River Basin: development of a GIS-based data browser. Computers & Geosciences 29 (5), 595–604.

NWS/OH, 1997. Stage III precipitation processing system, system guide. Hydrologic Research Laboratory. http://www.nws.noaa.gov/oh/hrl/pps/pps.htm.

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

Ogden, F.L., Garbrecht, J.F.L., Debarry, P.A., Maidment, A.R., 2001. GIS and distributed watershed models II: modules, interfaces, and models. Journal of Hydrologic Engineering 6, 515–523.

Reed, S.M., Maidment, D.R., 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), University of Texas at Austin Online Report 95-3 (119pp).

Reed, S.M., Maidment, D.R., 1999. Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. Journal of Hydrologic Engineering 4, 174–182.

Seo, D.J., Breidenbach, J.P., Johnson, E.R., 1999. Real-time estimation of mean field bias in radar rainfall data. Journal of Hydrology 223, 131–147.

Smith, J.A., Krajewski, 1991. Estimation of the mean field boas of radar rainfall estimates. Journal of Applied Meteorology 30, 397–412.

Smith, J.A., Seo, D.J., Baeck, M.L., Hudlow, M.D., 1996. An intercomparison study of NEXRAD precipitation estimates. Water Resources Research 31, 2035–2045.

West, R., 2001. Understanding ArcSDE, ESRI GIS. Redlands, CA (58pp).

Xie, H., Small, E., Hendrickx, J., Richmond, M., Zhou, X., 2003. GIS based NEXRAD precipitation (Stage III) database. In: Proceedings of the ESRI 2003 User Conference, July 7–11, San Diego, CA (16pp).

Young, C.B., Bradley, A.A., Krajewski, W.F., Kruger, A., Morrissey, M.L., 2000. Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. Journal of Hydrometeorology 1, 241–254.

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