1.0. INTRODUCTION

Accurate rainfall estimates during a storm event are invaluable to a forecaster responsible for flood warnings or reservoir operation. Typically concerned with "real-time" forecasting, a forecaster needs to predict actual flows within the next several hours as opposed to simulating a probabilistic design flood. Traditionally, rain gages have been used for measuring precipitation and telemetry for transmitting real-time records from remote gages to a forecast office. A significant drawback associated with gage information is that data is collected at a point and an interpolation scheme is required to produce a rainfall surface and to calculate watershed-average rainfall. As the time period for analysis diminishes, interpolation from gaging stations becomes less and less reliable because short duration rainfall data are less spatially correlated than are long duration data. In more recent years, ground-based radar has emerged as an effective tool for generating a rainfall surface with high temporal resolution (Smith, 1993). The National Weather Service (NWS) produces gridded precipitation estimates as part of its Next Generation Weather Radar (NEXRAD) program. Traditional lumped runoff model formulations need to be adapted to incorporate this new spatial rainfall information. This report describes a procedure for combining a gridded NEXRAD precipitation surface with a gridded description of surface topography from a digital elevation model (DEM) to generate input for a spatially distributed runoff model. Geographic information systems (GIS) provide standardized functions used to accomplish this task.

NEXRAD data are a significant improvement over traditional methods of estimating rainfall in both space and time. Traditional modeling with the unit hydrograph approach involves using watershed-average values for rainfall. Before NEXRAD, watershed-average rainfall values were as good an estimate as any because rain gage networks were typically sparse and the rainfall distribution between gages could not be accurately determined. National Weather Service River Forecast Centers typically spatially averaged precipitation inputs over areas of 750 to 2600 square kilometers and temporally averaged values over 6 to 24 hours (Lindsey, 1993). The NEXRAD StageIII product (described further below) offers precipitation estimates spatially averaged over grid cells of approximately 16 square kilometers and temporally averaged over 1 hour.

USGS Digital Elevation Models (DEMs) provide a description of the land surface. Arc/Info Grid processing of DEMs yields watershed boundaries and gridded estimates of travel length from each DEM cell in a watershed to the outlet. Arc/Info commands allow

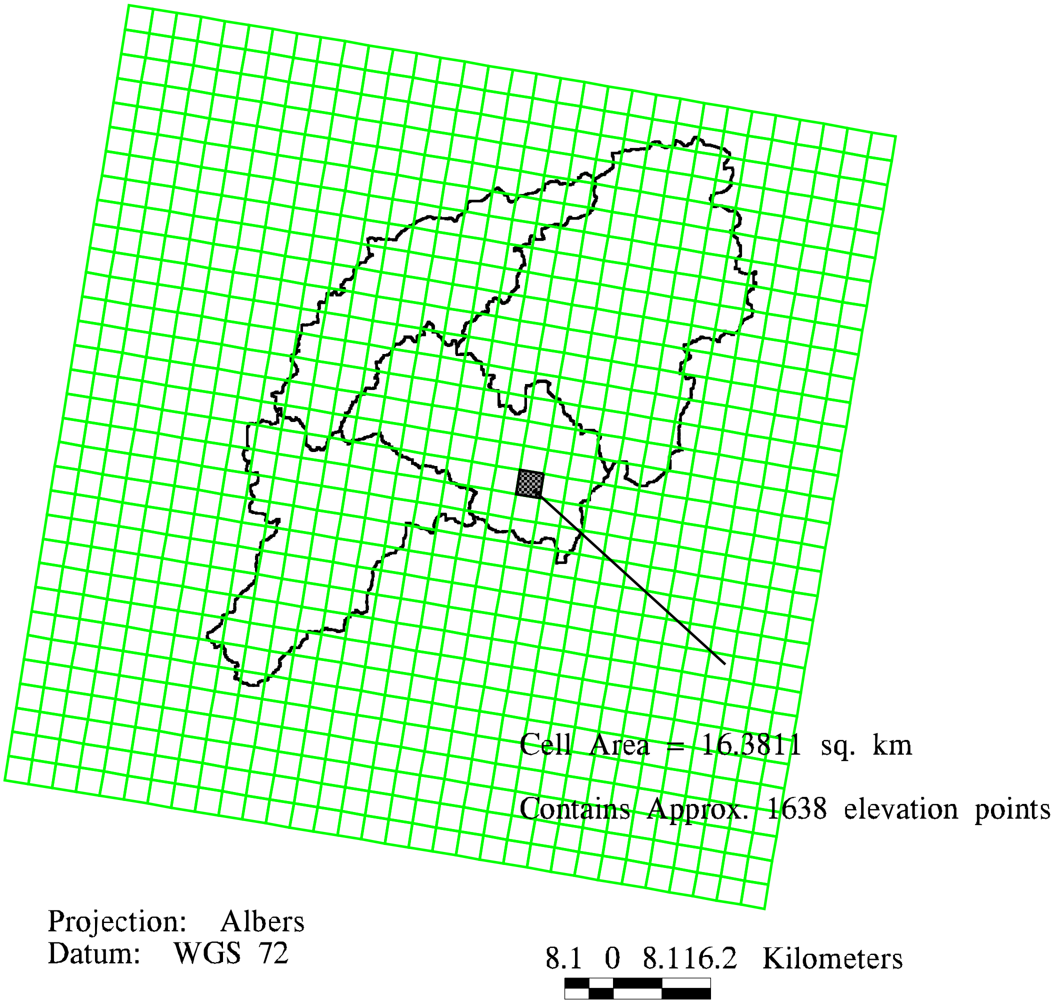

Figure 1.1: StageIII Cells Overlaid on Watersheds at Tenkiller

intersection of watershed boundaries with rainfall cells and subsequent computation of travel length statistics on DEM cells that lie within a given rainfall cell. NEXRAD StageIII rainfall cells are much larger than the DEM cells. For the location and projection parameters selected in this study, a rainfall cell contains about 1600 elevation values as indicated on Figure1.1 . Figure1.1 shows StageIII cells overlaid upon our study area — watersheds draining to the Tenkiller Reservoir in eastern Oklahoma and western Arkansas.

The United States Army Corps of Engineers Hydrologic Engineering Center (HEC) is developing "Next Generation Software" designed to replace HEC1, HEC2, and other codes. The Hydrologic Engineering Center is actively evaluating methods for using the StageIII product as input to the "Next Generation Software." The initial approach proposed by the Hydrologic Engineering Center is to use a modification of the Clark conceptual runoff model dubbed "modClark" (HEC, 1995). The modClark method involves computing rainfall excess for each rainfall cell, translating the rainfall excess to a watershed outlet, and routing flow through a linear reservoir at the outlet to generate a direct runoff hydrograph. Inputs required to run the modClark procedure that can be obtained with GIS are the area of each rainfall cell contained within a watershed and a measure of travel length or travel time to the outlet. Additional parameters such as average slope or travel time for each rainfall cell can be easily computed but do not play a role in modClark at this time. An attractive feature of the modClark approach is that the topographic parameters derived from GIS need to be computed only once. This report details the spatial analysis pre-processing needed for the modClark model.

1.1 EXECUTIVE SUMMARY

An Arc/Info GIS procedure has been developed to create a parameter file of information for input to the modClark distributed runoff program (HEC, 1995). This parameter file contains rainfall cell-ID, rainfall cell area, and average travel length from the rainfall cell to a watershed or subwatershed outlet. Three basic tasks required to obtain cell information are (1) to obtain and process DEM for watershed delineation and calculate travel lengths, (2) to properly define the geographic position of rainfall cells relative to the land surface, and (3) to determine the contributing area and travel length from each precipitation cell to watershed outlet(s). The procedure developed here is generic enough to be applicable anywhere in the conterminous United States; the results for a sample study area, the Tenkiller Reservoir drainage basin are presented in this report. Proper geo-positioning of NEXRAD radar estimates relative to GIS data sets describing land surface features is important and is sufficiently complicated to merit special attention. A working method for positioning NEXRAD radar estimates is presented.

Executing the commands and programs described in this report requires Arc/Info, Version 7.0 or higher with the Grid module, a FORTRAN 77 compiler, and the UNIX utilities *gunzip* and *dd*. The ArcViewII program is also useful but not essential. In addition to verbal descriptions of key commands, the syntax for key commands are printed in bold face as a user would enter them; computer responses are printed in italics. Particularly for tasks 2 and 3 above, command sequences have been automated using Arc Macro Language (AML) and FORTRAN programs. Listings of the codes for all programs are attached in the Appendix to this report. AML provides a framework through which this procedure could be further automated and made more user friendly. In fact, the Hydrologic Engineering Center has already modified the attached programs to develop a user friendly training module for generating a modClark parameter file (HEC, June 9, 1995); yet, the functionality remains the same.

1.2 DESCRIPTION OF DATA

Table1.1 summarizes the data sources used in this study and provides Internet addresses for obtaining the data.

Table 1.1: Internet Addresses for Data Sources

Data Source	Internet Address
ABRFC StageIII	http://gopherpc.abrfc.noaa.gov/abrfc
DEMs	http://sun1.cr.usgs.gov/eros-home.html
HUCs	http://h2o.er.usgs.gov/nsdi/wais/water/huc250.HTML
RF1	http://h2o.er.usgs.gov/nsdi/wais/water/rf1.HTML

In this report, both the raster and vector data structures are discussed. A raster data layer stores values in rectangular array of uniform cells and is referred to as a grid. A DEM is an example of raster data. A vector data layer stores points, lines, or polygons and is referred to as a coverage. A point is a single coordinate value, a line is a series of coordinate values, and a polygon is a connected sequence of lines. A point might be used to represent a stream gaging station, a line to represent a stream, and a polygon to represent a watershed. The RF1 and HUC files described below are examples of vector data. Key concepts used to communicate between these two data structures are that a single cell in a grid is equivalent to a point in a coverage, a line of cells in a grid is equivalent to a line in a coverage, and a zone of cells in a grid is equivalent to a polygon in a coverage. Arc/Info offers several functions that convert data layers between the raster and vector data structures.

1.2.1 NEXRAD Data

As part of the NEXRAD program, the National Weather Service, in conjunction with the Departments of Defense and Transportation, is in the process of deploying WSR-88D (Weather Surveillance Radar - 1988 Doppler) weather radars throughout the country. By 1996, 120 NEXRAD radars will be deployed in the United States (Smith, 1993). A single WSR-88D beam has an effective range of approximately 230 km (US Army Corps of Engineers, 1994), covering more than 166,000 km^2. A three stage data processing procedure incorporating information from radars, rain gages, and satellites has been developed by the National Weather Service to generate high quality hourly precipitation estimates. Different stages of processing are designed for different tasks ranging from flash-flood warnings to stream flow forecasting and reservoir operation. Stage I algorithms use radar information and a limited number of rain gage records to generate 1-hour and 3-hour storm totals on a 2 km rectilinear grid. Stage II processing incorporates additional gage data, resulting in a rainfall surface based on gage data alone which is merged with the radar surface. Surface temperature and satellite information are also used in StageII to eliminate anomalous radar echoes. In StageIII, estimates from several radars are mosaiced into a common grid system so that basin-wide stream flow forecasts can be made. StageIII also incorporates interactive quality control by the forecaster. Both StageII and StageIII products provide hourly estimates in the Hydrologic Rainfall Analysis Project (HRAP) grid system, a 4 km grid in a polar Stereographic map projection (Shedd and Fulton, 1993). StageIII hourly estimates for the Arkansas-Red River Basin can be obtained through Internet within 45 minutes after the hour of estimation (Lillie, personal communication, 1994).

Thirteen National Weather Service River Forecast Centers (RFCs) are responsible for forecasts in major river basins of the United States. The StageIII processing procedure is currently operational only at the National Weather Service Arkansas-Red Basin River Forecast Center (ABRFC) in Tulsa, Oklahoma, but will eventually operate at all thirteen River Forecast Centers. The Arkansas-Red Basin River Forecast Center's StageIII data cover all of Oklahoma, and parts of Kansas, Colorado, Arkansas, New Mexico, Texas, and Missouri (Shedd and Fulton, 1993). Although the focus of this study is on the use of StageIII data, StageII data is also useful in watersheds that are covered by the beam from a single radar, especially where StageIII data is not yet available. Because the StageII and StageIII products are both defined in the HRAP grid system, the procedures for positioning StageII or StageIII cells relative to land surface features in other coordinate systems are identical.

The Hydrologic Rainfall Analysis Project (HRAP) grid as defined by Greene and Hudlow (1982) is used to define the location of each average precipitation value in a StageII or StageIII data set. Flat map coordinate systems like HRAP are defined using a datum and a map projection. A datum consists of a three-dimensional mathematical surface (typically a sphere or spheroid) approximating the shape of the earth and a point of origin. A projection transforms features on a three-dimensional surface into a two-dimensional plane. Some projections can be visualized by imagining a beam of light passing through the earth and producing an image on the projection surface; other projections are purely mathematical. The HRAP coordinate system is defined in the polar Stereographic map projection with a spherical, earth-centered datum of radius 6371.2 km. A regular mesh in the polar Stereographic plane defines the HRAP cells. Rainfall estimates are referenced to the lower-left corner of HRAP cells. While the HRAP system is defined on a spherical datum, GIS data sets describing the land surface, including DEMs, are typically defined on an ellipsoidal datum. Chapter 3 of this report describes the transformations required to properly geo-reference HRAP cells and DEMs.

1.2.2 Digital Elevation Models

3-arc second (3”) DEMs, created by the Defense Mapping Agency (DMA) and distributed by the United States Geological Survey (USGS), are also readily available on Internet. These elevation data were generated by the Defense Mapping Agency from cartographic or photographic sources. For cartographic sources, hypsographic features from maps ranging in scale from 1:24,000 to 1:250,000 were digitized and then elevation data were processed into the desired matrix form. The DEMs are stored according to the names of the USGS 1 : 250,000 map sheets. These USGS map sheets cover a 1° x 2° area but elevation data are stored in 1°x 1° blocks. A user can download data for either the eastern or western half of a USGS map sheet. An elevation data point can be found every 3-arc seconds in both the North-South and East-West directions; thus, a 1° x 1° block of data contains 1201 rows and 1201 columns. As illustrated in “Digital Elevation Models” (USGS, 1990), 3” spacing does not correspond to uniform distances between elevation points on the surface of the earth. Because meridians of longitude converge at the poles, the spacing between data points along a parallel of latitude decreases as one moves north or south away from the equator. The spacing along a meridian of longitude is nearly constant and varies only slightly with the curvature of the earth. With a spherical datum, the length along the surface of the earth per radian of latitude (L_ϕ) and per radian of longitude (L_λ) are given by $L_\phi=R$ and $L_\lambda=R\cos\phi$ respectively. In meters, 3” spacing along a parallel of latitude or a meridian of longitude is equivalent to $(L_\lambda * \pi/180°)/1200$ and $(L_\phi * \pi/180°)/1200$ respectively. For example, at 30° N (Houston, Texas) the spacing between elevation points along a parallel of latitude is about 80 meters, but the spacing along a parallel of latitude at 40°N (Philadelphia, Pennsylvania) is about 70 meters. The spacing between data points along a meridian of longitude is about 92 meters at both locations. More complicated equations for an ellipsoidal datum given by Snyder, (1987), p.25, yield similar results. DEM data downloaded from Internet are typically defined on the WGS 72 datum.

In order to process digital elevation data for a hydrologic study, it must first be projected into a flat map coordinate system so that the coordinates are measured in units of distance rather than degrees. This is necessary because GIS functionality for computing area, distance, and slope depends upon the data being in a Cartesian coordinate system. Using Arc/Info functions to compute a slope where the horizontal dimensions are in degrees of latitude and longitude and the vertical dimension is elevation in meters makes no sense because the distance on the ground for a unit change of 3” of latitude is different from the corresponding distance for 3” of longitude and because the horizontal and vertical dimensions in the slope computation are different. During the map projection process, a new grid is created by resampling data from the original grid at uniform spacing in the projected domain, usually 100 m spacing for grids derived from 3” digital elevation data.

1.2.3 Gage Locations, HUCs, and RF1 Files

In addition to DEMs, watershed delineation requires specifying outlet locations. For the Tenkiller study, geographic coordinates of stream gaging stations provided by the Hydrologic Engineering Center were used to locate outlets.

Although not essential for the spatial analysis described in this study, USGS Hydrologic Cataloging Units (HUCs) and the Environmental Protection Agency's River Reach Files (RF1) provide a useful reference frame and a basis for checking watersheds and streams delineated from the DEMs. HUCs in digital form at 1:250,000 scale and RF1 digital line representation of streams in the United States at approximately 1:500,000 scale are available through Internet at the addresses given in Table1.1 .

1.3 SELECTING A STUDY AREA

Consultation with John Peters at the Hydrologic Engineering Center and engineers at the Corps of Engineers office in Tulsa, Oklahoma, led to the selection of the 4163 km^2 (1607 mi^2) Tenkiller Reservoir watershed as an appropriate study area for testing this GIS procedure. Several considerations went into choosing this watershed. Most importantly, the watershed to be studied had to be located where the StageIII product was available, limiting the study to within the Arkansas-Red River Basin. It was also important to choose a watershed large enough so that it could be broken into subwatersheds and each of these subwatersheds would be large enough to contain several precipitation cells — otherwise the effects of improved spatial precipitation resolution on runoff computations could not be tested. In addition, subwatersheds free of control structures were chosen to simplify routing considerations. An important question to consider is: Into how many subwatersheds should a watershed be divided? The Arc/Info Grid model allows flexibility in delineating watersheds so that subdivision can range from modeling the Tenkiller watershed as a single lumped unit down to treating each digital elevation model cell as a separate watershed.

1.4 SCALES OF ANALYSIS

As shown in Figure1.1 , there are four spatial scales at which hydrologic analysis can be conducted: the 100 m digital elevation cell (0.01 km^2 in area), the NEXRAD 4 km rainfall cell (approximately 16 km^2 in area), the four subwatersheds draining to flow gaging or estimation points (average area = 1040 km^2), or the watershed (4163 km^2) taken as a whole. In this study, the NEXRAD rainfall cell is taken as a hydrologic response unit and its properties are estimated by averaging the corresponding properties of the approximately 1600 digital elevation model cells within the rainfall cell. In particular, the geographic flow distance from each NEXRAD cell to the watershed or subwatershed outlet is estimated by averaging the flow lengths of all digital elevation model cells within the NEXRAD cell boundaries. Where a watershed or subwatershed boundary cuts through a NEXRAD cell, that cell is partitioned into components whose properties are calculated separately. Two test cases have been run — computing average travel lengths from each NEXRAD cell to the watershed outlet for the entire Tenkiller watershed and for Tenkiller divided into four subwatersheds. This report focuses on the more complicated case of analysis with subwatersheds.

1.5 CHOOSING A MAP PROJECTION AND DATUM FOR THE STUDY

Because this study required overlaying data from several different sources, it was necessary to select a standard map projection and datum for analysis. The hydrologic data sets listed in Table1.1 are available at different map scales and in different coordinate systems. Arc/Info GIS allows maps at different scales to be merged easily provided the maps have a common datum, projection, and map units. Arc/Info has built-in capability to transform data sets among many map projections and datums, although some of these transformations are approximate.

A national Albers Equal-Area Conic projection with the parameters listed in Table1.2 was chosen as the standard map projection for this study. This projection is commonly used for maps of the conterminous United States at scales of 1:2,500,000 and smaller (Snyder, 1987). The United States Department of Agriculture uses these same parameters for mapping its State Soil Geographic (STATSGO) database. Using these parameters, the scale error will be slightly less than 1 percent at the center of the United States with a maximum scale error of 1.25 percent at the northern and southern borders (Snyder, 1987). An equal-area projection seems appropriate for hydrologic modeling because drainage area on the globe is preserved in the projected space; therefore, precipitation depth-volume relationships are also preserved. Choosing standard parallels at a regional scale for Oklahoma and surrounding states, different than those listed in Table 1.2, could reduce scale error for the study, but the scale error of the national projection is not large relative to the uncertainty in hydrologic fluxes (i.e., precipitation, evaporation, loss rates). Also, using the national standard map projection allows for flexibility in expanding or contracting the study region without requiring additional coordinate transformations.

Table 1.2: Parameters of Albers Map Projection

First standard parallel	29.5° N
Second standard parallel	45.5° N
Longitude of central meridian	96° W
Latitude of the projection origin	23° N
False Easting	0.0
False Northing	0.0

Geographic datums are defined by an origin, a reference ellipsoid, and the orientation of the ellipsoid relative to the geoid. The two main ellipsoids used in the United States are the Geodetic Reference System of 1980 ellipsoid (GRS 80) which is used in the North American Datum of 1983 (NAD 83) for civilian mapping and in the World Geodetic System of 1984 (WGS 84) for military mapping, and the Clarke (1866) ellipsoid which is used in the earlier civilian datum, the North American Datum of 1927 (NAD 27). The WGS 72 datum based on the WGS 72 ellipsoid is also used; the DEMs downloaded from Internet for the Tenkiller area are defined on the WGS 72 datum. For the Tenkiller study, the decision was made to keep the DEM data in its original datum (WGS 72) and to transform other data sets into this datum. Although transformations of horizontal locations between coordinate systems can be made easily with Arc/Info, no easy adjustment can be made for vertical location. Although not a major concern for present purposes, transforming elevation values between datums might degrade the quality of the information.

For now, choice of an ellipsoidal datum when working with NEXRAD data is a moot point because a spherical earth datum is associated with the rainfall cells. Transformation between spherical and an ellipsoidal datums is discussed in depth in Chapter 3. No transformations between ellipsoidal datums were required at the scale of this study because horizontal shifts were small enough so that required information could be obtained by overlaying data sets defined on different ellipsoidal datums without having to make a datum conversion. For example, in the Tenkiller analysis, coordinates of the stream gaging stations provided by the Hydrologic Engineering Center were most likely taken from a map defined on the NAD 27 datum. An approximate transformation to WGS 72 could be made in Arc/Info by first transforming from NAD 27 to NAD 83 and then from NAD 83 to WGS 72. Experiments showed that horizontal differences between NAD 27 and NAD 83 points are on the order of 55 meters and differences between NAD 83 and WGS 72 are on the order of 10 meters in the Tenkiller area. Knowing the exact location of gaging stations is not required because their locations are only used to choose an outlet cell from a conceptual drainage network defined on a 100 m grid. Incidentally, the geographic coordinates of gaging station locations were reported in decimal degrees to four decimals, implying a North-South position of plus or minus 11 meters — there are 3600 arc seconds in a degree or 0.000277 degrees per arc second; therefore, 0.0001 degrees is equivalent to about 0.36 arc seconds and an arc second is equivalent to about 30 m on the earth's surface so 30 * 0.36 ~ 11 m. Further discussion of map projection and datum issues can be found in Chapter 3 which discusses transforming precipitation cells to the common map projection and datum.