NDOR Research Project Number SPR-1(2) P541
Transportation Research Studies

REGRESSION EQUATIONS

Branden J. Strahm and David M. Admiraal

University of Nebraska – Lincoln
W355 Nebraska Hall
Lincoln, NE 68588-0531
Telephone (402) 472-8568
FAX (402) 472-8934

Sponsored by
The Nebraska Department of Roads
1500 Nebraska Highway 2
Lincoln, Nebraska 68509-4567
Telephone (402) 479-4337
FAX (402) 479-3975

August 2005

Technical Report Documentation Page

1. Report No. SPR-1(2) P541	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Regression Equations		5. Report Date August 2005 6. Performing Organization Code
7. Author/s Branden J. Strahm and David M. Admiraal		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil Engineering University of Nebraska – Lincoln W348 Nebraska Hall Lincoln, NE 68588-0531		10. Work Unit No. (TRAIS) 11. Contract or Grant No. SPR-1(2) P541
12. Sponsoring Organization Name and Address U.S. Department of Transportation Research and Special Programs Administration 400 7th Street, SW Washington, DC 20590-0001		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Regional regression equations were developed to estimate peak-flow magnitudes using Geographic Information systems (GIS). Peak discharges were estimated at return intervals ranging from 2- to 500-years in Nebraska. Flow data from gaging stations located in or within 50 miles of Nebraska were collected. Regional regression analysis, using weighted-least squares (WLS) regression and data from 273 gaging stations, were used to develop equations for seven hydrologic regions. The WLS regression accounted for the differences in record lengths of the annual peak streamflows between sites. Contributing drainage areas ranged from 0.42 to 6,230 mi^2. The equations can be used to estimate peak discharges for selected return periods at sites without flow data.

Digital Elevation Models (DEMs) were the primary data used to extract basin characteristics. The DEMs used in this project are based on 30 m by 30 m data spacing intervals with a Universal Transverse Mercator projection, and are commercially available from the USGS. Morphometric basin characteristics were extracted using ArcInfo software. The DEMs reduced processing time and improved the accuracy of the physical basin characteristics. Soil characteristics were used to improve the accuracy of the regression equations while precipitation data were found to be of lower statistical importance than other characteristics.

Regression equations were developed for seven hydrologic regions in Nebraska. Two sets of regression equations were developed for each region: one representative of basins with drainage areas of less than 10 mi^2 and one for the complete range of drainage areas. The standard error of estimate for the 10- and 25-year frequency equations ranged from 24 to 93 percent for the complete range of drainage areas. The equations for basins with areas of less than 10 mi^2 had a standard error of estimate for the 10- and 25-year return period of 22 to 75 percent. Based on standard error estimates and comparison with other methods, the regression equations worked best for regions located in eastern Nebraska. The equations for western Nebraska regions do not estimate peak flows as accurately because of insufficient peak flow data and high spatial variability of basin attributes.

17. Key Words Nebraska regression equations, peak stream flow prediction, small watersheds, GIS	18. Distribution Statement No Restrictions. This document is available to the public from the sponsoring agency.		
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. Of Pages 145	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of form and completed page is authorized.**

ABSTRACT

Regional regression equations were developed to estimate peak-flow magnitudes using Geographic Information systems (GIS). Peak discharges were estimated at return intervals ranging from 2- to 500-years in Nebraska. Flow data from gaging stations located in or within 50 miles of Nebraska were collected. Regional regression analysis, using weighted-least squares (WLS) regression and data from 273 gaging stations, were used to develop equations for seven hydrologic regions. The WLS regression accounted for the differences in record lengths of the annual peak streamflows between sites. Contributing drainage areas ranged from 0.42 to 6,230 mi^2. The equations can be used to estimate peak discharges for selected return periods at sites without flow data.

Digital Elevation Models (DEMs) were the primary data used to extract basin characteristics. The DEMs used in this project are based on 30 m by 30 m data spacing intervals with a Universal Transverse Mercator projection, and are commercially available from the USGS. Morphometric basin characteristics were extracted using ArcInfo software. The DEMs reduced processing time and improved the accuracy of the physical basin characteristics. Soil characteristics were used to improve the accuracy of the regression equations while precipitation data were found to be of lower statistical importance than other characteristics.

Regression equations were developed for seven hydrologic regions in Nebraska. Two sets of regression equations were developed for each region: one representative of basins with drainage areas of less than 10 mi^2 and one for the complete range of drainage areas. The standard error of estimate for the 10- and 25-year frequency equations ranged from 24 to 93 percent for the complete range of drainage areas. The equations for basins with areas of less than 10 mi^2 had a standard error of estimate for the 10- and 25-year return period of 22 to 75 percent. Based on standard error estimates and comparison with other methods, the regression equations worked best for regions located in eastern Nebraska. The equations for western Nebraska regions do not estimate peak flows as accurately because of insufficient peak flow data and high spatial variability of basin attributes.

DISCLAIMER

ACKNOWLEDGMENTS

This research project would not have been possible without the funding provided by the Nebraska Department of Roads. We would also like to thank the engineers at the Department of Roads for their assistance with this project, and especially Kevin Donahoo and Robert Carnazzo for providing us with project data and administrative advice.

We also appreciate the assistance of the Department of Natural Resources of the State of Nebraska and The Conservation and Survey Division of the University of Nebraska, who provided us with much of the electronic data and technical advice that we needed to complete this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Peak flow characteristics such as magnitude and frequency of peak discharges, are important considerations in the design of highway bridges and culverts. A majority of the literature about peak flow predictions contains information on large, perennial streams for the design of major drainage structures. Obtaining an accurate estimate of the relationship between extreme flows and recurrence interval (Q-T relationship) is difficult if there is no flow record at the site of interest.

Regional peak flow frequency analysis makes it possible to estimate extreme flow values in locations with limited flow data using data from watersheds with similar hydrologic responses. Historically, regional peak flow frequencies have been used to improve the accuracy of extreme flows at gaged sites and to estimate flows at sites where no stream flow record is available. Generally these relations do not work as well for very small watersheds, particularly for watersheds with ephemeral streams. This is because of the topographic resolution used and the lack of flow data for small watersheds.

With the recent introduction of high resolution digital maps, relevant physical basin characteristics can be delineated with improved accuracy. The problem of assigning a flood risk to a particular flow value has received substantial attention in the literature (Beckman, 1976; Cordes and Hotchkiss, 1993; and Soenksen et al., 1999a). Estimating flood risk through peak flow frequency is limited by the lack of available data necessary to predict the risk associated with return periods greater than the period of record. A common regional peak flow frequency technique is to transfer information from surrounding gaging stations, and apply it to events in a given location.

1.1 Background

Using GIS, the United States Geological Survey (USGS) in cooperation with the Nebraska Department of Roads (NDOR) have developed an extensive set of peak-flow frequency relations for the entire state of Nebraska (Soenksen et al., 1999a). But the drainage areas and maps used to develop the relations were larger than those typical of NDOR roadway construction projects. Because of the low topographic resolution, the accuracy of the basin characteristics is not as useful for small NDOR projects. The topographic resolution used in the USGS equations was significantly lower than the 7.5-minute resolution normally used for NDOR projects. Although the USGS regression equations work for large basins they are difficult to apply to most NDOR projects because of how they were developed, and they are not as accurate for many of the smaller basins typical of NDOR projects.

The latest update in Nebraska's regression equations divided the state into seven hydrologic regions. Regionalization assumes homogeneity within each region, and should increase the predictive accuracy of the regression equations. Nebraska's regions were created based on permeability, percent non-contributing drainage area and watershed divides (Soenksen et al., 1999a). Major basins include the Big Blue River, Elkhorn River, Salt Creek, Big Nemaha River, and the Missouri River tributaries.

1.2 Regression Analysis

A relationship was developed between basin characteristics and peak-flow characteristics using a weighted-least squares (WLS) regression. WLS accounted for the differences in record lengths of the annual peak stream flows between sites. Basin characteristics were chosen based on minimizing the standard error between the observed and predicted values determined from the regression analysis. Each region has equations developed to predict peak discharge for the recurrence intervals of 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-years in Nebraska.

With the exception of the High Permeability region, two sets of regression equations were developed for each region: one representative of basins with areas of less than 10 mi^2 and one for the complete range of drainage areas. The High Permeability region does not have equations for basins less than 10 mi^2 because of the small amount of regional data. Contributing drainage area had the highest statistical relationship to peak flows and was used in all equations.

Geographic Information Systems (GIS) datasets used in Water Resources have provided a consistent method for watershed and stream delineation. Updates in GIS technology have allowed for the extraction of basin characteristics that previously were undefined. This included delineating stream networks and contours within the basin. The improved spatial resolution will influence some of the calculated basin characteristics because they are a function of the data scale. Twenty-five morphometric characteristics were extracted from the 7.5-minute Digital Elevation Models (DEM). The DEMs reduced processing time and improved the consistency of physical basin characteristic calculations. Improved resolution allowed for the analysis of watersheds smaller than 1.0 mi^2.

The performance of the updated regression equations was then evaluated by comparing them with other methods of determining peak flows for twelve Nebraska Department of Roads (NDOR) projects. The twelve locations represent six hydrologic regions used in the development of the regression equations for areas of less than 10 mi^2. A site description and discussion was given for each ungaged stream.

1.3 Purpose

The goal of this project is to redevelop the USGS regression equations using the 7.5-minute quad maps so that the equations are more appropriate for smaller watersheds. The regression equations are based on variables similar to those used by the USGS using lower resolution data, but they have been developed for small basins as well as large basins. The expected benefits of the project are:

- A new set of regression equations will make it possible to take advantage of new GIS technologies to rapidly calculate accurate estimates of peak flows for both small and large watersheds.
- The regression equations and software will reduce processing time and improve peak flow predictions.
- The procedures developed will allow regression equations to be more easily updated as new flow data and Geographic Information Data become available.
- The use of 7.5-minute Digital Elevation Models will improve the spatial resolution so that the revised equations will be appropriate for maps with higher resolution.

1.4 Outline of Other Chapters

The remainder of this report includes a literature review, the methods and procedures used to obtain the regression equations, the resulting equations, comparisons with alternative methods of computing peak flows, and conclusions. In chapter 2, the literature review presents the regionalization procedures, Nebraska’s peak flow history, National flood frequency programs, Geographic Information Systems and the regression models. The methods and procedures used to develop the regional regression equations are explained in Chapter 3. The regional regression equations and a discussion are given in Chapter 4. The regional regression equations are compared with existing Nebraska Department of Roads methods in Chapter 5. A summary and conclusions of the research are presented in Chapter 6.

2. LITERATURE REVIEW

An estimate of peak streamflow frequency is useful for floodplain management and for cost effective design of highway bridges and culverts. Most literature on the subject contains information about larger, perennial streams for the design of major drainage structures. Accurate estimates of return period peak discharges are difficult when there is no flow record at the site of interest. Regional peak flow frequency analysis makes it possible to estimate extreme flow values in locations with limited flow data using data from watersheds with similar hydrologic characteristics.

The first section of this literature review discusses the regionalization procedures and methods used to develop regional regression equations. This section includes Nebraska's regions, data splitting, probability-weighted moments, and the region-of-influence approach. The second section examines Nebraska's peak flow history based on reports from the USGS and others. In this section, the application and limitations of Nebraska's peak flow frequency analysis is discussed. The third section contains information and data from the USGS National Flood Frequency program through the application of regional regression equations. Analyses done for states that are close in proximity to Nebraska are considered the most important, and the methods used for relevant projects are discussed. The fourth section presents the application of Geographic Information Systems to hydrologic predictions. A discussion is given on the relevant datasets, resolution effects and the extraction of basin characteristics. The last sections introduce the Log Pearson Type III distribution, multiple regression models, least squares regression and standard error of estimate.

2.1 Regionalization

Regionalization is used in peak flow frequency predictions to improve the accuracy of estimating equations. Gaging stations can be grouped by geographic location, flow characteristics or by basin attributes. Grouping stations increases the homogeneity within, while also increasing the heterogeneity between groups. The homogeneity within each region improves the accuracy of prediction within that region.

2.1.1 Nebraska Regions

To accurately predict peak flow frequency, knowledge of peak flow characteristics and basin attributes is needed. Beckman (1976) investigated peak flows in Nebraska for recurrence intervals of up to 100 years for natural flows. Regionalization was accomplished by using basin and climatic characteristics for selected watersheds. Five regions were created for the entire state of Nebraska.

Western and central Nebraska was divided into two regions based on soil type. Region 1 is widely scattered, while region 2 is made up of the sandhill terrain. Figure 2.1 shows the generalized areas of soils in Nebraska. Sandhill streams are predominately groundwater fed and have small contributing drainage areas, giving them relatively steady flows. Depressions, lakes and soils with large infiltration rates result in large differences between total and contributing drainage area in the sandhills. The eastern part of the state is divided into three regions by watershed divides. Region 3 includes almost the entire eastern side of the state. Region 4 is the loess-hill area which contains the lower portion of the Loup River system originating from sandhill streams. The Big Blue River basin is Region 5 that extends into Kansas. The regional divisions that were created divided the state into five logical hydrologic regions.

In the last update of Nebraska's regression equations, the state was subdivided into seven hydrologic regions. Western Nebraska was regionalized by permeability and the percent of noncontributing drainage area. The Upper Republican River basin was used in the southwest corner of the state. The central and south-central region was developed from Loup River tributaries and streams located in the Platte River and Republican River floodplain. The eastern regions were developed from watershed divides. Major basins included the Big Blue River, the Elkhorn River, Salt Creek, and the Big Nemaha River. Figure 2.2 illustrates the seven hydrologic regions used in Nebraska's regional regression equations.

2.1.2 Data Splitting

Tasker (1982) compared methods of hydrologic regionalization for gaging stations in Arizona. Data for 221 stations were used to demonstrate the usefulness of data splitting for model comparison and deciding which scheme for best defining sub-regions. Data splitting is used to compare methods of determining homogeneous hydrologic regions, which increases predictive accuracy. Data splitting uses statistics to divide data into groups to improve peak flow predictions. Cluster analysis of stations is based on characteristics of a stream's drainage area. Clustering is done to improve the peak flow estimates by grouping statistically relevant characteristics. This is a more objective method of creating regions, which creates clusters with similarities.

Different clusters have different characteristics such as: drainage area, mean annual precipitation, basin elevation, and the soils index for each basin. Tasker concluded that large aerial regions can be sub-divided into hydrologically homogeneous regions to improve peak flow prediction for ungaged streams. Also, cluster analysis can be effective if some form of validation can be done, such as data splitting, to decide on several possible groups.

The grouping of basins for regional peak flow frequency analysis can be based on basin characteristics instead of geographical regions (Wiltshire, 1985). Geographical homogeneity cannot be guaranteed because neighboring basins can be physically different. To reduce the bias in regionalization, grouping basins on measurable characteristics was applied. Wiltshire concluded that basins with high annual rainfall totals, which usually have soils near field capacity, yield large peak flows with small variability in magnitudes. Drier regions with impermeable soils yield variable peak flow responses. For Wiltshire, grouping regions based on basin characteristics resulted in substantial improvements in peak flow estimates compared to previous projects.

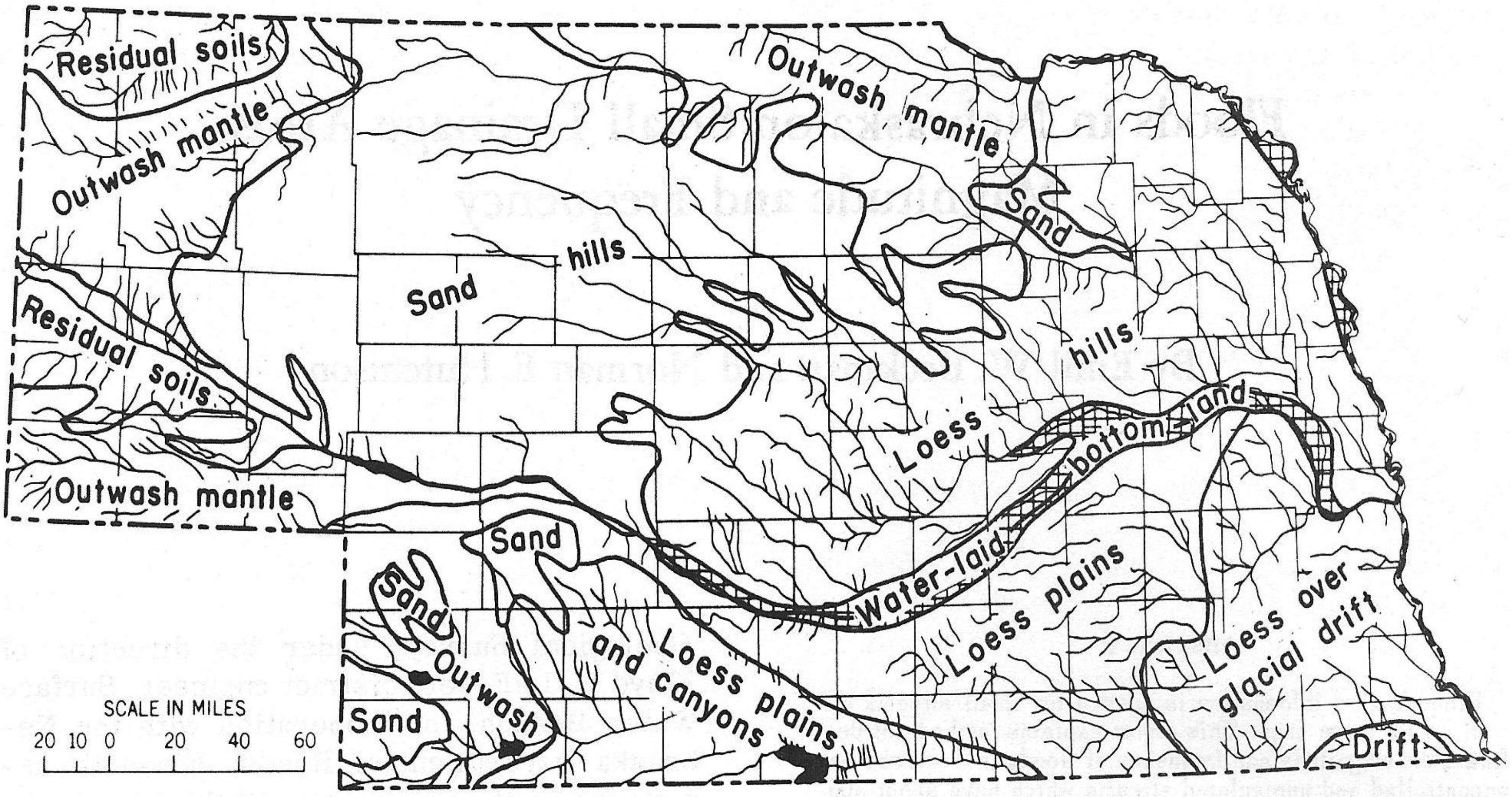


Figure 2.1: Generalized areas of soils in Nebraska (Furness 1955).

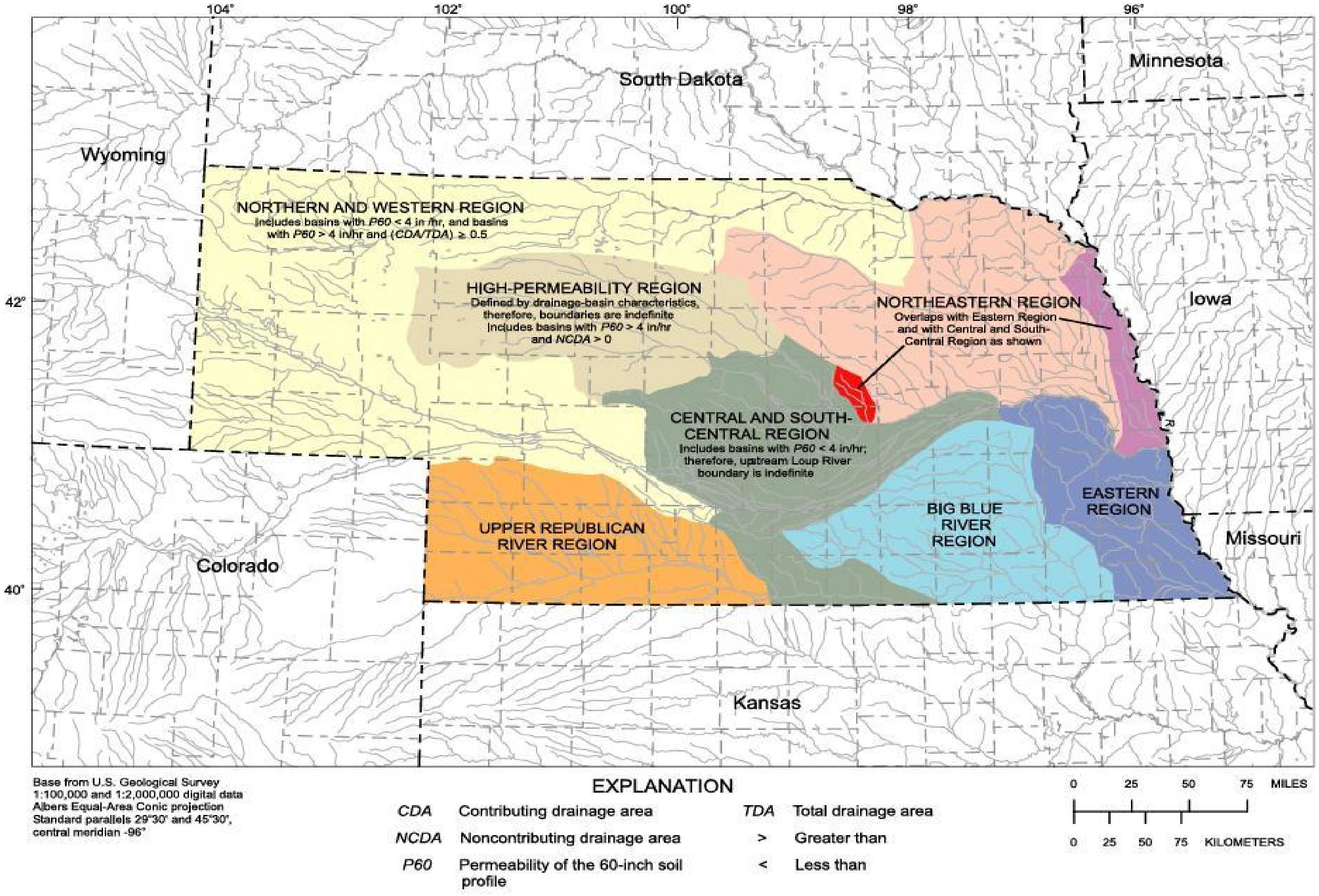


Figure 2.2: Hydrologic regions in Nebraska for unregulated peak-flow frequency equations (Soenksen et al., 1999a).

A procedure to classify and optimize statistics in drainage basins for homogeneous groups was examined by Wiltshire (1986a, 1986b, 1986c). A dimensionless relationship of peak flows and mean annual peak flow is commonly used for hydrologically homogeneous regions. The relationship gives a regional average frequency curve for the combined data. Preferably, a group of basins should be different from other groups and have homogeneity of peak flow frequency characteristics to allow group average curves to be accurately defined. The proposed method for forming groups involves first identifying a basin characteristic. Then the station data are divided into groups based on a reference value of the basin characteristic. An optimum solution is found by iteratively adjusting the reference value of the basin characteristic. This process may involve subdividing the regions into three or more significantly different groups. It also can utilize multiple basin characteristics at each possible arrangement of group boundaries. Wiltshire subdivided Scotland into five geographical regions.

Wiltshire (1986c) also examined cluster analysis to achieve homogeneity for regional peak flow frequency analysis. An alternative to forming regions in geographical space is to identify groups of basins that have similar morphometric or soils characteristics. Heterogeneity is expected of regions which contain variable basin characteristics and peak flow frequency curves. Cluster analysis was used to subdivide basins in Scotland into ten homogeneous clusters. The clusters had similar mean annual specific discharges, similar basin characteristics, and peak flow series with similar coefficients of variance (a measure of flood variability from year to year). The ten groups show a high degree of homogeneity due to the basins being grouped by hydrological similarity.

The identification of hydrologically homogeneous regions can be achieved on the basis of physical characteristics of the drainage area. Acerman and Sinclair (1986) classified watersheds according to basin characteristics for peak flow analysis. A likelihood ratio test was applied to 168 basins to test the homogeneity for a single regional frequency relationship. A dimensionless peak flow frequency curve was used to relate an estimated mean annual peak flow to peak flows of less frequent occurrence. Acerman and Sinclair found that neighboring basins within a region may be physically and hydrologically very different or similar. Regionalization is done with respect to physical characteristics without referring to stream discharge. The dominant characteristics affecting peak flow predictions were found to be the basin area, stream density, soils, stream storage, main channel slope, and climatic data. Cluster analysis was also used to provide a systematic method to a multivariable problem. It allows for two geographically close basins to have completely different sets of stations that are representative of them.

2.1.3 Regionalization Methods

Regional peak flow frequency estimation has shown that accurate peak flow relations are possible when peak flow frequency distributions are identical at all sites in a region (Lettenmaier et al., 1987). Regional estimation methods using the three parameter generalized extreme value (GEV) distribution are insensitive to some regional heterogeneity in the coefficient of variation. The objective of Lettenmaier et al. was to explore the strength of regional peak flow methods using probability-weighted moments (PWM) and sensitivity of selected methods to check the performance of regional peak flow methods in moments higher than first order. After the region of influence (ROI) for a site has been found, the PWM estimation is an efficient way of combing flow data in regional peak flow frequency analysis. It can be found from:

$$M_{jr} = \frac{1}{n_j} \sum p_i^r x_i \tag{2.1}$$

$$p_i = \frac{(i - 0.35)}{n_j} \tag{2.2}$$

Where M_{jr} is a sample estimate for the order *r* PWM for site *j*, p_i is the probability for the ith peak flow, x_i is the peak flow maximum and n_j is the number of annual maximum flow values for site *j*. The quality of the available estimates should improve with increased record length. Lettenmaier et al. (1987) found that two-parameter regional peak flow frequency estimation method can perform well if it is assumed that the sample distribution is similar to the population. Also, three-parameter GEV tends to give large variance in peak flow estimates for at-site applications.

Delineation of groups for regional peak flow frequency analysis using Monte Carlo simulations were researched by Burn (1988). Basin characteristics for rivers in Manitoba were explored and applied to the regional Monte Carlo simulation. The purpose of regionalization is to identify a group of stations that are similar with respect to peak flow events and frequency (Q-T relationship). When defining homogeneity, there is a trade off between quality and quantity of data. Adding more stations to a region illustrates that there is more information available, but additional information may be poor and dissimilar from other stations. The result is a trade off between how many regions should be included in the regional analysis. The Connection between the Q-T relationships is due to rainfall patterns and physical basin characteristics of the streams. A total of 41 stations with at least 25 years of record were applied for the final analysis. An f-test was used to compare the coefficient of variation (CV) between regions at the 1% level of significance, resulting in three regions. The Monte Carlo simulation was also used to further evaluate the characteristics of each region using probability weighted moments and GEV distributions. The accuracy of predicting extreme flows was improved through regionalization. Grouping regions that have similar CV's will result in regions with an ideal homogeneity.

Peak flow frequency analysis for ungaged sites was examined by Zrinji and Burn (1994) using a region of influence approach (ROI). The ROI approach incorporates a homogeneity test while selecting stations that are part of a region. The purpose is to develop a new approach to regional analysis at ungaged sites. The regionalization approach uses basin characteristics to find the similarity between regions and an ungaged stream. Through the regionalization process the characteristics that are most influential in the regression equations can be determined. The ROI provides no fixed boundaries, because each site is first considered its own region. Regions were generated by a Euclidean distance:

$$D_{jk} = \left[\sum W_i (X_{ji} - X_{ki})^2\right]^{\frac{1}{2}} \tag{2.3}$$

Where D_{jk} is the weighted Euclidean distance from site *j* to *k*, W_i is the weight associated to the basin characteristic *i*, and X_{ji} is a standardized value for basin characteristic *i* for site *j*. Weights are important when defining which basin characteristics are important when defining similarity between regions. The ROI process uses a combination of basin characteristics of all sites and flow data from gaged stations to create a flexible homogeneous region. This leads to improved extreme flow estimates for regions without data.

2.2 Nebraska Peak-flow History

Several projects by the USGS and others in Nebraska have been done to improve prediction of peak flows at given recurrence intervals. These projects include studies by Furness (1955), Beckman and Hutchison (1962), Patterson (1966), Matthai (1968), Beckman (1976), Cordes and Hotchkiss (1993) and Soenksen et al. (1999a). The application and limitations of Nebraska's regional peak flow frequency analyses are discussed below.

2.2.1 Furness Frequency Relations

Furness (1955) developed peak flow relations for two regions, with drainage areas greater than 100 mi^2. The relations are used to define the average magnitude of peak flows for return periods of up to 50 years. It was determined that Nebraska peak flows are caused by a combination of physiographic factors, climate, and regulation. Soils were used to regionalize Nebraska into two subregions: the sandhills region and everything except the sandhills region. Figure 2.1 illustrates the generalized areas of Nebraska's soils.

Even with the significant regional variation of Nebraska's climate, there is very little correlation between precipitation amounts and peak flow magnitudes. Regulation of Nebraska's streams, at that time, resulted in over one million acres of affected peak flow runoff. The most influential factor in peak flow magnitude is stream drainage area. All other peak flow factors were lumped into a coefficient:

$$Q_{2.33} = CA^{0.7} \tag{2.4}$$

Where $Q_{2.33}$ is the mean annual peak flow (cfs), C is the peak flow coefficient, and A is the contributing drainage area (mi^2).

A nomograph was used to compute the peak flow discharge for the frequency desired. The report includes the necessary figures and maps for defining peak flow frequency for any stream in Nebraska with a drainage area of greater than 100 mi^2.

2.2.2 Circular 458 Method

The prediction of peak flows for small ungaged watersheds in Nebraska is important for the design of control structures. Peak flow magnitudes and frequencies on small watersheds in Nebraska were first examined by Beckman and Hutchison (1962). Peak flow discharges of watersheds with drainage areas of less than 300 mi^2 were compared. Soils and climate data vary widely across the entire state. Annual rainfall amounts gradually increase from 14 inches in the west to 34 inches in the southeastern corner. Thunderstorms provide a large percentage of the rainfall from May through July. Soils range from very high permeability in the sand hills, to Loess deposits to the east. The annual peak flow series was used to examine peak flow records. The state was divided into 10 regions based on soil type and watershed divides (Figure 2.3).

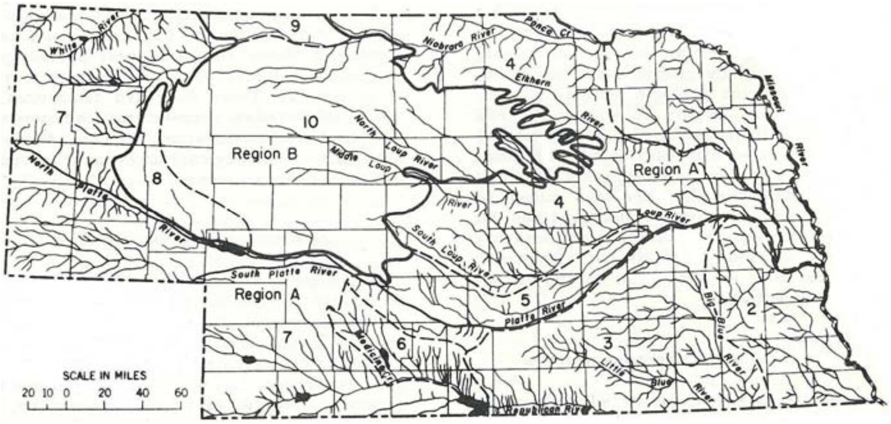


Figure 2.3: Map of Nebraska showing flood-frequency regions and hydrologic areas (USGS Circular 458).

The gaging stations for each region were then grouped and analyzed. Frequency relations were developed by investigating the maximum peak flows for 142 gages in Nebraska. Discharges for given return periods were graphically compared to contributing drainage area. It was found that the accuracy of peak flow magnitudes for selected return periods is dependent upon the number of stations and the length of record. The regional frequency curves should not be considered with confidence beyond a return period of 25 years.

The first step when using Beckman's method is to determine which of the 10 hydrologic regions is applicable to the design. Second the mean annual peak flow can be determined from the contributing drainage area using Figure 2.4. Finally, the mean annual peak flow is related to the recurrence interval by a ratio shown in Figure 2.5. Curves A and B are representative of the sandhills region and everything but the sandhills region.

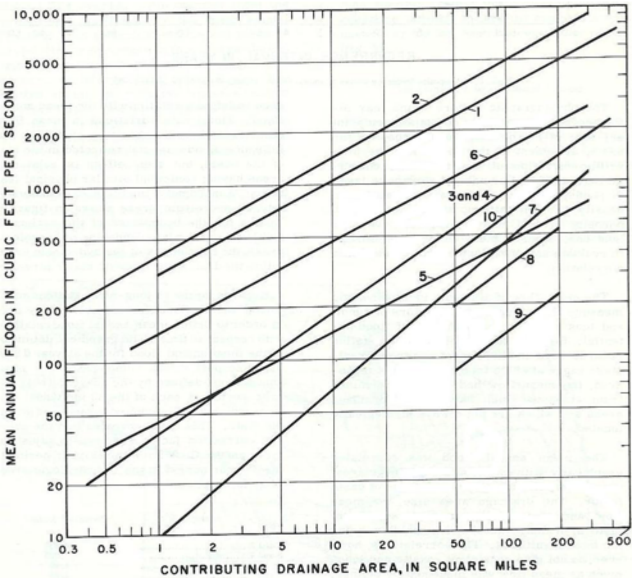


Figure 2.4: Variation of mean annual peak flow with contributing drainage area in hydrologic areas 1 – 10 (USGS Circular 458).

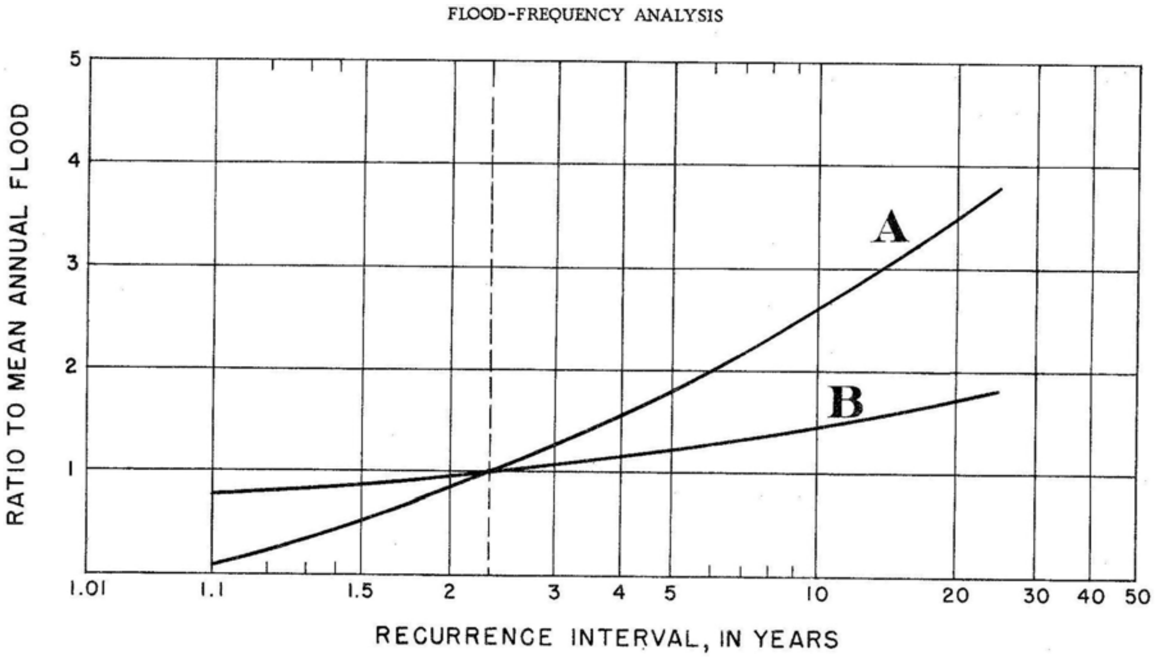


Figure 2.5: Composite frequency curves of annual peak flows (USGS Circular 458).

2.2.3 Water Supply Papers 1609 and 1670

Patterson (1966) and Matthai (1968) developed curves for estimating the magnitude of peak flows for frequencies between 1.1 and 50 years. The entire Missouri River basin peak flow predictions can be determined from a set of curves. The curves use a dimensionless frequency curve and a basin characteristic relation to predict peak discharge. The methods are similar to those of Furness (1955) and Beckman and Hutchison (1962). Figure 2.6 shows the hydrologic area numbering system in Nebraska.

2.2.4 Beckman Regression Equations, WRI 76-106

The first to use multiple regression techniques to predict magnitude and frequency was Beckman (1976). Equations for recurrence intervals of up to 100 years were developed based on selected basin characteristics. Nebraska was subdivided into five hydrologic regions based on regression techniques (Figure 2.7). Station data were analyzed using the Log-Pearson Type III distribution method, recommended by the WRC Bulletin 15. Consequently, these equations do not reflect the most current methods in Bulletin 17B. The updated bulletin provides revised procedures for the weighting of station skews, dealing with outliers, making station comparisons, and defining confidence limits. Flow peaks at 303 gaging stations with 13 or more years of record were used. Five sets of equations were developed for each region based on three basin characteristics. Significant basin characteristics were the contributing drainage area, slope, precipitation, and temperature. Standard errors of estimate ranged from 60 to 102 % in the western part of Nebraska to as low as 22 % in the Big Blue region. The equations provided by Beckman for the 2-, 10-, 50-, and 100-year return periods are listed by region in Table 2.1.

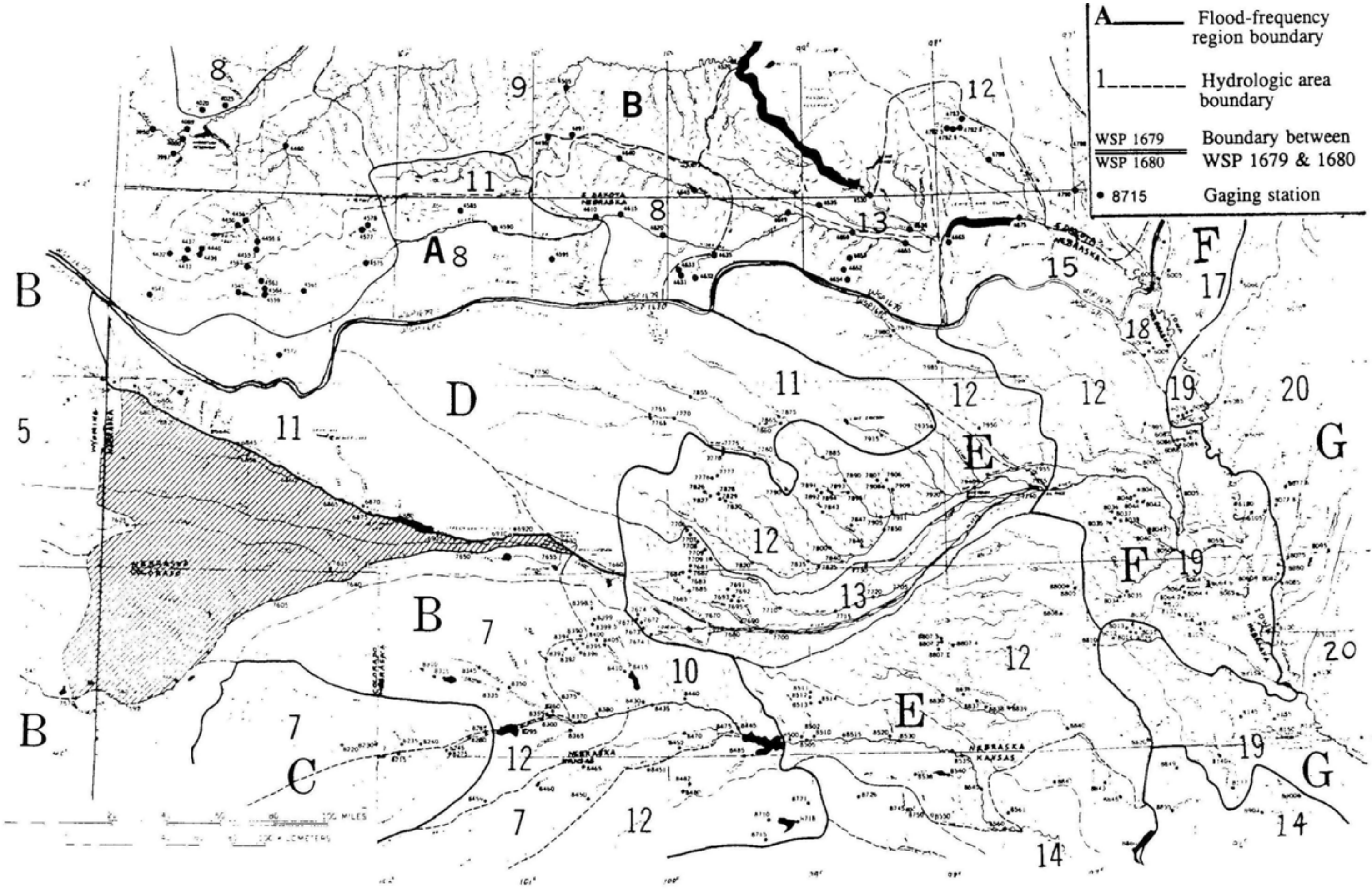


Figure 2.6: Nebraska hydrologic areas (WSP 1679 and WSP 1680).

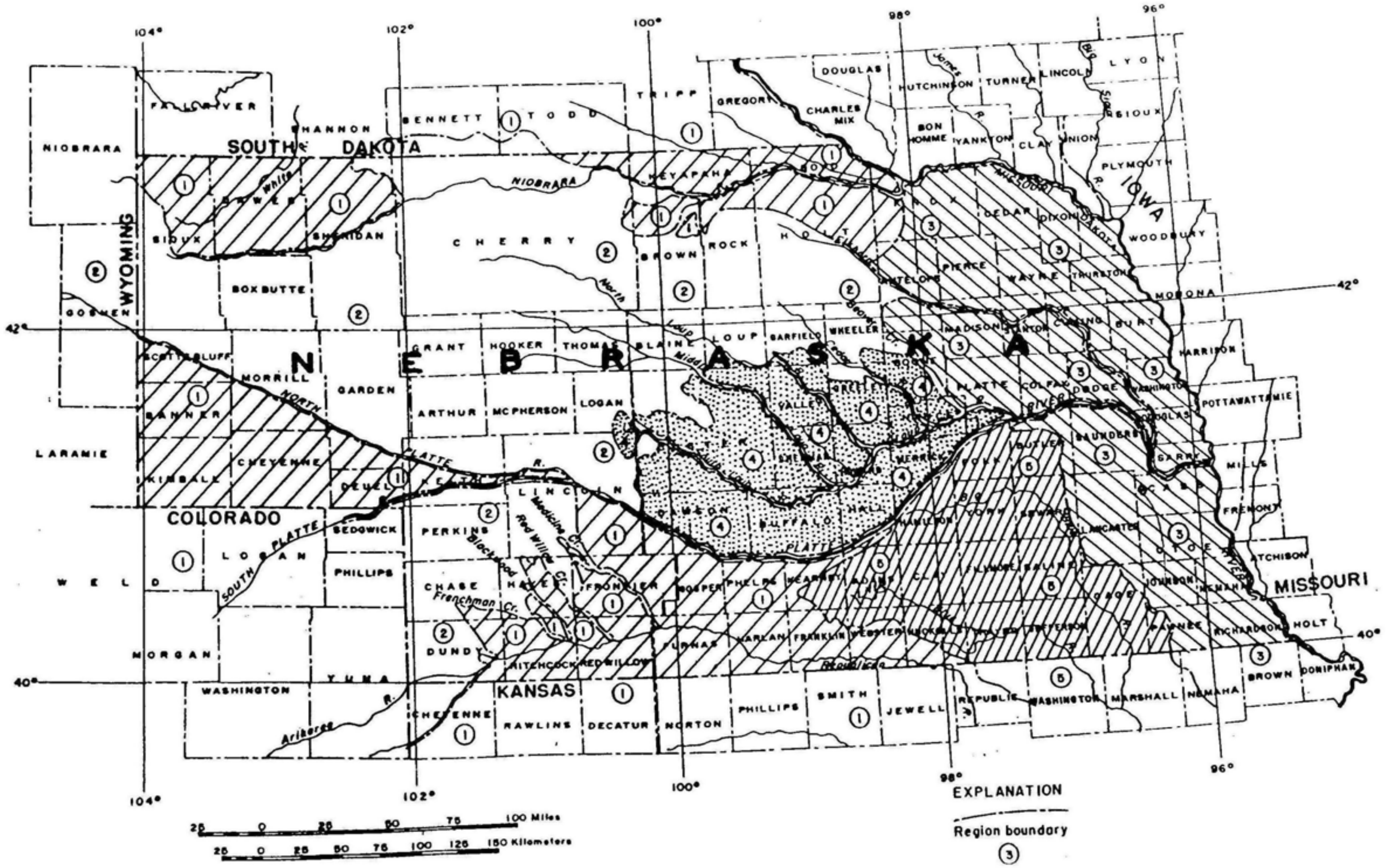


Figure 2.7: USGS hydrologic regions of Nebraska (Beckman, 1976)

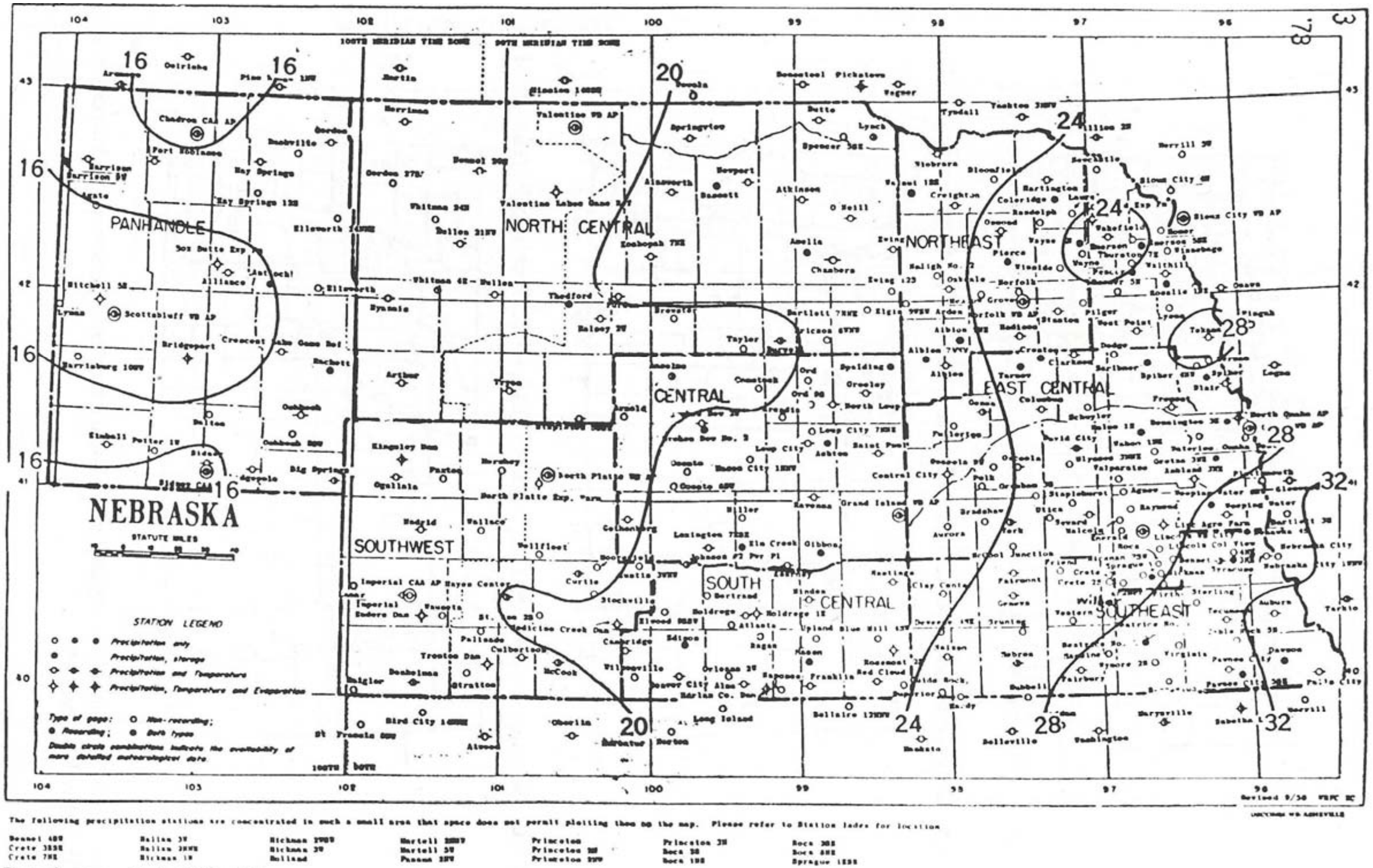
Table 2.1: The USGS regression equations in Nebraska for the 2-, 10-, 50-, and 100-year return period (Beckman, 1976).

Region 1	***SEE (%)***	***Region 4***	***SEE (%)***
$Q_2 = 1.6\ A_c^{0.997} (P-13)^{1.952} L^{-0.795}$	102	$Q_2 = 1774\ A_c^{1.226} (I_{24,50}-5)^{1.831} L^{-1.380}$	54
$Q_{10} = 67.0\ A_c^{0.737} (P-13)^{1.149} L^{-0.609}$	65	$Q_{10} = 8475\ A_c^{1.451} (I_{24,50}-5)^{1.491} L^{-1.783}$	45
$Q_{50} = 491.0\ A_c^{0.656} (P-13)^{0.742} L^{-0.544}$	85	$Q_{50} = 22301\ A_c^{1.650} (I_{24,50}-5)^{1.382} L^{-2.081}$	57
$Q_{100} = 997.0\ A_c^{0.624} (P-13)^{0.588} L^{-0.513}$	98	$Q_{100} = 31454\ A_c^{1.724} (I_{24,50}-5)^{1.365} L^{-2.184}$	65
Region 2		***Region 5***	
$Q_2 = 0.63\ A_c^{0.797} S^{0.427} (I_{24,50}-3)^{2.863}$	76	$Q_2 = 0.94\ A_c^{0.831} (T_1-11)^{1.606} S^{0.501}$	35
$Q_{10} = 0.49\ A_c^{0.839} S^{0.814} (I_{24,50}-3)^{3.320}$	60	$Q_{10} = 13.30\ A_c^{0.721} (T_1-11)^{1.114} S^{0.443}$	22
$Q_{50} = 0.51\ A_c^{0.864} S^{1.008} (I_{24,50}-3)^{3.632}$	75	$Q_{50} = 44.10\ A_c^{0.687} (T_1-11)^{0.845} S^{0.521}$	32
$Q_{100} = 0.55\ A_c^{0.872} S^{1.063} (I_{24,50}-3)^{3.731}$	84	$Q_{100} = 63.90\ A_c^{0.680} (T_1-11)^{0.741} S^{0.572}$	37
Region 3			
$Q_2 = 103\ A_c^{1.231} (T_3-37)^{0.798} L^{-1.230}$	51	A_c, contributing drainage area (mi^2); L, main stream length (mi); S, main stream slope (ft/mi); P, mean annual precipitation (in); $I_{25,50}$, maximum 24-hour 50-year rainfall (in); T_1, mean minimum January temperature (°F); T_3, normal daily maximum March temperature (°F).	
$Q_{10} = 412\ A_c^{1.026} (T_3-37)^{0.741} L^{-0.948}$	37		
$Q_{50} = 887\ A_c^{0.891} (T_3-37)^{0.703} L^{-0.745}$	46		
$Q_{100} = 1162\ A_c^{0.843} (T_3-37)^{0.686} L^{-0.671}$	52		

2.2.5 Updated USGS Equations

The original USGS study was completed in 1976 (Beckman, 1976) and included peak flow data through the 1972 water-year. Beckman's regression equations were updated to account for an additional 19-years of peak flow data (Cordes and Hotchkiss, 1993). The Log Pearson Type III distribution with the additional flow data will provide more accurate peak flow predictions. The basin characteristics used in the updated regression equations are:

A_c = Contributing drainage area (mi^2)
L = Length from station to basin divide along main channel (mile)
S = Slope, measured from the elevations at 0.10 and 0.85 of the channel length, divided by L (ft/mile)
P = Average annual precipitation (inches) (Figure 2.8)
$I_{24,\,2}$ = Rainfall intensity for a two-year, 24-hour event (in/hr) (Figure 2.9)
SN10 = Equivalent moisture content of snow (in) as of March 15 (Figure 2.10)
T_3 = Normal daily maximum March temperature (^{o}F) (Figure 2.11**)**



Based on period 1931-55

Figure 2.8: Mean annual precipitation, P (Cordes and Hotchkiss, 1993)

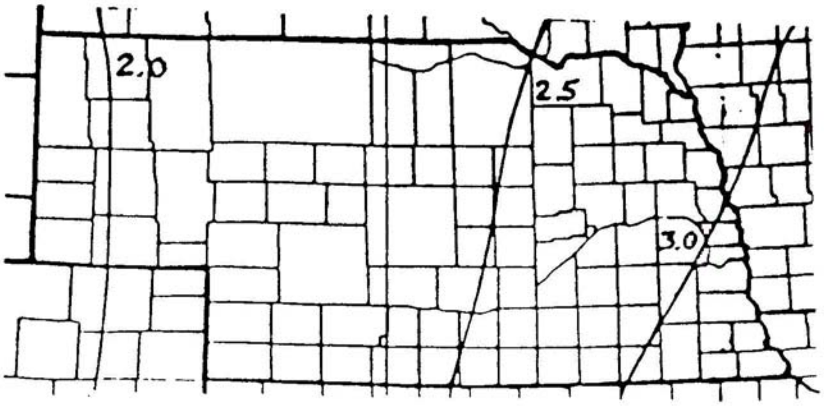


Figure 2.9: 2-year, 24-hour rainfall intensity, $I_{24,2}$ (Cordes and Hotchkiss, 1993)

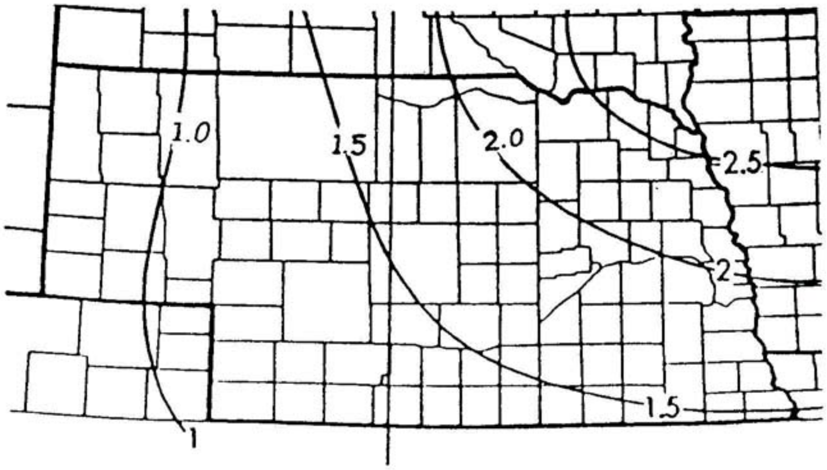


Figure 2.10: 10%-probability-equivalent moisture content of snow as of March 15, SN10 (Cordes and Hotchkiss, 1993)

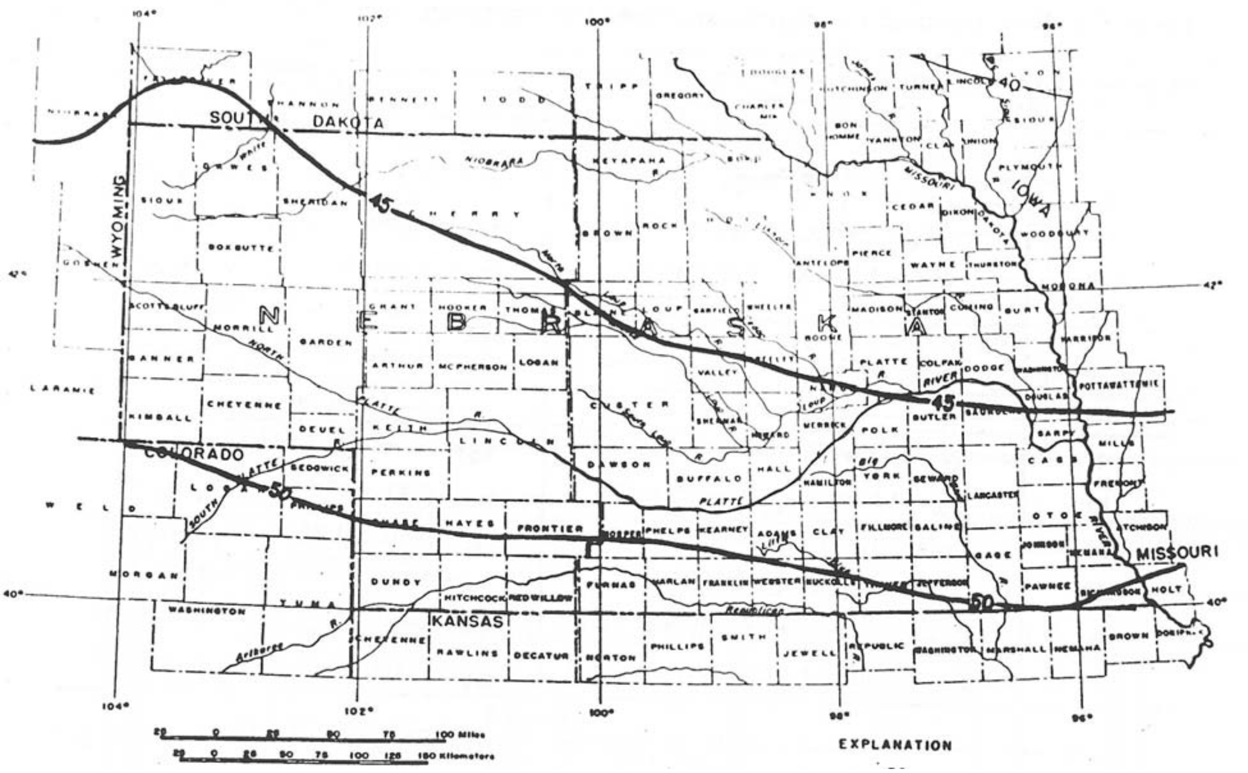


Figure 2.11: Normal daily March temperature, T_3 (Cordes and Hotchkiss, 1993)

After the collection of basin characteristics, the stations were divided into Beckmans five hydrologic regions (Figure 2.7). Six regression equations were developed for each region based on three basin characteristics. Significant basin characteristics were the contributing drainage area, main channel length, slope, precipitation, moisture content and temperature. The updated equations did not use some of the variables used in Beckmans study, because the original variables were no longer statistically significant. The updated USGS equations are shown in Table 2.2.

Table 2.2: Updated USGS regional regression equations for Nebraska (Cordes and Hotchkiss, 1993)

Return Period	**REGION 1**	**REGION 2**
2	$Q_2 = 1.965\ A_c^{0.493}\ (P-13)^{1.44}$	$Q_2 = 0.269\ A_c^{0.912}\ S^{0.967}\ SN10^{2.337}$
10	$Q_{10} = 211.7\ A_c^{0.324}\ (P-13)^{0.314}$	$Q_{10} = 0.109\ A_c^{0.9917}\ S^{1.653}\ SN10^{2.607}$
50	$Q_{50} = 6366\ A_c^{0.211}\ (P-13)^{-0.630}$	$Q_{50} = 0.0845\ A_c^{1.036}\ S^{2.005}\ SN10^{2.632}$
100	$Q_{100} = 23553\ A_c^{0.170}\ (P-13)^{-1.011}$	$Q_{100} = 0.0816\ A_c^{1.051}\ S^{2.119}\ SN10^{2.615}$
200	$Q_{200} = 82183\ A_c^{0.131}\ (P-13)^{-1.382}$	$Q_{200} = 0.0816\ A_c^{1.064}\ S^{2.216}\ SN10^{2.587}$
500	$Q_{500} = 400713\ A_c^{0.082}\ (P-13)^{-1.863}$	$Q_{500} = 0.0844\ A_c^{1.079}\ S^{2.326}\ SN10^{2.536}$
	REGION 3	**REGION 4**
2	$Q_2 = 7.57*10^{-10}\ A_c^{0.815}\ S^{0.599}\ P^{7.099}$	$Q_2 = 341.4\ A_c^{0.443}\ L^{0.126}\ (T_3-43)^{-2.062}$
10	$Q_{10} = 2.55*10^{-8}\ A_c^{0.722}\ S^{0.505}\ P^{6.657}$	$Q_{10} = 4741\ A_c^{0.914}\ L^{-0.783}\ (T_3-43)^{-1.960}$
50	$Q_{50} = 8.19*10^{-7}\ A_c^{0.688}\ S^{0.492}\ P^{5.908}$	$Q_{50} = 19516\ A_c^{1.285}\ L^{-1.411}\ (T_3-43)^{-1.903}$
100	$Q_{100} = 3.26*10^{-6}\ A_c^{0.681}\ S^{0.497}\ P^{5.581}$	$Q_{100} = 31008\ A_c^{1.433}\ L^{-1.648}\ (T_3-43)^{-1.876}$
200	$Q_{200} = 1.37*10^{-5}\ A_c^{0.677}\ S^{0.504}\ P^{5.226}$	$Q_{200} = 46677\ A_c^{1.573}\ L^{-1.871}\ (T_3-43)^{-1.850}$
500	$Q_{500} = 9.20*10^{-5}\ A_c^{0..673}\ S^{0.516}\ P^{4.740}$	$Q_{500} = 75811\ A_c^{1.752}\ L^{-2.148}\ (T_3-43)^{-1.819}$
	REGION 5	
2	$Q_2 = 0.00137\ A_c^{0.790}\ S^{0.777}\ I_{24,2}^{8.036}$	
10	$Q_{10} = 0.00126\ A_c^{0.687}\ S^{0.683}\ I_{24,2}^{10.037}$	
50	$Q_{50} = 0.00240\ A_c^{0.632}\ S^{0.640}\ I_{24,2}^{10.467}$	
100	$Q_{100} = 0.00335\ A_c^{0.615}\ S^{0.628}\ I_{24,2}^{10.491}$	
200	$Q_{200} = 0.00464\ A_c^{0.599}\ S^{0.618}\ I_{24,2}^{10.490}$	
500	$Q_{500} = 0.00755\ A_c^{0.581}\ S^{0.606}\ I_{24,2}^{10.393}$	

Problems were encountered in the development of the Region 1 equation set. Region 1 has the highest variability between stations and only has two statistically relevant basin characteristics. In addition, as found by Beckman, this region had the highest standard error. Region 5 had the lowest standard error, which was also consistent with Beckmans regression equations. The results showed that the updated USGS equations statistically improve peak flow prediction.

2.2.6 Soenksen Regression Equations, WRI 99-4032

The latest evaluation of peak-flow frequency relations in Nebraska was completed by Soenksen et al. (1999a). The objective of the report was to update Beckman's peak-flow relations, develop a new set of equations, and evaluate Nebraska's gaging station network. The relations between the peak-flow and return period for individual basins were developed following the guidelines of Bulletin 17B of the IACWD (Interagency Advisory Committee on Water Data, 1982). With the addition of new technology in GIS, previously undefined basins could be delineated. Previously some basin attributes were either too difficult or too timeconsuming to compute manually. Digital data has made rapid delineation of drainage basin characteristics possible.

As a part of this study, Provaznik (1997) investigated regional peak flow frequency using L-moments for possible improvements and efficiency over Bulletin 17B. Regions were created by geographic proximity, basin attribute, and the Region of Influence (ROI) approach. The ROI method reduced the heterogeneity, but did not create homogeneity in all regions. The statistics showed significant differences between the at-site estimates and Bulletin 17B estimates. The differences between estimates can be attributed to the treatment of outliers by moments and the different distributions.

Basin characteristics were quantified using the software *Basinsoft* written by Harvey and Eash (1996). *Basinsoft* utilizes ArcInfo to generate GIS data layers from digital cartographic data. The instructions and verification are given to quantify 27 selected morphometric characteristics.

Eight sets of regional regression equations were developed for seven regions in Nebraska. The generalized least squares procedure was used to relate basin characteristics to annual peak flows. The standard error of estimate (SEE) for the 100-year return period discharge ranged from 12 to 64 percent. Soenksen et al. (1999a) expanded Beckman’s regions to seven distinct regions (Figure 2.2).

The high-permeability region also had a composite analysis to account for large parameter variations in the sandhills. The composite analysis used half the amount of gaging stations and an additional basin characteristic in the regression equations. Equations for 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year return periods for each region are shown in Table 2.3a and b.

2.3 National Peak Flow Frequency

The USGS has developed a national peak flow frequency program to estimate the peak flows at ungaged locations. Each state has developed a set of regional regression equations. A state to state comparison is made between the regional equations for selected projects.

2.3.1 Colorado Regression Equations, WRI 99-4190

This report presents the regression equations and methods used to develop the magnitude and frequency of peak flows in Colorado by Vaill (1999). The regression equations are based on at least 10 or more years of stream flow records for 328 gaging stations. A generalized least-squares (GLS) regression was used to estimate the 2- through 500-year recurrence interval. Colorado was sub-divided into five hydrologic regions. The basin characteristics with the highest statistical significance were the drainage area, the mean annual precipitation, and the mean basin slope. The highest standard error of estimate (SEE) was found in the plains regions, which generally ranged from 200 to 300 percent. The lowest SEE ranged from 40 to 80 percent in the mountainous region. The method that was developed to determine peak discharge was dependent on whether the site was gaged, on a stream with a gaging station, or ungaged. Sites near a gaging station can be estimated by using a ratio of drainage area (Equation 2.5). This method is valid for drainage area ratios between 0.5 and 1.5.

Table 2.3a: USGS regression equations for the seven Nebraska regions at the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year return periods (WRI 99-4032). [†]

Big Blue River Region (32 stations with 10 or more years of record)	SEE (%)	**Northeastern Region** (40 stations with 15 or more years of record)	SEE (%)
$Q_{2} = 54\ CDA^{0.627}\ MSS^{0.425}\ TTP^{1.69}\ SD^{0.468}$	39	$Q_{2} = 132\ CDA^{0.676}\ SF^{-0.335}\ DF^{0.295}\ PLP^{-0.592}$	46
$Q_{5} = 160\ CDA^{0.580}\ MSS^{0.492}\ SF^{-0.220}\ SD^{0.533}\ TTP^{1.05}$	18	$Q_{5} = 395\ CDA^{0.652}\ SF^{-0.421}\ DF^{0.323}\ PLP^{-0.514}$	36
$Q_{10} = 267\ CDA^{0.546}\ MSS^{0.534}\ SF^{-0.264}\ SD^{0.511}\ TTP^{0.790}$	10	$Q_{10} = 715\ CDA^{0.633}\ SF^{-0.469}\ DF^{0.338}\ PLP^{-0.443}$	35
$Q_{25} = 463\ CDA^{0.500}\ MSS^{0.618}\ SF^{-0.360}\ SD^{0.631}$	10	$Q_{25} = 1{,}360\ CDA^{0.612}\ SF^{-0.518}\ DF^{0.356}\ PLP^{-0.352}$	36
$Q_{50} = 607\ CDA^{0.491}\ MSS^{0.638}\ SF^{-0.372}\ SD^{0.617}$	10	$Q_{50} = 2{,}070\ CDA^{0.597}\ SF^{-0.548}\ DF^{0.370}\ PLP^{-0.286}$	38
$Q_{100} = 764\ CDA^{0.483}\ MSS^{0.656}\ SF^{-0.382}\ SD^{0.601}$	12	$Q_{100} = 3{,}000\ CDA^{0.583}\ SF^{-0.573}\ DF^{0.384}\ PLP^{-0.223}$	40
$Q_{200} = 936\ CDA^{0.477}\ MSS^{0.672}\ SF^{-0.389}\ SD^{0.584}$	14	$Q_{200} = 5{,}240\ CDA^{0.562}\ SF^{-0.667}\ DF^{0.452}$	42
$Q_{500} = 1{,}190\ CDA^{0.469}\ MSS^{0.692}\ SF^{-0.396}\ SD^{0.557}$	17	$Q_{500} = 7{,}030\ CDA^{0.551}\ SF^{-0.655}\ DF^{0.440}$	45

Eastern Region (42 stations with 10 or more years of record)	SEE (%)	**Central and South-Central Region** (37 stations with 15 or more years of record)	SEE (%)
$Q_{2} = 5.7\ CDA^{0.558}\ BS^{0.655}\ PLP^{-0.470}$	46	$Q_{2} = 54.8\ CDA^{0.994}\ RR^{1.00}\ (TTP\text{-}2)^{4.24}\ SF^{-0.738}$	68
$Q_{5} = 21.1\ CDA^{0.533}\ BS^{0.551}\ PLP^{-0.528}$	30	$Q_{5} = 73.4\ CDA^{0.942}\ RR^{1.32}\ (TTP\text{-}2)^{3.98}\ SF^{-0.647}$	47
$Q_{10} = 42.1\ CDA^{0.519}\ BS^{0.495}\ PLP^{-0.537}$	25	$Q_{10} = 80.8\ CDA^{0.931}\ RR^{1.51}\ (TTP\text{-}2)^{3.92}\ SF^{-0.614}$	45
$Q_{25} = 90.2\ CDA^{0.504}\ BS^{0.433}\ PLP^{-0.520}$	24	$Q_{25} = 89.4\ CDA^{0.923}\ RR^{1.71}\ (TTP\text{-}2)^{3.88}\ SF^{-0.587}$	48
$Q_{50} = 151\ CDA^{0.494}\ BS^{0.390}\ PLP^{-0.498}$	25	$Q_{50} = 96.4\ CDA^{0.918}\ RR^{1.83}\ (TTP\text{-}2)^{3.84}\ SF^{-0.572}$	52
$Q_{100} = 242\ CDA^{0.485}\ BS^{0.349}\ PLP^{-0.474}$	27	$Q_{100} = 104\ CDA^{0.914}\ RR^{1.93}\ (TTP\text{-}2)^{3.83}\ SF^{-0.560}$	56
$Q_{200} = 377\ CDA^{0.476}\ BS^{0.310}\ PLP^{-0.450}$	29	$Q_{200} = 111\ CDA^{0.910}\ RR^{2.02}\ (TTP\text{-}2)^{3.81}\ SF^{-0.549}$	61
$Q_{500} = 650\ CDA^{0.465}\ BS^{0.260}\ PLP^{-0.417}$	32	$Q_{500} = 121\ CDA^{0.906}\ RR^{2.12}\ (TTP\text{-}2)^{3.80}\ SF^{-0.538}$	68

[AWC, available water capacity (in/in); BS, basin slope (ft/mile); CR, compactness ratio (dimensionless), CDA, contributing drainage area (mi^2); DF, drainage frequency, (number of first order streams per mile); MAP, mean annual precipitation (inches), MCS, main channel slope (ft/mile), MSS, maximum soil slope (percent); PLP, permeability of least permeable layer (in/hr); SD, stream density (mi/mi^2); RR, relative relief (ft/mile); SF, shape factor (dimensionless); TTP, 2-year, 24-hour precipitation (inches); Q, peak discharge (cfs); SEE, standard error of estimate]

[†] Note: according to a USGS errata sheet, in the first three equations of Table 2.3a (Q_2 – Q_{10} of the Big Blue River Region) TTP should be replaced with TTP-2 (Soenksen, 1999b).

Table 2.3b: USGS regression equations for the seven Nebraska regions at the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year return periods (WRI 99-4032).

Upper Republican River Region **(33 stations with 15 or more years of record)**	**SEE (%)**	**High Permeability Region** **(49 stations with 25 or more years of record)**	**SEE (%)**
$Q_2 = 1.97\ CDA^{0.545}\ MCS^{1.19}\ CR^{-0.735}$	52	$Q_2 = 0.066\ CDA^{0.750}\ (MAP-15)^{0.548}\ BS^{0.934}$	42
$Q_5 = 3.67\ CDA^{0.570}\ MCS^{1.32}\ CR^{-0.895}$	46	$Q_5 = 0.408\ CDA^{0.777}\ (MAP-15)^{0.525}\ BS^{0.653}$	41
$Q_{10} = 4.93\ CDA^{0.583}\ MCS^{1.39}\ CR^{-0.937}$	48	$Q_{10} = 8.76\ CDA^{0.736}\ (MAP-15)^{0.527}\ AWC^{0.835}\ BS^{0.539}$	42
$Q_{25} = 6.58\ CDA^{0.597}\ MCS^{1.46}\ CR^{-0.946}$	52	$Q_{25} = 14.8\ CDA^{0.773}\ (MAP-15)^{0.695}\ AWC^{1.17}\ MCS^{0.546}\ BS^{0.318}$	44
$Q_{50} = 7.84\ CDA^{0.606}\ MCS^{1.50}\ CR^{-0.931}$	55	$Q_{50} = 73.2\ CDA^{0.779}\ (MAP-15)^{0.756}\ AWC^{1.35}\ MCS^{0.766}$	46
$Q_{100} = 9.12\ CDA^{0.613}\ MCS^{1.54}\ CR^{-0.905}$	60	$Q_{100} = 119\ CDA^{0.777}\ (MAP-15)^{0.787}\ AWC^{1.56}\ MCS^{0.860}$	47
$Q_{200} = 10.4\ CDA^{0.619}\ MCS^{1.57}\ CR^{-0.868}$	64	$Q_{200} = 184\ CDA^{0.774}\ (MAP-15)^{0.816}\ AWC^{1.74}\ MCS^{0.942}$	49
$Q_{500} = 12.2\ CDA^{0.626}\ MCS^{1.61}\ CR^{-0.809}$	71	$Q_{500} = 313\ CDA^{0.769}\ (MAP-15)^{0.850}\ AWC^{1.94}\ MCS^{1.04}$	53

Northern and Western Region **(34 stations with 15 or more years of record)**	**SEE (%)**	**High Permeability Region - Composite Analysis** **(23 stations with 20 or more years of record)**	**SEE (%)**
$Q_2 = 0.176\ CDA^{0.762}\ RR^{0.878}\ (MAP-12)^{0.929}\ PLP^{-0.357}$	126	$Q_2 = 0.127\ CDA^{0.684}\ (MAP-15)^{0.715}\ DF^{0.456}\ BS^{0.968}$	35
$Q_5 = 0.686\ CDA^{0.642}\ RR^{0.932}\ (MAP-12)^{1.05}\ PLP^{-0.360}$	62	$Q_5 = 1.09\ CDA^{0.774}\ (MAP-15)^{0.590}\ DF^{0.454}\ BS^{0.576}$	42
$Q_{10} = 1.69\ CDA^{0.577}\ RR^{0.892}\ (MAP-12)^{1.08}\ PLP^{-0.337}$	55	$Q_{10} = 21.8\ CDA^{0.744}\ (MAP-15)^{0.626}\ AWC^{1.17}\ DF^{0.399}\ BS^{0.602}$	44
$Q_{25} = 5.06\ CDA^{0.508}\ RR^{0.802}\ (MAP-12)^{1.07}\ PLP^{-0.302}$	55	$Q_{25} = 159\ CDA^{0.805}\ (MAP-15)^{0.718}\ AWC^{1.40}\ DF^{0.637}\ MCS^{0.773}$	47
$Q_{50} = 10.7\ CDA^{0.464}\ RR^{0.731}\ (MAP-12)^{1.06}\ PLP^{-0.272}$	59	$Q_{50} = 368\ CDA^{0.817}\ (MAP-15)^{0.730}\ AWC^{1.76}\ DF^{0.637}\ MCS^{0.864}$	49
$Q_{100} = 35.2\ CDA^{0.213}\ BS^{0.589}\ (MAP-12)^{0.643}$	64	$Q_{100} = 776\ CDA^{0.828}\ (MAP-15)^{0.741}\ AWC^{2.07}\ DF^{0.641}\ MCS^{0.941}$	51
$Q_{200} = 37.4\ CDA^{0.192}\ BS^{0.629}\ (MAP-12)^{0.711}$	65	$Q_{200} = 1520\ CDA^{0.838}\ (MAP-15)^{0.752}\ AWC^{2.35}\ DF^{0.645}\ MCS^{1.01}$	55
$Q_{500} = 41.6\ CDA^{0.168}\ BS^{0.669}\ (MAP-12)^{0.786}$	70	$Q_{500} = 3390\ CDA^{0.851}\ (MAP-15)^{0.767}\ AWC^{2.67}\ DF^{0.654}\ MCS^{1.09}$	61

[AWC, available water capacity (in/in); BS, basin slope (ft/mile); CR, compactness ratio (dimensionless), CDA, contributing drainage area (mi^2); DF, drainage frequency, (number of first order streams per mile); MAP, mean annual precipitation (inches), MCS, main channel slope (ft/mile), MSS, maximum soil slope (percent); PLP, permeability of least permeable layer (in/hr); SD, stream density (mi/mi^2); RR, relative relief (ft/mile); SF, shape factor (dimensionless); TTP, 2-year, 24-hour precipitation (inches); Q, peak discharge (cfs); SEE, standard error of estimate]

$$Q_{T(u)} = Q_{T(g)} \left(\frac{A_u}{A_g} \right)^x \tag{2.5}$$

Where: $Q_{T(u)}$ = peak discharge at the ungaged site (cfs)
$Q_{T(g)}$ = weighted peak discharge at gaged site (cfs)
A_u = drainage area at ungaged site (mi^2)
A_g = drainage area at gaged site (mi^2)
x = average regional peak flow exponent

The relation is valid if the given locations have similar basin and climatic characteristics and are in the same region. Previous research by Livingston and Minges (1987) provided equations for estimating peak flow characteristics for rural watersheds with drainage areas of less than 20 mi^2 in the plains regions of eastern Colorado. Standard errors ranged from 36 to 57 percent and significant characteristics were the effective drainage area, the relief factor, and the 24-hour, 100-year rainfall intensity.

2.3.2 South Dakota Regression Equations, WRI 98-4055

Peak-flow equations were developed for recurrence intervals of 2- through 500-years for seven hydrologic regions in South Dakota (Sando, 1998). The equations are applicable to natural streams with drainage areas of less than 1,000 mi^2. The generalized least-squares (GLS) regression analysis was based on 197 streamflow gaging stations that had 10 or more years of record. The purpose of this project was to create Log-Pearson Type III distributions and to create a relationship between peak flows and basin characteristics. Basin and climatic characteristics in the final regression equations include contributing drainage area, main channel slope, and the precipitation intensity index. For the 100-year recurrence interval, the SEE of the peak-flow equations ranged from 22 to 110 percent. Generally, peak flows in the Black Hills regions are highly variable and difficult to regionalize due to the fractured limestone outcrops. Previous research by Becker (1980) used data from 115 stations with drainage areas ranging from 0.05 to 100 mi^2. Statistically relevant parameters were the area, main channel slope, and soil infiltration index.

2.3.3 Missouri Regression Equations, WRI 95-4231

Estimation of unregulated stream discharges in rural Missouri has been researched by Alexander and Wilson (1995). Generalized least-squares (GLS) regression was applied to return period discharges for three hydrologic regions. Ordinary least-squares (OLS) regression was used to demonstrate that basin area and main channel slope were statistically relevant. The basin and climatic significance was based on a 95-percent confidence level, where the standard error was minimized. Missouri was subdivided into three regions, which included 278 gaging stations. Errors for the GLS regression ranged from 30 to 49 percent. The study provided techniques for estimating peak flow discharges at unregulated streams in Missouri. Frequency relations followed Bulletin 17B, while basin information was compiled from a combination of 1:24,000 scale topographic maps, and 1:100,000 and 1:250,000 digital data.

2.3.4 Iowa Regression Equations, WRI 00-4233

Techniques for estimating peak flow frequency discharges for streams in Iowa were developed by Eash (2000). Three hydrologic regions were developed using generalized least squares regression. GLS regression was used to weight regression variables to improve the predictive accuracies of the peak flow frequency equations. Recurrence intervals of the 2- to 500-year discharge were used in the regression analysis. Gaging stations with at least 10 years of streamflow record in Iowa and out-of-state stations with 25 years of record were used. The predictive accuracy of each equation was based on root-mean-square error calculations. The multi-variable equations were developed using basin area, main channel slope, and the Des Moines Lobe (DML) landform. The DML variable is the ratio of basin area within the Des Moines Lobe landform to total area of the basin. One-variable equations were developed for quick calculations, when determining the peak flow frequency discharge. The final regression analysis included 241 gaging stations. The standard error of prediction (SEP) for one-variable equations ranged from 34 to 45 percent. Only two regions had multi-variable equations, and had SEP's ranging from 31 to 42 percent.

2.3.5 Kansas Regression Equations, WRI 00-4079

Peak streamflows were estimated for return periods of the 2- to 200-year peak discharge using GLS regression (Rasmussen and Perry, 2000). The regression equations are based on at least 10 years of stream flow records at 253 gaging stations in Kansas. Instead of sub-dividing Kansas into hydrologic regions, it was grouped according to drainage area. The best results were obtained when the contributing drainage area ranged from 30 to 9,100 mi^2. Compared to all stations, a reduction of 12 to 20 percent in the standard error of prediction (SEP) was achieved. The SEP for basins ranging from 0.17 to 30 mi^2 had equal to or slightly greater SEP than equations developed using all the stations. Significant basin characteristics were the contributing drainage area, mean annual precipitation, average soil permeability, and slope of the main channel. Overall the SEP of Kansas's regression equations ranged from 31 to 62 percent. The general climate of Kansas varies from semiarid in the western two-thirds to a more humid climate to the east. Peak flows on small streams in Kansas are usually the result of high intensity thunderstorms.

2.3.6 Ohio Regression Equations, WRI 86-4354

Multiple regression equations to estimate low flow characteristics were developed for ungaged streams in Ohio by Kortun and Schwartz (1986). The equations include basin area, main channel slope, forested area, and average annual precipitation. Data from 132 stations at recurrence intervals of 2- and 10-years were used. Ohio was divided into five regions, with different regression parameters. The general form of the multiple-regression model used was:

$$Y = CX_1^{b_1} X_2^{b_2} ... X_n^{b_n} - 0.1 \tag{2.6}$$

The constant 0.1 was added to the equation to include flows equal to zero. Where, *C* is the regression constant, X_1, X_2,…X_n are the selected basin characteristics, and b_1, b_2,…b_n are the regression coefficients to the *n* number of independent variables. SAS was used in the multiple regression analysis. It helped in determining an optimum set of regression variables to be used in the equations. Regional boundaries were adjusted to improve the fit to the data, determined from the standard error and coefficient of determination (R^2). Standard error of estimate is a measure of the average variation of the observed values of the dependent variable from the regression line, which is a crude indicator of the level of accuracy. The coefficient of determination is a measure of the effectiveness of the independent variables in explaining observed variations. Tests were also done for constant residual variance, a sensitivity analysis, and co-linearity which can cause round off errors in the regression equations. It was concluded that the most accurate estimate of low-flows should use long term records from a gaging station near the site. Only if there is no available information should one use the developed low-flow regression equations.

2.3.7 Texas Regional Equations, WRI 96-4307

Asquith and Slade (1996, 1999) investigated techniques to develop regional regression equations of peak streamflow frequencies at ungaged sites in Texas. The peak streamflow data for Texas was subdivided into 11 regions that resulted in 16 sets of equations from the 559 gaging stations. The basis for this project was the comparison of the statistical relationship between peak streamflow frequency and basin characteristics. The equations were developed using weighted least squares (WLS) regression.

In WLS regression each data point can be given a different weight, which is dependent on the period of record. In some regions equations were developed for drainage areas less than 32 mi^2. A break line was developed by visual inspection from the 100-year peak discharge and contributing drainage area. A region of overlap was used to increase the accuracy of the equations. The purpose of the report was to update and present regional equations that accurately predict peak flow frequency for natural basins. A natural basin is defined as having less than 10 percent of its drainage area controlled by reservoirs or man made structures that affect flow. A Log-Pearson Type III distribution was fit to collected and historical peak stream flow data. Historical information is critical for evaluating peak stream flow frequency estimates for larger recurrence intervals. The results showed that contributing drainage area had the highest statistical significance, while stream slope proved to be second.

2.4 Geographic Information Systems

The application of Geographic Information Systems (GIS) datasets to water resources has provided a consistent method for watershed and stream network delineation. Hydrologists may use datasets to assess water quality, determine water supply, prevent flooding, and manage water resources. The following sections discuss relevant datasets, resolution effects, and hydrologic software and how they apply to peak flow predictions.

2.4.1 Source Data Sets

The United States Geological Society (USGS) has developed digital cartographic and geographic data as part of the National Mapping Program. Available digital products include Digital Elevation Models (DEMs), digital land cover data, and digital line graphs (DLG’s). Advantages of digital data include the coverage of large areas with reasonable resolution, quick and repeatable techniques, and the flexibility to address a variety of problems (Loveland and Ramey, 1986). Previous research used 1:250,000-scale DEMs, but new 1:24,000-scale DEMs have since become available.

Characteristics of the 1:250,000-scale DEM data are that they consist of geographic coordinates, have regularly spaced arrays of elevations every 3 arc-seconds, include 1,201 pixels, and coverages consist of a 1 degree blocks. They are created by interpolating elevations digitized from topographic maps. Three arc-seconds is approximately 90 meters on the north-south axis but is a variable distance on the east-west axis. The accuracy of the 1:250,000-scale data is dependent on the scale of the source material used to create it. Contour intervals change depending on the terrain, flat regions use 50-foot intervals while steep terrains use 200-foot intervals.

The 7.5-minute quadrangles are derived from existing contour maps, manual profiling from stereomodels and from digitizing using orthophoto equipment. A majority of 7.5-minute DEMs are created from orthophoto equipment with elevation values spaced every 30 meters. The manual scanning and digitizing of photographs are taken from photos at an altitude of 40,000 ft. The DEM data is a regularly spaced array, referenced in the Universal Transverse Mercator (UTM) coordinate system. The accuracy of a 7.5-minute DEM depends on the aerial photographs or contours on the 7.5-minute topographic map. The overall accuracy is improved significantly compared to the 1:250,000-scale DEMs (Elassal and Caruso, 1983).

2.4.2 Resolution Effects

Digital representations of topographic surfaces are considered to be a mathematically continuous surface. There are high standards for topographic maps because they can be compared with actual surfaces and are considered to be stable over time. Large scale digital representations of topographic maps do a good job of representing areas of high relief. But, when relative relief is small, digital representations do a poor job of representing the topography.

According to Carter (1998), DEMs are preferred to digital terrain models (DTM) because they contain only elevation data, while DTM's include landscape attributes. It is difficult to measure the differences between the digital topography accuracy and real terrain. But it was found that at least 90% of elevations determined from continuous contours were within one-half the contour interval. To eliminate errors in the digitizing process the land surface should be surveyed. It is important to select points that are along ridgelines, stream channels, and valleys. If a highly detailed DEM database is needed, the precision will be expensive. Depending on the purpose of the digital model, one should consider the resolution, accuracy and precision of the data.

A study was conducted that compared the resolutions of digital elevations models to hydrologic parameters of peak discharge. Moglen and Hartman (2001) compared DEMs with cell sizes of 12 ft, 36 ft, 60 ft, 96 ft, 30 m, and 90 m. The drainage areas, flow length, relief, slope, and peak discharges were examined at each scale. The highest resolution of 12-foot grid cells was used as a reference value compared to the other results. Peak discharge was estimated using the Natural Resource Conservation Service TR-55 model from a given precipitation depth and return period event.

The results showed that the relative error in drainage area calculations decreases as the basin area increases. But, for areas of less than 5 km^2 (2 mi^2) relative errors range from 10 to 40

% for the 30 and 90 meter DEMs. Flow length measurements showed a linear bias, with higher resolution data having the longest flow paths. This happens because coarser resolution data can not represent small scale meandering of the stream channel. Slope was also found to be effected by resolution differences. The watershed relief and average slope was found to be smaller using lower resolution data. It was concluded that the coarser resolution data systematically overestimated the peak discharge in the NRCS model. When using existing regression equations, hydrologic engineers should expect smaller peak discharge predictions as higher resolution DEM data become available.

The effect of DEM scale on the accuracy of hydrologic prediction was evaluated from three grids with a basin area of 7.2 km^2 (2.8 mi^2). The reference DEM was created from low altitude aerial photography and compiled into a 30 meter grid. The second grid was created from NASA's images from space using Spaceborne Imaging Radar (SIR-C). Finally the third grid used was a 7.5-minute DEM, representing the current product available from the USGS. A statistical analysis was used to examine differences in the watershed area, point elevations, and topographic parameters. The analysis found that the USGS DEM had systematic errors associated with the processes by which it was created. The hydrologic predictions of the USGS DEM typically gave 10 % reductions in the peak runoffs. It was found that vertical accuracy does affect hydrological modeling. Grids that were directly derived reduced the amount of spatial clutter of the data (Kenward et al., 2000).

A comparison of drainage networks from DEMs were evaluated for a wide range of areas. A total of 20 basins ranging from 150 to 1000 km^2 (58 to 386 mi^2) were delineated from both 1:250,000 and 1:24,000 DEMs in West Virginia. A commercial GIS package was used to derive the drainage networks. A comparison was made between the two scales, with a constant stream density. The stream density was controlled by the total stream length at each scale. They found that the sensitivity of extracted basin parameters to grid size of the DEM varied from parameter to parameter. The stream order frequency analysis showed that with increasing stream order the difference between scales also increased. When stream order was increased, scatter increased if the lower resolution DEM was used. Statistically the errors between the two scales were reduced with an increase in terrain complexity. Deficiencies in 1:250,000-scale networks are partially due to the spatial and vertical resolution. Also, the basin size doesn't affect the accuracy of the extracted drainage parameters (Wang and Yin, 1998).

Wiche (1992) also examined the accuracy of DEMs at different scales. The first DEM was created from digitized USGS topographic maps with 5- to 10-foot contour intervals. The second DEM was extracted from aerial photographs at an elevation of 4,800 feet. The photos were encoded and a two-dimensional vector file was created with 2-foot contours. Five streams with areas ranging from 2.62 to 10.2 mi^2 were examined on the James River basin in North Dakota. It was found that differentiating between contributing and non-contributing drainage areas was difficult. Due to errors in resolution within the DEMs it was difficult to distinguish between natural and programmatic depressions.

2.4.3 Hydrologic Software

A program was written in FORTRAN to extract topographic structure from DEMs (Jensen and Domingue, 1988). The algorithms traditionally include raster processing systems using neighborhood operations. By defining cells relative to their neighboring cells one can calculate the slope, aspect and relief. To delineate watershed networks, a procedure was written to create three general grid datasets. The first step is to create a depressionless DEM in which cells that are sinks are filled to match neighboring cells. The second step is to assign a flow direction to each grid cell, in one of eight directions. The last step is to create a flow accumulation data set in which each cell is assigned a weight corresponding to the number of cells that flow into it. The resulting datasets can be used to delineate watershed boundaries and to define stream channels in raster format. Computer generated watersheds had areas that were within 97% of manually delineated watersheds. Visual comparisons of manually delineated and computer generated stream networks show that main channels are identical. Jenson concluded that the creation of a program that will derive morphologic information for large numbers of watersheds would be useful.

The standard for geographic information systems (GIS) software packages is ArcInfo. It is a complete GIS mapping and analysis system when used with the Spatial Analyst package. Spatial Analyst is integrated into ArcGIS and allows surface, terrain and algebraic analysis. Terrain analysis tools can model slope, hillshade, watershed delineation, contour generation and viewshed. Algebraic functions can reclassify values, assign weighted values to grids, and sum grid values within polygons and multiple grids. Cell based raster datasets, or DEMs are well matched for geographic systems. ArcGIS Spatial Analyst supports hydrologic modeling features used in water resources.

ArcGIS allows users to create and analyze cell based maps and integrate raster data with vector data sources. Another application used for coverage processing and analysis functions within ArcInfo is ArcToolbox. It has over 150 geoprocessing functions used for data conversion, map management, overlay analysis, and map projections. Most of the tools are used to manipulate ArcInfo coverages, which are the preferred GIS data type. ArcInfo Workstation has a standard user interface which includes basic geoprocessing functionality (Environmental Systems Research Institute, 2002).

2.4.4 GIS Applications

The application of digital data in GIS has revolutionized the construction of hydrologic data structures in water resources. Large scale hydrologic modeling and analysis can now be done with increased accuracy and speed. The identification of basins and subbasin boundaries can be described by GIS data layers derived from DEMs. A connection was made between the stream networks and subbasin data layers by Hellweger and Maidment (1999). A conceptual model of large watersheds was generated by defining the subbasins, stream reaches, reservoirs, junctions, diversions, sources, and sinks. The hydrologic modeling program used was HEC-HMS, developed by the U.S. Army Corps of Engineers Hydrologic Engineering Center. Currently the identification of U.S. basins is done with the USGS Hydrologic Unit System (HUC). This divides the country into 21 major regions composed of 222 subregions. With the use of GIS a system of identifying North American rivers basins was proposed (Verdin, 1999). A system of delineation and codification of basins on the basis of topography was constructed, and a symbolic coverage showing where the elements are located and how they are connected was created. The model is applicable to a large range of watershed sizes, suggesting that hundreds of flow elements can be connected.

2.4.5 Basin Characteristic Extraction

The process of extracting hydrologic information from digital data has been thoroughly examined. To be useful, the derived networks have to be extracted at the correct length scale and drainage density. The method given by Tarboton et al. (1991), extracts the highest drainage density network corresponding to traditional scaling laws. Threshold values were examined for 21 DEMs with varying scales throughout the U.S. A combination of the constant drop analysis and slope-area scaling showed breaks in the scaling process, which were used to define a threshold value. They provided a successful technique used for estimating drainage density. The result gave derived stream networks at varying scales similar to traditional digitized topographic maps.

An automated method was created to quantify physical basin characteristics from DEMs by Majure and Eash (1991). The method includes the combination of two existing software packages. The first incorporates FORTRAN programming, which produces three ArcInfo coverages that represent the drainage basin. The derived coverages include the stream network, elevation contours and the drainage basin (Martz and Garbrecht, 1993). The second software package uses Arc Macro Language (AML) and INFO programs to quantify basin characteristics. Eleven actual basin characteristics are directly measured, but 27 parameters are calculated. To verify the accuracy of the automated method results were compared to manual and digitized measurements for three watersheds. From preliminary comparisons it was concluded that the automated method produced reliable results. The approximate amount of time required to define the basin characteristics was also examined. The automated method required 6 hours of processing time, while manual and digitizing methods took 16 and 13 hours respectively.

Another procedure was developed to quantify drainage basin characteristics with varying scales (Eash, 1994). The Basin Characteristic System (BCS) uses digitized maps, digital line graph (DLG) data and DEMs. Software was also developed to assign attributes to specific features and quantify 24 basin characteristics. Recent developments in cartographic data have improved the processing of digital data using GIS. The purpose of this project was to describe the BCS process used in peak flow estimation studies. Due to edge matching problems encountered with 1:250,000-scale DEMs, both 1:250,000 topographic maps and 1:100,000 DLG hydrography data were used.

Basin area was manually digitized into GIS from topographic maps. Stream networks were extracted from DLG data using GIS software. The DEMs were used to create elevation contours with at least five contours per basin. The accuracy was quantified by comparing the results to manual measurements from topographic maps. Measurements of the main channel slope, basin slope, and basin relief produced the largest errors. The errors are due to the large scale used in the DEM, which is data-scale dependent. Other morphometric basin characteristics were affected by the scale of cartographic data used in the measurement. Basin comparisons of BCS calculations may be unreliable if the same scales are not used when creating the drainage divides, stream networks and elevation contours.

The use of commercial software to delineate watersheds using USGS DEMs has become common practice. Brown et al. (2000) described an automated procedure to delineate basins using ArcInfo GRID software. The development of watershed boundaries from digital sources removes the subjective nature of defining divides. The final product optimizes the efficiency of the computer with the judgment of a hydrologist to produce high quality delineation. A detailed explanation of the grid functions are given, when user input is needed. After careful review of the computer generated boundaries, improvements were made to correct for man-made features that were not shown in the DEM. The ArcInfo hydrologic software combines the ideas of previous software written by Majure and Eash (1991), Martz and Garbrecht (1993), Eash (1994) and others.

2.4.6 Supplemental Information

Two excellent books explaining GIS in water resources are available from the ESRI Press. The first book by Djokic and Maidment (2000) explains the *Hydrologic and Hydraulic Modeling Support in GIS*. It describes the importance and accuracy of DEMs in water resources modeling. It also describes the link between GIS and hydraulic modeling software. *Arc Hydro GIS for Water Resources* by Maidment (2002) explains the process of extracting information about river networks, watersheds, and water bodies. The Arc Hydro data model incorporates information about streams, gaging stations, drainage basins, hydrography, channels, surface elevation, rainfall, and aerial photography. The process of integrating all these data layers is the reason GIS has become a powerful tool in water resources.

2.5 Log Pearson Type III Distribution

The Log Pearson Type III (LP3) distribution is commonly used in peak flow analysis. Bulletin 17B contains guidelines for the development of peak-flow frequency relations as recommended by the IACWD (Interagency Advisory Committee on Water Data, 1982). The method of moments is used to determine the statistical parameters of the distribution from station data. The three parameters included in the distribution are the mean, standard deviation, and the skew coefficient.

$$\overline{X} = \frac{\sum X}{N} \qquad \textbf{(2.7)}$$

$$S = \left[\frac{\sum (X - \overline{X})^2}{(N-1)} \right] \qquad \textbf{(2.8)}$$

$$G = \frac{N \sum (X - X)^3}{(N-1)(N-2)S^3} \qquad \textbf{(2.9)}$$

Where: X = logarithm of annual peak flows
N = number of items in data set
$\overline{X}$ = mean logarithm
S = standard deviation of logarithms
G = skew coefficient of logarithms

The skew coefficient (station skew) is sensitive to extreme events, which makes it difficult to obtain an accurate skew for small samples. The accuracy of the skew coefficient can be improved by weighting the station skew from nearby sites. The generalized skew can be obtained by a generalized skew map, a skew equation, or the mean of the station skew values.

$$G_w = \frac{MSE_{\overline{G}}(G) + MSE_G(\overline{G})}{MSE_{\overline{G}} + MSE_G} \qquad \textbf{(2.10)}$$

Where: G_w = weighted skew coefficient
G = station skew
$\overline{G}$ = generalized skew
MSE_G = mean-square error of generalized skew
$MSE_{\overline{G}}$ = mean-square error of station skew

The distribution is then fitted to:

$$Log(Q) = \overline{X} + KS \qquad \textbf{(2.11)}$$

Where: Q = discharge
$\overline{X}$ = mean logarithm of peak flow peaks
K = frequency factor based on skew coefficient and return period
S = standard deviation of logarithms

Bulletin 17B is the recommended method to determine peak-flow frequency distributions. It is assumed that peak flows altered by reservoir regulation, or the possibility of an unusual event are not covered by the LP3 distribution as described in Bulletin 17B. Potential errors can arise from the randomness of events, land cover changes, and the reliability of the flow estimates.

2.6 Multiple Regression Model

The most commonly used relation between flow statistics and the watershed characteristics is the power-form function. It is based upon the assumption that the model can be linearized by a logarithmic transformation. The regression model used in regional peak flow frequency analysis is:

$$Q_T = \alpha_o A_1^{\alpha_1} A_2^{\alpha_2} \ldots A_n^{\alpha_n} \qquad \textbf{(2.12)}$$

Where, Q_T is the return period discharge, $\alpha_o, \alpha_1 \ldots, \alpha_n$ are the estimated model parameters, and $A_1, A_2, \ldots A_n$ are the watershed characteristics. Riggs (1973) provides some background information on this technique. The return period discharges estimated from the LP3 distribution are used as the dependent variables. The morphometric, soils and precipitation characteristics of the basin are the independent variables.

Pandey and Nguyen (1999) assessed the performance of regression models when estimating peak flows at ungaged locations. The performance was based on the accuracy of the predicted quantities. Non-linear techniques provided much better estimates than linear models at ungaged sites. Linear models have a tendency to under-predict the peak flow approximations and are more biased. The regression equation can then be applied to ungaged basins, using the appropriate basin characteristics.

2.7 Least Squares Regression

In Ordinary least-squares regression (OLS), the parameters are determined such that the squared sum of errors between observed and predicted peak flows are minimized. The OLS method gives unbiased and minimum variance estimates of parameters provided they are normally distributed (Draper and Smith, 1981).

Generally, discharge data used in regional analysis come from stations having varying conditions and unequal lengths. The data becomes heteroscedastic from the variations in conditions and the length of record, making some flow estimates less reliable (Tasker, 1982). Heteroscedasticity is when there are large differences in flow record lengths (e.g., one station may have twenty five years of record while another may only have fifteen) or the flows are cross-correlated. The problem caused by heteroscedasticity can be overcome by scaling or weighting the observed flow data when estimating regression parameters. Weighted least-squares (WLS) regression accounts for the differences in record length of the annual peak stream flow between sites. The WLS regression minimizes the squared sum of the weighted residuals, instead of the residuals.

Stedinger and Tasker (1986) used generalized least squares (GLS) regression in regional hydrologic analysis to account for heteroscedasticity and inter-site correlations. From Monte-Carlo experiments they demonstrated that, when record lengths vary widely and flows are cross-correlated, the GLS regression provides better estimates of the regressed parameters. When compared to WLS and OLS, GLS also produces less biased estimators of the variance from the residuals.

2.8 Standard Error of Estimate

The standard error of estimate (SEE) is used to compare LP3 discharges with the regression estimates. The logarithmic transformation of variables is useful in hydrology problems (Tasker, 1978). SEE is based on model error and will only change when the regression equation is changed. The SEE expressed in logarithmic units is:

$$SEE = \sqrt{\frac{\sum\left(Log(Q_{LP3}) - Log(Q_{REG})\right)^2}{N}} \qquad \textbf{(2.13)}$$

Where: SEE = standard error of estimate in log units
Q_{LP3} = Log Pearson Type III discharge
Q_{REG} = regression equation discharge
N = number of gaging stations

To express SEE in percent, a natural log conversion is used:

$$SEE(\%) = \left[\left(e^{\left[\ln(10)\bullet SEE\right]^2}\right) - 1\right]^{1/2} \qquad \textbf{(2.14)}$$

The standard error reported in log units or percent is usually followed by a statement that two-thirds of the observations are within one standard error of the regression equation. In this project the SEE of estimate was used to compare the accuracy of the regression equations.

3. TOOLS FOR DEVELOPING AND USING REGRESSION EQUATIONS

The methods and procedures used to develop the regional regression equations are explained in this chapter. In order to use the equations to estimate streamflow from a watershed, the necessary parameters will need to be calculated using the methods described. The first section of this chapter discusses the datasets used to develop the basin characteristics. A combination of Digital Elevation Models (DEM), the State Soil Geographic Data Base (STATSGO), and precipitation data was used. The second section examines the Geographic Information Systems (GIS) software used to extract basin characteristics. ArcInfo was used to manipulate and extract relevant basin characteristics. The third section contains information on the procedures used to develop the watershed database. A step by step procedure is given to extract hydrological information from DEMs. The significant basin characteristics extracted from each set of DEMs are:

- Total drainage area (TDA)
- Relative relief (RR)
- Basin slope (BS)
- Main channel slope (MCS)

The fourth section of this chapter discusses the peak flow frequency analysis. Collection of peak flow records and how the records are related to basin characteristics are presented. The regression method and a discussion of the selection of relevant basin characteristics are given. The last section examines the graphical relationship between basin characteristics and peak flows. It demonstrates how each basin characteristic might influence peak flow magnitudes.

3.1 Source Datasets

Digital Elevation Models (DEM) were the primary sources of data from which basin characteristics were extracted. DEMs are commercially available from the USGS. Soils and precipitation information was obtained from Soenksen et al. (1999a), for Nebraska and adjacent states. The soils and precipitation data were developed by the Natural Resources and Conservation Service (NRCS) and National Climatic Data Center (NCDC), respectively.

3.1.1 Digital Elevation Models

Watersheds were delineated from 1:24,000-scale Digital Elevation Models (DEM), with 30-meter resolution. For each gaging station the stream networks and elevation data were all developed from the same dataset. DEM data files are a digital representation of topographic data in raster format. They consist of an array of elevations representing ground positions at regularly spaced horizontal intervals. The DEMs used in this project are based on 30 m by 30 m data spacing intervals with a Universal Transverse Mercator projection, and are commercially available from the USGS. The DEMs were collected and processed to produce hydrologic derivative datasets to compute watershed characteristics. The use of a single dataset simplifies and increases processing speed of the basin network analysis. It also produces a uniform dataset with seamless basin measurements. With the proper software, watershed elevation models are relatively simple to produce. However, many of the important basin characteristics are scale dependent; and require that a 1:24,000-scale DEM be used if the characteristics are to be implemented in regression equations. Examples of scale-dependent characteristics include main channel length, main channel slope, and basin slope. Unfortunately, a number of basin properties that have a strong influence on peak flows, are also impossible to extract from electronic data without displaying some form of scale dependence.

3.1.2 Soil Characteristics

Soil information was defined by Dugan (1984) and interpreted using ArcInfo polygon coverages. A thorough soils database for each station was created in WRI 99-4032, and is considered unchanged since its development. The soil characteristics for Nebraska were delineated from 1:24,000-scale maps; 1:250,000-scale topographic maps were used for basins outside of Nebraska. The soil database for each basin was developed from a digital data layer of the State Soil Geographic Data Base (STATSGO), (NRCS, 1994).

3.1.3 Precipitation Characteristics

Two average precipitation characteristics were collected: the mean annual precipitation (MAP) and the two-year, 24-hour precipitation event (TTP). The two-year, 24-hour precipitation event was digitized into an ArcInfo layer, with a contour interval of 1-inch. Mean annual precipitation was based on data from the National Oceanic and Atmospheric Administration and the National Climatic Data Center for the period of 1961-1990. Information was collected at each weather station and Thiessen polygons were created. As with the soils, precipitation data was collected from Soenksen et al. (1999a) for each watershed.

3.2 GIS Software

Basin characteristics were quantified using the hydrologic modeling functions within ArcInfo. ArcInfo was used to manipulate the DEMs into useable hydrologic information. The Spatial Analyst package was required to process the DEMs. A majority of the DEM transformation was done using the Arc Workstation command line. Arc Workstation was used to manipulate the DEMs into seamless elevation grids and to develop hydrologic derivatives. Editing was also done with ArcToolbox and ArcGIS to get a final product. ArcToolBox was used for editing line and polygon coverages and defining projections. All basin characteristics were extracted using ArcGIS. The process of extracting hydrologic information, such as watershed boundaries and stream networks, from DEMs was done using a combination of the programs within ArcInfo.

3.3 Methods Used to Create Watershed Database

In this section the methods used to develop a watershed database from Digital Elevation Models (DEM) are presented. The first step involves processing the DEMs into depressionless grids for proper basin delineation. In the second step, two hydrologic derivatives are created from the DEMs. The derivatives are then used to delineate watersheds and stream networks. The extraction and explanation of relevant basin characteristics is also discussed. Many of the commands were executed from the **Grid:** command line, so examples discussed in this section are preceded by the **Grid:** identifier. In the examples, user selected input and output files are italicized (e.g., *Mosaic*), and Spatial Analyst commands are usually capitalized (e.g., MOSAIC). Some commands require data from the user (e.g., x-coordinate and y-coordinate).

3.3.1 DEM Processing

If basin characteristics determined from DEMs are to be used in the development of regression equations, the first step is to locate the gaging station of interest on the DEM. Geographic coordinates are needed to determine the location and size of the basin. Previous knowledge of the selected basin is helpful when collecting applicable DEMs. The coordinates are converted into point coverages, which can be displayed with the DEM in ArcGIS. Creating point coverages allows the stations to be viewed spatially with the other GIS resources. The United States Geological Survey (USGS) has divided 1-degree blocks into 64, 7.5-minute quadrangles, each identified by an alpha-numeric code. Multiple DEMs are often used to represent large watersheds since each DEM has a coverage area of only one 7.5-minute quadrangle. Selecting the necessary DEMs can be done by trial and error or by using good engineering judgment. The DEMs selected should include the entire watershed boundary, and can be selected by inspecting the DEMs in ArcGIS. Figure 3.1 is an example of a gaging station and the collection of upstream DEMs that make up the associated watershed. In Figure 3.1, the entire watershed upstream of the gaging station is covered by the DEMs provided, and watershed divides are clearly visible.

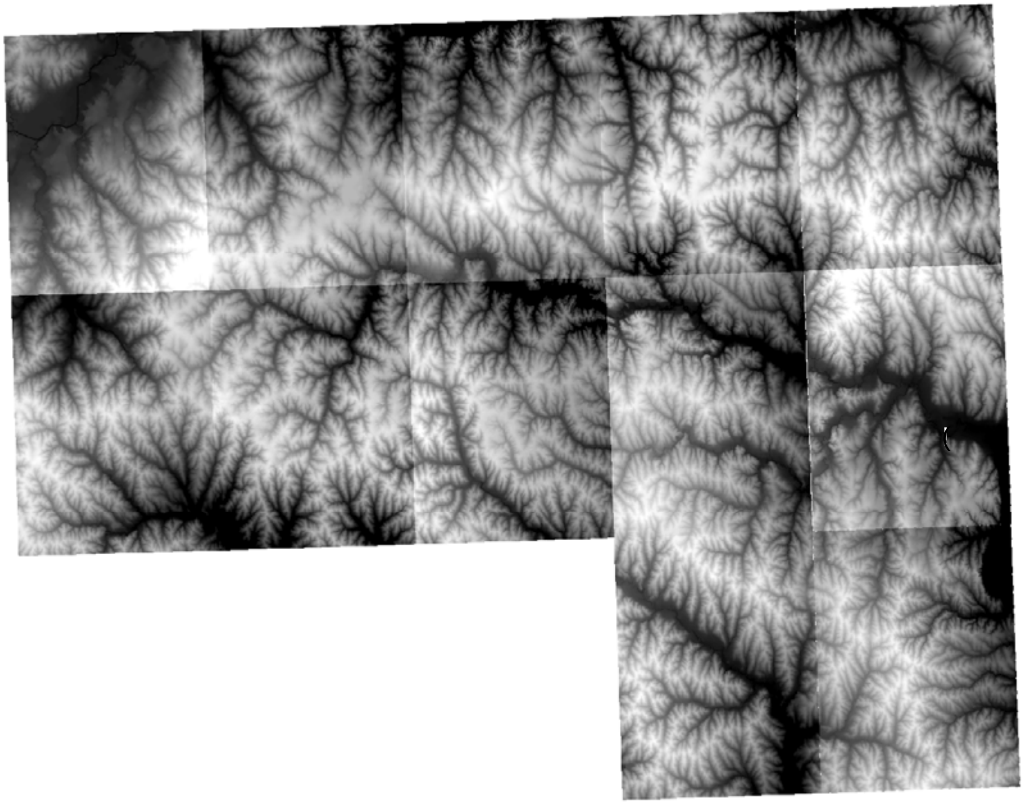


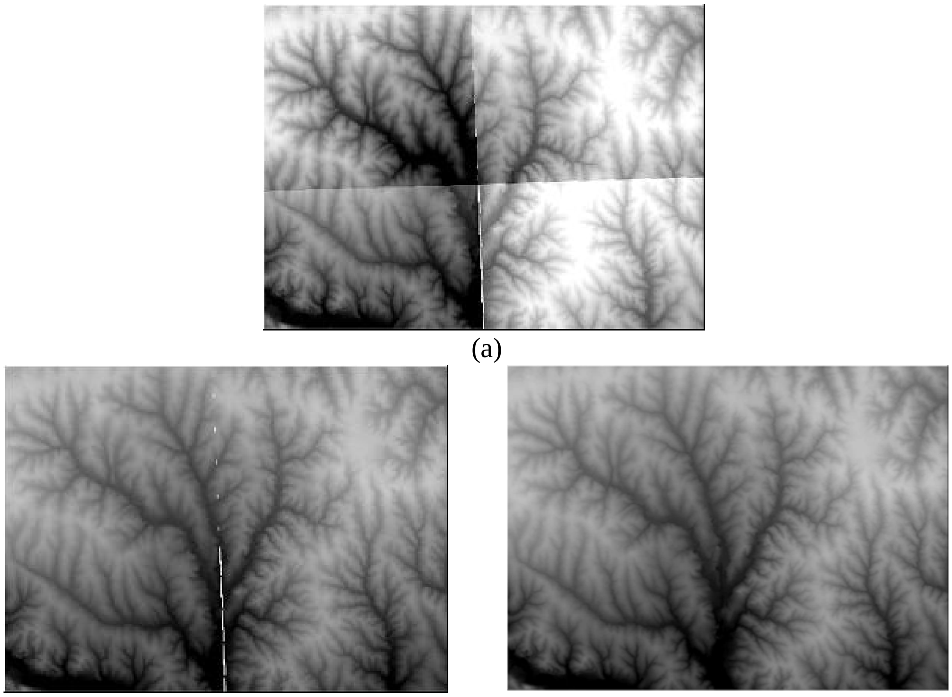
Figure 3.1: Gaging station 06806500 located on Weeping Water Creek at Union, Nebraska, illustrated with DEMs.

Once the appropriate DEMs have been identified they can be manipulated using the Arc Workstation command line. First, the MOSAIC function in the Arc Workstation grid module is used to combine all adjacent 7.5-minute grids (DEMs) to form one grid. As shown in Equation 3.1, the name of each grid that makes up the watershed is provided to the MOSAIC function.

Grid: *Mosaic* = MOSAIC (*grid1*, *grid2 …*) **(3.1)**

In Arc Workstation, the Mosaic command is limited to processing 50 grids at one time, including the one created. Mosaic creates a smooth transition between the overlapping areas of neighboring grids. The result of the Mosaic command is demonstrated in Figure 3.2b after the command is applied to the four adjacent grids shown in Figure 3.2a. In some cases, there are small gaps of missing data between adjacent grids. A majority of data gaps occur at the intersections of 1-degree blocks. The MOSAIC function does not interpolate and fill in any missing data. Thus, a Grid expression, Equation 3.2, is used to fill gaps of missing data by interpolating elevations from neighboring cells (the gaps may be as wide as three rows or three columns). Figure 3.2b shows a seam that remains after the MOSAIC function has been applied, but the seam can be filled to match surrounding data as shown in Figure 3.2c.

Grid: *Seams* =
con (isnull (*Mosaic*), focalmean (*Mosaic*, rectangle, 5, 5), *Mosaic*) **(3.2)**


(a)

(b) (c)

Figure 3.2: The steps involved in the mosaic process: (a) original DEMs before processing, (b) output of the MOSAIC function with a seam, and (c) final result after the seam is filled.

Once a seamless grid has been created, all sinks within the DEM have to be removed so that basin boundaries can be determined. Sinks are depressions within the DEM caused either by error in the grids or by natural depressions. Sinks result in discontinuous stream networks within the watershed. There are natural sinks, but most sinks are caused by the resolution of the DEMs. These errors are often due to sampling effects and the rounding of elevations. There are some regions where natural sinks are common, and caution should be exercised when applying regression equations in these regions. However, naturally occurring sinks in elevation data are rare if the cell size is 30 meters or larger, and can usually be considered to be errors. In the DEMs, basins with large areas of natural sinks are considered part of the total drainage basin and are included within the watershed. It was found by Tarboton, et al. (1991) that from 0.9 to 4.7 percent of cells in a DEM consisted of natural and false sinks. The most common sinks had depths with a range of 2.6 to 4.8 meters (8.5 to 15.7 ft). To correct for errors in the data the FILL command, Equation 3.3, is used to fill all sinks in the watershed boundary.

Grid: FILL *Seams Fillgrid* sink 50 **(3.3)**

Because of errors in resolution, filling the sinks is also done to ensure proper delineation of the basins and streams. Equation 3.3 fills all sinks with a depth of less than 50 feet. A large z-limit or depth is chosen to ensure that all sinks are filled within the DEM, whether or not they are natural. The output file *Fillgrid* provides a depressionless DEM ready for hydrologic development.

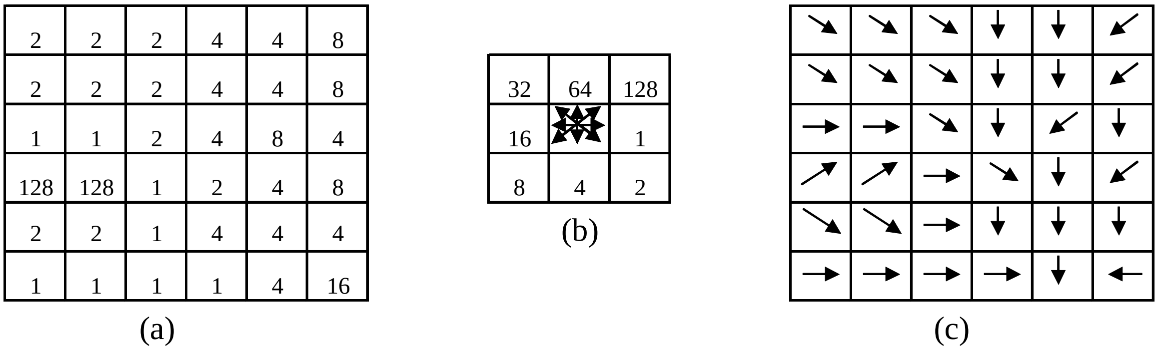
3.3.2 Hydrologic derivatives

The next step is to create two hydrologic derivative datasets with the Arc Workstation grid module: grids containing flow direction and flow accumulation. The FLOW DIRECTION function is the first important grid derived from the DEM. Figure 3.3 is an example of the DEM grid cells with their representative elevations (ESRI, 2002). Equation 3.4 creates a grid of flow direction from each elevation cell into its steepest down slope neighbor. As input, it requires the depressionless elevation grid created by the FILL command.

Grid: *Flowdir* = FLOWDIRECTION (*Fillgrid*) **(3.4)**

78	72	69	71	58	49
74	67	56	49	46	50
69	53	44	37	38	48
64	58	55	22	31	24
68	61	47	21	16	19
74	53	34	12	11	12

Figure 3.3: Representative digital elevation grid (numbers shown are elevations).



2	2	2	4	4	8
2	2	2	4	4	8
1	1	2	4	8	4
128	128	1	2	4	8
2	2	1	4	4	4
1	1	1	1	4	16

(a)

32	64	128
16		1
8	4	2

(b)

(c)

Figure 3.4: Flow-direction grid with physical representation of the flow direction (Environmental Systems Research Institute, 2002)

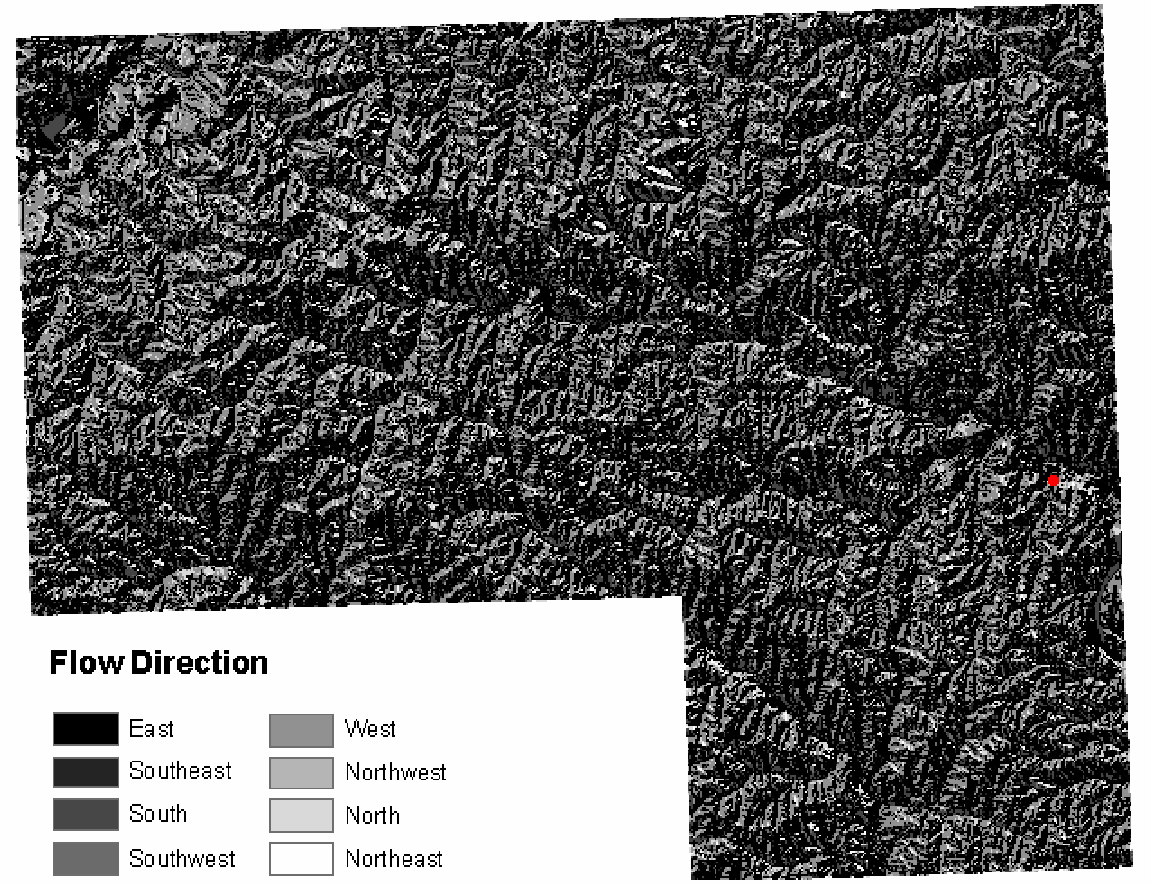


Figure 3.5: Visual illustration of the flow-direction grid

Each cell in the elevation grid is assigned a value that identifies the direction of flow out of the cell (Figure 3.4b) - there are eight possible flow directions from each cell. In Figure 3.4a, numbers have been assigned that identify the flow direction from each cell. Figure 3.4c is the vector representation of the flow direction (based on the numbers in Figure 3.4a). The flow direction grid is used to delineate watersheds by compiling all of the cells that concentrate to an outlet. Figure 3.5 is a visual illustration of the flow direction grid for the dataset in Figure 3.1.

The second hydrologic derivative map is the flow accumulation grid, and is calculated from the flow direction grid. It is used to develop the stream network and to identify watershed outlets. The FLOW ACCUMULATION grid records the number of cells that drain (both directly and indirectly) to an individual cell in the grid (Figure 3.6). Cells located on the watershed divide have a flow accumulation of 0 (since no cells drain into them), while the cell with the highest flow accumulation is located at the watershed outlet. Equation 3.5 is the command used to derive a flow accumulation grid from a flow direction grid.

Grid: *Flowacc* = FLOWACCUMULATION (*Flowdir*) **(3.5)**

0	0	0	0	0	0
0	1	1	2	2	0
0	3	7	5	4	0
0	0	0	20	0	1
0	0	0	1	24	0
0	2	4	7	35	2

Figure 3.6: Flow-accumulation grid showing the cumulative number of cells that drain into a given cell in the flow network (Environmental Systems Research Institute, 2002)

Figure 3.7 is a graphical representation of the output of the flow accumulation function in ArcGIS. The darkest cells have the largest number of cells draining to them. A synthetic stream network is developed by identifying cells with high flow accumulations. The flow accumulation grid was also created using the Arc Workstation Grid module. A stream network can then be created by applying a threshold value to the flow accumulation grid. Only cells with an accumulation value that is greater than the threshold are included in the stream network. If a threshold was not chosen, individual streams could not be identified (i.e. a threshold of 0 would include every grid cell as part of the stream network).

Grid: *Threshold* = con (*Flowacc* > 750, 1) **(3.6)**

After comparing grids by trial and error, a threshold value of 750 cells was chosen. This threshold value generates stream networks similar to those of previous research (WRI 99-4032) on Nebraska streams.

Figure 3.7: Flow-accumulation grid of Weeping Water Creek.

3.3.3 Watershed Delineation

Finally, after the hydrologic derivatives have been created, the watershed is delineated. The flow direction and flow accumulation grids, along with the gaging station coordinates, are used to determine watershed area. The actual location of the station will most likely not fall directly on a stream created from the flow accumulation grid. Figure 3.8 shows a gaging station located close to the main channel. In order for the watershed to be delineated properly, the gaging station must be located in one of the cells that the main channel is comprised of. To correct this, the station can be manually interpreted or snapped to the nearest stream. Manually selecting the watershed outlets closest to the gaging station minimizes errors when selecting basin outlets. The most likely error is snapping the station to the wrong stream. It is important to know the station coordinates within the main channel because they must be manually inputted into the grid module. The SELECT POINT function, Equation 3.7, selects the outlet cell from which the watershed is created.

Grid: *Gage* = SELECTPOINT (*Fillgrid*, x-coordinate, y-coordinate) **(3.7)**

The output of Equation 3.7 is a point grid located at the coordinates that were entered into the Arc Workstation grid module. Once the coordinates of the basin outlet are identified, it can be entered into the WATERSHED function in Arc Workstation, Equation 3.8. This function relies on the flow direction grid and the station point grid.

Grid: *Basin* = WATERSHED (*Flowdir*, *Gage*) **(3.8)**

Equation (3.8**)** determines the total number of cells flowing into a given outlet. The output file contains the drainage area upstream from the gaging station. The SELECTPOINT and WATERSHED functions can be combined into one expression, Equation 3.8a.

Grid: *Basin* = WATERSHED (*Flowdir*, SELECTPOINT (*Fillgrid*, x-, y-)) **(3.8a)**

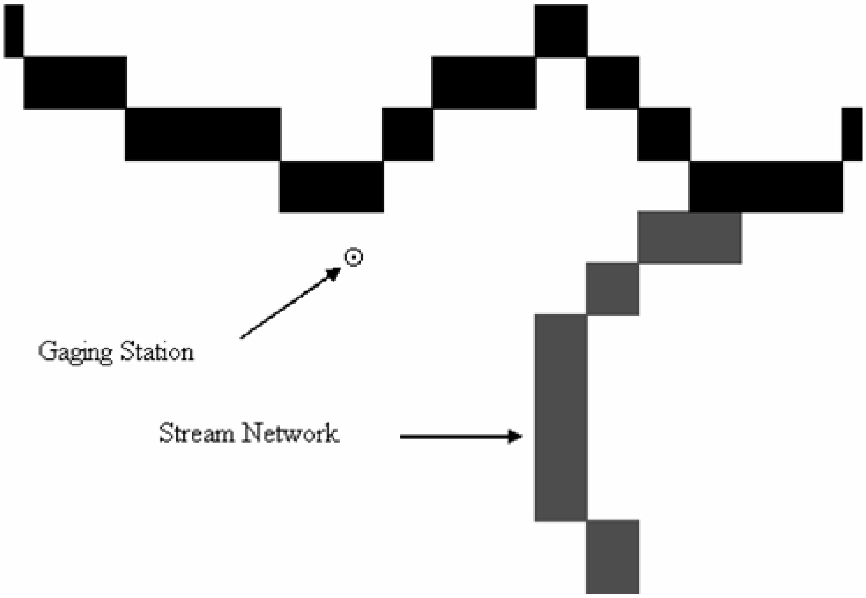


Figure 3.8: The position of the gaging station pictured is near the stream channel but not directly located in the stream bed.

The grid file *Basin*, determined using Equation 3.8a, can be projected in ArcGIS but lacks vector information. To extract spatial information, the grid file needs to be converted to a polygon coverage. This can be done in both Arc Workstation and ArcToolBox. The function GRIDPOLY, in the grid module, converts a grid file into a coverage with spatial attributes, as demonstrated in Equation 3.9.

Grid: *Watershed* = GRIDPOLY (*Basin*) **(3.9)**

The output to Equation 3.9 contains an attribute table with the total watershed area (TDA) and basin perimeter (BP) in map units. ArcToolBox can also produce a polygon coverage using the *import to coverage* menu item, resulting in the same attribute information. Some cleaning and engineering judgment will be required to reach the final result in ArcGIS. When converting from a grid file the coverages occasionally lose their coordinate system. By cleaning the file you can reestablish the coordinate system from the source grid. Figure 3.9 shows the original DEM with the watershed coverage superimposed on it. Simultaneously plotting the two coverages provides a good assessment of whether the watershed boundaries agree with the elevation data.

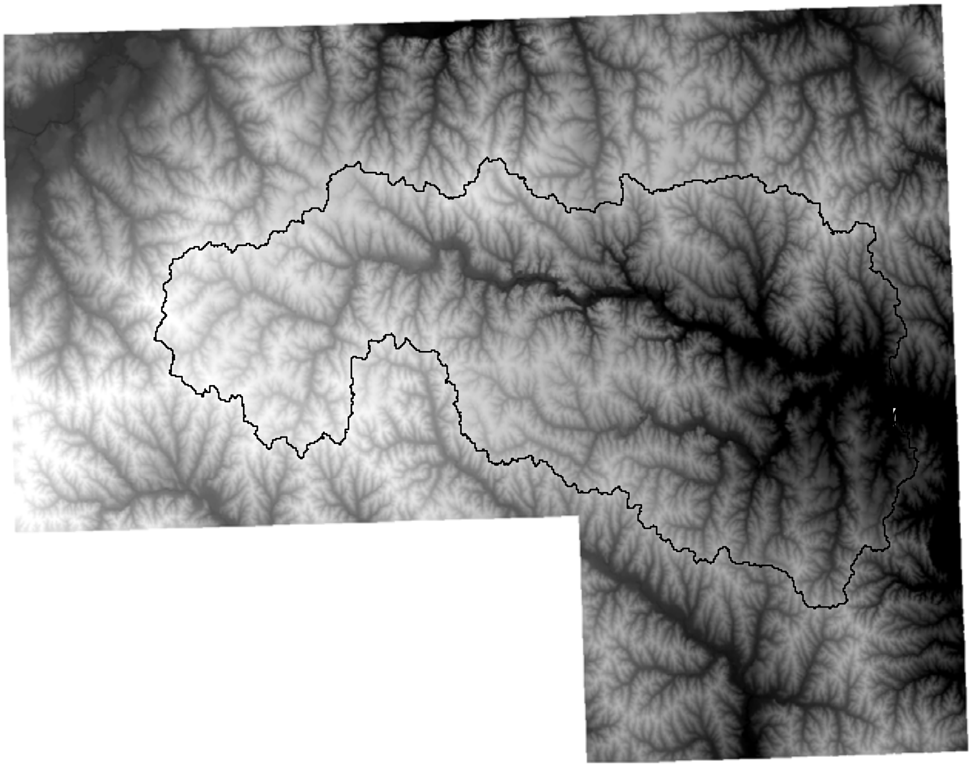


Figure 3.9: The watershed coverage and gaging station as displayed on the DEM.

3.3.4 Stream Development

After the watershed has been created, it can be used to determine the main channel length. The main channel is measured from the basin outlet to the intersection of the main channel and the basin boundary. The FLOWLENGTH function can determine the length of the longest reach, but the process involves multiple steps. The input files to Equation 3.10 require the elevation grid, the flow direction grid, channel coordinates and the WATERSHED and SELECTPOINT functions.

Grid: *Step1* = FLOWLENGTH (*Flowdir*,
WATERSHED (*Flowdir*, SELECTPOINT
(*Fillgrid*, x-coordinate, y-coordinate)), upstream) **(3.10)**

When using the FLOWLENGTH command in ArcGIS, it is recommended to overlay the flow accumulation grid with the watershed polygon coverage. First visually locate the main channel on the flow accumulation grid and locate the point halfway upstream. Record the coordinates and input them into the FLOWLENGTH function. The output will give a single channel below the midway point and a stream network above it. The upstream network created from the first FLOWLENGTH output (Figure 3.10), is used to find the next point upstream.

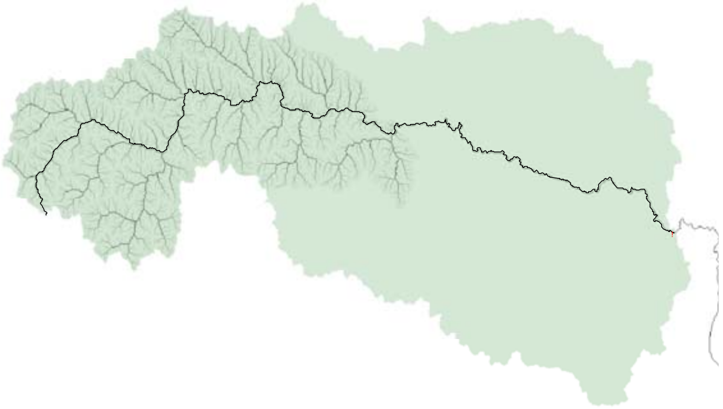


Figure 3.10: The first output of the flow length function.

Continue to move upstream until the longest flow path from the basin outlet contained in the watershed is found. The flow length command is an iterative process that may take several steps. The end result is a single channel that starts at the basin outlet and goes upstream to the intersection of the main channel and the basin boundary. The next step is to assign a threshold value to the grid line, Equation 3.11.

Grid: *Length* = con (*Step* > 100, 1) **(3.11)**

A threshold of 100 cells was chosen to represent the main channel. The threshold in Equation 3.11 creates a grid in which all cells that are part of the main channel are filled with the value “1”, and all cells that are not part of the main channel are empty (filled with “NODATA”). This process is necessary before converting to line coverage. Finally the *Length* grid needs to be converted into a line coverage to extract the attribute information. This again can be done with both Arc Workstation and ArcToolBox. GRIDLINE from the grid module converts a grid file into a line coverage with spatial attributes as shown in Equation 3.12.

Grid: *Mainchannel* = GRIDLINE (*Length*) **(3.12)**

The output of Equation 3.12 is an attribute table with the length of the main channel. ArcToolBox also produces a line coverage from the *import to coverage* menu, which results in the same attribute information. A visual representation of the main channel extended is given in Figure 3.11. Some cleaning and engineering judgment will be required to reach the final result in ArcGIS. The *Mainchannel* coverage occasionally contains loops in areas of low relief. These loops, which artificially lengthen the main channel, can be removed by manually cleaning the coverage.

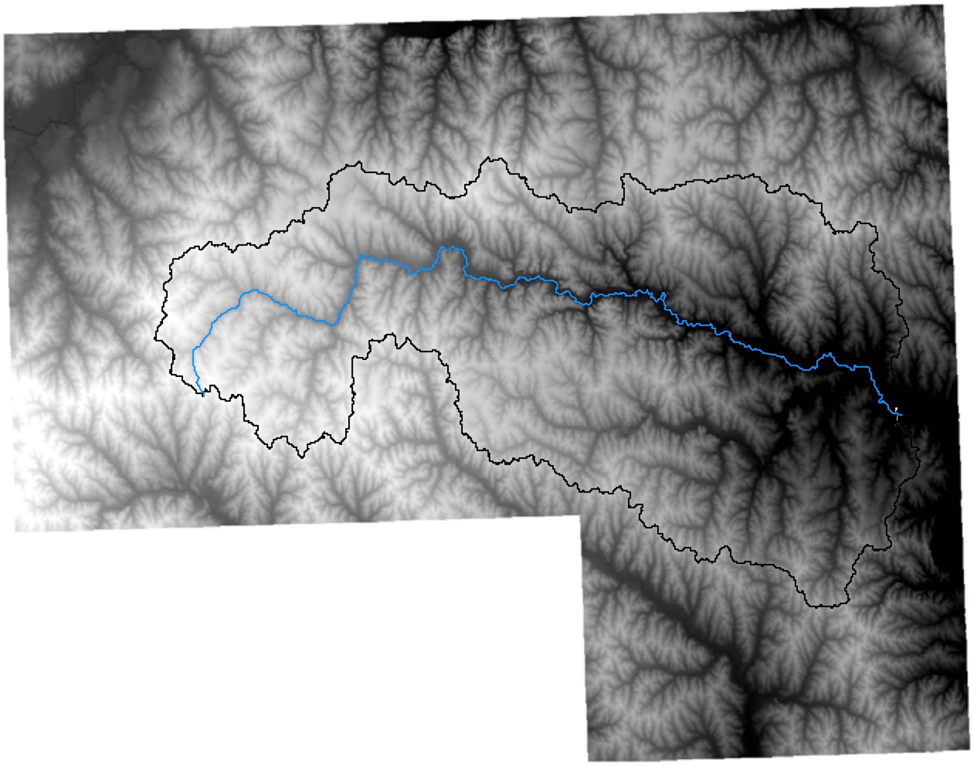


Figure 3.11: The main channel length measure from the basin outlet to the main channel extended and the basin boundary.

The next major step is to determine the stream order and total stream length of the network. The most common and standard method to assign stream order is the Strahler method. The Stream network utilizes the STREAMLINE and STREAMORDER function as well as the flow direction and flow accumulation threshold.

Grid: *Network* = STREAMLINE (STREAMORDER (*Threshold*, *Flowdir*, Strahler), *Flowdir*) **(3.13)**

The STREAMORDER, Equation 3.13, assigns a numeric order to line segments of a grid representing branches of a linear network (Figure 3.12). The STREAMLINE function converts a grid representing a raster linear network to a line coverage. A threshold value was assigned, as discussed for the flow accumulation grid.

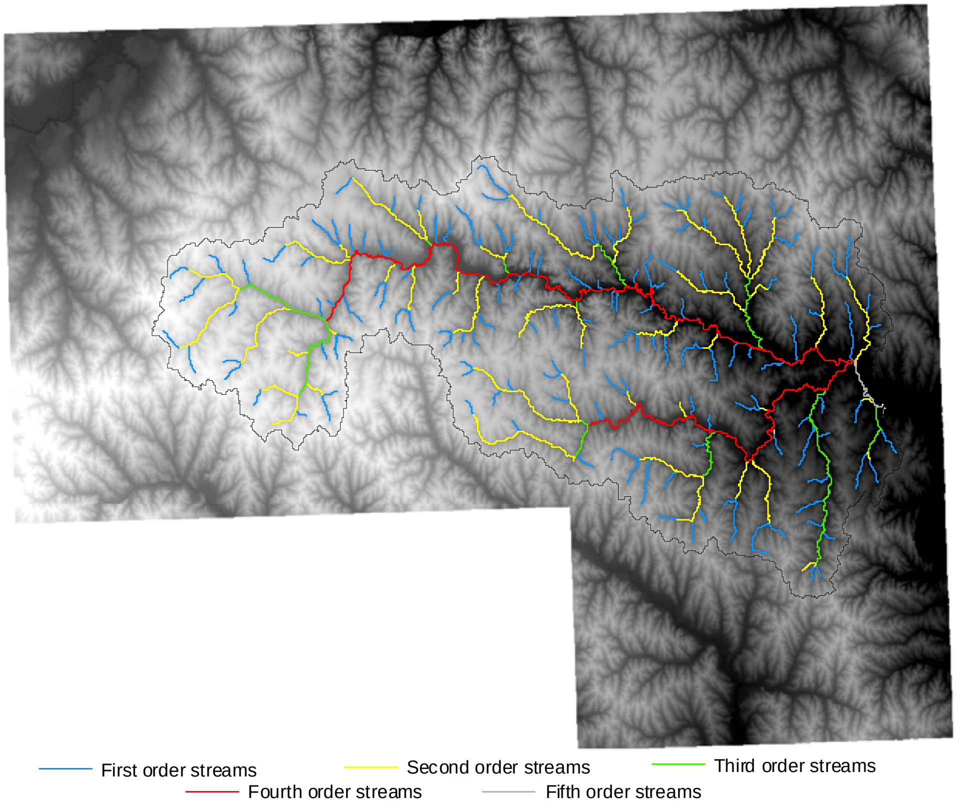


Figure 3.12: Shrahler method of numbering stream network.

The last step in the grid module is clipping the elevation grid (DEM) with the watershed polygon. GRIDCLIP uses the DEM and the watershed polygon coverage as shown in Equation 3.14.

Grid: GRIDCLIP *Fillgrid Clipped* COVER *Watershed* **(3.14)**

Equation 3.14 clips the grid within the constraints of the watershed. Now the highest and lowest elevation grid cells can be found by sorting the attribute table. The DEM contains elevation data, and can be directly measured. COVER is used to identify that the clipping will be using a polygon coverage. The resulting output file *Clipped* contains only the elevation data for the selected basin. The elevation of the basin outlet can also be found from the intersection of the main channel and watershed boundary.

Clipping the elevation grid allows the extraction of the basin relief. Basin relief is the elevation difference between the highest grid cell (E_{max}) and the grid cell at the basin outlet (E_{min}). This allows for the calculation of the Relative Relief (RR) of the drainage basin.

$$RR = \frac{E_{\max} - E_{\min}}{BP} \qquad \textbf{(3.15)}$$

Where, BP is the length of the perimeter of the basin. The relative relief is a significant basin characteristic that can be related to peak flow frequencies. Figure A.1 in the appendix shows the locations of the basin relief quantification.

3.3.5 Using the Arc Module

The Workstation Arc module can be used to create contours with 10-foot contour intervals from the elevation grid (DEM). The command LATTICECONTOUR is used to create contours that are representative of the 7.5-minute DEM, and is demonstrated in Equation 3.16.

```
Arc: LATTICECONTOUR Fillgrid Contour10 10                                   (3.16)
```

The output file from Equation 3.16, *Contour10*, will contain some extra lines that should be discarded within ArcGIS. These errors are often due to sampling effects and the rounding of elevations within the DEM. To correct this, all contour lines less than 200 meters in length were deleted (this length, though somewhat arbitrary, appears to eliminate erroneous contour lines). After weeding out the stray contours, remaining contours should be clipped within the given watershed polygon coverage. The CLIP command requires the contours that need to be clipped (*Contour10*) and the clip coverage (*Watershed*) as inputs.

```
Arc: CLIP Contour10 Contours Watershed LINE                                 (3.17)
```

Equation 3.17 will remove all of the contour lines outside of the watershed of interest. In Equation 3.17, the LINE command tells ArcInfo that the output file is a line-coverage. The output will be a line-coverage with lengths and elevations of each line segment. Equations similar to 3.17 (Equations 3.18 and 3.19) can be used to clip the stream network and the main channel extended, removing all streams that are outside of the region of interest.

```
Arc: CLIP Network Streams Watershed LINE                                    (3.18)

Arc: CLIP Mainchannel MCL Watershed LINE                                    (3.19)
```

In Equation 3.18, the *Network* coverage is clipped with the *Watershed* and results in a stream network for the watershed. In Equation 3.19, the file *Mainchannel* is clipped and contained within the watershed.

Creating contours that have length and elevation attributes is helpful when calculating the average Basin Slope (BS). Basin slope is quantified using the "contour band" method and is computed as:

$$BS = \frac{\left(\sum Contour_{Lengths}\right)(10\,feet)}{TDA} \tag{3.20}$$

Where, $Contour_{Lengths}$ is the length of each 10-foot elevation contours within the watershed, and TDA is the total drainage area. Average basin slope was determined to be a significant basin characteristic when correlated with peak flow quantities.

At this point, the only hydrologic modifications needed from ArcInfo are the elevations at 10 and 85 percent of the distance along the *MCL* upstream from the basin outlet. It is recommended to create two copies of the *MCL* to be edited in ArcGIS. Using the trim function, edit each line separately to get the adjusted lengths. The final result will allow for manual extraction of the elevation data at the end of each line segment.

The purpose of finding the elevations at 10 and 85 percent of the distance along the main channel upstream from the basin outlet is to calculate the Main Channel Slope (MCS); this is done using Equation 3.21:

$$MCS = \frac{E_{85} - E_{10}}{0.75MCL} \tag{3.21}$$

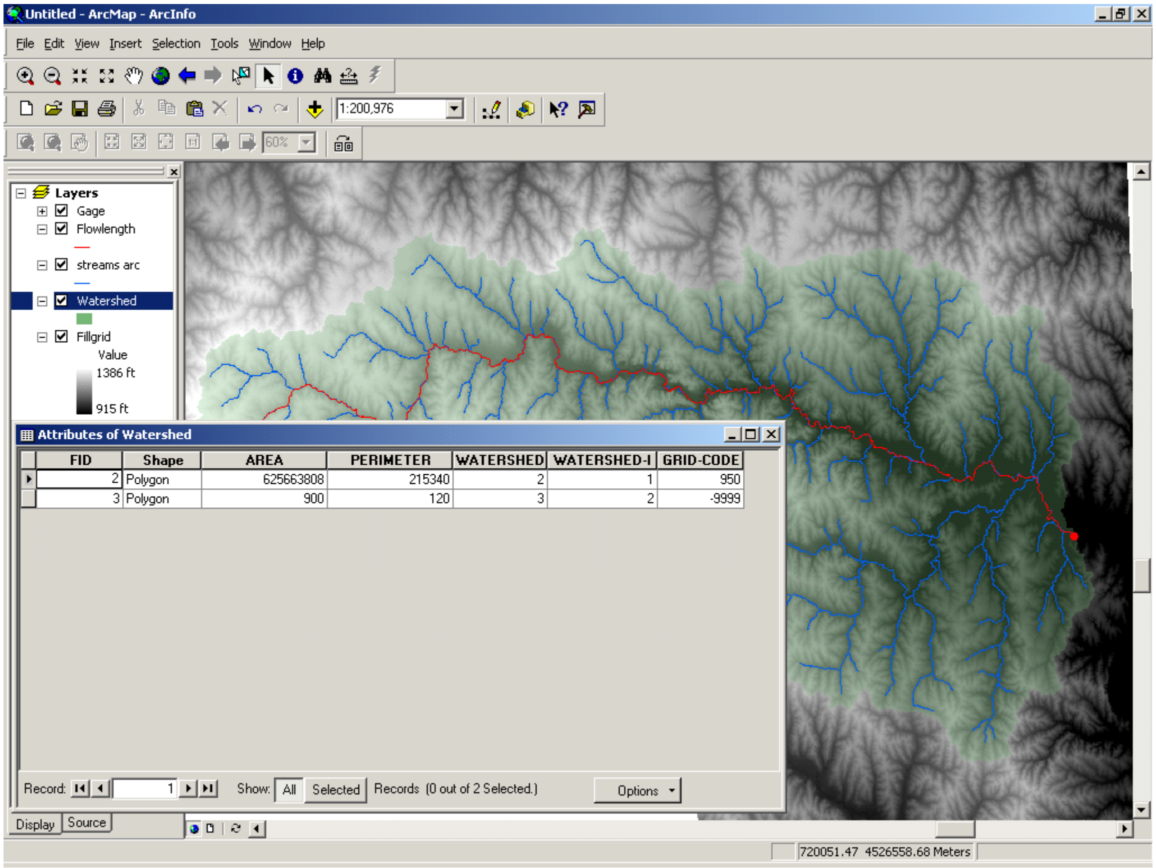
Where, E_{10} and E_{85} are the respective elevations and MCL is the main channel length. Main channel slope was found to be statistically relevant when related to peak flow estimates. Figure A.2 in the appendix shows the contour lines and the locations of the MCS variables.

3.3.6 Data Extraction

After all of the files have been created in Arc Workstation, they must be opened in ArcGIS to extract their attributes. There are 12 measured morphometric basin attributes that are used for the calculation of other basin attributes. Attributes extracted directly from the DEMs are the:

- Total drainage area (TDA)
- Basin perimeter (BP)
- Main channel length (MCL)
- Total stream length (TSL)
- Number of first order streams (FOS)
- Basin stream order (BSO)
- Highest elevation grid cell (E_{max})
- Elevation at the basin outlet (E_{min})
- Total length of elevation contours
- Contour interval (10 feet)
- Elevation at 10% of the upstream distance along the main channel (E_{10})
- Elevation at 85% of the upstream distance along the main channel (E_{85})

The *Watershed* coverage gives the basin area and perimeter in map units in the attribute table (Figure 3.13). The *MCL* attribute table gives lengths to each poly line that make up the entire main channel extended. The total length of the main channel can be calculated by summing the lengths within the attribute table. The *Stream* network attribute table contains stream lengths and stream order. Total stream length (TSL) can be found by summing the lengths of each stream segment. Stream order can be found by sorting the stream order column, and the largest number is the basin stream order (BSO). From the same column the total number of first order streams (FOS) can be found by summing all first order streams. The highest and lowest elevation can be found by sorting the *Clipped* elevation grid attribute table. The *Contours* poly line also contains lengths that can be summed in the attribute table. The total length of the contours at a known contour interval is used in the calculation of average basin slope (BS). The last piece of necessary information is from the 10 and 85 percent lengths of the main channel. The elevations at the ends of the 10 and 85 percent line can be recorded from each respective grid cell. These 12 extracted characteristics are used to define and calculate 25 Morphometric characteristics. Table A.1 in the appendix gives an explanation of the basin characteristics quantified using ArcInfo.



Attributes of Watershed

FID	Shape	AREA	PERIMETER	WATERSHED	WATERSHED-I	GRID-CODE
2	Polygon	625663808	215340	2	1	950
3	Polygon	900	120	3	2	-9999

Record: 1 Show: All Selected Records (0 out of 2 Selected.) Options

Figure 3.13: Attribute table of watershed coverage in ArcGIS (ESRI, 2002).

Soil and Precipitation information were collected from the WRI 99-4032 data (Soenksen et al., 1999a). Four soil characteristics were collected for each watershed. The average permeability rate of a 60-inch soil profile (P60), average minimum permeability of the least permeable layer (PLP), average available water capacity (AWC) and the average maximum soil slope (MSS) were collected. Soil values were calculated by taking an area weighted value within each watershed. The average soil characteristics are representative of the upper 60-inch soil profile. Precipitation data were also area-weighted for each drainage basin. Two average precipitation characteristics were obtained, the mean annual precipitation (MAP) and the two-year, 24-hour precipitation event (TTP). Although the two precipitation characteristics listed above are tabulated, these characteristics did not appear to have a strong influence on the current set of regression equations and thus was not used in the present analysis.

3.4 Peak Flow Frequency Analysis

Peak flow discharges at recurrence intervals of 2, 5, 10, 25, 50, 100, 200, and 500 years were collected for 273 gaging stations in and around Nebraska. The relationship between peak flows and the frequency of occurrence for individual drainage basins were used in the development of regional analysis. A relationship was established between observed annual peak discharges and the annual exceedance probability. For all the peak flows collected, Bulletin 17B of the IACWD (Interagency Advisory Committee on Water Data, 1982) was followed in the peak flow frequency analysis. This bulletin contains guidelines for the development of these relations using Log-Pearson Type III frequency distributions. Peak flows for individual stations have been developed from peak stage and continuous record gages. Topographic information was also collected for individual watersheds. Basin characteristics, soil types, and climate all affect the annual peak discharges. The final regression equations relate peak flows to basin characteristics for each recurrence interval.

3.4.1 Peak-Discharge

In this report, peak flow frequencies were collected for Nebraska and its surrounding states. Peak stream-flows were estimated at selected recurrence intervals, ranging from 2- to 500-years, using Log-Pearson Type III distributions, since Log Pearson Type III distributions are sometimes used to predict recurrence intervals greater than the period of record. Only unregulated streams were used in the regression analysis (Soenksen et al., 1999a) as noted in Bulletin 17B. Human activities that alter flow conditions include urbanization, channelization, construction of reservoirs, diversions and changes in land cover. Despite an attempt to avoid watersheds with strong human influence, none of the watersheds used in the regression analysis are void of human influence, and changes in land use within the watersheds of interest (e.g. tillage practices) may have a significant influence on the results. Detailed peak flow frequency analyses for each gaging station were conducted in previous research projects and are considered unchanged since their completion. These analyses were utilized in the present work to save time. Only records with relatively constant watershed conditions were used in the frequency analyses, and gaging stations were required to have at least 10 years of peak flow records to be used. Regional skews were developed that could be used to detected outliers, make station comparisons, and compute confidence limits for a frequency curve. Gaging stations close to Nebraska were also used in the peak flow frequency analysis. Out-of-state peak flow data were collected from South Dakota, Iowa, Missouri, and Kansas. The number of stations with drainage areas less than 10 mi^2 in Nebraska was small, and the out of state stations helped to boost the number of stations used in the regression analysis. Stations were selected based on similar peak flow characteristics, topography, and location. The out of state stations were also included in the regional analysis. Peak flows used for the regression analysis are given in Table D.1 of Strahm (2003).

3.4.2 Nebraska Stream Data

Nebraska stream flow data were collected from Soenksen et al. (1999a). Recurrence intervals of 2 to 500 years were developed for each station. Also, the length of record was noted, which is used in the weighted least-squares regression analysis. Since the work of Soenksen et al., a small number of peak flow records have been collected. However, due to the lack of additional discharge information, the peak flow frequency analysis developed by Soenksen et al. has not changed appreciably, and the corresponding discharge predictions are considered accurate. Gaging station information through the 1994 water year was used to develop the peak discharge frequencies.

3.4.3 South Dakota Stream Data

Data from six South Dakota gaging stations used in the development of Nebraska’s generalized skew were gathered. Peak flows with recurrence intervals of 2 to 500 years were collected (Sando, 1998). Three stations in the Dry Branch Creek watershed near Parkston, SD were used, with drainage areas ranging from 25 to 100 mi^2. Also, three stations in the Saddlerock Creek basin near Beresford, SD were used; drainage areas of these stations ranged from 2 to 23 mi^2. The report by Sando (1998) has the latest analysis of peak discharges in South Dakota, including gaging station information through the 1994 water year.

3.4.4 Iowa Stream Data

Two southwest Iowa stations were used in the Nebraska regression analysis. Discharges for return periods of 2 to 500 years were collected (Eash, 2000). The Maple Creek watershed near Alta, IA, with a drainage area of 15 mi^2, was used. Also, the Soldier River basin at Pisgah, IA, with a drainage area of 440 mi^2, was used. Gaging station information through the 1997 water year was used to develop the peak flow data.

3.4.5 Missouri Stream Data

Four northwest Missouri gaging stations were used in the Nebraska regression analysis. Peak flow discharges were gathered for the 2- to 500-year recurrence intervals, but the 200-year peak flow was not available (Alexander and Wilson, 1995). A curve fit was used to estimate the 200-year peak discharge for each station. The Tarkio River basin at Fairfax, MO had the largest drainage area (470 mi^2). The other three basins were Mill Creek, White Cloud Creek, and Jenkins Branch, all with areas of less than 6.0 mi^2. All Missouri stations had at least 25 years of recorded peak discharge information. The report by Alexander and Wilson (1995) has the latest analysis of peak discharges in Missouri, including gaging station information through the 1992 water year.

3.4.6 Kansas Stream Data

Finally, gaging station information from sixteen stations in northern Kansas was collected and used. Recurrence intervals of peak discharges ranged from 2 to 200 years (Rasmussen and Perry, 2000). A curve fit was used to extrapolate the 500-year discharge for each station. Drainage areas of the Kansas stations ranged from 0.9 to 1,700 mi^2; six stations had a drainage area of less than 10 mi^2. Basins included the South Fork Sappa Creek, Beaver Creek, and the Solomon River. The northern Kansas stations have systematic records of at least 28 years. Gaging station information through the 1997 water year was used to develop the peak discharge frequencies.

3.4.7 Drainage Basin Characteristics

After a database of peak-streamflow frequencies at gaging stations was collected, hydrologic characteristics of each individual basin were collected. Twenty-five morphometric attributes were quantified using ArcInfo 8.0. The total drainage area (TDA) was collected for each basin, along with slope and stream characteristics. For basins with known non-contributing drainage areas, published contributing drainage area data were used. The stream networks were developed for the total drainage area. Soils and precipitation characteristics were collected from Soenksen et al. (1999a). The soil database was quantified from the State Soil Geographic Data Base (STATSGO), (Natural Resources Conservation Service, 1994). Precipitation data were collected by the National Oceanic and Atmospheric Administration from 1961-1990. No improvements in the resolution have been done to the soils and precipitation data layers. The detailed soils and precipitation data were considered unchanged from the previous report (Soenksen et al., 1999a). Morphometric characteristics used for the regression analysis are given in Table C.1 of Strahm (2003).

3.4.8 Nebraska Regions

The state was sub-divided into seven hydrologic regions for unregulated peak-flow frequency equations. Regionalization was based on watershed divides, soil permeability and the percentage of contributing drainage area. The seven Nebraska regions are the Big Blue, Eastern, Northeastern, Central and South-Central, Upper Republican, Northern and Western, and High Permeability region. They are a modification of Beckman’s Regions, created by Soenksen et al. (1999a). The same regions are used but with additional gaging stations within each region. Each region was then sorted by contributing drainage areas, for additional analysis. One of the purposes of this project was to emphasize small drainage basins. But, the number of streams with peak-flow records decreases for smaller basins. A cut-off of areas less than 10 mi^2 was used for an additional analysis. This allowed for an examination of small basins that had long enough peak flow records. The number of gaging stations was lower, but the analysis provided a better representation of the peak flow frequencies for small basins. Regression equations were then developed using basins with areas of less than 10 mi^2 for all of the hydrologic regions except for the high permeability region. The high permeability region lacked sufficient records to analyze watersheds with areas of less than 10 mi^2.

3.4.9 WLS Regression

Two files were created in Excel, one contained watersheds with less than approximately 10 mi^2 of area and the other had the entire range of basin areas. Each group was examined separately using the statistical program SPSS 11.5. SPSS is a windows based program that accepts imported Excel files. The non-linear function within SPSS was used to relate discharge to several basin characteristics. The peak discharge was the dependent variable, while the independent variables were the basin characteristics. For each model parameter a starting value was manually entered. This allowed the program to more quickly converge on a result and reduced the chances of erroneous convergence. The multi-variable regression model used was:

$$Q_T = A\left(CDA^x BC_1{}^y BC_2{}^z\right) \qquad \textbf{(3.22)}$$

Where Q_T is the peak discharge at a given recurrence interval, CDA is the contributing drainage area, and BC_1 and BC_2 are other basin characteristics. Model parameters of A, x, y, and z were found by regression using SPSS.

Each region was individually examined to determine which basin attributes had the greatest influence on peak flows. After each station was grouped into its respective hydrologic region, the stations were sorted by contributing drainage area. 25 morphometric, four soils and two precipitation characteristics were tabulated for each station. Individual stations also had peak-flow frequency distributions ranging from 2 to 500 years.

A weighted least-squares (WLS) procedure was used to account for the fact that the stations used for each regression had unequal record lengths. The WLS procedure was developed to deal with situations in hydrology where a regression model is heteroscedastic. The WLS procedure gives greater weight to stations with longer periods of record, while assigning less weight to stations with small periods of record. The procedure does not account for cross-correlation of peak flow data. Cross-correlation of peak flows is found when rainfall events affect multiple stations, and the resulting peak flow data are not entirely independent. The WLS model performs better than ordinary least-squares (OLS), because WLS provides distorted estimates of model error and the precision at which the parameters are being estimated (Stedinger and Tasker, 1986). The differences between the OLS and WLS procedures are primarily associated with the varying periods of record of multiple stations. OLS regression assigns the same weight to each station, regardless of the length of its flow record. The WLS model gives more weight to stations with larger records, by counting them multiple times. For example, if a station has 25 years of record, it is weighted 25 times; if it has a 10-year period of record it is weighted 10 times, etc. This creates improvements in the precision of the parameters of the hydrologic regression model when sites have varied lengths of record.

3.4.10 Basin Characteristic Selection

Before a statistical correlation was made between basin characteristics and peak flow frequencies a few guidelines were established: First, all regression equations would include the contributing drainage area (CDA) as the first basin characteristic because drainage area is directly related to the magnitude of the stream discharge. Second, the regression equations would be limited to three basin characteristics for each return period. Including more than three variables adds complexity to the equations, while only slightly improving the predictability. Third, the equations would include at least one slope or soil characteristic. Preferably, the regression equations would include both one slope and one soil characteristic. Possible slope characteristics are the basin slope (BS), main channel slope (MCS), and relative relief (RR). Soils characteristics include: available water capacity (AWC), permeability (PLP and P60), and the maximum soil slope (MSS) of the soil type. Finally, exponents greater than a power of two were avoided. When the number of stations on which the regression is based is small, large exponents can cause unwarranted emphasis of one of the basin characteristics.

As a starting point, the basin characteristics used in previous research projects were first used in the regression analysis. To simplify the equations the same groups of characteristics were repeated for each regional regression analysis. To discover which characteristics had the highest statistical correlation, each independent variable was plotted against the peak-flow frequencies. Variables with the lowest sum of squares were noted, eliminating low-correlation characteristics. The next step was to group statistically relevant variables together within the drainage area. Multiple combinations were used for each return period discharge. A trial and error process was used to eliminate combinations that did not improve the equation's accuracy. As a rule of thumb, combinations that had an R^2 value greater than 0.70 were selected.

The next step was to put all possible combinations into Excel and to compare the predicted flows to the Log-Pearson discharges. Each combination used the respective basin characteristics to predict each return period discharge. Then, the standard error of estimate (SEE) was used for equation comparisons. The *SEE* in logarithmic units was found for each recurrence interval using Equation 3.23:

$$SEE = \left[\frac{\left(\sum \left(\log Q_{REG} - \log Q_{LP3} \right)^2 \right)}{n} \right]^{1/2} \quad \textbf{(3.23)}$$

Where Q_{REG} is the predicted discharge using a regression equation, Q_{LP3} is the Log-Pearson Type III discharge, and n is the number of stations used. The combination with the lowest SEE was selected for the regression equation. The final result is a set of regression equations for the 2 to 500-year recurrence intervals that predict peak flow discharges at ungaged locations. This process was repeated for each hydrologic region in Nebraska. Six sets of regional regression equations were developed for watersheds with areas of less than 10 mi^2 and seven sets of equations were developed for regions that represent the whole range of basin areas.

3.5 Basin Characteristic Analysis

The following analysis examines the relationship between basin characteristics and how they influence the magnitude of peak flows. A Graphical comparison is made between basin characteristics and unit discharge (discharge per unit area). Contributing drainage area was used to normalize the peak flows for a wide range of drainage areas. Basin characteristics from various regions were chosen to demonstrate how unit discharges and basin characteristics might be correlated.

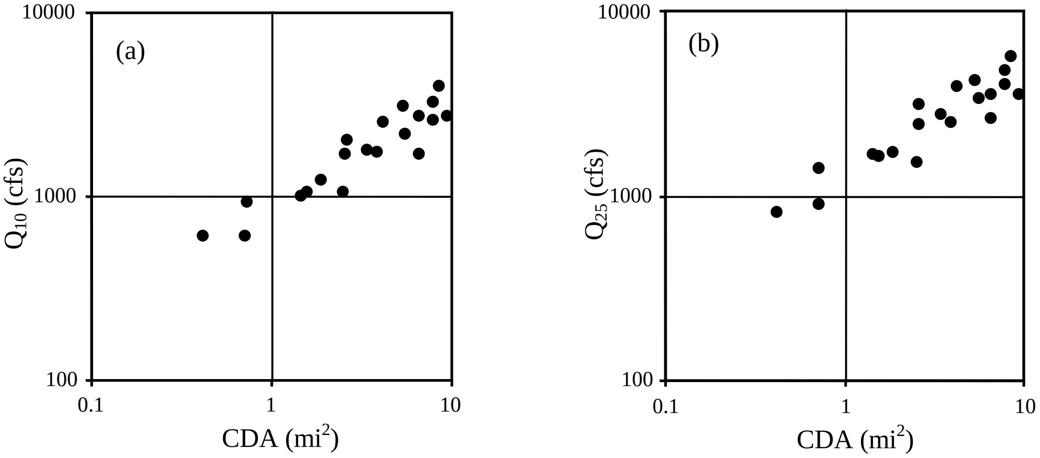
Discharges with return periods of 10- and 25-years were examined to determine the importance of each basin characteristic. The Nebraska Department of Roads (NDOR) designs some culverts with return periods of 25-years, though many designs are for 50- and 100-year events. Figures shown are:

- CDA vs. Discharge
- Relief Quantifications vs. unit discharge
- Shape Quantifications vs. unit discharge
- Soil Characteristics vs. unit discharge

The number of basin characteristics was reduced based on their statistical significance in SPSS 11.5. Nine basin characteristics are graphically compared to unit discharge for both the 10- and 25-year return periods. Examples of the plots of the 10- and 25-year return period peak flow comparisons are shown in Figures 3.14 – 3.22.

3.5.1 Contributing Drainage Area

Figures 3.14a and 3.14b show the relation between contributing drainage area and peak flow for the Eastern region. Compared to other basin characteristics, basin area has the highest correlation with peak flow discharges for all of the regions. As expected, with an increasing contributing drainage area the magnitude of the peak flow generally increases.

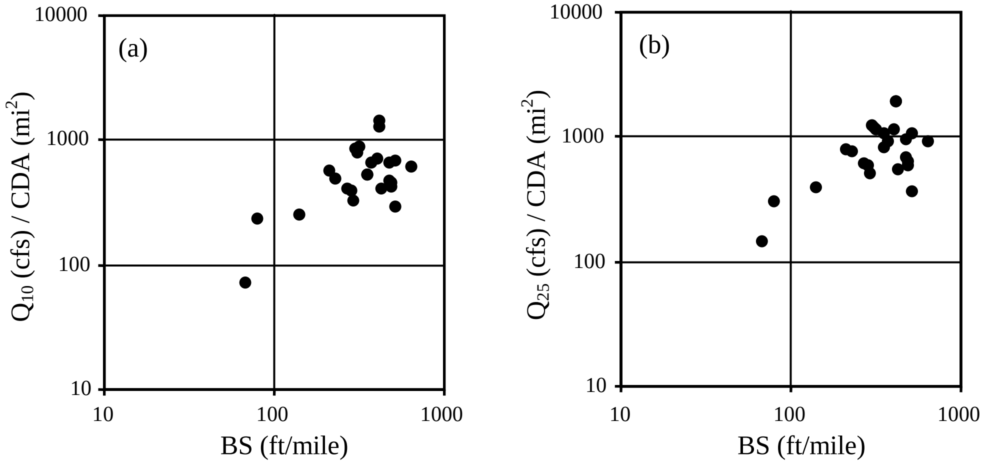


Figures 3.14: Peak discharge vs. the contributing drainage area (CDA) for the (a) 10-year and (b) 25-year peak flows in the Eastern region.

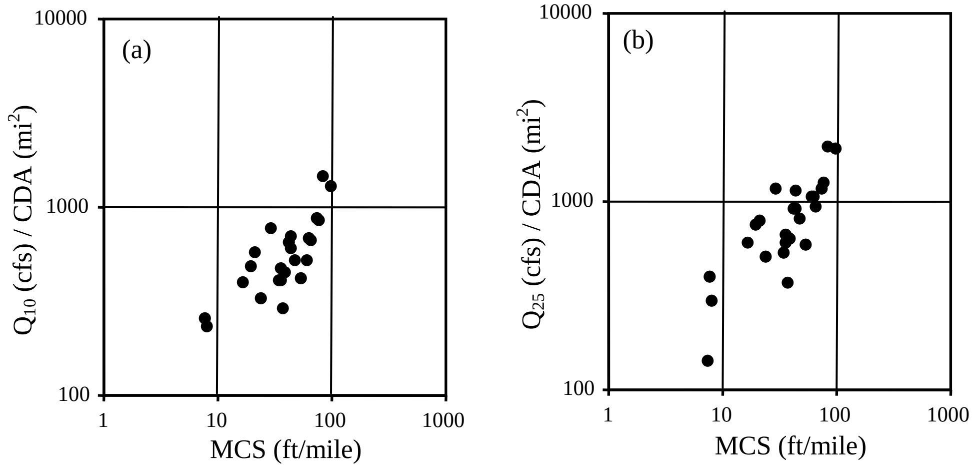
3.5.2 Relief Quantifications

Slope characteristics affect the magnitude and the time of concentration of peak flows. Significant relief quantifications were the average basin slope, the main channel slope and the relative relief. The average basin slope (BS) is a function of the lengths of elevation contours at a given interval for the drainage area. Large basin slopes represent watersheds with steep topography, and result in shorter times of concentration that generally increase the peak flow discharge. Figures 3.15a and b graphically show the relationship between basin slope and the discharge per unit area for the Eastern region. The figures show a direct relationship between unit discharge and basin slope.

The Main Channel Slope (MCS) represents the average slope of the main channel. Larger main channel slopes result in increased stream velocities in the channel. Increased velocities could potentially increase peak flow magnitudes. Figures 3.16a and b demonstrate the relationship between main channel slope and discharge per unit area for Eastern region. As expected, the figures show that with an increase in MCS the discharge per unit area also increases.

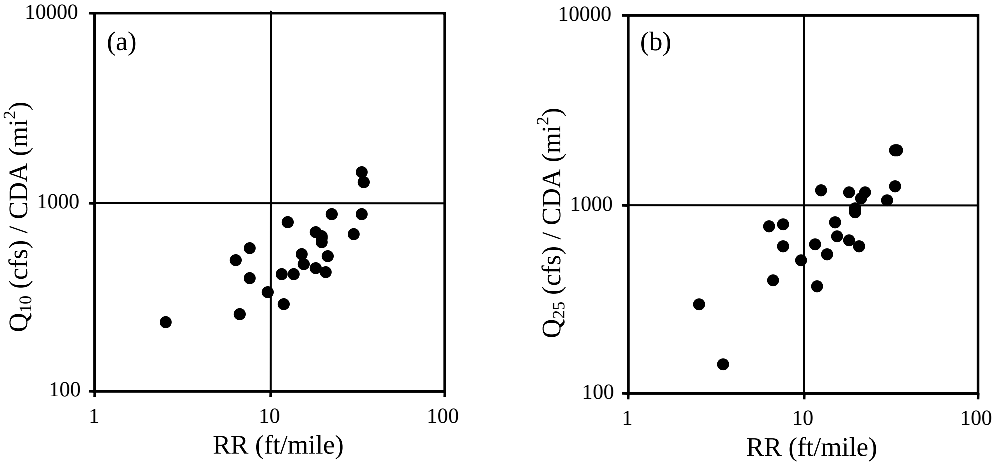


Figures 3.15: Average basin slope (BS) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in the Eastern region.



Figures 3.16: Main channel slope (MCS) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Eastern region.

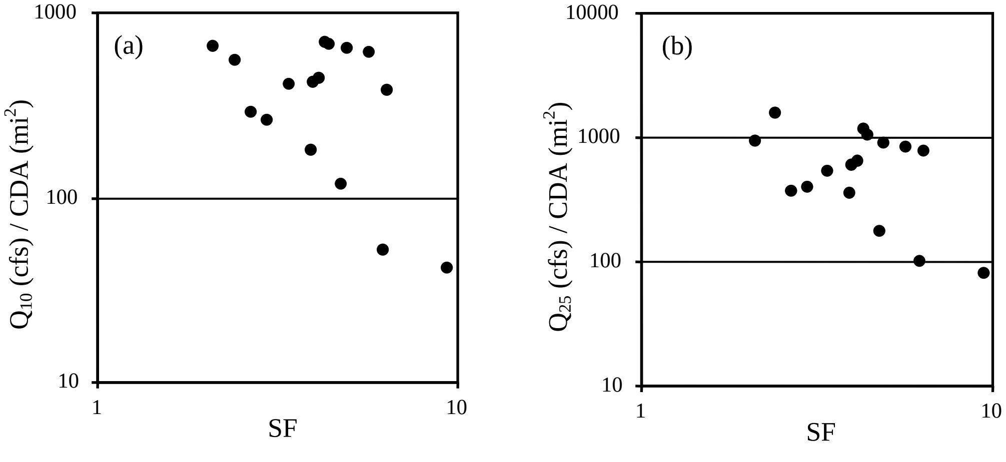
The relative relief (RR) is a function of the maximum elevation difference within the watershed over the basin perimeter. Relative relief is a slope attribute and increasing RR should cause an increase in peak discharge. For example, Figures 3.17a and 3.17b show an increasing RR with larger unit discharges for the Eastern region. The figures illustrate a correlation between relative relief and discharge per unit area in this region.



Figures 3.17: Relative relief (RR) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Eastern region.

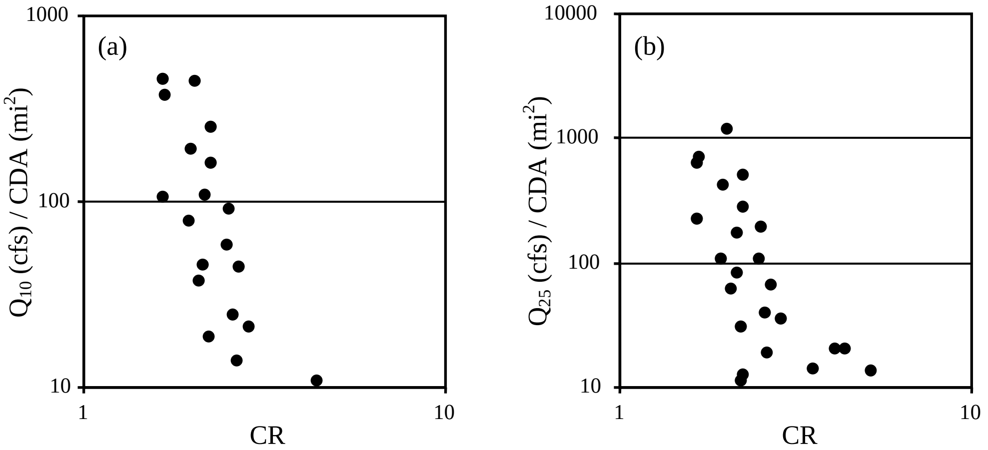
3.5.3 Shape Quantifications

Shape attributes represent the geometry of the drainage basin. Two shape characteristics investigated for use in the regression equations were the shape factor and the compactness ratio. The shape factor (SF) is a function of the main channel length and the drainage area. A large shape factor is indicative of a meandering stream. Figures 3.18a and 3.18b show a slight correlation between the unit discharge and the shape factor for the Northeastern region. Large shape factors potentially cause a decrease in the peak discharge because total rainfall runoff at a gaging station is distributed over a longer period of time (i.e., the time of concentration is longer).



Figures 3.18: The shape factor (SF) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Northeastern region.

The Compactness Ratio (CR) is a function of the basin perimeter and the drainage area. A circular basin has the smallest possible compactness ratio (1.0); the ratio increases as the ratio of the basin perimeter and the basin area increases. Figures 3.19a and 3.19b show that the compactness ratio has a slight correlation with the unit discharge in the Upper Republican region. Large compactness ratios appear to cause a decrease in the unit discharge.



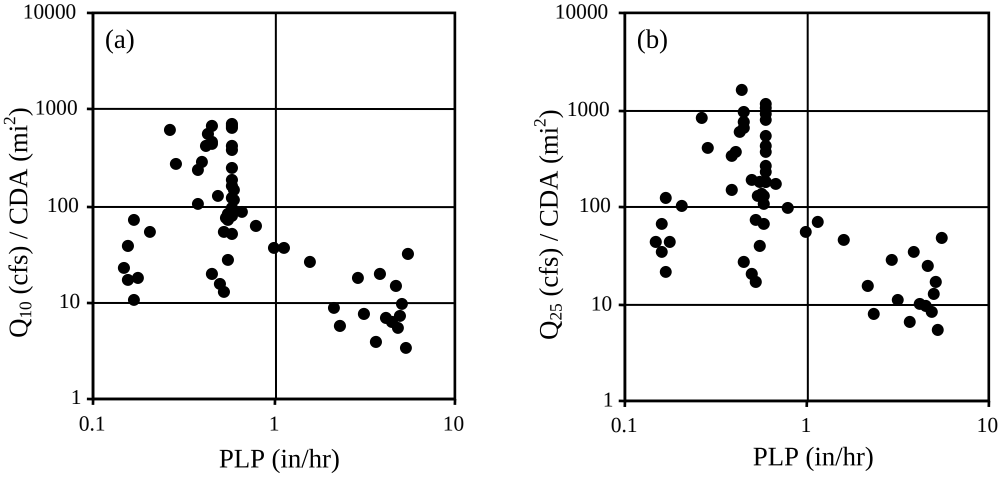
Figures 3.19: The compactness ratio (CR) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Upper Republican region.

3.5.4 Soil Characteristics

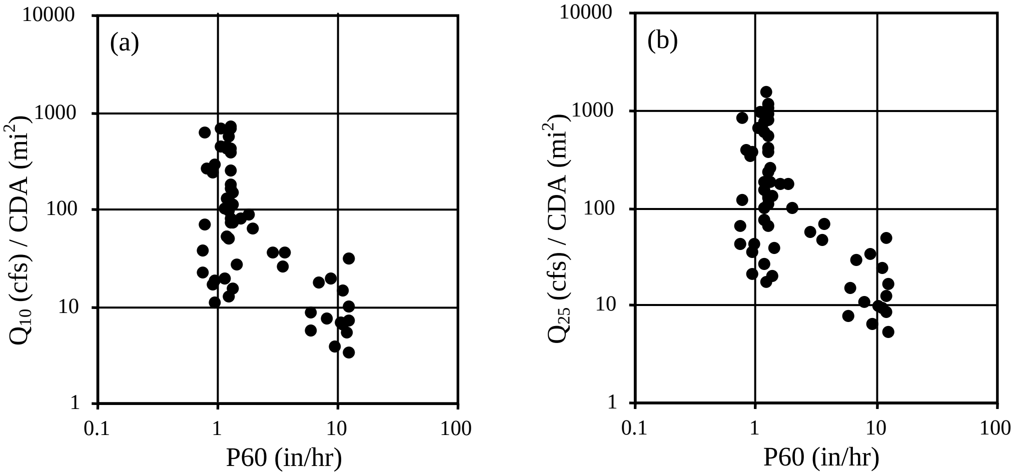
Soils within the watershed can affect the peak flow magnitude through infiltration rates and typical soil slopes. The three soil characteristics investigated for use in the regression equations are the permeability of the least permeable layer, the average permeability rate of a 60-inch soil profile, and the average maximum soil slope.

If the permeability of the least permeable layer (PLP) is low, infiltration is also low, leading to higher peak flows. Figures 3.20a and 3.20b show a definite correlation between peak flows and PLP in the Northeastern region. In some cases, the large spatial variability between soil types and the limited resolution of soils data can result in poor correlation between PLP and peak flow magnitudes. Also, the range of variation of permeabilities within a region may not be large enough to result in a strong correlation. Again, high permeability rates should decrease peak flow magnitudes, while low permeabilities should increase peak flows.

The average permeability rate of a 60-inch soil profile (P60) can also affect peak flow magnitudes. Figures 3.21a and 3.21b show an inverse relationship between P60 and the unit discharge for the Northeastern region. The expected result is that with decreasing permeability the peak discharge increases. However, incorrect estimation of the dominant soil type sometimes leads to poor correlation between P60 and the unit discharge. Furthermore, since the state has been regionalized, the influence of soil type on peak flow is significantly reduced (soil types vary more strongly between regions than within regions).

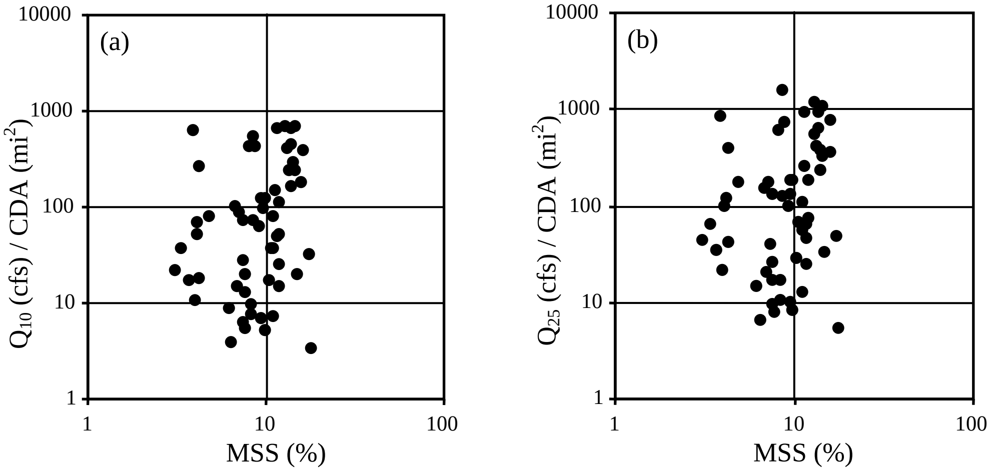


Figures 3.20: Permeability of the least permeable layer (PLP) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Northeastern region.



Figures 3.21: Average permeability rate of a 60-inch soil profile (P60) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Northeastern region.

The average maximum soil slope (MSS) is also a statistic that represents the soil type. An increase in soil slope is expected to increase the discharge per unit area. Figures 3.22a and 3.22b show the relationship between soil slope and unit discharge for the Northeastern region. The data is highly clustered with little apparent correlation.



Figures 3.22: Average maximum soil slope (MSS) vs. the unit discharge for the (a) 10-year and (b) 25-year peak flows in Northeastern region.

The graphical comparisons demonstrate how basin characteristics are related to peak discharges. Drainage area has the highest statistical correlation with peak flows. Also, increased relief causes increases in peak flow magnitudes with increasing slopes. Shape quantifications do not appear to have a strong correlation with peak discharge. Large soil permeabilities and low soil slopes can decrease peak discharges, but correlation between soil properties and unit discharge is not always strong because the peak flow data has already been regionalized. Graphical comparisons of the basin characteristics and peak discharge only take one basin characteristic into account at a time, and poor correlation does not mean that there is no correlation. For example, Figure 3.21 shows a lot of scatter, but it may be because basin relief dominates the effects of soil permeability. In other words, soil permeability may be important, but not as important as basin relief.

4. REGIONAL EQUATIONS

The regions developed by the USGS WRI 99-4032 were used in the development of the new regression equations. Nebraska was sub-divided into seven hydrologic regions based on geography and hydrology. Within each region a weighted least-squares (WLS) regression model was used to estimate peak streamflows for sites without flow data. WLS accounted for the differences in record lengths of the annual peak streamflows between sites. The basin characteristics used in each equation were selected by minimizing the sum of squares from the model output and using the standard error of estimate (SEE) to evaluate sensitivity. The sum of squares provides an estimate of the difference between the observed value and the predicted value, determined from regression analysis. Final comparisons were done using the SEE, for each hydrologic region.

SPSS 11.5 regression software was used to develop the regression equations. A nonlinear regression method was used when finding a relation between the dependent variables and the set of independent variables. The Log-Pearson Type III frequency discharges are the observed data and, the predicted values were generated using basin characteristics. Basin characteristics that produced the lowest sum of squares were then used in the SEE analysis. The regional analysis eliminated basins with a SEE larger than two log units. When basins with drainage areas of less than 0.5 mi^2 were used to develop regression equations for the complete range of drainage areas, large SEEs were observed. Thus, extremely small basins were excluded from the regression analysis when developing regression equations using all of the available gages. However, the extremely small basins were not excluded when regression equations were developed for small watersheds (<10 mi^2).

Regional equations were developed for all seven hydrologic regions in Nebraska. Equations were also developed for watersheds smaller than 10 mi^2 in six of the seven regions. The separate analysis of small watersheds is expected to improve the accuracy of prediction for smaller watersheds. Equations were not developed for small watersheds in the high-permeability region, due to the small amount of regional data. Tables B.1 and B.2 list the stations used in the development of the regression equations. Previous projects excluded some basins less than 1.0 mi^2 because of the low-resolution topographic maps used in the regression analysis. The 7.5-minute DEM data made it possible to delineate watersheds with drainage areas of less than 1.0 mi^2. The small watersheds were then used in the regression analysis to strengthen peak flow estimates.

Regional regression analysis, using WLS regression and data from 273 gaging stations, were used to develop equations for each hydrologic region. Each region in Nebraska had annual peak flow estimates for recurrence intervals of 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-years. The recurrence interval discharges were designated by Q_2, Q_5, Q_{10}, Q_{25}, Q_{50}, Q_{100}, Q_{200}, and Q_{500} respectively. In this chapter, tables of regression equations and statistics are given along with a discussion of each region. The tables also give the ranges of variables used to develop the equations. The regions overlap in some instances, and in some cases multiple equations can be used to predict peak flows. The tables in this chapter also provide the standard error of estimate (SEE) for each regression equation. The SEE parameter is based on the model error, which will only change if the equation changes.

4.1 Big Blue Region

The Big Blue Region contains the Big Blue River basin, including parts of southeastern Nebraska and northeastern Kansas. The Big Blue Region contains the Big Sandy Creek, Turkey Creek and the Little Blue River drainage areas. Two sets of equations were developed for the Big Blue Region to improve the accuracy of prediction for smaller watersheds. Table 4.1 gives the range of contributing drainage areas used in the regression analysis.

Table 4.1: Number of stream-gaging stations and average length of record for stations in the Big Blue Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	2	24.5
Less than 10	7	25.0
Less than 25	12	23.8
Less than 50	15	25.7
Less than 100	20	24.8
Less than 300	24	27.0
Less than 1,000	32	29.0
Less than 10,000	43	40.0

4.1.1 Small Basin Analysis

Equations developed for watersheds less than 10 mi^2 are given in Table 4.2. The regression equations are based on 8 stations with at least 11 years of record. The statistically relevant basin characteristics used in the regression are the contributing drainage area (CDA), main channel slope (MCS), and permeability of the least permeable layer (PLP). The main channel slope is data-scale dependent. Basin characteristics that are data-scale dependent are influenced by the resolution of the topography. Drainage areas ranged from 0.90 to 10.1 mi^2. Four gaging stations had peak flow records of greater than 30 years, including one basin with an area of less than 1.0 mi^2. The other four stations had less than 12 years of peak flow data. In the equation, the CDA exponent remained nearly constant with increasing recurrence intervals, while the MCS exponent increased. The PLP became less significant with a decreasing negative exponent.

The SEE was calculated for each return period and compared to WRI 99-4032. The SEE for each return period is higher than the USGS Big Blue Region equations. There are a few possible reasons for the discrepancy. First, the USGS equations used up to five basin attributes, compared to the three used in the new equations. Adding more basin attributes adds complexity to the equations, and because only a small number of data are used in the regression, the additional complexity does not necessarily result in improved predictive accuracy. Second, the smallest watershed used for the USGS equations was 2.0 mi^2, compared to 0.9 mi^2 for the new equations. The addition of watersheds smaller than 2.0 mi^2 will aid in the peak flow prediction of small watersheds. The improved resolution of the DEMs allowed for the extraction of basin characteristics in smaller watersheds that were not possible with previous maps.

Table 4.2: Peak-flow equations for basins less than 10 mi^2 in the Big Blue Region.

Parametric Equation	*SEE (log_{10} units)*	*SEE (percent)*
$Q_2 = 13\ CDA^{0.138}\ MCS^{0.243}\ PLP^{-1.276}$	0.253	64%
$Q_5 = 29\ CDA^{0.180}\ MCS^{0.308}\ PLP^{-1.151}$	0.159	38%
$Q_{10} = 45\ CDA^{0.193}\ MCS^{0.351}\ PLP^{-1.045}$	0.123	29%
$Q_{25} = 75\ CDA^{0.199}\ MCS^{0.406}\ PLP^{-0.896}$	0.093	22%
$Q_{50} = 109\ CDA^{0.196}\ MCS^{0.441}\ PLP^{-0.776}$	0.080	19%
$Q_{100} = 154\ CDA^{0.186}\ MCS^{0.473}\ PLP^{-0.665}$	0.073	17%
$Q_{200} = 215\ CDA^{0.173}\ MCS^{0.502}\ PLP^{-0.555}$	0.070	16%
$Q_{500} = 432\ CDA^{0.134}\ MCS^{0.519}\ PLP^{-0.320}$	0.075	17%
Applicable ranges of variables: CDA 0.90 – 10.09 MCS 12.2 – 46.2 PLP 0.14 – 0.42	8 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); MCS, main channel slope (ft/mile); PLP, permeability of least permeable layer (in/hr)]. **Note: MCS is data-scale dependent.**

4.1.2 Large Basin Analysis

A second group of equations was developed for a larger range of basin areas in the Big Blue River region and are given in Table 4.3. The regression equations are based on 41 stations with at least 11 years of record each. The statistically relevant basin characteristics used in the regression are the contributing drainage area (CDA), shape factor (SF), and maximum soil slope (MSS). The shape factor is data-scale dependent, and the drainage areas ranged from 0.90 to 4,370 mi^2. Over 70 percent of the gaging stations had at least 30 years of data. Six of the stations had historical records of greater than 50 years. The significance of CDA decreased with larger return periods, while the importance of MSS increased. The SF exponent was most significant for smaller peak flows.

The SEE is higher compared to the smaller watersheds and the USGS equations. A majority of the basins included in the regression are greater than 100 mi^2, and the equations provide good estimates for large drainage areas, but they do not represent basins less than 10 mi^2 well. The USGS equations have significantly lower SEE, but include fewer drainage areas and use additional variables. The USGS equations also use climatic and data-scale dependent basin characteristics that differ from those of the new equations. Figure 4.1 shows the locations of the gaging stations used in the regression analysis. Figures like Figure 4.1 show how the stream gages are distributed throughout the region, and give some idea of how well regression equations should be expected to perform when applied to a particular site in the region.

Table 4.3: Peak-flow equations for the Big Blue Region

Parametric Equation	*SEE (log$_{10}$ units)*	*SEE (percent)*
$Q_{2} = 46\ \text{CDA}^{0.831}\ \text{SF}^{-0.954}\ \text{MSS}^{0.744}$	0.250	63%
$Q_{5} = 174\ \text{CDA}^{0.687}\ \text{SF}^{-0.716}\ \text{MSS}^{0.661}$	0.176	42%
$Q_{10} = 296\ \text{CDA}^{0.603}\ \text{SF}^{-0.528}\ \text{MSS}^{0.664}$	0.160	38%
$Q_{25} = 418\ \text{CDA}^{0.518}\ \text{SF}^{-0.287}\ \text{MSS}^{0.721}$	0.163	39%
$Q_{50} = 453\ \text{CDA}^{0.470}\ \text{SF}^{-0.121}\ \text{MSS}^{0.793}$	0.170	41%
$Q_{100} = 444\ \text{CDA}^{0.436}\ \text{SF}^{0.024}\ \text{MSS}^{0.886}$	0.180	43%
$Q_{200} = 413\ \text{CDA}^{0.412}\ \text{SF}^{0.141}\ \text{MSS}^{0.987}$	0.190	46%
$Q_{500} = 349\ \text{CDA}^{0.391}\ \text{SF}^{0.275}\ \text{MSS}^{1.147}$	0.211	51%
Applicable ranges of variables: CDA 0.90 – 4,370 SF 3.31 – 53.05 MSS 1.9 – 14.5	41 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); SF, shape factor (dimensionless); MSS, maximum soil slope (%)]. **Note: SF is data-scale dependent.**

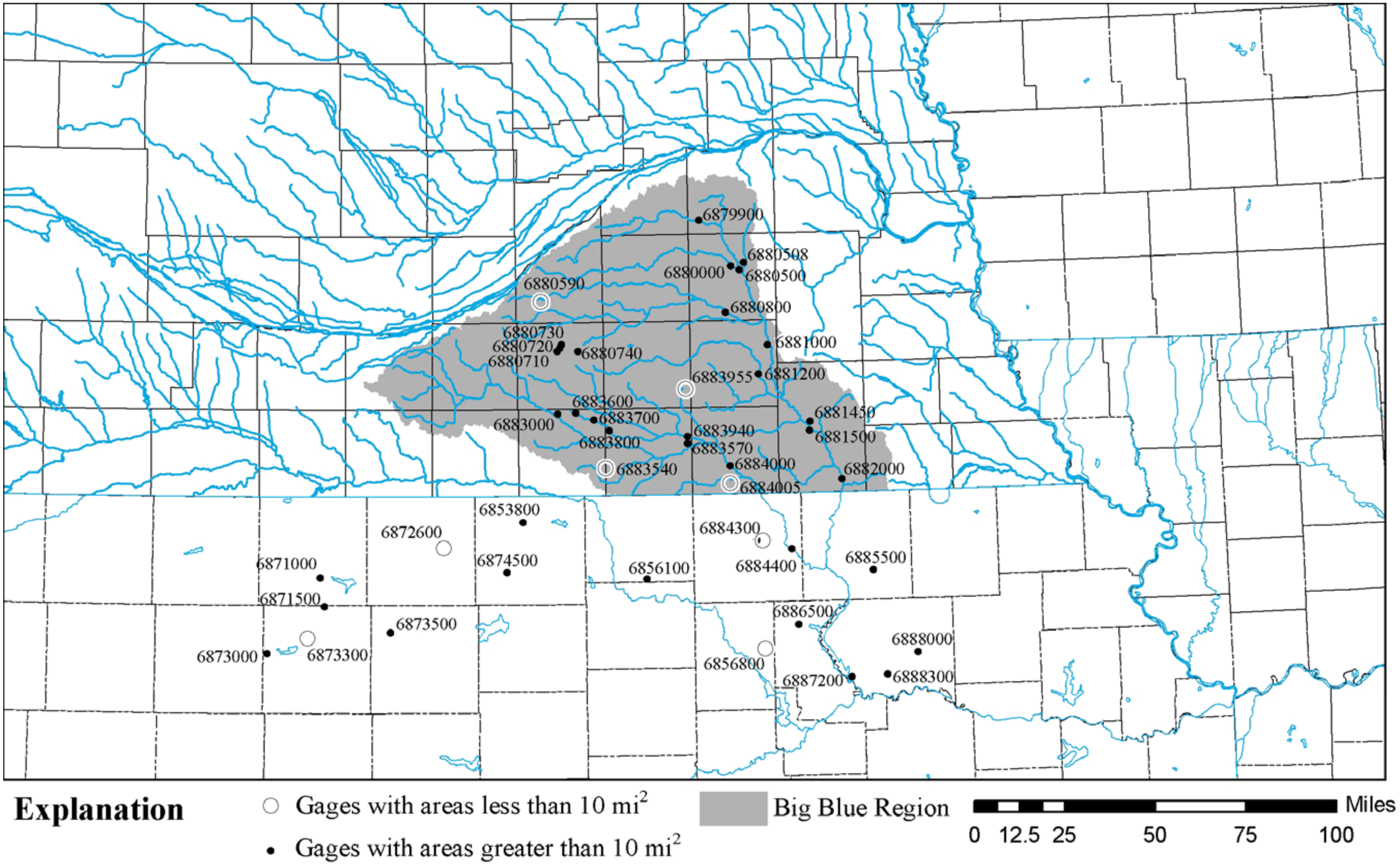


Figure 4.1: Location of streamflow gaging stations used in the Big Blue regression analysis.

4.2 Eastern Region

This region includes the Missouri River tributaries, in eastern Nebraska, northeastern Kansas, northwestern Missouri and western Iowa. The region includes the Missouri tributaries south of Omaha Creek in Nebraska. It also includes the Salt Creek watershed and low lying areas adjacent to the Platte River. The Nemaha River and the Missouri River tributaries located in Missouri were also used. Two sets of equations were developed for the Eastern Region to improve the accuracy of prediction for smaller watersheds. Table 4.4 gives the range of contributing drainage areas used in the regression analysis.

Table 4.4: Number of stream-gaging stations and average length of record for stations in the Eastern Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	9	24.9
Less than 10	21	28.8
Less than 25	29	35.9
Less than 50	34	44.6
Less than 100	38	36.5
Less than 300	47	58.1
Less than 1,000	51	54.1
Less than 10,000	53	51.2

4.2.1 Small Basin Analysis

Equations for the Eastern Region developed for watersheds with drainage areas of less than 10 mi^2 are given in Table 4.5. The equations are based on 21 stations with at least 11 years of record. The statistically relevant basin characteristics used in the regression are the contributing drainage area (CDA), the basin slope (BS), and the average permeability of a 60-inch soil profile (P60), where the basin slope is data-scale dependent and the drainage areas ranged from 0.42 to 10.3 mi^2. Fourteen gaging stations had at least 25 years of peak flow record. There were five stations with areas less than 1.6 mi^2 with 29 years of peak flow data. The exponents for the CDA and the BS decreased with an increasing return period. Also, the P60 exponent remains relatively constant over the range of return periods.

The USGS standard error of estimate statistic is on average higher than for the current set of regression equations, but return period intervals greater than 10-years gave similar SEE results. The USGS regression equations are similar in form to the small basin equations. The main differences are the number of stations used for the regression and the basin slope (different scales were used for the two sets of regression equations). Also, the smallest watershed the USGS used was 1.6 mi^2, compared to five less than 1.6 mi^2 in the current set of equations. Compared to the USGS exponents, trends for the basin attribute exponents are the same for increasing recurrence interval. The additional smaller watersheds should improve the peak flow prediction for the eastern region.

Table 4.5: Peak-flow equations for basins less than 10 mi^2 in the Eastern Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 8.1\ CDA^{0.613}\ BS^{0.589}\ P60^{-0.499}$	0.101	24%
$Q_5 = 32\ CDA^{0.586}\ BS^{0.489}\ P60^{-0.613}$	0.095	22%
$Q_{10} = 65\ CDA^{0.579}\ BS^{0.434}\ P60^{-0.629}$	0.100	23%
$Q_{25} = 138\ CDA^{0.573}\ BS^{0.372}\ P60^{-0.625}$	0.109	25%
$Q_{50} = 227\ CDA^{0.567}\ BS^{0.330}\ P60^{-0.606}$	0.116	27%
$Q_{100} = 365\ CDA^{0.563}\ BS^{0.287}\ P60^{-0.586}$	0.123	29%
$Q_{200} = 556\ CDA^{0.563}\ BS^{0.249}\ P60^{-0.551}$	0.129	30%
$Q_{500} = 1008\ CDA^{0.552}\ BS^{0.190}\ P60^{-0.517}$	0.137	32%
Applicable ranges of variables: CDA 0.42 – 10.3 BS 143 – 641 P60 0.44 – 1.32	21 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); P60, permeability of 60-inch profile (in/hr)]. **Note: BS is data-scale dependent.**

4.2.2 Large Basin Analysis

Equations were also developed in the eastern region using data from all of the gaging stations (including large basins). The regression equations are based on 51 stations with 11 or more years of record. Table 4.6 lists the regression equations and SEE statistic. The statistically relevant basin characteristics used in the regression are the contributing drainage area (CDA), basin slope (BS), and the permeability of the least permeable layer (PLP). Drainage areas ranged from 0.70 to 1,563 mi^2 for each recurrence interval. Two-thirds of the gaging stations have at least 25 years of record. Thirteen stations had historical records of greater than 50 years.

The SEE is higher than both the USGS and small watershed regression equations. The eastern region analysis included more gaging stations and improved accuracy DEMs. A majority of the watersheds are smaller than 100 mi^2, and include a wide range of basin characteristics. Because of the quality of the stations used to develop the regression equations, the Eastern region regression equations should provide reasonable discharge estimates. The locations of the gaging stations used are shown in Figure 4.2.

Table 4.6: Peak-flow equations for the Eastern Region.

Parametric Equation	*SEE (log$_{10}$ units)*	*SEE (percent)*
$Q_2 = 1.6\ CDA^{0.661}\ BS^{0.673}\ PLP^{-0.518}$	0.201	49%
$Q_5 = 21\ CDA^{0.550}\ BS^{0.450}\ PLP^{-0.527}$	0.119	28%
$Q_{10} = 63\ CDA^{0.495}\ BS^{0.375}\ PLP^{-0.531}$	0.104	24%
$Q_{25} = 196\ CDA^{0.441}\ BS^{0.293}\ PLP^{-0.521}$	0.119	28%
$Q_{50} = 425\ CDA^{0.408}\ BS^{0.230}\ PLP^{-0.510}$	0.139	33%
$Q_{100} = 905\ CDA^{0.381}\ BS^{0.163}\ PLP^{-0.491}$	0.160	38%
$Q_{200} = 1801\ CDA^{0.357}\ BS^{0.097}\ PLP^{-0.513}$	0.187	45%
$Q_{500} = 3903\ CDA^{0.336}\ BS^{0.015}\ PLP^{-0.480}$	0.196	48%
Applicable ranges of variables: CDA 0.70 – 1,563 BS 93.7 – 640.7 PLP 0.13 – 0.60	51 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); PLP, permeability of least permeable layer (in/hr)]. **Note: BS is data-scale dependent.**

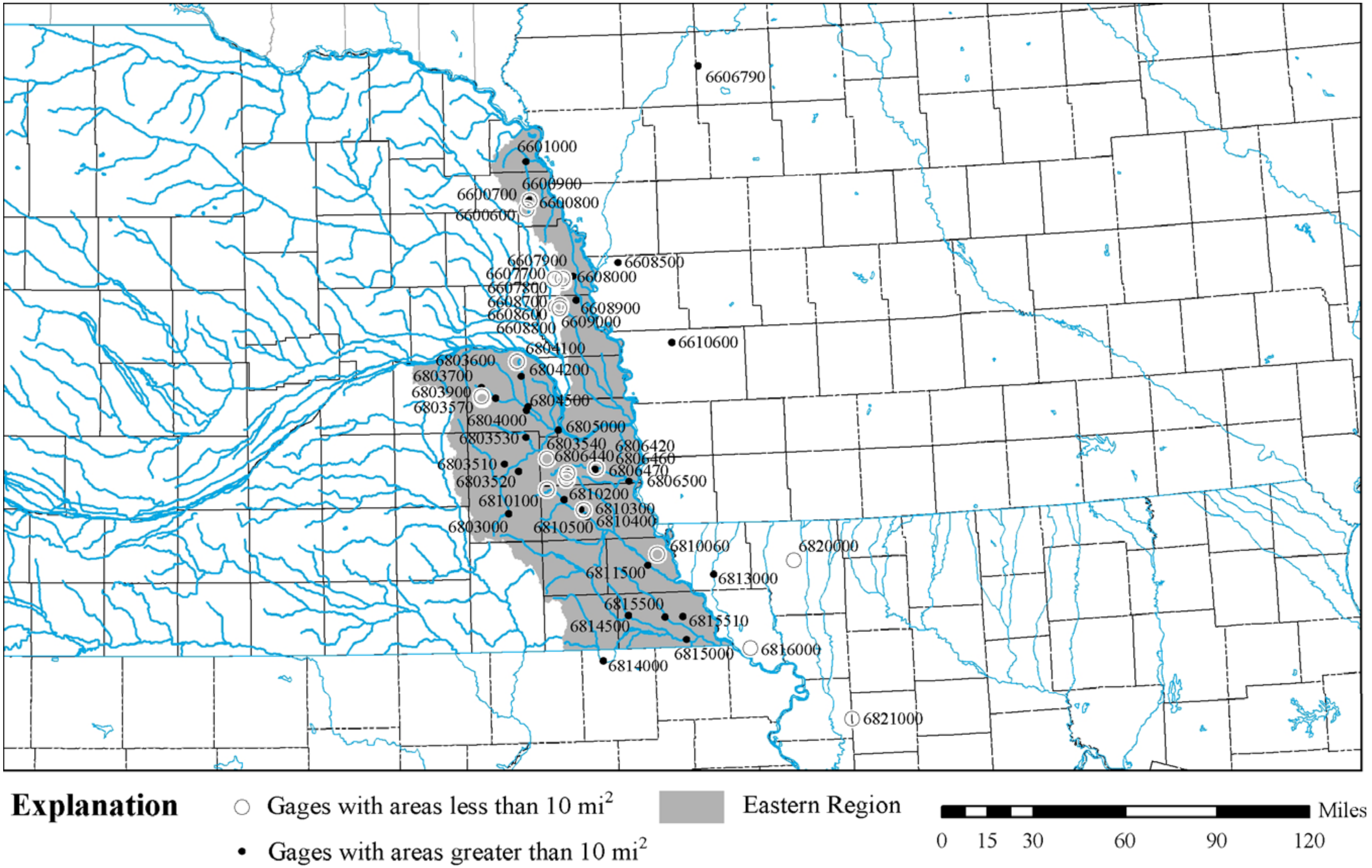


Figure 4.2: Location of streamflow gaging stations used in the Eastern regression analysis.

4.3 Northeastern Region

This region includes most of northeastern Nebraska, parts of southeastern South Dakota, and parts of northwestern Iowa. The majority of the area in Nebraska is composed of the Elkhorn River drainage basin. It also includes basins east of the North Loup River and areas north of the lower Platte River. Missouri River tributaries from the Platte River to the mouth of the Niobrara River were also included. The Northeastern region uses some of the same gages used in the Eastern region regression equations. A majority of the stations in the Northeastern region are located on the eastern edge of Nebraska. Because of the lack of flow data and increased variability on the western side of the region, the equations become less accurate in the west. Two sets of equations were developed for the Northeastern Region to improve the accuracy of prediction for smaller basins. Table 4.7 gives the range of contributing drainage areas used in the regression analysis.

Table 4.7: Number of stream-gaging stations and average length of record for stations in the Northeastern Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	8	21.5
Less than 10	15	19.6
Less than 25	21	19.7
Less than 50	25	24.4
Less than 100	30	22.1
Less than 300	37	34.8
Less than 1,000	45	35.7
Less than 10,000	50	42.9

4.3.1 Small Basin Analysis

Equations for watersheds with drainage areas of less than 10 mi^2 in the northeastern region are given in Table 4.8. They are based on data from 13 stations with at least 11 years of record. The statistically significant basin characteristics used in the regression are the contributing drainage area (CDA), shape factor (SF), and the permeability of the least permeable layer (PLP), where the shape factor is data-scale dependent. Drainage areas ranged from 0.50 to 9.5 mi^2. Five stations had peak flow records greater than 25 years, including three with areas of less than 1.6 mi^2. Four stations had less than 12 years of flow data.

The USGS standard error of estimate (SEE) statistic is higher than the small basin watershed regression. Standard errors for return periods of greater than 10 years were on average 10% lower than USGS standard errors. The USGS regression included one parameter more than was used in the equations for small basins, and different basin characteristic combinations were used. Also, the smallest watershed used for the USGS regression equations was 1.5 mi^2, compared to three less than 1.5 mi^2 in the present analysis.

Table 4.8: Peak-flow equations for basins less than 10 mi^2 in the Northeastern Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 2406\ CDA^{0.636}\ SF^{-1.478}\ PLP^{0.759}$	0.250	63%
$Q_5 = 3561\ CDA^{0.490}\ SF^{-1.135}\ PLP^{0.631}$	0.159	38%
$Q_{10} = 4480\ CDA^{0.424}\ SF^{-0.961}\ PLP^{0.623}$	0.123	29%
$Q_{25} = 5879\ CDA^{0.359}\ SF^{-0.780}\ PLP^{0.655}$	0.100	23%
$Q_{50} = 7047\ CDA^{0.320}\ SF^{-0.659}\ PLP^{0.692}$	0.100	23%
$Q_{100} = 8335\ CDA^{0.285}\ SF^{-0.556}\ PLP^{0.728}$	0.109	26%
$Q_{200} = 9806\ CDA^{0.255}\ SF^{-0.464}\ PLP^{0.766}$	0.125	29%
$Q_{500} = 12082\ CDA^{0.219}\ SF^{-0.364}\ PLP^{0.808}$	0.150	36%
Applicable ranges of variables: CDA 0.50 – 9.5 SF 2.11 – 6.41 PLP 0.21 – 0.60	13 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); SF, shape factor (dimensionless); PLP, permeability of least permeable layer (in/hr)]. **Note: SF is data-scale dependent.**

4.3.2 Complete Basin Analysis

A second group of equations was developed for the entire range of basin sizes in the Northeastern region and are given in Table 4.9. The regression equations are based on 49 stations with at least 11 years of record. The statistically relevant basin characteristics used in the regression equations are the contributing drainage area (CDA), the basin slope (BS), and the maximum soil slope (MSS). The basin slope is data-scale dependent. Drainage areas used in the regression equations ranged from 0.50 to 5,870 mi^2. Two-thirds of the gaging stations had at least 25 years of peak flow record. Seven of the stations had historical records longer than 50 years.

The SEE statistic is much higher than for the small basin regression, but similar to the USGS analysis. The equations for the northeastern region have a large number of gaging stations, with a wide variation of basin characteristics. This could account for the differences in SEE, for the given regression equations. Figure 4.3 shows the locations of the gaging stations used in the regression analysis. The locations of the stations are biased to the east side of the region; this is especially true for the small basins, and should be considered when using the regression equations.

Table 4.9: Peak-flow equations for the Northeastern Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_{2} = 0.4\ CDA^{0.613}\ BS^{0.958}\ MSS^{0.030}$	0.275	70%
$Q_{5} = 1.5\ CDA^{0.571}\ BS^{0.871}\ MSS^{0.040}$	0.203	50%
$Q_{10} = 3.1\ CDA^{0.550}\ BS^{0.809}\ MSS^{0.090}$	0.178	43%
$Q_{25} = 6.6\ CDA^{0.526}\ BS^{0.733}\ MSS^{0.189}$	0.164	39%
$Q_{50} = 11\ CDA^{0.511}\ BS^{0.678}\ MSS^{0.273}$	0.165	39%
$Q_{100} = 16\ CDA^{0.497}\ BS^{0.629}\ MSS^{0.368}$	0.172	41%
$Q_{200} = 23\ CDA^{0.482}\ BS^{0.574}\ MSS^{0.465}$	0.192	46%
$Q_{500} = 40\ CDA^{0.460}\ BS^{0.504}\ MSS^{0.589}$	0.228	56%
Applicable ranges of variables: CDA 0.50 – 5,870 BS 60.2 – 738 MSS 3.1 – 17.6	49 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); MSS, maximum soil slope (%)]. **Note: BS is data-scale dependent.**

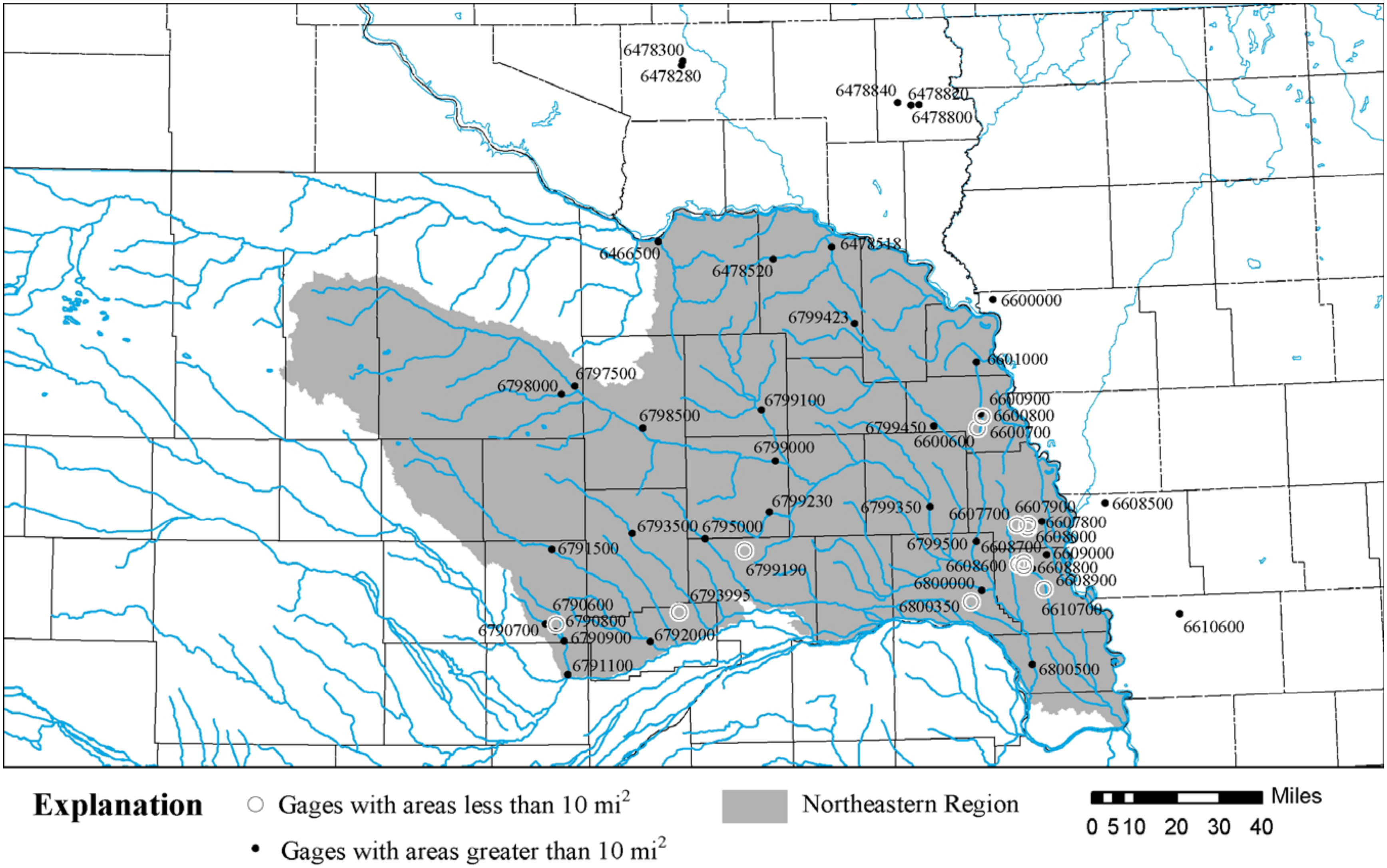


Figure 4.3: Location of streamflow gaging stations used in the Northeastern regression analysis.

4.4 Central and South-Central Region

The Central and South-Central Region includes basins south and east of the central sandhills, which contain tributaries of the middle Platte, Loup and middle Republican Rivers. Basins in this region generally have a P60 of less than 4 in/hr. The region overlaps somewhat with the northeastern region. Stations in the Republican River basin downstream of the Harlan County Reservoir are also used in the regression. Again, two sets of equations were developed for the Central and South-Central Region. Table 4.10 gives the range of contributing drainage areas used in the regression analysis.

Table 4.10: Number of stream-gaging stations and average length of record for stations in the Central and South-Central Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	7	24.1
Less than 10	14	24.6
Less than 25	24	24.2
Less than 50	32	22.5
Less than 100	38	21.5
Less than 300	42	22.7
Less than 1,000	46	30.1
Less than 10,000	47	42.0

4.4.1 Small Basin Analysis

Equations developed for watersheds with areas of less than 10.6 mi^2 are given in Table 4.11. The regression equations are based on 11 stations with at least 11 years of record. Basin characteristics used in the regression are the contributing drainage area (CDA), basin slope (BS), relative relief (RR) and shape factor (SF). The basin slope, relative relief and shape factor are data-scale dependent. Drainage areas ranged from 0.80 to 10.6 mi^2 in size. Two stations with areas of less than 1.5 mi^2 had more than 27 years of peak flow record; the smallest watershed had a drainage area of 0.8 mi^2 and 40 years of record. Over half the stations had at least 25 years of peak flow record. The regression equations have a decreasing constant with an increasing exponent for the CDA. The CDA exponent for the 2- and 5-year recurrence interval should not be trusted. The equations are developed from a wide range of areas, but there is large variability between basin area and peak discharge for small return intervals. For the 2- and 5-year recurrence interval the regression equations developed for larger basins is recommended. The RR attribute becomes significant at recurrence intervals greater than 200-years.

For the equations for small basins, the standard error of estimate (SEE) was lower than for the USGS equations for return intervals of greater than 10-years. This set of equations is the only set that uses two different slope characteristics: BS and RR. The equations for return periods greater than 200-years used RR. Due to the limited number of gaging stations and years of record, the largest recurrence intervals had high SEE.

Table 4.11: Peak-flow equations for basins with drainage areas of less than 10 mi^2 in the Central and South-Central Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 74528\ CDA^{-0.162}\ BS^{-0.807}\ SF^{-1.008}$	0.363	100%
$Q_5 = 15457\ CDA^{0.039}\ BS^{-0.363}\ SF^{-1.059}$	0.238	59%
$Q_{10} = 7064\ CDA^{0.139}\ BS^{-0.159}\ SF^{-1.018}$	0.172	41%
$Q_{25} = 3166\ CDA^{0.218}\ BS^{0.030}\ SF^{-0.877}$	0.141	33%
$Q_{50} = 1767\ CDA^{0.260}\ BS^{0.143}\ SF^{-0.695}$	0.148	35%
$Q_{100} = 931\ CDA^{0.290}\ BS^{0.245}\ SF^{-0.450}$	0.172	41%
$Q_{200} = 242\ CDA^{0.430}\ RR^{0.858}$	0.194	47%
$Q_{500} = 174\ CDA^{0.422}\ RR^{1.141}$	0.214	53%
Applicable ranges of variables: CDA 0.80 – 10.6 BS 73 – 925 SF 2.44 – 7.51 RR 4.6 – 21.9	11 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); SF, shape factor (dimensionless); RR, relative relief (ft/mile)]. **Note: BS, SF, and RR are data-scale dependent.**

4.4.2 Complete Basin Analysis

A second group of equations was developed for the entire set of gaging stations in the Central and South-Central region and are given in Table 4.12. The regression equations are based on 46 stations with at least 11 years of record. The statistically relevant basin characteristics used in the regression are the contributing drainage area (CDA), main channel slope (MCS), and the average permeability of 60-inch soil profile (P60), where the main channel slope is data-scale dependent. Drainage areas of 0.60 to 1,590 mi^2 were used in the regression. Over 60% of the gaging stations used in the regression had over 25 years of record. Two stations have peak flow records of over 50 years.

The SEE statistic is much higher for the equations given in Table 4.12 than for either the equations for small watersheds or the USGS regression equations. There is a lot of variability in the basin characteristics and the annual peak flow discharges. One reason for the large SEE was because exponents were limited to powers of less than 1.5. The USGS equations have exponents as large as 3.0, which do improve the SEE. But, large exponents give most of the weight to individual basin characteristics, leading to potential prediction errors. Figure 4.4 shows the locations of the gaging stations used in the regression analysis.

Table 4.12: Peak-flow equations for the Central and South-Central Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 40\ CDA^{0.376}\ MCS^{0.158}\ P60^{0.871}$	0.470	149%
$Q_5 = 189\ CDA^{0.280}\ MCS^{0.036}\ P60^{1.018}$	0.369	103%
$Q_{10} = 346\ CDA^{0.246}\ MCS^{0.020}\ P60^{1.111}$	0.342	93%
$Q_{25} = 577\ CDA^{0.221}\ MCS^{0.036}\ P60^{1.219}$	0.339	92%
$Q_{50} = 734\ CDA^{0.213}\ MCS^{0.068}\ P60^{1.292}$	0.349	95%
$Q_{100} = 871\ CDA^{0.209}\ MCS^{0.107}\ P60^{1.351}$	0.364	101%
$Q_{200} = 980\ CDA^{0.210}\ MCS^{0.153}\ P60^{1.426}$	0.383	108%
$Q_{500} = 1041\ CDA^{0.218}\ MCS^{0.226}\ P60^{1.504}$	0.408	119%
Applicable ranges of variables: CDA 0.60 – 1,590 MCS 4.1 – 46.7 P60 1.11 – 4.28	46 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); MCS, main channel slope (ft/mile); P60, permeability of 60-inch profile (in/hr)]. **Note: MCS is data-scale dependent.**

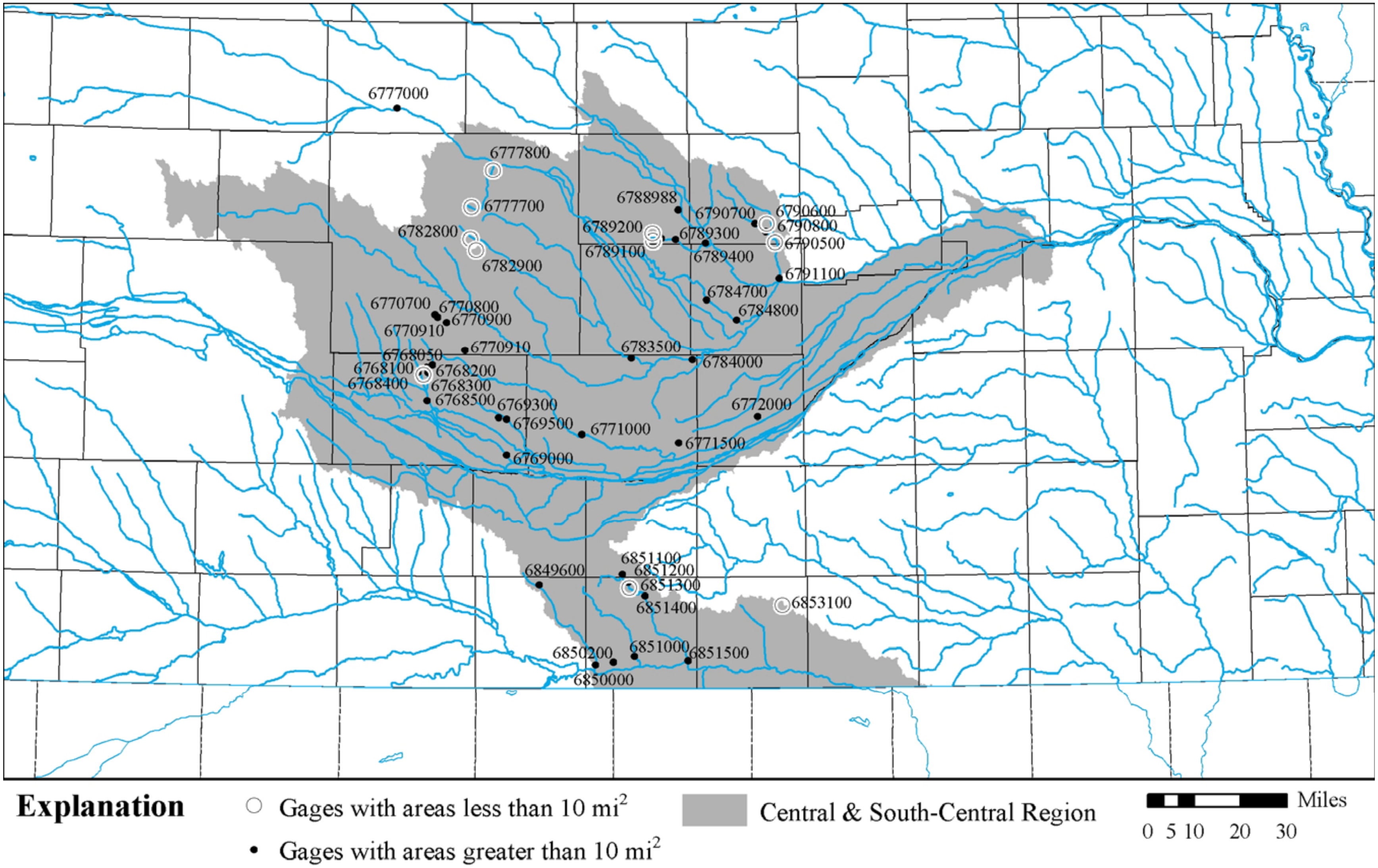


Figure 4.4: Location of streamflow gaging stations used in the Central and South Central regression analysis.

4.5 Upper Republican Region

The Upper Republican Region includes the Republican River upstream of the Harlan County Reservoir. It covers a large portion of southwestern Nebraska, northwestern Kansas and northeastern Colorado. Two sets of equations were developed for the Upper Republican Region. Table 4.13 gives the range of contributing drainage areas used in the regression analysis.

Table 4.13: Number of stream-gaging stations and average length of record for stations in the Upper Republican Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	3	35.0
Less than 10	7	33.6
Less than 25	12	35.3
Less than 50	15	37.6
Less than 100	21	33.5
Less than 300	25	39.8
Less than 1,000	29	40.3
Less than 10,000	39	48.8

4.5.1 Small Basin Analysis

Equations developed for watersheds of less than 10.0 mi^2 are given in Table 4.14. The regression equations are based on 7 stations with at least 11 years of record. Basin characteristics used in the regression are the contributing drainage area (CDA), main channel slope (MCS), and the average permeability of the 60-inch soil profile (P60), where the main channel slope is data-scale dependent. Drainage areas ranged from 0.70 to 9.1 mi^2. Six of the stations had over 25 years of peak flow record. Two stations with drainage areas of less than 1.5 mi^2 had over 27 years of record. Trends in the regression equations show the exponents for CDA and MCS increasing for larger return periods. But, the regression constant and the P60 exponent decrease with increasing recurrence interval.

For the small watershed analysis, the standard error of estimate (SEE) was lower than the USGS equations on average by 20%. The small watershed analyses used a smaller number of stations, which resulted in a lower SEE than for the USGS equations. The smallest watershed used in the development of the USGS equations was 6.78 mi^2, while three stations in the new set of equations had smaller drainage areas.

Table 4.14: Peak-flow equations for basins less than 10 mi^2 in the Upper Republican Region.

Parametric Equation	*SEE (log$_{10}$ units)*	*SEE (percent)*
$Q_2 = 278\ CDA^{0.267}\ MCS^{-0.184}\ P60^{-0.996}$	0.287	74%
$Q_5 = 115\ CDA^{0.523}\ MCS^{0.257}\ P60^{-1.092}$	0.165	39%
$Q_{10} = 57\ CDA^{0.667}\ MCS^{0.550}\ P60^{-1.062}$	0.116	27%
$Q_{25} = 22\ CDA^{0.816}\ MCS^{0.909}\ P60^{-0.974}$	0.092	21%
$Q_{50} = 11\ CDA^{0.912}\ MCS^{1.180}\ P60^{-0.879}$	0.103	24%
$Q_{100} = 4.6\ CDA^{1.001}\ MCS^{1.460}\ P60^{-0.769}$	0.127	30%
$Q_{200} = 1.8\ CDA^{1.090}\ MCS^{1.759}\ P60^{-0.642}$	0.159	38%
$Q_{500} = 0.3\ CDA^{1.246}\ MCS^{2.276}\ P60^{-0.406}$	0.224	55%
Applicable ranges of variables: CDA 0.70 – 9.1 MCS 11.7 – 87.7 P60 1.29 – 13.01	7 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); MCS, main channel slope (ft/mile); P60, permeability of 60-inch profile (in/hr)]. **Note: MCS is data-scale dependent.**

4.5.2 Complete Basin Analysis

A second group of equations was developed for the entire range of basin areas in the Upper Republican region and are given in Table 4.15. The regression equations are based on 36 stations with at least 11 years of record. Basin characteristics used in the regression are the contributing drainage area (CDA), main channel slope (MCS), and the compactness ratio (CR), where the main channel slope and compactness ratio are data-scale dependent. Drainage areas used in the regression ranged from 1.1 to 1,590 mi^2. Over 90% of the stations used have over 25 years of peak flow record. Historical records for ten stations were greater than 50 years. Trends in the regression equations show the exponents for CDA, MCS and CR decreasing with increasing return period. But, the regression constant increases with increasing recurrence interval.

The SEE is higher for the complete analysis than for either the smaller watershed regression or the USGS regression equations. The basin characteristics used are the same ones used in the USGS equations, but results differ in this analysis because two of the variables (MCS and CR) are data-scale dependent. The data-scale dependent variables are affected by the topographic resolution used to delineate them. The high resolution maps used in the present analysis to assess the basin characteristics are considered to be more accurate. Also, three stations were added to the regression dataset compared to the USGS equations. Figure 4.5 shows the locations of the gaging stations used in the regression analysis.

Table 4.15: Peak-flow equations for the Upper Republican Region.

Parametric Equation	*SEE (log_{10} units)*	*SEE (percent)*
$Q_2 = 0.5\ CDA^{0.674}\ MCS^{1.636}\ CR^{-1.018}$	0.266	67%
$Q_5 = 2.8\ CDA^{0.604}\ MCS^{1.467}\ CR^{-0.783}$	0.240	60%
$Q_{10} = 7.2\ CDA^{0.558}\ MCS^{1.348}\ CR^{-0.606}$	0.253	64%
$Q_{25} = 22\ CDA^{0.503}\ MCS^{1.196}\ CR^{-0.409}$	0.286	74%
$Q_{50} = 46\ CDA^{0.467}\ MCS^{1.087}\ CR^{-0.289}$	0.314	83%
$Q_{100} = 91\ CDA^{0.436}\ MCS^{0.986}\ CR^{-0.187}$	0.342	93%
$Q_{200} = 169\ CDA^{0.411}\ MCS^{0.895}\ CR^{-0.103}$	0.369	103%
$Q_{500} = 381\ CDA^{0.375}\ MCS^{0.786}\ CR^{-0.024}$	0.395	113%
Applicable ranges of variables: CDA 1.1 – 4,450 MCS 7.8 – 47.1 CR 1.66 – 8.25	36 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); MCS, main channel slope (ft/mile); CR, compactness ratio (dimensionless)]. **Note: MCS and CR are data-scale dependent.**

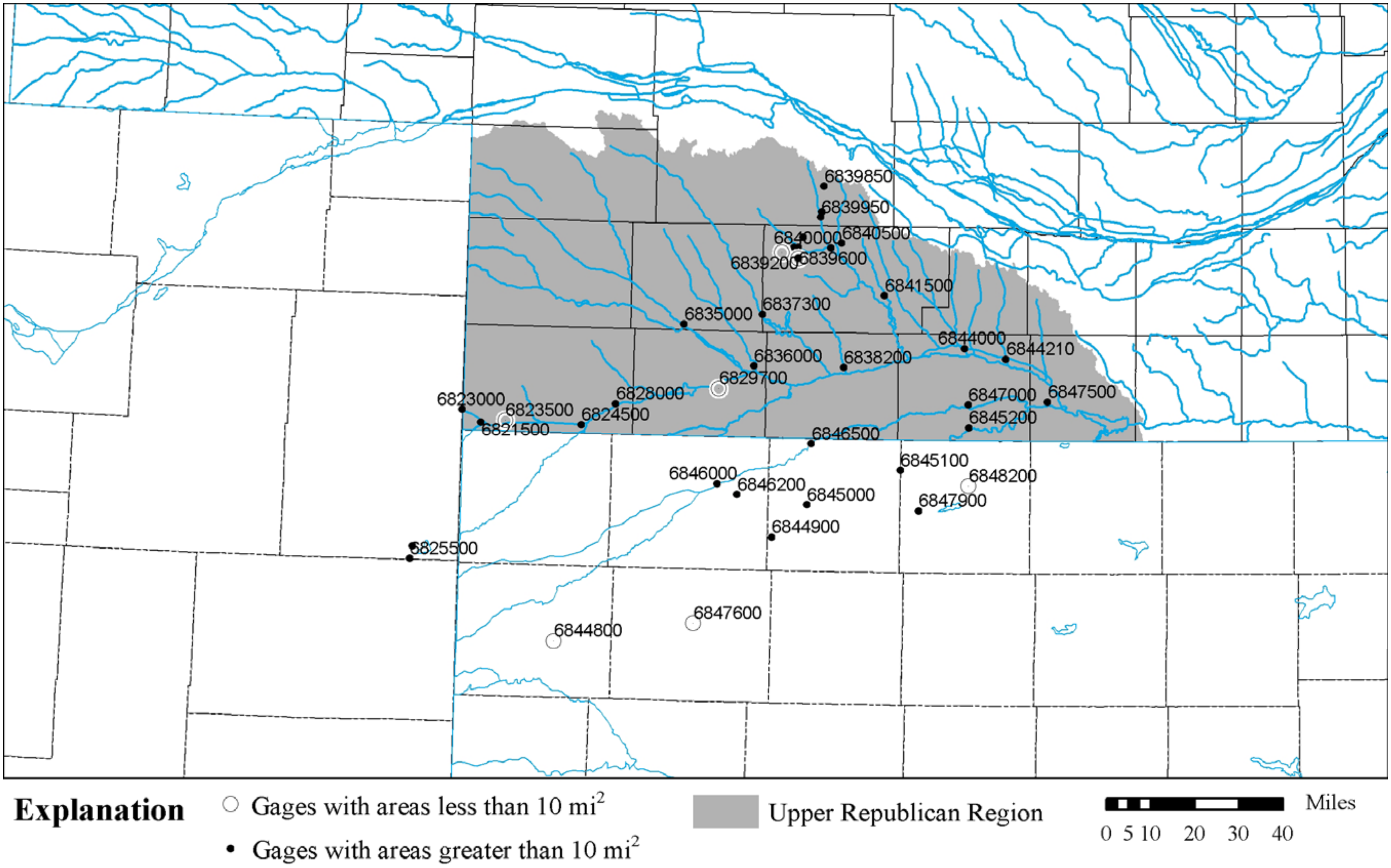


Figure 4.5: Location of streamflow gaging stations used in the Upper Republican regression analysis.

4.6 Northern and Western Region

This region was developed from stations in Wyoming, southern South Dakota and northwestern Nebraska. Major basins include the Cheyenne, White, Niobrara, and North Platte River drainage areas. Also, stations with a P60 greater than 4 in/hr were used if the ratio of CDA to TDA was at least 50 percent. These limitations were used to better represent typical basins within the Northern and western region. Two sets of equations were developed for the Northern and Western Region to improve the accuracy of prediction for smaller watersheds. Table 4.16 gives the range of contributing drainage areas used in the regression analysis.

Table 4.16: Number of stream-gaging stations and average length of record for stations in the Northern and Western Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	3	19.7
Less than 10	11	19.6
Less than 25	17	21.7
Less than 50	20	25.0
Less than 100	21	22.5
Less than 300	25	34.5
Less than 1,000	32	46.1
Less than 10,000	37	42.2

4.6.1 Small Basin Analysis

Equations for basins with drainage areas of less than 10 mi^2 in the Northern and Western region are given in Table 4.17. The regression equations are based on 12 stations with at least 11 years of record. Basin characteristics used in the regression are the contributing drainage area (CDA), basin slope (BS), and maximum soil slope (MSS). The drainage areas ranged from 1.8 to 10.5 mi^2. Five stations had over 25 years of peak flow record. The smallest watershed had the longest flow record used in the regression equations. The regression constant increased with increasing return periods, while the exponents for the CDA, BS and MSS all decreased.

The small watershed equations had SEE values that were lower than for the USGS regression equations for return periods of greater than 25-years. The USGS equations utilized different basin attributes and used one additional basin characteristic in their regression. The USGS equations used four variables in their regression equations, two of which were data-scale dependent.

Table 4.17: Peak-flow equations for basins less than 10 mi^2 in the Northern and Western Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 338\ CDA^{1.135}\ BS^{-1.049}\ MSS^{1.005}$	0.579	221%
$Q_5 = 449\ CDA^{1.610}\ BS^{-1.208}\ MSS^{1.087}$	0.345	94%
$Q_{10} = 658\ CDA^{1.712}\ BS^{-1.141}\ MSS^{1.032}$	0.291	75%
$Q_{25} = 974\ CDA^{1.438}\ BS^{-0.939}\ MSS^{0.907}$	0.208	51%
$Q_{50} = 1352\ CDA^{1.096}\ BS^{-0.737}\ MSS^{0.785}$	0.146	35%
$Q_{100} = 1722\ CDA^{0.815}\ BS^{-0.550}\ MSS^{0.663}$	0.136	32%
$Q_{200} = 2071\ CDA^{0.622}\ BS^{-0.383}\ MSS^{0.529}$	0.163	39%
$Q_{500} = 3637\ CDA^{0.428}\ BS^{-0.161}\ MSS^{0.182}$	0.226	56%
Applicable ranges of variables: CDA 1.8 – 10.5 BS 77.9 – 1085 MSS 7.6 – 46.3	12 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); MSS, maximum soil slope (%)]. **Note: BS is data-scale dependent.**

4.6.2 Complete Basin Analysis

A second group of peak flow frequency relations was developed for the complete range of basin areas. The regression equations in Table 4.18 are based on 36 stations with 11 or more years of record. Where, the contributing drainage area (CDA), the relative relief (RR) and the permeability of least permeable layer (PLP) were used. Drainage areas ranged from 1.8 to 2,157 mi^2. Twenty-four gaging stations used in the regression analysis had more than 25 years of peak flow record. Five stations had historical records of longer than 50 years. The level of significance for RR and PLP remained relatively constant for equations Q_2 through Q_{500}, but the CDA exponent decreased for higher intervals. The regression constant increased with an increasing return period.

The USGS regression equations and the small watershed regression equations produced lower magnitudes of SEE for all return period peak flow predictions when compared to the equations developed for the complete range of drainage areas. The SEE was quite high for the northern and western region and can be related to the high spatial variability, lack of stream flow data and semi arid climate. Figure 4.6 shows the locations of the gaging stations used in the regression analysis.

Table 4.18: Peak-flow equations for the Northern and Western Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_{2} = 144\ CDA^{0.437}\ RR^{-0.668}\ PLP^{-0.122}$	0.445	136%
$Q_{5} = 525\ CDA^{0.372}\ RR^{-0.724}\ PLP^{-0.151}$	0.348	95%
$Q_{10} = 1115\ CDA^{0.310}\ RR^{-0.668}\ PLP^{-0.162}$	0.325	87%
$Q_{25} = 2645\ CDA^{0.225}\ RR^{-0.562}\ PLP^{-0.174}$	0.310	81%
$Q_{50} = 4577\ CDA^{0.174}\ RR^{-0.500}\ PLP^{-0.180}$	0.321	85%
$Q_{100} = 7512\ CDA^{0.131}\ RR^{-0.459}\ PLP^{-0.182}$	0.344	93%
$Q_{200} = 11797\ CDA^{0.094}\ RR^{-0.429}\ PLP^{-0.181}$	0.372	104%
$Q_{500} = 20921\ CDA^{0.050}\ RR^{-0.407}\ PLP^{-0.175}$	0.415	122%
Applicable ranges of variables: CDA 1.8 – 2,157 RR 2.8 – 41.9 PLP 0.10 – 5.26	36 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^{2}); RR, relative relief (ft/mile); PLP, permeability of least permeable layer (in/hr)]. **Note: RR is data-scale dependent.**

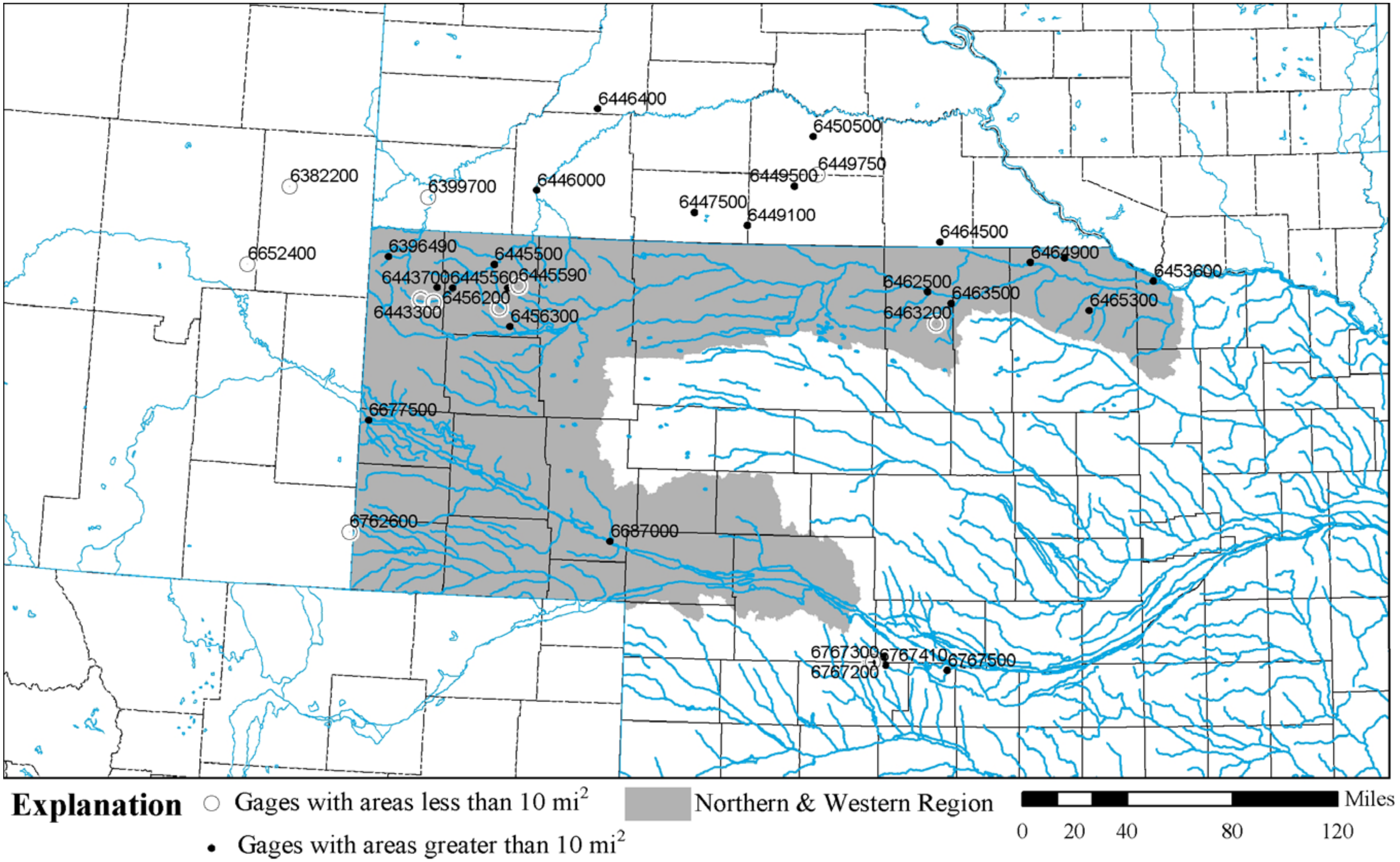


Figure 4.6: Location of streamflow gaging stations used in the Northern and Western regression analysis.

4.7 High Permeability Region

The High Permeability Region consists of drainage basins centrally located in the Nebraskan sandhills. Watersheds with a P60 that was greater than 4 in/hr and basins with large non-contributing drainage areas were used. The region includes a large area of central Nebraska, but not all of it is continuous. There are small areas in Colorado, South Dakota and Wyoming that have similar characteristics. Only one set of equations was developed for the high permeability region, because only a small number of basins had areas of less than 10 mi^2. Table 4.19 gives the range of contributing drainage areas used in the regression analysis.

Table 4.19: Number of stream-gaging stations and average length of record for stations in the High Permeability Region based on Contributing Drainage Area (CDA).

CDA (mi^2)	*Number of stations*	*Average length of record (years)*
Less than 3	0	0.0
Less than 10	2	41.5
Less than 25	3	34.7
Less than 50	7	24.7
Less than 100	13	32.1
Less than 300	22	39.6
Less than 1,000	39	40.0
Less than 10,000	51	41.4

4.7.1 Complete Basin Analysis

Equations for the high permeability region are based on data from 51 stations with at least 11 years of record. Table 4.20 shows the regression equations with the standard error for each return period discharge. Contributing drainage areas ranged from 8.6 to 6,230 mi^2. The most significant basin characteristics were the contributing drainage area (CDA), basin slope (BS) and permeability of the least permeable layer (PLP). The PLP and CDA's significance decreased with an increase in recurrence intervals, while the significance of BS remained relatively constant for all return periods.

The USGS subdivided the high permeability region into two subregions. Both sets of equations have lower SEE, but include two more basin characteristics. On average the high permeability region gave the highest SEE for all recurrence intervals. High SEE can be related to the spatial variability, climatic effects, large non-contributing areas, and the effects of groundwater fed streams. Figure 4.7 shows the locations of the gaging stations used in the regression analysis.

Table 4.20: Peak-flow equations for the High-Permeability Region.

Parametric Equation	***SEE (log$_{10}$ units)***	***SEE (percent)***
$Q_2 = 0.43\ CDA^{0.943}\ BS^{0.192}\ PLP^{0.874}$	0.262	66%
$Q_5 = 1.8\ CDA^{0.961}\ BS^{0.077}\ PLP^{0.665}$	0.286	74%
$Q_{10} = 3.6\ CDA^{0.941}\ BS^{0.072}\ PLP^{0.559}$	0.304	80%
$Q_{25} = 7.4\ CDA^{0.901}\ BS^{0.103}\ PLP^{0.446}$	0.338	91%
$Q_{50} = 12\ CDA^{0.865}\ BS^{0.139}\ PLP^{0.369}$	0.370	103%
$Q_{100} = 19\ CDA^{0.826}\ BS^{0.181}\ PLP^{0.297}$	0.409	119%
$Q_{200} = 28\ CDA^{0.789}\ BS^{0.223}\ PLP^{0.234}$	0.450	139%
$Q_{500} = 47\ CDA^{0.737}\ BS^{0.278}\ PLP^{0.154}$	0.510	172%
Applicable ranges of variables: CDA 8.6 – 6,230 BS 43.2 – 601.4 PLP 1.32 – 5.80	51 stations with 11 or more years of record	

[Q, peak discharge (cfs); CDA, contributing drainage area (mi^2); BS, basin slope (ft/mile); PLP, permeability of least permeable layer (in/hr)]. **Note: BS is data-scale dependent.**

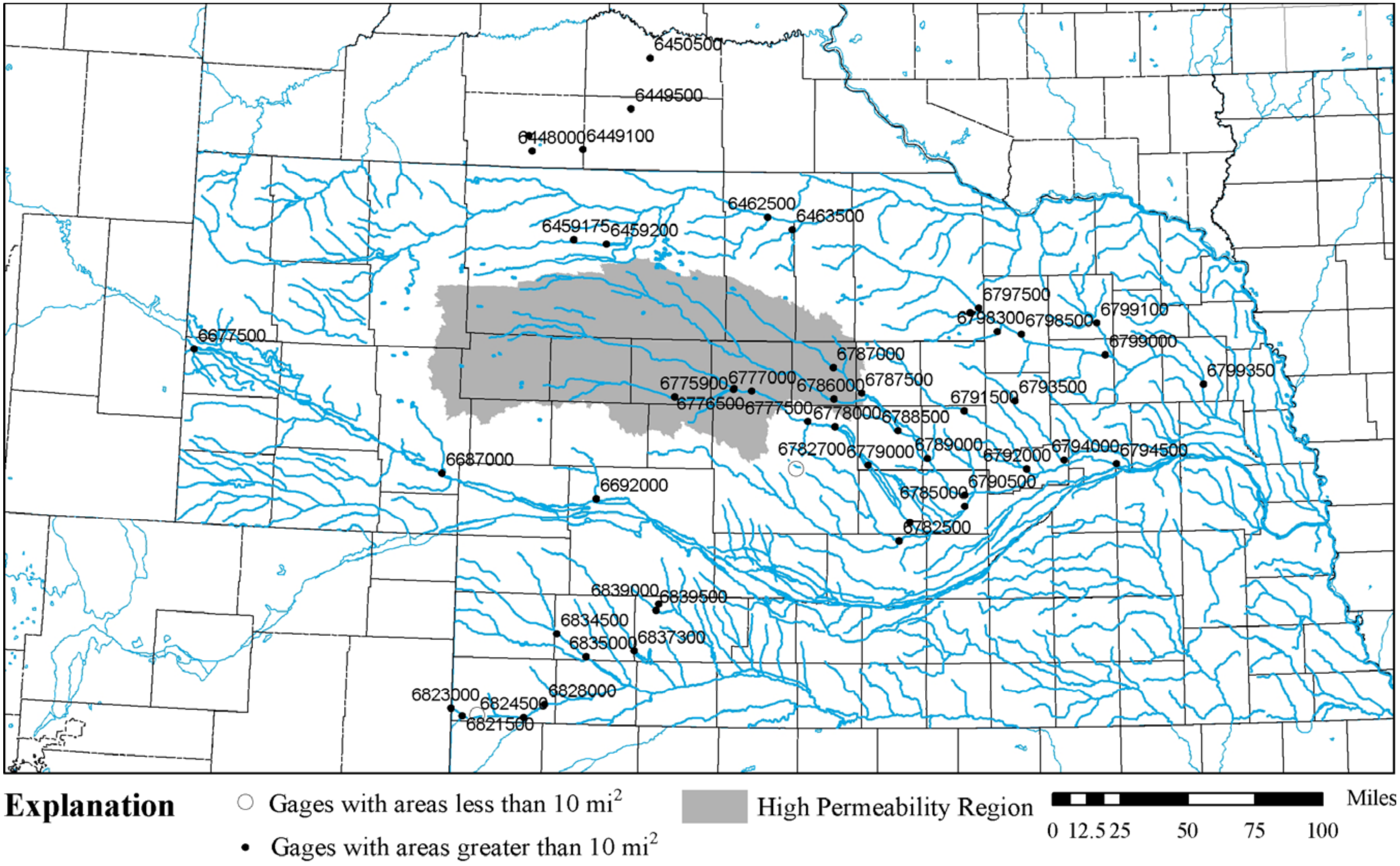


Figure 4.7: Location of streamflow gaging stations used in the High Permeability regression analysis

5. COMPARISON WITH NDOR METHODS

Regression equations were developed from 7.5-minute DEMs for seven hydrologic regions in Nebraska. Six regions had a set of equations to predict discharges for watersheds with areas of less than 10 mi^2. To provide an assessment of the updated regression equations for smaller streams, a comparison was made between the new equations and existing methods. The updated equations were compared to the TR55 method (NRCS), the Hydraflow Hydrograph method (Intelisolve), the Rational method, Beckmans 1976 regression equations (Beckman, 1976), and the updated USGS regression equations (Cordes and Hotchkiss, 1993).

The Nebraska Department of Roads (NDOR) supplied peak discharge estimates at 12 sites in Nebraska; the estimates were based on existing peak flow calculation methods. Peak discharges were calculated for the 10-, 25-, 50-, and 100-year return periods at each of the sites. Calculations made using existing methods were determined by Carnazzo and Donahoo (2003). The methods used included the TR55 method, Hydraflow Hydrograph, and the Rational method. Six of the hydrologic regions are represented by the 12 example sites, and a description and discussion is provided for each ungaged stream. The updated equations that were developed for basins with drainage areas less than 10 mi^2 are designated by "< 10 mi^2" and the equations that were developed using basins of all sizes are designated by "complete". Table 5.1 shows the basin characteristics of the watersheds delineated from the 7.5-minute Digital Elevation Models. Soil characteristics were interpreted using Water Supply Paper 2222 (Dugan 1984). Figure 5.1 shows the locations of all NDOR projects and the hydrologic regions in which the projects are located.

Table 5.1: The basin characteristics of the watersheds analyzed.

Description	*CDA (mi^2)*	*MCS (ft/mile)*	*BS (ft/mile)*	*RR (ft/mile)*	*SF*	*CR*	*PLP (in/hr)*	*P60 (in/hr)*	*MSS (%)*
Deshler	0.70	37.4	145.9	13.8	6.08	1.98	0.45	0.72	2.7
Liberty	1.79	47.7	243.3	17.4	2.45	1.52	0.17	0.45	7.6
Nebraska City	1.12	93.8	453.0	37.9	3.45	1.56	0.60	1.28	10.4
Winnebago	2.28	96.6	678.2	33.6	2.42	1.61	0.73	1.35	8.8
Hartington	2.30	56.5	302.8	25.9	3.91	1.75	1.30	4.50	9.3
Broken Bow	0.33	48.5	357.5	31.1	2.03	1.71	0.60	0.72	25.6
Sargent	1.38	40.3	1,120.1	28.2	4.09	1.91	3.00	5.60	23.4
McCook	3.63	41.0	556.4	20.5	3.65	1.68	0.60	1.31	19.3
Gering	4.11	47.8	162.1	57.8	7.93	1.85	0.60	1.30	9.0
O'Neill	0.91	52.9	148.2	11.5	2.27	1.59	2.00	13.88	12.4
Sidney	0.98	72.2	433.3	32.4	4.19	1.79	0.47	1.89	44.0
Sunol	0.35	96.4	281.4	49.8	5.30	1.75	0.47	1.89	44.0

Note: MCS, BS, RR, SF, and CR are data-scale dependent

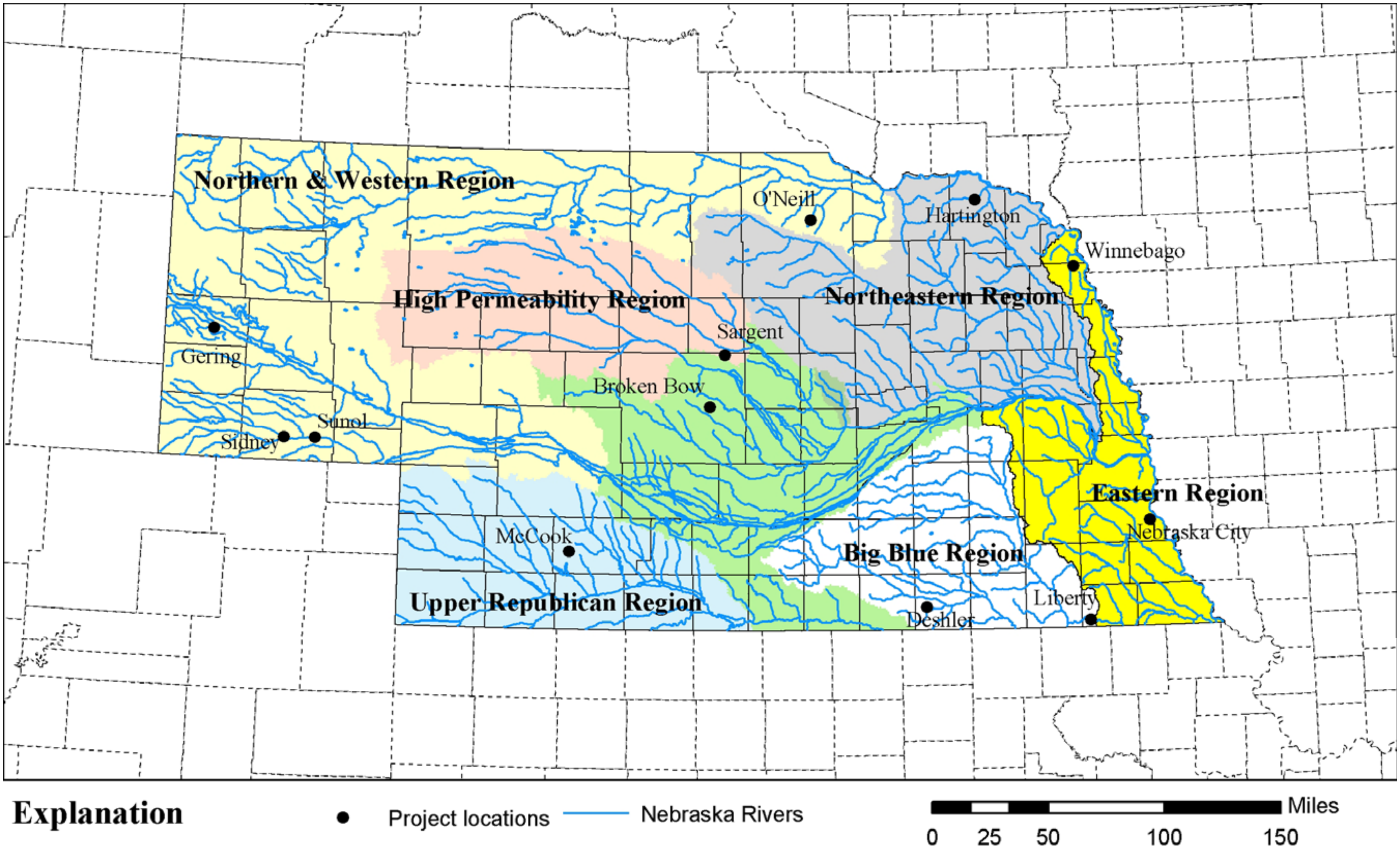


Figure 5.1: The Nebraska Department of Roads existing projects used to compare updated regression equations with existing methods of determining peak flows.

5.1 Deshler, Nebraska – Thayer County

The stream of this project example is located on US Highway 136 in Deshler, and its drainage area is rural farmland. The drainage area is 0.70 mi^2 and is located in the Big Blue Region. Table 5.2 gives the peak discharge estimates for existing methods and the updated regression equations.

Table 5.2: Updated regression equations compared to existing methods for a watershed near Deshler, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		471	578	681	783
Cordes Regression		440	762	1,143	1,575
Beckmans Regression		246	478	765	1,164
Rational Method		365	470	575	678
Updated Regression	(<10 mi^2) – BB	345	622	934	1,361
	(complete) – BB	179	424	678	958

On average the regression equations give higher peak discharges than the TR55 and Rational methods. The largest discharges are predicted using Cordes regression equations. The TR55 and Rational methods give the lowest discharges, and they predict similar peak flows. For basins with drainage areas of less than 10 mi^2, peak flows calculated using the updated equations are a little higher than predicted when methods prescribed by NDOR are used, but they compare reasonably well. The complete updated regression equations also gave comparable discharge estimates. None of the methods produce peak discharge estimates that are unreasonably different than the peak discharges produced using the recommended methods. The location of the watershed is shown in Figure 5.2.

5.2 Liberty, Nebraska – Pawnee County

The second ungaged location is west of Liberty on Nebraska Highway 8. The drainage area is 1.79 mi^2 and is located in the Big Blue Region. Table 5.3 gives the peak discharges estimated using existing methods and the updated regression equations.

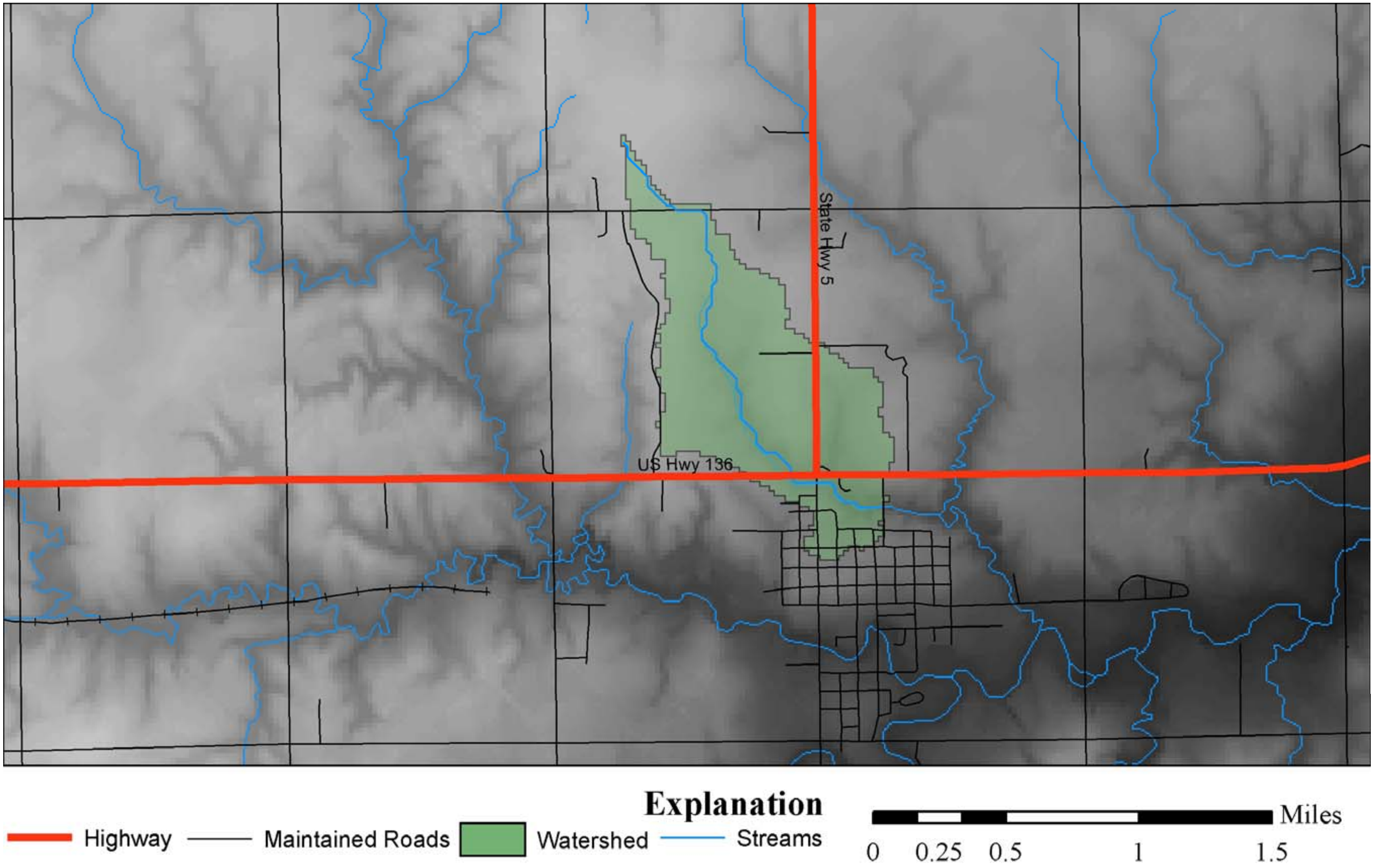


Figure 5.2: Drainage area upstream of a culvert located on U.S. Highway 136 in Deshler, Nebraska.

Table 5.3: Updated regression equations compared to existing methods for a watershed near Liberty, Nebraska.

Method	Q_{10} *(cfs)*	Q_{25} *(cfs)*	Q_{50} *(cfs)*	Q_{100} *(cfs)*
Hydraflow Hydrograph	2,267	2,753	3,231	3,843
Cordes Regression	5,396	8,889	13,333	17,892
Beckmans Regression	1,002	1,918	3,076	4,790
Rational Method	1,035	1,313	1,598	1,873
Updated Regression (<10 mi^2) – BB	1,246	1,979	2,658	3,470
(complete) – BB	1,008	1,887	2,671	3,524

The largest discharges were predicted by Cordes regression equations; these discharges were consistently much higher than the other methods and do not appear to be accurate. The smallest discharges are given by the Rational Method, but the peak flow estimates using the Hydraflow method, Beckman's regression equations and the updated regression equations are similar. The four methods provide predictions that are in agreement with each other. The location of the watershed is shown in Figure 5.3.

5.3 Nebraska City, Nebraska – Otoe County

The ungaged stream is located west of Nebraska City on US Highway 2. The drainage area is 1.17 mi^2 and is located in the Eastern Region. Table 5.4 gives the peak discharge estimates for existing methods and the updated regression equations.

Table 5.4: Updated regression equations compared to existing methods for a watershed in Otoe County, Nebraska.

Method	Q_{10} *(cfs)*	Q_{25} *(cfs)*	Q_{50} *(cfs)*	Q_{100} *(cfs)*
TR55 Method	1,331	1,669	1,949	2,255
Hydraflow Hydrograph	997	1,251	1,460	1,687
Cordes Regression	1,758	3,327	4,158	5,561
Beckmans Regression	1,407	2,267	3,099	4,065
Rational Method	1,024	1,320	1,582	1,865
Updated Regression (<10 mi^2) – East	843	1,226	1,566	1,945
(complete) – East	866	1,613	2,360	3,282

The highest predicted peak flows were calculated by Cordes and Beckmans regression equations. They were both significantly higher than the other methods. Peak flows predicted using TR55, Hydraflow, Rational and the updated regression equations were in agreement with each other. The updated equations for the complete range of areas produced slightly higher results than the other four methods. The location of the watershed is shown in Figure 5.4.

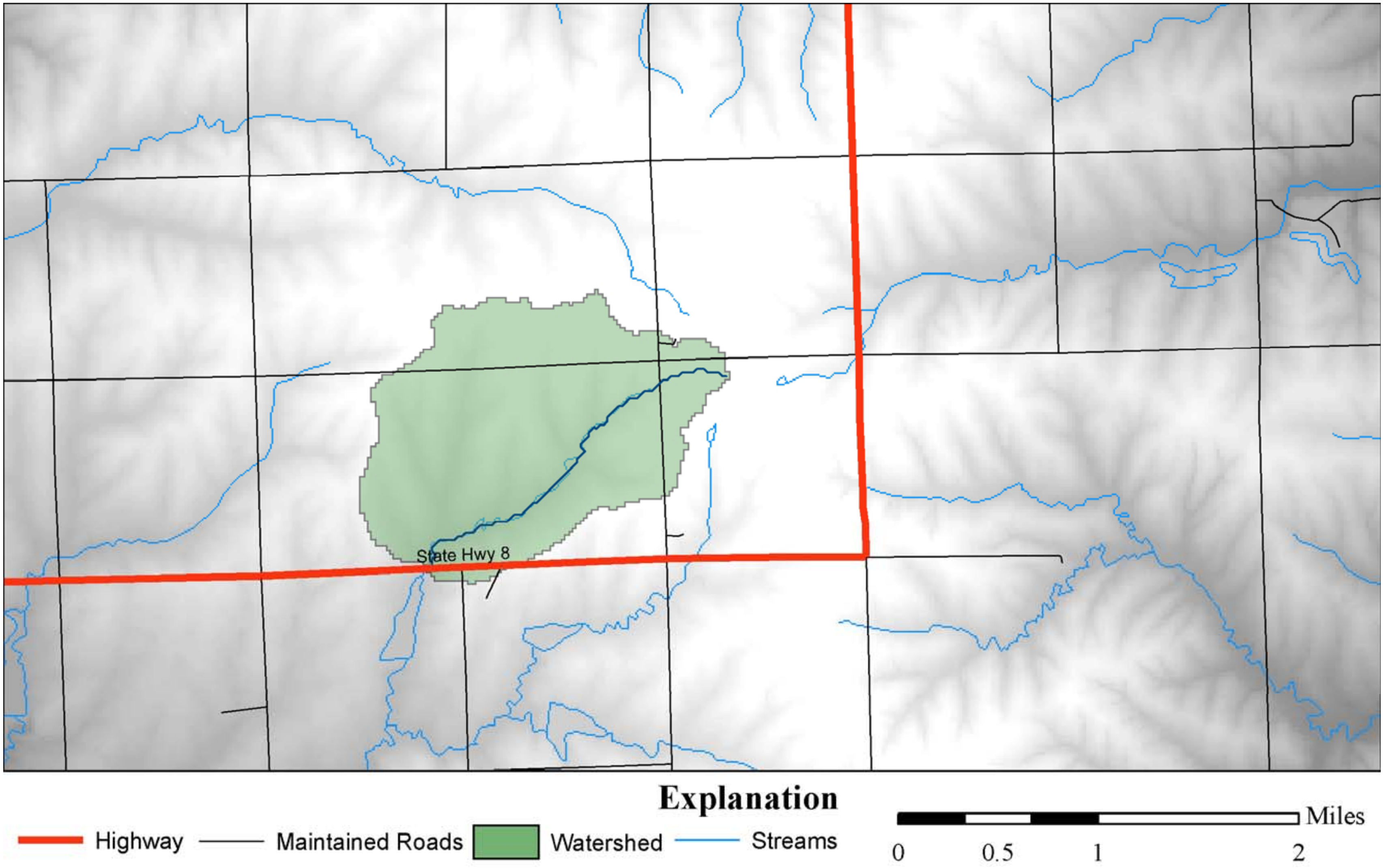


Figure 5.3: Site located west of Liberty, Nebraska on Nebraska Highway 8.

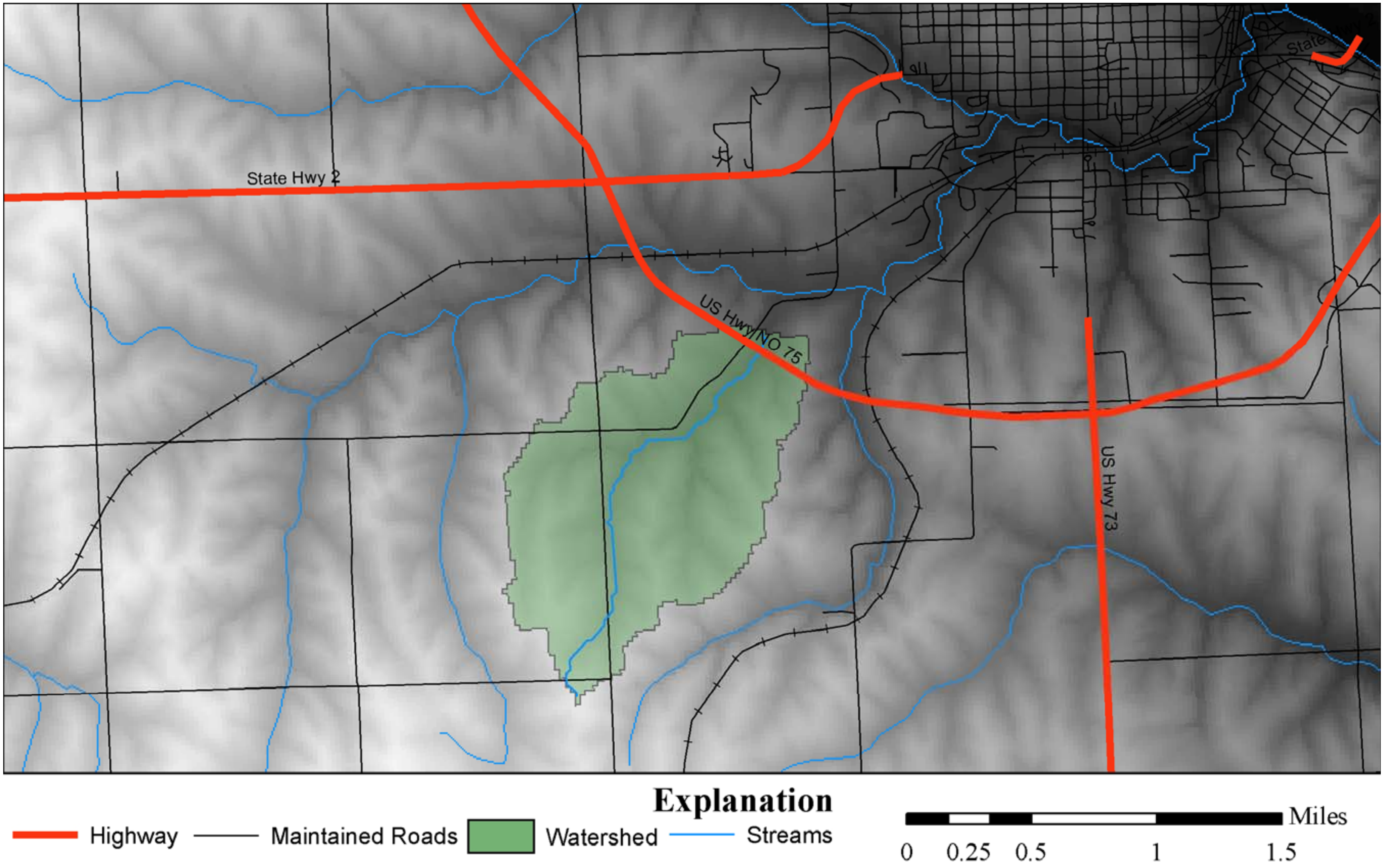


Figure 5.4: Site located west of Nebraska City, Nebraska near U.S. Highway 75.

5.4 Winnebago, Nebraska – Thurston County

The ungaged stream is located northeast of Winnebago on a county road. The drainage area is 2.28 mi^2 and is located on the border between the Eastern and Northeastern Region. Table 5.5 gives the peak discharge estimates for existing methods and the updated regression equations.

Table 5.5: Updated regression equations compared to existing methods for a watershed near Winnebago, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
Hydraflow Hydrograph		1,683	2,250	2,800	3,202
Cordes Regression		763	1,328	2,072	2,969
Beckmans Regression		1,766	2,764	3,714	4,802
Rational Method		1,583	2,087	2,446	2,914
Updated Regression	(<10 mi^2) – East	1,466	2,071	2,593	3,159
	(complete) – East	1,291	2,243	3,133	4,172
Updated Regression	(<10 mi^2) – NE	2,232	3,225	4,119	5,126
	(complete) – NE	1,162	1,821	2,423	3,137

The largest predicted peak flows were calculated from the Beckmans regression equations and the updated Northeastern (NE) regression equations (< 10 mi^2). They were substantially higher than the other methods, and may over-predict peak discharge. Peak flow estimates from the Hydraflow method, Cordes regression equations, the Rational method, and the updated regression equations located in the East (< 10 mi^2) and NE (complete) Region are in agreement. Cordes regression equations may underestimate discharges with return periods less than 50-years. Peak discharges predicted using Hydraflow, the Rational method and the updated equations from the Eastern (< 10 mi^2) and Northeastern (complete) region are in agreement for the most part. The method used by NDOR was the Hydraflow Hydrograph, which produced results that were in agreement with the updated regression estimates for the Eastern and Northeastern region. The location of the watershed is shown in Figure 5.5.

5.5 Hartington, Nebraska – Cedar County

The site is located north of Hartington on Nebraska Highway 15. The drainage area is 2.3 mi^2 and is located in the Northeastern Region. Table 5.6 gives the peak discharge estimates using existing methods and the updated regression equations.

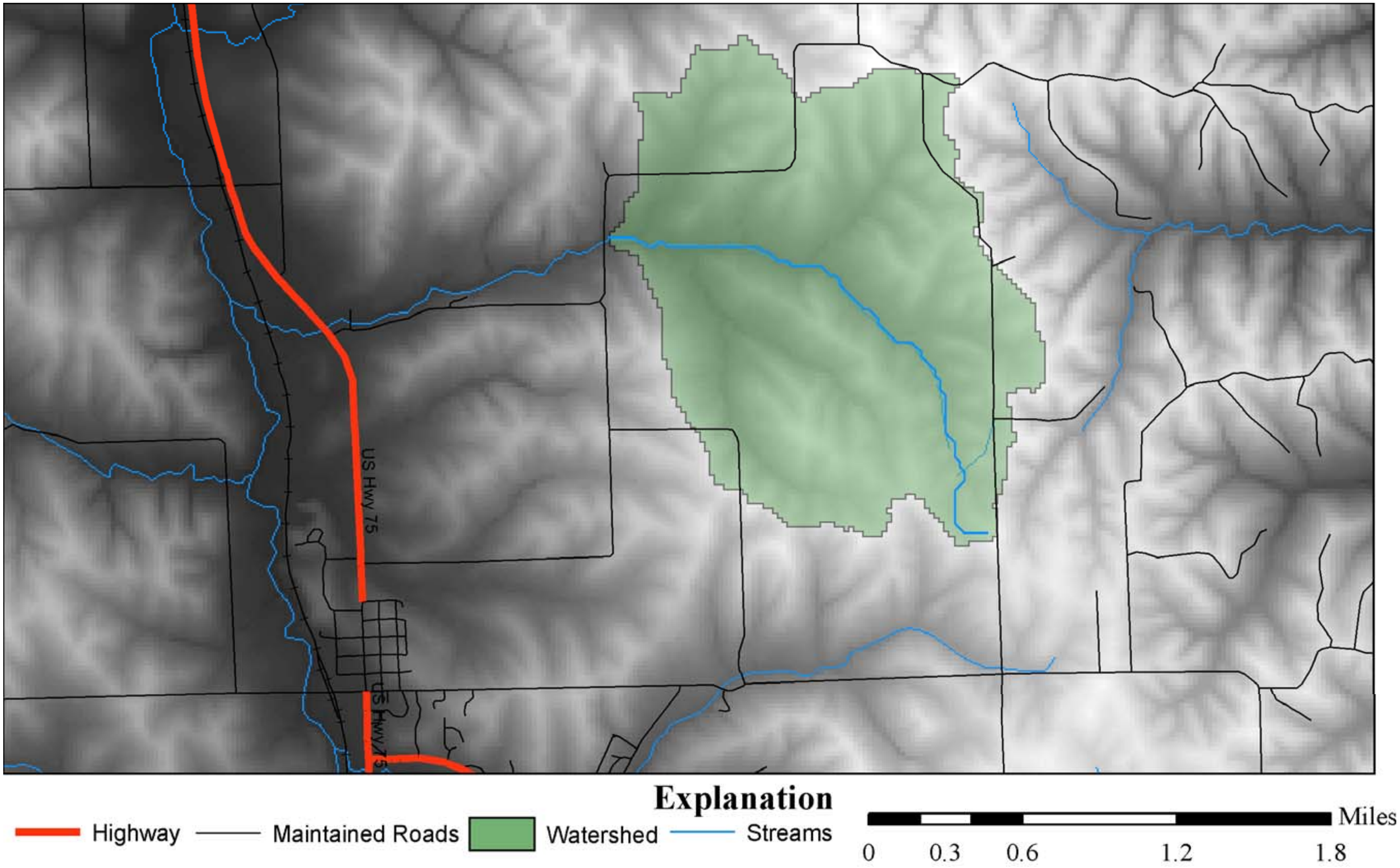


Figure 5.5: Site located north of Winnebago, Nebraska upstream of U.S. Highway 75.

Table 5.6: Updated regression equations compared to existing methods for a watershed near Hartington, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		1,597	2,166	2,709	3,232
Hydraflow Hydrograph		1,923	2,594	3,234	3,790
Cordes Regression		932	1,604	2,406	3,347
Beckmans Regression		1,390	2,208	3,011	3,937
Rational Method		1,046	1,320	1,657	1,990
Updated Regression	(<10 mi^2) – NE	2,025	3,250	4,491	5,993
	(complete) – NE	612	1,025	1,432	1,939

The existing peak flow estimation methods are similar, with the updated regression equations (< 10 mi^2) having higher discharges, especially for long return periods. The recommended method used by NDOR was the TR55 method, which estimated lower peak discharges than were estimated by the updated regression equations. The Hydraflow method over predicted the TR55 method by approximately 400 cfs. Compared to the recommended method the Rational method gave peak flows slightly lower for each recurrence interval. The updated regression equations (< 10 mi^2) produced peak flows that were much higher than peak flows predicted with existing methods. The updated equations for the complete range of drainage areas gave results comparable to results of the Rational method.

There are two possible reasons why the updated regression equations (< 10 mi^2) are slightly larger than the other methods. First, Figures 4.3 and 5.1 show that the closest station with an area of less than 10 mi^2 is located 60 miles to the southeast. Secondly the Northeastern equations were developed with stations that had a PLP range of 0.21-0.60 in/hr. The PLP for the Hartington site is 1.30 in/hr. The location of the watershed is shown in Figure 5.6.

5.6 Broken Bow, Nebraska – Custer County

The ungaged location is east of Broken Bow on Nebraska Highway 70. The drainage area is 0.34 mi^2 and is located in the Central & South Central Region. Table 5.7 gives the peak discharge estimates using existing methods and the updated regression equations.

Table 5.7: Updated regression equations compared to existing methods for a watershed near Broken Bow, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		160	235	297	347
Hydraflow Hydrograph		155	229	289	339
Rational Method		293	383	460	546
Updated Regression	(<10 mi^2) – C&SC	1,159	1,599	1,885	2,082
	(complete) – C&SC	197	346	487	656

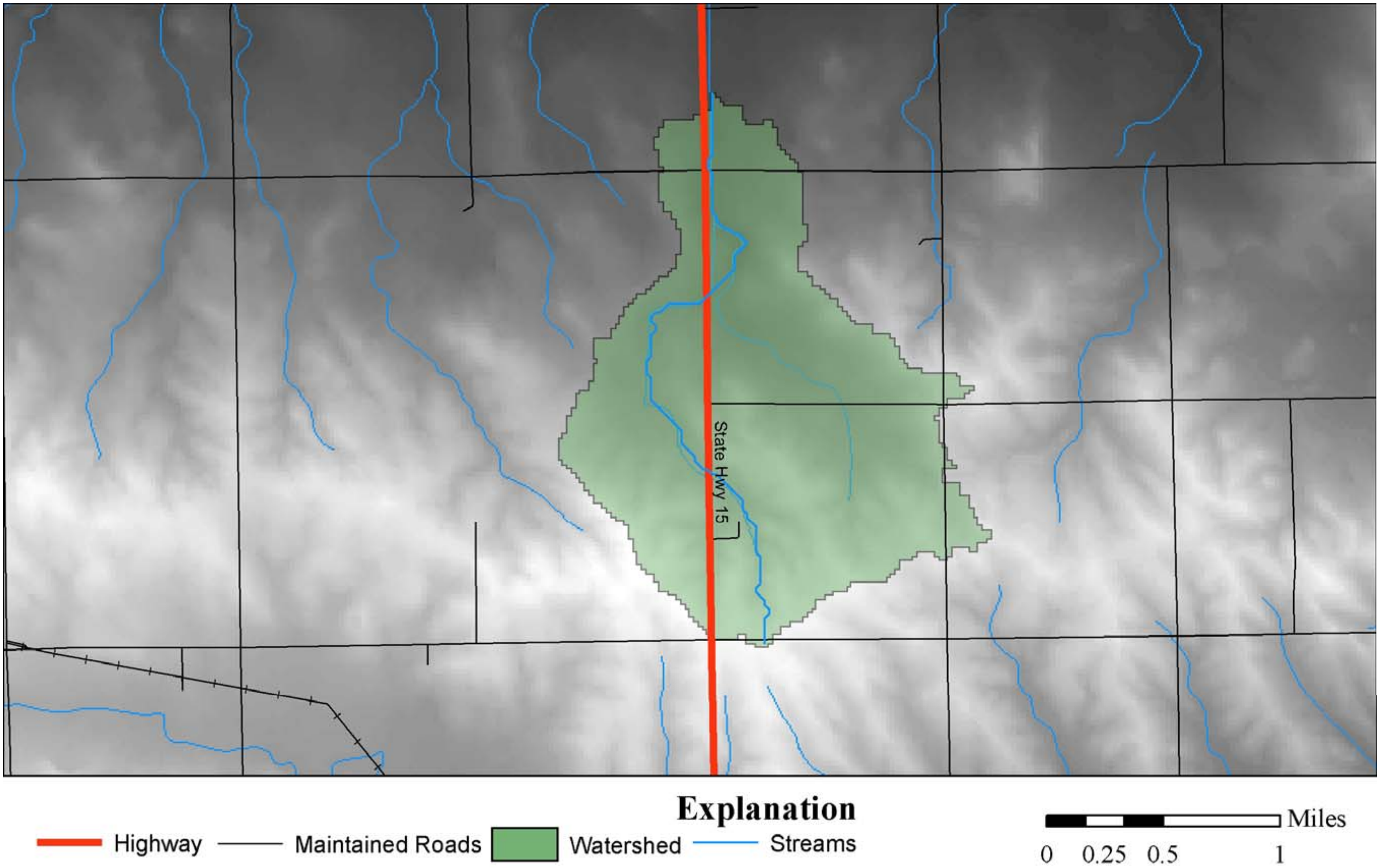


Figure 5.6: Site located north of Hartington, Nebraska on Nebraska Highway 15

The largest peak flow estimates are generated from the updated regression equations (< 10 mi^2). One reason for the large discrepancy may be that the updated regression equations for basins less than 10 mi^2 were not developed from basins smaller than 0.5 mi^2. The smallest basin used in the development of the Central & South Central (C&SC) equations is double that size. Also, the equations were only developed from the basins' morphometric characteristics, excluding soils and climatic effects. Another possible reason the C&SC equations are consistently higher is because they were developed from a SF range of 2.44-7.51. The SF for the Broken Bow site is 2.03.

The other methods provide consistent peak flow estimates for all return periods. The recommended method used by NDOR was the Rational method, which estimated peak flows that were a fourth of the estimates calculated using the updated regression estimates (< 10 mi^2). The updated equations for the complete range of areas provided discharge estimates that were more consistent with the recommended NDOR methods. The location of the watershed is shown in Figure 5.7.

5.7 Sargent, Nebraska – Custer County

The ungaged stream is located north of Sargent on US Highway 183. The drainage area is 1.38 mi^2 and is located in the Central & South Central Region. Table 5.8 gives the peak discharge estimates using existing methods and the updated regression equations.

Table 5.8: Updated regression equations compared to existing methods for a watershed near Sargent, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		623	890	1,165	1,391
Hydraflow Hydrograph		596	915	1,171	1,394
Cordes Regression		884	1,600	2,576	3,616
Beckmans Regression		297	537	781	1,068
Rational Method		992	1,337	1,602	1,925
Updated Regression	(<10 mi^2) – C&SC	577	1,219	1,971	3,031
	(complete) – C&SC	1,367	2,707	4,178	6,116

Cordes and the updated regression equations (complete) predicted the largest peak flows, while Beckmans equations predicted the lowest discharge estimates. The TR55 and Hydraflow methods produced similar peak discharges for the entire range of return periods; these were the recommended NDOR methods. The updated regression equations (< 10 mi^2) also predicted discharges higher than the TR55 and Hydraflow method. Compared to the updated equations the 10- and 25-year discharge estimates were approximately the same. For larger return periods, the separation between peak flow estimates increased. The Rational method produced peak flows that were a little higher than the recommended NDOR method. One reason for the differences between the updated equations and the recommended methods is the lack of peak flow data. Central Nebraska is highly variable and does not have enough gaging stations to accurately represent the region. The location of the watershed is shown in Figure 5.8.

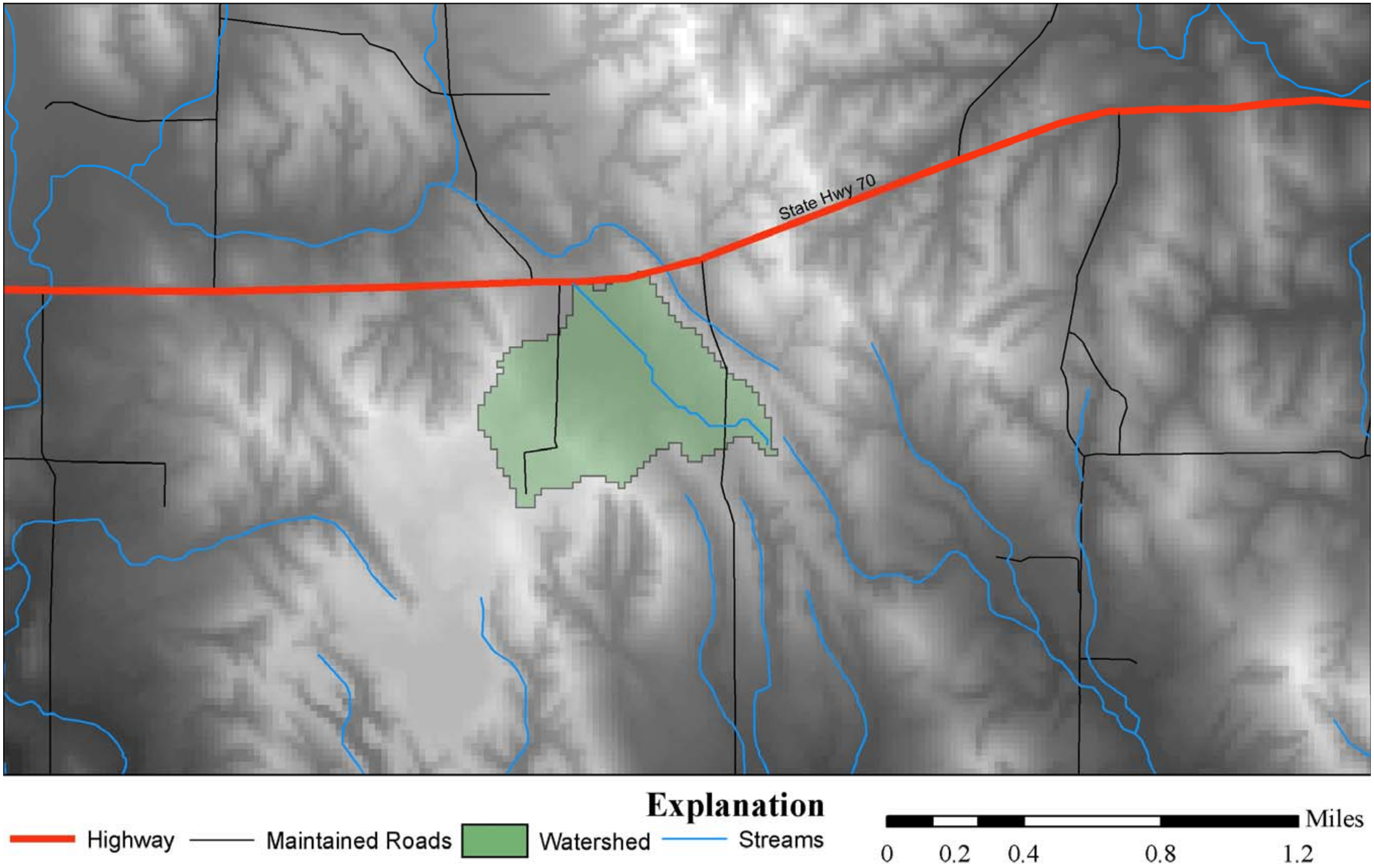


Figure 5.7: Site located east of Broken Bow, Nebraska on State Highway 70.

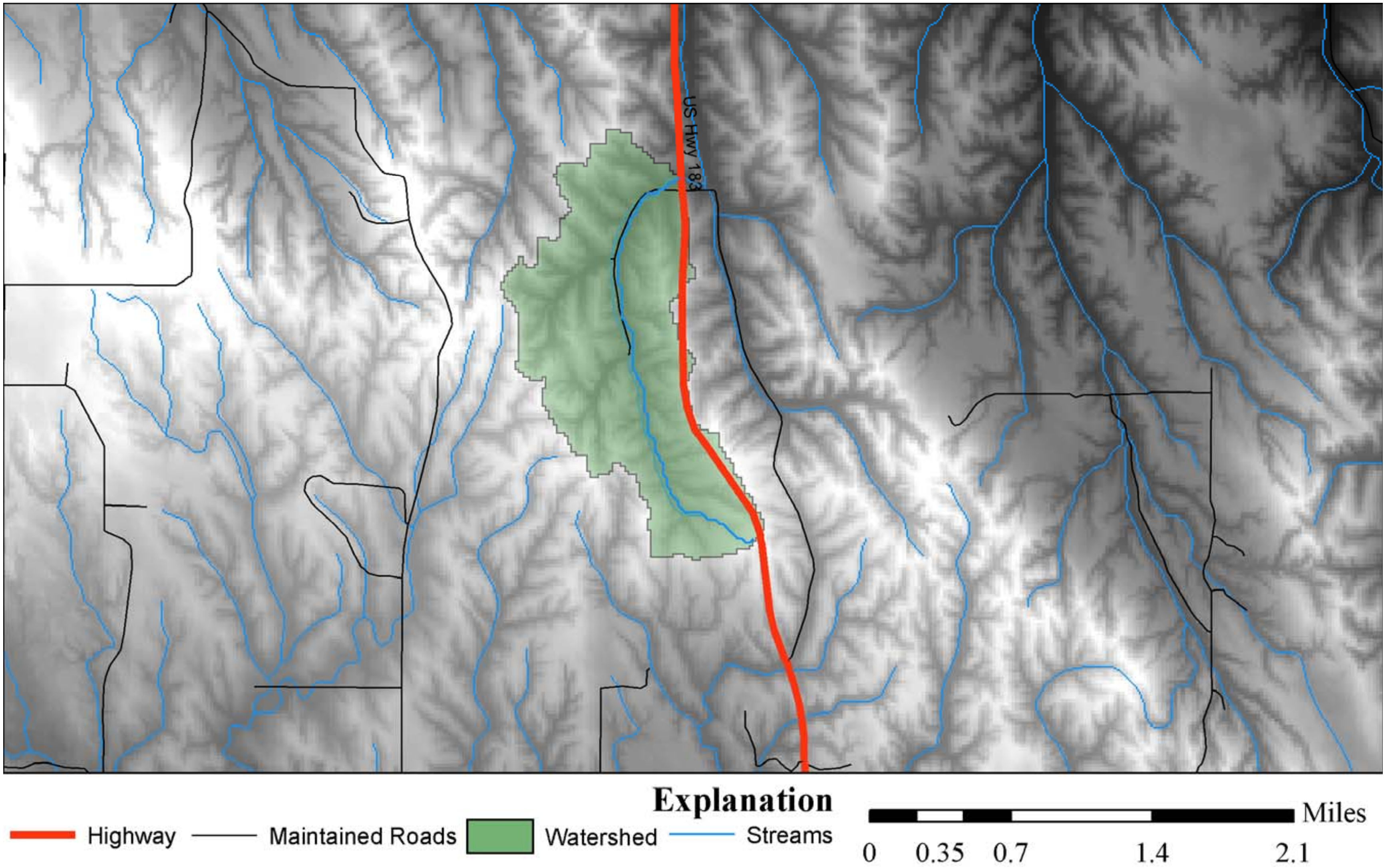


Figure 5.8: Site located north of Sargent, Nebraska on U.S. Highway 183.

5.8 McCook, Nebraska – Frontier County

This site is located between McCook and Maywood on US Highway 83. The drainage area is 3.63 mi^2 and is located in the Upper Republican Region. Table 5.9 gives the peak discharge estimates for existing methods and the updated regression equations.

Table 5.9: Updated regression equations compared to existing methods for a watershed near McCook, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method – Poor/Poor		2,265	3,111	3,911	4,611
TR55 Method – Fair/Good		1,436	2,115	2,758	3,334
Hydraflow – 78/48.0 min		2,383	3,230	3,983	4,649
Hydraflow – 70/48.0 min		1,514	2,237	2,900	3,501
Hydraflow – 78/78.5 min		1,683	2,288	2,826	3,302
Cordes Regression – Region 1		590	1,631	2,447	4,093
Cordes Regression – Region 2		455	976	1,464	2,228
Beckmans Regression – Region 1		749	1,521	2,425	3,643
Beckmans Regression – Region 2		332	608	937	1,352
Rational Method		1,790	2,368	2,850	3,443
Updated Regression	(<10 mi^2) – UR	780	1,416	2,250	3,075
	(complete) – UR	1,600	2,831	4,086	5,649

Many methods were used to estimate peak discharge, due to the location of the culvert. The recommended method by NDOR is the Hydraflow method with a curve number (CN) of 70 and time of concentration (t_c) of 48 minutes. The other two Hydraflow methods gave similar estimates, for the range of return periods. The TR55 methods, varying by CN, produced results that were similar to the Hydraflow estimates. Cordes and Beckmans methods used equations from two different regions. Region 1, which represented the Upper Republican region, yielded approximately the same peak flows as the recommended method for high return periods. Compared to the Hydraflow method, region 2 underpredicted discharge estimates for all return periods. When compared with the TR55 and Hydraflow results, the updated regression equations (< 10 mi^2) under predicted discharges for the 10- and 25-year return periods. But, the peak flow estimates were comparable for recurrence intervals of greater than 50-years. The location of the watershed is shown in Figure 5.9.

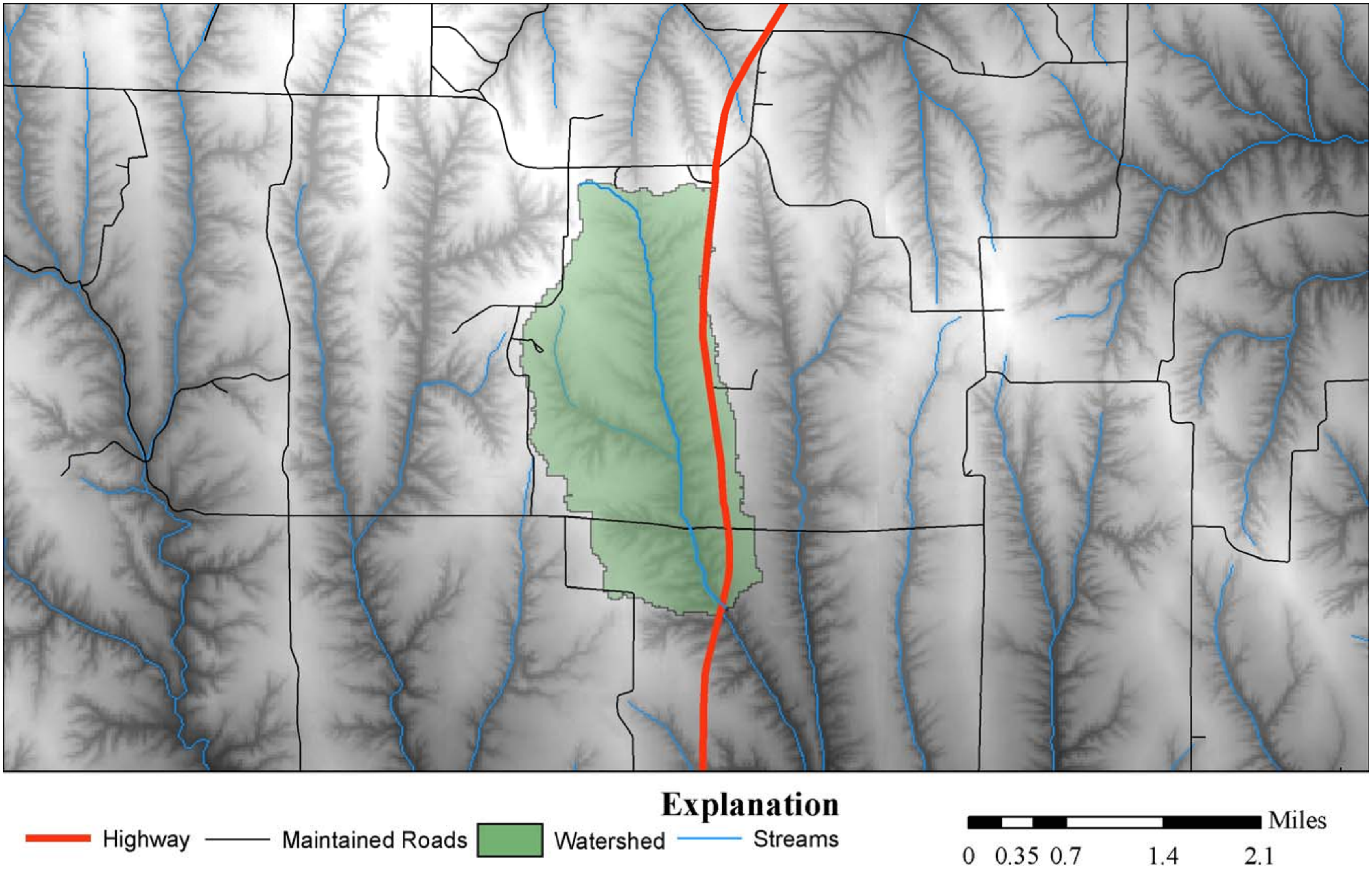


Figure 5.9: Site located north of McCook, Nebraska on U.S. Highway 83.

5.9 Gering, Nebraska – Scottsbluff County

The ungaged stream is located south of Gering near the Gering drain. The drainage area is 4.1 mi^2 and is located in the Northern & Western Region. Table 5.10 gives the peak discharge estimates using existing methods and the updated regression equations.

Table 5.10: Updated regression equations compared to existing methods for a watershed near Gering, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		396	597	764	904
Hydraflow Hydrograph		383	589	775	931
Cordes Regression – P(14)		246	5,609	7,012	25,460
Cordes Regression – P(16)		347	2,808	3,510	8,385
Beckmans Regression – P(14)		43	146	329	685
Beckmans Regression – P(16)		152	394	743	1,306
Rational Method		644	881	1,036	1,277
Updated Regression	(<10 mi^2) – NW	290	596	1,054	1,725
	(complete) – NW	125	407	844	1,543

The TR55 and Hydraflow analysis provided almost identical peak discharge estimates, but the Hydraflow method was recommended by NDOR. The Rational method produced estimates slightly greater than the Hydraflow output. Cordes and Beckmans regression equations appear to overpredict peak flows, except for Beckmans P(14). The updated equations produced results that were similar to the Hydraflow estimates, except for the 100-year return period. Overall the updated equations appear to predict peak flows relatively well for all recurrence intervals. The location of the watershed is shown in Figure 5.10.

5.10 O’Neill, Nebraska – Holt County

The site is located north of O’Neill on US Highway 281. The drainage area is 0.91 mi^2 and is located in the Northern & Western Region. Table 5.11 gives the peak discharge estimates using existing methods and the updated regression equations.

The site is located in the Northern and Western (NW) region but is extremely close to the Northeastern region (NE). Regression equations of both regions were used to estimate peak flows. Cordes and the updated NW regression equations (complete) were almost identical to each other but were larger than NDOR recommended methods. The updated NE regression equations (< 10 mi^2) gave the largest flows for high return periods. A possible reason why NE equations are consistently higher is because they were developed from a PLP range of 0.21-0.60 in/hr. The PLP for the O’Neill site is 2.0 in/hr. Figures 4.3 and 5.1 show that the closest gaging station with a drainage area of less than 10 mi^2 is located approximately 60 miles to the southeast.

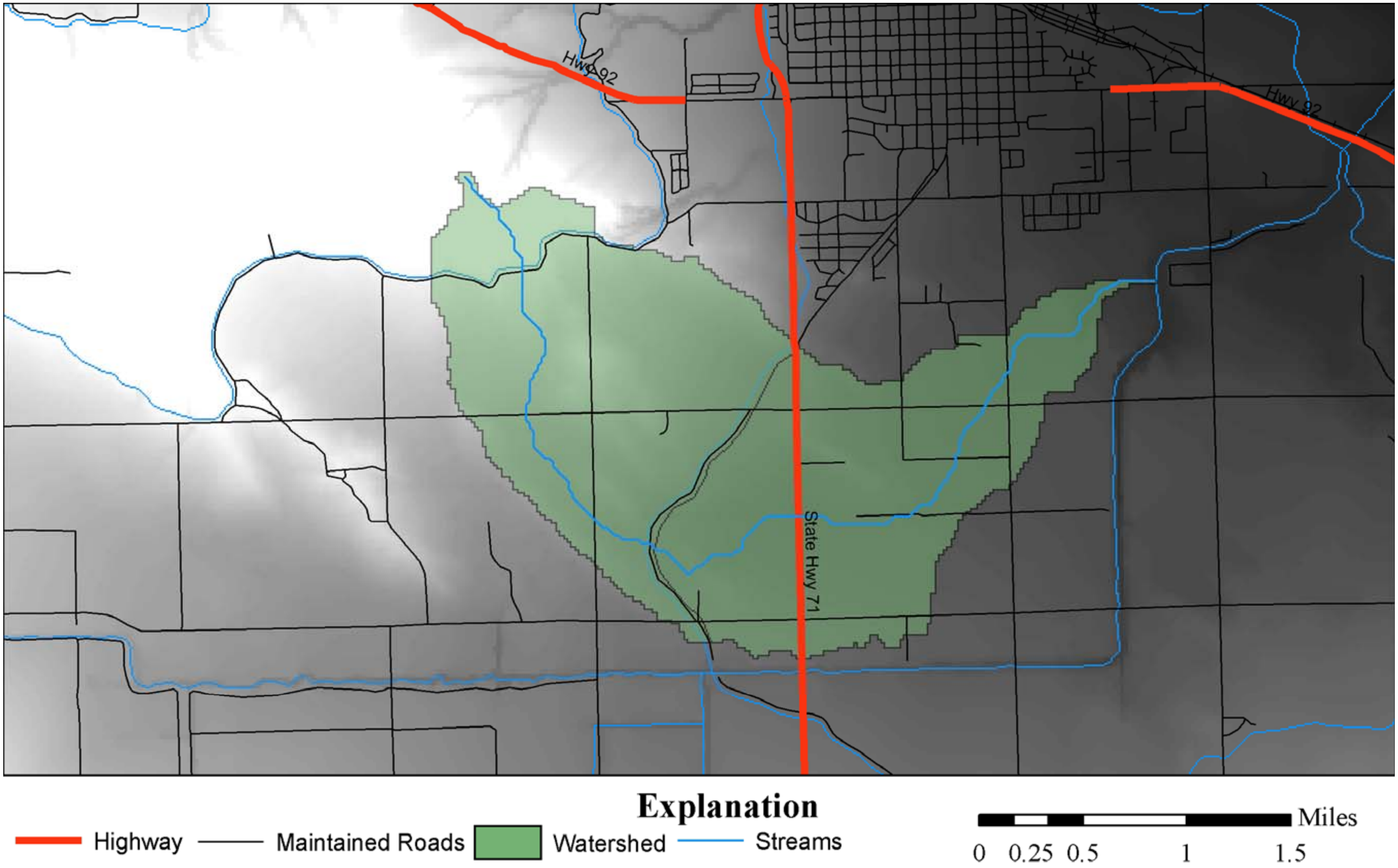


Figure 5.10: Site located south of Gering, Nebraska near State Highway 92.

Table 5.11: Updated regression equations compared to existing methods for a watershed near O'Neill, Nebraska.

Method		Q_{10} *(cfs)*	Q_{25} *(cfs)*	Q_{50} *(cfs)*	Q_{100} *(cfs)*
Hydraflow Hydrograph		237	340	435	518
Cordes Regression – Region 1		408	1,050	1,560	2,509
Cordes Regression – Region 2		464	1,050	1,503	2,266
Rational Method		509	671	842	983
Updated Regression	(<10 mi^2) – NW	25	76	221	541
	(complete) – NW	189	580	1,169	2,127
Updated Regression	(<10 mi^2) – NE	3,010	4,715	6,428	8,511
	(complete) – NE	211	393	594	867

The Hydraflow, Rational and updated NW regression equations (< 10 mi^2) predicted lower discharges. The updated NW equations predicted the lowest peak flows for the given return periods. Peak flows were significantly underpredicted for return periods of less than 50-years. But, the larger return period peak flows compared favorably with results of the NDOR recommended methods. A possible reason that the NW equations are lower is because they were developed for a CDA range of 1.8-10.5 mi^2. The CDA for the O'Neill site is 0.91 mi^2. Furthermore, Figures 4.6 and 5.1 illustrate that the majority of gaging stations with an area of less than 10 mi^2 are at least 150 miles to the west of the test site. The updated equations for the NE region (complete) gave results comparable to NDOR recommended methods. The location of the watershed is shown in Figure 5.11.

5.11 Sidney, Nebraska – Cheyenne County

The site is located west of Sidney on County Road 22. The drainage area is 0.98 mi^2 and is located in the Northern & Western Region. Table 5.12 gives the peak discharge estimates computed using existing methods and the updated regression equations.

Table 5.12: Updated regression equations compared to existing methods for a watershed near Sidney, Nebraska.

Method		Q_{10} *(cfs)*	Q_{25} *(cfs)*	Q_{50} *(cfs)*	Q_{100} *(cfs)*
TR55 Method		412	590	789	961
Hydraflow Hydrograph		413	592	783	947
Cordes Regression		299	1,300	3,186	7,757
Beckmans Regression		223	540	966	1,617
Rational Method		655	884	1,062	1,282
Updated Regression	(<10 mi^2) – NW	31	98	295	741
	(complete) – NW	123	426	919	1,745

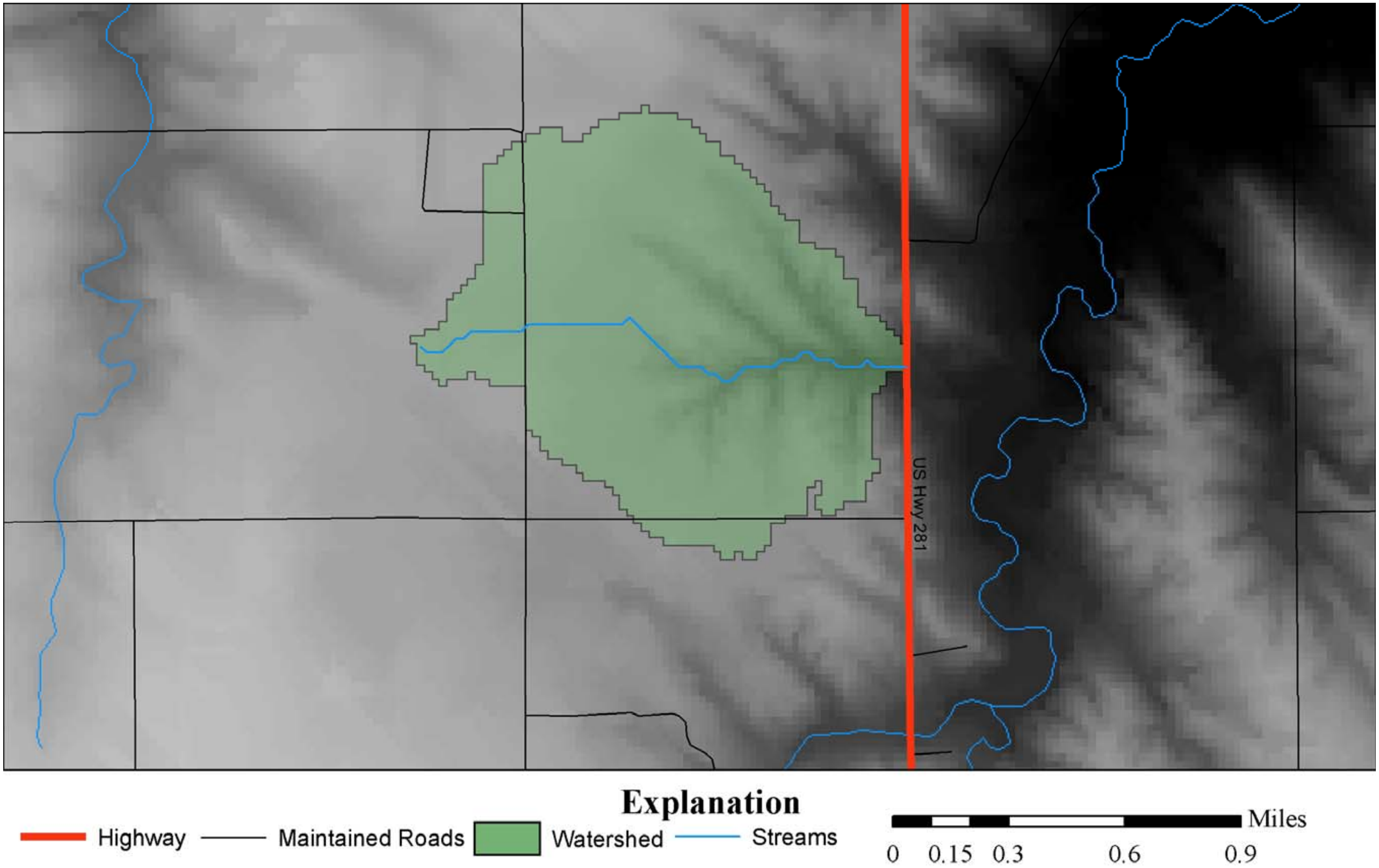


Figure 5.11: Site located north of O’Neill, Nebraska on U.S. Highway 281.

The largest peak flow estimates are given by Cordes regression equations and smallest by the updated regression equations (< 10 mi^2). The updated regression equations are likely too low for return periods of less than 50-years. The recommended method used by NDOR was the Rational method. Beckmans and the updated regression equations (complete) gave peak flow estimates similar to the Rational method. Almost identical peak flow predictions were given by the TR55 and Hydraflow methods, but these discharges are slightly lower than the Rational method estimates. Overall the updated equations (< 10 mi^2) did not compare well to existing NDOR methods. The updated equations for the NW regions were developed from watersheds larger than 1.8 mi^2, which is twice as large as the Sidney basin. Also, figures 4.6 and 5.1 illustrate that a majority of gaging stations with an area less than 10 mi^2 are north of the current site. The location of the watershed is shown in Figure 5.12.

5.12 Sunol, Nebraska – Cheyenne County

The site is located in Sunol on US Highway 30. The drainage area is 0.35 mi^2 and is located in the Northern & Western Region. Table 5.13 gives the peak discharge estimates using existing methods and the updated regression equations.

Table 5.13: Updated regression equations compared to existing methods for a watershed near Sunol, Nebraska.

Method		*Q_{10} (cfs)*	*Q_{25} (cfs)*	*Q_{50} (cfs)*	*Q_{100} (cfs)*
TR55 Method		116	185	264	356
Hydraflow Hydrograph		117	191	281	380
Cordes Regression		235	1,100	2,724	6,837
Beckmans Regression		173	426	771	1,302
Rational Method		448	603	717	856
Updated Regression	(<10 mi^2) – NW	9	33	129	401
	(complete) – NW	67	264	618	1,249

The highest peak flows were predicted using Cordes regression equations. The large regression slopes give low estimates for the 10-year peak flow but high discharges for the 100-year return period. Beckmans and the updated regression equations (complete) produced peak flows that were lower than those predicted by Cordes but were still large for high return periods. The recommended method used by NDOR was the Hydraflow method, which produced results that were nearly identical with those produced using the TR55 method. The Rational method produced results that were higher than those produced using the recommended method, but the results were reasonable.

The updated equations (< 10 mi^2) predicted peak flows that were extremely low compared to peak flows computed using the existing methods. Return periods of 10- and 25-years were inaccurate, but for larger return periods the results compared favorably. The small estimates can be partially attributed to the size of the drainage basin. The area is five times smaller than any stream gage used in the development of the updated equations. Also, figures

4.6 and 5.1 illustrate that a majority of gaging stations with an area less than 10 mi^2 are north of the current site. A lack of stream gages near Sunol makes the peak flow estimates less accurate. The location of the watershed is shown in Figure 5.13.

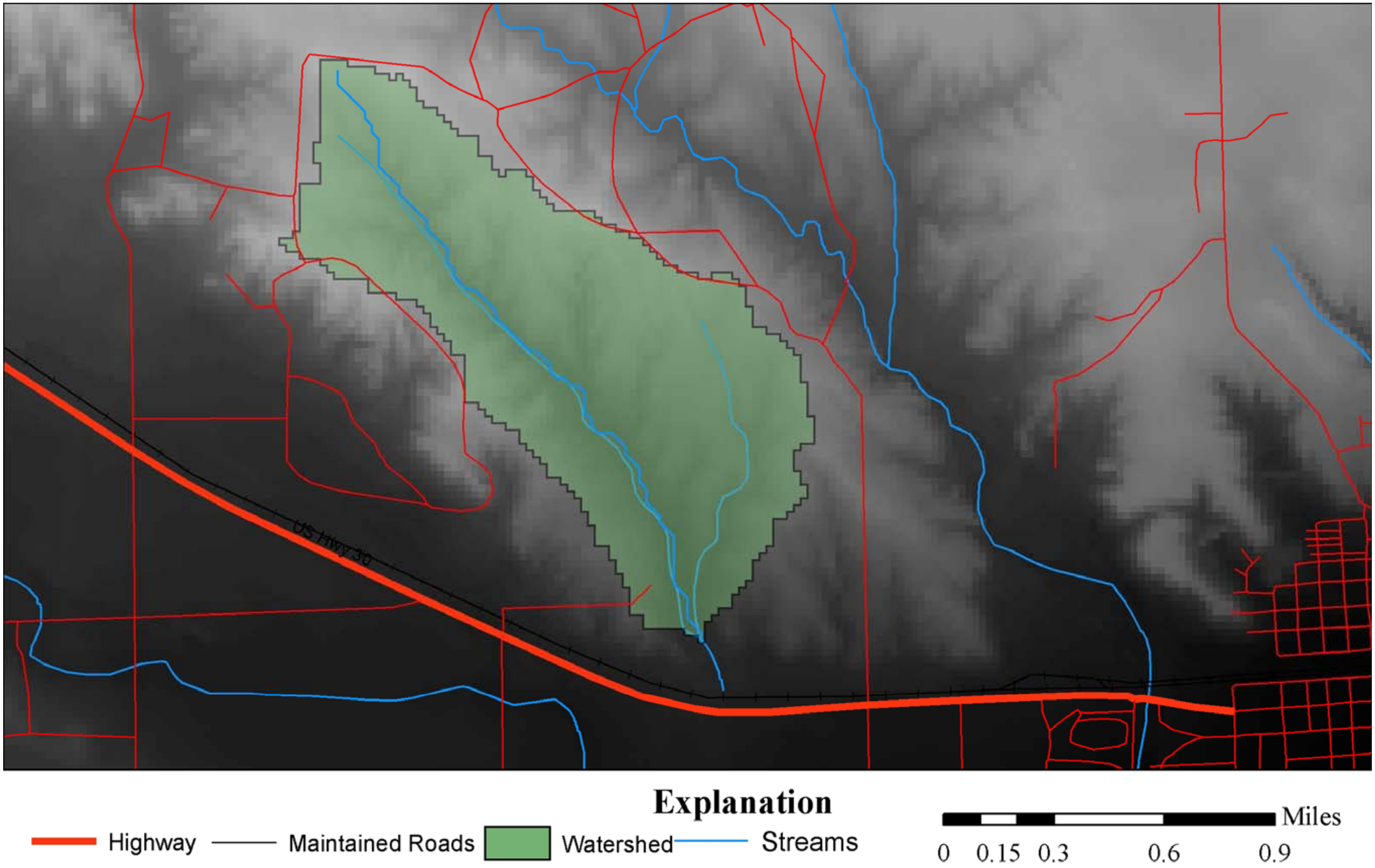


Figure 5.12: Site located northwest of Sidney, Nebraska near U.S. Highway 30.

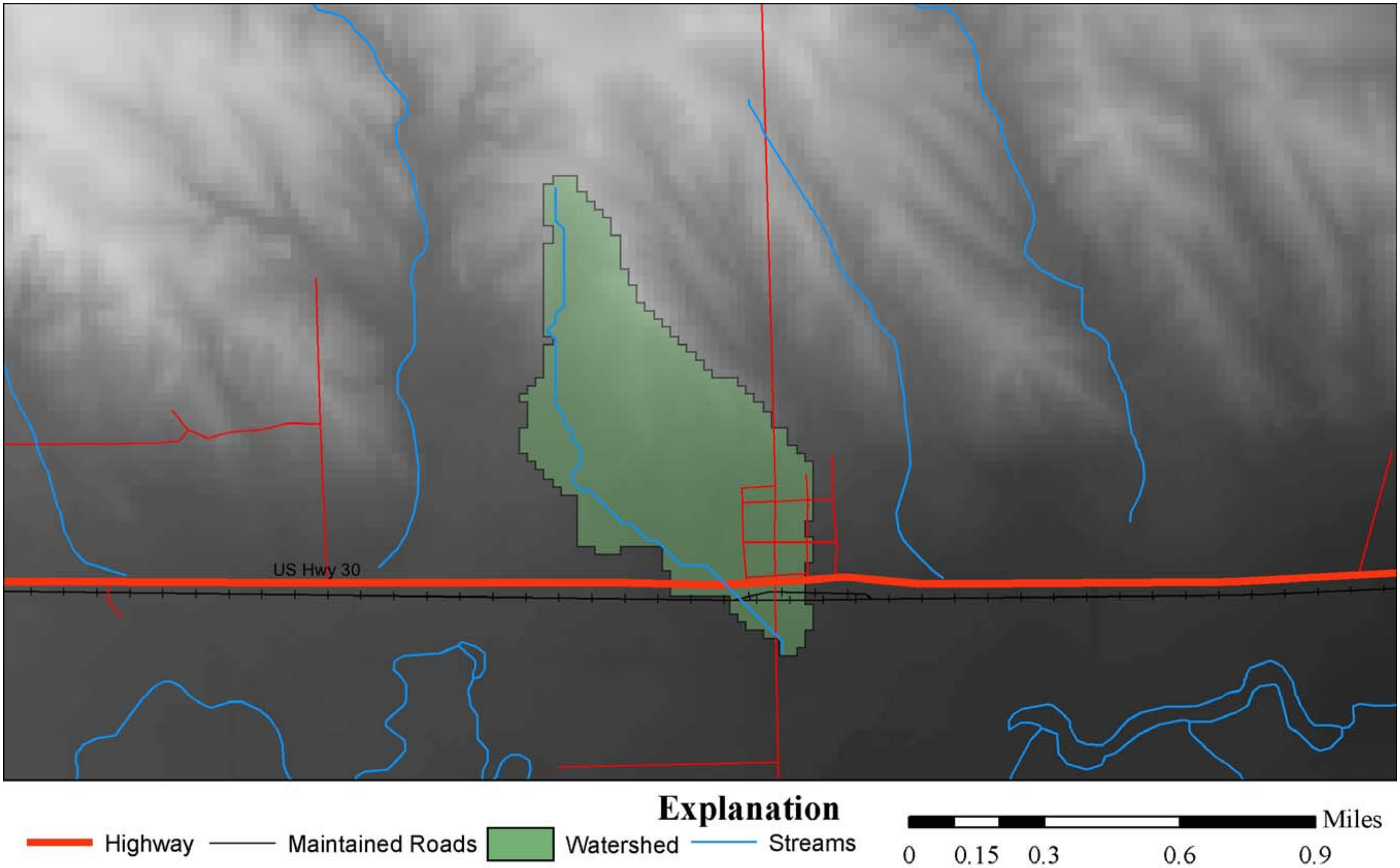


Figure 5.13: Site located at Sunol, Nebraska on U.S. Highway 30.

6. SUMMARY AND CONCLUSIONS

In this chapter, a summary of the methods and procedures used in the project is given. Conclusions of the work are outlined, and future research and implementation that might be of interest based on findings from this research effort are discussed.

6.1 Summary

The objective of this research project was to develop a set of regression equations that allow the Nebraska Department of Roads (NDOR) to rapidly estimate peak flow discharges for both large and small watersheds. The new equations take advantage of new Geographic Information Systems (GIS) technology to reduce processing time and to improve peak flow predictions. The use of 7.5-minute Digital Elevation Models (DEM) improved the spatial resolution so that the revised equations are applicable for high resolution maps.

Regional peak flow frequency analysis made it possible to estimate extreme flow values in locations with limited flow data using data from watersheds with similar hydrologic responses. Using a GIS and digital spatial data, drainage-basin characteristics were quantified. Peak discharges were estimated at return intervals ranging from 2- to 500-years in Nebraska. The regional regression analysis used a weighted-least squares (WLS) regression and data from 273 gaging stations to develop peak flow equations for seven hydrologic regions.

Twenty-five morphometric characteristics were extracted from the 7.5-minute DEMs. The improved DEM resolution allowed for the extraction of characteristics from previously undefined watersheds. The basin characteristics were extracted using ArcInfo software. A basin characteristic database was created using ArcInfo software. ArcInfo was used to manipulate the DEMs into useable hydrologic information. There are twelve measured morphometric basin characteristics which were used in the development of other calculated basin characteristics.

Peak-flow frequency data were gathered for unregulated streams with at least 10 years of annual peak-flow records. Nebraska’s return period discharge estimates were collected from Soenksen et al. (1999a), who used the Log-Pearson Type III frequency distribution and the guidelines in Bulletin 17B of the Interagency Advisory Committee on Water data to determine the peak flows. The gaging station information through the 1994 water year was used to develop the peak discharge frequencies. In addition, the most recent peak-flow frequencies were collected from selected basins in South Dakota, Iowa, Missouri and Kansas. All of the out-of-state stations used in the analysis had flow data at least through the 1992 water year.

Regionalization was used to improve the accuracy of peak flow predictions in Nebraska. In the latest update of Nebraska’s regression equations, the state was subdivided into seven hydrologic regions (Soenksen et al., 1999a). Western Nebraska was regionalized based on permeability and the percent of noncontributing drainage area. The Upper Republican River basin was used in the southwest corner of the state. The central and south-central region was developed from Loup River tributaries and streams located in the Platte River floodplain. The eastern regions were based on watershed divides. Major basins included the Big Blue River, Elkhorn River, Salt Creek, Big Nemaha River, and the Missouri River tributaries. The seven hydrologic regions are the Big Blue, Eastern, Northeastern, Central and South-Central, Upper Republican, Northern and Western, and the High Permeability region.

A weighted least-squares (WLS) regression model was used to develop a relationship between basin characteristics and peak-flow data. The WLS regression model takes into consideration the length of record at each site. Basin characteristics were chosen by minimizing the standard error between observed and predicted peak discharge values, as determined from the regression analysis. Each region had an annual peak flow estimate for the recurrence intervals of 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-years in Nebraska.

All regional regression equations included contributing drainage area (CDA) which is likely to be the most important basin characteristic. Drainage area is directly related to the magnitude of the stream discharge. The regression equations were limited to three basin characteristics for each return period. Also, the equations included at least one slope or soil characteristic in them, and preferably both a slope and soil characteristic. Important slope characteristics include the average basin slope (BS), main channel slope (MCS), and relative relief (RR). Statistically relevant soil characteristics were the average permeability of the least permeable layer (PLP), average permeability rate of the 60-inch soil profile (P60) and the average maximum soil slope (MSS). In addition, a correlation was established between peak flows and the compactness ratio (CR) and shape factor (SF). In the regression equations, exponents with powers of greater than two were avoided. Large exponents can cause the significance of a basin attribute to be over-represented.

6.2 Conclusions

Regional equations were developed for seven hydrologic regions in Nebraska. Two sets of regression equations were developed for each region: one representative of basins with areas less than 10 mi^2 and one for the complete range of drainage areas except for the High Permeability region. The elimination of large watersheds increased the accuracy of prediction for smaller watersheds, but because the number of gages used in the analysis was necessarily reduced, the level of confidence in the resulting equation is also lower.

The Big Blue region is primarily the Big Blue River drainage area in southeastern Nebraska. The equations developed for basins with areas of less than 10 mi^2 gave reasonable estimates when compared to two Nebraska Department of Roads (NDOR) projects. The standard error of the equations ranged from 16 to 64 percent. Equations developed for the complete range of drainage areas had a standard error of 38 to 63 percent. When compared with NDOR predictions the regression equations developed for the Big Blue Region produced consistent results.

The Eastern region represents the Missouri River tributaries in northeastern Nebraska and the southeastern corner of the state. The equations developed for basins with areas of less than 10 mi^2 gave reasonable estimates when compared to the results of two NDOR projects. The standard error of the equations ranged from 22 to 32 percent. Equations developed for the complete range of drainage areas had a standard error of 24 to 49 percent. The equations developed for the complete range of drainage areas also were in agreement with the peak flow estimates determined for the NDOR projects.

The Northeastern region includes most of the Elkhorn River drainage area in Nebraska. The equations developed for basins with areas of less than 10 mi^2 gave reasonable results when compared to three NDOR sites. But, a majority of the stations are located in the eastern part of the region. When compared to the NDOR results of sites located on the west side of the region, the regression equations produced estimates that were high. The standard error of the equations ranged from 23 to 63 percent. Equations developed for the complete range of drainage areas had a standard error of 39 to 70 percent. The gaging stations used in the development of the complete range of drainage areas have a spatially uniform representation in the Northeastern region.

The Central and South-Central region represents the middle Platte, Loup and middle Republican Rivers in Nebraska. The equations developed for basins with areas of less than 10 mi^2 gave variable results. The region has a wide variety of soils and morphological characteristics and lacks representative peak flow data. The standard error of the equations developed for small basins ranged from 33 to 100 percent. Equations developed for the complete range of drainage areas had a standard error of 92 to 149 percent. The equations for the complete range of drainage areas gave reasonable estimates when compared to results of the NDOR methods.

The Upper Republican region represents the southwestern corner of Nebraska. The standard error of the equations developed for basins with areas of less than 10 mi^2 ranged from 21 to 74 percent. Equations developed for the complete range of drainage areas had a standard error of 60 to 113 percent. Both sets of equations produced estimates that were in agreement with the peak flow estimates determined for recent NDOR projects.

The Northern and Western region includes a majority of northwestern Nebraska. The equations developed for basins with areas of less than 10 mi^2 did not compare well to NDOR estimates. The region covers a large area, is highly variable, and has a majority of the gaging stations located in the northwestern corner of the state. The standard error of the equations ranged from 32 to 221 percent. Equations developed for the complete range of drainage areas had a standard error of 81 to 136 percent. Neither set of equations accurately predicts peak discharge, but the equations developed using all of the gaging stations may be more trustworthy because of the shortage of gaging stations on watersheds with small drainage areas.

The High Permeability region is representative of basins centrally located in the Nebraska sandhills. Only equations for the complete range of drainage areas were created. The High Permeability region is highly variable and has high permeability rates. The standard error of estimate ranged from 66 to 172 percent. The regression equations are not likely to be as accurate in this region, but no NDOR sites were available for comparison with existing methods.

With the use of 7.5-minute Digital Elevation Models the spatial resolution used to develop the regression equations was improved. The Big Blue, Eastern and Northeastern region (regions on the eastern side of the state) compared the best with existing NDOR peak-flow estimates. Due to the lack of peak flow data and the higher spatial variability of basin attributes, western Nebraska regions do not accurately estimate peak flows. The greatest concern is that there is only a limited number of streamflow gages, and a much smaller number of streamflow gages for small watersheds.

6.3 Recommendations

The accuracy of the regional regression equations is dependent upon the datasets used to develop them. Regional analysis creates homogeneity within regions, which improves the accuracy of the peak flow estimates. However, when the state is subdivided into regions the number of stations within each region is limited. Additional gaging stations, uniformly distributed throughout the state, are critical to provide a better representation of each region. Additional flow data at current locations and at sites that currently are not recording flow characteristics will strengthen the regression equations. Since many roadway and construction projects require flow information for small basins, it would be beneficial to gage a larger number of basins with areas of less than 10 mi^2. With recent improvements in technology, creating a simple, low cost, and durable recording device may be practical on a statewide basis. Also, satellite data is improving and it may eventually become possible to remotely obtain higher resolution flow data. These things should be considered for future improvements in peak flow prediction.

The locations of additional gaging stations should also be carefully considered. It is important to create a uniform distribution of stations. Each region should be populated with stations that cover the entire region. In addition, the locations should consider a wide range of topography and soil characteristics. The basin slope, main channel slope and relative relief of each basin can be easily extracted. Updates in Nebraska’s GIS databases have made the collection of soil characteristics quicker.

The procedures used to develop Nebraska’s updated regression equations will be helpful as updates in GIS technology and new data become available. Recently 7.5-minute, 10 meter Digital Elevation Models were released by the USGS for Nebraska and updated, high resolution soil maps will be released in the future. Improved topographic resolution and soil properties will improve representations of basin properties and should make basin delineation more accurate, but it should also be recognized that many of the variables used in the regression equations are data-scale dependent, and the equations will need to be adjusted if new scales are introduced.

As a final note, the introduction of high resolution mapping and other GIS capabilities has made it desirable to look into relating precipitation to stream-flow. In future research, it will be beneficial to focus on methods of using precipitation data to predict peak flows, rather than using a statistical representation of the peak flows themselves. Improvements in Doppler radar and other measurement techniques have made it much easier to gather precipitation data, and although humans have an influence on precipitation amounts, they have a much stronger influence on land-use. High resolution elevation data, accurate soils data, and real-time land use monitoring will all contribute to more accurate coupling between precipitation data and peak flow data. Statistical peak flow data, on the other hand, do not take changes in land use into account; this can lead to gross inaccuracies (e.g., if urbanization or changes in tillage practices occur). This methodology may be easier to develop for small watersheds where storm coverage is often 100% and the impact of base flow is not as great.

REFERENCES

Acerman, M.C., and Sinclair, C.D. (1986). “Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland,” *Journal of Hydrology*, **84**, pp. 365 – 380.

Asquith, W.H., and Slade, R.M, Jr. (1996). Regional equations for estimation of peak-stream flow frequency for natural basins in Texas, U.S. Geological Survey Water-Resources Investigations Report 96-4307, 68 p.

Asquith, W.H., and Slade, R.M, Jr. (1999). Site specific estimation of peak stream flow frequency using generalized least-squares regression for natural basins in Texas, U.S. Geological Survey Water-Resources Investigations Report 99-4172, 19 p.

Alexander, T.W., and Wilson, G.L. (1995). Techniques for estimating the 2- to 500- year flood discharges on unregulated streams in rural Missouri, U.S. Geological Survey Water-Resources Investigations Report 95-4231, 33 p.

Becker, L.D. (1980). Techniques for estimating flood peaks, volumes, and hydrographs on small streams in South Dakota, U.S. Geological Survey Water-Resources Investigations Report 80-80, 82 p.

Beckman, E.W. (1976). Magnitude and frequency of floods in Nebraska, Water-Resources Investigations Report 76-109, 128 p.

Beckman, E.W., and Hutchison, N.E. (1962). Floods in Nebraska on small drainage areas, magnitude and frequency, Geological Survey Circular 458, 33 p.

Brown, J.R., Ulery, R.L., and Parcher, J.W. (2000). Creating a standardized watersheds database for the Lower Rio Grande/Rio Bravo, Texas, U.S. Geological Survey Open-File Report 00-065, 17p.

Burn, D.H. (1988). “Delineation of groups for regional flood frequency analysis,” *Journal of Hydrology*, **104**, pp. 345 – 361.

Carnazzo, R. and Donahoo, K. (2003) Written communication, February.

Carter, J.R. (1998). “Digital representations of topographic surfaces,” Photogrammetric Engineering and Remote Sensing,” **54 (11)**, pp. 1577-1580.

Cordes, K and Hotchkiss, R. (1993). Design discharge of culverts: final report, Nebraska Department of Roads, NDOR Research Project Number RES-1 (0099) P466 Transportation Research Studies, December.

Dietsch, B. (2005). Written communication, March.

Djokic, D., and Maidment, D.R. (2000). Hydrologic and hydraulic modeling support with geographic information systems, Redlands California: ESRI Press.

Draper, N.R., and Smith, H. (1981). Applied regression analysis, Wiley, New York.

Dugan, J.T. (1984). Hydrologic characteristics of Nebraska soils, U.S. Geological Survey Water-Supply Paper 2222, 19 p. and 12 pls.

Eash, DA. (1994). “A geographic information system procedure to quantify drainage-basin characteristics,” *Water Resource Bulletin*, **30 (1)**, pp. 1-8.

Eash, D.A. (2000). Techniques for estimating flood-frequency discharges for streams in Iowa, U.S. Geological Survey Water-Resources Investigations Report 00-4233, 88 p.

Elassal, A.A., and Caruso, V.M. (1983). USGS digital cartographic data standards: digital elevation models, U.S. Geological Survey Circular 895-B.

Environmental Systems Research Institute. (2002). ArcINFO 8.x.

Furness, L.W. (1955). Floods in Nebraska, magnitude and frequency, Lincoln, Nebraska, Nebraska Department of Roads and Irrigation, 103p.

Harvey, C.A., and Eash, D.A. (1996). Description, instructions, and verification for Basinsoft, a computer program to quantify drainage-basin characteristics, U.S. Geological Survey Water-Resources Investigations Report 95-4287, 25 p.

Hellweger, F.L., and Maidment, D.R. (1999). “Definition and connection of hydrologic elements using geographic data,” *Journal of Hydrologic Engineering*, **4 (1)**, pp. 10-18.

Interagency Advisory Committee on Water Data. (1982). Guidelines for determining flood flow frequency. Bulletin 17B, U.S. Department of the Interior, U.S. Geological Survey, Office of Water Data Coordination. Reston, VA.

Jensen, S.K., and Domingue, J.O. (1988). “Extracting topographic structure from digital elevation data for geographic information systems analysis,” *Photogrammetric Engineering and Remote Sensing*, **54 (11)**, pp. 1593 – 1600.

Kenward, T., Lettenmaier, D.P., Wood, E.F., and Fielding, E. (2000). "Effects of digital elevation model accuracy on hydrologic predictions," *Remote Sensing of Environment*, **74 (3)**, pp. 432-444.

Kortun, G.F., and Schwartz, R.R. (1986). Multiple-regression equations for estimating low flows at ungaged stream sites in Ohio, U.S. Geological Survey Water-Resources Investigations Report 86-4354, 39 p.

Lettenmaier, D.P., Wallis, J.R., and Wood, E.F. (1987). "Effects of regional heterogeneity on flood frequency estimation," *Water Resources Research*, **23 (2)**, pp. 313 – 323.

Livingston, R.K., and Minges, D.R. (1987). Techniques for estimating regional flood characteristics of small rural watersheds in the Plains Region of Eastern Colorado, U.S. Geological Survey Water-Resources Investigations Report 87-4094, 72 p.

Loveland, T.R., and Ramey, B. (1986). Applications of U.S. Geological Survey digital cartographic products, U.S. Geological Survey Bulletin, 1583.

Maidment, D.R. (2002). Arc Hydro GIS for water resources, Redlands California: ESRI Press.

Majure, J.J., and Eash, D.A. (1991). An automated method to quantify physical basin characteristics, U.S. Geological Survey Water-Resources Investigations Report 91-4034, 44 p.

Martz, L.W., and Garbrecht, J. (1993). "Automated extraction of drainage network and watershed data from digital elevation models," *Water Resource Bulletin*, **29 (6)**, pp. 901-908.

Matthai, H.F. (1968). Magnitude and frequency of floods in the United States, part 6-B, Missouri River Basin below Sioux City, Iowa, U.S. Geological Survey Water-Supply Paper 1680, 471 p.

Moglen, G.E., and Hartman, G.L. (2001). "Resolution effects on hydrologic modeling parameters and peak discharge," *Journal of Hydrologic Engineering*, **6 (6)**, pp.490-497.

Natural Resource Conservation Service (1994). State soil geographic (STATSGO) data base, data use information, U.S. Department of Agriculture, National Soil Survey Center, Miscellaneous Publication 1492, 110 p.

Pandey, G.R., and Nguyen, V.T.V. (1999). "A comparative study of regression based methods in regional flood frequency analysis," *Journal of Hydrology*, **225 (1-2)**, pp. 92-101.

Patterson, J.L. (1966). Magnitude and frequency of floods in the United States, part 6-A, Missouri River Basin above Sioux City, Iowa, U.S. Geological Survey Water-Supply Paper 1679, 491 p.

Provaznik, M.K. (1997). Flood frequency for Nebraska using L-moments, Lincoln, Nebraska, University of Nebraska, unpublished thesis, 58 p. and 16 appendices.

Rasmussen, P.P, and Perry, C.A. (2000). Estimation of peak stream flows for unregulated rural streams in Kansas, U.S. Geological Survey Water-Resources Investigations Report 00-4079, 33 p.

Riggs, H.C. (1973). Regional analysis of streamflow characteristics: techniques of water resources investigations, Book 4, Chap B3, U.S. Geological Survey, Washington DC.

Sando, Steven K. (1998). Techniques for estimating peak-flow magnitude and frequency relations for South Dakota streams, Water-Resources Investigation Report 98-4055, 24 p.

Soenksen, P.J., Miller, L.D., Sharpe, J.B., and Watton, J.R. (1999a). Peak-flow frequency relations and evaluation of the peak-flow gaging network in Nebraska, Water-Resources Investigation Report 99-4032, 47 p.

Soenksen, P.J., Miller, L.D., Sharpe, J.B., and Watton, J.R. (1999b). Peak-flow frequency relations and evaluation of the peak-flow gaging network in Nebraska, Errata Sheet 2, Water-Resources Investigation Report 99-4032, pg 28.

SPSS, Inc. (2001). SPSS regression models 11.0., SPSS Inc.

Stedinger, J. and Tasker, G. (1986). "Regional hydrologic analysis, 2, model-error estimators, estimation of sigma and log-Pearson type 3 distributions," *Water Resources Research*, **22 (10)**, pp. 1487-1499.

Strahm, B.J. (2003). Development of Regression Equations for Hydrologic Analysis Using GIS, University of Nebraska, M.S. Thesis.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I. (1991). "On the extraction of channel networks from digital elevation data," *Hydrological Processes*, **5**, pp. 81-100.

Tasker, G.D. (1978). "Relation between standard errors in log units and standard errors in percent," *Water Resources Bulletin*, Jan-Mar-Apr-June.

Tasker, G.D. (1982). "Comparing methods of hydrologic regionalization," *Water Resources Bulletin*, **18 (6)**, pp. 965 – 970.

Tasker, G.D., Stedinger, J.R. (1986). "Regional skew with weighted LS regression," *Journal of Water Resources Planning and Management*, ASCE, **112 (2)**, pp. 225 – 237.

Tasker, G.D., Stedinger, J.R. (1989). "An operational GLS model for hydrological regression," *Journal of Hydrology*, **111**, pp. 361-375.

Vaill, J.E. (1999). Analysis of the magnitude and frequency of floods in Colorado, Water-Resources Investigation Report 99-4190, 12 p.

Verdin, K.L., and Verdin, J.P. (1999). "A topological system for delineation and codification of the earth's river basins," *Journal of Hydrology*, **218**, pp. 1-12.

Wang, X., Yin, Z. (1998). "A comparison of drainage networks derived from digital elevation models at two scales," *Journal of Hydrology*, **210**, pp. 221-241.

Wiche, G.J. (1992). Application of digital elevation models to delineate drainage areas and compute hydrologic characteristics for sites in the James River Basin, North Dakota, U.S. Geological Survey Water-Supply Paper 2383, 23 p.

Wiltshire, S.E. (1985). "Grouping basins for regional flood frequency analysis," *Hydrological Sciences Journal – des Sciences Hydrologiques*, **30 (1)**, pp. 151 – 159.

Wiltshire, S.E. (1986a). "Identification of homogeneous regions for flood frequency analysis," *Journal of Hydrology*, **84**, pp. 287 – 302.

Wiltshire, S.E. (1986b). "Regional flood frequency analysis I: homogeneity statistics," *Hydrological Sciences Journal – des Sciences Hydrologiques*, **31 (3)**, pp. 321 – 333.

Wiltshire, S.E. (1986c). "Regional flood frequency analysis II: multivariate classification of drainage basins in Britain," *Hydrological Sciences Journal – des Sciences Hydrologiques*, **31 (3)**, pp. 335 – 347.

Zrinji, Z., and Burn, D.H. (1994). "Flood frequency analysis for ungaged sites using a region of influence approach," *Journal of Hydrology*, **153 (1-4)**, pp. 1-21.

APPENDIX A. DEFINITIONS OF MORPHOMETRIC AND SOILS PARAMETERS

Table A.1: Drainage basin characteristics quantified using Arcinfo

Morphometric Characteristics

Morphometric characteristics were delineated from 1:24,000-scale Digital Elevation Models (DEM), with 30-meter resolution. The DEM's are an array of elevations representing ground positions at regularly spaced horizontal intervals. The use of a single dataset simplified and increased the processing speed of the basin network analysis. It also produced a uniform dataset with seamless basin measurements.

Basin-Area Quantifications

Total drainage area (TDA): in square miles, the WATERSHED function in GRID was used to determine the area. TDA includes all areas that will potentially contribute to surface runoff, based on topography.

Contributing Drainage Area (CDA): in square miles, the area within the TDA that contributes directly to surface runoff. If NCDA exists the CDA was determined from published data.

Non-contributing drainage area (NCDA): in square miles, all areas in the basin that do not directly contribute to surface runoff. NCDA was considered as SINKS within the DEM's, but errors in resolution limited its use. TDA and published CDA were used to calculate non-contributing areas.

$CDA = TDA - NCDA$

Basin-Length Quantifications

Basin Perimeter (BP): in miles, determined from the PERIMETER value in the INFO file of the watershed polygon coverage. Basin perimeter is a measure of the length around the entire total drainage area.

Basin Width (BW): Effective basin width, in miles.

$$BW = \frac{CDA}{MCL}$$

Basin-Relief Quantifications

Average Basin Slope (BS): in feet per mile, the contour-band method was used to determine the basin slope of the TDA.

BS = [(total length of all selected elevation contours) (contour interval)] / TDA

Basin Relief (BR): if feet, measured as the elevation difference between the highest grid cell (E_{max}) and the elevation of the watershed outlet (E_{min}).

Table A.1: (*continued*)

Maximum basin elevation (E_{max}): Taken from statistics INFO file of the elevation grid.

Minimum basin elevation (E_{min}): Taken from statistics INFO file of the elevation grid.

Basin Quantifications

Compactness Ratio (CR): dimensionless,

$$CR = \frac{BP}{2\sqrt{\pi CDA}}$$

Elongation Ratio (ER): dimensionless,

$$ER = \sqrt{\frac{4CDA}{\pi MCL^2}} = 1.13\sqrt{\frac{1}{SF}}$$

Rotundity of Basin (RB): dimensionless,

$$RB = \frac{\pi MCL^2}{4CDA} = 0.785SF$$

Relative Relief (RR): in feet per mile,

$$RR = \frac{BR}{BP}$$

Shape Factor (SF): dimensionless,

$$SF = \frac{MCL}{BW}$$

Channel or Stream Quantifications

Main Channel Length (MCL): in miles, the FLOWLENGTH command was used to determine the length of the longest reach. Flow path was measured from the basin outlet to the watershed divide in the TDA.

Total Stream Length (TSL): in miles, summing the lengths of all stream segments within the total drainage area. Using the INFO table from the STREAMORDER coverage, TSL can be found by summing the LENGTH column.

Channel-Relief Quantification

Main-Channel Slope (MCS): in feet per mile, Computed from the difference in elevations at 10 percent (E_{10}) and 85 percent (E_{85}) of the distance along the main channel from the pour point to the basin divide.

Table A.1: (*continued*)

$$MCS = \frac{(E_{85} - E_{10})}{0.75MCL}$$

Channel or Stream Quantification

Main-Channel Sinuosity Ratio (MCSR): dimensionless,

$$MCSR = \frac{MCL}{BL}$$

Stream Density (SD): in miles per square miles,

$$SD = \frac{TSL}{CDA}$$

Constant of Channel Maintenance (CCM): in square miles per mile,

$$CCM = \frac{CDA}{TSL} = \frac{1}{SD}$$

Main-Channel Slope proportion (MCSP): dimensionless,

$$MCSP = \frac{MCL}{\sqrt{MCS}}$$

Ruggedness Number (RN): in feet per mile,

$$RN = \frac{(TSL)(BR)}{CDA} = (SD)(BR)$$

Slope Ratio (SR): dimensionless,

$$SR = \frac{MCS}{BS}$$

Stream-Order Quantifications

First Order Streams (FOS): dimensionless, a STREAMORDER grid was created using the Strahler method option in GRID. GRID summary statistics are used to compute the number of first order streams.

Basin Stream Order (BSO): dimensionless, stream order of the main channel at the basin outlet.

Drainage Frequency (DF): in number of first order streams per mile,

$$DF = \frac{FOS}{CDA}$$

Table A.1: (*continued*)

Relative Stream Density (RSD): dimensionless,

$$RSD = \frac{(FOS)(CDA)}{TSL^2} = \frac{DF}{SD^2}$$

Soil Characteristics

Soils were based on characteristics defined by Dugan (1984) for Nebraska Stations and by State Soil Geographic Data Base (STATSGO) for stations outside of Nebraska (Natural Resources Conservation Service, 1994).

Average Permeability (P60): in inches per hour, the average permeability rate of the soil horizon. Where, PERMH and PERML are the maximum and minimum value for a range in permeability and FA is the fractional area of the drainage basin occupied by the soil series.

$$PAvgH = \frac{(PERMH + PERML)}{2}$$

$$P60 = \sum (PAvgH \bullet FA)$$

Average Available Water Capacity (AWC): in inches per hour, where AWCH is the maximum value for the range of available water capacity for the soil horizon.

$$AWC = \sum (AWCH \bullet FA)$$

Average Minimum permeability (PLP): in inches per hour, where PERML is the minimum value for the range in permeability rate for the soil layer.

$$PLP = \sum (PERML \bullet FA)$$

Average Maximum Soil Slope (MSS): in percent, the maximum value for the range of slope (SLOPEH) of a soil series.

$$MSS = \sum (SLOPEH \bullet FA)$$

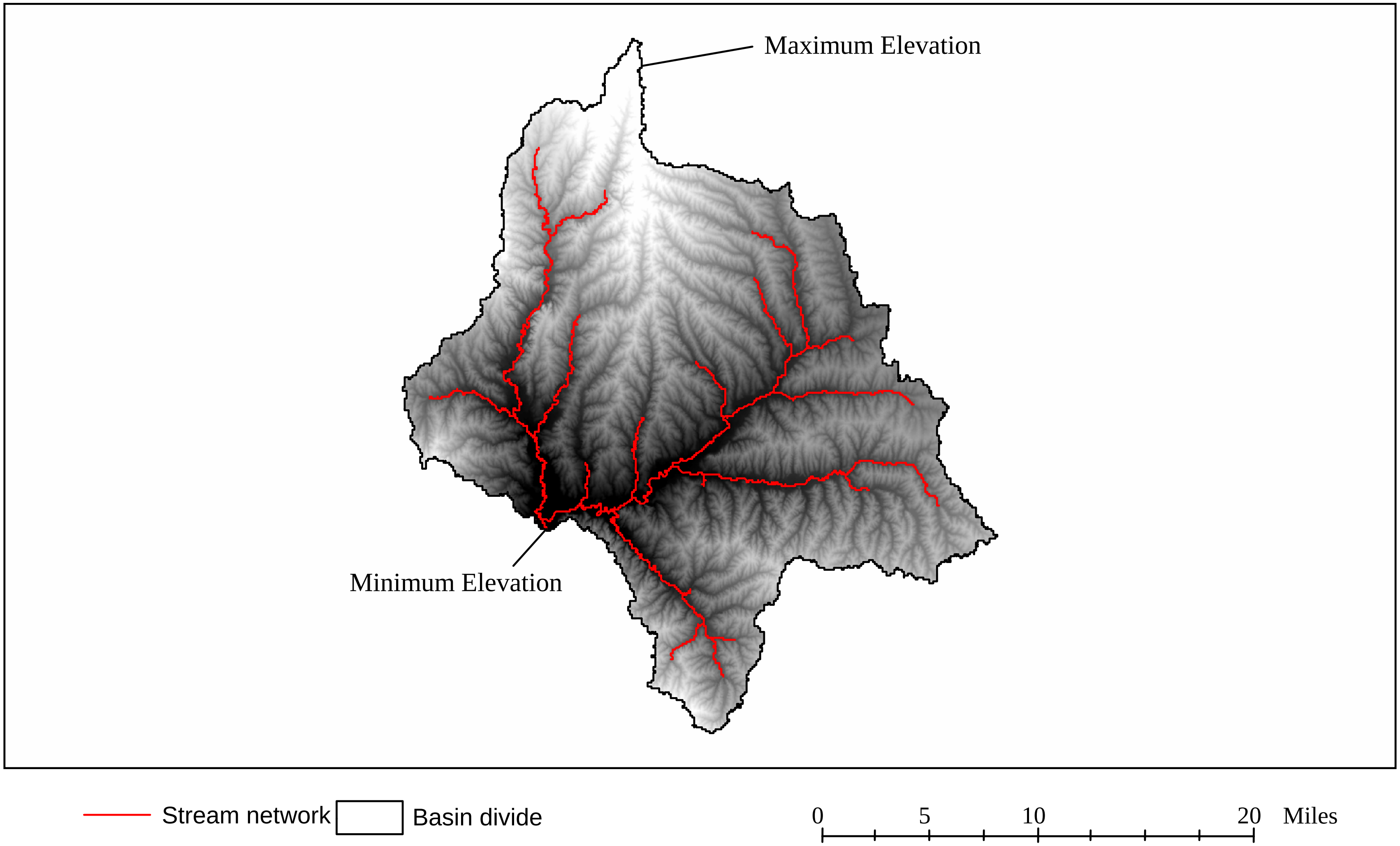


Figure A.1: Locations of the basin relief quantifications.

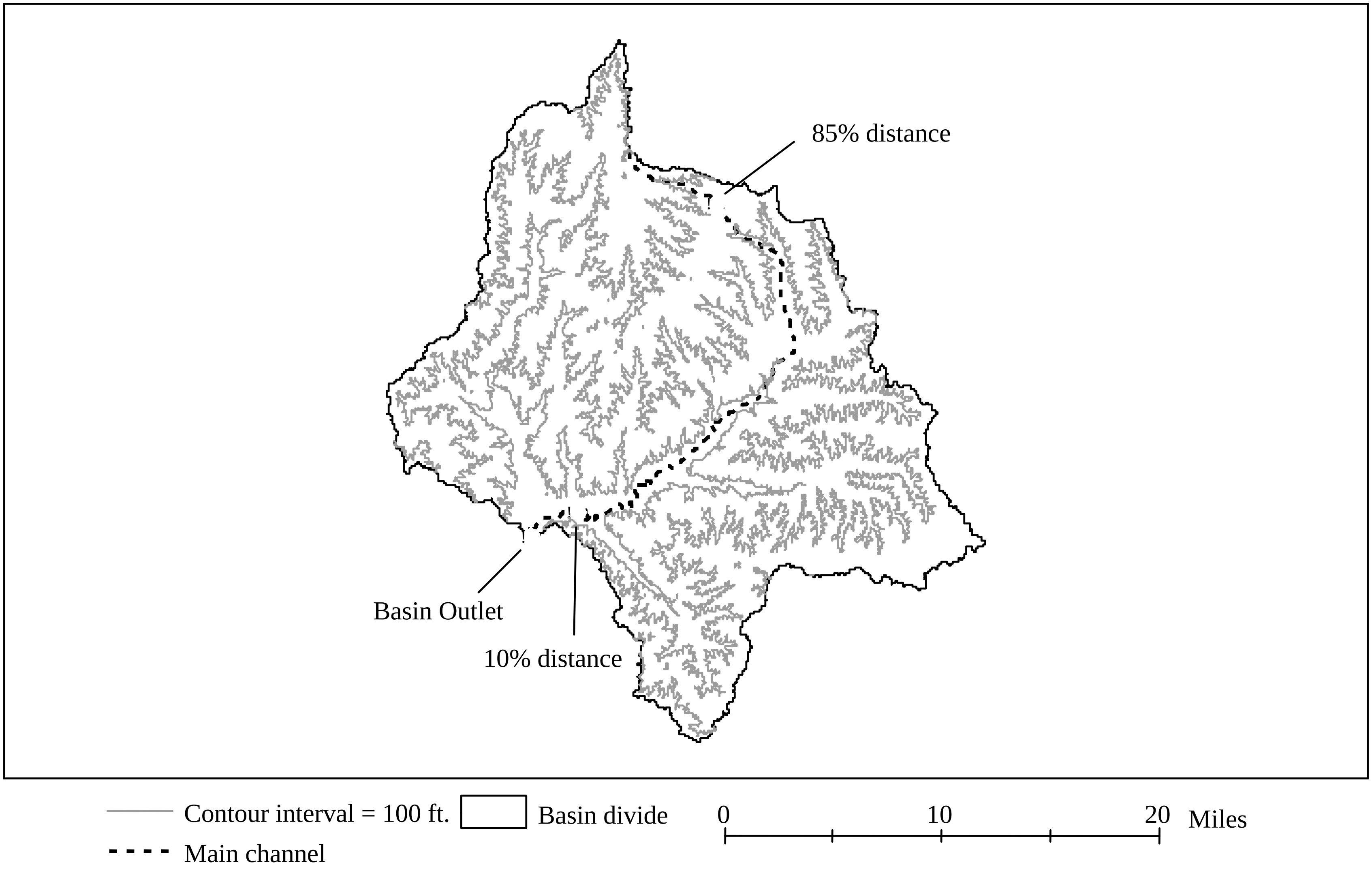


Figure A.2: Example of graphical output from ArcGIS.

APPENDIX B. GAUGING STATIONS USED IN REGRESSION ANALYSIS

Table B.1: Gaging stations with drainage areas of less than 10 mi^2.

Region	Gaging stations used in regression analysis									
Big Blue	6856800	6872600	6873300	6880590	6883540	6883955	6884005	6884300		
Eastern	6600600	6600800	6607700	6607800	6607900	6608600	6608700	6608800	6803540	6803570
	6803700	6804100	6806420	6806440	6806470	6810060	6810100	6810400	6816000	6820000
	6821000									
Northeastern	6600600	6600800	6607700	6607800	6607900	6608600	6608700	6608800	6610700	6790600
	6793995	6799190	6800350							
Central & South Central	6768300	6777700	6777800	6782800	6782900	6789100	6789200	6790600	6790900	6851300
	6853100									
Upper Republican	6823500	6829700	6839200	6839700	6844800	6847600	6848200			
Northern & Western	6382200	6399700	6443200	6443300	6445590	6449750	6456200	6463200	6652400	6762600
	6767100	6767200								

Table B.2: Gaging stations used in complete regression analysis.

Region	Gaging stations used in regression analysis
Big Blue	6853800 6856100 6856800 6871000 6871500 6873000 6873300 6873500 6874500 6879900 6880000 6880500 6880508 6880710 6880720 6880730 6880740 6880800 6881000 6881200 6881450 6881500 6882000 6883000 6883540 6883570 6883600 6883700 6883800 6883940 6883955 6884000 6884005 6884200 6884300 6884400 6885500 6886500 6887200 6888000 6888300
Eastern	6600600 6600700 6600800 6600900 6601000 6606790 6607700 6607800 6607900 6608000 6608500 6608600 6608700 6608800 6608900 6609000 6610600 6803000 6803510 6803520 6803530 6803540 6803600 6803700 6803900 6804000 6804100 6804200 6804500 6805000 6806400 6806420 6806440 6806460 6806470 6806500 6810060 6810100 6810200 6810300 6810500 6811500 6813000 6814000 6814500 6815000 6815500 6815510 6816000 6820000 6821000
Northeastern	6466500 6478280 6478300 6478518 6478520 6478800 6478820 6478840 6600000 6600600 6600700 6600800 6600900 6601000 6607700 6607800 6607900 6608000 6608500 6608600 6608700 6608800 6608900 6609000 6610600 6610700 6790600 6790700 6790800 6790900 6791100 6791500 6792000 6793500 6793995 6795000 6797500 6798000 6798500 6799000

Table B.2: Gaging stations used in complete regression analysis *(Continued)*.

Region	Gaging stations used in regression analysis									
Central & South Central	6768050	6768100	6768200	6768400	6768500	6769000	6769100	6769200	6769300	6769500
	6770700	6770800	6770900	6770910	6771000	6771500	6772000	6777700	6777800	6782800
	6782900	6783500	6784000	6784700	6784800	6788988	6789100	6789200	6789300	6789400
	6789500	6790600	6790700	6790800	6790900	6791100	6849600	6850000	6850200	6851000
	6851100	6851200	6851300	6851400	6851500	6853100				
Upper Republican	6821500	6823000	6824500	6825000	6825500	6828000	6835000	6836000	6837300	6838200
	6839000	6839200	6839400	6839500	6839600	6839850	6839900	6839950	6840000	6840500
	6841500	6844000	6844210	6844800	6844900	6845000	6845100	6845200	6846000	6846200
	6846500	6847000	6847500	6847600	6847900	6848200				
Northern & Western	6382200	6396490	6399700	6443200	6443300	6443700	6444000	6445500	6445560	6446000
	6446400	6447500	6449100	6449500	6449750	6450500	6453500	6453600	6456200	6456300
	6462500	6463500	6464500	6464900	6465300	6652400	6677500	6687000	6767100	6767200
	6767300	6767400	6767410	6767500						
High Permeability	6447500	6448000	6449100	6449500	6450500	6459175	6459200	6462500	6463500	6677500
	6687000	6692000	6775500	6775900	6776500	6777000	6777500	6778000	6779000	6780000
	6782500	6782700	6785000	6786000	6787000	6787500	6788500	6789000	6790500	6791500
	6792000	6793500	6794000	6794500	6797500	6798000	6798300	6798500	6799000	6799100
	6799350	6821500	6823000	6823500	6824500	6828000	6834500	6835000	6837300	6839000
	6839500									

APPENDIX C. IMPLEMENTATION PLAN

By Kevin Donahoo and David Admiraal

Following the completion of Water-Resources Investigations Report 99-4032, it became apparent to the potential users at the Nebraska Department of Roads that the procedures outlined within that report would be difficult to duplicate using available office resources. The purpose of this research is to develop new equations and procedures that would enable designers at the Department of Roads to use the updated GIS-based Regression Equations using available office means.

Since the Regression Equations developed through this research are also of interest to organizations outside of the Nebraska Department of Roads, coordination with other agencies and organizations has already been initiated. Dr. Admiraal has presented the results of this research to Nebraska Department of Natural Resources and Nebraska Department of Roads personnel. As a result, in addition to the automation processes developed herein for the Arcinfo software, additional automated procedures were developed by staff members at Nebraska Department of Natural Resources for use within the Arcview GIS software. This has enhanced the information-sharing capabilities between the two agencies, and has allowed both agencies to compare the results from the new equations with other regression equations.

It is considered standard operating procedure for several hydrology methods to be used on large scale drainage studies so that each method can be considered for suitability at the study site. The new equations contained within this report are already being implemented as one of those methods.