TEXAS TECH UNIVERSITY

Multidisciplinary Research in Transportation

Empirical Flow Parameters - A Tool for Hydraulic Model Validity Assessment

Theodore G. Cleveland, Kyle B. Strom,
Hatim Sharif, Xiaofeng Liu

Texas Department of Transportation

Research Project #: 0-6654
Research Report #: 0-6654 -1
www.techmrt.ttu.edu/reports.php

May 2014

NOTICE

The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-14/0-6654-1		

4. Title and Subtitle	5. Report Date
Empirical Flow Parameters -- A Tool for Hydraulic Model Validity Assessment	August 2013
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Theodore G. Cleveland, Kyle B. Strom, Hatim Sharif, Xiaofeng Liu	0-6654-1

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Texas Tech University, Water Resources Center 10th and Akron, Room 203 Lubbock, TX 79409-1022	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080	Technical Report 1 Sep 2010 – 31 Aug 2013
	14. Sponsoring Agency Code

15. Supplementary Notes
Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

16. Abstract

Data in Texas from the U.S. Geological Survey (USGS) physical stream flow and channel property measurements for gaging stations in the state of Texas were used to construct relations between observed stream flow, topographic slope, mean section velocity, and several other hydraulic factors, to produce three methods to assess "if hydraulic model results are way off."

A generalized additive modeling technique was used to develop two equations, QGAM and VGAM, that estimate discharge and velocity at a location in Texas from a cross sectional flow area, topwidth, location (latitude and longitude), and mean annual precipitation.

The data were also analyzed to develop conditional distributions of hydraulic parameters based on location related values to extend the utility of the QGAM and VGAM models

The data, and ancillary properties developed in the research were used to develop dimensionless bankfull depth-discharge and similar charts as a way to evaluate anticipated behavior in streams in Texas.

17. Key Words	18. Distribution Statement
Empirical velocity distributions; Generalized additive models; conditional distributions; dimensionless hydraulic charts	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161, www.ntis.gov

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	166	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

0-6654 Empirical Flow Parameters – A Tool for Hydraulic Model Validity Assessment

Texas Tech Center for Multidisciplinary Research in Transportation (TechMRT)
Texas Tech University, Lubbock, Texas

Research Project: 0-6654
Research Report: 0-6654-1 Final Report

Final Report
Texas Department of Transportation
Research Project Number 0-6654

August, 2013

Empirical Flow Parameters –
A Tool for Hydraulic Model Validity Assessment

by

Theodore G. Cleveland, Associate Professor,
TechMRT, Texas Tech University

Kyle B. Strom, Associate Professor,
Civil and Environmental Engineering, University of Houston

Hatim Sharif, Associate Professor
Civil and Environmental Engineering,University of Texas – San Antonio

Xiaofeng Liu, Associate Professor,
Civil and Environmental Engineering,University of Texas – San Antonio

Research Report 0-6654-1 (Final Report)
Research Project 0-6654

Research Project Title: Empirical Flow Parameters – A Tool for Hydraulic Model Validity Assessment

Sponsored by the Texas Department of Transportation
August, 2013

Texas Tech University
Texas Tech Center for Multidisciplinary Research in Transportation (TechMRT)
Department of Civil and Environmental Engineering
10th and Akron
Box 41023
Lubbock, Texas 79409-1023

Disclaimer

The contents of this report reflect the views of the author(s), who is (are) responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration (FHWA) or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation. This report is not intended for construction, bidding, or permit purposes. The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers names appear herein solely because they are considered essential to the object of this report. The researcher in charge of this project was Dr. Theodore G. Cleveland at Texas Tech University.

There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract (to date), including any art, method, process, machine, manufacture, design, or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Acknowledgments

The support of the TxDOT project advisory committee members: Wade Odell (RTI Research Engineer), Amy Ronnfeldt, Stan Hopfe, Jaime Villena-Morales, David Zwernemann, and Jorge Millan is greatly appreciated.

Substantial contributions to the research were provided by the Texas Water Science Center of the U.S. Geological Survey.

0-6654 Empirical Flow Parameters –
A Tool for Hydraulic Model Validity Assessment
Final Report

Contents

1 Introduction

1.1 Scope of work

Water-surface profile modeling assembles models based on generalizations of parameter values from textbooks, professional literature, computer program documentation, and from engineering experience. Stage-discharge relations or measurements of streamflow at or adjacent to the modeling locale are seldom available for use in refining model parameters. In streamflow measurement at least three components are important; depth, width, and velocity. At field scale depth and width are straightforward but the velocity measurement is a significant contributor to overall uncertainty, complicated because a mean section velocity (as reported in a model) requires a spatial integration of the measured velocity field. As a result, modeling efforts by even experienced engineers are assembled and often judged to be valid based entirely on experiences from earlier modeling efforts for hydraulically similar settings.

This situation often leads engineers in good faith to report velocities (needed for assessing forces on bridge piers, and assessing erosion and scour potential) that are unusually large and in some instances absurd. *This research developed independent ways to assess computed velocities based on prior, authoritative, observational experience.*

The results permit an engineer to rapidly evaluate or review modeling effort and determine if the modeled results are comparatively common or unusual, with the explicit caveat that unusual results could very well be reliable, but that additional explanation should be expended in these unusual situations. The results of this research (graphs and statistical distributions) provide an assessment of modeling risk that could be used to balance the cost of additional modeling with the cost of accepting an unusual result for design.

1.2 Research Approach

The observational basis for developing the independent comparison tools came from U.S. Geological Survey (USGS) physical stream flow and channel property measurements for gaging stations in the state of Texas. These discharge measurements reside within the USGS National Water Information System (NWIS) and are readily obtained U.S. Geological Survey (2009b) by streamgage number.

1.3 Project Objectives

The objectives of this project were:

1. To determine and present from existing data in Texas, relations between observed streamflow, topographic slope, mean section velocity, and other hydraulic factors, to produce charts such as Figure 1 and to produce empirical distributions of the various

flow parameters to provide a methodology to "check if model results are way off!"

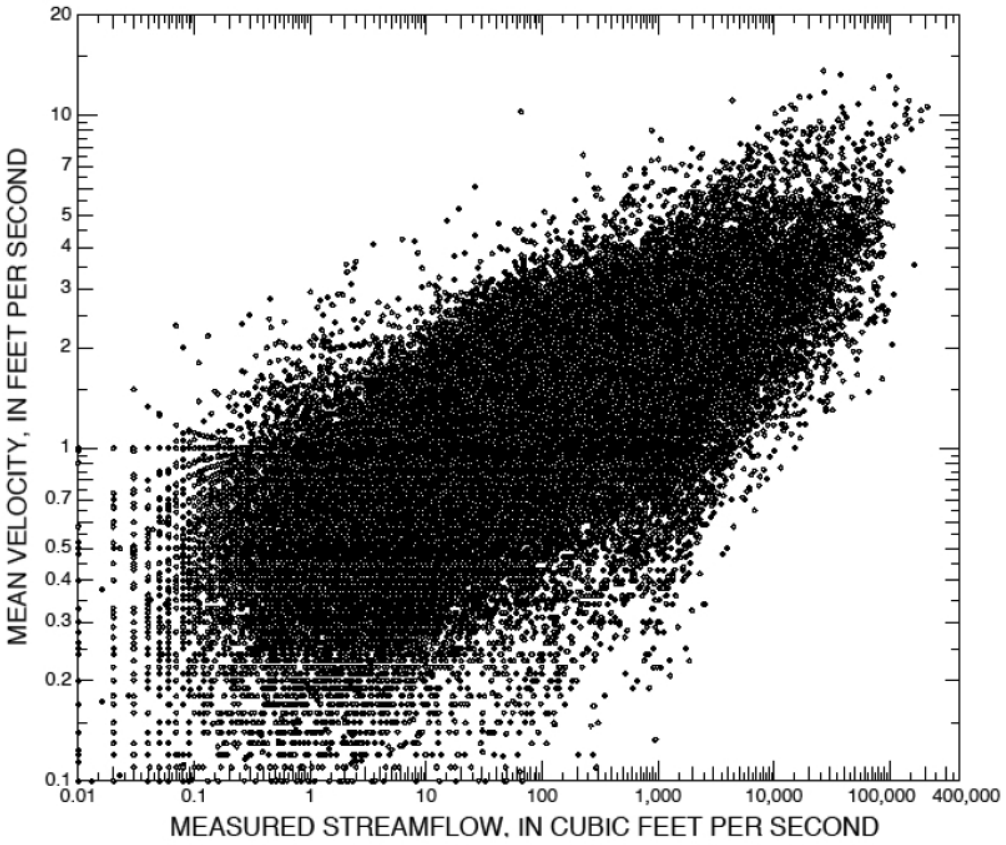

Figure 1: Mean Section Velocity versus Streamflow from U.S. Geological Survey Streamflow Measurement Database in Texas (adapted from Asquith and Herrmann, 2009).

Figure 1, while dimensional, contains information that permits rapid checking of computed velocities for a given discharge any vertical slice of the figure (a specific value of Q) returns an empirical velocity distribution. For example, at $Q = 100$ cfs, observed velocities range between about 0.2 and 5 feet per second (fps). Hence a computed velocity above 5 fps, would be unusual, and if unexplainable in terms of contraction coefficients or other engineered cause, would be suggestive of a modeling error and would warrant further investigation would be warranted.

2. To produce a statistical regional tool to estimate mean velocity or other selected parameters for storm flows or other conditional discharges at ungauged locations (most bridge crossings) in Texas to provide a secondary way to compare such values to a conventional hydraulic modeling approach.

3. To present ancillary values such as Froude number, stream power, Rosgen channel classification, sinuosity, and other selected characteristics (readily determinable from

existing data) to provide additional information to engineers concerned with the hydraulic-soil-foundation component of transportation infrastructure.

1.4 Work Plan

The project objectives were met through completion of the following eight specific tasks.

Task 1: Review literature on the subject of maximum discharge, velocity, slope, ancillary parameters, and regional dependence of Texas river hydraulics.

Task 2: Identify and quantify additional gaging stations and watershed characteristics that can be added to the existing velocity-discharge database to provide additional information that could be helpful in building the statistical tools.

Task 3: Merge the existing flow database and the characteristics database of Task 2 into the primary working database for the project analysis.

Task 4: Compile a GIS database of physiographic provinces of Texas and their subdivisions based on geologic structure, soil, land cover, and climate.

Task 5: Tasks 3 and 4 will result in functional databases suitable for analysis and interpretation. The research team will analyze the database to determine whether or not universal hydraulic geometry relations hold in general for Texas rivers. If so, an alternative method for checking to see if a simulation result is "way off" will be developed based on these universal hydraulic geometry relations.

Task 6: Ancillary properties such as stream power, Manning n-values, Froude number, and Darcy-Weisbach friction factor will be computed from the database for both Texas as a whole and by physiographic region. The distribution of these ancillary parameters will be reported in a meaningful way as supplementary information regarding the general hydraulics of Texas rivers.

Task 7: Empirical distributions will be constructed from the results of Tasks 3, 4, and 5. These distributions (equations relating a value and a cumulative frequency) will constitute the fundamental component of a statistical regional tool to estimate mean velocity or other selected parameters for storm flows or other conditional discharges at un-gauged locations.

Task 8: Report findings through tech memos, semi-annual reports, and a final report.

1.5 Final Report Overview and organization

Stand-alone technical memoranda for each of the first seven tasks listed above have been submitted to TxDOT over the course of the project. In this final report, we synthesize work and findings from all aspects of the project to address the three project objectives (section

1.3). In the next section, the general task 1 literature review is given (section 2). This is followed by a description of the original USGS database along with additions made and the procedures followed to refine the database for further analysis (section 3). Methods for the products developed that make use of the database in providing an independent check on hydraulic model velocity output are detailed in sections 4-7. Three methods were developed. These methods include: a statistical regression model (method 1, sections 4 and 5), conditioned and unconditioned empirical distributions of observed data and ancillary parameters (method 2, section 6), and application of universal hydraulic geometry relations (method 3, section 7). After explaining each of the product methods, examples of how to use each are given (section 8). Conclusions are given in section 10.

2 Literature Review

2.1 Scope of the Literature Review

The literature review covers the subject of maximum discharge, velocity, slope, ancillary relationships (stream power, Fr, etc.), and information from previous work specifically related to streams and rivers in the state of Texas. The focus of this review was to locate, in the literature, documentation of prior studies that produced findings that relate characterizations and classifications of short reaches to maximum velocity and flow geometry. Specific questions that were asked to guide the review included:

1. What is typical natural channel geometry, and how are top-width and discharge related in such channels?
2. What kind of generic classifications of existing river form are appropriate to distinguish different channel types?
3. Can energy slope be inferred from readily available geomorphic and discharge data?
4. Is there a well-documented relationship between topographic slope and energy slope in particular flow systems?
5. Are there particular physiographic elements of Texas that might influence any of these relationship?

The remainder of this literature review briefly examines the researcher-determined relevant literature and the current answers to these research questions. The review focused on three broad related themes: (1) resistance equations and their relationship to channel geometry, (2) energy, water surface, and topographic (map) slope, and (3) past work focusing on hydraulics of Texas streams and rivers. These themes are interrelated to each other and to the overall geomorphic structures that control flow.

2.2 Resistance Equations

The hydraulic engineering community has a long history of developing simple equations that relate the depth and cross-sectionally averaged flow velocity under steady, uniform flow with flow resistance properties of the channel. Such relationships are known as resistance equations and are often used to give estimates of depth or velocity given either discharge or depth respectively. A basic review of these equations is presented here because the various equations and work done in the area likely hold clues about how the parameters in the dataset of observed hydraulic and channel properties could be arranged to develop the velocity index tool.

For natural rivers, resistance equations can be broadly categorized as being either, 1) fixed bed, or 2) mobile bed relations. For fixed bed equations, the friction effects at the interface

of the flow and bed are characterized by grain roughness and parameterized by a length scale that is of the order of the diameter of the grains in the bed. For the mobile bed case, the frictional effects can be dominated by pressure drag caused by flow separation around bedforms, and the roughness length scale may be significantly larger than that of the individual grains. In this review, we will first present a brief introduction to resistance relationships in general and then summarize pertinent fixed and mobile bed equations for various river types.

2.2.1 Basic Form

The basic form of resistance equations comes from either a bulk force balance analysis on the flow, or a depth-averaging of the vertically varying velocity profile (e.g., Keulegan, 1938). A bulk force balance on the water within a reach of steady, uniform flow reveals that the bed shear stress, τ_B is related to the channel slope and hydraulic radius as,

$$\tau_B = \rho g R S \tag{1}$$

where ρ is the density of water, g is the acceleration of gravity, R is the hydraulic radius, and S is the channel slope; note that for sufficiently wide channels, $R \approx h$ where h is the flow depth, and that the channel slope is equal to the friction slope for uniform flow. Furthermore, the velocity of the channel can be related to the hydraulic radius and channel slope through what is known as the Chezy relationship,

$$V = C_Z \sqrt{RS} \tag{2}$$

where V is the cross sectionally averaged velocity, $V = (1/A) \int_A u dA$, and C_z is the Chezy coefficient. Taking $C_z = R^{1/6}/n$ results in the well-known Manning resistance equation in SI units,

$$V = \frac{1}{n} R^{2/3} S^{1/2} \tag{3}$$

where n is the Manning coefficient. Both the Chezy and Manning equations (Eq. 32 and 30) are forms of resistance relationships and they indicate that velocity, flow depth, channel slope, and some friction property of the bed, which is tied up in the coefficients, are linked with a particular functionality for steady, uniform open channel flow. Other resistance relations come from considering the log-law velocity distribution for fully-rough flow, and these relationships help to give a better feel for the physical scales involved in the resistance coefficients C_z and n. The rough-wall log-law describes the velocity distribution over the vertical:

$$\frac{u}{u_*} = \frac{1}{\kappa} \ln \left(\frac{z}{k_s} \right) + 8.5 \tag{4}$$

where u_* is the friction velocity ($u_* = \sqrt{\tau_B/\rho}$), κ is the von Karman constant ($\kappa \approx 0.4$), z is the vertical coordinate, and k_s is a roughness length scale. Integrating $u = u(z)$ in

Eq. 4 over the depth and dividing by the flow depth, i.e. depth averaging, results in the so-called Keulegan (1938) resistance equation,

$$\frac{V}{u_*} = \frac{1}{\kappa} \ln\left(\frac{h}{k_s}\right) + 6 = \frac{1}{\kappa} \ln\left(11\frac{h}{k_s}\right) \tag{5}$$

Alternatively, the Keulegan (1938) equation can also be approximated with a Manning-Strickler form power law as,

$$\frac{V}{u_*} = 8.1\left(\frac{h}{k_s}\right)^{1/6} \tag{6}$$

Other ways of expressing resistance relationships of the form of Eq. 5 and 6 can be developed by using the following definition for bed shear stress τ_B at a point,

$$\tau_B = \rho C_f V^2 \tag{7}$$

where C_f is the friction coefficient, and the definition of the Darcy-Weisbach friction factor,

$$f = \frac{8gRS}{V^2} \tag{8}$$

Making use of these (Eq. 7 and 8) and the definition for $u_* = \sqrt{\tau_B/\rho}$, the ratio of the depth-averaged velocity and the friction velocity is related to the friction coefficient and the Darcy-Weisbach friction factor as,

$$\frac{V}{u_*} = C_f^{-1/2} = \sqrt{\frac{8}{f}} \tag{9}$$

which leads to the basic form of most resistance equations other than the Chezy and Manning equations in the form of,

$$\frac{V}{u_*} = C_f^{-1/2} = \sqrt{\frac{8}{f}} = \alpha_1 \ln\left(\frac{R}{z_R}\right) + \beta_1 \tag{10}$$

or,

$$\frac{V}{u_*} = C_f^{-1/2} = \sqrt{\frac{8}{f}} = \alpha_2 \left(\frac{R}{z_R}\right)^{\beta_2} \tag{11}$$

where z_R is some general roughness length scale and α_1, α_2, β_1, and β_2 are constants that are reflective of the particular roughness conditions of a given flume or river. These forms may, at first glance, seem to be significantly different than the Manning and Chezy equations, but they are actually quite similar except for the non-dimensionality of the friction related terms in Eq. 9 compared to the dimensionality of the Chezy coefficient and Manning's n value. The similarity is clearer when u_* is replaced with its definition, $u_* = \sqrt{gRS}$, and the power-law form of the equation is used (Eq. 6 or 11). In fact, comparing the Manning equation with the Manning-Strickler power law equation (Eq. 6) allows for better insight into a possible functional form of the Manning's n value. Using the definition for the bed shear stress (Eq. 1) along with the Chezy equation (Eq. 32) shows

that the Manning coefficient is related to the roughness length scale and can be expressed in terms of k_s as,

$$n = \frac{k_s^{1/6}}{8.1 g^{1/2}} \tag{12}$$

or more generally,

$$n = \frac{z_R^{1/6}}{a_1 g^{1/2}} \tag{13}$$

where a_1 is a constant coefficient. These relationships are known as the Manning-Strickler relationships for n (Brownlie, 1983); in general, the roughness length scale k_s is taken to be proportional to the diameter of the grains within the bed, $k_s \propto d$ where d is the diameter of a representative grain.

A myriad of equations of the form in Eq. 10 and 11 with various α_1, α_2, β_1, and β_2 and z_R values have been developed for alluvial rivers and flume studies under both fixed and mobile bed conditions. Many of these are listed and summarized in Brownlie (1981), Buffington and Montgomery (1999), García (2008), and Recking et al. (2008b). One of the main reasons for the large variety in these developed semi-empirical relationships is that the complex bathymetry, grain size distributions, and grain arrangements that control the frictional characteristics of a reach are parameterized with a single roughness length scale, z_R. This is reasonable to do, but results in each z_R value being somewhat unique to the particular grain size distribution, grain arrangement, and overall channel bathymetry for which the value was originally obtained.

2.2.2 Fixed Beds

For fixed-bed roughness, the roughness length scale z_R is typically taken as some multiple of a characteristic grain size,

$$z_R = n_s d_x \tag{14}$$

where d_x is a characteristic grain size such as d_{50} (Keulegan, 1938; Meyer-Peter and Müller, 1948; Bray, 1979), d_{84} (Hey, 1979), or d_{90} (Kamphuis, 1974) and n_s is a constant multiplier historically calculated as being between 1 and 6.6 depending on the particular river/flume conditions (García, 2008). This n_s value also depends on the grain size statistic used for d_x. For example, López and Barragán (2008) analyzed resistance data from several gravel bed rivers and reported that n_s was 2.54 times greater when d_{50} was used for d_x when compared to d_{90} being used for d_x. A nice summary table for $k_s = nd_x$ dependancies over a large variety of sand and gravel alluvial channel can be found in Table 2-1 of García (2008). While each particular stream may have a slightly different z_R functionality, a reasonable value to use in the case of gravel bed rivers is the relation proposed by Kamphuis (1974),

$$z_R = 2d_{90} \tag{15}$$

in combination with the Manning-Strickler relation given in Eq. 6 (Parker, 1991; Wong and Parker, 2006).

Although the size of the grains in the river does play a major role in setting the frictional characteristics of a flow, the grain size is not the only physical factor affecting roughness. Pressure drag around larger object such as larger stable grains (Bathurst, 1985; Ferro, 1999; Bathurst, 2002; Pagliara and Chiavaccini, 2006), wood deposits (Manga and Kirchner, 2000; Curran and Wohl, 2003; Wilcox and Wohl, 2006), vegetation (Kean and Smith, 2005; Rameshwaran and Shiono, 2007; Kean and Smith, 2010), and larger scale bed forms and channel geometry irregularities (Heritage et al., 2004; Comiti et al., 2009) can add to the grain roughness in setting the frictional characteristics of a stream. In fact, contributions from these non-grain-scale roughness elements can in many cases be very significant (Millar, 1999; Afzalimehr et al., 2010). The contributions from grain and form, or pressure, drag can be conceptually decomposed as,

$$f = f^{'} + f^{''} + f^{'''} \quad \text{or} \quad n = n^{'} + n^{''} + n^{'''} \tag{16}$$

where f is the total integrated friction factor, $f^{'}$ is the friction factor due to grain roughness, $f^{''}$ is the friction due to large obstructions such as bedforms and trees, and $f^{'''}$ is friction due to larger-scale bathymetric forcing such as meander beds, and pool-riffle sequences (e.g., Eaton and Church, 2004). This conceptual decomposition of f or n is useful in thinking about the roughness characteristics of the channel and how changes in the within and overall channel properties impact the relationship between depth, velocity, and slope. However, the decomposition is still of limited practical use when it comes to predicting the roughness properties of the bed because little is known about "how much" each particular process contributes and how to parameterize that in a way that is useful at a practical level.

Because of the difficulty in relating the various roughness components to measurable parameters in the field, the integrated roughness-length scale value in relationships of the form of Eq. 10, or the Mannings n value, become parameters that integrate a wide variety of processes contributing to the frictional characteristics of the particular stream for which the resistance equation was developed. Some of the more promising work in moving away from site specific resistance type relations has come from the theoretical stage-discharge relationships developed by Jason Kean and Dungan Smith with the USGS (Kean and Smith, 2005, 2010). Their work in particular will be examined during the development of the velocity indexing tool because of their focus on using USGS data.

2.2.3 Mobile Beds

Mobile alluvial beds can impact the resistance of the river bed by either the creation and destruction of bedforms and/or by the presence of rolling and saltating grains in the bedload layer along a flat bed. The first is more important in larger sand-bed rivers such as the Brazos, and the second is more relevant for gravel bed rivers such as those found along the valley floors in the Hill Country region of Central Texas.

Sand-Bed Rivers: In sand-bed rivers, most flow conditions will cause at least partial grain motion. When this happens, the movable and stationary grains in the bed self-

organize into repeating bedforms which modify the frictional characteristics of the bed. Increasing flowrate from some baseline value can cause transitions in bedform size and spacing which in term can make the roughness properties of the bed dynamic within a given flow event. Resistance models which account for the development of bedforms in sand-bed rivers are known as stage-discharge relations. Some of the more notable historic relations are the Einstein and Barbarossa (1952), Engelund (1966), and Brownlie (1983) relationships. The Einstein and Barbarossa (1952) and Engelund (1966) relationships are based on the concept that resistance can be decomposed into grain and form resistance. In these relations, grain resistance is set to some function of the grain size, as in the case with fixed bed resistance, and form resistance is set to some function of the streams capacity to move sediment and develop bedforms. These concepts are supported by the experimental study of Fedele and García (2001), which concluded frictional characteristics depended not only on the grain size, but also on the flow intensity relative to the mobility of the grains and the relative submergence of the grains themselves Fedele and García (2001). Brownlie (1983) takes a slightly different approach in the development of a stage-discharge relation by not explicitly considering the decomposition between form and grain resistance and by conceptually treating bedforms as larger equivalent roughness elements. The advantage of the Brownlie (1983) method is that it was developed with a large data set and is put in a form that is easy to integrate into computer-based calculations. More recently, Wright and Parker (2004) improved upon the work of Brownlie (1983) by incorporating density stratification effects due to suspended sediment. The improvements they made were found to be most beneficial for large, low-sloped rivers, such as the Atchafalaya River, but the method was found to also work well on small and moderate sized sand-bed rivers such as the Rio Grande.

Many other stage-discharge relations have also been developed, and good summaries of most of these relations can be found in Vanoni (1975) and García (2008). However, even with the large number of proposed equations, no single equation is universally applicable. Each formulation tends to work best for conditions that are most similar to those for which the experimental data used in the development of the relation was obtained (Chang, 1988; García, 2008). However, of the various stage discharge relations, the Wright and Parker (2004) model is likely the most applicable over the largest range of conditions for sand-bed rivers.

Gravel-Bed Rivers: For small drainage areas in central Texas, first and second order mixed alluvial-bedrocks streams often have complex bathymetries and various interlocked grains and stepped morphologies similar to those in mountainous environments. For these streams, flow conditions which cause motion of the bed are typically limited in time and most flows will be governed by fixed bed resistance under low relative submergence. However, larger gravel-bed alluvial and mixed alluvial/bedrock rivers with plane bed and pool-riffle morphologies could see sustained motion during periods of high flows. When this occurs the motion of the gravel impacts the frictional or resistance characteristics of the river in a slightly different way than in the case of sand-bed rivers. In sand-bed rivers, the change in the frictional characteristics comes primarily from the development of ripples,

dunes, and anti-dune bedforms. In gravel-bed rivers, bedforms created during periods of motion are far less pronounced in the vertical and less regular in the longitudinal. During these periods of motion, roughness is still dominated by grain roughness and larger bathymetric properties of the channel, but the stream can experience and increase in the resistance properties of the river due to the momentum extraction in the layer of moving grains (Smith and McLean, 1977; Wiberg and Rubin, 1989; Whiting and Dietrich, 1990; Campbell et al., 2005; Recking et al., 2008a).

Recking et al. (2008b) presents the most recent summary of the work on resistance in mobile bed cases and concludes that there are 3 stages of resistance in gravel bed flows. For low flow conditions, grains are stationary and the roughness length scale can be set equal to the diameter of the grains in the bed, i.e. fixed bed resistance. At high flow conditions, where the flow is 2.5 times that needed to initiate motion, Recking et al. (2008b) concluded that a constant roughness length scale value was sufficient and that it was 2.6 times greater than the roughness length scale for the fixed bed case. The third condition identified was a transitions zone between the low and high flow conditions where grains were more sporadically entrained in space and time. Within this zone, Recking et al. (2008b) proposed that the roughness length scale is a flow-dependent parameter.

2.3 Resistance Equations – Application

Resistance equations, coupled with the continuity equation ($Q = VA$), offer the quickest momentum-based method for estimating mean velocity and depth at a particular cross section under the assumption of uniform steady flow given the stream properties of: approximate channel cross sectional geometry, grain size, and slope. However, due to the difficulty in representing the integrated roughness effects with a single roughness length scale, the particular equation and roughness length scale used should technically be carefully chosen and then calibrated for each particular site. Because this is not possible to do for all streams and locations in Texas, the researchers will likely select a few key rivers and check to see which of the resistance equations best describes the measured data for that particular site. Because the measured USGS dataset does not contain slope and grain size information, doing this may involved obtaining the grain size and slope information from other sources or making field measurements. Finding the basic form of a resistance relation which works reasonably well for stereotypical river cases could be fruitful in defining the mean trend in the data.

In addition, the review of resistance equations suggests that it would likely be fruitful to split the aggregate database for measured river flow properties in Texas by grain size, relative submergence, and relative stream size when developing the velocity index tool. For example, it would be good to split sand-bed and gravel-bed rivers into different categories and then further split the gravel-bed rivers between larger streams with relatively high depth to grain size ratios from small streams with likely low depth to grain size ratios. However, measured data at gaging stations does not always include grain size. In fact, the USGS database, "QVFAB.txt" only contains information pertaining to discharge, velocity,

flow area, flow top width, and the Froude number. For this reason, categorization of different streams will need to be done by physiographic regions and/or by using known properties of a particular stream or watershed.

In the context of the research questions posed in the introduction, two can be addressed at this point in the review.

1) "What is typical natural channel geometry, and how are top-width and discharge related in such channels?" The literature reviewed in general support the notion that a "typical" geometry may exist and the top-width, area, and discharge relationships should be discoverable using existing knowledge. This "typical" geometry will likely exist in regionalized classifications, that themselves need examination, and foreshadow the next question.

2) "What kind of generic classifications of existing river form are appropriate to distinguish different channel types?" Less guidance was found to answer this question – our research focus is how to answer this question using the tools typically available to a practicing hydrologist (e.g. conventional topographic maps,digital elevation models, Google Earth, and/or possibly site-specific surveys). At the time of this writing, this question remains unanswered; the concept appears to be supported by the literature reviewed but such a classification scheme may be one of the principal outcomes of this research.

2.4 Energy, water surface, and map slopes

The resistance equations relate velocity to a frictional resistance term and the slope of the energy grade line (the energy slope). The energy slope S_f is usually unavailable and in practice the topographic slope S_{MAP}, sometimes called the map slope, is substituted. In some cases the water surface slope S_{WS} is available from stage information at two gages on the same stream and is also substituted.

2.4.1 Wisconsin, Unknown-bed

Magilligan (1988) examined the relationship between energy slope, water surface slope, and map slope for a set of large magnitude discharges in Wisconsin, with the understanding that the physically correct slope for computing critical tractive force, stream power, or discharge is the energy slope. The focus of the work was to assess the validity of map slope substitution for water slope and energy slope in indirect discharge estimation techniques. Regression analysis was used to infer a relationship between map slope and water surface slope construction a regression model structurally similar to Equation 17. The regression coefficient β_1 on map slope was smaller than unity.

$$S_{\mathrm{WS}} = \beta_0 + \beta_1 S_{\mathrm{MAP}} \tag{17}$$

Magilligan (1988) further extended the analysis to relate energy slope to map slope and reported a quadratic relationship structurally similar to Equation 18 with the regression

coefficient β_1 on the square of map slope nearly 100.

$$S_f = \beta_0 + \beta_1 S_{\text{MAP}}^2 \tag{18}$$

Magilligan (1988) concluded that water surface slope should be used in lieu of map slope for paleo-hydraulic studies and energy slope, if computable, is the physically correct slope to choose in discharge computations. The author further stated that there is substantial variability between events even on the same stream system and knowledge of the flow producing mechanism and flood wave arrival times was vital.

2.4.2 Western USA, Gravel-bed

Prestegaard (1983) studied 12 rivers on the West coast of the United States. The study used energy and water-surface slope interchangeably and the focus of the work was to explain the relationship between resistance terms and the water-surface slope as various scales. These rivers were all gravel-bed rivers and the study reported sufficient detail to allow the authors of the technical memorandum estimate water-surface slope S_{WS} and compare these estimates to the reported map slope S_{MAP}. Coincidentally the regression equation reported in Magilligan (1988) was a remarkably good predictor of the back-computed water-surface slopes in the 12 West coast rivers. These predicted water-surface slopes and the back-computed water slopes are plotted in Figure 2

2.4.3 Utah, Rigid-bed

O'Conner et al. (1986) examined a pool-riffle-pool sequence in a stream system in Utah. The low flow water surface slope was reported as 0.006, the topographic slope was not directly reported, but a channel elevation profile was presented in O'Conner et al. (1986). The technical memorandum authors estimated the topographic slope as 0.003 by computing an average slope along the profile from the pool-to-pool elevations depicted in the profile. O'Conner et al. (1986) also presented stream power computations based on HEC-2 computer simulations where the discharge was systematically adjusted until the computed water surface profile matched silt lines preserved in the sandstone walls and slack-water sediments. These computations were reported in sufficient detail to back-compute the energy slope S_f. The average value of the back computed energy slope along the entire profile was 0.002. The regression equation reported in Magilligan (1988) was a fair predictor of the average back-computed energy slope, but there was insufficient detail in O'Conner et al. (1986) to really test the relationship. Furthermore, the Utah stream is a comparatively rigid-boundary stream, with deeply incised Navaho sandstone forming the walls of the stream at the modeled flood flows.

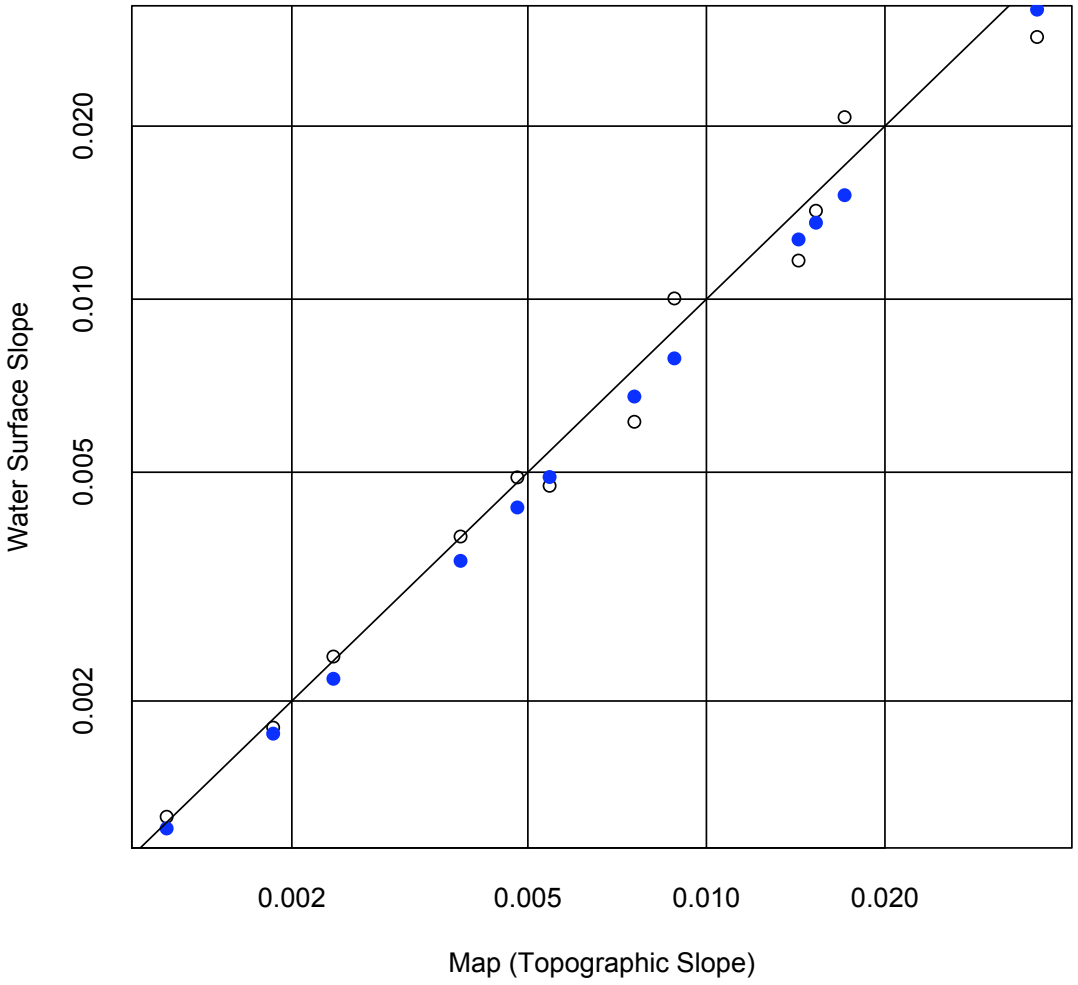

Figure 2: Computed water-surface slope S_{WS} and reported map slope S_{MAP} from Prestegaard (1983) [Open Markers] and estimated water-surface slope using regression equation reported in Magilligan (1988) [Solid Markers] for 12 rivers in California, Oregon, and Washington State. The solid line is an equal value line.

2.4.4 Entire USA, Slope-Area Methods

A conceptual precursor for the objectives of the research project is provide by Riggs (1976). Riggs suggests a "simplified" slope-area method for estimation of high magnitude peak streamflow in natural channels. The slope-area method is a widely used produce within the USGS and other entities for post-flood computation of peak streamflow based on high-water marks. These marks are used to define the downstream slope of the water surface, and when these are combined with topographic survey providing multiple cross-sectional area and other hydraulic properties, an estimate of peak streamflow results. Unfortunately, the slope-area method is expensive to conduct for high magnitude events. Perhaps the greatest source of uncertainty in the slope-area method is the roughness (Manning's) coefficient. Riggs (1976) describes further historical background and professional dialog concerning the general accuracy and applicability of the slope-area method. Riggs sought a quick, reproducible, and inexpensive alternative or compliment to the slope-area method.

Figure 3 is the plot of the relation between roughness coefficient and the square root of water-surface slope from Riggs (1976). These data are attributed to Barnes (1967) but augmented by Riggs. For the research project, these data were digitized from a scanned image of the Riggs figure. The relation is reasonably well defined as shown in Figure 3, and Riggs concludes that "a relation of this type might be used to modify an estimated n [roughness] according to slope; or one might conclude that the two variables are so highly related that only one of the two [slope or roughness] is needed [for] computing [peak streamflow]." Further importance of the figure is that it suggests that there is a distinct relationship between energy slope and the roughness term, hence there is evidence left by a system as to what energy can be expected. This evidence will be vital for application in ephemeral systems.

Riggs proceeds with arguing that Manning's equation can be reduced to

$$Q = aA^bS^c \tag{19}$$

where Q is streamflow in cubic feet per second, A is cross-sectional area in square feet, S is water-surface slope ("redefined" from friction slope in Riggs' approach), and a, b, c are regression coefficients. Riggs fits two regression equations to the data in figure 3 and the preferred model is quadratic and is

$$\log Q = 0.366 + 1.33 \log A + 0.05 \log S - 0.056(\log S)^2 \tag{20}$$

where the variables are as previously defined. It must be pointed out that at this point: no Texas data were included in Riggs (1976).

Riggs continues and shows that equation 20 was no less or no more reliable for streamflow computation than the slope-area method. Riggs continues with a "further simplification" and ignores the contribution of the water-surface slope term of the regression. Riggs continues that this new method "can reliably compute [streamflow] from mean cross-sectional area alone." This simpler model is applicable to "nearly full channels of a certain type." These channel types are largely those in the Pacific Northwest, USA.

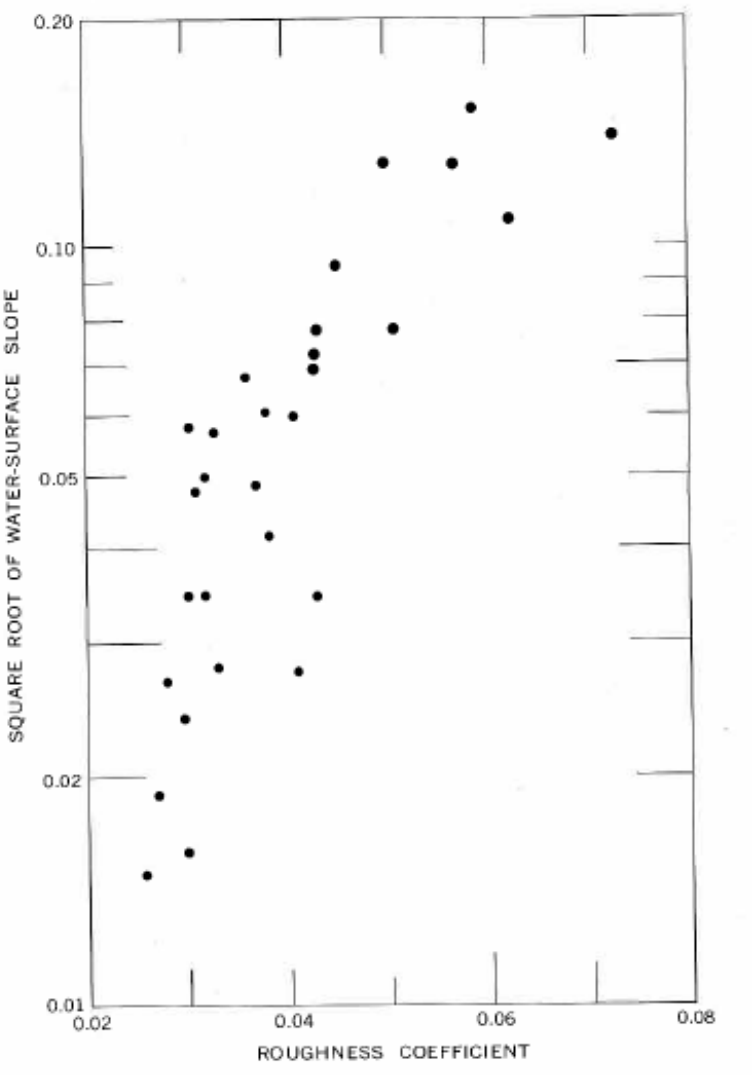

Figure 3: Relation of square root of water surface slope and the computed Manning's roughness coefficient from Riggs (1976, fig. 1)

The authors of the current (2011) research project have the luxury of over 300 times more data concerning streamflow, cross-sectional area, and other select hydraulic and watershed properties in which to explore statistical relations for prediction of streamflow in lieu of the slope-area method. However, unlike the data considered by Riggs (1976), water-surface slope (or frictional slope) is lacking. A major thematic element of statistical evaluation of the USGS streamflow measurement database for the current (2011) research project is how to estimate streamflow using regression from hydraulic and watershed properties to the extent that a suitable "slope" is absent. Another phase will subsequently attempt to use a measurement of a "channel slope" near a site for inclusion in another regression model. A comparison between the two regression equations will be made and the role of

slope evaluated. The focus on slope is made because objective measurement of slope in the absence of defined water-surface profile for a given stream reach is fraught with ambiguities and difficulties.

2.5 Energy, water surface, and map slopes – Application

The studies reviewed support the notion that energy slope can be approximated from readily available geomorphic and discharge data, most likely following similar procedures as in O'Conner et al. (1986). Furthermore, because the research database is larger, a correlation between map slope and energy slope is likely to be found. The lack of "slope" in the existing Texas-centric database was already discussed, but even with ambiguities, a procedure to approximate a slope is amenable to written explanation.

As in the resistance equations discussion, in the context of the guiding research questions, the remaining two can be answered here.

3) "Can energy slope be inferred from readily available geomorphic and discharge data?" The research team believes the answere to this question is yes. The presence of studies in the literature that ask and answer a very similar question implies that the notion itself is important, and the method to proceed is documented.

4) "Is there a well-documented relationship between topographic slope and energy slope in particular flow systems?" The answer to this question is **maybe**. The relationship for Wisconsin seemed to explain the Western USA results, but a numerical experiment with only 12 entries is a bit lacking for general application. The likely approach is some regional metric to account for variation in the energy-map slope relationship (i.e. we expect the relationship to differ between the coastal plains and central Texas.)

2.6 Review of Texas Stream Channels

The following sections discuss some findings from the literature particular to streams and rivers in Texas.

2.6.1 Federal Highway Administration Literature

The Federal Highway Administration (FHWA) often serves as an excellent, authoritative information source for various aspects of highway design. In a report regarding channel stability for highway design, the FHWA provides some general descriptions of stream channels for physiographic regions of the United States (2006). However, the report is not intended to cover each region in great detail, but encourages local surveying. One is not certain how consistently the reports descriptions would apply to Texas streams. In regard to what the USGS identifies as the Coastal Plain of the United States (the Texas portion of which

approximately coincides with what the TBEG identifies as the Gulf Coastal Plains), the FHWA report, relying on data outside of Texas, defines Coastal Plain as having streams with slopes that are "primarily gentle with bed material consisting of sand to fine gravel" (p. 29). In regard to what the USGS identifies as the Great Plains of the United States (the Texas portion of which approximately coincides with what the TBEG identifies as the Edwards Plateau, the Central Texas Uplift, the Grand Prairie, the North-Central Plains, and the High Plains), the FHWA report refers to the work of Simon and Rinaldi (2000) as describing the portion of the Great Plains (USGS classification) lying in the Midwest United States as having thousands of miles of "highly unstable streams due to the combination of easily erodible soils and extensive human disturbance" (p. 30). Finally, in regard to what the USGS identifies as Basin and Range of the United States (the Texas portion of which coincides approximately with what TBEG also refers to as Basin and Range), the report describes streams as being intermittent with drainage that is internal, i.e., the streamflow is absorb by or seeps through the channel bed itself before reaching the basin outlet. Also alluvial fans are developed from streams transporting sediment down steep mountainous channel slopes to shallower slopes at the base of mountains. In summary, while these descriptions provide suggestions as to what we might find in any particular Texas stream channel, we seek additional information to help discern how Texas might be subdivided for better prediction of mean velocities in Texas streams.

2.6.2 Channel Characteristics for North-Central Plains, Grand Prairie, Blackland Prairie, Central Texas Uplift, and Edwards Plateau

One of the most interesting studies of Texas stream channel characteristics is a geomorphic investigation of the Brazos River from Waco to Knox City (Stricklin, 1961). Immediately upstream of Graham (in the North-Central Plains of Figure 2), the channel is typically wide and shallow "because of the caving tendency of bedrock and alluvial banks, both of which contain large proportions of sand." The stream has a "prevailing tendency to braid upstream from Graham" because of "a high degree of bedrock erodibility and a consequent overloading of the stream." From Graham to Waco, the channel is less erodible because it passes through "resistant Pennsylvanian and Cretaceous strata." Braiding is replaced by incised, ingrown meanders (i.e., path curves having an undercut bank on one side and a gentle slope on the other). The slope is less steep (approximately 2 ft per mile as opposed to the roughly 3.5 ft per mile upstream of Graham) as the decline in elevation is spread out over a more tortuous and hence longer path. Finally, another noticeable change occurs near Waco:

> In passing from its incised meander belt at Waco, the Brazos enters a valley that has been readily widened in Taylor marl, decreases in channel gradient from 2 to 1.5 feet per mile, and winds along a rout of smoothly curved meanders developed on an expanded flood plainAlthough the mechanics of their formation may obey the same natural laws, these meanders contrast strongly with the incised meanders: their relief is low and represents the relief of the flood plain,

> not the valley; subject to rapid change in width and configurationTheir capacity for rapid change is impressively demonstrated by comparing aerial photographs taken several years apartchannel exposures of bedrock are limited to only a few places, consequently the stream, at least during low-water stage, flows primarily in its own alluvium. But during flood, the river probably sweeps up the thin veneer of coarse debris that ordinarily forms its bed, carrying it along as bed load and eroding the valley floor of soft bedrock. (pp. 22-24)

The Brazos stream presents a challenge to mapping stream channel characteristics based on TBEG physiographic provinces. The stream channel displays no noteworthy change as it passes over the North-Central Plains /Grand Prairie physiographic boundary. Rather, a substantial change occurs at Graham, some 80 km northwest of the point where the stream crosses the boundary. The transition point appears to coincide well with the change in lithology, where the Markley Formation (upstream of Graham, and having sandstone and mudstone as its primary and secondary rock types, respectively) yields to the Thrifty and Graham Formations (having shale and sandstone as the primary and secondary rock types, respectively within the North-Central Plains at Graham (Figure 20; a legend for the figure is located at the end of this chapter in figures 21 and 22). More encouraging is that the widening that occurs at Waco coincides with the transition to the Blackland Prairies.

Another study which emphasizes the relationship of stream channel characteristics more with lithology than with physiographic region is that of Allen et al. (2002), who examined the erodibility and geometry of urban channels in the Dallas-Fort Worth area. The study area covers over 20 miles of streams, and straddles the Grand Prairie / Black Prairies boundary while also including portions of the Austin Chalk and Eagle Ford Shale Formations. The physiographic region is not incorporated as a factor contributing to erodibility and geometry, but geotechnical properties of Austin Chalk and Eagle Ford Shale are measured to compute erosion thresholds, with a view towards assessing the varying impact of urbanization of channel widening. The authors found that as discharge increases with urbanization, alluvial bottom channels will tend to deepen more than rock bed and gravel bed channels, and present regression equations for width-depth relationships as shone in Figure 4. However, the authors did not present maps or coordinates of tested points along streams, and, in any case, it appears doubtful that the bank and bed type would be predictable based on a lithologic map alone, due to the "highly variable nature of bed and bank lithologies, vegetation, and in channel structures in urban areas" (p. 1490).

Conyers and Fonstad (2005) calculated an effective Mannings n for high flow conditions at 35 USGS gage stations in the Texas Hill Country, which occupies the southeasternmost portion of the Edwards Plateau, after justifying the assumption of a parabolic stream cross-section for calculating hydraulic radius. The authors provide the following description of the channels:

> The river channels that flow through the Texas Hill Country and cut through the Balcones Escarpment begin as wide, shallow gravel-bedded streams in the west. As they flow toward the east and southeast, they cut deeply into the

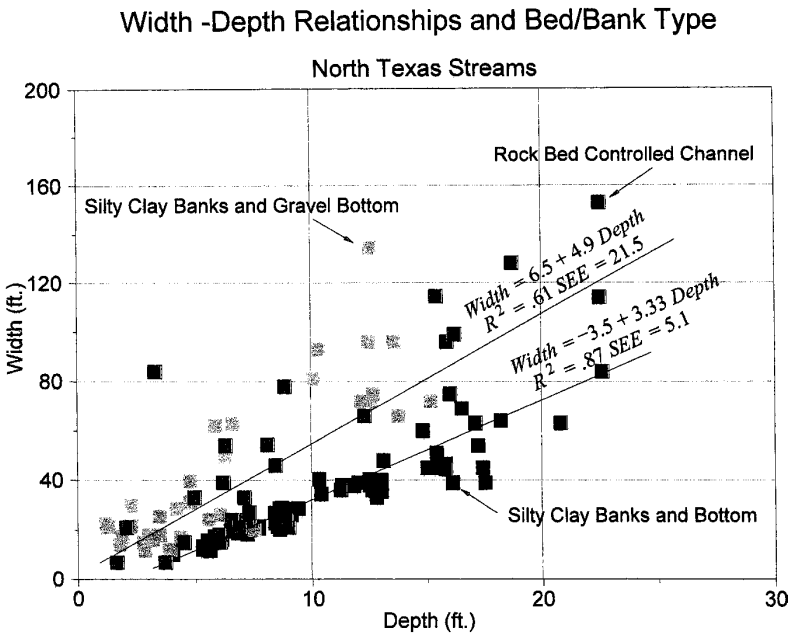

Figure 4: A comparison of width-depth channel geometry for alluvial vs. gravel and rock channel beds in the Dallas-Fort Worth area (Allen et al., 2002)

> plateau, exposing the streams to large bedrock surfaces and changing the character of these streams to high-walled channels with some surfaces covered in alluvium and some exposed to bedrock[L]arge bedrock stepsand large cylindrical potholesgenerate energy-reducing flow structurescreating a higher overall hydraulic resistance than in neighboring alluvial reaches. (pp. 385-386)

Although the conditions for the validity of the Manning equation are not fulfilled, the widely varying effective n values displayed in Figure 5 with no detectable spatial correlation suggests the difficulty in predicting stream velocity at an ungaged point based on statistical analyses. Utilizing differences in lithological categories shown in Figure 20 rather than merely physiographic region to develop a narrower range of anticipated velocities offers little promise. The Texas Hill Country is almost exclusively of the Glen Rose Limestone Formation.

Slightly overlapping the Texas Hill Country study area of Conyers and Fonstad (2005) is that of Heitmuller and Hudson (2009), who observed downstream trends in sediment size and composition of deposits in the Llano River watershed, the majority of which lies within the Central Texas Uplift (check if Central Texas Uplift is identical to the Llano Uplift). Near Junction, where the North and South Llano Rivers converge in the southwestern portion of the Central Texas Uplift, the Valley Confinement Index (VCI), which is the ratio of the channel valley width to the bankfull channel width, is 8.5. That ratio decreases to approximately 1.0 near where the Llano River discharges into the Colorado River in the southeastern portion of the Central Texas Uplift (Phillips et al., 2005), approximately 70 km

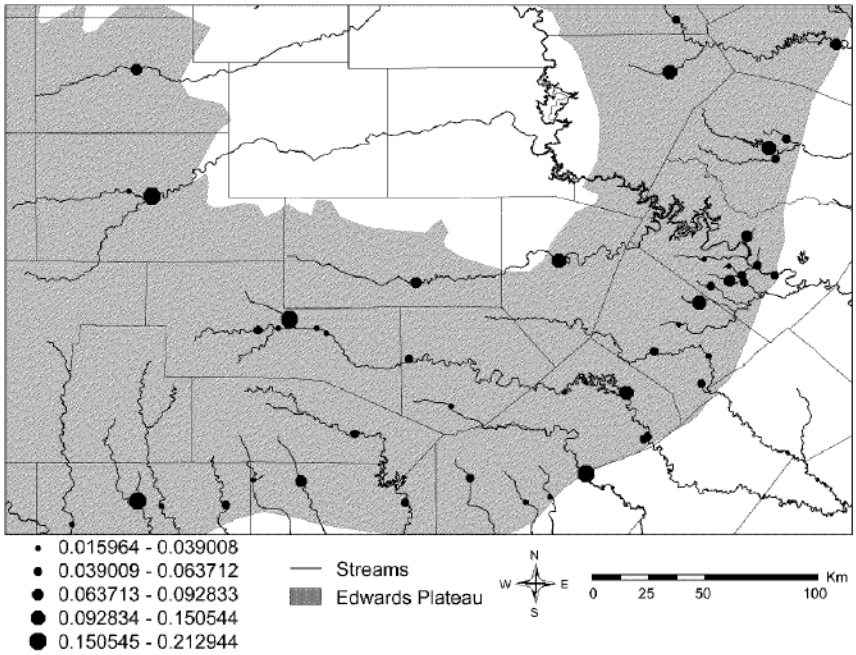

Figure 5: patial distribution of the high-flow observed n values in Texas Hill Country streams (Conyers and Fonstad, 2005).

east of Junction (Heitmuller and Hudson, 2009). Also, "an abrupt gravel-to-sand transition occurs [in channel-bar deposits] immediately downstream of a distinct lithologic change from mostly carbonate rocks to igneous and metamorphic rocksdespite an increasingly constricted alluvial valley" (p. 246). How this transition in bedload material and channel roughness may impact V is beyond the scope of the study.

Shepherd (1975) examined the Sandy Creek watershed, contiguous to the southeastern portion of the Llano River watershed. He observed an increase in channel width-depth ratio and a decrease in pool-riffle thalweg patterns and maximum size of channel-bed sediment as the dominant bedrock changed from limestone to schist to sandstone to gneiss to granite, all within the 1,025 km^2 watershed in the Llano region of central Texas.

2.6.3 Channel Characteristics for the Coastal Prairie and the Interior Coastal Plain

Various studies observe geometry and other characteristics of stream channels in the Coastal Prairies and the Interior Coastal Plains. Stream channels studied include those of the Nueces River immediately south of Uvalde (Gustavson, 1978), the Guadalupe River from near Victoria to the Gulf of Mexico (Morton and Donaldson, 1978), and the southeast Texas portions of the Brazos, Navasota, Neches, Sabine, and Trinity Rivers (Phillips et al., 2005; Phillips and Lutz, 2008; Phillips, 2011; Slattery and Phillips, 2011). In general, as

one moves downstream from the Interior Coastal Plains/Blackland Prairies boundary to the Gulf of Mexico, the sediment load becomes less, the mean grain size of bed material decreases, and the gradient decreases.

A fairly common exception to the gradient decrease downstream is that of profile convexity in tributaries as they approach trunk streams. This convexity may be due to the tributary responding to increasing depth of the trunk stream incision (Phillips and Lutz, 2008). Such convexity might not be easily captured by digital elevation models, and the resulting underestimated slope may lead to underprediction of velocities within the tributary on the downstream side of the convexity. Figure 6 displays a convex profile for Big Creek as it enters the Brazos River within the Coastal Prairies region.

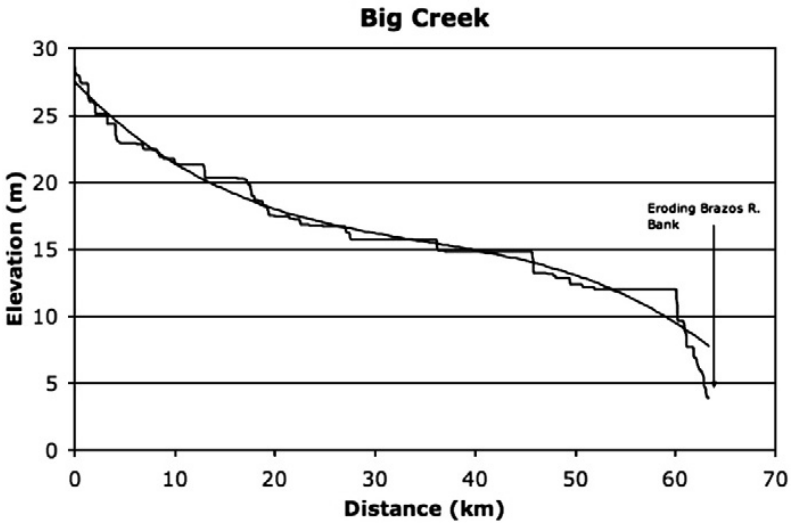

Figure 6: The longitudinal profile of Big Creek includes a convex portion immediately upstream of convergence with the Brazos River. The smooth line represents the best-fit polynomial equation [from Phillips and Lutz (2008)].

Along the Trinity River, from the Trinity delta to about halfway to the Coastal Prairies / Interior Coastal Plains boundary, the stream channel bed is actually below sea level, and the Holocene sea level rise (i.e., the 60 m or so rise in sea level that occurred in the earlier portion of the Halocene epoch, which began 12,000 years ago) exerts control on the fluvial system (Phillips et al., 2005). One might expect similar control to be exerted on other major Texas streams for tens of kilometers inland, as a comparison of their water surface profiles with that of the Trinity River in Figure 7 (Slattery and Phillips, 2011) would indirectly but strongly suggest.

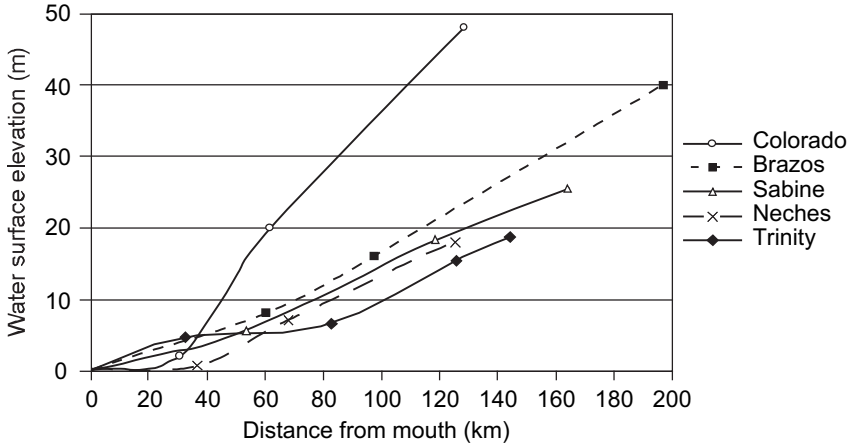

Figure 7: Water surface profiles for the lower reaches of five major rivers of southeast Texas at on July 12, 2004 at 0:00 hr (July 9, 2004, at 0:00 hr for the Sabine River) at which time discharge was higher than average but not rare [after Slattery and Phillips (2011)].

2.6.4 Regions not Well-Represented in the Literature

For some regions of Texas, no literature beyond that of FHWA to relate stream channel characteristics to physiography or lithology was found. These regions are the Basin and Range and the High Plains physiographic provinces. Not surprisingly, in these provinces there are also relatively few USGS gages to be used in the present study (Figure 8).

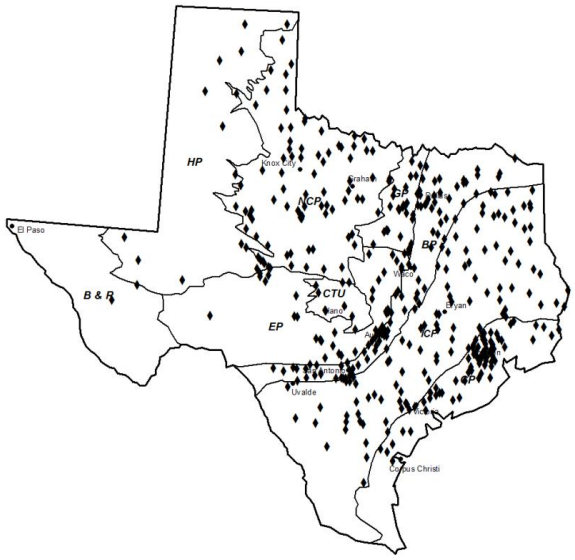

Figure 8: USGS gage site locations (indicated by diamonds) used in this study. There are few sites in the High Plains (HP) and Basin and Range (BR) physiographic regions.

2.7 Summary and conclusions from the review

Based on review of the technical literature, a number of findings are reported.

1. Resistance equations will contribute heavily to the indexing tool envisioned for this project, regional classification schemes will be invented to provide sufficient information to approximate the actual resistance term vital to their use.
2. Energy slope can be approximated from map slope and the existing `QVFAB.txt` database can be interrogated to recover information to relate these terms (recall the database does not contain slope, but does contain location information).
3. The notion of drainage system classification is supported by the literature, and this classification will be vital in recovering resistance terms for application in the index tool.

3 Database

3.1 USGS database

For the operational support of the streamflow gaging-station network in Texas, the USGS has collected and digitally archived about 140,000 discharge measurements (including zero-flow values) form more than 600 streamgages over the approximate time span of Dec. 1897–Feb. 2009. These discharge measurements, which are actually individual summaries of extensive field-collected data, reside within the USGS National Water Information System (NWIS) and are readily obtained U.S. Geological Survey (2009b) by streamgage number (a unique numerical identifier). The vast majority of the data represent discharges Q measured from current-meter-based (velocity-meter) techniques Turnipseed and Sauer (2010). For most of the discharge measurements concomitant hydraulic properties are also available, these are cross-section flow area A, water-surface top width B, reported mean velocity V, and other details. The basic relation between Q, A, and V is $Q = AV$. The basic relation between hydraulic (mean) depth D and A and B is $D = A/B$.

A preliminary screening of the complete Texas dataset of these 140,00 discharge measurements, produced a base observational dataset of 89,874 discharge records from 437 streamgages. Further details on the criteria used in the refinement are given below in section 3.4.1.

3.2 Gage station characteristics

One major objective of the research project was the development of a core database of watershed and climatic characteristics for the 437 U.S. Geological Survey streamflow-gaging stations. This database will be used in regional statistical analysis, such as regression, to relate discharge and mean velocity to selected or statistically important predictor variables.

This section outlines the additional gage station and watershed characteristics computed by USGS researchers, led by T.E. Burley, using a geographic information system with many custom algorithms of the USGS and authoritative datasets (Homer et al., 2004). These algorithms are reliable and were primarily developed as part of a comprehensive in number and scale, watershed processing campaign conducted for the Texas Commission on Environmental Quality. Additionally, T.E. Burley and his staff have completed FGDC[1]-compliant metadata documentation for their analysis are intermediate geographic information system files.

The gage station characteristics derived for each of the 437 station from geographic information system processing are outlined below:

[1]Federal Geographic Data Committee. The FGDC is tasked by Executive Orders to develop procedures and assist in the implementation of a distributed discovery mechanism for national digital geospatial data.

- `DELINID`: A "short" site/watershed ID for processing purposes;
- `SITENO`: The original NWIS 8 digit site ID;
- `STATIONNM`: The NWIS descriptive name for the site;
- `HUC8`: The eight digit hydrologic unit code (HUC);
- `WSAREA`: The area of the watershed in square kilometers;
- `WSPERIM`: The perimeter of the watershed in kilometers;
- `BASINCIRCULARITY`: Indicator of basin shape. The circularity ratio of the watershed to the area of a circle having the same area as the watershed, which serves as a dimensionless index to indicate the form of the basin compared to a circle of equivalent area (Miller, 1953)

  $$R_c = \frac{4\pi A}{P^2} \tag{21}$$

  where R_c is circulatory ratio, P is watershed perimeter in kilometers, and A is area of watershed in square kilometers;
- `DRAINAGEDENSITY`: Ratio of total length of streams in a watershed to the total area of the watershed (distance (kilometers) of channels per square kilometer) (Gordon et al., 2004);
- `STRAHLER`: The Strahler stream order of the outlet point associated with the watershed (Strahler, 1952);
- `SHREVE`: The Shreve stream order of the outlet point associated with the watershed (Shreve, 1967);
- `MCSLOPE`: Slope of the main channel from elevation of longest mapped channel to the elevation of the outlet (station);
- `MCSLOPE1KM`: Slope of the channel one kilometer up from the outlet (station);
- `SINUOSITYRATIO`: Approximate measurement of the degree to which a stream meanders within it's valley, computed by dividing channel length by valley length;
- `SLOPE1085`: The "10–85" slope of the channel (Gordon et al., 2004);
- `PRECIP`: The mean 30 year (1980–2009) precipitation for the land surface area covered by the watershed (units: millimeters), derived from PRISM data;
- `PCT11`: Percent open water;
- `PCT21`: Percent developed, open space;
- `PCT22`: Percent developed, low intensity;
- `PCT23`: Percent developed, medium intensity;

- `PCT24`: Percent developed, high intensity;
- `PCT31`: Percent barren land;
- `PCT41`: Percent deciduous forest;
- `PCT42`: Percent evergreen forest;
- `PCT43`: Percent mixed forest;
- `PCT52`: Percent shrub/scrub;
- `PCT71`: Percent grassland/herbaceous;
- `PCT81`: Percent pasture/hay;
- `PCT82`: Percent cultivated crops;
- `PCT90`: Percent woody wetlands; and
- `PCT95`: Percent emergent herbaceous wetland.

The PCT fields are all derived from the NLCD 2001 ver. 2 land cover data: `http://www.mrlc.gov/nlcd01_data.php` and `http://www.epa.gov/mrlc/classification.html`

Graphical visualization of selected combinations of gaging station characteristics are shown in Figures 9 through 18. These figures provide important information for guidance in what variables might be useful or redundant in statistical analysis of mean velocity and discharge from the USGS discharge measurement database. Most of the figures use watershed drainage area on the horizontal axis. This plotting axis is selected intentionally because of the great influence of drainage area or "watershed scale" for statistical analysis as well as the influence of watershed scale on other watershed characteristics[2].

[2]This influence can be substantial. Many measures of watershed length (main channel length, perimeter, centroid to outlet length, etc.) are so highly correlated to area as to be interchangeable in a statistical analysis.

For example, Figures 9 and 10 show substantial differences. In Figure 9, the main channel slope decreases with increasing drainage area and hence provides redundant statistical information. However, Figure 10 suggests independence between watershed area and the proximal channel slope.

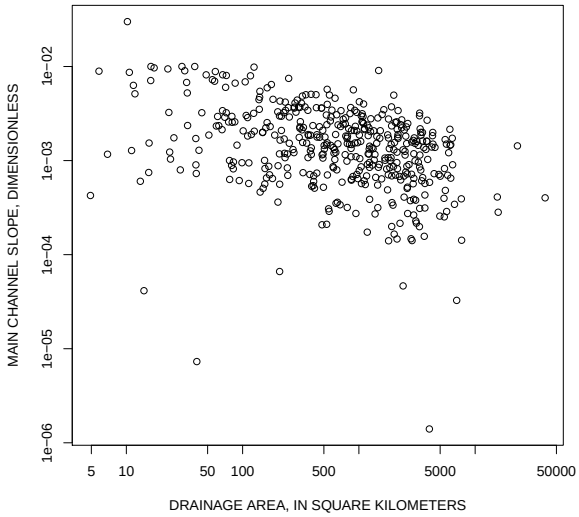

Figure 9: Relation between watershed drainage area and main channel slope

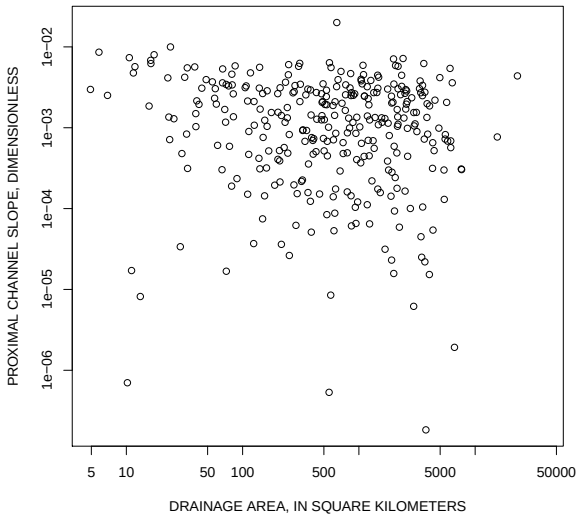

Figure 10: Relation between watershed drainage area and proximal channel slope

This independence is further reinforced by inspection of Figure 11, where knowledge of MCL provides little predictability of the proximal channel slope.

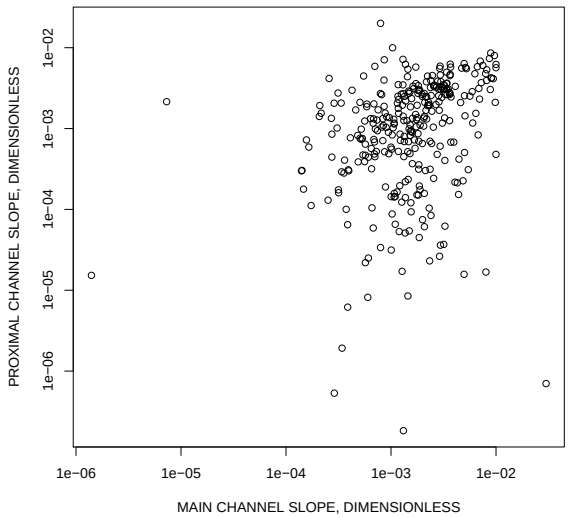

Figure 11: Relation between main channel slope and proximal channel slope

The relation between main channel slope and the 10-85 slope is shown in Figure 12. Figure 12 shows extreme correlation between the two slope definitions and shows that it makes little computational or statistical sense to expend the effort to determine a 10-85 slope.

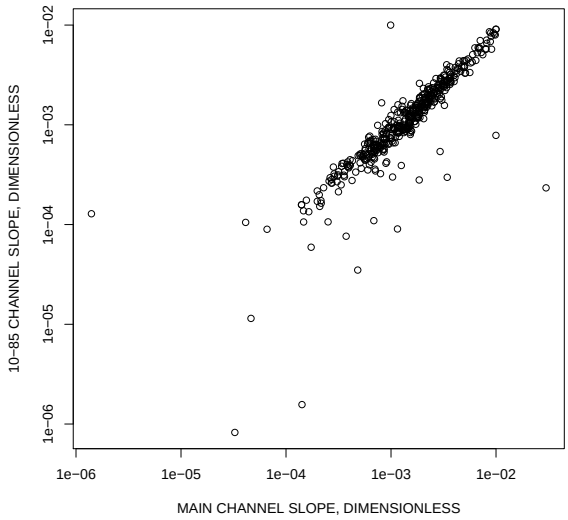

Figure 12: Relation between main channel slope and 10-85 channel slope

Figures 13 and 14 show considerable correlation between watershed circularity and sinuosity ratio with watershed drainage area. Hence, in a statistical analysis, there is little likelihood that either watershed circularity and sinuosity ratio would be a contributing[3] predictor variable of mean velocity or discharge.

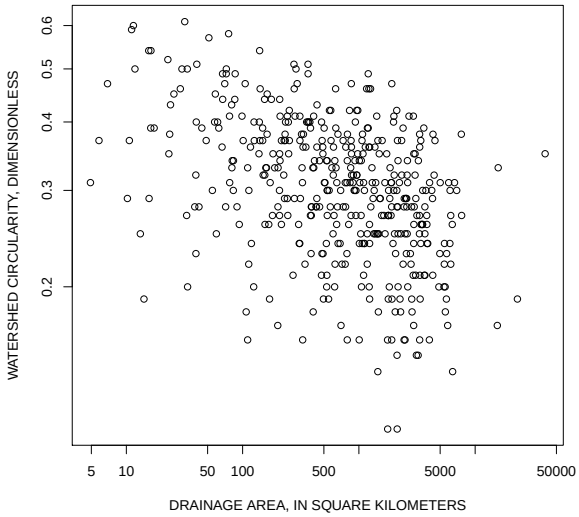

Figure 13: Relation between watershed drainage area and watershed circularity

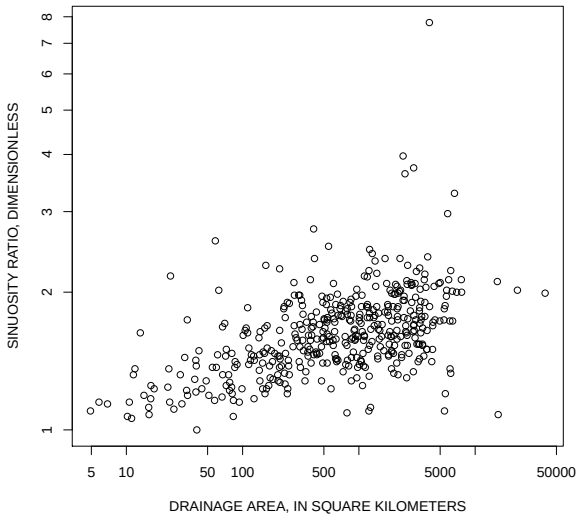

Figure 14: Relation between watershed drainage area and sinuosity ratio

[3]Adding substantive explanation of variance.

Figures 15 and 16 show considerable correlation between drainage density and watershed perimeter with watershed drainage area. Hence, it is unlikely in statistical analysis that either drainage density or watershed perimeter would be important predictor variables of mean velocity or discharge. This finding is particularly true of watershed perimeter.

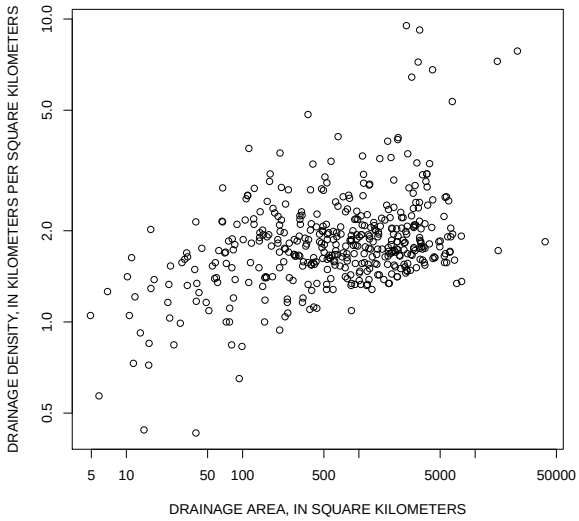

Figure 15: Relation between watershed drainage area and drainage density

Watershed perimeter is so highly correlated with watershed area (Fig. 16) that watershed area could be effectively computed from watershed perimeter. Stream order was computed

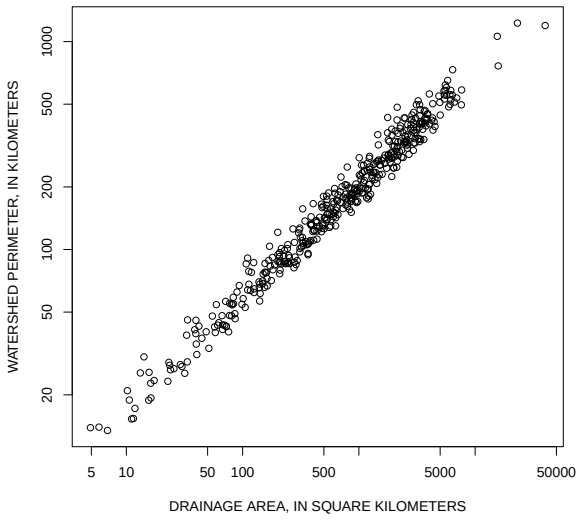

Figure 16: Relation between watershed drainage area and watershed perimeter

according to two definitions. The Shreve and Strahler definitions of stream order were both used. Figure 17 shows that the Shreve stream order also is highly correlated with watershed area; whereas, the Strahler stream order appears to be less so.

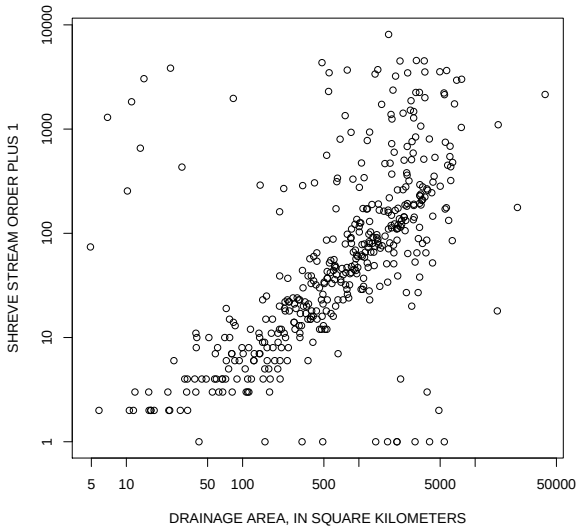

Figure 17: Relation between watershed drainage area and Shreve stream order

Unfortunately, from a data analysis perspective, the Shreve stream order is more attractive for regression analysis because of the great range or resolution of the numerical values.

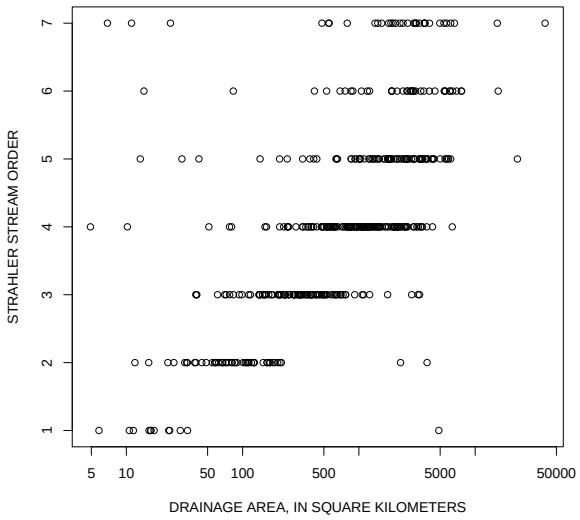

Figure 18: Relation between watershed drainage area and Strahler stream order

This resolution is in contrast with the apparently discrete numberical values of the Strahler stream order (Fig 18), which only acquires 7 separate values or simply "classifications" across an large range in watershed drainage areas.

As note, some gaging station characteristics are highly correlated and the engineer should be aware of such correlation when using various predictive tools. There appear to be some outliers in the figures — an unanticipated, but typical result. The characteristics shown are the result of objective but analyst-guided processing.

The characteristics referenced in this technical memorandum are those to be used for the remainder of the research project. These characteristics are available to the research team (and PMC) via the research server (`http://cleveland2.ce.ttu.edu`)

3.3 GIS database

The gaging station characteristics of the previous section were added to the original USGS data based (section 3.1). The expanded databased was then further built upon by providing estimates of channel slope obtained by an alternate method to those described in section 3.2, and provides physiographic data for each station. This physiographic data was assigned to each station as a possible additional descriptor for predicting the mean velocity , V, and possibly aiding in the development of regional regression equations, or a statewide regression equation containing a variable to account for differences in physiographic region, thereby narrowing the range of probable V values at any particular site.

Physiography may also prove helpful in predicting the effective channel roughness, which influences energy loss and therefore V. Mannings n is often used as a measure of roughness of piping and concrete channels. Mannings n is also assigned to natural channel beds and side slopes. However, in actual streams, energy losses may be due also to a variety of other factors as well, and we consider the effective channel roughness, hereby denoted ne , to take such factors into account. Major factors contributing to this effective channel roughness, identified earlier by Cowan (1956) and later by others in greater detail (e.g., Chen et al., 2011), include sediment size, degree of surface irregularity, variation of channel cross-section, effect of obstructions, and vegetation.

Seven physiographic provinces, three of which are further divided into subdivisions, have been identified by the Texas Bureau of Economic Geology (TBEG) at the University of Texas at Austin (1996), as shown by the colored regions of Figure 19, and as listed in Table 1. The United States Geological Survey (USGS) has also developed a physiographic map of the United States including Texas (tapestry of time and terrain, `http://tapestry.usgs.gov`), based on the work of Fenneman and Johnson (1946). In general the Bureau of Economic Geologys divisions are more detailed and will be used to facilitate the present discussion.

Figure 20 displays the lithology identified by the USGS within each physiographic region (`http://mrdata.usgs.gov/geology/state/state.php?state=TX`). Lithology is worthy

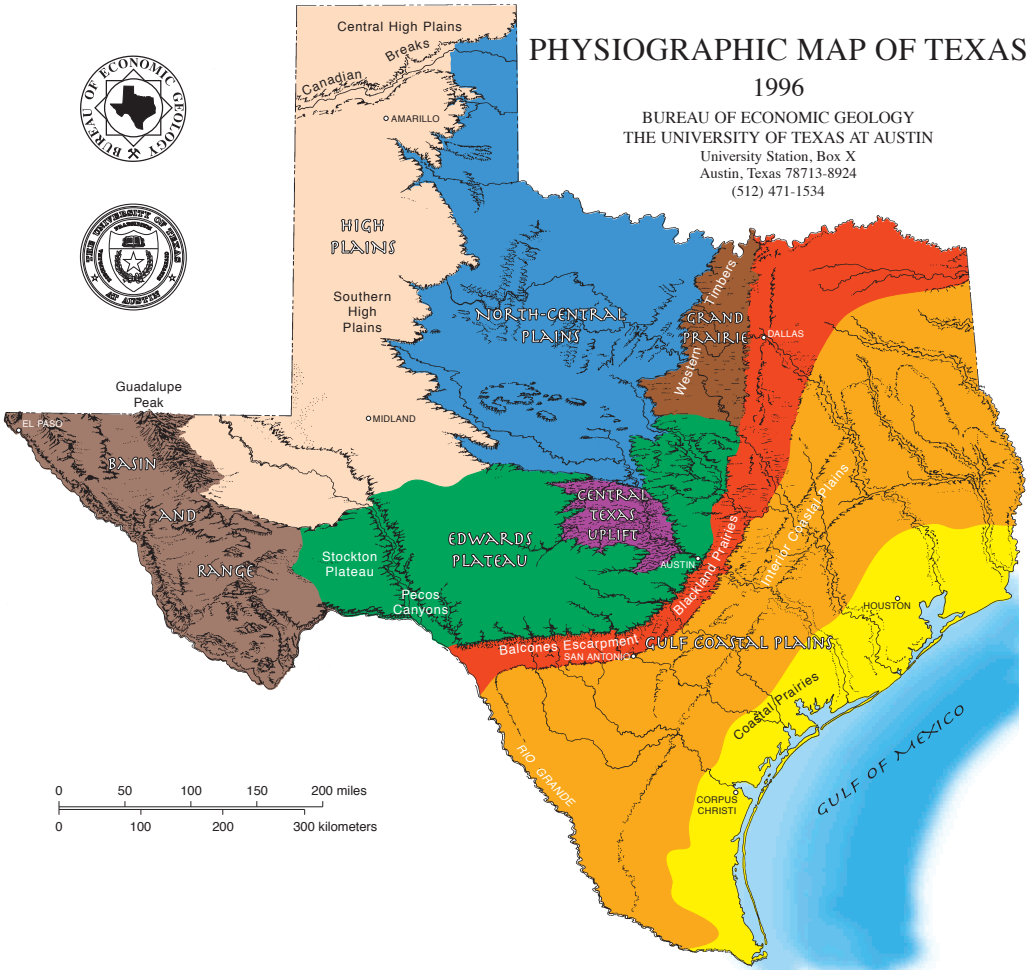

Figure 19: Physiographic map of Texas produced by the Bureau Of Economic Geology.

PROVINCE	MAX. ELEV. (ft)	MIN. ELEV. (ft)	TOPOGRAPHY	GEOLOGIC STRUCTURE	BEDROCK TYPES
Gulf Coastal Plains					
Coastal Prairies	300	0	Nearly flat prairie, 1-4 mi to Gulf	Nearly flat strata	Deltaic sands and muds
Interior Coastal Plains	800	300	Parallel ridges (questas) and valleys	Beds tilted toward Gulf	Unconsolidated sands and muds
Blackland Prairies	1000	450	Low rolling terrain	Beds tilted south and east	Chalks and marls
Grand Prairie	1250	450	Low stairstep hills west, plains east	Strata dip east	Calcareous east, sandy west
Edwards Plateau					
Principal	3000	2350	Flat upper surface with box canyons	Beds dip south; normal faulted	Limestones and dolomites
Pecos Canyons	2000	1200	Steep-walled canyons		Limestones and dolomites
Stockton Plateau	4200	1700	Mesa-dotted terrain rises to west	Unfaulted, near-horizontal beds	Carbonates and alluvial sediments
Central Texas Uplift	2000	800	Knobby plain; surrounded by questas	Centripetal dips, strongly faulted	Granites; metamorphics; sediments
North-Central Plains	3000	900	Low north-south ridges (questas)	West dip; minor faults	Limestones; sandstones; shales
High Plains					
Central	4750	2900	Flat prairies slope east and south	Slight dips east and south	Eolian silts and fine sands
Canadian Breaks	3800	2350	Highly dissected; local solution valleys		
Southern	3800	2200	Flat; many playas; local dune fields		
Basin and Range	8750	1700	North-south mountains and basins	Some complex folding and faulting	Igneous; metamorphics; sediments

Table 1: Discriptions of physiographic provinces and subdivisions in Texas (Texas Bureau of Economic Geology, 1996)

of special attention here because the other factors that help define physiography primarily vegetation, climate, and topography may be covered to a substantial extent by the database developed under Tasks 2 and 3. Finkenbinder (2008) explored the possibility of using varying lithology within physiographic regions of Pennsylvania as opposed to using physiographic region alone to predict channel geometry, but with inconclusive results due to small sample size Dividing the state into lithologic rather than physiographic regions for development of regression equations is not immediately anticipated. The display of the lithographic regions is intended to facilitate discussion of channel geometry and channel properties found in the literature.

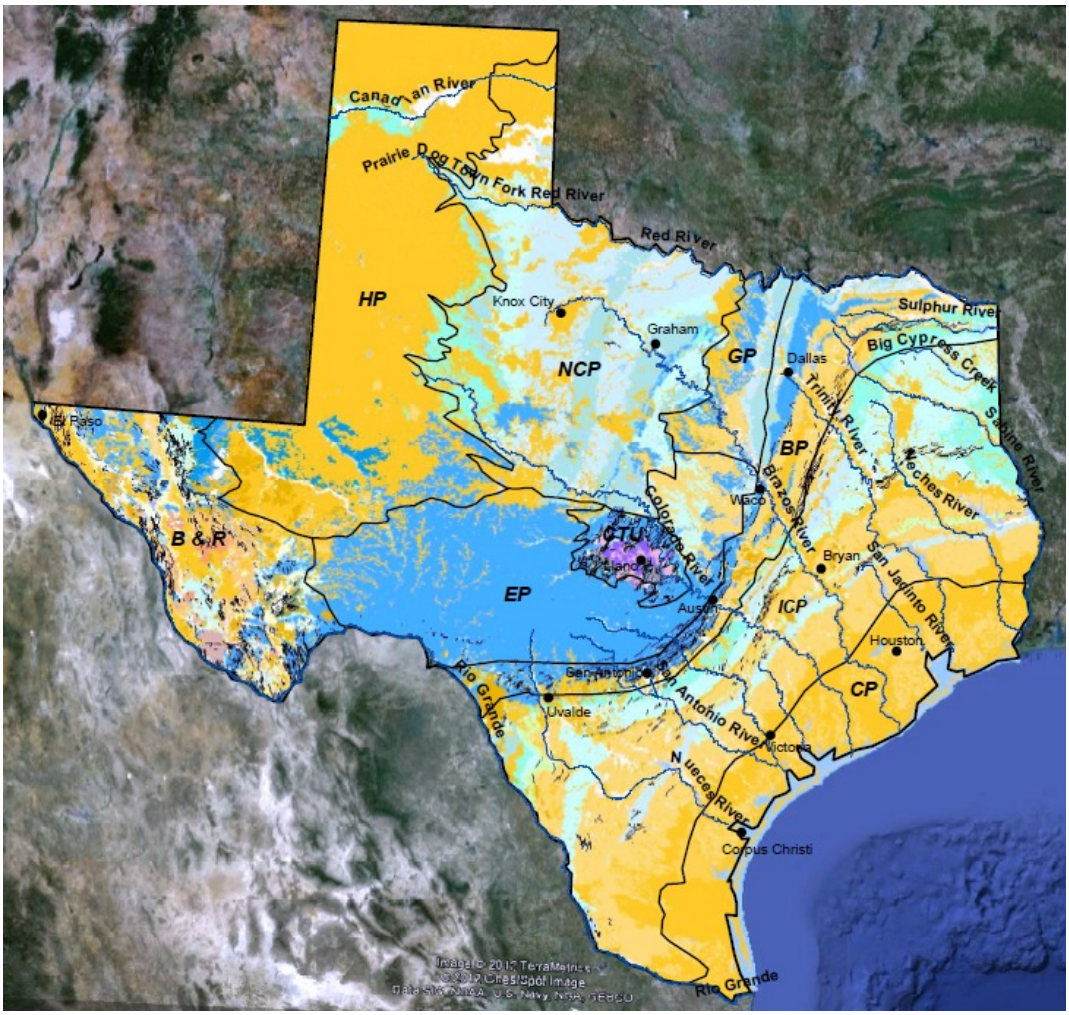

Figure 20: Physiographic provinces (outlined in black) and lithology for Texas. Adapted from the Texas Bureau of Economic Development (1996) and Google Earth. Physiographic names are abbreviated due to space limitation: Basin and Range (BR), Blackland Prairies (BP), Central Texas Uplift (CTU), Coastal Prairies (CP), Edwards Plateau (EP), Grand Prairie (GP), High Plains (HP), Interior Coastal Plains (ICP), and North-Central Plains. (A legend for the image can be found at: `http://mrdata.usgs.gov/catalog/lithclass-color.php`, in the figures 21 and 22.)

Figure 21: Part 1 of the the legend for Figure 20.

Figure 22: Part 2 of the the legend for Figure 20.

In recent years advances in remote sensing technology and algorithms have become useful in identifying vegetation in the riparian zone, cross-sectional channel geometry, and in helping to assess effective channel roughness. For example,Hutton and Brazier (2012) present an algorithm for better separating the elevation of the ground itself from the elevation of top of vegetation in airborne Light Detection And Ranging (LiDAR) data. Bertoldi et al. (2011) use LiDAR data and aerial photographs to determine riparian vegetation extent, height, and structure, and to map elevations of submerged channels. Hostache et al. (2010) utilize Synthetic Aperture Radar (SAR) images to measure flood water levels within an average vertical accuracy of 40 cm, and to help identify areas in the floodplain and in the channel where the Manning roughness coefficients are homogeneous. Utilization of such advances in research are not practically implemented at present within TxDOT, but are briefly here mentioned to provide an awareness of future possibilities. The present focus is to consider physiographic region to help predict V.

TBEG does not provide shapefiles of its physiographic regions map. To create the shapefile which is displayed in Figures 20 and 5, first TBEGs map in pdf format was converted into jpeg format, and then imported into ArcGIS. Prominent points of the Texas political boundary and shoreline were then used as control points to georeference the image to a Texas state shapefile. Physiographic boundary lines of the georeferenced jpeg file were then traced manually to create the shapefile.

The literature review suggests that lithologic heterogeneity within several physiographic regions may contribute substantially to variations in channel geometry and effective roughness, thereby impacting V. Meanwhile, the Coastal Prairies and Interior Coastal Plains physiographic regions cover a substantial portion of the USGS stations included in this project, and appear relatively homogeneous. Regions for curves may thus be worth identifying not only by differences in V (after adjusting for other factors) but also by the width of the confidence interval.

The estimation of channel bed slopes was achieved by the following means:

1. A shapefile of the 21 major Texas river basins (cite) was superimposed on 30m resolution digital elevation model (DEM) raster files (cite). For each basin, raster files visually identified as falling within the basin were combined into a mosaic. However, as it was later learned that DEMs of 30 m resolution caused the computer to crash in basin autodelineation, the cells were aggregated to a 90 m resolution.

2. For each major river basin, ArcSWAT (which was used as a plug-in to ArcGIS) was used for autodelineation. One of the user-defined parameters in the ArcSWAT autodelineation process is the threshold subbasin area. By trial-and-area, it was discovered that a value of 1,000 hectares for this parameter would lead to a stream channel appearing at all but 6 of the 437 gaging stations, and that substantially lower values would still not capture these few sites. The threshold of 1,000 hectares was set for all autodelineations to capture 431 stations.

3. For each of the 431 station sites, a distance of 500 m was traced upstream along the

SWAT-generated stream and the elevation of the DEM raster cell at that distance was copied into an Excel spreadsheet. The same was repeated for 500 m downstream. The upstream elevation minus the downstream elevation, divided by 1,000 m, was taken as the slope.

The above process generated some estimates of channel bed slopes that are negative, suggesting that the water would be flowing uphill. A negative channel bed slope, though unlikely, is possible. Relatively dead water would rest in the depressed zone, but the water surface itself would no doubt decline in the direction of streamflow. Phillips et al. (2005) identified the channel bed of the Trinity River to be below sea level in a large portion of its length in the Coastal Prairies. Yet, many of the negative slopes estimates cannot be thought of as representing actual channel bed slopes. Too many of them, approximately 14%, were found to be negative, and most of these are not near the coast line. We believe that errors in estimation are due to the fact that some DEM cells lying along the SWAT-delineated streams are average elevations over a 90m by 90m area. In some cases that average may be closer to the stream bed channel elevation, while in other cases that average may be closer to that of the river bank or other topography. A negative slope estimate could easily arise when the upstream elevation estimate is more representative of the channel bed, but the downstream elevation estimate is more representative of higher topography in the area.

With some care, the distribution of negative slope estimates may be helpful in modeling the distribution of errors for the entire set of slope estimates. This error distribution might prove helpful in accounting for a portion of error variances in the regional curves.

In an effort to reduce the errors in slope estimation, R script was written to extract the SWAT-estimated subbasin channel slope for each ArcSWAT-delineated subbasin containing a gaging site. All such estimates are positive. The disadvantage of this methodology is that the slopes are measured over varying lengths. Nearly all are under 10 km, but 20 are in the 10 km to 30 km range. Varying the subbasin threshold size such that all subbasin main channels are of the same length is not possible.

3.4 The final database

In order to proceed with the statistical analyses that will result in a class of interpretive end products for the research project, it is necessary to assemble a database for which the records are believed to be accurate or otherwise reliable. Whereas, there is a natural expectation that odd values, outliers, and erroneous data still exist, considerable effort has been made through iterative statistical processing. The final resulting database for the analysis therefore consists of the merging of a refined USGS discharge measurement database with the database containing watershed properties as described in section 3.2.

3.4.1 Refinement

The database contains 89,874 discharge records for the approximate period Dec. 1897–Feb. 2009 for 437 selected Texas streamgages. The 437 streamgages were selected as a prerequisite for this paper based on preliminary screening of more than 600 streamgages and select preprocessing that included factors such as consideration of streamflow data type, record length, number of discharge measurements, and regional setting or location of the streamgage. In general, a candidate streamgage needed to be a continuous-record type and represent streamgages that are considered examples of conventional (traditional) USGS streamgaging operation and not special projects (perhaps streamgages operated with theoretical weir stage-discharge relations), partial duration streamgages (perhaps flood hydrograph, or conversely, low-flow streamgages), or peak-only streamgages.

This unified discharge measurement database provides the foundational basis for the analysis reported here and contains the following attributes: discharge, reported mean velocity, cross-section flow area, water-surface top width, Froude number, and estimated flow-duration probability of the discharge. Unlike the approach by Castro and Jackson (2001) in their study of regional bankfull relations, no site visits to any of the 437 Texas streamgages were made for this study. The unified discharge measurement database was assembled through the following steps:

1. Daily Mean Streamflow Values—For the large and reasonably comprehensive list (437) of continuous-record (daily mean values of streamflow) streamgages in Texas, the daily mean streamflow values were retrieved from U.S. Geological Survey (2009a);

2. Streamflow Measurements—For the 437 streamgages, the discharge measurement file for each streamgage was retrieved from U.S. Geological Survey (2009b);

3. Complete Records—The measured discharge Q in cubic meters per second (m^3/s), "channel velocity" (referred to herein as "reported mean velocity") V in meters per second (m/s), "channel area" (referred to herein as "cross-section flow area") A in square meters (m^2/s), and "channel width" (referred to herein as "water-surface top width" or just "top width") B in meters (m) were extracted and only those records with $Q > 0$ were retained;

4. Computed Mean Velocity—Computed mean velocity $\overline{V}$ in m/s was computed by $\overline{V} = Q/A$. The adjective "computed" (as opposed to "reported" mean velocity in Step 3) in this paper refers to Q divided by A irrespective of the source of Q or A;

5. Velocity Consistency—The computed $\overline{V}$ was compared to the reported V, and if the absolute difference was greater than 0.03 m/s (chosen by the authors), then the record (a single discharge measurement) was rejected for inclusion in the unified discharge measurement database and thus not retained for the analysis reported here;

6. Froude Number—The Froude number was computed by $\mathrm{Fr} = V(gA/B)^{-1/2}$ where g is acceleration of gravity. For this paper, Fr is not used but is retained in the unified

discharge measurement database;

7. Flow-Duration Curve—The entire period of record of daily mean streamflow for each streamgage referenced in Step 1 was converted to a streamgage-specific, flow-duration curve Vogel and Fennessey (1994); and

8. Individual Discharge Probabilities—The probability of each Q was determined from the respective streamgage-specific, flow-duration curve of daily mean streamflow values using linear interpolation as necessary.

Further discussion of selected details of the 8 steps is needed to provide additional context for various decisions or observations that are important to communicate:

- On Greater than Zero Discharge—Step 3 excludes reverse flow ($Q < 0$) in tidal and zero-flow conditions ($Q = 0$);

- On Incomplete Attributes—To clarify, any discharge measurements (direct or indirect) lacking any core attributes (Q, V, A, and B) or in violation of Step 5 were not retained for the unified discharge measurement database;

- On Streamflow Probability—Step 7 states that the *entire* record of each streamgage was used to compute each streamgage-specific, flow-duration curve. This explicitly means that no attempt was made to define periods of stationary (unchanging statistical properities) streamflow or more importantly statistics of hydraulic relations. For example, no differentiation between pre- and post-reservoir conditions (if applicable) for a given streamgage was made. Such streamgage-specific investigation is beyond the scope of this paper;

- On Streamgage Location—USGS streamgages are only very rarely located in settings in which backwater conditions occur because a unique stage-discharge relation is desired. Also, a given streamgage is not anticipated to permanently exist at the exact same location along a stream during the course of the streamgage's operational time frame; however, many streamgages remain more-or-less sited at their original locations. Streamgage locations are referenced to the nearest town or locality with a postal code, for example, USGS streamgage 08167000 Guadalupe River near Comfort, Texas. Streamgages are periodically relocated to nearby locations, but adjustments to identity (number and name) are not made, because of channel migration; channel rectification/restoration; bridge maintenance, decommission, and new construction; property access (landowner changes); and changes in safety policy and practices. Changes in bridge characteristics are likely the most common cause of relocation because many streamgages in Texas often are located along Texas Department of Transportation right-of-way;

- On Measurement Location—A fact, which likely hampers many streamgage-specific investigations of geomorphic processes using USGS measurement databases, is that the precise cross-section location of an individual discharge measurement is neither reported or fully documented in USGS discharge measurement summaries used herein.

Furthermore, the measurement location is not expected to coincide with the same location either over the years or over a range of discharge conditions. There are many discipline- and technically-specific reasons discharge measurements might be not made at precisely the same geographic stream location because of discharge magnitude and year-over-year streamgage operation;

- On Bankfull Conditions and Floodplain Engagement—The discharge measurements (summaries) available from U.S. Geological Survey (2009b) do not provide consistent and, even when available, only limited details identifying whether the measurement summary is applicable for a partially to full channel or whether the floodplain (if it exists in a classical sense) is engaged by the water surface near the measurement location. Because of generally more favorable conditions for measurement, discharge measurements are often performed, whenever possible, in places with flow conditions lacking substantial floodplain inundation. Also, many streamgages are located near bridges because of the more favorable conditions for truck-mounted-crane, high-magnitude discharge measurement;

A discussion is needed that concerns components of the well known Manning's equation for computation of simplified open-channel hydraulics in the context of USGS discharge measurement databases. Manning's equation is:

$$Q = \frac{c_o}{n} A (A/WP)^{2/3} S^{1/2} \tag{22}$$

where c_o is 1 for SI units and 1.486 for US customary units, n is the Manning n-value, WP is the wetted perimeter, and S is the friction slope. The Manning equation provides a useful mathematical structure to statistically evaluate Q and V through intrinsic relations between A, B, WP, and S. However, several limitations excluded application of Manning's equation in a statistical context for this paper:

- Friction slope—Friction slope is indisputably an important parameter because Q and V are proportional to the square root of slope. However, the friction slope is not available from U.S. Geological Survey (2009b). Channel slope often is used in place of friction slope in Manning's equation; channel slope also is not available from U.S. Geological Survey (2009b).

- Manning's n-value—The Manning n-value also is indisputably an important parameter in Equation 22. Unfortunately, n-values, which are not direct measures of roughness, or other roughness parameters, such as median grain sizes, influencing channel hydraulics are not readily available for any of the streamgages in general or for individual discharge measurements across time in particular; and

- Wetted perimeter—The wetted perimeter, which is used to compute the hydraulic radius (the $R = A/WP$ term in Manning's equation), likely is useful as a direct predictor variable on Q or V or is useful as a predictor variable when expressed as hydraulic radius. The field-measured data for direct measurements of discharge by

the USGS contain horizontal stationing and vertical sounding (depth) information. From these raw data, *WP* for individual measurements could be estimated. Unfortunately at the present time (2012), the USGS discharge measurement database U.S. Geological Survey (2009b), being summaries of the field observations, lack either *WP* values or the raw data to compute them. Hence, *WP* values are not available for this study.

The unified discharge measurement database of 87,874 records for 437 streamgages in Texas was subsequently filtered or reduced to contain discharge measurements that could be reasonably associated with direct-runoff conditions. Specifically, discharge measurements exceeding the 90th-percentile daily mean streamflow as determined by the streamgage-specific, flow-duration curves were retained for the analysis reported herein. This 90th-percentile discharge measurement database, is the database used for the statistical analysis. The 90th-percentile database contains 17,753 discharge records for 424 of the original 437 streamgages. Each of the 424 streamgages has at least one measurement greater than the 90th-percentile daily mean streamflow for that streamgage.

Summary statistics of A, Q, V, Fr, and B of the 90th-percentile discharge measurement database are listed in Table 2. After filtering for high-magnitude discharge considerable variation or range remains in A (about 6 orders of magnitude), Q (about 7 orders of magnitude), V (about 2 orders of magnitude), Fr (about 2 orders of magnitude), and B (about 5 orders of magnitude). These tabulated statistics of their respective distributions could be used for additional data screening and record rejection prior to regionalization. For example, the maximum $B = 14{,}000$ m is almost certainly too large, the minimum Fr $= 0.00610$ is almost certainly too small, and the maximum Fr > 1 (indicative of supercritical flow conditions) is seemingly high for natural channel flow. Additional data screening and record rejection was not made prior to statistical analysis except for the removal of a few extreme outliers.

Statistic	Cross section area, A (m^2)	Discharge, Q (m^3/s)	Mean velocity, V (m/s)	Froude number, Fr (--)	Water-surface top width, B (m)
Minimum	0.00372	0.000283	0.0152	0.00610	0.0671
1st quartile	10.9	5.94	.436	.125	15.9
Median	47.6	30.6	.658	.187	33.8
Mean	167	154	.750	.231	78.2
3rd quartile	165	125	.951	.296	79.6
Maximum	6,940	7,610	4.22	2.37	14,000

Table 2: Summary statistics of selected hydraulic parameters from 90th-percentile discharge measurement database in Texas

3.4.2 Merging

The refined database was merged with a small subset of the watershed properties identified in section 3.2 to produce the final databased used in the statistical analysis. Only a few, and not all, parameters from the gaging station characterization were added to the refined database because few characteristics were found to have power in improving predictions of Q or V. The gaging station characteristics added to the refined database include: the contributing drainage area, the main channel slope, the proximal channel slope, the 10-85 main channel slope, and the regionalization parameter OmegaEM from Asquith and Roussel (2009); OmegaEM is discussed in more detail in the following section. Overall the final database contains about 87,874 records with 17,753 records for which the flow exceeded the 90th-percentile discharge.

database_txdot0-6654.txt

```
STATION|LATDEG|LONDEG|CDA|MCS|PCS|MCS1085|MAP|OMEGAEM|Q|A|V|B|FDC
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|373|117|3.19|104|0.8844
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|961|300|3.2|285|0.952
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|1240|370|3.35|285|0.9619
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|3240|613|5.29|259|0.9852
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|2790|564|4.95|244|0.983
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|15200|2020|7.52|365|0.9986
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|1020|333|3.06|231|0.9553
7227500|35.47028|101.87917|584|0.000993058|-0.00130157471|0.01|48636.71|-0.071|3760|628|5.99|292|0.9871
```

Figure 23: Screen capture of first few rows of the refined database

Figure 23 is a screen capture of the first few rows of the database. The data are arranged in columns using the pipe symbol " | " as the delimiter. The column headings in the figure correspond to the following descriptions:

1. `STATION` is 8-digit USGS station identification code. These codes can be entered into the NWIS web interface to recover textural description of the particular station and other locational information.
2. `LATDEG` is the latitude in degrees and decimal degrees (dd.ddddd).
3. `LONDEG` is the longitude in degrees and decimal degrees (ddd.ddddd).
4. `CDA` is the contributing drainage area to the station in square miles (mi^2).
5. `MCS` is the main channel slope (Asquith and Slade, 1997).
6. `PCS` is the proximal channel slope.
7. `MCS1085` is the 10-85 main channel slope (Gordon et al., 2004).
8. `OMEGAEM` is the OmegaEM parameter (Asquith and Roussel, 2009).
9. `Q` is the observed discharge in cubic feet per second (ft^3/s).
10. `A` is the cross sectional flow area (at the above observed discharge) in square feet

(ft^2).

11. `V` is the mean section velocity (ratio of discharge to flow area) in feet per second (ft/s).

12. `B` is the topwidth (at the observed discharge) in feet (ft).

13. `FDC` is the flow duration curve for the associated station. There are hundreds of stations represented in the database, and a flow duration curve was computed for each station. The exceedance probabilities of flow for that station were maintained to provide a useful conditioning capability.

4 Statistical Approach

4.1 Generalized Additive Models

Complex relations between both Q and V and available predictor variables (described in Section 4.2) were anticipated. Therefore, in lieu of conventional multi-linear regression modeling Faraway (2005), generalized additive modeling (GAM) Hastie and Tibshirani (1990); Wood (2006) was chosen. A GAM is a statistical model between a response variable and an additive combination of various parametric terms and smooth terms (functions). The incorporation of smooth functions can be an advantage to GAMs over simpler multi-linear regression because appropriately configured smooth functions accommodate otherwise difficult to "linearly model" components of a prediction-response model. A Gaussian family for the generalized linear model Faraway (2006) was used to estimate the GAM models reported here using mostly default arguments of the `gam` function in the R environment R Development Core Team (2011) from the `mgcv` package by Wood (2009). The model fitting is based on maximum likelihood (not conventional least-squares) for parameter fitting (optimization). The basic form of a GAM model:

$$y_i = \mathbf{X}_i\Theta + f_1(x_{1i}, x_{2i}) + f_2(x_{3i}) + \ldots + \epsilon_i \tag{23}$$

where y_i is a suitably transformed response variable for the ith observation, $\mathbf{X}_i$ is a model matrix for strictly parametric and suitably transformed predictor variables, Θ is a parameter matrix, the f_k are "smooth functions" of the predictor variables x_{ik}, and ϵ_i are error terms taken as independently and identically distributed $N(0, \sigma^2)$ (Gaussian distribution or normal distribution) random variables. The $\mathbf{X}_i\Theta$ term is the familiar multi-linear regression component of a GAM.

In this analysis, separate GAM analyses of Q and V were conducted. The GAM model of Q is referred to as QGAM, and similarly, the GAM model of V is referred to as VGAM. As further described and justified in Section 4.2, the basic form of the QGAM is:

$$\begin{aligned} \log(Q) &= b_1 + a_1\log(A) + a_2\log(B) + a_3\Omega \\ &\quad + f_5(\text{longitude}, \text{latitude}) + f_6(P) \end{aligned} \tag{24}$$

and the basic form of the VGAM is:

$$\begin{aligned} V^{1/5} &= b_2 + a_4\log(Q) + a_5\log(B) + a_6\Omega \\ &\quad + f_9(\text{longitude}, \text{latitude}) + f_{10}(P) \end{aligned} \tag{25}$$

where log is base-10 logarithm, Q is discharge in m^3/s, V is mean velocity in m/s, b_k are intercepts, a_k are regression coefficients, A is cross-section flow area in m^2/s, B is

top width in m, Ω is the OmegaEM parameter from Asquith and Roussel (2009) and is described in Section 4.2, f_k are smooth functions in one or two dimensions as indicated and the numerical value of the subscript references the applicable figure of this report, and P is mean annual precipitation in millimeters (mm) and is described in Section 4.2. The QGAM and VGAM are respectively presented in Sections 5.1 and 5.2. Lastly, the predictive potential of watershed drainage area was found to be unsuitable as a predictor variable for the Q and V regionalization of the 90th-percentile discharge measurement database. Select predictor variables are discussed in the next section along with choice of variable transformation.

4.2 Preprocessing and Preliminary Analysis

4.2.1 OmegaEM Parameter

Asquith and Roussel (2009) developed regional equations to estimate annual peak-streamflow frequency for undeveloped watersheds in Texas. As part of their analysis, they created a generalized residual of the 10-year (0.10 AEP) discharge equation that is referred to as the OmegaEM parameter. This parameter represents a generalized terrain and climate index that expresses peak-streamflow potential not otherwise represented in the watershed characteristics of drainage area, main-channel slope, and P. The OmegaEM parameter is gridded by 1-degree quadrangles (Asquith and Roussel, 2009, p. 14) and is reproduced and shown in Figure 24. Although developed from analysis of undeveloped watersheds, the parameter captures generalized terrain and climate influences on channel conveyance properties affecting discharge magnitude.

The authors hypothesize that OmegaEM should be a useful, but minor, predictor of Q and V because OmegaEM expresses regional variation in otherwise difficult to quantify variations in high-magnitude discharge. Using the latitude and longitude of each of the 424 streamgages, the OmegaEM parameter was computed for each streamgage by bilinear interpolation from the gridded values in Figure 24.

4.2.2 Mean Annual Precipitation

Climatological conditions in Texas are diverse. Bomar (1994) provides a review of Texas weather and climate and details historically important rainfall and resulting floods, the characteristics of the atmosphere, and general weather statistics for Texas. For the 424 streamgages, P ranges from about 292 mm for a streamgage in the extreme western part of Texas to 1,571 mm for a streamgage in the extreme southeastern part of Texas.

Using the latitude and longitude of each of the 424 streamgages, mean annual precipitation P in mm was retrieved for each streamgage from PRISM Climate Group (2010) for the 1971–2000 normals. The PRISM Climate Group (2010) source was chosen for expediency. Given the many sources of uncertainty both in GAM development and implementation

ian might exist. General erosional and attendant geomorphologic settings as represented by stream channel shapes are also affected by P. Channel shape in turn influences relations between discharge and mean velocity through the hydraulic characteristics of cross-section flow area and top width.

The authors also considered other climate normals available from PRISM Climate Group (2010) including mean July high and mean January low temperatures and their difference. These climate indices seem to be no better predictors or contributors to the explanation Q or V variance than P.

4.2.3 Variable Transformation

The authors hypothesize for the objective of Q regionalization that the hydraulic parameters of A and B should be critically important parameters. A preliminary issue at hand is the choice of transformation in the GAM analysis. Analysis through multi-linear regression, Box-Cox power transformations Box and Cox (1964) (the `boxcox` function in R from the `MASS` package by Venables and Ripley (2002)), and preliminary GAM analysis showed that logarithmic transformation on Q, A, and B was appropriate.

The authors also hypothesize for the objective of V regionalization that the hydraulic parameters of Q and B should be critically important parameters. The use of A is not appropriate or even possible in the context here because the reported V values are effectively, if not exactly, the ratio of Q to A. A preliminary issue at hand is the choice of transformation in the GAM analysis. Analysis through multi-linear regression, Box-Cox power transformations, and preliminary GAM analysis showed that fifth-root transformation on reported V (or $V^{1/5}$) and logarithmic transformation on Q and B was appropriate.

Preliminary QGAM and VGAM were fit following the algebraic structure of Equations 24 and 25 and were used to identify a few extreme outliers. The minimum of the absolute value of the range of the residuals was separately computed for the preliminary QGAM and VGAM. The Q and V records having residuals in absolute value greater than the respective minimums subsequently were removed; summary of these removals (very few) is made in Sections 5.1 and 5.2. The effect of outlier removal was to enhance the centering of the residuals in the final QGAM and VGAM models.

4.2.4 Main Channel and Proximal Slope

An open question with this work was whether main-channel slope (MS) or proximal slope (PS) were statistically important variables for regionalization of discharge measurements. That is, do either of these slopes enhance the prediction of mean velocity or discharge values from USGS discharge measurements associated with direct-runoff conditions in Texas (i.e., discharge measurements exceeding the 90th-percentile daily mean streamflow at a given USGS streamgage)? Based on the presence of S in the Manning equation (eq. 22), one

might expect slope to be an important explanatory variable for Q and V. One might further expect that the proximal slope would be more powerful than the main channel slope since the Manning equation is a more localized reach-averaged hydraulic equation. These two expectations were tested as part of the project and a self-contained detailed report of the analysis can be found in the appendix (Appendix A). The overall conclusion of the analysis was that neither the proximal or main channel slope where very useful in adding explanatory power to the regression models, and that the proximal channel slope was no more powerful as an explanatory variable than the main channel slope.

In the analysis, and in the report as a whole, the main channel slope, MS is defined as the change in elevation in feet (ΔE) between the two end points of L (main-channel length) divided by L in feet. Hence, MS is a "watershed slope" based on the principal channel network as represented in L. L is defined as the length in stream-course feet of the longest defined channel represented by the particular data used by an analyst from the approximate watershed headwaters to the outlet (streamgage). The proximal slope, PS, on the other hand, is defined as the change in elevation in feet (ΔE) between points located at the watershed outlet (streamgage) and about 0.62-miles (1 km) upstream from the streamgage. These two slopes are defined in the gaging station characteristic section (section 3.2) and the database as `MCSLOPE` and `MCSLOPE1KM`.

5 Products: Method 1 - Regression Analysis

5.1 Generalized Additive Model of Discharge

The final QGAM in R output is shown in Figure 25. For the QGAM, each of the predictor variables is statistically significant. The adjusted R-squared value is about 0.95, and the residual standard error is about $s = 0.22$ base-10 logarithm of m^3/s, which is the square root (Wood, 2006, p. 61) of the "Scale est." because a Gaussian family was used for this GAM. For the final QGAM model, 26 discharge measurements for 13 streamgages (USGS station numbers: 07295500, 08018730, 08047500, 08080700, 08110325, 08129300, 08166000, 08185000, 08186500, 08190500, 08197500, 08202700, and 08210400) were removed but the overall streamgage count remained at 424. The QGAM with the coefficients shown in Figure 25 can be written as:

$$\begin{aligned} \log(Q) &= -0.2896 + 1.269\log(A) - 0.2247\log(B) + 0.2865\Omega \\ &\quad + f_5(\text{longitude}, \text{latitude}) + f_6(P) \end{aligned} \tag{26}$$

where log is base-10 logarithm, Q is discharge in m^3/s, A is cross-section flow area in m^2, B is top width in m, Ω is the OmegaEM parameter from Figure 24, P is mean annual precipitation in mm, and f_5 and f_6 are "smooth functions" of the indicated predictor variables in Figures 26 and 27, respectively. For Figure 26, the base map and superimposed smooth lines were created in R using graphic capabilities of packages by Minka (2011) and Wood (2009), and Figure 27 was created using graphic features by Wood (2009). The red, green, and black lines as ensembles of three for each numerical value shown in Figure 26 are not all shown for reasons such as grid resolution for the graphic, nonuniform distribution of streamgages, and general statistical magnitude of the two-dimensional smooth surface.

The `k=14` argument (shown in Figure 25) to the $f_5(\text{longitude}, \text{latitude})$ or $f_5(l, k)$ smooth function of location represents the dimension of the isotropic thin plate regression spline (Wood, 2006, p. 225). The `bs="cr", k=5` arguments (shown in Figure 25) to the $f_6(P)$ smooth function represent cubic regression splines (`bs="cr"`) with the dimension `k=5` representing "knots" of the spline (Wood, 2006, p. 226). The spline dimensions were chosen through visual evaluation of figures similar to Figures 26 and 27.

The residuals of the discharge model are shown in Figure 28, and summary statistics of the residuals are shown in Figure 25. Because of overplotting, gray transparency was used for Figure 25 to enhance visual density of the data point distribution. The Akaike Information Criterion is a measure of information content of a regression model. The statistic accounts for a trade off between the number of parameters and the fit of the model; small values are sought. The Akaike Information Criterion is −3,830 for the model in Equation 26 but −281 for the model lacking f_5 and f_6. The percent change in residual standard error from

```
DISCHARGE GENERALIZED ADDITIVE MODEL (QGAM), SI UNITS

Select Abbreviations:
log = base-10 logarithm used on Q, A, and B
Q   = discharge in cubic meters per second
A   = cross-section area in square meters
B   = water-surface top width in meters
oem = OmegaEM parameter (Asquith and Roussel, 2009)

Family: gaussian
Link function: identity

Formula:
logQ ~ logA + logB + oem +
       s(LongitudeDegrees, LatitudeDegrees, k = 14) +
       s(MeanAnnualPrecipMillimeters, bs = "cr", k = 5)

Parametric coefficients:
             Estimate Std. Error t-value Pr(>|t|)
(Intercept) -0.289609   0.006156  -47.05   <2e-16
logA         1.269194   0.004927  257.59   <2e-16
logB        -0.224712   0.007641  -29.41   <2e-16
oem          0.286524   0.028057   10.21   <2e-16
---

Approximate significance of smooth terms:
                                       edf Ref.df      F p-value
s(LongitudeDegrees,LatitudeDegrees) 12.87  13.00 187.19  <2e-16
s(MeanAnnualPrecipMillimeters)       4.00   4.00  25.96  <2e-16
---

R-sq.(adj) =  0.949   Deviance explained = 94.9%
GCV score = 0.047158  Scale est. = 0.047103  n = 17727
Residual Standard Error (gaussian family) = 0.217032

RESIDUAL SUMMARY
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-1.04100 -0.12800  0.01848  0.00000  0.14320  1.05000
```

Figure 25: Summary in R output of generalized additive model of base-10 logarithm of discharge based on statistical relations between the base-10 logarithms of discharge and water-surface top width, OmegaEM parameter by Asquith and Roussel (2009), and separate smooth functions of longitude and latitude $f_5(l, k)$ (figure 26) and mean annual precipitation $f_6(P)$ (figure 27); select acronyms are defined in Notation section

the model lacking f_5 and f_6 to the model in Equation 26 is -9.6 percent. A preference for the more complex model involving the smooth functions f_5 and f_6 is made.

Lastly, loose interpretation of the parametric coefficients can be made that are consistent with well-known hydraulic constraints. The positive coefficient on A shows that Q increases with increasing A; the negative coefficient on B shows that Q decreases with increasing B. The positive coefficient on OmegaEM indicates that Q increases in proportion to

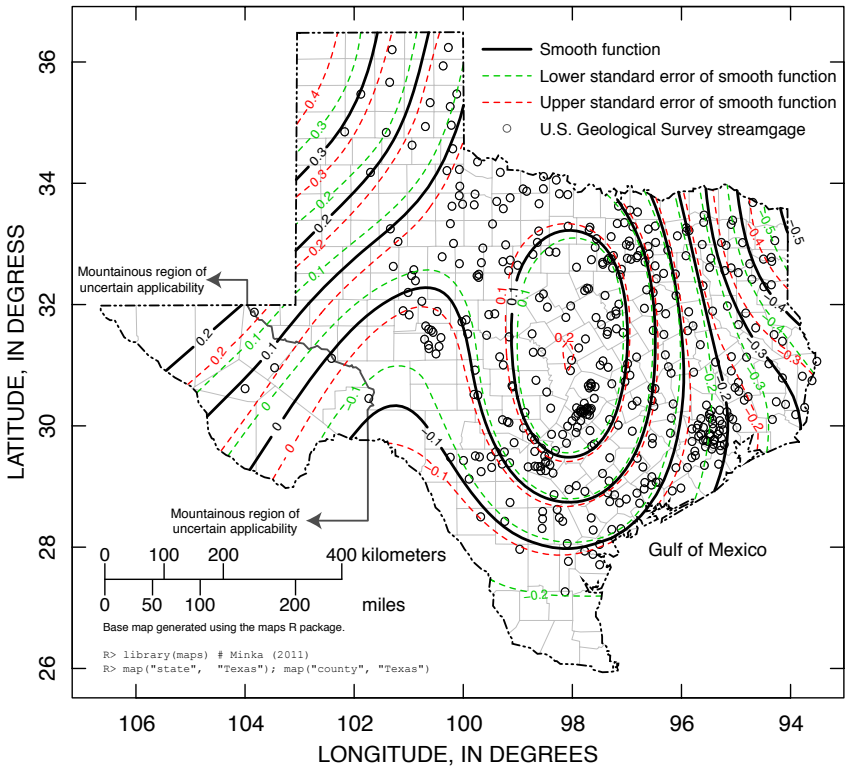

Figure 26: Smooth function $f_5(l, k)$ of location in Texas for the discharge model shown in Figure 25

OmegaEM. OmegaEM takes on a positive value in the central part of Texas (the region demarked by positive OmegaEM values) and is greatest along the Balcones escarpment in south central Texas. (O'Connor and Costa, 2003, p. 9) identify this region (Balcones escarpment) of the nation as having "concentrations of large floods." (Asquith and Roussel, 2009, p. 23) provide further and relevant discussion. Thus, OmegaEM acts to increase Q in QGAM near the central part of Texas and reduce Q in other parts. The smooth function $f_5(l, k)$ of location also shows a tendency for larger Q in the central part of Texas. The smooth function $f_6(P)$ shows that there is a subtle relation between P and Q that is difficult to interpret given the presence of the two other spatially varying parameters (OmegaEM and f_5).

5.2 Generalized Additive Model of Mean Velocity

The final VGAM in R output is shown in Figure 29. For VGAM, each of the predictor variables is statistically significant. The adjusted R-squared value is about 0.67, and the residual standard error is about $s = 0.063$ fifth root of m/s, which is the square root of the "Scale est." because a Gaussian family was used for this GAM. For the final VGAM model

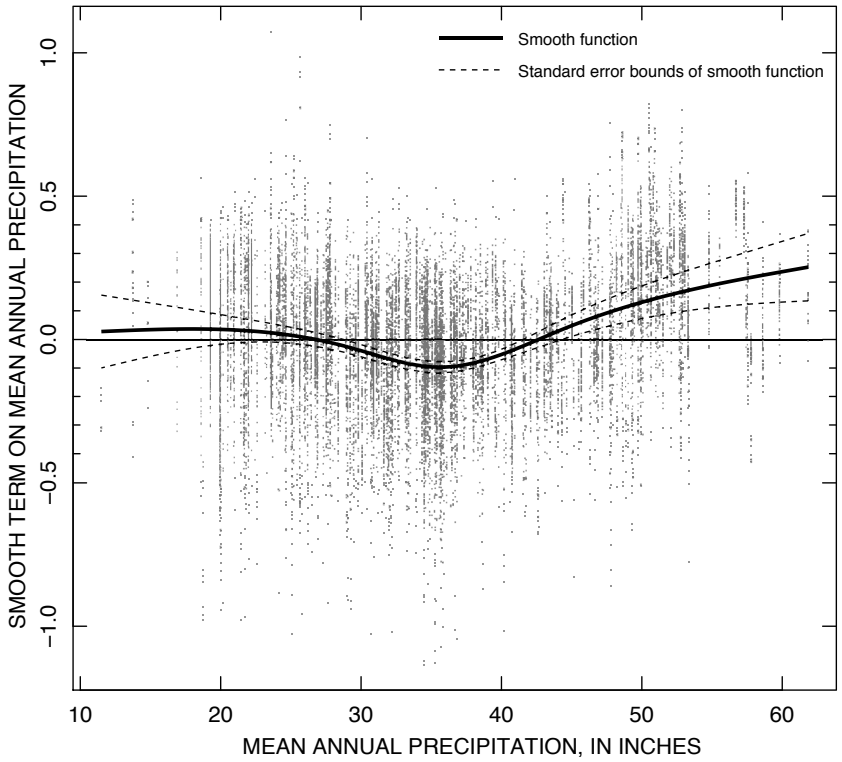

Figure 27: Smooth function $f_6(P)$ of mean annual precipitation for the discharge model shown in Figure 25

reported here, two discharge measurements for two streamgages (USGS station numbers: 08105000 and 08176500) were removed but the overall streamgage count remained at 424. The VGAM with the coefficients shown in Figure 29 can be written as:

$$\begin{aligned} V^{1/5} &= 0.9758 + 0.1588\log(Q) - 0.1820\log(B) + 0.0854\Omega \\ &\quad + f_9(\text{longitude}, \text{latitude}) + f_{10}(P) \end{aligned} \tag{27}$$

where log is base-10 logarithm, V is mean velocity in m/s transformed by the fifth root, Q is discharge in m^3/s, B is top width in m, Ω is the OmegaEM parameter from Figure 24, P is mean annual precipitation in mm, and f_9 and f_{10} are "smooth functions" of the indicated predictor variables in Figures 30 and 31, respectively. For Figure 30, the base map and superimposed smooth lines were created in R using graphic capabilities of packages by Minka (2011) and Wood (2009), and Figure 31 was created using graphic features by Wood (2009). The red, green, and black lines as ensembles of three for each numerical value shown in Figure 30 are not all shown for reasons such as grid resolution for the graphic, nonuniform distribution of streamgages, and general statistical magnitude of the two-dimensional smooth surface.

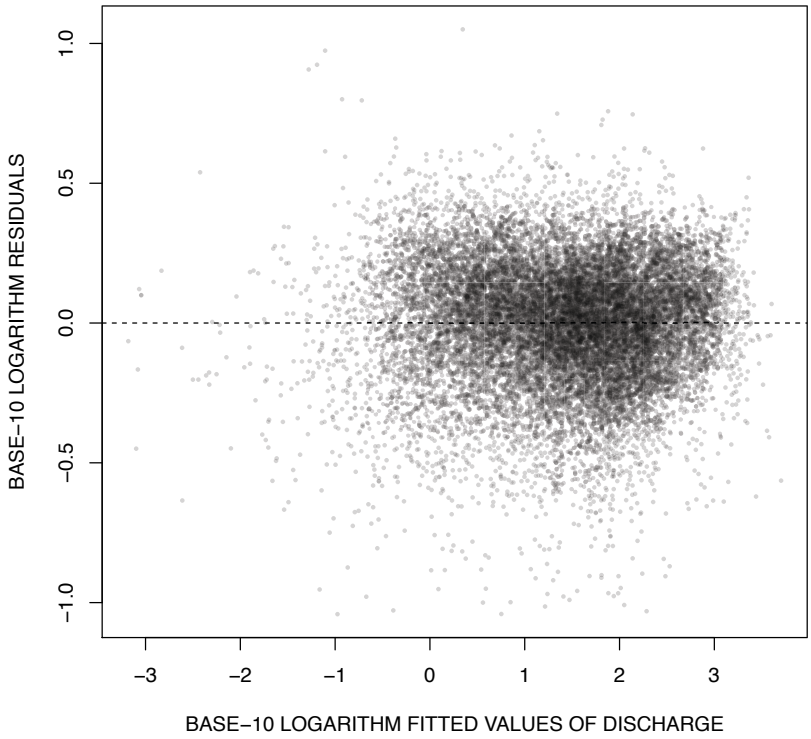

Figure 28: Residuals for the discharge model shown in Figure 25

The `k=14` argument (shown in Figure 29) to the $f_9(\text{longitude}, \text{latitude})$ or $f_9(l, k)$ smooth function of location represents the dimension of the isotropic thin plate regression spline (Wood, 2006, p. 225). The `bs="cr", k=5` arguments (shown in Figure 29) to the $f_{10}(P)$ smooth function represent cubic regression splines (`bs="cr"`) with the dimension `k=5` representing "knots" of the spline (Wood, 2006, p. 226). The spline dimensions were chosen through visual evaluation of figures similar to Figures 30 and 31.

The residuals of the mean velocity model are shown in Figure 32, and summary statistics of the residuals are shown in Figure 29. Because of overplotting, gray transparency was used for Figure 29 to enhance visual density of the data point distribution. The Akaike Information Criterion is −47,700 for the model in Equation 27 but −42,100 for the model lacking f_9 and f_{10}. The percent change in residual standard error from the model lacking f_9 and f_{10} to the model in Equation 27 is −14 percent. A preference for the more complex model involving the smooth functions f_9 and f_{10} is made.

Again, loose interpretation of the parametric coefficients can be made that are consistent with well-known hydraulic constraints. The positive coefficient on Q shows that V

```
VELOCITY GENERALIZED ADDITIVE MODEL (VGAM), ENGLISH UNITS

Select Abbreviations:
tV  = fifth-root of mean velocity in feet per second
log = base-10 logarithm used on Q and B
Q   = discharge in cubic feet per second
B   = water-surface top width in feet
oem = OmegaEM parameter (Asquith and Roussel, 2009)

Family: gaussian
Link function: identity

Formula:
tV ~ logQ + logB + oem + s(LongitudeDegrees, LatitudeDegrees,
    k = 14) + s(MeanAnnualPrecipInches, bs = "cr", k = 5)

Parametric coefficients:
             Estimate Std. Error t-value Pr(>|t|)
(Intercept)  1.044800   0.002666  391.95   <2e-16
logQ         0.201457   0.001267  158.98   <2e-16
logB        -0.230772   0.002318  -99.54   <2e-16
oem          0.108277   0.010280   10.53   <2e-16
---

Approximate significance of smooth terms:
                                       edf Ref.df      F  p-value
s(LongitudeDegrees,LatitudeDegrees) 12.72  12.99 205.02  < 2e-16
s(MeanAnnualPrecipInches)            4.00   4.00  11.33 3.52e-09
---

R-sq.(adj) =  0.671   Deviance explained = 67.1%
GCV score = 0.0063971  Scale est. = 0.0063897  n = 17751
Residual Standard Error (gaussian family) = 0.079936

RESIDUAL SUMMARY
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.477100 -0.049340 -0.001011  0.000000  0.051530  0.511400
```

Figure 29: Summary in R output of generalized additive model of fifth root of mean velocity based on statistical relations between the base-10 logarithms of discharge and water-surface top width, OmegaEM parameter by Asquith and Roussel (2009), and separate smooth functions of longitude and latitude $f_9(l, k)$ (figure 30) and mean annual precipitation $f_{10}(P)$ (figure 31); select acronyms are defined in Notation section

increases with increasing Q; the negative coefficient on B shows that V decreases with increasing B. The positive coefficient on OmegaEM indicates that V increases in proportion to OmegaEM. This finding was anticipated (see discussion in Section 5.1). The smooth function $f_9(l, k)$ of location also shows a tendency for smaller V in the eastern part of Texas. The authors hypothesize that this observation is consistent with greater vegetation density in the riparian zones in the eastern parts of Texas than in the western parts, and vegetation is associated with larger P and other physiographic factors. The smooth

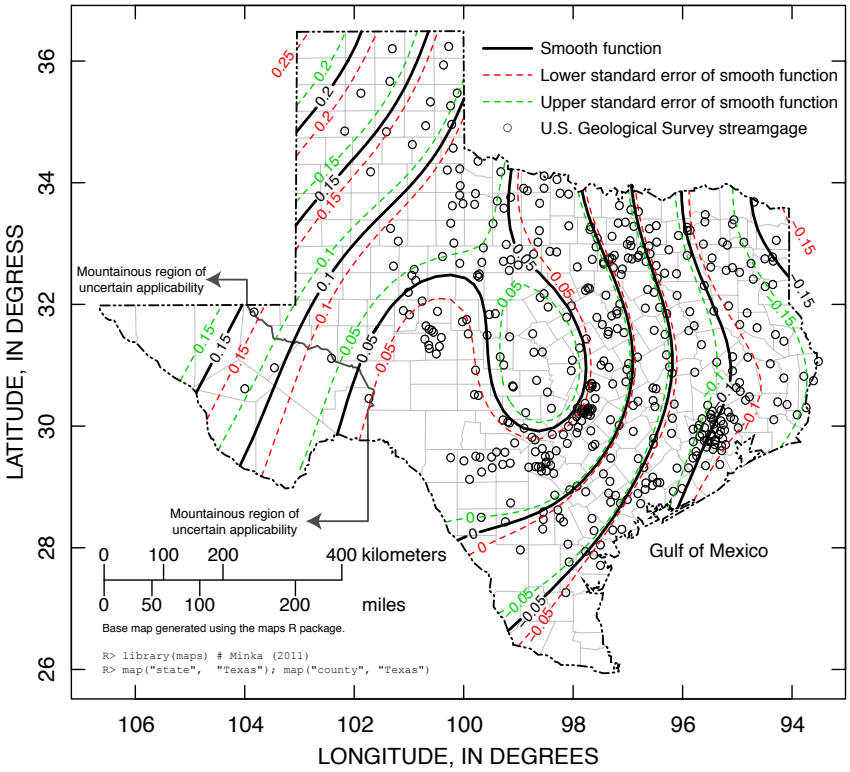

Figure 30: Smooth function $f_9(l, k)$ of location in Texas for the mean velocity model shown in Figure 29

function $f_{10}(P)$ shows that there is a subtle relation between P and V that is difficult to interpret given the presence of two other spatially varying parameters (OmegaEM and f_9).

5.3 Limitations of QGAM and VGAM and Thoughts for Improvement

According to the (National Research Council, 2004, p. 123) "a limitation of [the discharge measurement database] is that [streamgages] are chosen to have particular channel characteristics, such as the existence of a control section that will ensure a unique rating curve." The (National Research Council, 2004, p. 123) continues, "the channel characteristics of [streamgage] locations may thus not be representative of randomly selected locations at any point along the entire length of a stream or river." This last statement is particularly relevant for regional analysis of discharge measurement databases in that many high-magnitude discharge measurements are made at bridge crossings; the primary end-user application for VGAM is foreseen to be at or near bridge crossings in Texas. The general applicability or unapplicability of QGAM and VGAM for other cross sections of streams in Texas is difficult to quantitatively assess.

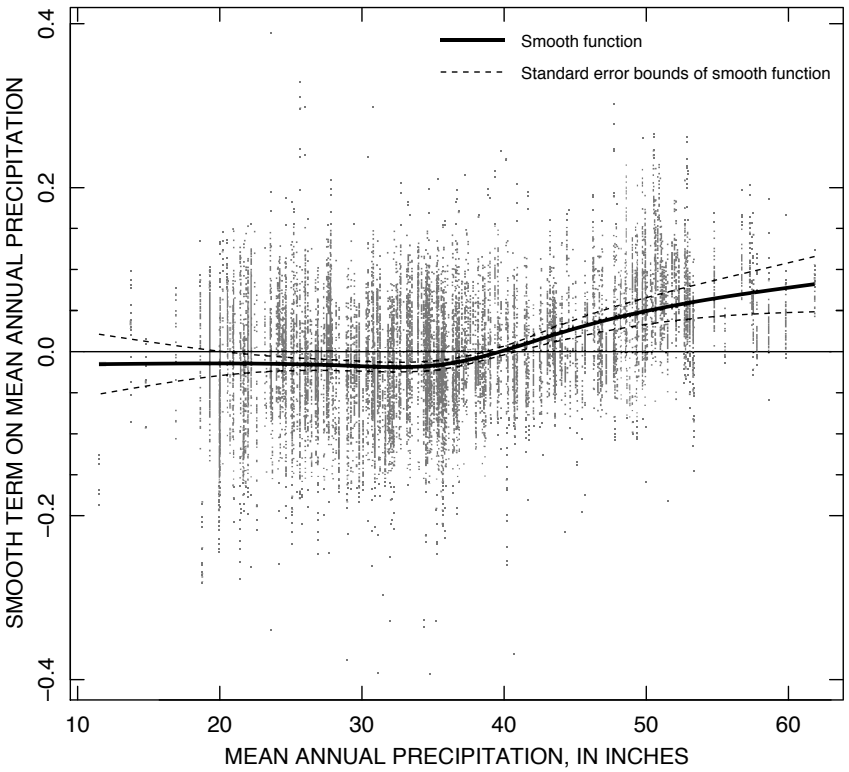

Figure 31: Smooth function $f_{10}(P)$ of mean annual precipitation for the mean velocity model shown in Figure 29

Assuming that the QGAM and VGAM do have acceptable applicability for other cross sections in Texas, additional discussion of applicability in terms of location is needed. The far western part of Texas is a mountainous region (Figures 26 and 30) with few USGS streamgages. The applicability of QGAM and VGAM is uncertain, but the models might retain some but difficult to quantify applicability in far western Texas. The number of streamgages diminishes rapidly towards the southernmost part of Texas; however because of the low-relief terrain, similarity in soils and vegetation, and orientation of the region with respect to the Gulf of Mexico, the authors suggest that QGAM and VGAM remain applicable. Lastly, the far north-northwestern parts of Texas also have few streamgages. By consideration of the physiographic features and the preponderance of branded sand channels in that general region, the authors suggest that QGAM and VGAM might retain some but difficult to quantify applicability.

General enhancement to the GAM diagnostics should be attainable through deliberate and systemic review of the summary statistics of A, Q, V, Fr, and B (recall such statistics for the entire database listed in Table 2). It might be possible for analysts to select particular variable thresholds. For example, all discharge measurements with $0.1 \leq \text{Fr} < 1$ or $1 \leq B \leq 2{,}000$ m could be retained and the regional analysis proceeding from there.

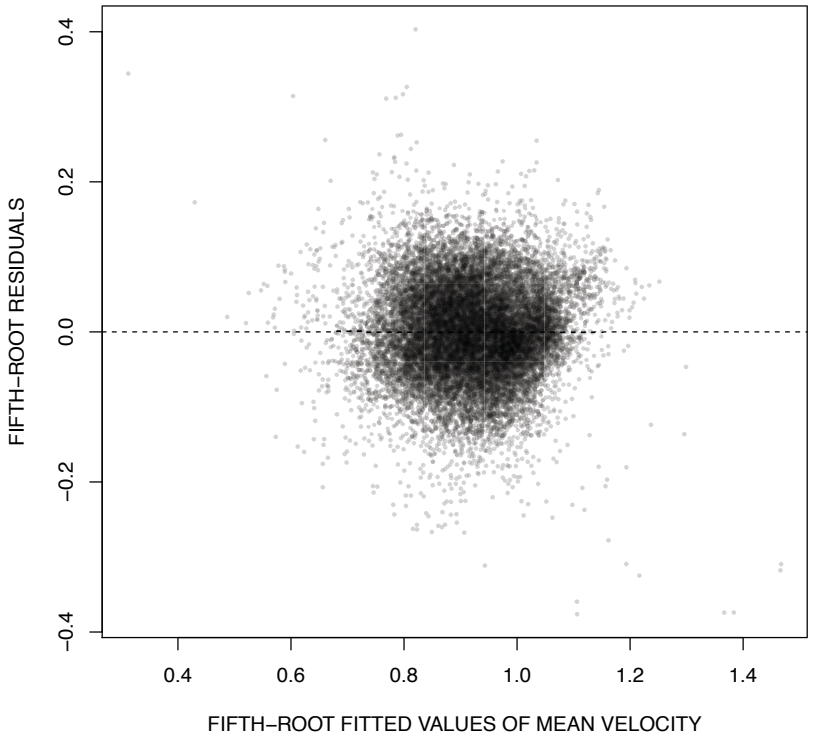

Figure 32: Residuals for the mean velocity model shown in Figure 29

A suggested approach beyond conventional residual or standardized residual plots would be an evaluation of the inherently coupled relations between Q and V on a per-streamgage basis. For example, it is known that the Q and V for most streamgages show positive association (Q increasing with V and vice versa); however, a not insubstantial number of streamgages do show negative association between Q and V. Could the generalized association (positive or negative) of Q and V for a given streamgage be used for further statistical enhancement?

6 Products: Method 2 - Empirical Distributions

6.1 Primary Variables

Empirical distributions are a statistical tool to estimate mean section velocity or other selected parameters for storm flows or other conditional discharges at ungauged locations. These distributions are an alternative to a regional regression or regression-like approach such as the one presented in section 5 (Asquith et al., 2013) that provides an equation for estimation of the expected value for mean velocity and/or discharge for an ungaged location and the prediction limits of that estimate. Prediction limits are analogous to empirical distributions presented in this section.

The empirical distributions presented herein are based on the refined database (section 3.4). Accessing the distributions is accomplished using R, and examples of such use are provided herein. A particularly powerful feature of R as the tool to access the database is the ability to rapidly construct conditional distributions, where the distributional information is conditioned on some criteria in the database. Conditioning addresses concerns such as where the 95th percentile discharge from all observations may be less meaningful than the 95th percentile discharge for all observations from drainage areas less than 40 square miles (Discharge conditioned on drainage area).

6.2 Unconditioned Distributions

Unconditioned distributions are empirical distributions based on the entire database without regard to any of the other retained variables. They are constructed using either the `quantile()` function in R or using a Weibull plotting position formula. Loading the database into R is illustrated in Listing 1.

Listing 1: R code demonstrating loading the database and preparing some useful plot labels

```
# EXAMPLE 1 # ** Loading the database into R, a useful function, and persistent plot
    labels
txdot0_6654.db <- read.table("database_txdot0-6654.txt",
                             header=TRUE, sep="|")
DB <- txdot0_6654.db; # shorten the database name considerably

"weibullpp" <- function(x, sort=TRUE) {
    denom <- length(x) + 1
    ranks <- rank(x, ties.method = "first")
    ifelse(sort, return((sort(ranks))/denom), return((ranks)/denom))
}
XLAB  <- "NONEXCEEDANCE PROBABILITY"
YLABV <- "VELOCITY, IN FEET PER SECOND"
YLABQ <- "DISCHARGE, IN CUBIC FEET PER SECOND"
```

The three approaches are illustrated in using the database to generate useful plots, tabulations, and finally direct access using the `quantile()` function.

6.2.1 Discharge (Q)

The entire database was analyzed to produce an empirical discharge distribution for discharge for the measured values in Texas. Figure 33 results from computing the plotting position for each entry in the database.

The script in R for generating the figure is shown in Listing 2.

Listing 2: R code demonstrating constructing an empirical distribution function plot for discharge

```
# EXAMPLE 2 #  ** Empirical Unconditioned Distribution of Discharge
plot(weibullpp(DB$Q),sort(DB$Q),log="y",xlab=XLAB,ylab=YLABQ,type="s",lwd=3,tck=1, main="
     Empirical CDF for Discharge")
```

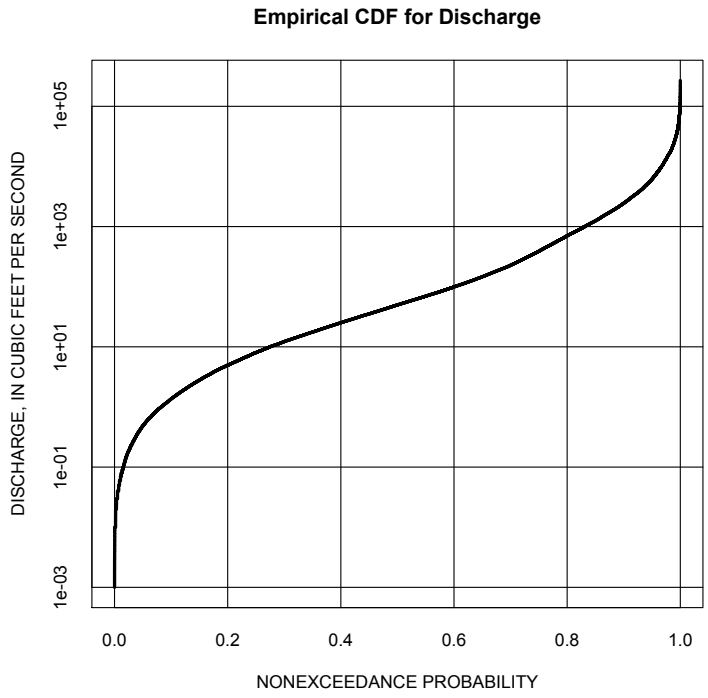

Figure 33: Empirical cumulative distribution for discharge in Texas

Table 3 is a tabular representation of the curve displayed in Figure 33.

The script in R for generating the table in Table 3 is shown in Listing 3. The tabular output in R was copied into an Excel worksheet, then pasted into the typesetting program to generate Table 3. This step (cut-paste-reformat) is not needed to use the tools, but was used for this particular table to highlight that the `quantile()` function returns the smallest value in the database at the 0th percentile and the largest value at the 100th percentile level.

Both these representations can be used to answer questions like what is the probability of

observing a discharge less than 12.2 cubic feet per second (without regards to where in the state we may be)? The answer would be to find the value of interest and then read the associated non-exceedance probability either from the graph or the tabulation. In this case about 30 percent of the observations are smaller than 12.2, so one could expect to observe such a discharge in a random measurement about 30 percent of the time.

Listing 3: R code demonstrating constructing an empirical distribution tabulation for discharge

```
# EXAMPLE 3 #  ** Empirical Unconditioned Distribution of Discharge -- Tabular
    Representation
# Useful quantiles for tabular presentations
> EMPQUANT<-c
    (0,0.0001,0.001,0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95,0.99,0.999,0.9999,1.0)

> cbind(quantile(DB$Q,EMPQUANT)) # output from R below
               [,1]
0%     1.000000e-03
0.01%  2.715900e-03
0.1%   1.000000e-02
1%     6.000000e-02
5%     4.800000e-01
10%    1.360000e+00
20%    4.950000e+00
30%    1.220000e+01
40%    2.510000e+01
50%    4.980000e+01
60%    9.930000e+01
70%    2.270000e+02
80%    7.010000e+02
90%    2.440000e+03
95%    6.160000e+03
99%    2.680000e+04
99.9%  7.916820e+04
99.99% 1.479887e+05
100%   2.686000e+05
>
```

Lastly, instead of using a chart or a tabulation, the result can be recovered directly from the database using the `quantile()` function in R. Listing 4 illustrates the use of R to directly locate a value based on a desired non-exceedance probability.

Listing 4: R code demonstrating use of the quantile() function for discharge

```
# EXAMPLE 4 #  ** Empirical Unconditioned Distribution of Discharge -- Quantile Function
    to Assess a Particular Value
> quantile(DB$Q,0.3) #  output from R below
 30%
12.2
```

Table 3: Empirical cumulative distribution for discharge – Tabular representation.

Percent Non-Exceedance	Discharge (cfs)
0.0000*	0.001
0.0001	0.003
0.0010	0.010
0.0100	0.060
0.0500	0.480
0.1000	1.360
0.2000	4.950
0.3000	12.200
0.4000	25.100
0.5000	49.800
0.6000	99.300
0.7000	227.000
0.8000	701.000
0.9000	2440.000
0.9500	6160.000
0.9900	26800.000
0.9990	79168.200
0.9999	147988.700
1.0000**	268600.000

* Smallest observed value in the database.
** Largest observed value in the database.

6.2.2 Mean section velocity (V)

Mean section velocity distributions are constructed in the same fashion as discharge for the measured values in Texas. Figure 34 results from computing the plotting position for each entry in the database. The script in R for generating Figure 34 is shown in Listing 5, the only change is to replace `DB$Q` with `DB$V` in Listing 2.

Listing 5: R code demonstrating constructing an empirical distribution function plot for velocity

```
# EXAMPLE 5 #  ** Empirical Unconditioned Distribution of Velocity
plot(weibullpp(DB$V),sort(DB$V),log="y",xlab=XLAB,ylab=YLABQ,type="s",lwd=3,tck=1, main="
     Empirical CDF for Velocity")
```

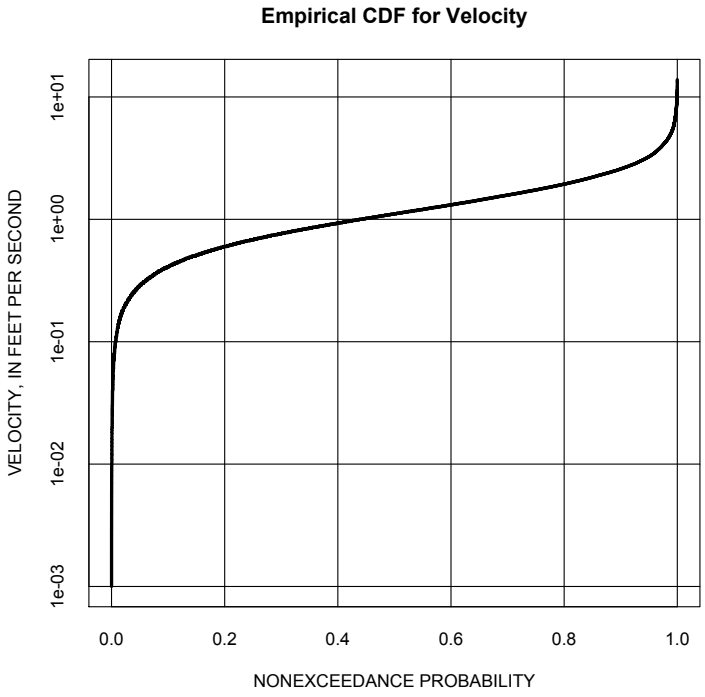

Figure 34: Empirical cumulative distribution for velocity in Texas

The tabular form for Figure 34 is presented in Table 4. Again the low and high values are identified directly in the table.

Using a similar question to that of discharge, if a computer model indicates a mean section velocity of 10 feet per second, that value corresponds to approximately a non-exceedance probability of over 99 percent. Hence in any random measurement of velocity, one would expect to observe a value less than 10 feet per second almost always.

Using the `quantile()` function and some trial-and-error one can determine that the non-exceedance probability for 10 feet per second is 99.95 percent (or 0.9995); the R code to find

Table 4: Empirical cumulative distribution for velocity – Tabular representation.

Percent Non-Exceedance	Mean Section Velocity (ft/s)
0.0000*	0.001
0.0001	0.004
0.0010	0.020
0.0100	0.120
0.0500	0.290
0.1000	0.410
0.2000	0.600
0.3000	0.760
0.4000	0.930
0.5000	1.110
0.6000	1.310
0.7000	1.580
0.8000	1.940
0.9000	2.580
0.9500	3.250
0.9900	5.190
0.9990	8.897
0.9999	11.350
1.0000**	13.850

* Smallest observed value in the database.
** Largest observed value in the database.

the values is illustrated in Listing 6. The interpretation is unchanged, computed velocities above 10 feet per second are unusual and not supported by the database.

Listing 6: R code demonstrating use of the quantile() function for velocity

```
# EXAMPLE 6 #  ** Finding quantile for velocity of 10 feet per second
> quantile(DB$V,0.9995) # output from R below
99.95%
    10
```

6.2.3 Topwidth (B)

Figure 35 is the empirical distribution for the hydraulic variable topwidth. The majority of observed topwidth values in Texas (99.9 percentile) are less than one mile; hence a computed topwidth in excess of one mile is unusual. Such wide flows would be anticipated in flood flow with substantial overbank involvement.

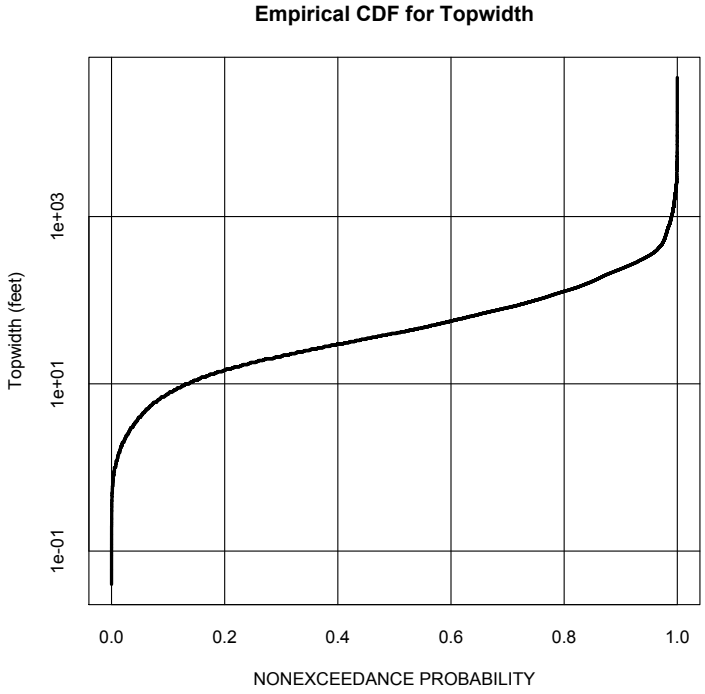

Figure 35: Empirical cumulative distribution for topwidth in Texas

The tabular form for Figure 35 is presented in Table 5. Again the low and high values are identified directly in the table. The listing to generate the topwidth graph and table is shown in Listing 7

Listing 7: R code demonstrating constructing an empirical distribution function for topwidth

```
# EXAMPLE 7 # ** Empirical distribution for topwidth
> plot(weibullpp(DB$B),sort(DB$B),log="y",xlab=XLAB,ylab="Topwidth (feet)",type="s",lwd=3,
    tck=1, main="Empirical CDF for Topwidth")
> cbind(quantile(DB$B,EMPQUANT)) # output from R below
            [,1]
0%         0.040
0.01%      0.200
0.1%       0.500
1%         1.200
5%         4.000
10%        7.600
20%       14.500
30%       21.300
40%       30.000
50%       40.000
60%       56.000
70%       81.000
80%      128.000
90%      235.000
95%      341.000
99%     1002.820
99.9%   2686.000
99.99%  8970.241
100%   45983.000
```

Table 5: Empirical cumulative distribution for topwidth – Tabular representation.

Percent Non-Exceedance	Topwidth (ft)
0.0000*	0.040
0.0001	0.200
0.0010	0.500
0.0100	1.200
0.0500	4.000
0.1000	7.600
0.2000	14.500
0.3000	21.300
0.4000	30.000
0.5000	40.000
0.6000	56.000
0.7000	81.000
0.8000	128.000
0.9000	235.000
0.9500	341.000
0.9900	1002.820
0.9990	2686.000
0.9999	8970.241
1.0000**	45983.000

* Smallest observed value in the database.
** Largest observed value in the database.

6.2.4 Flow Area (A)

Flow area is the final unconditioned empirical distribution presented; Figure 36 is the empirical distribution for the hydraulic variable flow area.[4] The R commands to generate the chart or table are shown in Listing 8.

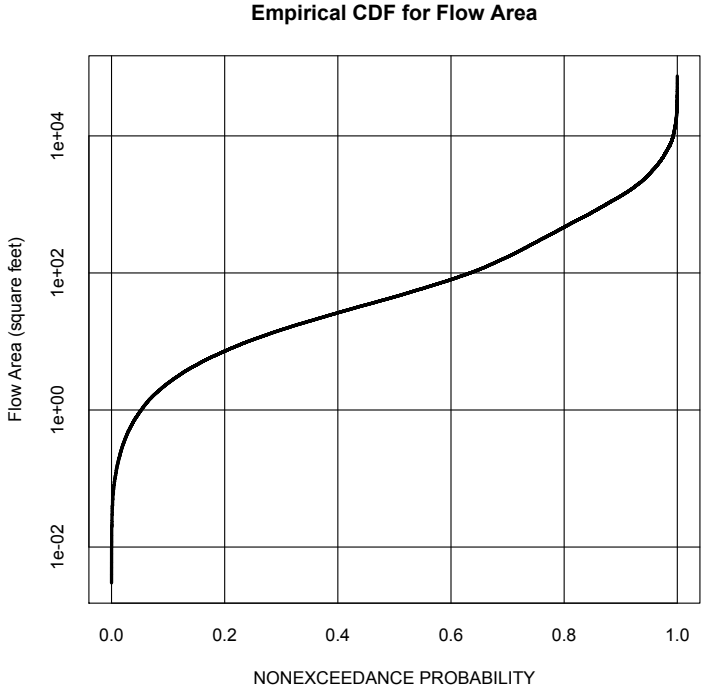

Figure 36: Empirical cumulative distribution for cross-section flow area in Texas

Listing 8: R code demonstrating building an empirical distribution function for flow area

```
# EXAMPLE 8 # ** Empirical distribution for flow area
> plot(weibullpp(DB$A),sort(DB$A),log="y",xlab=XLAB,ylab="Flow Area (square feet)",type="s
    ",lwd=3,tck=1, main="Empirical CDF for Flow Area")
> cbind(quantile(DB$A,EMPQUANT)) # output from R below
           [,1]
0%        0.003
0.01%     0.010
0.1%      0.030
1%        0.150
5%        0.940
10%       2.480
20%       7.220
30%      14.800
40%      26.200
50%      44.700
60%      80.000
70%     171.000
80%     465.000
90%    1320.000
95%    2750.000
99%    8472.870
```

[4]Distinct from the hydrologic variable of drainage area which will be examined in the conditioned distributions.

```
99.9%  22484.100
99.99% 41953.430
100%   74700.000
```

Table 6 is the tabular representation of Figure 36.

Table 6: Empirical cumulative distribution for flow area – Tabular representation.

Percent Non-Exceedance	Flow Area (ft^2)
0.0000*	0.003
0.0001	0.010
0.0010	0.030
0.0100	0.150
0.0500	0.940
0.1000	2.480
0.2000	7.220
0.3000	14.800
0.4000	26.200
0.5000	44.700
0.6000	80.000
0.7000	171.000
0.8000	465.000
0.9000	1320.000
0.9500	2750.000
0.9900	8472.870
0.9990	22484.100
0.9999	41953.430
1.0000**	74700.000

* Smallest observed value in the database.
** Largest observed value in the database.

6.3 Conditioned Distributions

The real value of the database and empirical distributions accessed using R, is the ability to condition the distributions on other variables in the database. This conditioning is presented by example, but in essence is a filtering process.

6.3.1 Discharge conditioned on contributing drainage area (`Q | CDA`)

A logical conditioning is to ask from the database what is a certain non-exceedance probability associated with all discharges recorded from drainage areas less than some value (or even bracketed). This question is a conditioned probability statement. Operationally we would search the database and exclude all records associated with drainage areas larger

that the prescribed conditioning value, then compute the empirical distribution from the remaining values.

As an example consider what is the empirical nonexceedance discharge distribution for gages having contributing drainage areas of 100 square miles or less? The drainage area is the variable `CDA` and building the conditional distribution is illustrated in Listing 9.

Listing 9: Building conditioned empirical distributions, for discharge conditioned on drainage area

```
# EXAMPLE 9
#   ** Empirical Distribution of Discharge Conditioned on Contributing Drainage Area
# Get all measurements for watershed area less than 100 square miles
> QQ<-DB[DB$CDA < 100,]$Q  # Filter the database, put results in QQ
# Plot an empirical distribution (sneaky here, actually don't use QQ just yet!)
> plot(weibullpp(DB[DB$CDA < 100,]$Q),sort(DB[DB$CDA < 100,]$Q),log="y",xlab=XLAB,ylab=
    YLABQ,type="s",lwd=3,tck=1, main="Conditional Empirical CDF for Discharge for CDA <
    100 sq.mi.")
# Compute the 50th percentile using the quantile() function of R---this is the median
> print(quantile(QQ,0.5)) # output from R below
  50%
7.735
# Generate a tabulation
> cbind(quantile(QQ,EMPQUANT)) # output from R below
              [,1]
0%     1.00000e-03
0.01%  1.37310e-03
0.1%   1.00000e-02
1%     4.00000e-02
5%     1.70000e-01
10%    4.10000e-01
20%    1.12000e+00
30%    2.24000e+00
40%    4.19400e+00
50%    7.73500e+00
60%    1.50000e+01
70%    3.07000e+01
80%    6.82000e+01
90%    2.46000e+02
95%    8.29350e+02
99%    3.70380e+03
99.9%  1.02807e+04
99.99% 1.53000e+04
100%   1.65000e+04
>
```

The result of conditioning is apparent in the median value of discharge. When all drainage areas were considered, the median discharge was about 50 cfs, but when conditioned on contributing drainage area less than 100 square miles, the median is around 8 cfs — about six times smaller. The result is anticipated. Smaller drainage areas should produce smaller discharges for similar weather conditions.

Figure 37 is a plot of the conditioned empirical distribution of discharge. As a guideline, the author suggests that when conditioning, the analyst check the array sizes and try to maintain about 100 records after conditioning; with this suggestion each retained record represents about 1 percent of any empirical distribution subsequently generated.

6.3.2 Multiple conditioning (`Q | CDA, B, FDC`)

Multiple conditioning based on several variables is feasible. Listing 10 is an example of a multiple conditioning empirical distribution where the analyst seeks the 95th percentile of discharges from contributing drainage areas less than 100 square miles, with topwidth less

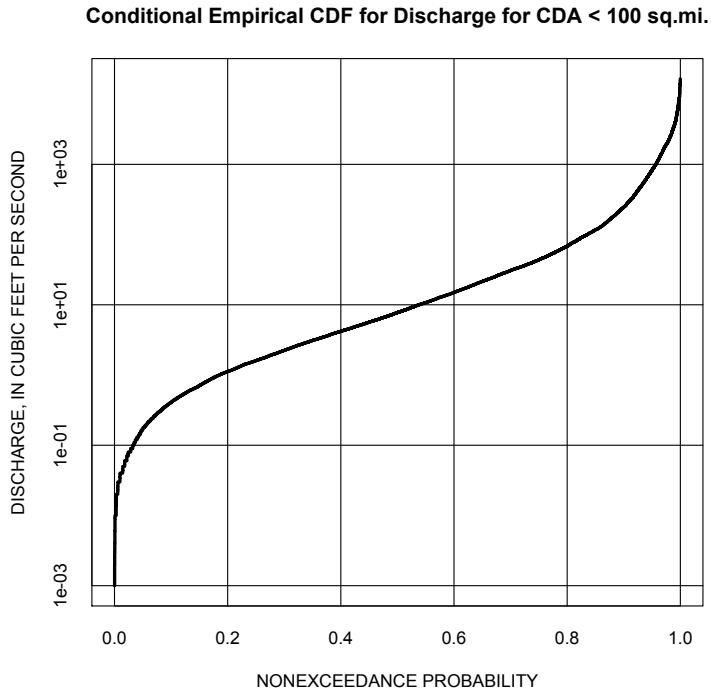

Figure 37: Empirical cumulative distribution for discharge for drainage areas less than 100 square miles in Texas

than 40 feet, and discharges greater than the 80th percentile on the station's individual flow duration curve.

Listing 10: R code demonstrating a multiple condition, empirical distributional analysis for discharge

```
# EXAMPLE 10 #   ** Empirical distribution of discharge for multiple conditions
# Get all measurements for watershed area less than 100 square miles
#   and topwidth less than 40 feet
#   and greater than 80th percentile on the flow-duration curve
# Compute the 75th percentile of these measurements
> print(quantile(DB[DB$CDA < 100 & DB$B < 40 & DB$FDC > 0.80, ]$Q, 0.95)) # output from R
    below
  95%
  224.6
```

The result in this case is a discharge of about 225 cfs.

6.3.3 Mean section velocity conditioned on contributing drainage area (V | CDA)

Listings illustrating conditioning velocity on other variables are presented below.

Consider the question as to what is the 90th percentile velocity for contributing drainage

areas less than 20 square miles? Listing 11 illustrates the R commands that answer this question.

Listing 11: Building conditioned empirical distributions, for velocity conditioned on drainage area

```
# EXAMPLE 11
#   ** Empirical Distribution of Velocity Conditioned on Contributing Drainage Area
# Get all measurements for watershed area less than 20 square miles
> VV<-DB[DB$CDA < 20,]$V  # Filter the database, put results in VV
> length(VV) # check length, want to be more than 100
[1] 2066
> print(quantile(VV,0.9)) # output from R below
  90%
2.415
# Generate a tabulation
> cbind(quantile(VV,EMPQUANT)) # output from R below
            [,1]
0%     0.0070000
0.01%  0.0076195
0.1%   0.0200000
1%     0.0800000
5%     0.1900000
10%    0.2700000
20%    0.4000000
30%    0.5000000
40%    0.6000000
50%    0.7300000
60%    0.8800000
70%    1.1200000
80%    1.5200000
90%    2.4150000
95%    3.7975000
99%    5.7965000
99.9%  7.1135000
99.99% 8.0279550
100%   8.2200000
```

The result is a velocity of 2.4 feet per second represents the 90th percentile for drainage areas of 20 square miles or less. Thus if a modeler has computed a velocity in excess of 2.4 feet per second for a 20 square mile watershed, that is an unusual value and would indicate some additional explanation.

6.3.4 Multiple conditioning (V | CDA, B, FDC)

As with discharge multiple conditions for velocity can be considered. For example suppose the analyst seeks the empirical velocity distribution for drainage area less than 10 square miles, topwidth less than 30 feet, and flows (in the retained records) that are at the 90th percentile or greater on the individual station flow duration curves. Listing 12 illustrates conditioning the velocity on drainage area, topwidth, and station flow duration curve values. The result is tabulated using the `quantile()` function approach presented in earlier examples.

Once the output is examined, the median value for these conditions is about 1.5 feet per second, the largest value retained after the conditioning is 6 feet per second. We can now consider what the distribution is conveying.

If a modeler were to calculate a velocity for a contributing drainage area of 10 square miles, for a topwidth at the point of interest of 30 feet and arrive at a value of say 7 feet per second, the modeler should be concerned.

The database suggests that such a value has not yet been observed in Texas streamflow, even when considering flows at the 90th percentile on the individual station flow duration curves. Hence that value, 7 feet per second, unless otherwise explained would be disturbing.[5]

Listing 12: Building multiple conditions for mean section velocity

```
# EXAMPLE 12 #   ** Empirical Distribution of Velocity Conditioned on Contributing
    Drainage Area, Topwidth, and Flow Duration
# Get all measurements for watershed area less than 10 square miles
#   topwidth less than 30 feet
#   and greater than 90th percentile on the station flow-duration curve
VV<-DB[DB$CDA < 10 & DB$B < 30 & DB$FDC > 0.9,]$V  # filter database, put result in VV
> length(VV) # check length -- OK but getting to be on the small side!
[1] 132
# Build a tabular empirical distribution.
> cbind(quantile(VV,EMPQUANT)) # output from R below
           [,1]
0%     0.340000
0.01%  0.341179
0.1%   0.351790
1%     0.433100
5%     0.530000
10%    0.662000
20%    0.830000
30%    1.006000
40%    1.344000
50%    1.485000
60%    1.886000
70%    2.217000
80%    2.772000
90%    3.453000
95%    4.368000
99%    5.531600
99.9%  5.964630
99.99% 5.996463
100%   6.000000
```

6.4 Ancillary parameter (derived) distributions (Fr)

The database can be used to compute ancillary (or derived) values, such as Froude number,

$$Fr = \frac{V}{\sqrt{gA/B}} \tag{28}$$

where V is the mean section velocity, A is the cross sectional flow area, B is the topwidth, and g is gravatitional acceleration.

Listing 13 illustrates the construction of a derived distribution for Froude number for contributing drainage areas less than 10 square miles. The result is tabulated using the `quantile()` function. For this example the median Froude number for such conditions is about 0.25, a decidedly subcritical flow. In fact, based on the database, supercritical flow is unusual occurring above the 99.9th percentile.

[5]The value from the model may indeed correct, but based on observations it is unusual. The whole point of the tool is to guide when a value is unusual and help identify potential data entry errors that would otherwise go unnoticed.

Supercritical flow is indicated for some measurements, but very few computed in this flow regime. It is unlikely that considerable frequency of supercritical flows actually exist and uncertainties in the values of `A` or `B` as reported during the measurement might not be representative enough and the apparent computation of supercritical flow is an artifact of sampling.

Listing 13: Building a derived (ancillary) empirical distribution

```
# EXAMPLE 13 #   ** Empirical Distribution of Froude number conditoned on drainage area
    less than 10 square miles.
# Get all measurements for watershed area less than 10 square miles, compute the Froude
    number, tabulate the result
FR <-DB[DB$CDA < 10,]$V/sqrt(32.2*DB[DB$CDA < 10,]$A/DB[DB$CDA < 10,]$B)  # compute Fr for
     CDA < 10, put into FR
> cbind(quantile(FR,EMPQUANT)) # output from R below
                [,1]
0%     0.007855307
0.01%  0.008281195
0.1%   0.012114189
1%     0.024655607
5%     0.059134675
10%    0.079623847
20%    0.122351979
30%    0.164806117
40%    0.204549940
50%    0.249712049
60%    0.296877395
70%    0.354723991
80%    0.421251480
90%    0.519535661
95%    0.596747258
99%    0.728635974
99.9%  0.950658157
99.99% 1.211274071
100%   1.240231395
```

6.5 Distributions of Additional Ancillary Parameters

Ancillary properties are properties such as stream power, Froude number, Darcy-Weisbach friction factor, bed shear stress, and other properties that are of secondary importance for checking the validity of HEC-RAS model output. While not of primary importance for checking HEC-RAS output, such parameters are important in general fluvial hydraulics, e.g. in support of scour and sediment transport concerns (HEC-18 and similar guidance), and are therefore calculated and provided as supplementary material.

A larger number various parameters such as fluvial reach classifications (e.g., Rosgen, 1994; Church, 2006), sediment transport parameters (dimensionless shear stress, transport rates, bed states, etc. (García, 2008)), and hydraulic parameters were investigated to see what could be calculated from the database. The lack of station-wide sediment grains size data was found to be the greatest limitation to the number of ancillary parameters that could be calculated. In the end, five ancillary parameters were calculated. Although they are not the primary quantities used for verifying HEC-RAS model output, they can provide general information regarding the range of observed values throughout Texas.

The primary measurements in the database of discharge, velocity, area and top width. The only ancillary parameter that could be calculated using purely measured quantities from

the database was the Froude number,

$$Fr = \frac{V}{\sqrt{gA/B}} \tag{29}$$

where V is average channel velocity, A is cross sectional flow area, B is the top width, and g is the gravitational acceleration.

All other ancillary parameters calculated required the use of the so-called main-channel slope, S_{mc}, which is defined as the change in elevation between the two end points of the main-channel divided by the distance, L (Asquith and Slade, 1997). S_{mc} was calculated by USGS researchers and added to the database of measured values at each station. In the calculation of S_{mc}, L is defined as the longest defined channel shown in a 10-meter digital elevation model (DEM) from the approximate watershed headwaters to the point of interest, and the elevation change between the two points is extracted directly from the 10-meter DEM. The main-channel slope is therefore more of a watershed slope based on the channel network than it is a local reach slope. Because of its calculation method, it is suspected that the main channel slope values will be, on average, somewhat higher than the local reach slopes at the stations because the main channel slope by definition incorporates elevation change further up in the watershed where slopes are likely higher. The main-channel slope was assumed constant for each station at all discharges. The additional ancillary properties that could be calculated with the dataset were the Manning roughness coefficient, n, the channel stream power, Ω, the Darcy-Weisbach friction factor, f, and the Chezy roughness coefficient, C. Each of these parameters are defined as:

$$n = \frac{c_o}{V} R_h^{2/3} S^{1/2} \tag{30}$$

$$f = \frac{8gR_hS}{V^2} \tag{31}$$

$$C = \frac{V}{\sqrt{R_hS}} \tag{32}$$

$$\Omega = \rho g Q S \tag{33}$$

where Q is the channel discharge, c_o is the unit system coefficient equal to 1.486 for US units and 1 for SI (1.486 in our calculations), R_h is the hydraulic radius (equal to the cross sectional flow area A divided by the wetter perimeter, P). To calculated these parameters from the database, it was assumed that the hydraulic depth, $D = A/B$ was equal to the hydraulic radius and that the local friction slope, S, was equal to the main channel slope.

Each of the five ancillary properties (equations 29-33) were computed from a filtered subset of the main database for which all zero or negative values of the main channel slope were removed; in all, 791 points were removed due to the filtering, leaving 89,083 measurements in the database. For presentation of the ancillary parameters, the complete database was broken down by mean daily flow percentile range. Two discharge percentile bands

were examined. The first was for discharges greater than the 90th percentile. These discharges were pulled out and analyzed as being representative of storm flow events (17,615 measurements in all). The second band examined was the set of flows which fell between the 40th and 60th percentiles. Flows within these percentiles are representative of mean flows (16,962 measurements in all). Calculation of the percentiles for each measurement was done by USGS personnel over the entire period of record without regard for changes in flow frequency due to such changes as the installation of a dam upstream (see TM2).

6.5.1 Nonexceedance curves for mean and storm flow events

Exceedance curves of the five ancillary parameters calculated for flows greater than the 90th percentile flow and for those between the 40th and 60th percentile are shown below in figures 38-42.

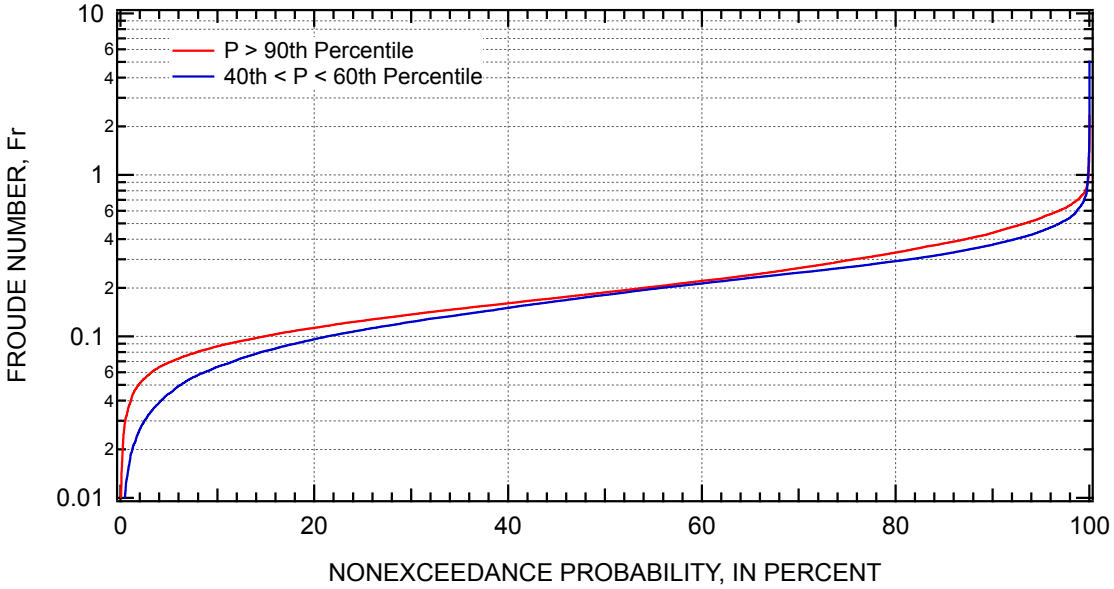

Figure 38: Nonexceedance probability for the Froude number Fr for flows above the 90th percentile and those between the 40th and 60th.

Relatively small differences in the nonexceedance curves are present for Fr, n, f, and C compared to the differences in stream power. The Froude number and the Manning n value are slightly higher for the majority of the >90th flows, though the very largest values of n in the mean flow band (between the 40th and 60th percentiles) surpasses those of the storm flow for similar nonexceedance values near the upper end. The same is true for the Darcy Weisbach friction factor (figure 40).

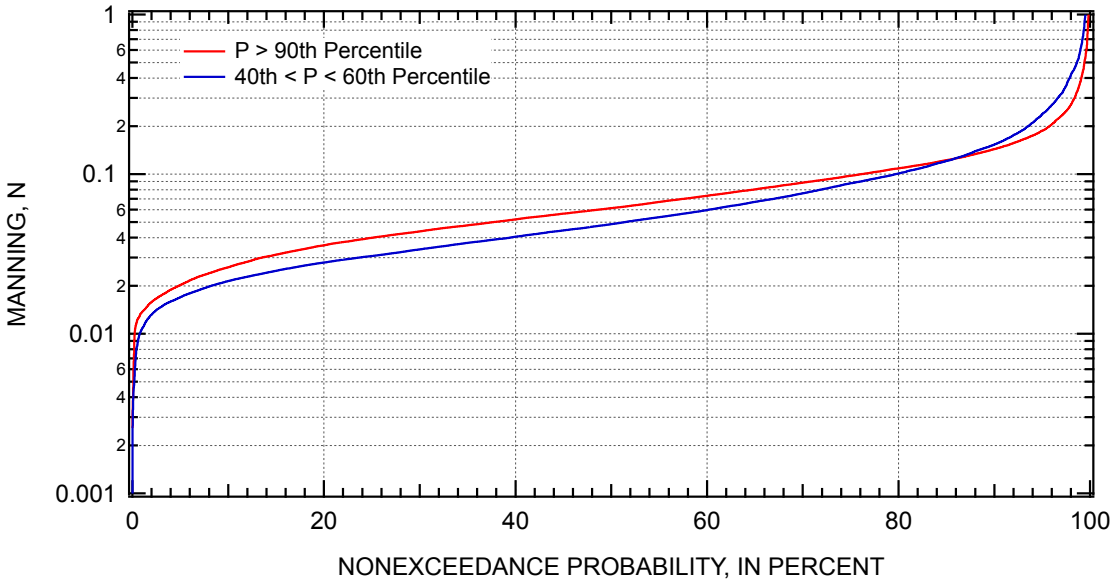

Figure 39: Nonexceedance probability for the Manning n for flows above the 90th percentile and those between the 40th and 60th.

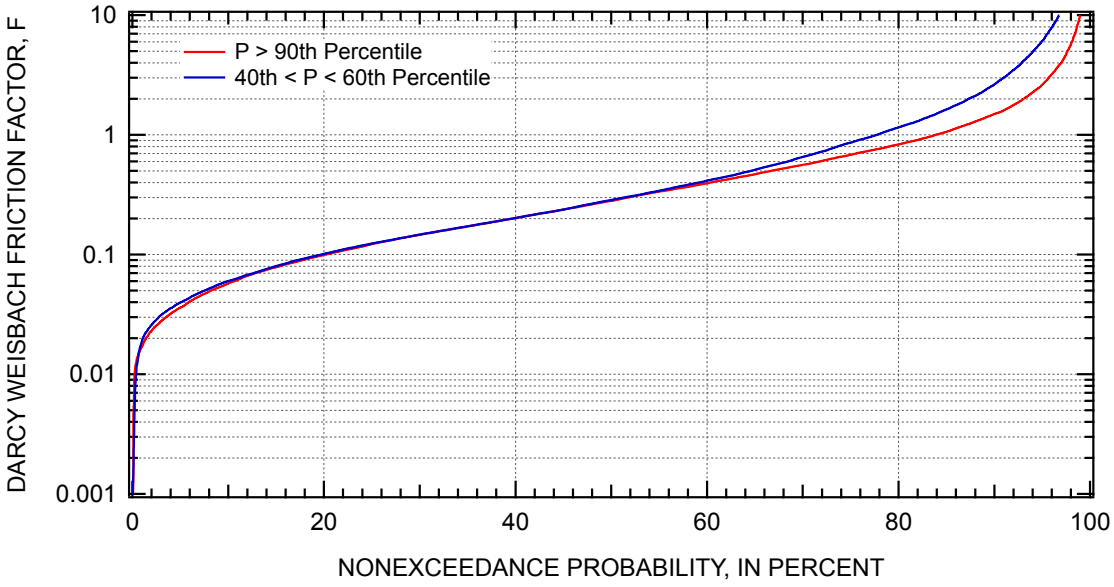

Figure 40: Nonexceedance probability for the Darcy-Weisbach f for flows above the 90th percentile and those between the 40th and 60th.

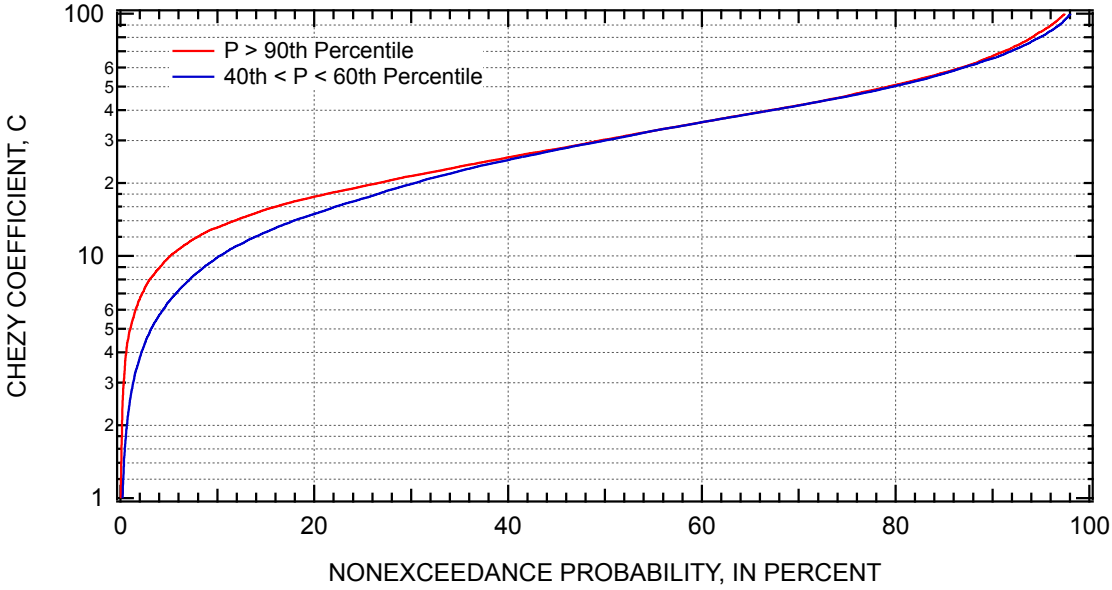

Figure 41: Nonexceedance probability for the Chezy C for flows above the 90th percentile and those between the 40th and 60th.

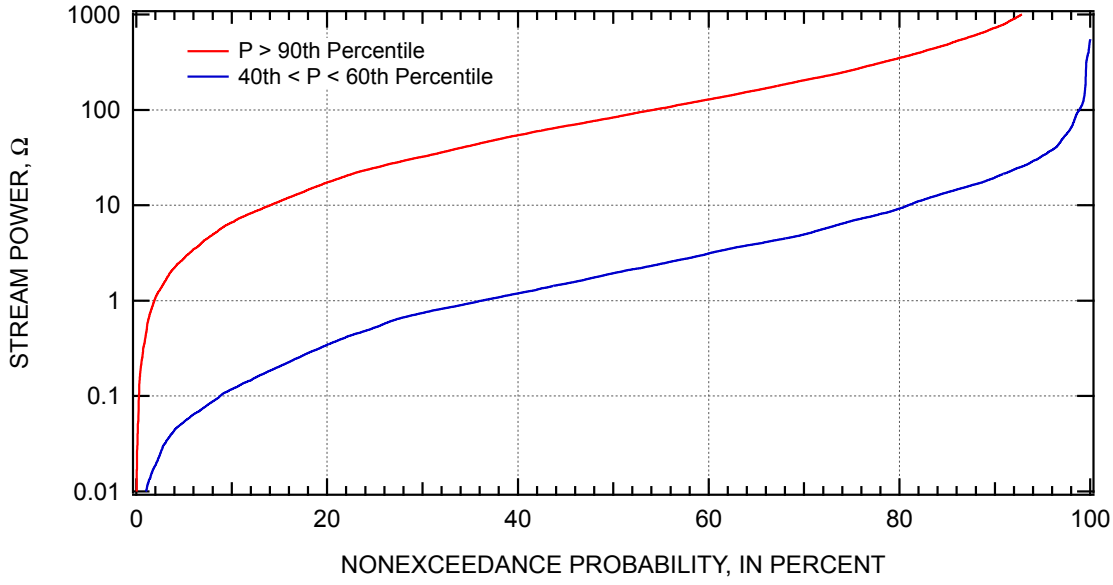

Figure 42: Nonexceedance probability for the stream power Ω for flows above the 90th percentile and those between the 40th and 60th.

6.5.2 Notes on the quality of the calculated parameters

It should be noted that the accuracy of the calculated ancillary data is suspect in places. This is not the case for Froude number, which can be calculated from the measurements in the database without any assumptions or use of a calculated slope. However, it is the case for the frictional parameters, n, f, and C and to a lesser degree the stream power Ω.

For example, there are 32 out of 17,615 (0.2%) calculated Manning n values greater than 1 in the storm flow database alone, and 4,218 values greater than $n = 0.1$ (24% of the database). While the Manning n value is a sites specific roughness factor that accounts for a wide variety of boundary and form resistance, values such as these are suspect. It is likely that the overly large resistance coefficients come from assuming that the hydraulic depth, $D = A/B$, is equal to the hydraulic radius, $R = A/W$, where WP is the wetted perimeter of the cross section, and from use of the Main Channel Slope for the local friction slope. In both cases, it is likely that each of the surrogate values D and S_{mc} will, on average, be greater than their corresponding true values of R and S_f. If this is true, then the calculated resistance values of n and f using surrogate values will be biased high (equation 30 and 31); the opposite will be true for the Chezy coefficient (equation 32). It is likely that biases introduced in assuming that $D = R$ will be strongest for relatively narrow channels with non-rectangular cross sections. The assumption is less limiting the more rectangular and wide the channel is since D and R are approximately equivalent for wide rectangular channels.

The stream power calculations do not assume that $R = D$, but they do make use of the computed main channel slope. Based on the method of main channel slope calculation, it is again thought that the calculated stream power values, Ω, are somewhat higher than the actual values.

6.5.3 Regionalization of ancillary parameters

In an attempt to tease out any possible regional dependence in the ancillary parameters, each gaging station was given three different physiographic region tags baed on the physiographic map of Texas produced by the Bureau Of Economic Geology (BEG) at the University Of Texas at Austin in 1996 (figure 19) and two additional more finely discretized physiographic region classifications based off of maps provided by the Texas Parks and Wildlife Department (TPWD). Maps and region tags for these two classifications along with box plots of the ancillary parameter broken down according to these three regionalization schemes are presented Appendix B. Plots of the ancillary parameters in the appendix are given two sets of floc conditions, with those two sets being for flows greater than 90th percentile flows (storm flows) and those falling between the 40th and 60th percentiles (mean flow).

7 Products: Method 3 - Hydraulic Geometry

7.1 Background

Alluvial rivers tend to construct their channel geometries and flood plains in consistent ways in terms of bankfull characteristics (García, 2008, chap. 3). Bankfull condition are attained when the river just start to spill out of its channel banks and onto its flood plains. On the rating curves, the bankfull condition is clearly defined by a slope change (see Figure 43).

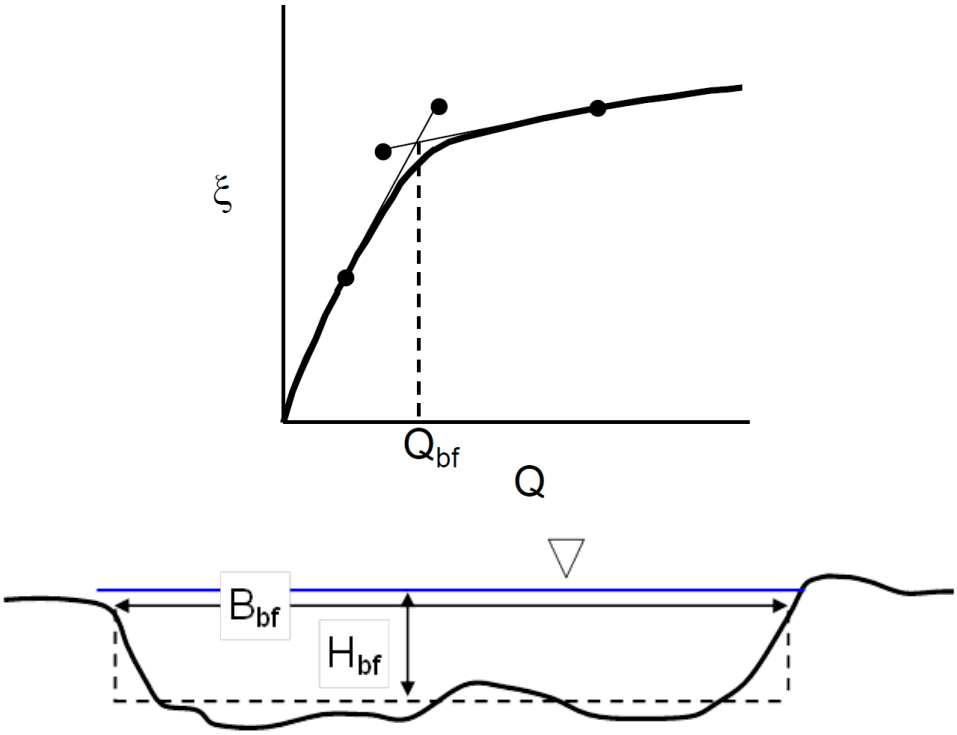

Figure 43: Bankfull conditions: (a) Schematic of the definition for bankfull discharge in terms of the stage discharge relation (García, 2008, chap. 3), (b) Definition of bankfull geometry.

The most important geometrical characteristics of rivers are their bankfull width and depth. Due to the large variation of rivers in nature, these characteristics seems random. However, rivers establish their bankfull width and depth through the co-evolution of the river channel and the floodplain (Parker, 2012). Statistically, they follow some general rules. The bankfull cross-sections and floodplains of alluvial rivers are created by the coupled interactions of flows and sediment movements. The river bed and lower banks are constructed from bed material load. The middle and upper banks are usually constructed predominantly out of wash load. In some cases, some bed material load can also be found in the floodplains. As the river avulses and shifts, this wash load material is spread out

across the floodplain.

As in another other discipline, there are many avenues to find the governing laws behind natural phenomena. For example, large part of this project tries to use statistical analysis to find possible significant relationships. On the other hand, dimensional/dimensionless analysis is another common and important method. Put into dimensionless forms, existing data have shown clear clustering for both sand and gravel rivers. The dimensionless parameters contain the bankfull width, depth, discharge, bed material, etc.

7.1.1 Bankfull conditions

A common definition of bankfull condition was provided by Dunne and Leopold (1978). They defined that bankfull condition at which the channel maintenance is the most effective, moving sediment, forming or removing bars, forming or changing bends and meanders, and generally doing work that results in the average morphologic characteristics of channels. Bankfull discharge is usually defined as the discharge with a recurrence interval of 1 to 2 years. Note the relationship between recurrence interval and probability of exceedence:

$$\text{Recurrence Interval} = \frac{1}{\text{Probability of exceedence}} \tag{34}$$

There are several ways to determine the bankfull discharge based on what information is available. The first is to use the flow duration curve and the find the corresponding discharge. In this research, bankfull discharge was defined with a recurrence interval of 1.5 years, corresponding to a 66.7% probability of exceedence. This discharge can also be cross-checked using the discharge-stage curve. As previously mentioned, there should be a significant slope change at bankfull condition in this curve reflecting the overspilling of water into much wider flood plain.

The second method, which is also of great importance, is through field observation. Bankfull indicators were observed in the field using guidelines set out by Dunne and Leopold (1978). Bankfull indicators include: (1) topographic break from vertical bank to flat floodplain, (2) topographic break from steep slope to gentle slope, (3) change in vegetation from bare to grass, moss to grass, grass to sage, grass to trees, or from no trees to trees, (4) textural change of depositional sediment, (5) elevation below which no fine debris (needles, leaves, cones, seeds) occurs, and (6) textural change of matrix material between cobbles or rocks. Bankfull width was measured at the bankfull elevation and bankfull depth was then determined as an average of the measured depths across the stream channel.

7.1.2 Example of bankfull discharge calculation

As an example, the flow duration curve and the discharge-topwidth curve for USGS station 8195000 (Frio River at Concan, TX) are plotted in Figure 44. The bankfull discharge was

determined to be around 50 cfs. This value was also confirmed by the discharge-topwidth curve.

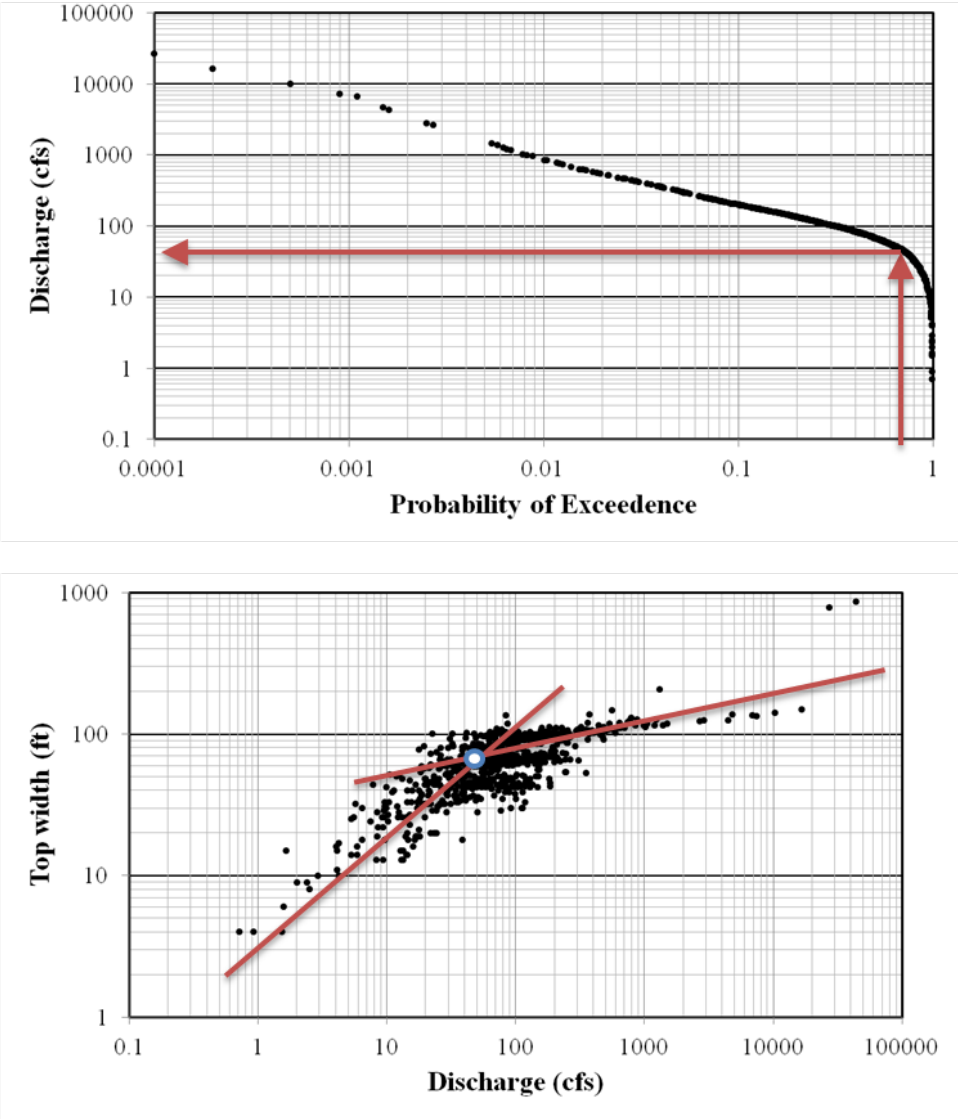

Figure 44: Example calculation of bankfull discharge for USGA gaging station 8195000 (Frio River at Concan, TX): (a) Flow duration curve, (b) Discharge as a function of top width. For this station, the bankfull discharge was determined to be about 50 cfs.

7.1.3 Definitions

Alluvial rivers can be broadly classified into two categories: sand-bed streams whose median size D_{50} of surface sediment is between 0.0625 mm and 2 mm, and gravel-bed streams whose

D_{50} falls in the range 2 mm and 256 mm. For a specific reach of a river at bankfull condition, the following parameters can be identified, the bankfull discharge Q_{bf}, the bankfull depth H_{bf}, the bankfull width B_{bf}, and the bankfull velocity U_{bf}. Using normal flow condition in open channels, the characteristic bankfull bottom shear stress τ_{bf} can be estimated as

$$\tau_{bf} = \rho g H_{bf} S \tag{35}$$

where ρ is the density of water, g is gravitational acceleration, S is the channel bed slope.

In the study of natural phenomena, a dimensional parameter such as those defined previously for bankfull conditions varies greatly from river to river. An appropriately defined dimensionless number (by grouping dimensional parameters together) can reveal more universal characteristics. The following dimensionless parameters are defined for alluvial streams at bankfull condition

$$\hat{Q} = \frac{Q_{bf}}{\sqrt{gD_{50}}{D_{50}}^2}, \quad \hat{B} = \frac{B_{bf}}{D_{50}}, \quad \hat{H} = \frac{H_{bf}}{D_{50}}, Fr_{bf} = \frac{U_{bf}}{\sqrt{gH_{bf}}} \tag{36}$$

$$\tau^*_{bf,50} = \frac{\tau_{bf}}{\rho g R D_{50}}, Cz_{bf} = \frac{U_{bf}}{\sqrt{gH_{bf}S}}, \quad R_{50} = \frac{\sqrt{gRD_{50}}D_{50}}{\nu}, \quad R = \frac{\rho_s}{\rho} - 1 \tag{37}$$

where ρ_s is the density of sediment, $R = (\rho_s/\rho - 1)$ is the sediment submerged specific gravity which has a value of 1.65 for natural sediment, and nu is the kinematic viscosity of water.

Of these dimensionless parameters, the bankfull Froude number Fr_{bf} characterizes a ratio of inertia to gravity force.

7.2 Analysis for Texas Rivers

While each river has its own unique characteristics, alluvial rivers show a considerable degree of commonality. This commonality is apparent in terms of dimensionless plots of bankfull characteristics. As statistical analysis (another part of TXDOT 0-6654), dimensionless analysis aims to find some general laws that rivers have to obey. We can use these general laws to check our hydraulic modeling results.

In this section, we will first present the data sources which are the foundation for the analysis. Then the results for Texas rivers will be presented and discussed.

7.2.1 Data sources

From the USGS streamflow measurement database, the following flow characteristics are available:

- Discharge,
- Flow area, top width, Fr, velocity, etc.,
- Flow duration curve.

Information on sediment sizes and distributions is collected through:

- Literature (as part of TXDOT 0-6724),
- Liu group at UTSA (as part of TXDOT 0-6724),
- Strom group at UH.

Channel slopes are not part of the original USGS flow database. They were calculated by USGS and UTSA from DEM data. More details on the methodology and procedures to calculate the slopes can be found in the technical memorandum corresponding to these tasks.

To compare with rivers in other parts of the world, some widely cited data sets in the literature are included for comparison. These data sets have been used in for example Parker et al. (2007) and Parker (2012) to derive universal relations describing bankfull hydraulic geometries. These data sets include gravel-bed rivers with D_{s50} (surface median size) ranging from 27 mm to 157.5 mm from Britain (Charlton et al., 1978), Alberta, Canada (Kellerhals et al., 1972) and Idaho, USA (Parker et al., 2003). For sand-bed rivers with $D_{s50} < 0.5$ mm, including single-thread streams and multiple-thread streams, data from Church and Rood (1983) and Parker et al. (1998) are used. In the following, we will use "Gravel British", "Gravel Alberta", "Gravel Idaho", "Sand Multiple", and "Sand Single" to distinguish these data sets. The names are self-explanatory.

7.2.2 Results for Texas Rivers

Based on the collected data, the dimensionless parameters defined in Equations 36 and 37 are calculated and plots are generated (Figures 45 to 51).

In Figure 45, the dimensionless discharge $\widehat{Q}$ is plotted against dimensionless width $\widehat{B}$. Both types of river types (sand and gravel) plots in a consistent way.

In Figure 46, the dimensionless discharge $\widehat{Q}$ is plotted against dimensionless depth $\widehat{H}$. Again, both types of river types (sand and gravel) plots in a consistent way.

In Figure 47, the dimensionless discharge $\widehat{Q}$ is plotted against the channel bed slope S. The relation for slope versus dimensionless bankfull discharge shows general consistency but much more scatter. This probably reflects the fact that much more time is required to change a river's bed slope than its width or depth Parker (2012). Plus, the definition of slope and the method to calculate are yet to be refined.

In Figure 48, the channel bed slope S is plotted against the dimensionless discharge $\widehat{Q}$.

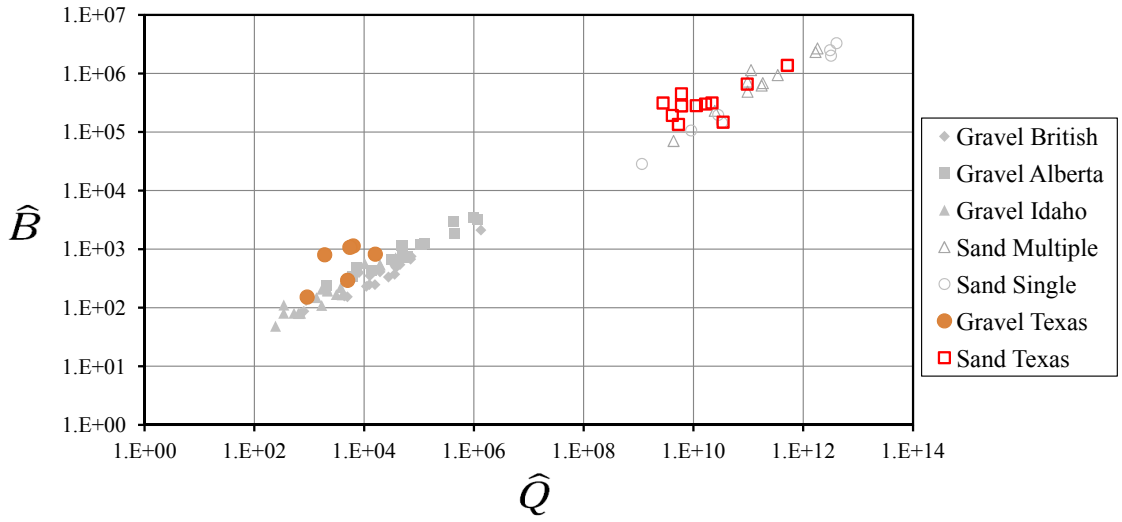

Figure 45: Dimensionless bankfull discharge $\widehat{Q}$ vs. dimensionless top width $\widehat{B}$

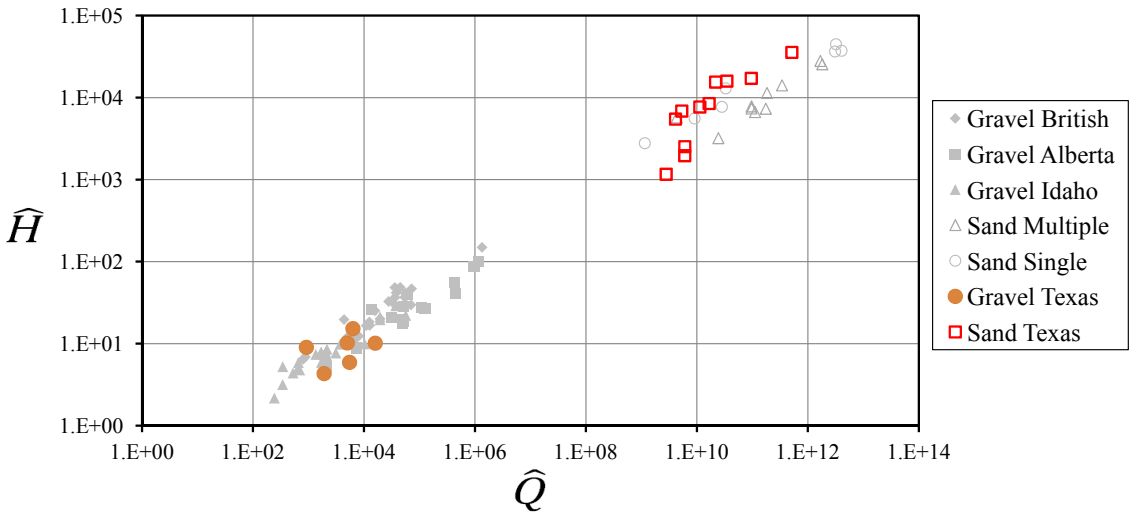

Figure 46: Dimensionless bankfull discharge $\widehat{Q}$ vs. dimensionless depth $\widehat{H}$

Sand-bed streams generally have lower bankfull Froude numbers. Majority of the streams are in the subcritical range (Fr_{bf} ¡ 1) at bankfull flow. Texas sand rivers have relatively lower Fr_{bf} than others. That might be due to the method to calculate of depth.

In Figure 49, the channel bed slope S is plotted against the dimensionless Chezy friction coefficient Cz_{bf}. The resistance coefficient is generally larger in the sand-bed streams. This

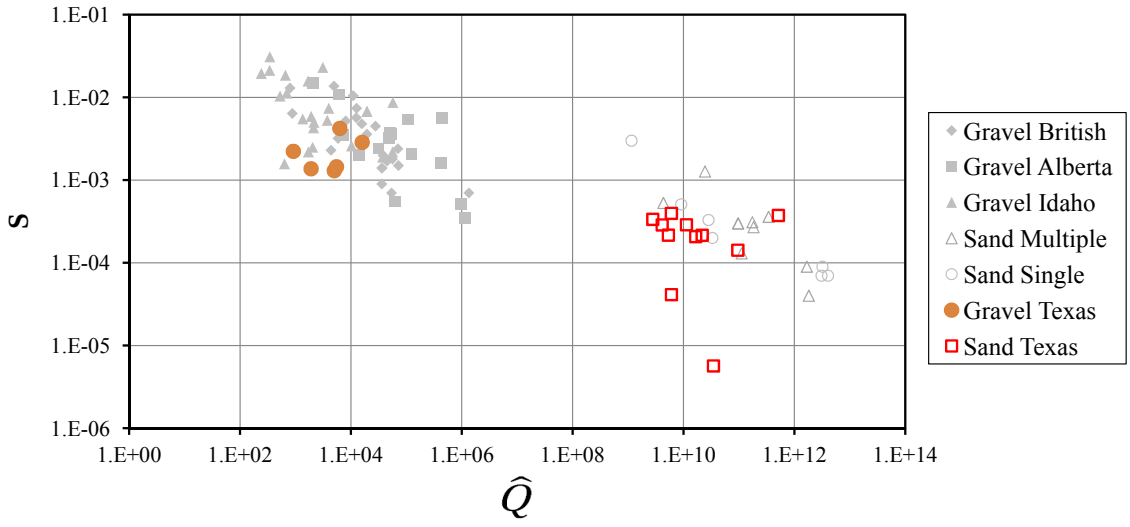

Figure 47: Dimensionless bankfull discharge $\widehat{Q}$ vs. channel slope S

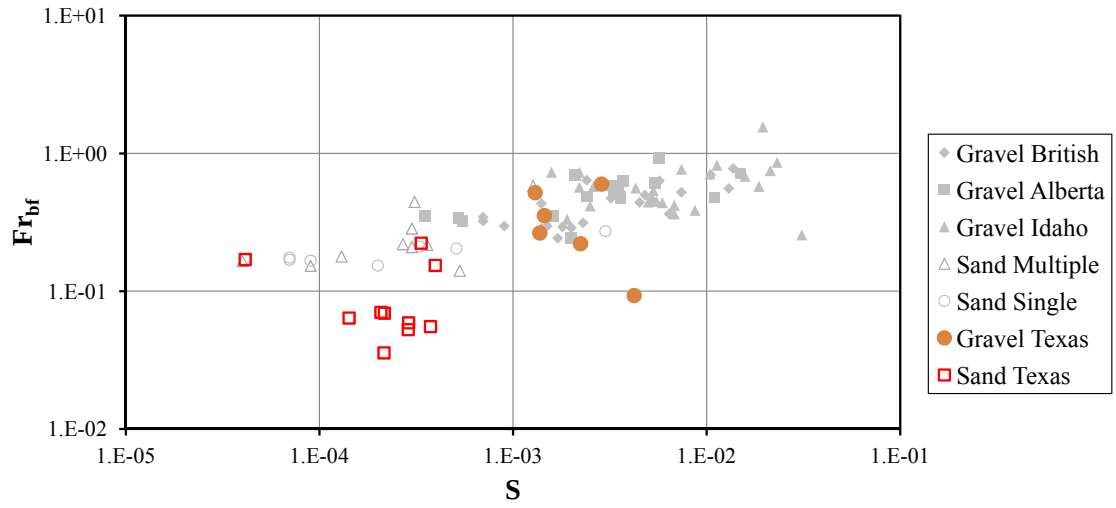

Figure 48: Channel slope S vs. dimensionless bankfull Froude number Fr_{bf}

probably reflects the effect of dunes Parker (2012).

In Figure 50, the dimensionless depth $\widehat{H}$ is plotted against the dimensionless Chezy friction coefficient Cz_{bf}. The sand-bed and gravel-bed sets plot in different regions, largely because in sand-bed streams resistance is more dependent on bedform characteristics. Texas sand-bed rivers add more scatter to the plot.

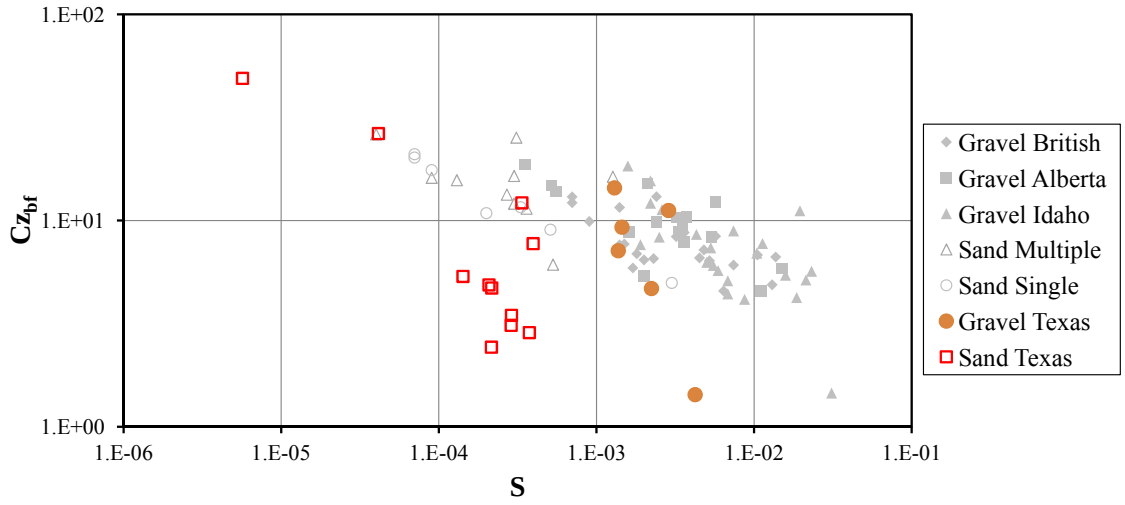

Figure 49: Channel slope S vs. dimensionless bankfull Chezy resistance coefficient Cz_{bf}

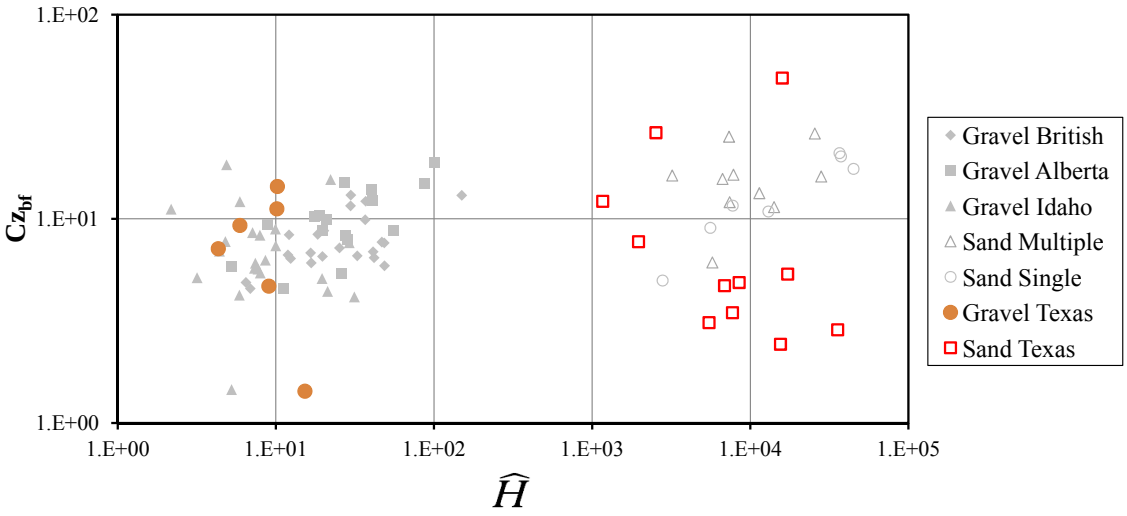

Figure 50: Dimensionless bankfull depth $\widehat{H}$ vs. dimensionless bankfull Chezy resistance coefficient Cz_{bf}

In Figure 51, the dimensionless discharge $\widehat{Q}$ is plotted against the bankfull Shields number τ^*_{bf50}. Gravel-bed streams maintain a bankfull Shields stress that is loosely about 0.05. Sand-bed streams maintain a bankfull Shields stress that is loosely about 1.9. The distinction between gravel and sand streams are clear. However, Texas sand-bed rivers add more scatter to the plot.

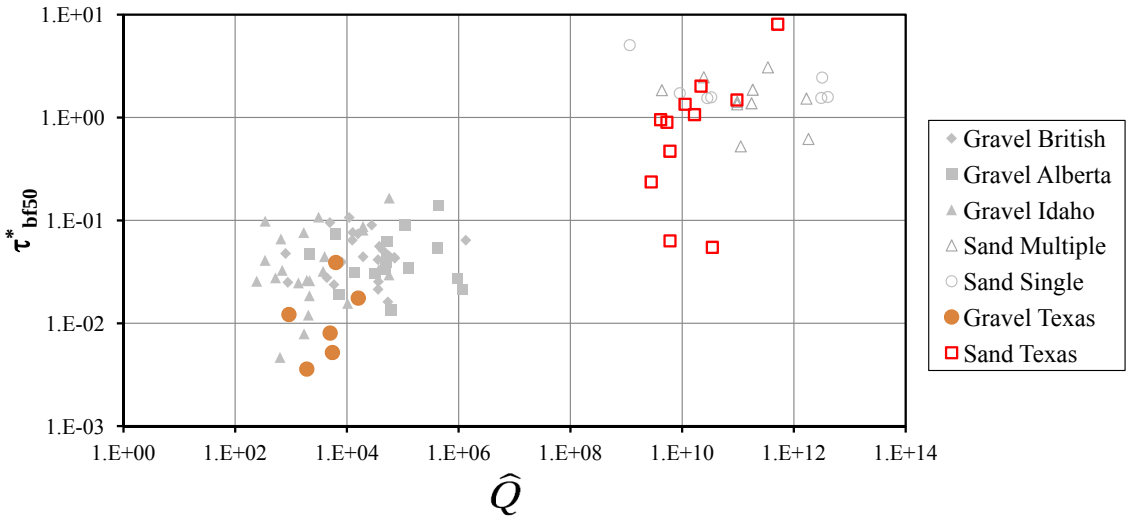

Figure 51: Dimensionless bankfull discharge $\widehat{Q}$ vs. banfull Shields number $\tau^*_{bf,50}$

8 Examples of product use

In this section, examples are given for how each of the three methods developed in sections 5, 6, and 7 can be used to make estimates of discharge or velocity, or to check how often a particular discharge has been observed at a given location.

8.1 Method 1: Regression Analysis

The following two subsections outline two different ways that the developed regression analysis (method 1 of section 5) can be used. The first example shows how the method can be used to estimate discharge at a site using the geographic location of the site of interest and the measured peak-flow top width and average depth (section 8.1.1). The second example shows how the methods could be used to check whether or not the velocity calculated by a hydraulic model is, "way off" (section 8.1.2).

8.1.1 Post-Event Discharge Estimation

Two example applications of the QGAM and VGAM are presented in this section. Suppose that a direct-runoff event occurred and an analyst is interested in estimating the Q_p for a particular stream located at about 31.5° north and −98.5° west. A postdirect-runoff event survey measures that the top width of the peak water surface at about 100 m and the average depth is estimated as 4.5 m. The estimated cross section area is thus 450 m^2. The P for the location is about 744 mm PRISM Climate Group (2010), and the OmegaEM parameter in Figure 24 for the location is about −0.106.

The smooth function $f_5(l, k)$ of the location for QGAM is judged to be about 0.15 from Figure 26 using interpretation and interpolation of the smooth function lines (black lines) and the lower and upper standard error lines (green and red lines, respectively) as available. The smooth function $f_6(P)$ of P for QGAM is about −0.02 from Figure 27. The Q_p can now be readily computed by variable substitution in Equation 26:

$$\begin{aligned} \log(Q_p) &= -0.2896 + 1.269\log(450) - 0.2247\log(100) \\ &\quad +0.2865(-0.106) + 0.15 - 0.02 \qquad (38) \\ \log(Q_p) &= 2.728 \qquad (39) \\ Q_p &= 535 \text{ m}^3/\text{s} \qquad (40) \end{aligned}$$

For this estimate of Q_p, the $\overline{V}$ is:

$$\overline{V} = \frac{535 \text{ [m}^3/\text{s]}}{450 \text{ [m}^2\text{]}} = 1.19 \text{ m/s} \qquad (41)$$

The VGAM provides an alternative estimate of V for a Q of 535 m^3/s. The smooth function $f_9(l, k)$ of the location for VGAM is judged to be about 0.06 from Figure 30 using interpretation and interpolation of the smooth function lines (black lines) and the lower and upper standard error lines (green and red lines, respectively) as available. The smooth function $f_{10}(P)$ of P for VGAM is about -0.02 from Figure 31. The V can be readily computed by variable substitution in Equation 27:

$$\begin{aligned} V^{1/5} &= 0.9758 + 0.1588\log(535) - 0.1820\log(100) \\ &\quad +0.0854(-0.106) + 0.06 - 0.02 \qquad (42) \\ V^{1/5} &= 1.076 \qquad (43) \\ V &= 1.44 \text{ m/s} \qquad (44) \end{aligned}$$

Lastly, the authors observe that the two estimates of V (1.19 m/s versus 1.44 m/s) are seemingly consistent with each other. Consistency between either a computed (from known or design discharge and known cross sectional area) or modeled $\overline{V}$ and V predicted by VGAM is the subject of the next section.

8.1.2 Review of Mean Velocity from a Hydraulic Model

The previous example application guides a user in computing Q given cross-section properties and other characteristics. The focus of the computations was on QGAM. For another example application, the focus is on VGAM. Suppose for the same location that an analyst has a design discharge Q_T of 800 m^3/s for a 0.02 AEP or recurrence interval of $T = 50$ years, and a hydraulic model predicts a B of 100 m and an A of 450 m^2 as used in the previous example for simplicity. The hydraulic model is thus predicting a computed $\overline{V}$ of 1.78 m/s. The VGAM can be used to independently evaluate the $\overline{V}$ from the hydraulic model. The V estimate from VGAM is 1.44 m/s as computed in the previous example.

Wood (2009) provides the `predict.gam` function (Wood, 2006, p. 243), which is designed for use in R. This function computes standard errors of a prediction for a GAM using a Bayesian posterior covariance matrix. However, without a digital presentation of the GAM object from R as well as R running on a host computer, the computations of standard error are tedious and error prone for desktop application by anticipated end users. A convenient means for end-user implementation to only approximate the distribution of a prediction from VGAM (or QGAM by association) thus is needed.

The prediction percentile for a multi-linear regression (Helsel and Hirsch, 2002, p. 295–322) can be computed by:

$$y(\Pi/100) = y_o + s \times t_{[\Pi/100, n-p]}\sqrt{1 + h_o} \qquad (45)$$

where $y(\Pi/100)$ is the predicted response for the Π percentile, y_o is a prediction from the regression model, n is the sample size, p is the number of parameters, $t_{[\Pi/100,n-p]}$ is the quantile distribution function (qdf) of the t-distribution, s is the residual standard error, and h_o is the leverage of the prediction. The sample size for VGAM is large (n =17,751, Figure 29) and the parameter count is small (p =7, Figure 29); as a result, the qdf of the standard normal distribution $\Phi(F)$ for nonexeedance probability F can be substituted for the t-distribution.

Although the specific leverage or its equivalence of a GAM for ungaged locations is extremely difficult to represent or approximate, the average leverage of a conventional multilinear regression model is p/n. The average leverage for VGAM is effectively zero because the ratio 7/17,751 is small. Therefore, h_o is approximately zero because of the enormous degrees of freedom and thus $\sqrt{1+h_o} \approx 1$. The residual standard error is $s = 0.0630$ (Figure 29). The prediction percentile of the 1.78 m/s velocity can thus be loosely approximated, recalling use of the fifth-root transformation and Equation 45, by:

$$y(\Pi/100) \approx 1.78^{1/5} \approx 1.44^{1/5} + 0.0630\ \Phi(\Pi/100) \quad (46)$$

$$\Phi(\Pi/100) \approx 0.739 \quad (47)$$

$$\Pi/100 \approx \phi(0.739) \approx 0.77 \quad (48)$$

where $\phi(x)$ is the cumulative distribution function of the standard normal distribution for value x. The results show that $\overline{V}$ of the hydraulic model is at the 77th percentile. The project reviewer would naturally conclude that the $\overline{V}$ of the hydraulic model is consistent with VGAM.

To further demonstrate VGAM application, suppose that an analyst wants to apply for the same location a design Q_T of 2,100 m^3/s. Suppose also that the analyst has run or is reviewing a hypothetical hydraulic model predicting B of 100 m and A of 450 m^2 (as used in previous examples for simplicity). The hydraulic model is thus predicting a computed $\overline{V}$ of 4.67 m/s. The prediction percentile for 4.67 m/s can be estimated, recalling use of the fifth-root transformation, by:

$$y(\Pi/100) \approx 4.67^{1/5} \approx 1.44^{1/5} + 0.0630\ \Phi(\Pi/100) \quad (49)$$

$$\Phi(\Pi/100) \approx 4.53 \quad (50)$$

$$\Pi/100 \approx \phi(4.53) > 0.999 \quad (51)$$

The results show that the hydraulically modeled V is in excess of the 99.9th percentile of VGAM. The analyst running or reviewing the hydraulic model would naturally conclude that the $\overline{V}$ is inconsistent with VGAM and by extension is inconsistent with more than 17,700 measurements of high-magnitude discharge in Texas. The apparent absence of congruence between the two V values could be a sign that enhancements to the reliability

of the hydraulic model through changes in model assumptions, parameter values, or select cross-section representations might be possible.

The previous computations considered a large hydraulically modeled $\overline{V}$. The problem could also be in the opposite direction. Suppose for the same location that the design Q_T is 210 m^3/s and again a hydraulic model is predicting a B of 100 m and an A of 450 m^2. The hydraulic model is thus predicting a computed $\overline{V}$ of 0.467 m/s. The prediction percentile for 0.467 m/s can be estimated, recalling use of the fifth-root transformation, by:

$$y(\Pi/100) \approx 0.467^{1/5} \approx 1.44^{1/5} + 0.0630\ \Phi(\Pi/100) \quad (52)$$

$$\Phi(\Pi/100) \approx -3.44 \quad (53)$$

$$\Pi/100 \approx \phi(-3.44) < 0.0003 \quad (54)$$

The results show that the hydraulically modeled $\overline{V}$ is less than the 0.03th percentile. Again, the analyst running or reviewing the hydraulic model would naturally conclude that the $\overline{V}$ is inconsistent with VGAM and by extension is inconsistent with more than 17,700 measurements of high-magnitude discharge in Texas. The apparent absence of congruence between the two V values could be a sign that enhancements to the reliability of the hydraulic model through changes in model assumptions, parameter values, or select cross-section representations might be possible.

The procedures shown to compute the distribution of a prediction from VGAM in this section are also applicable by association to the distribution of a prediction from QGAM although example computations are not shown in this paper.

8.2 Method 2: Empirical Distributions

The primary information that is gained from the various empirical distributions of the primary and ancillary parameters is the non-exceedance value of the parameter of interest under conditional constraint. Several different examples of how to generate and use the empirical distributions can be found throughout section 6. For example, one could ask what the non-exceedance value is for a velocity of 5 ft/s for all of Texas under any measured flow condition. This information could be obtained from the database using the R code provided in section 6.2.2 to extract graphical and tabular information. In this particular case, $V = 5$ ft/s has approximately a 0.99 non-exceedance probability (only exceeded 0.1% of the time). Any number of constraints such as only pulling from storm flows (flows greater than the 90th percentile) or constraining the samples by drainage basin area, can be used to construct the non-exceedance curves and tables.

The distributions for the ancillary parameters can be generated and used in much the same way. Section 6.4 explains how to use R to calculate the ancillary parameters and then pull conditioned or unconditioned distribution distributions that can then be used to check the

non-exceedance of the value of interest. Definitions of possible ancillary parameters that can be calculated with database are given in section 6.5. Alternatively, the figures for Fr, n, f, C, and Ω in section 6.5.1 can be used to quickly obtain non-exceedance values for storm and mean flow without any regional conditioning. For example, from figure 38, a Froude number of $Fr = 0.5$ has a 0.93 non-exceedance probability for storm flows (flow greater than the 90th percentile) and a 0.975 no non-exceedance probability for mean flows (flows between the 40th and 60th percentiles).

8.3 Method 3: Hydraulic Geometry

This section outlines the general procedure to follow for using the dimensionless hydraulic geometry relationships outlined in section 7 (Method 3). Assuming that Texas rivers follow the nearly universal hydraulic geometry functionality for alluvial rivers, hydraulic models of Texas rivers should yield results that reasonably follow this general functionality. If the modeling results show evidence of strong disagreement with other Texas streams in non-dimensional space, further investigation of the model is warranted.

In order to apply the results to check any hydraulic modeling output, the following data are needed:

- Flow duration curve or other alternative to determine the bankfull discharge,
- Channel geometry,
- Channel slope estimation,
- Sediment characteristics (sizes and distribution).

Of the needed inputs, the sediment characteristics at the site of interest will likely be the most difficult to obtain. For this project, we have simply surveyed the literature for sediment data on Texas rivers, and have also collected several sediment samples as part of another TxDOT project. In real engineering design, if budget and time allow, sediment samples should be taken at the design location and they should be analyzed following standard sieve analysis procedure.

A hydraulic model (e.g., HEC-RAS) can be checked by the following steps:

- Use the bankfull discharge as the inflow,
- Calculate the corresponding water depth and width,
- Calculate the dimensionless discharge, width, depth, etc.,
- Plot the result in the dimensionless figures shown before,
- If the predicted data points collapse well with the database, the hydraulic model passes the check. Otherwise, it raises a red flag and the modeler needs to do further investigation.

As an example, Figure 52 demonstrates two situations of modeled bankfull discharge and top width in dimensionless space. Non-dimensional discharge and top width from runs plotted with green markers would be considered "reasonable" and giving a passing mark. Whereas non-dimensional values from runs plotted with red markers would be considered "un-reasonable" and given failing marks since the data points fall well outside of the trend of measured river sections. Models that don't follow the measured trends in non-dimensional space should be flagged and checked for errors.

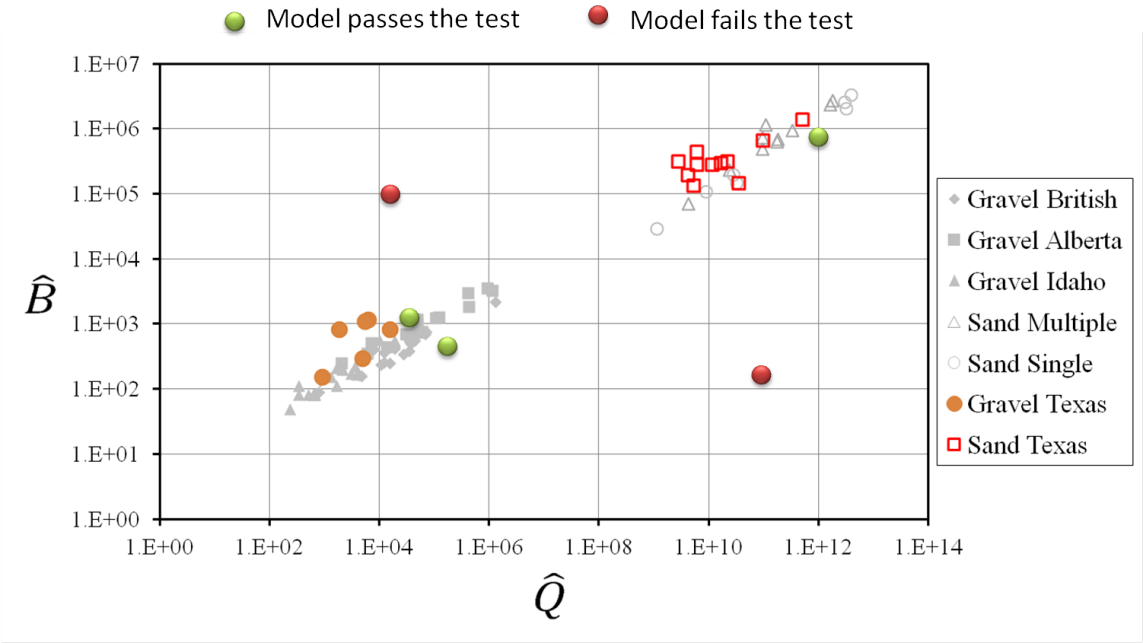

Figure 52: Application demonstration using the dimensionless plots

9 Examples of product use on a HEC-RAS model of Guadalupe Arroyo

In this section, each of the three methods developed in sections 5, 6, and 7 is used to evaluate a HEC-RAS hydraulic model of an un-gaged location in Texas. The location is at the bridge on Guadalupe Arroyo on US 62/180; an ephemeral stream in the El Paso District. This location, in fact, stimulated the initial investigations that led to this research and accompanying toolkit for model evaluation.

The area has few gaging stations and consequently few observation sets to evaluate hydrologic and hydraulic models; exactly the situation envisioned by the team for application of empirical velocity distributions.

9.1 The HEC-RAS model

A HEC-RAS model was constructed using cross sections approximated from Huitt-Zollars (2008). The model in this research report was constructed by the research supervisor. The model approximates the cross section information and some other values from the source report but is otherwise unrelated. The results herein are for illustrative purposes only.

Figure 53 is a Google Earth aerial photograph from 2009. The flow in the stream, when it occurs, is from right to left in the image. The streambed near the bridge at the time of the photograph is about 100 feet wide, and the banks that appear in the imagery are about 8 feet above the bottom of the streambed. The channel bed slope is about 0.013. Using Rosgen's riffle-to-riffle concept Rosgen (1996) as a forensic approach, the estimated friction slope is about 0.01.

The drainage area contributing to the stream at the bridge was estimated to be about 40 square miles, much of that area is steep terrain near Guadalupe Peak. The Texas regional regression equations Asquith et al. (1996) applied to the location return a discharge of about 2700 cfs for a 50-percent Annual Exceedance Probability.

Figure 54 is a screen capture of a HEC-RAS simulation for steady flow of 2700 cfs. The simulation is for a portion of the channel 300 feet upstream and downstream of the bridge location. The cross section spacing is 25 feet[6]. The upstream and downstream boundary conditions were set as normal depth using the friction slope estimate of 0.01. Program defaults were used. The left and right bank locations were approximated from the source document and manually entered into the program. The resistance terms were set at n=0.05 for the overbank, and n=0.04 for the channel as in the source document.

The results are initially unremarkable, the program ran, produced warnings regarding

[6]This spacing is similar to that of the source document, although no attempt to preserve cross sectional skew was made by the researcher.

Figure 53: Google Earth aerial photograph of US 62/180 bridge over Guadalupe Arroyo, Culberson County, Texas

conveyance ratios at two sections (where the two high velocities are reported), reasonable diligence was taken to enter the cross sections, select an appropriate discharge, and select appropriate boundary conditions. However the two excursions into supercritical flow at sections 14 and 9 are interesting as is the flow velocity being in excess of 6 ft/second everywhere and the two velocity spikes of about 11 feet per second. The engineer at this point might want to know how common are such velocities in a mostly natural channel (there is considerable human modification of this channel, but it is comprised of native materials at the time of the photographs).

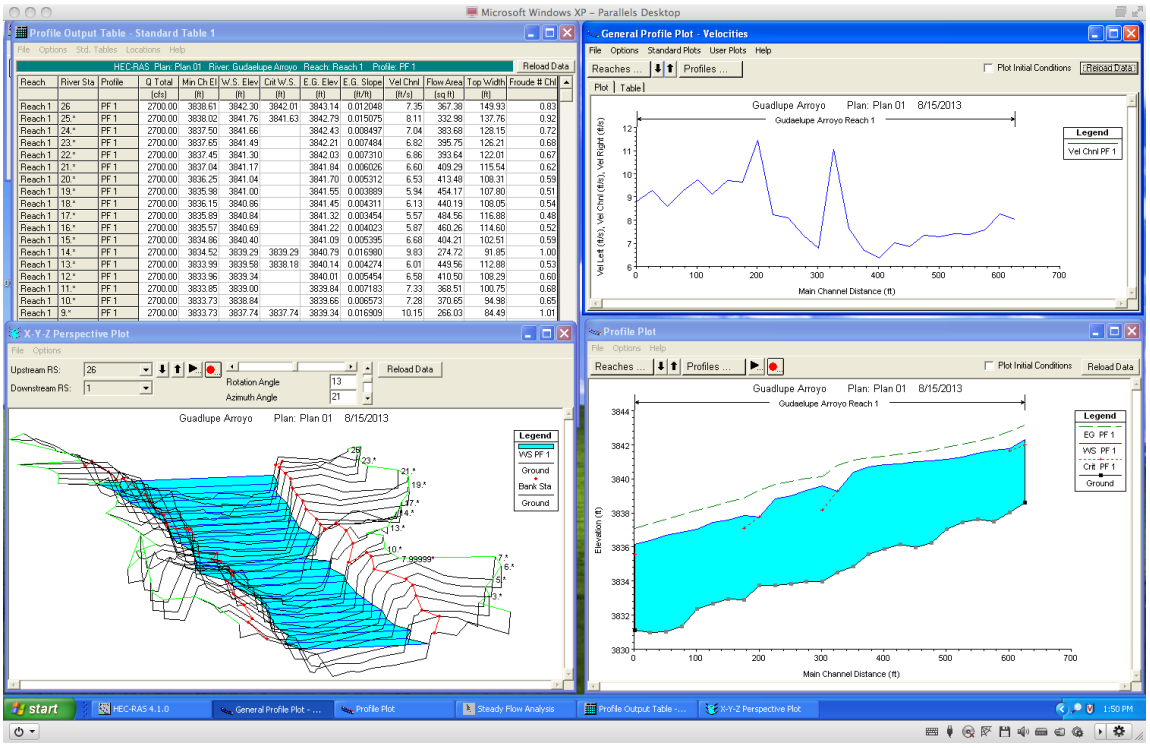

Figure 54: HEC-RAS results from example conditions

9.2 Method 1: Regression Analysis

The regression approach is illustrated first. Examining the upper left hand panel of Figure 54 the supercritical flow sections have cross sectional flow are of about 270 ft^2 and topwidth of about 85 ft.

The regression equations use SI units, so the flow area and topwidth are converted to their SI values; $270ft^2 \approx 25m^2$ and $85ft \approx 25.9m$. OmegaEM parameter in Figure 24 for the location is about -0.102. The smooth function $f_5(l,k)$ of the location for QGAM is judged to be about 0.25 from Figure 26. The smooth function $f_6(P)$ of P for QGAM is about 0.2 from Figure 27 using the upper standard error line in the figure. Thus the parameters for the regression equation are $\Omega = -0.102$,$f_5 = 0.2$,$f_6 = 0.05$. Substituting these numerical values results in the estimate of discharge for the location based upon the hydraulic variables of topwidth and flow area.

$$\begin{aligned} \log(Q_p) &= -0.2896 + 1.269\log(25) - 0.2247\log(25.9) \\ &\quad +0.2865(-0.102) + 0.25 + 0.2 \qquad (55) \\ \log(Q_p) &= 1.58 \qquad (56) \\ Q_p &= 38 \text{ m}^3/\text{s} = 1371 \text{ ft}^3/\text{s} \qquad (57) \end{aligned}$$

The reported cross section mean velocity in the HEC-RAS model at the supercritical cross sections are in excess of 10 ft/s. The VGAM function can be used with either the discharge computed by QGAM or the model discharge to assess the model results. As with the QGAM equation, the VGAM equation uses SI units, so the flows are converted; $1371cfs \approx 38m^3/s$ and $2700cfs \approx 76m^3/s$.

The smooth function $f_9(l,k)$ of the location for VGAM is judged to be about 0.2 from Figure 26. The smooth function $f_10(P)$ of P for VGAM is about 0.025 from Figure 27 using the upper standard error line in the figure. $\Omega = -0.102$,$f_9 = 0.2$, $f_{10} = 0.05$. Substituting these numerical values results in the estimate of velocity for the location based upon the hydraulic variables of topwidth and flow area.

$$\begin{aligned} V^{1/5} &= 0.9758 + 0.1588\log(38) - 0.1820\log(25.9) \\ &\quad +0.0854(-0.102) + 0.2 + 0.025 \qquad (58) \\ V^{1/5} &= 1.187 \qquad (59) \\ V &= 2.35 \text{ m/s} \approx 7.7 \text{ ft/s} \qquad (60) \end{aligned}$$

$$\begin{aligned} V^{1/5} &= 0.9758 + 0.1588\log(76) - 0.1820\log(25.9) \\ &\quad +0.0854(-0.102) + 0.2 + 0.025 \qquad (61) \\ V^{1/5} &= 1.23 \qquad (62) \\ V &= 2.86 \text{ m}^3/\text{s} \approx 9.4 \text{ ft/s} \qquad (63) \end{aligned}$$

Thus the expected discharge for the hydraulic conditions in the model is considerably lower than 2700 cfs based on the QGAM model, the expected velocity using either the QGAM estimate or the original discharge is 7.7 ft/s and 9.4 ft/s, respectively – still smaller than the nearly 11 ft/s reported in the computer program. The interpretation of these findings are threefold. First, the model discharge seems large, even if the hydraulic cross section and topwidth of the upstream or downstream sections are applied the QGAM estimated discharge is substantially lower than the 2700 cfs. The velocity in the supercritical section in the HEC-RAS model is higher than the VGAM expected velocity for the same discharge at the same topwidth, although only by about 10 percent. Secondly, the QGAM finding suggests that to produce the 2700 cfs discharge the flow area needs to be greater, which can only be achieved by deeper flow (at least in the constructed cross section). Lastly these differences suggest that the contributing area for the hydrology estimates is smaller than the 40 square miles. Herrmann and Cleveland (2010) on a field reconnaissance determined that the bankfull cross sectional flow area from a rapid survey about a kilometer upstream of the location where evidence of human earth work is negligible is about 133 ft^2 and the topwidth is about 50 ft. QGAM estimates that the discharge for these conditions is about 630 cfs, with a VGAM estimate of velocity of 7.4 ft/s. This estimate of discharge is substantially smaller than the 2700 cfs initially used.

An added use of the QGAM and VGAM models are an ability to rapidly estimate what the discharge and velocity might be at any cross section. In the example presented here the bridge is around section 12 and 13. The distance from the low chord of the bridge to the channel bottom is about 10 feet. Using the cross sectional survey the flow area represented is about 900 ft^2 with a topwidth of about 140 ft. The QGAM flow estimate for this condition is 5600 cfs, with an associated mean section velocity from VGAM of 9.8 ft/s.

9.3 Method 2: Empirical Distributions

The Texas database and conditional distribution approach lets the engineer ask, for Texas for drainage areas up to the 40 square mile range, with individual flow duration curve values greater than the 90-th percentile, the following questions:

1. Is 2700 cfs an accessible discharge?
2. What is the anticipated water velocity?

3. What is a typical topwidth?

9.3.1 Using the Conditional Distributions

The conditional distributions can be used directly, but the conditioning on a particular contributing drainage area value is severe and often the database has few if any values to return. An alternative in this example is to instead build tabular representations to answer the various questions.

Listing 14 is the R script used to generate a table of drainage areas and associated mean, median, and maximum discharges from the Texas database. The drainage area for the HEC-RAS model is 40 square miles and the discharge used in the HEC-RAS simulation is 2700 cfs. Examining the table in the listing, the value of 2700 cfs is smaller than the maximum observed value in the database for drainage areas up (and greater than) 40 square miles. However, the 2700 cfs is supposed to represent a median discharge, and the value of 2700 cfs is considerably greater than the mean or median value of discharge from the database for contributing drainage areas of 40 square miles and greater.

Listing 14: R code for discharge conditioned on contributing drainage area

```
> # HEC-RAS EXAMPLE
> # Query the database, recover Discharge for different CDA at FDC > 90\%
>   CQ10 <- DB[DB$CDA <   10 & DB$FDC >= 0.9  ,]$Q
>   CQ20 <- DB[DB$CDA <   20 & DB$FDC >= 0.9  ,]$Q
>   CQ40 <- DB[DB$CDA <   40 & DB$FDC >= 0.9  ,]$Q
>   CQ80 <- DB[DB$CDA <   80 & DB$FDC >= 0.9  ,]$Q
>  CQ160 <- DB[DB$CDA <  160 & DB$FDC >= 0.9  ,]$Q
>  CQ320 <- DB[DB$CDA <  320 & DB$FDC >= 0.9  ,]$Q
>  CQ640 <- DB[DB$CDA <  640 & DB$FDC >= 0.9  ,]$Q
> CQ1280 <- DB[DB$CDA < 1280 & DB$FDC >= 0.9  ,]$Q
> CQ2560 <- DB[DB$CDA < 2560 & DB$FDC >= 0.9  ,]$Q
> # Build vectors for tabular output
> CDAVAL  <-c(10,20,40,80,160,320,640,1280,2560)
> QMEAN   <-c(mean(CQ10),mean(CQ20),mean(CQ40),mean(CQ80),mean(CQ160),mean(CQ320),mean(
    CQ640),mean(CQ1280),mean(CQ2560))
> QMEDIAN <-c(median(CQ10),median(CQ20),median(CQ40),median(CQ80),median(CQ160),median(
    CQ320),median(CQ640),median(CQ1280),median(CQ2560))
> QMAX    <-c(max(CQ10),max(CQ20),max(CQ40),max(CQ80),max(CQ160),max(CQ320),max(CQ640),max
    (CQ1280),max(CQ2560))
> cbind(CDAVAL,QMEAN,QMEDIAN,QMAX)
      CDAVAL     QMEAN QMEDIAN   QMAX
 [1,]     10  212.1221    66.8   2330
 [2,]     20  228.1187    51.5   5340
 [3,]     40  399.0833    87.8  10900
 [4,]     80  712.5963   151.5  16500
 [5,]    160  953.3692   236.0  30000
 [6,]    320 1406.7825   302.0  71300
 [7,]    640 1812.4593   388.0  97900
 [8,]   1280 2260.1852   510.0 124000
 [9,]   2560 2567.8698   619.0 144000
> # Find quantile for the 2700 cfs value for drainage areas up to 40 and up to 80 square
    miles
> quantile(CQ40,.9753)
  97.53\%
2705.135
> quantile(CQ80,.937)
  93.7\%
2719.51
```

The quantile for drainage areas up to 40 square miles, for 90-th percentile discharges is 97-percent, so 2700 cfs is a rare value based on the empirical database. Even if the drainage area is doubled to 80 square miles, the quantile is 93-percent, still rare. Hence the engineer

can conclude that the discharge value of 2700 cfs has been observed in Texas for similar sized drainage areas, but the value is rare.

A similar procedure is illustrated in Listing 15 for the velocity. The HEC-RAS maximum velocity values are not even observed until the drainage area is quite large, at least an order of magnitude larger than the area in the simulation model. The engineer can further conclude that the computed velocities are unusual for the modeled situation.

Listing 15: R code for velocity conditioned on contributing drainage area

```
> # HEC-RAS EXAMPLE
> # Query the database, recover Velocity for different CDA at FDC > 90\%
>   CD10 <- DB[DB$CDA <   10 & DB$FDC >= 0.9 ,]$V
>   CD20 <- DB[DB$CDA <   20 & DB$FDC >= 0.9 ,]$V
>   CD40 <- DB[DB$CDA <   40 & DB$FDC >= 0.9 ,]$V
>   CD80 <- DB[DB$CDA <   80 & DB$FDC >= 0.9 ,]$V
>  CD160 <- DB[DB$CDA <  160 & DB$FDC >= 0.9 ,]$V
>  CD320 <- DB[DB$CDA <  320 & DB$FDC >= 0.9 ,]$V
>  CD640 <- DB[DB$CDA <  640 & DB$FDC >= 0.9 ,]$V
> CD1280 <- DB[DB$CDA < 1280 & DB$FDC >= 0.9 ,]$V
> CD2560 <- DB[DB$CDA < 2560 & DB$FDC >= 0.9 ,]$V
> # Build vectors for tabular output
> VMEAN   <-c(mean(CD10),mean(CD20),mean(CD40),mean(CD80),mean(CD160),mean(CD320),mean(
    CD640),mean(CD1280),mean(CD2560))
> VMEDIAN <-c(median(CD10),median(CD20),median(CD40),median(CD80),median(CD160),median(
    CD320),median(CD640),median(CD1280),median(CD2560))
> VMAX    <-c(max(CD10),max(CD20),max(CD40),max(CD80),max(CD160),max(CD320),max(CD640),max
    (CD1280),max(CD2560))
      CDAVAL    VMEAN VMEDIAN  VMAX
 [1,]     10 2.468865   1.880  7.29
 [2,]     20 2.254921   1.785  8.22
 [3,]     40 2.130951   1.780  8.41
 [4,]     80 2.132903   1.865  9.84
 [5,]    160 2.031214   1.750  9.84
 [6,]    320 2.077431   1.800 10.44
 [7,]    640 2.207989   1.860 13.09
 [8,]   1280 2.245259   1.910 13.85
 [9,]   2560 2.302871   1.980 13.85
```

The conditioned database is also meant to be used directly on the hydraulic variables. In the HEC-RAS model, the topwidth is between 100 and 200 feet. Listing 16 is a script to generate a table of discharges as a function of topwidth. Such tables would be useful in forensic interpretation of what discharge might have been in a stream from a debris line or other marker, or even just an ecotone in a stream to estimate bankfull discharge. In this example, the 2700 cfs discharge is greater than both the mean and median observed values for topwidths less than 200 and 100 feet, although the value is an order of magnitude smaller than the maximum value for these widths. The quantile associated with 2700 cfs for 100 and 200 foot topwidths are about 97 and 87 percent, respectively – as in the other two tables, the database suggests that the value of 2700 cfs is unusual, more so than would be expected for what is supposed to be a median value.

Listing 16: R code for discharge conditioned on topwidth

```
> # HEC-RAS EXAMPLE
> # Query the database, recover Discharge for different Topwidths at FDC > 90\%
>   TQ50 <- DB[DB$B <   50  & DB$FDC >= 0.9 ,]$Q
>  TQ100 <- DB[DB$B <  100  & DB$FDC >= 0.9 ,]$Q
>  TQ200 <- DB[DB$B <  200  & DB$FDC >= 0.9 ,]$Q
>  TQ400 <- DB[DB$B <  400  & DB$FDC >= 0.9 ,]$Q
> TQ1000 <- DB[DB$B < 1000  & DB$FDC >= 0.9 ,]$Q
> # Build vectors for tabular output
> TWVAL <-c(50,100,200,400,1000)
> QMEAN   <-c(mean(TQ50),mean(TQ100),mean(TQ200),mean(TQ400),mean(TQ1000))
> QMEDIAN <-c(median(TQ50),median(TQ100),median(TQ200),median(TQ400),median(TQ1000))
> QMAX    <-c(max(TQ50),max(TQ100),max(TQ200),max(TQ400),max(TQ1000))
> cbind(TWVAL,QMEAN,QMEDIAN,QMAX)
```

```
     TWVAL     QMEAN QMEDIAN   QMAX
[1,]    50  175.9892   66.15  35100
[2,]   100  497.8758  195.00  35100
[3,]   200 1238.3178  412.00  35100
[4,]   400 2831.0539  722.00  88300
[5,]  1000 4320.7715  962.00 209000
> quantile(TQ100,.979)
  97.9\%
2712.13
> quantile(TQ200,.87)
 87\%
2690
```

Listing 17 is the script to recover velocity for various topwidths. The HEC-RAS maximum velocities are above 10 ft/sec, which is substantially larger than any of the mean or median velocities in the table in the listing. Furthermore, although the value of 10 ft/sec is contained in the database, the value represents rare behavior. The quantile associated with 10 feet per second for 100 and 200 foot topwidths are about 99.99 percent, respectively – the database suggests that the value of 10 ft/sec is unusual, far more so than would be expected for what is supposed to be a median value.

Listing 17: R code for velocity conditioned on topwidth

```
> # HEC-RAS EXAMPLE
> # Query the database, recover Velocity for different Topwidths at FDC > 90\%
>   TV50 <- DB[DB$B <   50  & DB$FDC >= 0.9 ,]$V
>  TV100 <- DB[DB$B <  100  & DB$FDC >= 0.9 ,]$V
>  TV200 <- DB[DB$B <  200  & DB$FDC >= 0.9 ,]$V
>  TV400 <- DB[DB$B <  400  & DB$FDC >= 0.9 ,]$V
> TV1000 <- DB[DB$B < 1000  & DB$FDC >= 0.9 ,]$V
> # build vectors for tabular output
> TWVAL <-c(50,100,200,400,1000)
> VMEAN   <-c(mean(TV50),mean(TV100),mean(TV200),mean(TV400),mean(TV1000))
> VMEDIAN <-c(median(TV50),median(TV100),median(TV200),median(TV400),median(TV1000))
> VMAX    <-c(max(TV50),max(TV100),max(TV200),max(TV400),max(TV1000))
> cbind(TWVAL,VMEAN,VMEDIAN,VMAX)
     TWVAL    VMEAN VMEDIAN  VMAX
[1,]    50 1.840698   1.675 10.10
[2,]   100 2.045531   1.870 11.04
[3,]   200 2.241832   2.020 11.29
[4,]   400 2.438318   2.160 13.85
[5,]  1000 2.479427   2.170 13.85
> quantile(TV100,.9999)
  99.99\%
10.30238
> quantile(TV200,.9999)
  99.99\%
11.02293
```

For the example provided the value of 2700 cfs is somewhere between the 90th and 95th percentile of the censored flow distribution. The first interpretation is first that the value is within in the observation set, so 2700 cfs for an approximately 40 square mile drainage area is not unheard of. The value is at a higher percentile, and thus somewhat rare. The ephemeral nature of the stream of interest may encourage us to retain this value, and acknowledge the rarity; alternatively we may judge our hydrology to be suspect. The empirical database has confirmed that 2700 cfs is a realizable value, it has been observed in Texas for a drainage area of the appropriate size, it is not an absurd value, but it is a relatively rare value.

Examining the velocity results the HEC-RAS model values all exceed 6 feet per second, which are substantially greater than the mean and median values in Listing 17 and Listing 15. The two values near 11 feet per second are close to maximum values that have been

recorded in the database and hence would be quite rare.

The modeler can now assess the utility of the HEC-RAS simulation, given that the flow values (discharge and velocity) are rare compared to values in the conditioned database. The modeler might conclude that if the 2700 cfs value is to be retained the water in the HEC-RAS model should perhaps be deeper; a reasonable adjustment would be to change the HEC-RAS boundary conditions [7]. The QGAM estimated discharge of 630 cfs is at about the 84-th percentile in the contributing drainage area conditioned database which is a more central value than the original 2700 cfs.

One final assessment is the cross sectional flow area. Listing 18 is R script that extracts cross sectional flow area, in this case conditioned on topwidth. The two values of 100 and 200 feet are selected, and the flow areas associated with those values are extracted.

Listing 18: R code for flow area conditioned on topwidth

```
> # HEC-RAS EXAMPLE
> TA100 <- DB[DB$B <  100  & DB$FDC >= 0.9 ,]$A
> TA200 <- DB[DB$B <  200  & DB$FDC >= 0.9 ,]$A
> TMEAN <-c(mean(TA100),mean(TA200))
> TMEDIAN <-c(median(TA100),median(TA200))
> TMAX <-c(max(TA100),max(TA200))
> TLBL<-c(100,200)
> cbind(TLBL,TMEAN,TMEDIAN,TMAX)
     TLBL    TMEAN TMEDIAN TMAX
[1,]  100 207.0263   104.1 5220
[2,]  200 446.6126   213.0 5220
```

In the HEC-RAS simulation the flow areas for 2700 cfs are about 300 ft^2, which is close to the values in the listing, although the values in the listing correspond to different discharge and velocity values.

Table 7 summarizes the relationships between the HEC-RAS values used and the associated mean values from the conditional distributions. Hence like the regression method, the conditional distribution method also suggests that the discharge and subsequent velocity values used in the HEC-RAS model are unusual and bear additional examination.

Table 7: Comparison of HEC-RAS and Conditional Distribution Results

Variable	HEC-RAS	CDA $\approx$ 40 mi^2	B $<$ 100 ft	B $<$ 200 ft
Q cfs	2700	$\approx$ 400	497	1238
V_{mean} ft/s	$\approx$ 8	2.13	2.04	2.24
V_{max} ft/s	$\approx$ 11	8.41	11.04	11.29
B ft	90–150	58	100	200
A ft^2	270–400	160	207	446

To conclude this portion of the example, the engineer returns to the original questions:

[7] In the example normal depth was used with a slope of 0.013. Reducing the downstream slope to 0.005 has substantial impact on the flow depth and drops the velocity considerably. An alternative would be to specify a water elevation at the downstream boundary condition; however at some point the engineer needs to be concerned with the ability to pass flow under the bridge.

1. Is 2700 cfs an accessible discharge? Yes, the value has been observed for drainage areas about the same as the area at the point of interest, however the value is comparatively rare and not anticipated for a "median" event.
2. What is the anticipated water velocity? Somewhere in the 2–8 feet per second range. The HEC-RAS values are high, suggesting either too much discharge (hydrology) or model boundary conditions are inappropriate.
3. What is a typical topwidth? For the location of interest, based on contribution drainage area, 58 feet is the anticipated value. The HEC-RAS values are larger, but not by too much.

9.3.2 Method 2: Addendum

During the development of the example, the researchers developed charts from the underlying database to help check the results. These charts seemed useful enough to include in this section as an addendum.

Figure 55 is a plot of the velocity associated with different values of discharge from the entire Texas database. The axes are rotated intentionally to make the chart into a potential tool. The red line on the chart is an ad-hoc "fit" through the data "cloud" that represents the upper limit of observed velocities, there are some markers above the line, but the majority of the observations lies at or below this line. The utility is simple, if for instance a discharge of 10,000 cfs is appropriate, the anticipated upper velocity is about 10 feet per second.

The blue line on the chart is an ad-hoc "fit" that represents a lower limit of observed velocities. For the same 10,000 cfs, the lower anticipated velocity value is slightly under 1 foot per second. These two values constitute an estimate of the range of velocity values that would be anticipated at a location; a model that produces results outside this range would require some explaination.

Figure 56 is a plot of the velocity associated with different values of topwidth from the entire Texas database. The red and blue lines are ad-hoc "fits" through the data cloud. The meaning of the red and blue lines are the same as in the previous figure; the red line is an upper limit, while the blue line is a lower limit. This chart is particularly useful because the topwidth value could represent a forensic value (from a post-event debris survey), a design value perhaps just below the low chord of a bridge, or an inferred value from a site survey at an ephemeral stream where the evidence of discharge may just be a change in vegetation.

Figure 57 is a plot of the velocity associated with different values of cross-section flow area from the entire Texas database. The red and blue lines are ad-hoc "fits" through the data cloud. The meaning of the red and blue lines are the same as in the previous figures; the red line is an upper limit, while the blue line is a lower limit. This chart has utility in

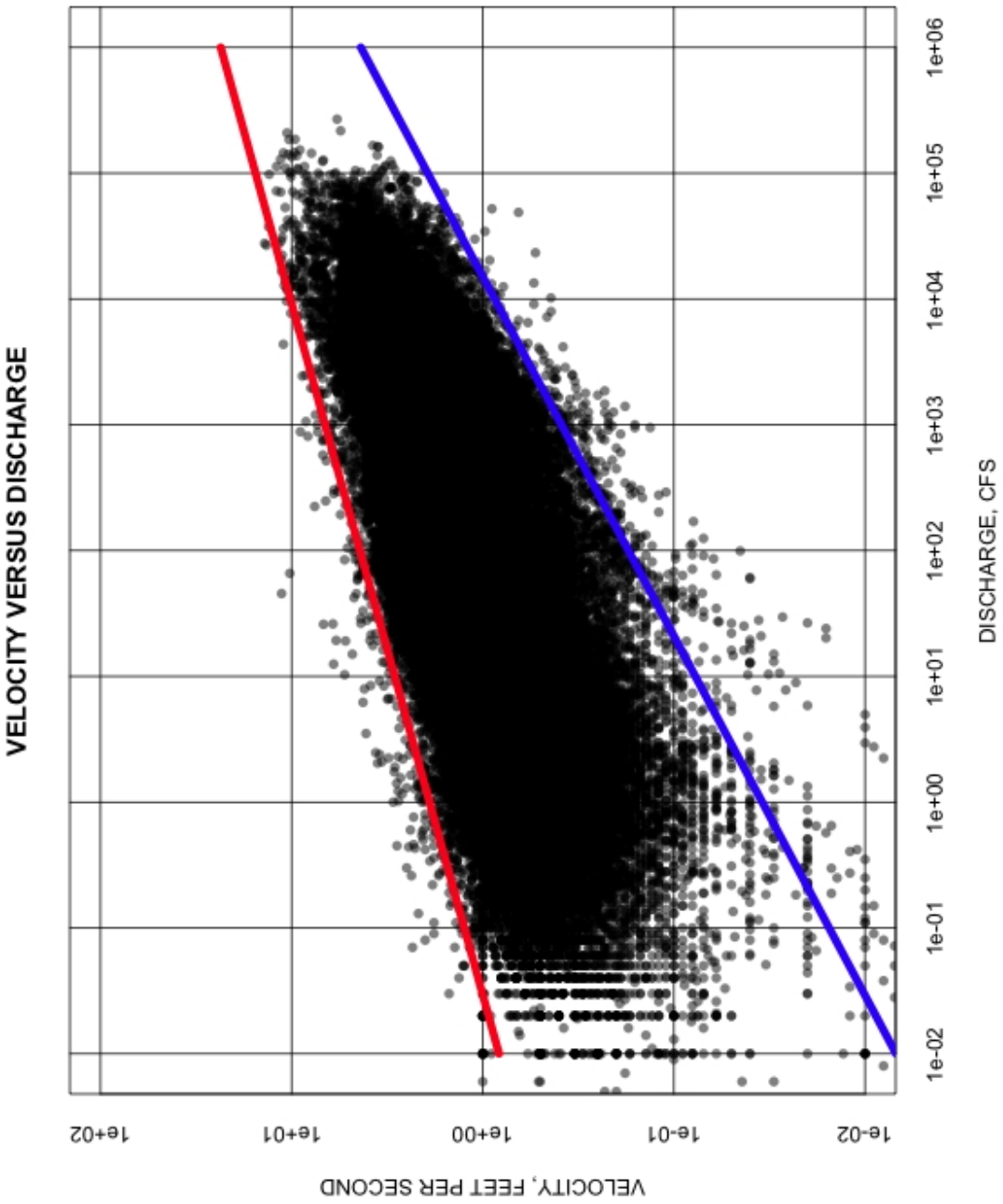

Figure 55: Velocity-discharge relationship.

estimating ranges of mean section velocity based upon an anticipated flow area.

Figure 58 is a plot of the velocity associated with different values of hydraulic depth from the entire Texas database. The red and blue lines are ad-hoc "fits" through the data cloud.

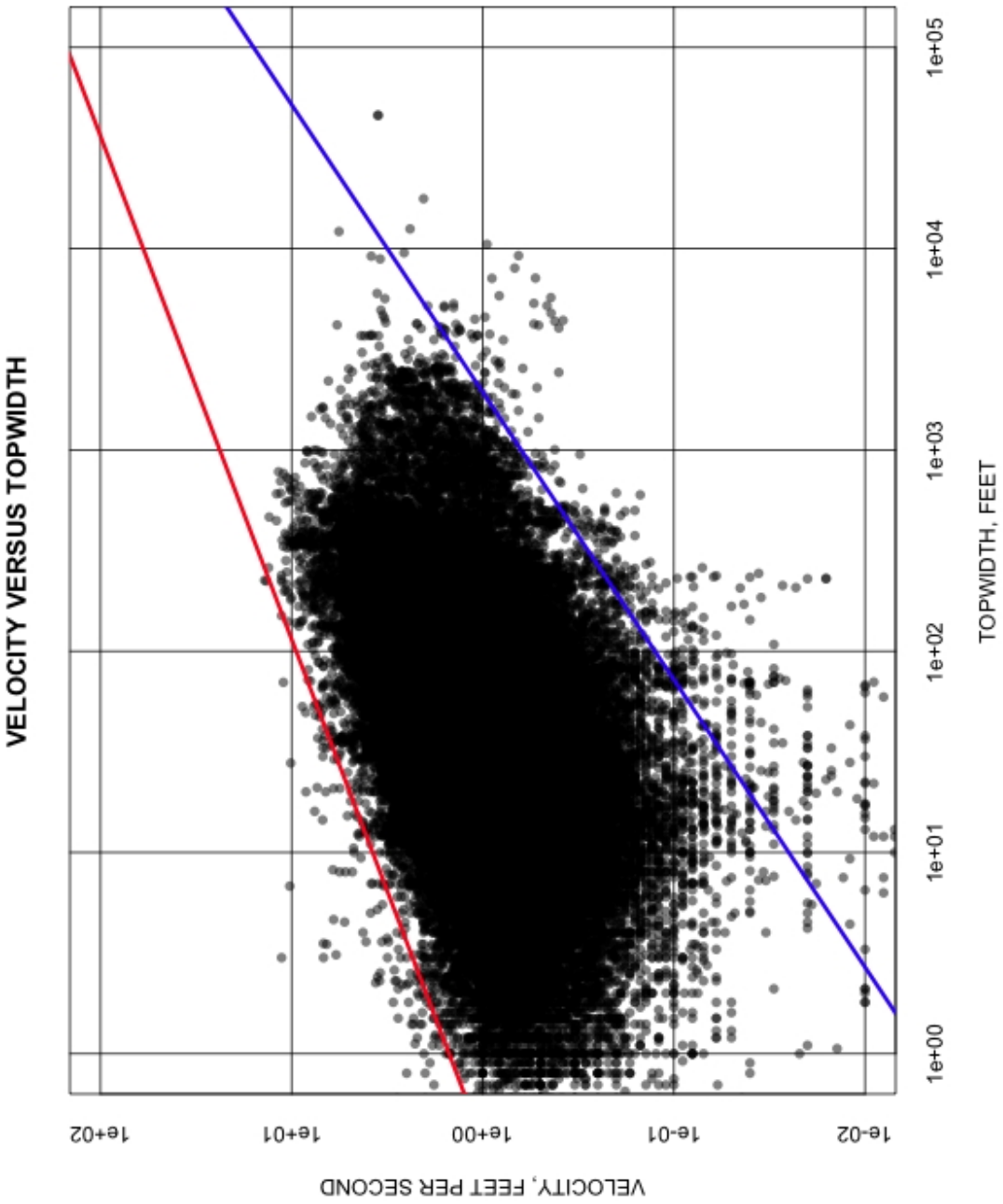

Figure 56: Velocity-topwidth relationship.

The meaning of the red and blue lines are the same as in the previous figures; the red line is an upper limit, while the blue line is a lower limit. Figure 58 represents a derived value, the hydraulic depth represented in the figure is the ratio of flow area and topwidth.

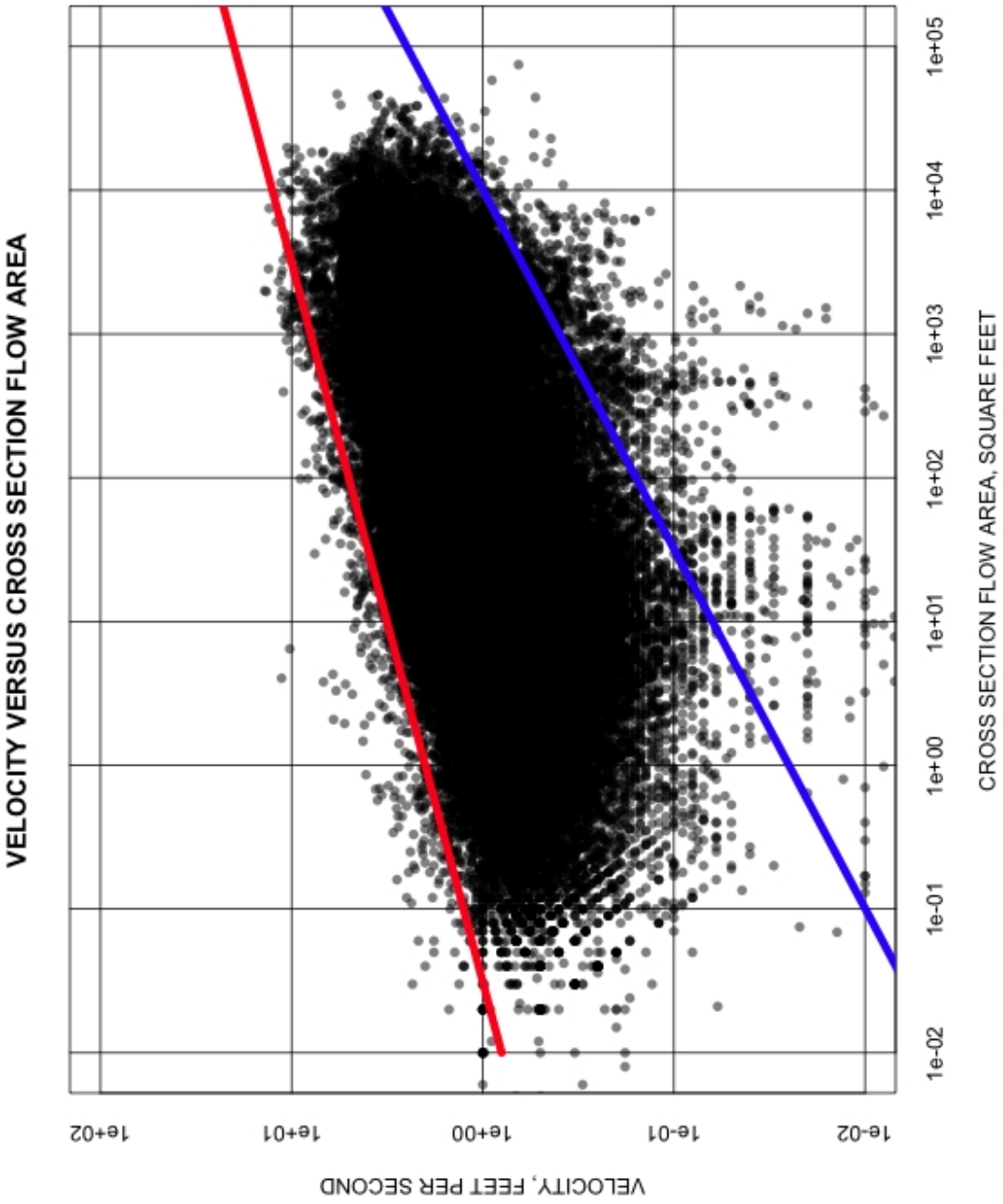

Figure 57: Velocity-area relationship.

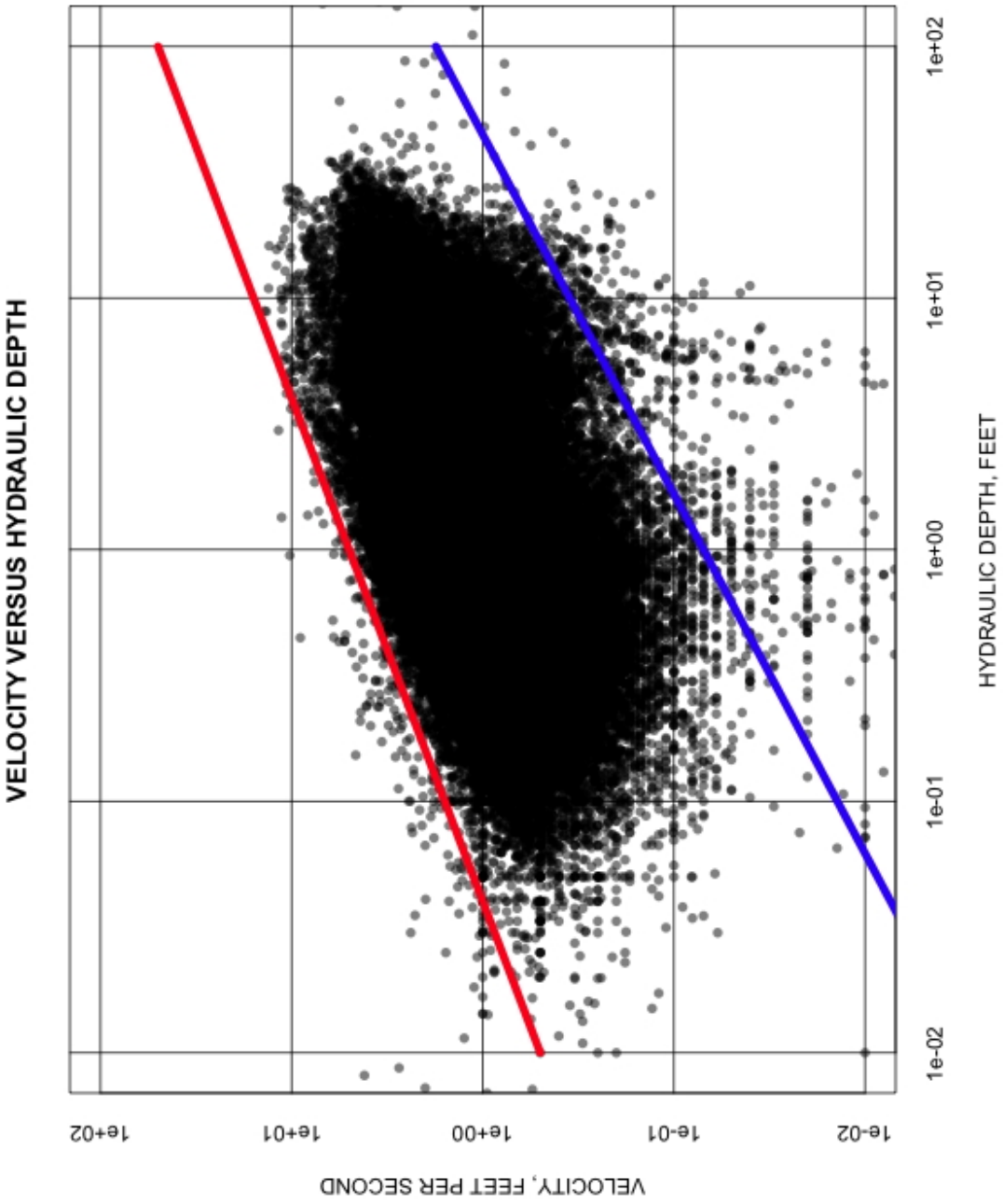

Figure 58: Velocity-depth relationship.

9.4 Method 3: Dimensionless Hydraulic Geometry

The hydraulic geometry approach assesses model suitability by comparing dimensionless characteristics estimated by the model with values from literature sources to determine if the model results are consistent with anticipated stream behavior.

In this example three values of bankfull discharge are postulated, 270 cfs, 2700 cfs, and 3966 cfs. These values represent a value anticipated to be too small (270 cfs), a value estimated by Asquith et al. (1996) for the location assuming about 40 square miles of contributing drainage area (2700 cfs), and a high value that was estimated by an alternate hydrologic approach (3966 cfs).

The method requires a specification of a mean grain diameter. A value of 0.4 millimeters was used as representative of a sand-sized material. The actual mean grain diameter is hard to quantify at the location, there is evidence of larger material embedded in the sand matrix Herrmann and Cleveland (2010), so in this example sand is assumed as appropriate.

The HEC-RAS simulation was run assuming normal depth at the downstream boundary with the slope set to 0.01. This condition results in the flow being nearly supercritical in the modeled section as explained previously. Using the output from HEC-RAS, dimensional values can be constructed from a subset of Equation 36 which is repeated below as Equation 64.

$$\hat{Q} = \frac{Q_{bf}}{\sqrt{gD_{50}}{D_{50}}^2}, \quad \hat{B} = \frac{B_{bf}}{D_{50}}, \quad \hat{H} = \frac{H_{bf}}{D_{50}} \tag{64}$$

Tables 8 and 9 are listings of the computed dimensionless hydraulic variables based on the HEC-RAS output and Equation 64 above. In each table the values in the first three columns are from HEC-RAS output, the next three columns are the result of application of Equation 64 to produce the corresponding dimensionless parameters based solely on the assumed mean grain diameter.

Table 8: Dimensionless hydraulic variable for Guadalupe Arroyo example, upstream section, $S_f \approx 0.012$

Q_{bf} (cfs)	B_{bf} (ft)	H_{bf} (ft)	$\hat{Q}$	$\hat{B}$	$\hat{H}$
270	85	0.85	7.63E8	6.48E4	6.48E2
2700	149	2.45	7.63E9	1.14E5	1.87E3
3966	153	3.23	1.12E10	1.17E5	2.46E3

These two sets of values are plotted on Figures 59 and 60. The solid square markers are the results for the approach cross section, and the filled circles are the results for the narrow section. The plots suggest that the two larger flow values are consistent with other sand bed rivers. The lowest discharge value plots sufficiently far away from the sand data that

Table 9: Dimensionless hydraulic variable for Guadalupe Arroyo example, narrow (bridge) section, $S_f \approx 0.012$

Q_{bf} (cfs)	B_{bf} (ft)	H_{bf} (ft)	$\hat{Q}$	$\hat{B}$	$\hat{H}$
270	71	0.76	7.63E8	5.41E4	5.79E2
2700	84	3.15	7.63E9	6.40E4	2.40E3
3966	86	4.96	1.12E10	6.55E4	3.78E3

it would be a potentially suspicious value if the system was indeed a sand bed river[8].

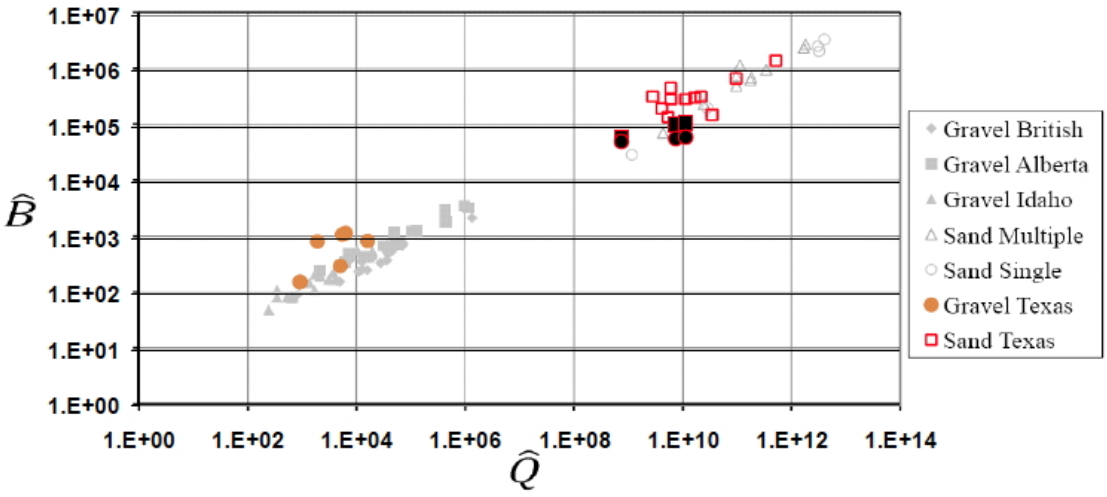

Figure 59: Dimensionless bankfull discharge $\widehat{Q}$ vs. dimensionless top width $\widehat{B}$ with three prospective discharge values.

[8]The site in the example is a sand-cobble river bed; the mean grain diameter used in the example is simply an assumption that half of a sample mass will pass a number 4 sieve and half will be retained on that sieve.

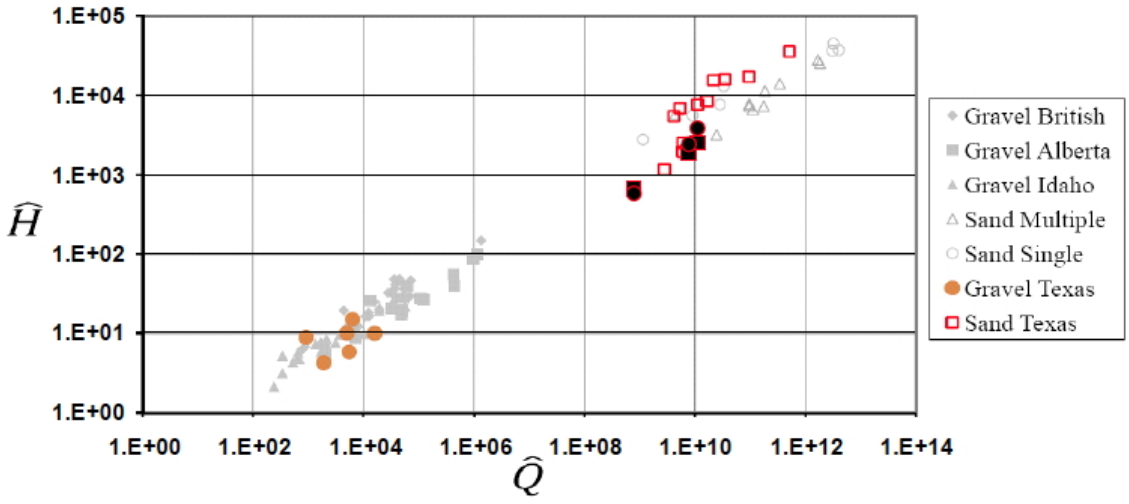

Figure 60: Dimensionless bankfull discharge $\widehat{Q}$ vs. dimensionless hydraulic depth $\widehat{H}$ with three prospective discharge values.

10 Conclusions

The project has made use of an extensive USGS flow measurement database to develop tools for estimating discharge at ungaged river crossing and/or for estimating how often a particular flow or velocity is observed in the state. The purpose of both tools is to aid TxDOT hydraulic engineers in making reasonable estimates of discharge or velocity at crossing without any data and for checking the "reasonableness" of velocity values calculated by hydraulic models.

The project started with the original USGS database of measured streamflow properties of discharge, section-averaged velocity, top width, and cross sectional flow area. Additional gaging station characteristics and spatial regional parameterization were added to the original USGS data for each record, and the combined database was then refined to remove low-quality data. The resulting database contains 87,874 records from 437 streamgages throughout Texas. This database is one of the primary products of the research projects, and details regarding its development and a listing of data available per record and station can be found in section 3. Filtering the data down to only those records for which the flow exceeded the 90th-percentile discharge (i.e, storm flow) resulted in a database of 17,753 records from 424 stations.

Three different methods were developed to make use of the new database and provide tools for checking the reasonableness of hydraulic model output. They are:

- Method 1: Statistical regression equations
- Method 2: Empirical distributions
- Method 3: Non-dimensional hydraulic geometry relations

The statistical approach (Method 1) uses generalized additive modeling techniques to develop regression equations for the dependent variables of discharge, Q, and section-averaged velocity, V, based on the independent variables of cross-sectional area, A, top width, B, the mean annual precipitation for the gage location, the latitude and longitude of the station, and the regional OmegaEM parameter of Asquith and Roussel (2009); the OmegaEM parameter is a generalized terrain and climate index expressing relative differences in peak-streamflow potential across Texas. Since both mean annual precipitation and the OmegaEM parameters are a function of latitude and longitude, the resulting regression equations for Q and V are essentially a function of cross-sectional flow area, top-width, and geographic location within Texas. These regression equations can then be used as independent predictors to either obtain unknown discharge estimates, or to obtain independent velocity estimates for comparison against hydraulic model output.

The empirical distribution method (Method 2) uses R to access the database and produce either conditioned or unconditioned non-exceedance curves and tables that tell the user how often a particular value of discharge, velocity, or other ancillary variable, has been observed in the database under the given constraints (say for example, how often as a velocity of 5

ft/s been observed for storm-flow conditions?). Values obtained from the hydraulic model could easily be used to lookup the non-exceedance value, and the exceedance value would give the user the needed information to see if the value they obtained was within reason. Non-exceedance plots are also given for the primary ancillary parameters that can be calculated from the database for quickly checking how often various values are observed in the database.

The third method developed for checking the reasonableness of hydraulic model output is Method 3. Method 3 makes use of the fact that most natural rivers, including those in Texas, have rather universal hydraulic geometries in non-dimensional space. Method 3 therefore provides plots and procedures for how to check the output from hydraulic models against observed trends. If, after non-dimensionalization, the model output does not follow the trend and scatter of measured observations, one would conclude that the model should be re-examined for accuracy.

Details of each of these methods can be found in sections 4-7, and examples outlining use of the methods is given in section 9.

References

Afzalimehr, H., V. P. Singh, and E. F. Najafabadi (2010). Determination of form friction factor. *Journal of Hydrologic Engineering 15*(3), 237–243.

Allen, P. M., J. G. Arnold, and W. Skipwith (2002). Erodibility of urban bedrock and alluvial channels, north texas. *Journal of the American Water Resources Association 38*(5), 1477–1492.

Asquith, W., G. Herrmann, and T. Cleveland (2013). Generalized additive regression models of discharge and mean velocity associated with direct-runoff conditions in Texas: The utility of the U.S. Geological Survey discharge measurement database. *in press, Journal of Hydrologic Engineering*.

Asquith, W. H. and M. C. Roussel (2009). Regression equations for estimation of annual peak-streamflow frequency for undeveloped watersheds in Texas using an L-moment-based, PRESS-minimized, residual-adjusted approach. Scientific Investigations Report 2009–5087, U.S. Geological Survey.

Asquith, W. H. and R. M. Slade (1997). Regional equations for estimation of peak- streamflow frequency for natural basins in Texas. Technical report, U.S. Geological Survey Water- Resources Investigations Report 96–4307.

Asquith, W. H., R. M. Slade, and J. Lanning-Rush (1996). Peak-flow frequency and extreme flood potential for streams in the vicinity of the Highland Lakes, central Texas. Water-Resources Investigations Report 96–4072, U.S. Geological Survey.

Barnes, H. H. (1967). Roughness characteristics of natural channels. *U.S. Geological Survey Water-Supply Paper 1849*, 213p.

Bathurst, J. C. (1985). Flow resistance estimation in mountain rivers. *Journal of Hydraulic Engineering 111*(4), 625–643.

Bathurst, J. C. (2002). At-a-site variation and minimum flow resistance for mountain rivers. *Journal of Hydrology 269*(1-2), 11 – 26.

Bertoldi, W., Gurnell, A. M., and N. A. Drake (2011). The topographic signature of vegetation development along a braided river: Results of a combined analysis of airborne lidar, color air photographs, and ground measurements. *Water Resources Research 47*, W06525.

Bomar, G. W. (1994). *Texas Weather*. Austin, Texas: University of Texas Press.

Box, G. E. P. and D. R. Cox (1964). An analysis of transformations (with discussion). *J. Royal Stat. Soc., Series B 26*, 211–252.

Bray, D. I. (1979, September). Estimating average velocity in gravel-bed rivers. *Journal of the Hydraulics Division, Proceedings of the ASCE 105*(HY9), 1103–1122.

Brownlie, W. R. (1981). Prediction of flow depth and sediment discharge in open channels. Report KH-R-43A, W. M. Keck Laboratory of Hydraulics and Water Resources, California Institute of Technology, Pasadena, California, USA.

Brownlie, W. R. (1983). Flow depth in sand-bed channels. *Journal of Hydraulic Engineering 109*(7), 959–990.

Buffington, J. M. and D. R. Montgomery (1999). Effects of hydraulic roughness on surface textures of gravel-bed rivers. *Water Resources Research 35*(11), 3507–3521.

Campbell, L., I. McEwan, V. Nikora, D. Pokrajac, M. Gallagher, and C. Manes (2005). Bed-load effects on hydrodynamics of rough-bed open-channel flows. *Journal of Hydraulic Engineering 131*(7), 576 – 585.

Castro, J. M. and P. L. Jackson (2001). Bankfull discharge recurrence intervals and regional hydraulic geometry relationships: Patterns in the pacific northwest, usa. *J. Amer. Wat. Res. Assoc. 37*(5), 1249–1262.

Chang, H. (1988). *Fluvial Processes in River Engineering*. Krieger.

Charlton, F. G., P. M. Brown, and R. W. Benson (1978). The hydraulic geometry of some gravel rivers in britain. Technical Report INT 180, Hydraulics Research Station, Wallingford, England.

Chen, S.-C., Y.-M. Kuo, and Y.-H. Li (2011). Flow characteristics within different configurations of submerged flexible vegetation. *Journal of Hydrology 398*(1-2), 124–134.

Church, M. (2006). Bed material transport and the morphology of alluvial river channels. *Annual Review of Earth and Planetary Science 34*, 32554.

Church, M. and K. Rood (1983). Catalogue of alluvial river data. Technical report, Dept, of Geography, University of British Columbia, Vancouver, B. C., Canada.

Comiti, F., D. Cadol, and E. Wohl (2009). Flow regimes, bed morphology, and flow resistance in self-formed step-pool channels. *Water Resources Research 45*.

Conyers, M. M. and M. A. Fonstad (2005). The unusual channel resistance of the texas hill country and its effect on flood flow predictions. *Physical Geography 26*(5), 379–395.

Curran, J. H. and E. E. Wohl (2003). Large woody debris and flow resistance in step-pool channels, Cascade Range, Washington. *Geomorphology 51*(1-3), 141–157.

Dunne, T. and L. Leopold (1978). *Water in Environmental Planning*. San Francisco, CA: W.H. Freeman and Co.

Eaton, B. C. and M. Church (2004). A graded stream response relation for bed load–dominated streams. *Journal of Geophysical Research 109*(F03011).

Einstein, H. A. and N. Barbarossa (1952). River channel roughness. *Transactions ASCE 117*, 1121–1146.

Engelund, F. (1966). Hydraulic resistance of alluvial streams. *Journal of the Hydraulics Division, Proceedings of the ASCE 92*(HY2), 315–326.

Faraway, J. J. (2005). *Linear Models with R*. Boca Raton, Florida: Chapman & Hall/CRC.

Faraway, J. J. (2006). *Extending the Linear Model with R—Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, Florida: Chapman & Hall/CRC.

Fedele, J. J. and M. H. García (2001). Hydraulic roughness in alluvial streams: A boundary layer approach. In G. Seminara and P. Blondeaux (Eds.), *Riverine, Coastal, and Estuarine Morphodynamics*, Berlin, Germany, pp. 37–60. Springer-Verlag.

Ferro, V. (1999). Friction factor for gravel-bed channel with high boulder concentration. *Journal of Hydraulic Engineering 125*(7), 771–778.

García, M. H. (Ed.) (2008). *Sedimentation Engineering: Processes, Measurements, Modeling, and Practice*. ASCE Manuals and Reports on Engineering Practice No. 110. ASCE.

Gordon, N., T. McMahon, B. Finlayson, C. Gippel, and R. Nathan (2004). *Stream Hydrology: An Introduction for Ecologists* (second ed.). The Atrium, Southern Gate, Chichester, West Sussex: John Wiley.

Gustavson, T. C. (1978). Bed forms and stratification types of modern gravel meander lobes, Nueces River, Texas. *Sedimentology 25*(3), 401–426.

Hastie, T. J. and R. J. Tibshirani (1990). *Generalized Additive Models*. Boca Raton, Florida: Chapman & Hall/CRC.

Heitmuller, F. T. and P. F. Hudson (2009). Downstream trends in sediment size and composition of channel-bed, bar, and bank deposits related to hydrologic and lithologic controls in the Llano River watershed, central Texas, USA. *Geomorphology 112*(3-4), 246–260.

Helsel, D. R. and R. M. Hirsch (2002). Statistical methods in water resources. Techniques of Water-Resources Investigations Book 4, Chapter A3, U.S. Geological Survey.

Heritage, G. L., B. P. Moon, L. J. Broadhurst, and C. S. James (2004). The frictional resistance characteristics of a bedrock-influenced river channel. *Earth Surface Processes and Landforms 29*(5), 611–627.

Herrmann, G. and T. Cleveland (2010). Moving substrate in an ephemeral stream: Case study in bridge survival. *Transportation Research Record: Journal of the Transportation Research Board 2*(2201), 3–9.

Hey, R. D. (1979). Flow resistance in gravel-bed rivers. *Journal of the Hydraulics Division, Proceedings of the ASCE 105*(HY4), 365–379.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan (2004). Development of a 2001

national land cover database for the United States. *Photogrammetric Engineering and Remote Sensing 70*(7), 829–840.

Hostache, R., X. Lai, J. Monnier, and C. Puech (2010). Assimilation of spatially distributed water levels into a shallow-water flood model. part ii: Use of a remote sensing image of Mosel River. *Journal of Hydrology 390*(3-4), 257–268.

Huitt-Zollars (2008). US-62/180 Bridge Drainage Analysis at Guadalupe Arroyo. Modeling report, Huitt-Zollars.

Hutton, C. and R. Brazier (2012). Quantifying riparian zone structure from airborne LiDAR: Vegetation filtering, anisotropic interpolation, and uncertainty propagation. *Journal of Hydrology 442*, 36–45.

Kamphuis, J. W. (1974). Determination of sand roughness for fixed beds. *Journal of Hydraulic Research 12*(2), 193–203.

Kean, J. W. and J. D. Smith (2005). Generation and verification of theoretical rating curves in the Whitewater River basin, Kansas. *J. Geophys. Res. 110*, F04012.

Kean, J. W. and J. D. Smith (2010). Calculation of stage-discharge relations for gravel bedded channels. *Journal of Geophysical Research 115*(F3), F03020.

Kellerhals, R., C. R. Neill, and D. I. Bray (1972). Hydraulic and geomorphic characteristics of rivers in alberta, river engineering and surface hydrology report. Technical Report 72-1, Research Council of Alberta, Canada.

Keulegan, G. H. (1938). Laws of turbulent flow in open channels. *Journal of Research of the National Bureau of Standards Research Paper 1151*(21), 707–741.

López, R. and J. Barragán (2008). Equivalent roughness of gravel-bed rivers. *Journal of Hydraulic Engineering 134*(6), 847–851.

Magilligan, F. J. (1988). Variations in slope components during large magnitude floods, wisconsin. *Annals of the Association of American Geographers 78*(3), 520–533.

Manga, M. and J. W. Kirchner (2000). Stress partitioning in streams by large woody debris. *Water Resources Research 36*(8), 23732379.

Meyer-Peter, E. and R. Müller (1948). Formulas for bed-load transport. In *Proceedings of the Second Meeting*, Stockholm, Sweden, pp. 39–64. IAHR.

Millar, R. G. (1999). Grain and form resistance in gravel-bed rivers. *Journal of Hydraulic Research 37*(3), 303–312.

Miller, V. (1953). A quantitative geomorphic study of drainage basin characteristics in the Clinch Mountain area, Virginia and Tennessee. Technical Report 3, Columbia University, Department of Geology.

Minka, T. P. (2011). maps: Draw geographic maps.

Morton, R. A. and A. C. Donaldson (1978). Hydrology, morphology, and sedimentology of Guadalupe fluvial-deltaic system. *Geological Society of America Bulletin 89*(7), 030–1036.

Motayed, A. K. and M. Krishnamurthy (1980). Composite roughness of natural channels. *Journal of the Hydraulics Division 106*(6), 1111–1116.

National Research Council (2004). *Assessing the national streamflow information program—Committee on review of the USGS national streamflow information program*. Washington, D.C.: The National Academies Press.

O'Conner, J. E., R. H. Webb, and V. R. Baker (1986). Paleohydrology of pool-and-riffle pattern development: Boulder creek, utah. *Geological Society of America Bulletin 97*(4), 410–420.

O'Connor, J. E. and J. E. Costa (2003). Large floods in the United States: Where they happen and why. Circular 1245, U.S. Geological Survey.

Pagliara, S. and P. Chiavaccini (2006). Flow resistance of rock chutes with protruding boulders. *Journal of Hydraulic Engineering 132*(6), 545–552.

Parker, G. (1991). Selective sorting and abrasion of river gravel ii: Applications. *Journal of Hydraulic Engineering 117*(2), 150–171.

Parker, G. (2012). 1d sediment transport morphodynamics with applications to rivers and turbidity currents. e-book.

Parker, G., C. Paola, K. Whipple, and D. Mohrig (1998). Alluvial fans formed by channelized fluvial and sheet flow: theory. *Journal of Hydraulic Engineering 124*(10), 1–11.

Parker, G., C. Toro-Escobar, M. Ramey, and S. Beck (2003). The effect of floodwater extraction on the morphology of mountain streams. *Journal of Hydraulic Engineering 129*(11), 885–895.

Parker, G., P. Wilcock, C. Paola, W. Dietrich, and J. Pitlick (2007). Physical basis for quasi-universal relations describing bankfull hydraulic geometry of single-thread gravel bed rivers. *Journal of Geopphysical Research 112*, F04005.

Phillips, J. D. (2011). Universal and local controls of avulsions in southeast Texas rivers. *Geomorphology 130*(1-2), 17–28.

Phillips, J. D. and J. D. Lutz (2008). Profile convexities in bedrock and alluvial streams. *Geomorphology 102*(3-4), 554–566.

Phillips, J. D., M. C. Slattery, and Z. A. Musselman (2005). Channel adjustments of the lower Trinity River, Texas, downstream of Livingston Dam. *Earth Surface Processes and Landforms 30*(11), 1419–1439.

Prestegaard, K. L. (1983). Variables influencing water-surface slopes in gravel-bed streams

at bankfull stage. *Geological Society of America Bulletin 94*, 673–678.

PRISM Climate Group (2010, July 1, 2010). PRISM Data Explorer.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rameshwaran, P. and K. Shiono (2007). Quasi two-dimensional model for straight overbank flows through emergent vegetation on floodplains. *Journal of Hydraulic Research 45*(3), 302–315.

Recking, A., P. Frey, A. Paquier, P. Belleudy, and J. Y. Champagne (2008a). Bed-load transport flume experiments on steep slopes. *Journal of Hydraulic Engineering 134*(9), 1302–1310.

Recking, A., P. Frey, A. Paquier, P. Belleudy, and J. Y. Champagne (2008b). Feedback between bed load transport and flow resistance in gravel and cobble bed rivers. *Water Resources Research 44*(5), W05412.

Riggs, H. C. (1976). A simplified slope-area method for estimating flood discharges in natural channels. *Journal of Research of U.S. Geological Survey 4*(3), 285–291.

Rosgen, D. (1996). *Applied River Morphology*. Pagosa Springs, CO: Wildland Hydrology.

Rosgen, D. L. (1994). A classification of natural rivers. *Catena 22*, 169–199.

Shreve, R. (1967). Infinite topologically random channel networks. *Journal of Geology 75*, 178–186.

Slattery, M. C. and J. D. Phillips (2011). Controls on sediment delivery in coastal plain rivers. *Journal of Environmental Management 92*(2), 284–289.

Smith, J. D. and S. R. McLean (1977). Spatially averaged flow over a wavy surface. *Journal of Geophysical Research 82*(12), 1735–1746.

Strahler, A. (1952). Dynamic basis of geomorphology. *Bulletin of the Geological Society of America 63*, 923–938.

Stricklin, F. L. (1961). Degradational stream deposits of the Brazos River, Central Texas. *Geological Society of America Bulletin 72*(1), 19–35.

Turnipseed, D. P. and V. B. Sauer (2010). Discharge measurements at gaging stations. Techni. and Meth. Book 3, Chapter A8, U.S. Geological Survey.

U.S. Geological Survey (2009a, March 1, 2009). Daily mean streamflow values for Texas.

U.S. Geological Survey (2009b, March 1, 2009). Streamflow measurements for Texas.

Vanoni, V. A. (Ed.) (1975). *Sedimentation Engineering*. ASCE Manuals and Reports on Engineering Practice No. 54. ASCE.

Venables, W. N. and B. D. Ripley (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

Vogel, R. M. and N. M. Fennessey (1994). Flow-duration curves I: New interpretation and confidence intervals. *J. Wat. Res. Plan. Managment 120*(4), 485–504.

Whiting, P. J. and W. E. Dietrich (1990). Boundary shear stress and roughness over mobile alluvial beds. *Journal of Hydraulic Engineering 116*(12), 1495–1509.

Wiberg, P. L. and D. M. Rubin (1989). Bed roughness produced by saltating sediment. *Journal of Geophysical Research 94*(C4), 5011–5016.

Wilcox, A. C. and E. E. Wohl (2006). Flow resistance dynamics in step-pool stream channels: 1. large woody debris and controls on total resistance. *Water Resources Research 42*(5), W05418.

Wong, M. and G. Parker (2006). Reanalysis and correction of bed-load relation of meyer-peter and müller using their own database. *Journal of Hydraulic Engineering 132*(11), 1159–1168.

Wood, S. N. (2006). *Generalized Additive Models—An Introduction with R*. Boca Raton, Florida: Chapman & Hall/CRC.

Wood, S. N. (2009, March 1, 2009). mgcv: GAMs with GCV/AIC/REML smoothness estimation and GAMMs by PQL.

Wright, S. and G. Parker (2004). Flow resistance and suspended load in sand-bed rivers: Simplified stratification model. *Journal of Hydraulic Engineering 130*(8), 796–805.

A Evaluation of Main-Channel Slope and Proximal Slope for Statistical Regionalization of U.S. Geological Survey Discharge Measurements Associated with Direct-Runoff Conditions in Texas

Suggested Citation: Asquith, W.H., and Burley, T.E., 2013, Evaluation of main-channel slope and proximal slope for statistical regionalization of U.S. Geological Survey discharge measurements associated with direct-runoff conditions in Texas, *in* Cleveland, T.G., Liu, X., Sharif, H.O., and Strom, K.B., 2013, Statistical properties and regional analysis of empirical flow parameters in Texas as obtained from U.S. Geological Survey discharge measurement databases—Tools for assessment of hydraulic models: Texas Department of Transportation Research Report 0–6654–1, chapter A, pp. 129–145

A.1 Background

Watersheds and stream channels can be quantified by various physical properties. A commonly used property in statistical analysis of streamflow and streamflow hydraulics is slope. The "main-channel slope," which is a slope characteristic of a watershed, is an often used definition of slope for statistical analysis. The slope of a stream channel near a U.S. Geological Survey (USGS) streamflow-gaging station (streamgage) or "proximal slope" also is a physically and intuitively important component in analysis of streamflow hydraulics. There is anticipation that proximal slope should be more statistically related to streamflow hydraulic properties such as mean velocity and discharge than main-channel slope based on principles of open-channel hydraulics (Sturm, 2010). The anticipation in part exists because the morphology of a stream channel is partly formed by local topography, and because through hydraulic feedback mechanisms, channel slope responds to bed and bank sediment properties (Leopold and others, 1964; Richards, 1982; Castro and Jackson, 2001; Metcalf and others, 2009; Sougnez and Vanacker, 2010; Sturm, 2010). Simplified mathematics of open-channel hydraulics with local friction slope also indicate that proximal slope should be more relevant than main-channel slope, which is intended to be representative of the watershed and not necessarily specific or arbitrary stream channel locations.

The USGS, in cooperation with the Texas Department of Transportation, evaluated main-channel slope and proximal slope for statistical regionalization of discharge measurements associated with the direct-runoff conditions in Texas. This Chapter addresses the mutual interest of the researchers and the Texas Department of Transportation (TxDOT) Research Oversight Committee to evaluate these two slopes in the context of the mean velocity and discharge. Specifically, the two slopes are evaluated for statistical regionalization of the USGS discharge measurement database in Texas associated with direct-runoff conditions. Direct-runoff conditions for this evaluation are identified by using USGS discharge measurements exceeding the 90th-percentile daily mean streamflow at a given USGS streamgage in Texas. Data for the period of record from about 420 streamgages in Texas were used to create the cumulative percentages of daily mean streamflow as described in Asquith and others (2013). The streamgage locations are shown in Asquith and others (2013, figs. 5 and 9). Discharge measurements are described in Rantz and others (1982a) and Turnipseed and Sauer (2010). Rigorous investigations of statistical regionalization of the USGS discharge measurement database in Texas that are particularly germane to this Chapter are described in Asquith (2013) and Asquith and others (2013). Asquith (2013) considers median-flow conditions, and Asquith and others (2013) consider direct-runoff conditions based on the same 90th-percentile exceedance criteria used for this evaluation.

A.2 Are Main-Channel Slope and Proximal Slope Statistically Important for Regionalization of Discharge Measurements?

An open question is whether main-channel slope (*MS*) or proximal slope (*PS*) are statistically important variables for regionalization of discharge measurements? That is, do either of these slopes enhance the prediction of mean velocity or discharge values from USGS discharge measurements associated with direct-runoff conditions in Texas?

The dimensionless *MS* is defined as the change in elevation in feet (ΔE) between the two end points of *L* (main-channel length) divided by *L* in feet. Hence, *MS* is a "watershed slope" based on the principal channel network as represented in *L*. The *L* is defined as the length in stream-course feet of the longest defined channel represented by the particular data used by an analyst from the approximate watershed headwaters to the outlet (streamgage).

The *PS* was used as an alternative definition of slope to *MS*. The dimensionless *PS* is defined as the change in elevation in feet (ΔE) between points located at the watershed outlet (streamgage) and about 0.62-miles (`1KM`, kilometer in R code herein) upstream from the streamgage.

USGS researchers, led by coauthor T.E. Burley (USGS Texas Water Science Center, Austin, Tex.) collaborated with other USGS spatial analysts (Archuleta and others, 2012) to characterize numerous watershed properties associated with about 420 USGS streamgages throughout Texas. These USGS streamgages are the same streamgages in the USGS discharge measurement database represented within the broader TxDOT Research Project Report containing this Chapter. For this Chapter, spatial data processing for the streamgages was done within a geographic information system (GIS) by using scripts written in the Python programming language by T.E. Burley and C.A. Archuleta. The Python scripts automated existing algorithms within the Environmental Systems Research Institute ArcGIS Hydrology Toolbox (Environmental Systems Research Institute, 2010). Previously published geospatial datasets were used (Homer and other, 2004; Gesch, 2007; U.S. Geological Survey, 2011). The Federal Geographic Data Committee (2013) (FGDC) coordinates the development of the National Spatial Data Infrastructure. FGDC-compliant metadata documentation has been completed for the intermediate GIS files produced by the Python scripts.

Much previous work by Asquith, Cleveland, and other authors for the TxDOT Research Program (Asquith and Slade, 1997; Asquith, 2001; Asquith and others, 2005; Asquith, 2011; Asquith and Roussel, 2007, 2009; Asquith and others, 2011; Cleveland and others, 2006) used a "main-channel slope" for regionalization and regression analysis of streamflow statistics but not for streamflow hydraulics as done in this Chapter. The definition of slope used for those prior investigations is conceptually defined as equivalent to *MS* considered in this Chapter. Specifically, a 10-meter resolution digital elevation model (DEM) was used to compute *MS* for this study (Gesch, 2007). Although conceptually equivalent, the values for *MS* used in this Chapter are likely to differ numerically from the *MS* values used for those previous studies. For this Chapter, the National Hydrography Dataset (NHD) was used to represent stream channels (U.S. Geological Survey, 2011). Heitmuller and others (2006) used algorithms for calculation of *L* or *S* that were precursors to those used for this Chapter and concluded from comparisons between *MS* estimates computed from 10-and 30-meter DEMs that *MS* estimates for these two terrain resolutions have numerical compatibility.

Visualization of select variables and their interrelations is useful as part of data exploration. The relation between watershed drainage area (contributing) and *MS* is shown in figure 1, and in contrast, the relation between watershed drainage area and *PS* is shown in figure 2. The figures indicate some visual differences between these slopes as they relate to drainage area. It is evident that *MS* has a tendency to decrease with increasing drainage area. In contrast to *MS*, *PS* seemingly has a considerably less discernible relation to drainage area. The relation between *MS* and *PS* is shown in figure 3. There is little discernible relation between *MS* and *PS* as evident from the figure. For this

Chapter, streamgages having estimated values of *PS* equaling zero were not used.

Manning's equation (eq. 1) is a widely-used, one-dimensional, steady-state, and uniform-flow model of open-channel flow. It is hypothesized that *PS* should provide more *explanatory power* than *MS* for predicting mean velocity V in feet per second or discharge Q in cubic feet per second where $Q = VA$ for cross-sectional area A in square feet based on inspection of Manning's equation:

$$V = \frac{1.486}{n\text{-}value} R^{2/3} \sqrt{S_f} \quad \text{[English units]}, \tag{1}$$

where R is hydraulic radius in feet and S_f is dimensionless local friction slope (the water-surface slope associated with energy lost caused by friction). The equation indicates the proportionality of velocity to the square root of friction slope ($V \propto \sqrt{S_f}$). It is often assumed that the channel slope is equivalent to friction slope ($PS \equiv S_f$). The quantity known as "Mannings n" or *n-value* is a coefficient (Barnes, 1967; Benson and Dalrymple, 1967; Dalrymple and Benson, 1967) that roughly corresponds to the general frictional losses in one-dimensional, open-channel flow attributable to geomorphic characteristics such as bed and bank material, channel sinuosity, extent and type of vegetation, and other "roughness" related geomorphic features. A prospective statistical model of V is a regression equation of the form

$$V = a \times S^b, \tag{2}$$

where a and b are regression coefficients, and from Manning's equation b could be expected to equal about 0.5 because of the "$\sqrt{\text{slope}}$ term" indicates that $a \propto (1.486/n\text{-}value)R^{2/3}$. The proportionality of a is taken as appropriate for this analysis because the USGS discharge measurement database does not provide a direct means for computation of R, and the *n-value* is unknown.

The computation of R in feet according to its definition ($R = A/P$) requires the wetted perimeter P in feet (Sturm, 2010), which can be computed from the original paper USGS discharge measurement sheets. Unfortunately, digital versions of these sheets are not generally available. It is outside the scope of the broader TxDOT Research Project to (1) extract thousands of paper discharge measurement sheets from Federal archives, (2) manually digitize those paper sheets, and (3) coordinate the inclusion of many modern measurements in which digital records of depths and lateral position in the stream channel often are recorded.

A.2.1 The 10-85 Main-Channel Slope—An Aside

A brief aside is needed regarding other definitions of "main-channel slope." Another definition is known as the 10-85 main-channel slope (10-85 slope), which is the slope of the main channel as defined for the path of L from 10 percent of L upstream along the channel from the outlet (streamgage) to the point at which 85 percent of L is reached (Gordon and others, 2004; Koltun and others, 2006). The relation between *MS* and the 10-85 slope is shown in figure 4. The figure indicates strong collinearity between the two slopes. The collinearity means that having one is statistically like having the other, except the 10-85 slopes are about 1/4-log cycles smaller. This difference is important and indicates that simple substitution of the 10-85 slope for *MS* in studies such as Asquith and Roussel (2009) and other studies by the authors and others (Asquith and Slade, 1997; Asquith, 2001; Asquith and others, 2005; Asquith, 2011; Asquith and Roussel, 2007; Asquith and others, 2011; Cleveland and others, 2006) is not advisable. The explanation for the difference is that the two slopes are not conceptually identical. The figure supports a conviction that there is little justification for using the 10-85 slope in lieu of *MS* in regional statistical models of Texas hydrology.

A.3 Preprocessing

The importance of these slopes is evaluated in this Chapter within the R Environment for Statistical Computing (R Development Core Team, 2012). As a consequence, listings of "R code" are shown. Efforts to describe R software input and output are made, and readers are directed to the documentation of the R software for more details. Comments are threaded within the R code listings and are identified by text and characters following the "#" (pound sign). Assignments to variables in the R code are identified by the "<-" character sequence. The "~" (tilde) symbol is to be read as "is a function of." Acronyms and other substantial abbreviations are separately identified as context dictates.

Listing 1 provides an algorithmic example of reprocessing of variables, base-10 logarithmic transformation, and filtering to determine those measurements associated with direct-runoff conditions (>90th-percentile daily mean streamflow) (Asquith and others, 2013). The filtering also removes those streamgages for which *PS* was computed as zero. The database (text) file is referred to in this Chapter is titled *QVFABmergePROXIMAL.txt* and is referred to elsewhere by this name in this TxDOT Research Report.

Listing 1. R code demonstrating the loading of data, subsequent filtering for the 90th-percentile daily mean streamflow, and base-10 logarithmic transformation of select variables.

```
# "QVFABmergePROXIMAL.txt" is a text file database.
DB <- read.table("QVFABmergePROXIMAL.txt", header=TRUE);
# DB is "assigned" the contents of the database.
DB <- DB[DB$ESTFLOWPROB > 0.90 & DB$MC_SLOPE_1KM != 0,]; # filter
# to acquire 90th percentile (direct-runoff conditions) of
# daily mean streamflow (as explained in the text) and in
# Asquith and others (2013) as well as not include those stations
# with an estimated proximal slope of zero.

# log10() = base-10 logarithmic transformation
logV  <- log10(DB$VELOCITY);  # mean velocity, feet / second
logQ  <- log10(DB$DISCHARGE); # discharge, cubic feet / second
logB  <- log10(DB$TOPWIDTH);  # top width, feet
logA  <- log10(DB$AREA);      # cross-sectional area, square feet
logMS <- log10(DB$MC_SLOPE);        # MS (see text), dimensionless
logPS <- log10(DB$MC_SLOPE_1KM);    # PS (see text), dimensionless
logR  <- log10(DB$AREA/Q$TOPWIDTH); # pseudo R (see text), feet
```

The last line in Listing 1 concerning a "pseudo R" requires discussion. Hydraulic depth D_h is equal to cross-sectional area A in square feet divided by topwidth B in feet or $D_h = A/B$. It is often assumed within engineering practice that R can be approximated as $R \approx D_h$ when P is unavailable to compute $R = A/P$. In other notation, $R \approx R_\circ$ (a "pseudo R") can be used as a first-order approximation of hydraulic radius in Manning's equation. Basic geometry and algebra requires that numerical differences between R and R_o diminish for wider, shallower stream channels.

Lastly, a caveat is needed regarding the streamflow metrology of the USGS discharge measurement database. From a data processing perspective, it is not possible to reliably differentiate "section control" measurements from "bed-and-bank control" (Rantz and others, 1982a,b; Sturm, 2010; Turnipseed and Sauer, 2010). Low-flow measurements are more likely to be made near a section control (such as a riffle or low-head dam [weir]) where non-uniform flow exists and Manning's equation does not wholly represent the hydraulics of open-channel flow (Rantz and others, 1982a, p. 10). Because of the filtering of the USGS discharge measurement database for the 90th-percentile daily mean streamflow, the measurements for "direct-runoff conditions" analyzed in this Chapter and in Asquith and others (2013) can be expected to mostly reflect bed-and-bank control. When measurements are made for bed-and-bank control circumstances, Manning's equation with its attendant

assumptions of steady-state, uniform, one-dimensional flow has increased applicability (Rantz and others, 1982b).

A.4 Analysis of Variance of Simple Linear Regression Models involving Slope

In Listing 2, two linear (base-10 logarithmic) regressions of mean velocity V (`logV`) with both MS and PS as variables (`logMS` and `logPS`), respectively, are presented. Analysis of variance (ANOVA) is made on each regression by using the `anova()` function. The ANOVA results indicate slightly larger residual sum of squares (RSS, `Residuals SumSq.` in Listing 2) for the use of PS as a variable in the regression (RSS=1,247) than for MS as a variable (RSS=1,232). From the general similarity of the RSS values, a conclusion can be drawn that PS does not have more explanatory power than MS for V estimation.

Listing 2. R code demonstrating various linear regression models of mean velocity with main-channel slope and proximal slope as defined in Listing 1 and subsequent analysis of variance (ANOVA).

```
# SELECT ABBREVIATIONS:
#   lm()    = linear model function, see R documentation
#   Pr      = probability
#   DF      = degrees of freedom
#   SumSq.  = sum of squares
#   MeanSq. = mean square
#   Pr(>F)  = probability (p-value) of the F-value of F-test
lmv1.1 <- lm(logV ~ logMS) # a log-linear model of V(MS)
anova(lmv1.1) # See documentation of R for output summary
   Analysis of Variance Table
   Response: logV
                DF  SumSq. MeanSq. F-value    Pr(>F)
   logMS         1   15.98 15.9763  227.35  <2.2e-16
   Residuals 17538 1232.41  0.0703

lmv1.2 <- lm(logV ~ logPS) # a log-linear model of V(PS)
anova(lmv1.2) # See documentation of R for output summary
   Analysis of Variance Table
   Response: logV
                DF  SumSq. MeanSq. F-value    Pr(>F)
   logPS         1    1.06 1.05571  14.844 0.0001172
   Residuals 17538 1247.33 0.07112
```

A similar analysis for discharge Q is presented in Listing 3. In Listing 3, two simple linear regressions of Q (`logQ`) with both MS and PS as variables `logMS` and `logPS`, respectively, are presented. The ANOVA results indicate a considerably larger residual sum of squares (RSS) for the use of PS as a variable in the regression (RSS=15,954) than for MS as a variable (RSS=12,405). The RSS values indicate that MS explains slightly more variation than PS. From the general similarity of the RSS values, a conclusion can be drawn that PS does not have more explanatory power than MS for Q estimation.

Listing 3. R code demonstrating various linear regression models of discharge with main-channel slope and proximal slope as defined in Listing 1 and subsequent analysis of variance (ANOVA).

```
# SELECT ABBREVIATIONS:
#   lm()    = linear model function, see R documentation
#   Pr      = probability
#   DF      = degrees of freedom
#   SumSq.  = sum of squares
#   MeanSq. = mean square
#   Pr(>F)  = probability (p-value) of the F-value of F-test
lmq1.1 <- lm(logQ ~ logMS) # a log-linear model of Q(MS)
```

```
anova(lmq1.1) # See documentation of R for output summary
   Analysis of Variance Table
   Response: logQ
               DF  SumSq. MeanSq. F-value    Pr(>F)
   logMS        1  3772.4  3772.4  5333.5  <2.2e-16
   Residuals 17538 12404.6     0.7

lmq1.2 <- lm(logQ ~ logPS) # a log-linear model of Q(PS)
anova(lmq1.2) # See documentation of R for output summary
   Analysis of Variance Table
   Response: logQ
               DF  SumSq. MeanSq. F-value    Pr(>F)
   logPS        1   222.6  222.57  244.66  <2.2e-16
   Residuals 17538 15954.3    0.91
```

A.5 Analysis in a Simplified Form of Manning's Equation

Linear regression models of V and Q are made by using a simplified form of Manning's equation in the two sections that follow. Select abbreviations for the linear regressions shown in Listings 4–7 are as follows:

```
lm()        = linear model function,
Min         = minimum,
1stQua      = first quartile,
3rdQua      = third quartile,
Max         = maximum,
Std.Error   = standard error in units of the response variable,
t-value     = T-statistic for the t-test,
Pr          = probability (p-value),
Pr(>|t|)    = probability of absolute value of t-value,
F-statistic = a statistic for the F-test, and
p-value     = p-value (a standard computed probability).
```

A.5.1 Linear Regression of Mean Velocity Using a Simplified Form of Manning's Equation

In this section, a linear regression of mean velocity V by using a simplified form of Manning's equation is presented. Manning's equation (eq. 1) provides for the relation between V, *n-value*, hydraulic radius R, and friction slope S_f. The *n-value* for individual discharge measurements is not determined. The R is the pseudo-hydraulic radius $R_\circ$. The S_f is considered as equal to either MS or PS.

The regression of V as a function of $R_\circ$ and MS along with a regression summary are shown in Listing 4:

Listing 4. R code showing a linear regression model of mean velocity with main-channel slope along with other variables as defined in Listing 1 by using a simplified form of Manning's equation.

```
summary(lm(logV ~ logR + logMS)) # summarize the regression

Call:
lm(formula = logV ~ logR + logMS)
Residuals:
     Min   1stQua   Median   3rdQua      Max
-1.56869 -0.14867  0.01944  0.16235  0.88036
Coefficients:
            Estimate  Std.Error t-value Pr(>|t-value|)
```

```
(Intercept) 0.329955   0.013217   24.96   <2e-16
logR        0.248853   0.004880   51.00   <2e-16
logMS       0.053644   0.004926   10.89   <2e-16
---
Residual standard error: 0.2474 on 17537 degrees of freedom (DF)
Multiple R-squared: 0.1403, Adjusted R-squared: 0.1402
F-statistic:  1431 on 2 and 17537 DF,  p-value: <2.2e-16
```

The results in Listing 4 indicate that both $R_\circ$ and MS appear to be statistically significant predictors of V. The residual standard error (RSE) is about 0.247 base-10 log feet per second. The adjusted R-squared (coefficient of determination) indicates that about 14 percent of the variation in V is explained by these variables. This is obviously not a large value. The coefficient for the base-10 logarithm of MS is about 0.05; this coefficient shows that a log-cycle change in MS changes the base-10 logarithm of V by 5 percent of a log cycle. Thus, with respect to V, MS contributes little variance explanation.

The regression of V as a function of $R_\circ$ and PS along with a regression summary are shown in Listing 5:

Listing 5. R code showing a linear regression model of mean velocity with proximal-channel slope along with other variables as defined in Listing 1 by using a simplified form of Manning's equation.

```
summary(lm(logV ~ logR + logPS)) # summarize the regression

Call:
lm(formula = logV ~ logR + logPS)
Residuals:
     Min   1stQua   Median   3rdQua      Max
-1.56158 -0.15009  0.01803  0.16353  0.84368
Coefficients:
            Estimate  Std.Error t-value Pr(>|t-value|)
(Intercept) 0.283012   0.007814   36.22   <2e-16
logR        0.231958   0.004307   53.86   <2e-16
logPS       0.030348   0.002346   12.93   <2e-16
---
Residual standard error: 0.247 on 17537 degrees of freedom (DF)
Multiple R-squared: 0.1427, Adjusted R-squared: 0.1426
F-statistic:  1459 on 2 and 17537 DF,  p-value: <2.2e-16
```

The results in Listing 5 indicate that $R_\circ$ and PS appear to be statistically significant predictors of V. The residual standard error (RSE) is about 0.247 base-10 log feet per second. The adjusted R-squared indicates that about 14 percent of the variation in V is explained by these variables. As with MS, this is obviously not a large value.

The two regression equations of V by using a simplified form of Manning's equation (Listings 4 and 5) are

$$V = 2.14 R_\circ^{0.249} MS^{0.0536} \qquad \text{RSE=0.247, adjusted R-squared} = 0.14, \text{ and} \tag{3}$$

$$V = 1.92 R_\circ^{0.232} PS^{0.0303} \qquad \text{RSE=0.247, adjusted R-squared} = 0.14. \tag{4}$$

Regression analysis of V indicates that MS or PS have very little explanatory power as indicated by the small adjusted R-squared values and near-zero values for the regression coefficients (about 0.05 and 0.03 for MS and PS, respectively).

The general applicability of eqs. 3 and 4 in Texas is difficult to assess and of minor interest because this Chapter concerns the evaluation of MS and PS. Summary statistics of the residuals (base-10 logarithm) are presented in Listings 4 and 5. Compared to eqs. 3 and 4, a more sophisticated regression model for estimation of V values associated with direct-runoff conditions is provided in Asquith and others (2013); in particular, that study included additional predictor variables along with

substantial documentation of data processing and interpretation.

A.5.2 Linear Regression of Discharge Using a Simplified Form of Manning's Equation

In this section, linear regression of Q by using a simplified form of Manning's equation is presented. Manning's equation for discharge (eq. 5 when eq. 1 is multiplied by cross-sectional area A in square feet) demonstrates the relation between discharge Q, *n-value*, hydraulic radius R, and dimensionless friction slope S_f:

$$Q = \frac{1.486}{\textit{n-value}} A R^{2/3} S_f^{1/2} \quad \text{[English units]}. \tag{5}$$

To reiterate some key aspects of variables in eq. 5 as discussed in Section A.5.1: (1) the *n-value* for individual discharge measurements is not determined; (2) the R is the pseudo-hydraulic radius $R_\circ$; and (3) the S_f is considered equal to either MS or PS. Because A is included in the linear regression of Q reported in this section, perspective is provided by considering the relation between drainage area (contributing) and A shown in figure 5. Although there is considerable variation between A and drainage area, the figure indicates that A increases with drainage area in a weakly linear fashion in double-logarithmic (log-log) space.

The regression of Q as a function of A, $R_\circ$, and MS along with a regression summary are presented in Listing 6:

Listing 6. R code demonstrating various linear regression models of discharge with other variables as defined in Listing 1 by using a simplified form of Manning's equation.

```
summary(lm(logQ ~ logA + logR + logMS)) # summarize the regression

Call:
lm(formula = logQ ~ logA + logR + logMS)
Residuals:
     Min   1stQua   Median   3rdQua      Max
-1.56053 -0.14892  0.02369  0.16372  0.84469
Coefficients:
            Estimate  Std.Error t-value Pr(>|t-value|)
(Intercept) 0.275901   0.013854   19.91   <2e-16
logA        1.052676   0.004241  248.23   <2e-16
logR        0.171124   0.007908   21.64   <2e-16
logMS       0.067720   0.005034   13.45   <2e-16
---
Residual standard error: 0.2462 on 17536 degrees of freedom (DF)
Multiple R-squared: 0.9343, Adjusted R-squared: 0.9343
F-statistic: 8.31e+04 on 3 and 17536 DF,  p-value: <2.2e-16
```

The results in Listing 6 indicate that A, $R_\circ$, and MS appear to be statistically significant predictors of Q. The residual standard error (RSE) is about 0.246 base-10 log cubic feet per second. The adjusted R-squared indicates that about 93 percent of the variation in Q is explained by these variables.

The regression of Q as a function of A, $R_\circ$, and PS along with a regression summary are presented in Listing 7:

Listing 7. R code demonstrating various linear regression models of discharge with other variables as defined in Listing 1 by using a simplified form of Manning's equation.

```
summary(lm(logQ ~ logA + logR + logPS)) # summarize the regression
Call:
lm(formula = logQ ~ logA + logR + logPS)
Residuals:
     Min   1stQua   Median   3rdQua      Max
```

```
-1.55355 -0.14983  0.02038  0.16496  0.80826
Coefficients:
            Estimate  Std.Error t-value Pr(>|t-value|)
(Intercept) 0.214183   0.010511   20.38   <2e-16
logA        1.040347   0.004132  251.80   <2e-16
logR        0.167222   0.007884   21.21   <2e-16
logPS       0.030571   0.002339   13.07   <2e-16
---
Residual standard error: 0.2463 on 17536 degrees of freedom (DF)
Multiple R-squared: 0.9342,      Adjusted R-squared: 0.9342
F-statistic: 8.305e+04 on 3 and 17536 DF,  p-value: <2.2e-16
```

The results in Listing 7 indicate that A, $R_{\circ}$, and MS appear to be statistically significant predictors of Q. The residual standard error (RSE) is about 0.246 base-10 log cubic feet per second. The adjusted R-squared indicates that about 93 percent of the variation in Q is explained by these variables. The two regression equations of Q by using a simplified form of Manning's equation (Listings 6 and 7) are:

$$Q = 1.888 A^{1.053} R_{\circ}^{0.171} MS^{0.0677} \qquad \text{RSE=0.246, adjusted R-squared} = 0.93, \text{ and} \qquad (6)$$

$$Q = 1.638 A^{1.040} R_{\circ}^{0.167} PS^{0.0306} \qquad \text{RSE=0.246, adjusted R-squared} = 0.93. \qquad (7)$$

Regression analysis of Q indicates that MS or PS have very little explanatory power as indicated by the near-zero values for the regression coefficients (about 0.07 and 0.03 for MS and PS, respectively).

The potential applicability of the two discharge equations (eqs. 6 and 7) is larger relative to the applicability associated the velocity equations (eqs. 3 and 4). This judgment is made in part because the addition of another predictor variable and the substantial increase in adjusted R-squared (about 14 percent of variation to 93 percent of variation). However, the error (as measured in RSE) in estimates of V obtained by using eqs. 6 or 7 by dividing an estimated Q by A would remain about 0.247 base-10 feet per second as is the RSE for eqs. 3 and 4.

The general applicability of eqs. 6 and 7 in Texas is difficult to assess and of minor interest because this Chapter concerns the evaluation of MS and PS. Summary statistics of the residuals (base-10 logarithm) are presented in Listings 6 and 7. Compared to eqs. 6 and 7, a more sophisticated and preferred regression model for estimation of Q values associated with direct-runoff conditions is provided in Asquith and others (2013); in particular, that study included additional predictor variables along with substantial documentation of data processing and interpretation.

The conclusion from the regression analysis of V and Q is that MS or PS have very little explanatory power. Also, the slopes MS and PS are not interchangeable in eqs. 3, 4, 6, and 7 because each slope has its own unique regression coefficient that results because each slope has its own characteristic magnitude and variation in Texas.

A.6 Commentary on Main-Channel Slope and Proximal Slope for Statistical Regionalization of Discharge Measurements Associated with Direct-Runoff Conditions in Texas

The analyses presented in this Chapter include ANOVA and linear regression examples evaluating the statistical importance or explanatory power of MS and PS for statistical regionalization of discharge measurements associated with direct-runoff conditions in Texas. The ANOVA results indicate that PS does not provide more explanatory power than MS for prediction of either V or Q. The hypothesis that the PS should provide more explanatory power than MS, which was evaluated by using a simplified form of Manning's equation, is therefore incorrect.

The answer to the question posed in Section A.2 as to whether either MS or PS are important

variables for predicting mean velocity or discharge values from USGS discharge measurements associated with direct-runoff conditions in Texas is no. This is substantiated by the near-zero magnitudes of the regression coefficients on *MS* and *PS* in eqs. 3, 4, 6, and 7. There is the possibility that, although shown to be statistically significant, both *MS* and *PS* are fitted to the noise of the underlying data. Neither *MS* or *PS* proved useful for the statistical regionalization of discharge measurements associated with direct-runoff conditions in Texas.

References

Archuleta, C.M., Gonzales, S.L., and Maltby, D.R., II, 2012, Automated delineation and characterization of watersheds for more than 3,000 surface-water-quality monitoring stations active in 2010 in Texas: U.S. Geological Survey Open-File Report 2012–1077, 33 p.

Asquith, W.H., 2001, Effects of regulation on L-moments of annual peak streamflow in Texas: U.S. Geological Survey Water-Resources Investigations Report 01–4243, 66 p., at `http://pubs.usgs.gov/wri/wri014243/`.

Asquith, W.H., 2011, A proposed unified rational method for Texas *in* Cleveland, T.G., Fang, X., and Thompson, D.B., 2011, Use of the rational and modified rational methods for TxDOT hydraulic design: Texas Department of Transportation Research Report 0–6070–1, section 3, pp. 18–58., at `http://library.ctr.utexas.edu/hostedPDFs/TechMRT_0-6070-1.pdf`.

Asquith, W.H., 2013, Regression models of discharge and mean velocity associated with median-streamflow conditions from a U.S. Geological Survey discharge measurement database in Texas: *ASCE Journal of Hydrologic Engineering*, 10.1061/(ASCE)HE.1943-5584.0000715.

Asquith, W.H., Thompson, D.B., Cleveland T.G., and Fang, Xing, 2005, Unit hydrograph estimation for applicable Texas watersheds: Texas Department of Transportation Research Report 0–4193–4, Texas Tech University Center for Multidisciplinary Research in Transportation, Lubbock, 71 p., at `http://library.ctr.utexas.edu/pdf/4193-4.pdf`.

Asquith, W.H., Cleveland, T.G., and Roussel, M.C., 2011, A method for estimating peak and time of peak streamflow from excess rainfall for 10- to 640-acre watersheds in the Houston, Texas, metropolitan area: U.S. Geological Survey Scientific Investigations Report 2011–5104, 38 p., at `http://pubs.usgs.gov/sir/2011/5104`.

Asquith, W.H., Herrmann, G.R., and Cleveland, T.G., 2013, Generalized additive regression models of discharge and mean velocity associated with direct-runoff conditions in Texas—The Utility of the U.S. Geological Survey discharge measurement database: ASCE Journal of Hydrologic Engineering, 10.1061/(ASCE)HE.1943-5584.0000635].

Asquith, W.H., and Roussel, M.C., 2007, An initial-abstraction, constant-loss model for unit hydrograph modeling for applicable watersheds in Texas: U.S. Geological Survey Scientific Investigations Report 2007–5243, 82 p., at `http://pubs.usgs.gov/sir/2007/5243/`.

Asquith, W.H., and Roussel, M.C., 2009, Regression equations for estimation of annual peak-streamflow frequency for undeveloped watersheds in Texas using an L-moment-based, PRESS-minimized, residual-adjusted approach: U.S. Geological Survey Scientific Investigations Report 2009–5087, 48 p., at `http://pubs.usgs.gov/sir/2009/5087/`.

Asquith, W.H., and Slade, R.M., 1997, Regional equations for estimation of peak-streamflow frequency for natural basins in Texas: U.S. Geological Survey Water-Resources Investigations Report 96–4307, 68 p., at http://pubs.usgs.gov/wri/wri964307/.

Barnes, H.H., 1967, Roughness characteristics of natural channels: U.S. Geological Survey Water-Supply Paper 1849, 213 p., at http://pubs.usgs.gov/wsp/wsp_1849/.

Benson, M.A., and Dalrymple, Tate, 1967, General field and office procedures for indirect measurements: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A1, 30 p., at http://pubs.usgs.gov/twri/twri3-a1/

Castro, J.M., and Jackson, P.L., 2001, Bankfull discharge recurrence intervals and regional hydraulic geometry relationships: Patterns in the Pacific Northwest, USA: Journal American Water Resources Association, v. 37, no. 5, pp. 1249–1262.

Cleveland, T.G., He, Xin, Asquith, W.H., Fang, Xing, and Thompson, D.B., 2006, Instantaneous unit hydrograph selection for rainfall-runoff modeling of small watersheds in north and south central Texas: ASCE Journal of Irrigation and Drainage Engineering, Sept.–Oct. 2006, pp. 479–485, at http://dx.doi.org/10.1061/(ASCE)0733-9437(2006)132:5(479).

Dalrymple, Tate, and Benson, M.A., 1967, Measurement of peak discharge by the slope-area method: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A2, 12 p., at http://pubs.usgs.gov/twri/twri3-a2/.

Environmental Systems Research Institute, 2010, ArcGIS Desktop, Release 10: Redlands, Calif., Environmental Systems Research Institute, accessed December 2012 at http://resources.arcgis.com/content/arcgisdesktop/10.0/about.

Federal Geographic Data Committee, 2013, National Spatial Data Infrastructure: Reston, Va., accessed April 2013 at at http://www.fgdc.gov/nsdi/nsdi.html.

Gesch, D.B., 2007, The National Elevation Dataset, in Maune, D., ed., Digital Elevation Model Technologies and Applications: The DEM Users Manual, 2nd edition: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, pp. 99–118.

Gordon, N.D., McMahon, T.A., Finlayson, B.L., Gippel, C.J., and Nathan, R.J., 2004, Stream hydrology—An introduction for ecologists (second edition): John Wiley, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 429 p.

Heitmuller, F.T., Asquith, W.H., Cleveland, T.G., Fang, Xing, and Thompson, D.B., 2006, Estimation of main channel length from basin length for selected small watersheds in Texas using 10- and 30-meter digital elevation models: Southwestern Geographer, Journal of the Southwestern Division of the Association of American Geographers, v. 10, pp. 23–35, at http://www2.geog.okstate.edu/swaag/SWgeographer.htm.

Homer, C., Huang, C., Yang, L., Wylie, B., and Coan, M., 2004, Development of a 2001 National land cover database for the United States: Photogrammetric Engineering and Remote Sensing, v. 70, no. 7, pp. 829–840, at http://www.mrlc.gov/nlcd2001.php.

Koltun, G.F., Kula, S.P., and Puskas, B.M., 2006, A Streamflow Statistics (StreamStats) Web Application for Ohio: U.S. Geological Survey Scientific Investigations Report 2006–5312, 62 p., at http://pubs.usgs.gov/sir/2006/5312/pdf/sir2006-5312.pdf

Leopold, L.B., Wolman, M.G., and Miller, J.P., 1964, Fluvial process in geomorphology: San Francisco, Freeman.

Metcalf, C.K., Wilkerson, S.D., and Harman, W.A., 2009, Bankfull regional curves for north and northwest Florida streams: Journal American Water Resources Association, v. 45, no. 5, pp. 1260–1272.

R Development Core Team, 2012, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, ISBN 3–900051–07–0, at http://www.R-project.org/.

Rantz, S.E., and others, 1982a, Measurement and computation of streamflow—Volume 1, Measurement of stage and discharge: U.S. Geological Survey Water-Supply Paper 2175, [variously paged], at http://pubs.usgs.gov/wsp/WSP2175/.

Rantz, S.E., and others, 1982b, Computation of discharge—Volume 2, Measurement of stage and discharge: U.S. Geological Survey Water-Supply Paper 2175, [variously paged], at http://pubs.usgs.gov/wsp/WSP2175/.

Richards, K.S., 1982, Rivers, Form and process in alluvial channels: London, Methuen.

Sougnez, N., Vanacker, V., 2010, Spatial variability in channel and slope morphology within the Ardennes Massif, and its link with tectonics: Hydrology and Earth System Sciences Discussions, v. 7, no. 5, pp. 6981–7006.

Sturm, T.W., 2010, Open channel hydraulics: 2nd ed., New York, McGraw-Hill.

Turnipseed, D.P., and Sauer, V.B., 2012, Discharge measurements at gaging stations: U.S. Geological Survey Techniques and Methods, Book 3, Chapter A8, [variously paged], at http://pubs.usgs.gov/tm/tm3-a8/.

U.S. Geological Survey, 2011, National hydrography dataset: U.S. Geological Survey, at http://nhd.usgs.gov/index.html.

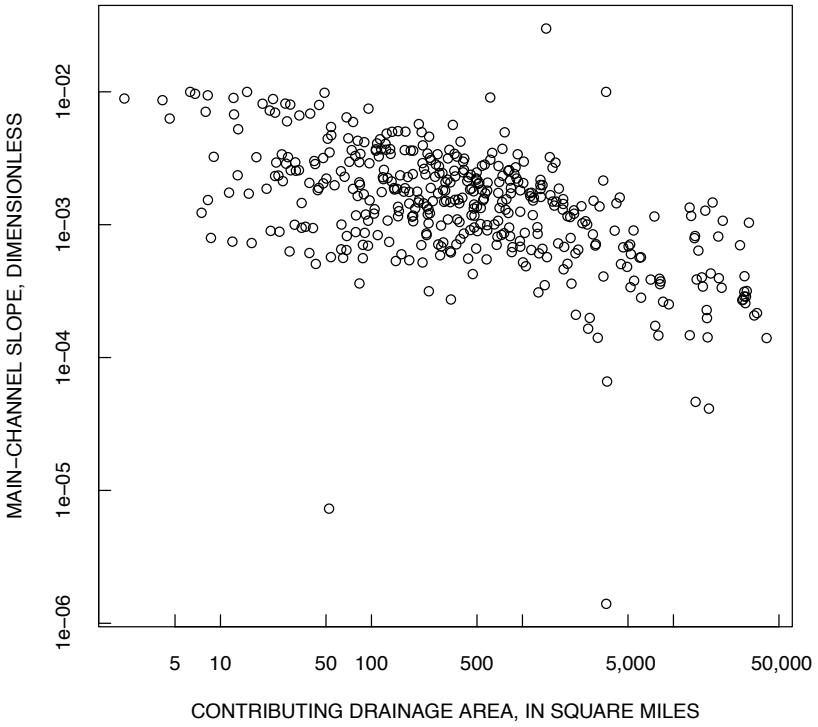

Figure 1. Relation between watershed drainage area and main-channel slope for about 420 selected U.S. Geological Survey streamflow-gaging stations (streamgages) in Texas.

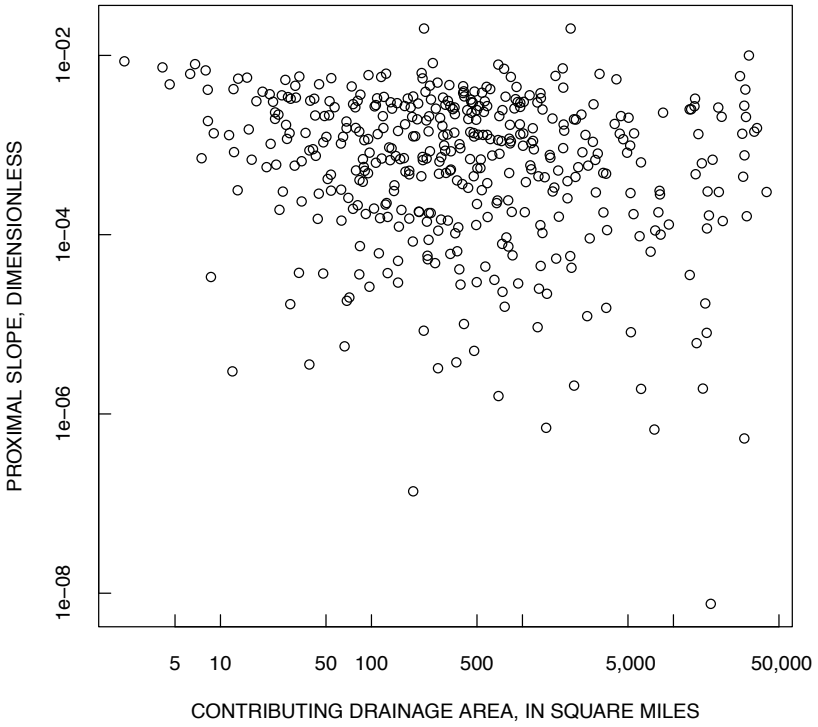

Figure 2. Relation between watershed drainage area and proximal channel slope for about 420 selected U.S. Geological Survey streamflow-gaging stations (streamgages) in Texas.

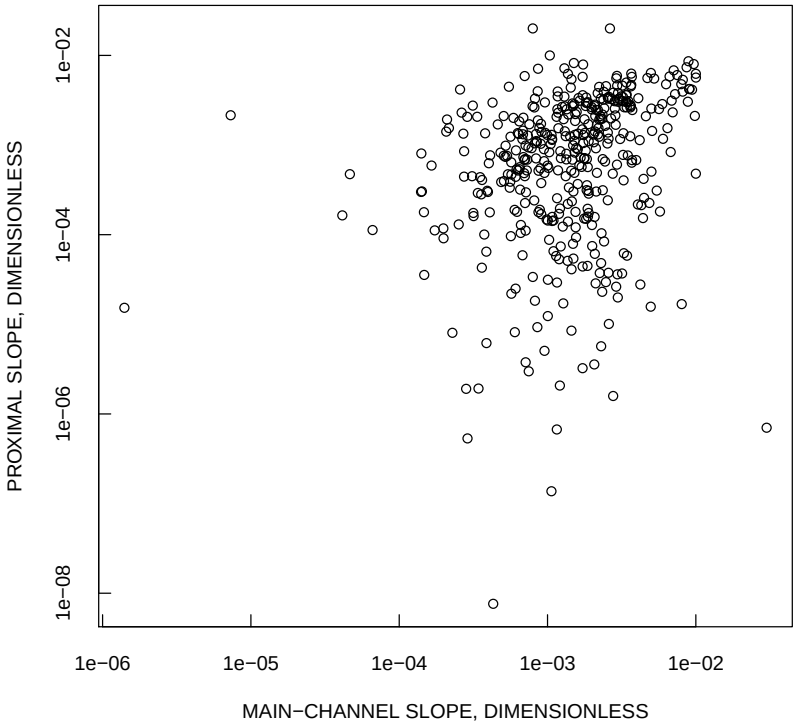

Figure 3. Relation between main-channel slope and proximal channel slope for about 420 selected U.S. Geological Survey streamflow-gaging stations (streamgages) in Texas.

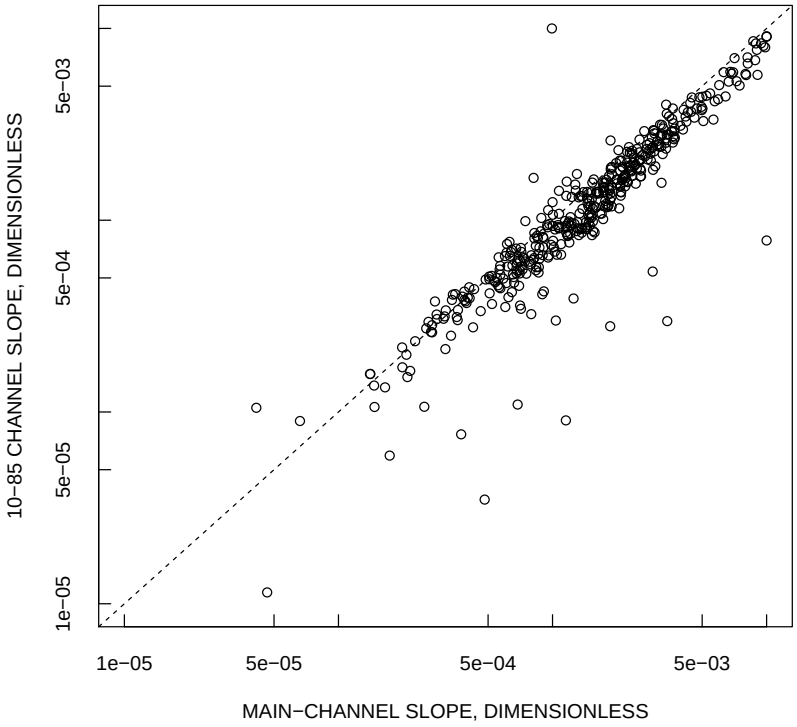

Figure 4. Relation between main-channel slope and 10-85 channel slope for about 420 selected U.S. Geological Survey streamflow-gaging stations (streamgages) in Texas. Dashed line represents an equal value line.

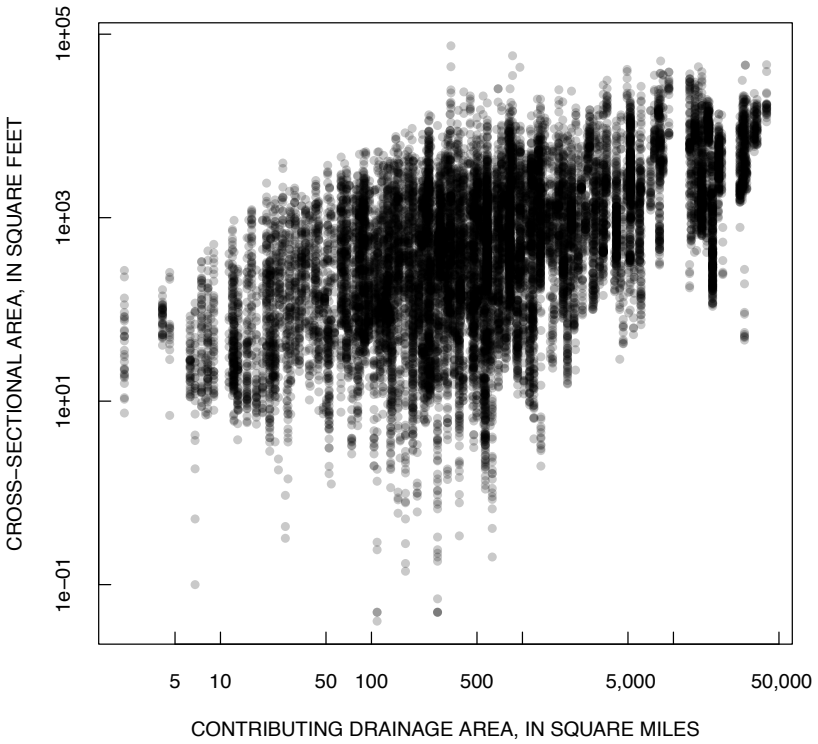

Figure 5. Relation between contributing drainage area and cross-sectional area of discharge measurement associated with direct-runoff conditions for discharge measurements made at about 420 selected U.S. Geological Survey streamflow-gaging stations (streamgages) in Texas.

B Regionalized Ancillary Parameters

In an attempt to tease out any possible regional dependence in the ancillary parameters, each gaging station was given three different physiographic region tags following three different classification schemes. The first classification used was based off of the physiographic map of Texas produced by the Bureau Of Economic Geology (BEG) at the University Of Texas at Austin in 1996. The map breaks Texas into 9 main physiographic regions (figure 61). Each gaging station in the database was tag with one of the region tags shown in figure 61 and added to the database by researchers at the University of Texas at San Antonio.

We provided two additional more finely discretized physiographic region classifications based off of maps provided by the Texas Parks and Wildlife Department (TPWD). The first TPWD classifications had 11 different regions and is similar to the BEG classification. The second TPWD classification has 35 Omernik labeled natural regions. Each gaging station in the database was then given an integer classification number to correspond with each of the two physiographic region classification schemes based off of the TPWD data. The physiographic region tags were added to the text file database for each measurement. Maps and region tags for these two classifications along with box plots of ancillary parameter broken down according to these two regionalization schemes are presented in sections B.2 and B.3.

The ancillary properties were sorted according to each of the three regional classification schemes and plotted as box plots for the two subset databases containing greater than 90th percentile flows (storm flows) and those falling between the 40th and 60th percentiles (mean flow).

Tag	Region
BR	Basin and Range
HP	High Plains
EP	Edwards Plateau
NCP	North-Central Plains
CTU	Central Texas Uplift
GP	Grand Prairie
BP	Blackland Prairies
ICP	Interior Coastal Plains
CP	Coastal Prairies

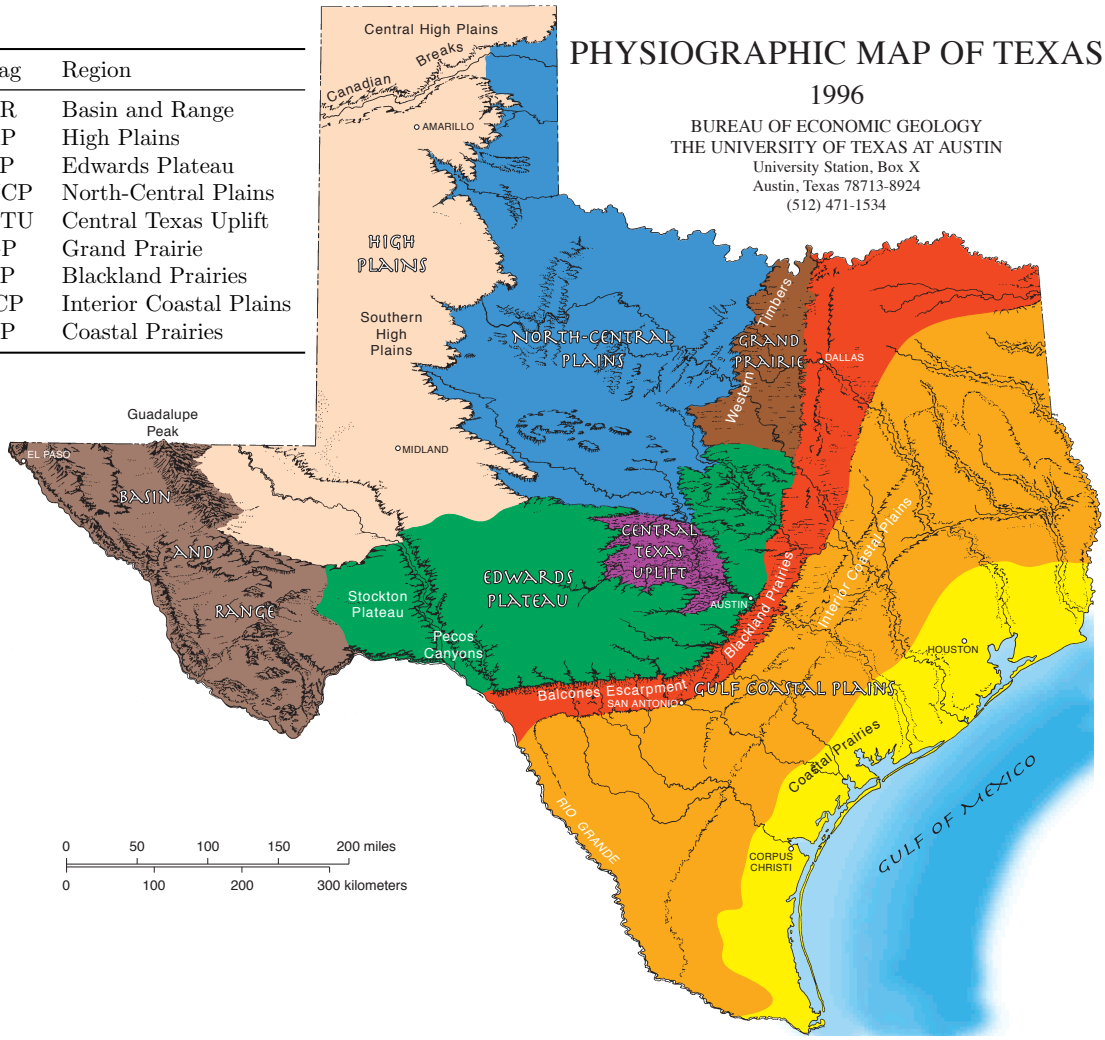

Figure 61: Physiographic map of Texas produced by the Bureau Of Economic Geology.

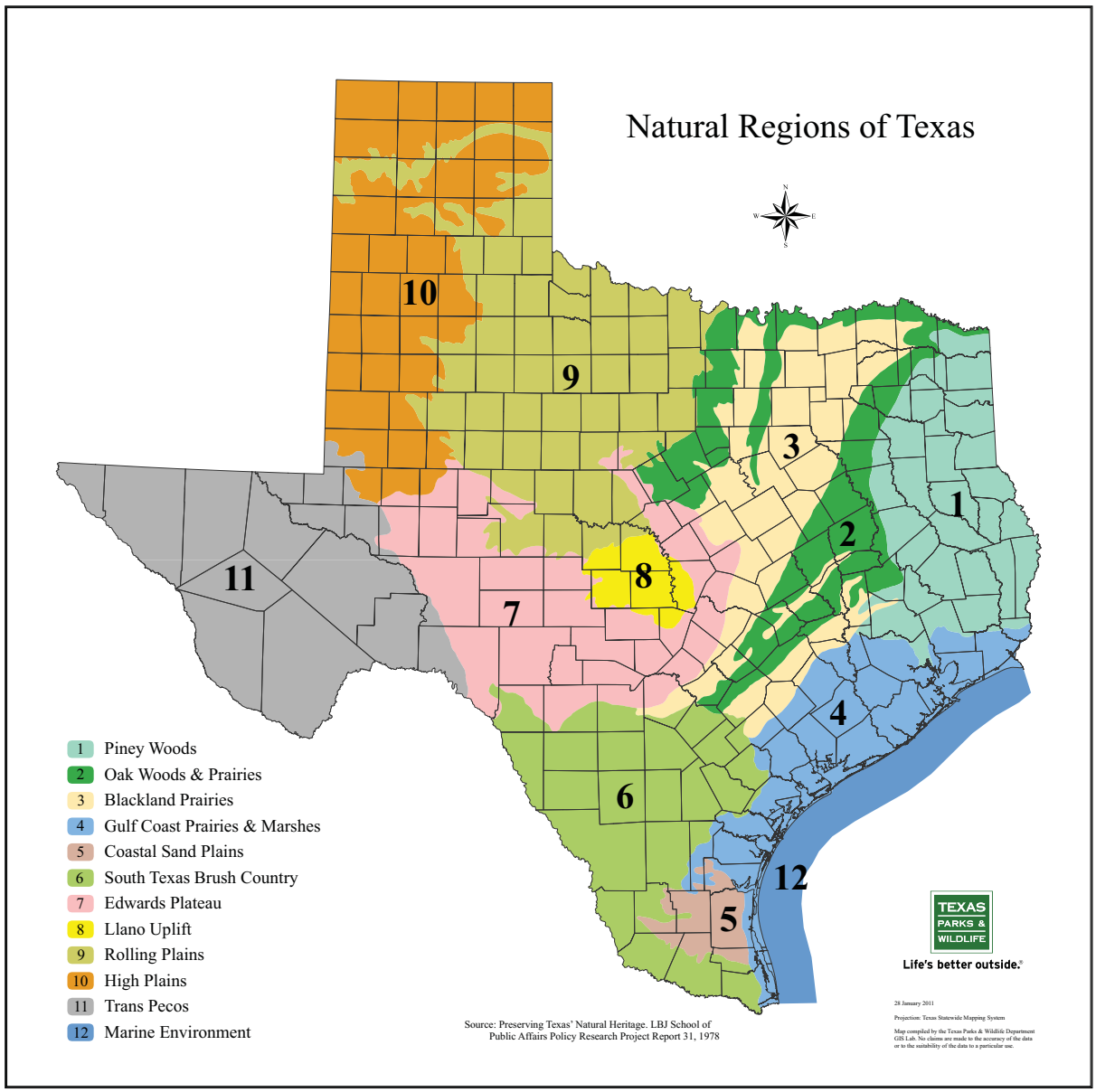

Figure 62: Texas Parks and Wildlife Department coarser discretized physiographic region map.

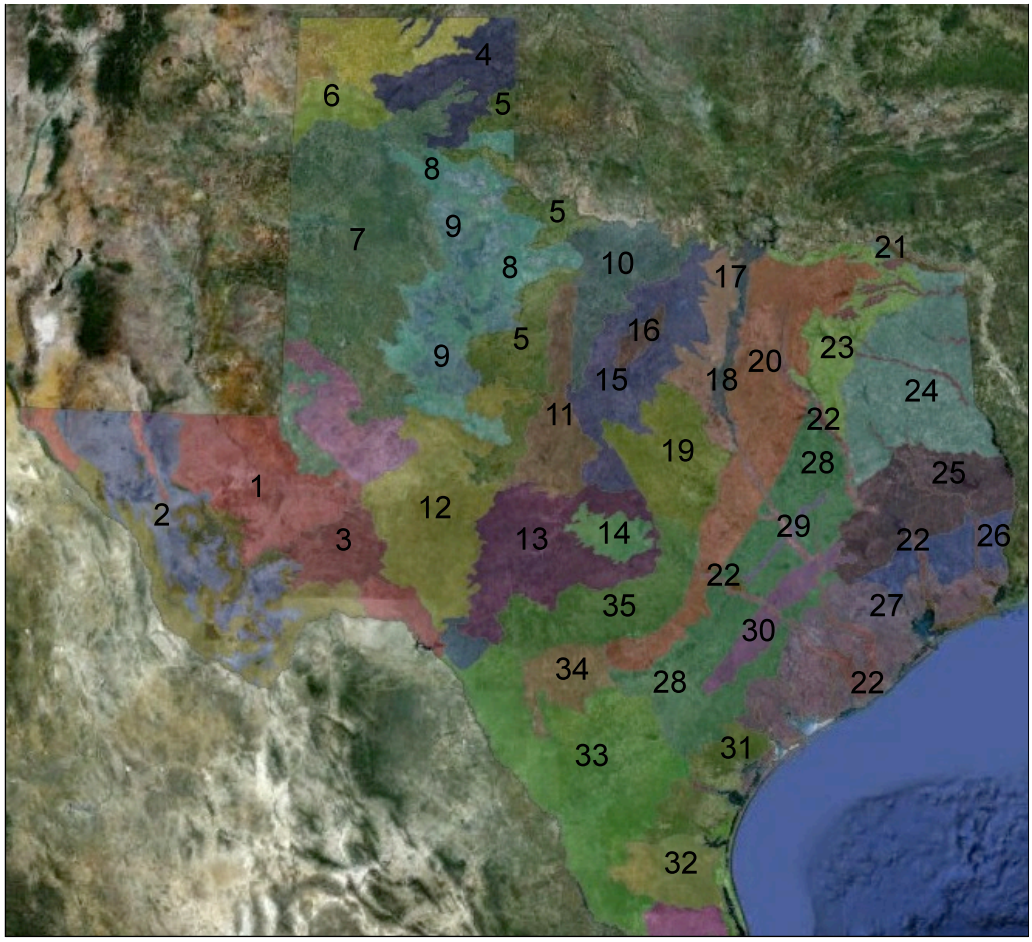

Figure 63: Texas Parks and Wildlife Department finely discretized physiographic region map.

Fine designated region number	Region description
1	Chihuahuan Basins and Playas
2	Chihuahuan Desert Grasslands
3	Stockton plateau
4	Canadian/Cimarron Breaks
5	Red Prairie
6	Semiarid Canadian Breaks
7	Llano Estacado
8	Caprock Canyons, Badlands, and Breaks
9	Flat Tablelands and Valleys
10	Broken Red Plains
11	Limestone Plains
12	Semiarid Edwards Plateau
13	Edwards Plateau Woodland
14	Llano Uplift
15	Western Cross Timbers
16	Carbonate Cross Timbers
17	Grand Prairie
18	Eastern Cross Timbers
19	Limestone Cut Plain
20	Northern Blackland Prairie
21	Red River Bottomlands
22	Floodplains and Low Terraces
23	Northern Post Oak Savanna
24	Tertiary Uplands
25	Southern Tertiary Uplands
26	Flatwoods
27	Northern Humid Gulf Coastal Prairies
28	Southern Post Oak Savanna
29	San Antonio Prairie
30	Southern Blackland/Fayette Prairie
31	Southern Subhumid Gulf Coastal prairies
32	Coastal Sand Plain
33	Texas-Tamaulipan Thornscrub
34	Northern Nueces Alluvial Plains
35	Balcones Canyonlands

Table 10: Numeric tags for the TPWD finely discretized physiographic regions.

B.1 Ancillary parameters by region using the BEG designation

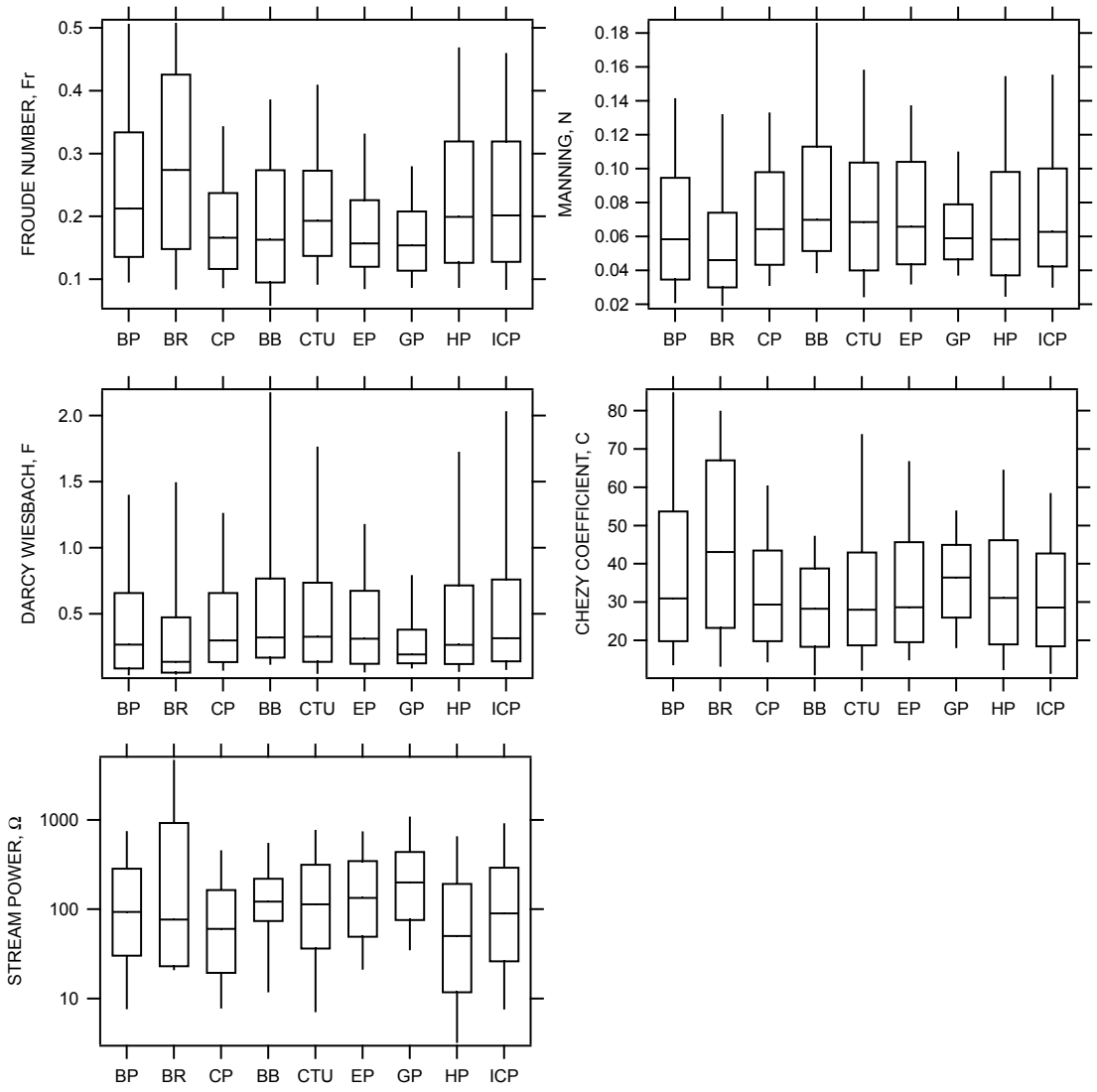

Figure 64: Box plots of ancillary parameters broken down by BEG regions (figure 61) for flows great than the 90th percentile (storm flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

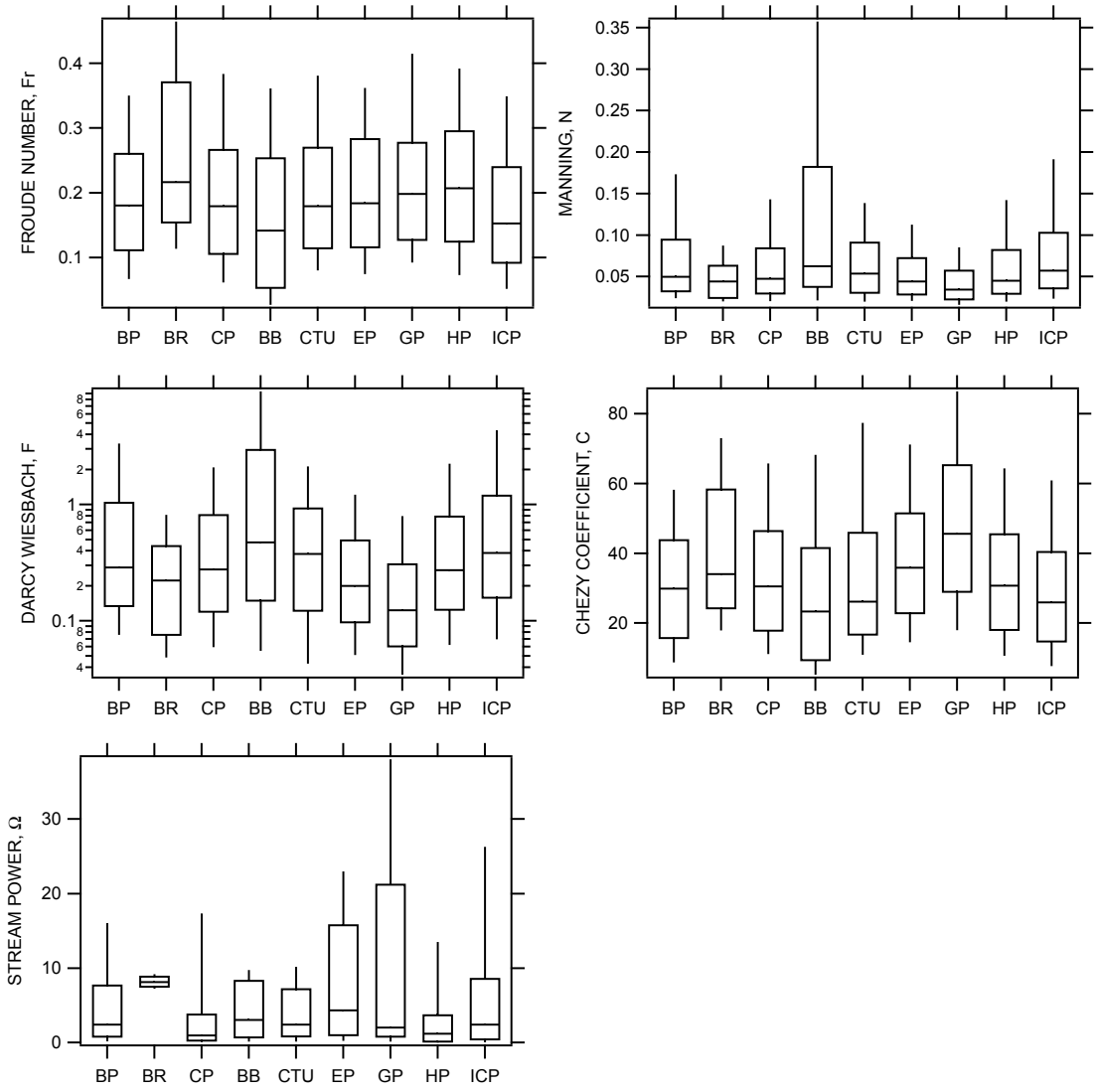

Figure 65: Box plots of ancillary parameters broken down by BEG regions (figure 61) for flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

B.2 Ancillary parameters by region using the coarse TWPD designations

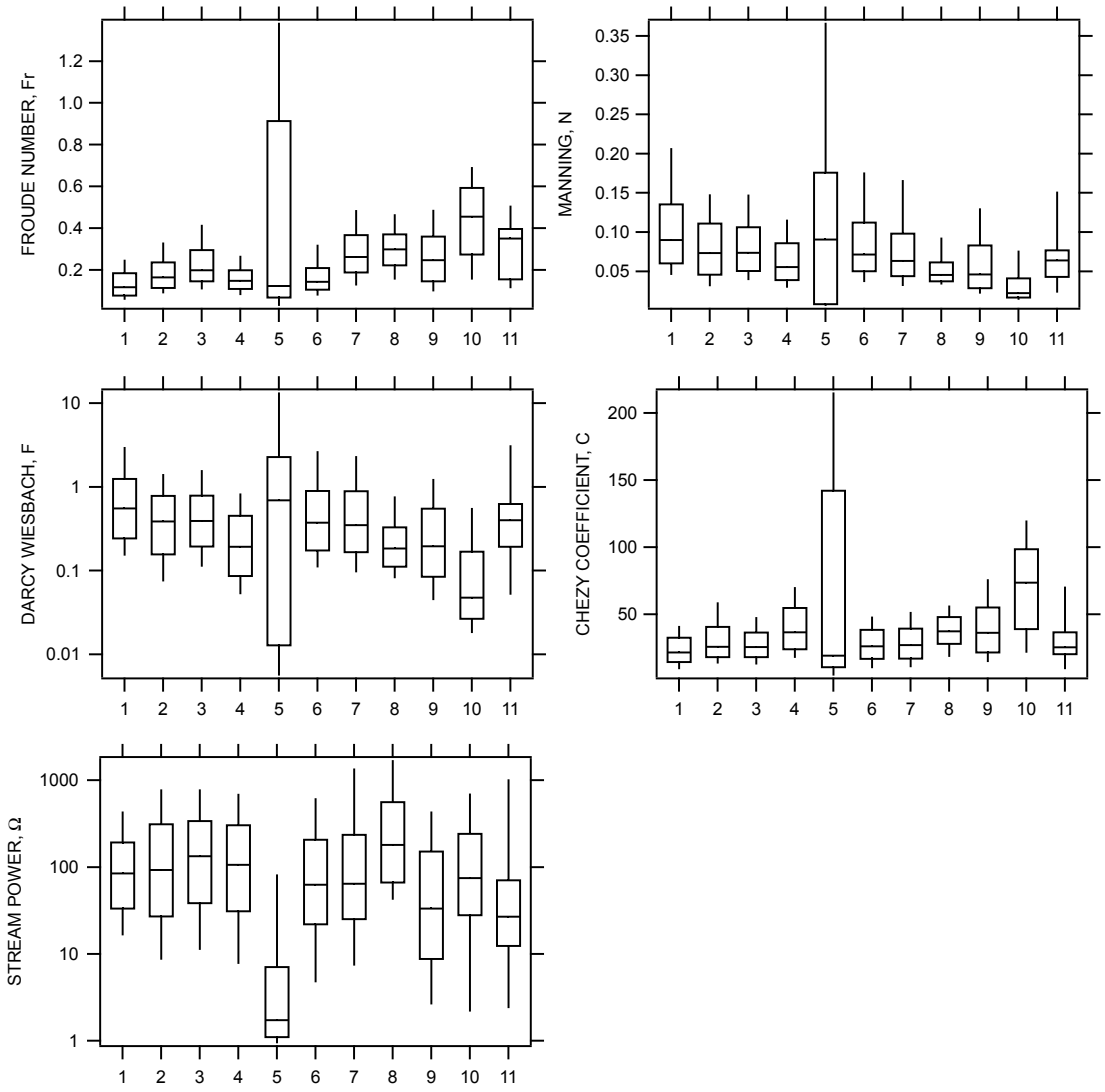

Figure 66: Box plots of ancillary parameters broken down by coarse TWPD regions (figure 62) for flows great than the 90th percentile (storm flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

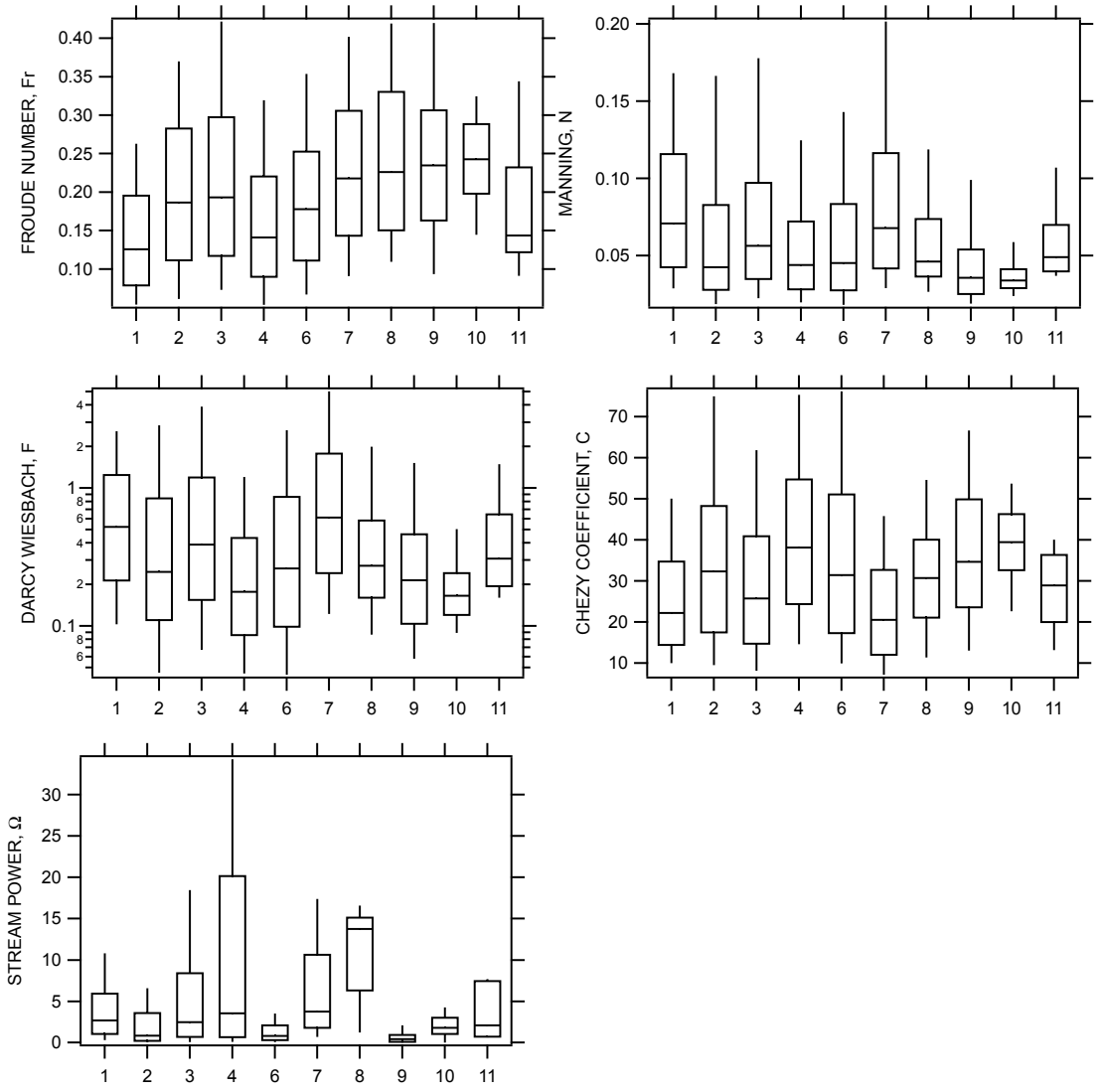

Figure 67: Box plots of ancillary parameters broken down by coarse TWPD regions (figure 62) for flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

B.3 Ancillary parameters by region using the fine TWPD designations

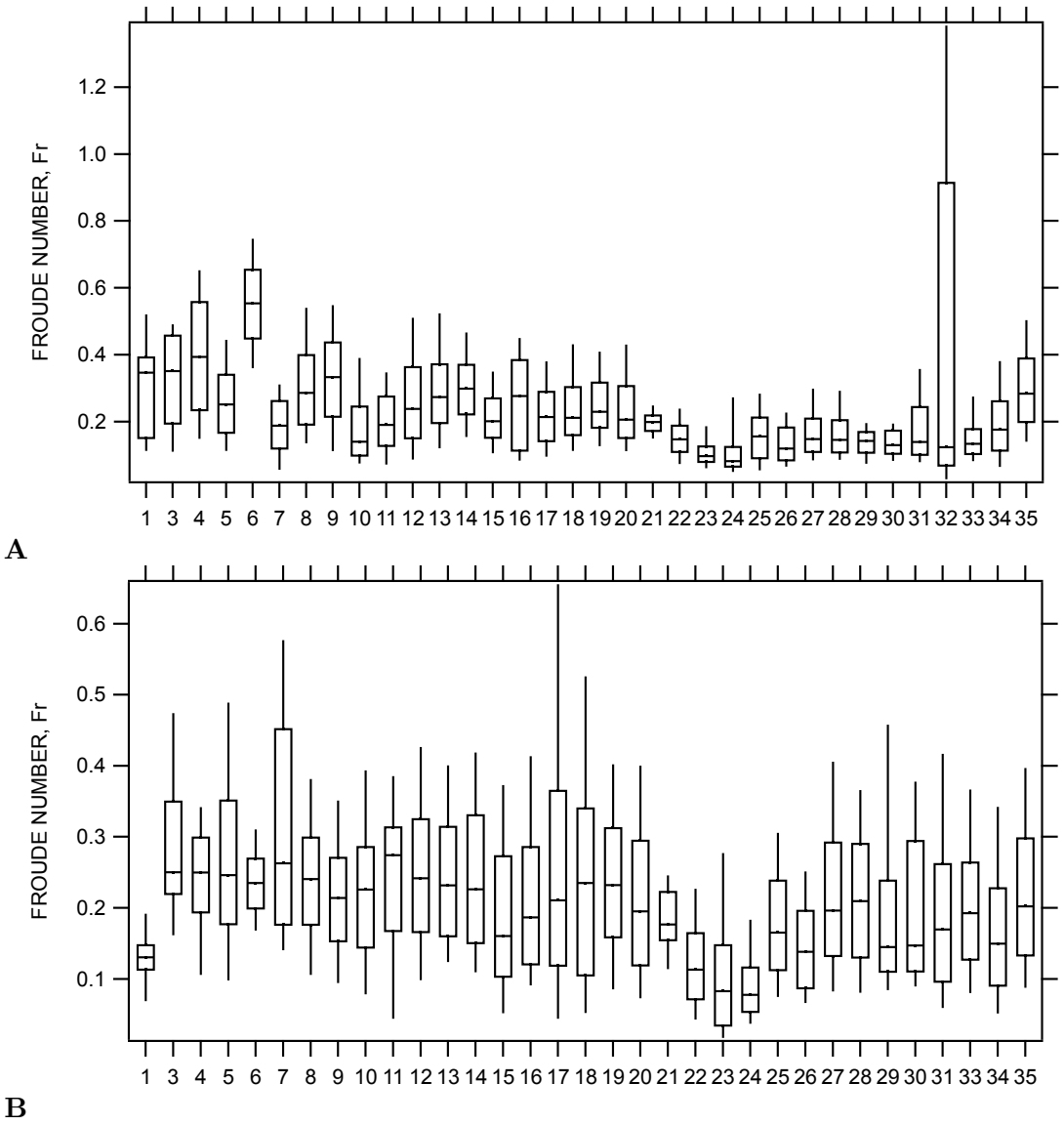

Figure 68: Box plots of Froude number broken down by fine TWPD regions (figure 63 and table 10) for (A) flows over the 90th percentiles (storm flow), and (B) flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

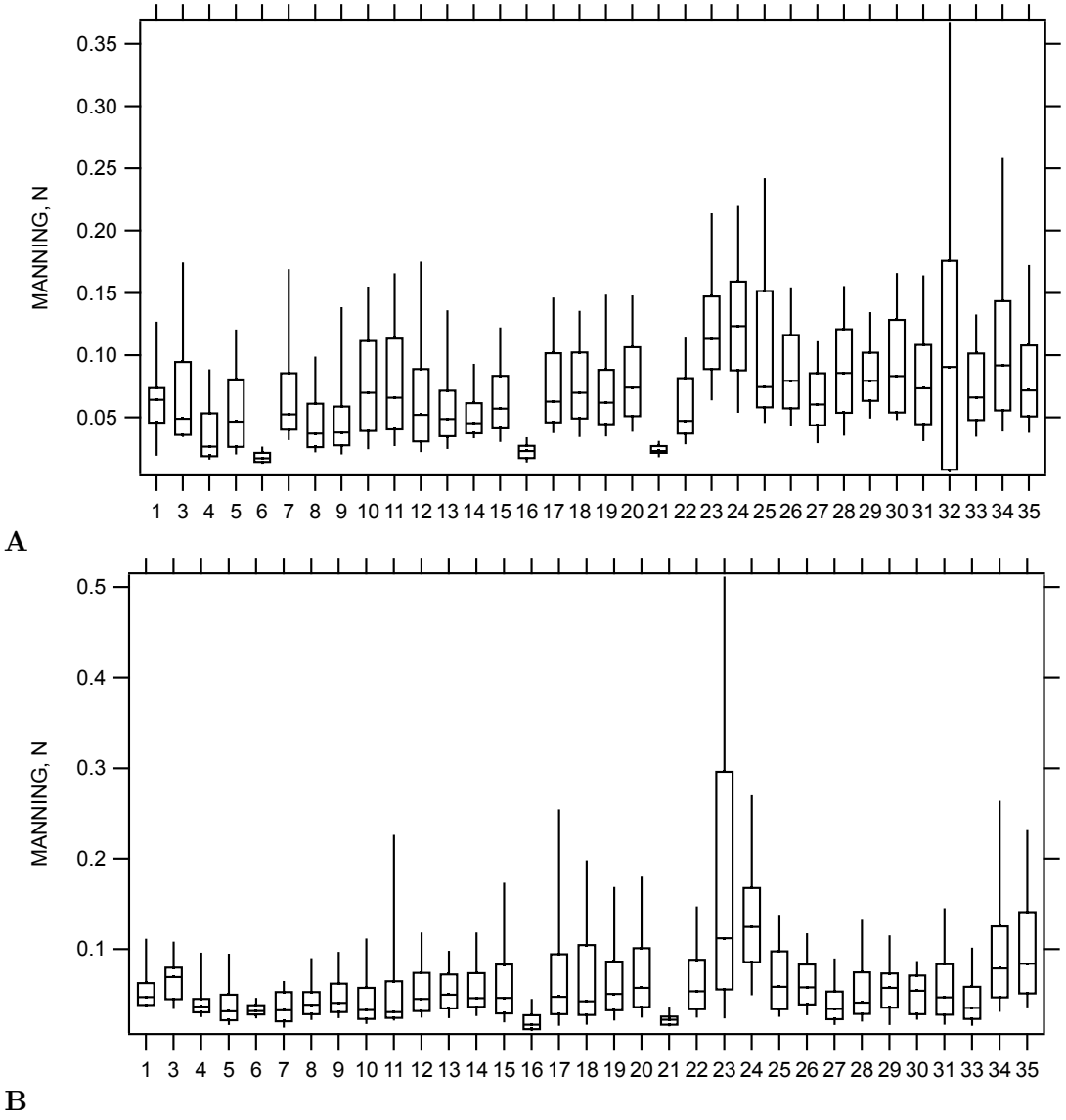

Figure 69: Box plots of Manning n broken down by fine TWPD regions (figure 63 and table 10) for (A) flows over the 90th percentiles (storm flow), and (B) flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

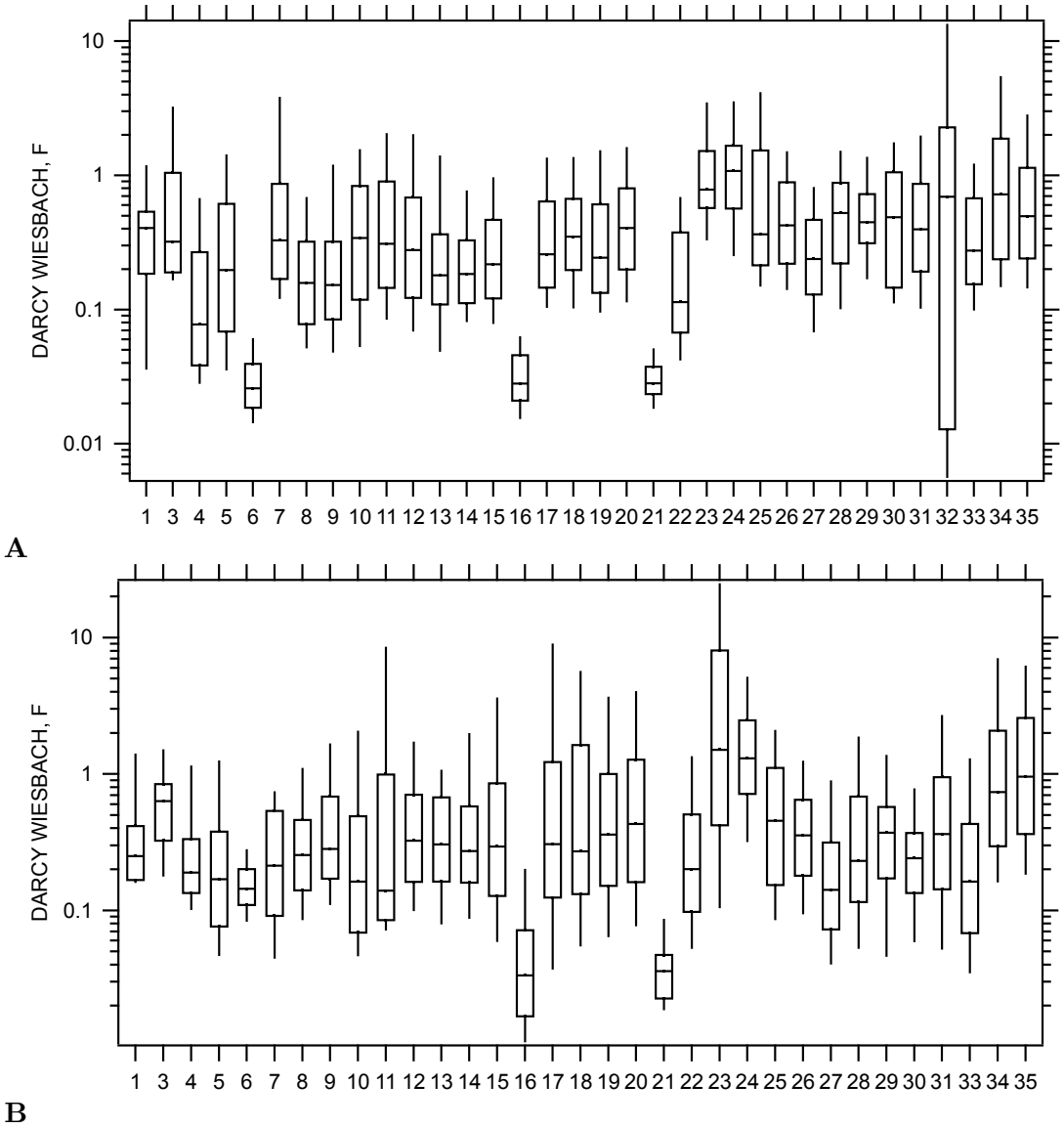

Figure 70: Box plots of Darcy Weisbach friction factor f broken down by fine TWPD regions (figure 63 and table 10) for (A) flows over the 90th percentiles (storm flow), and (B) flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

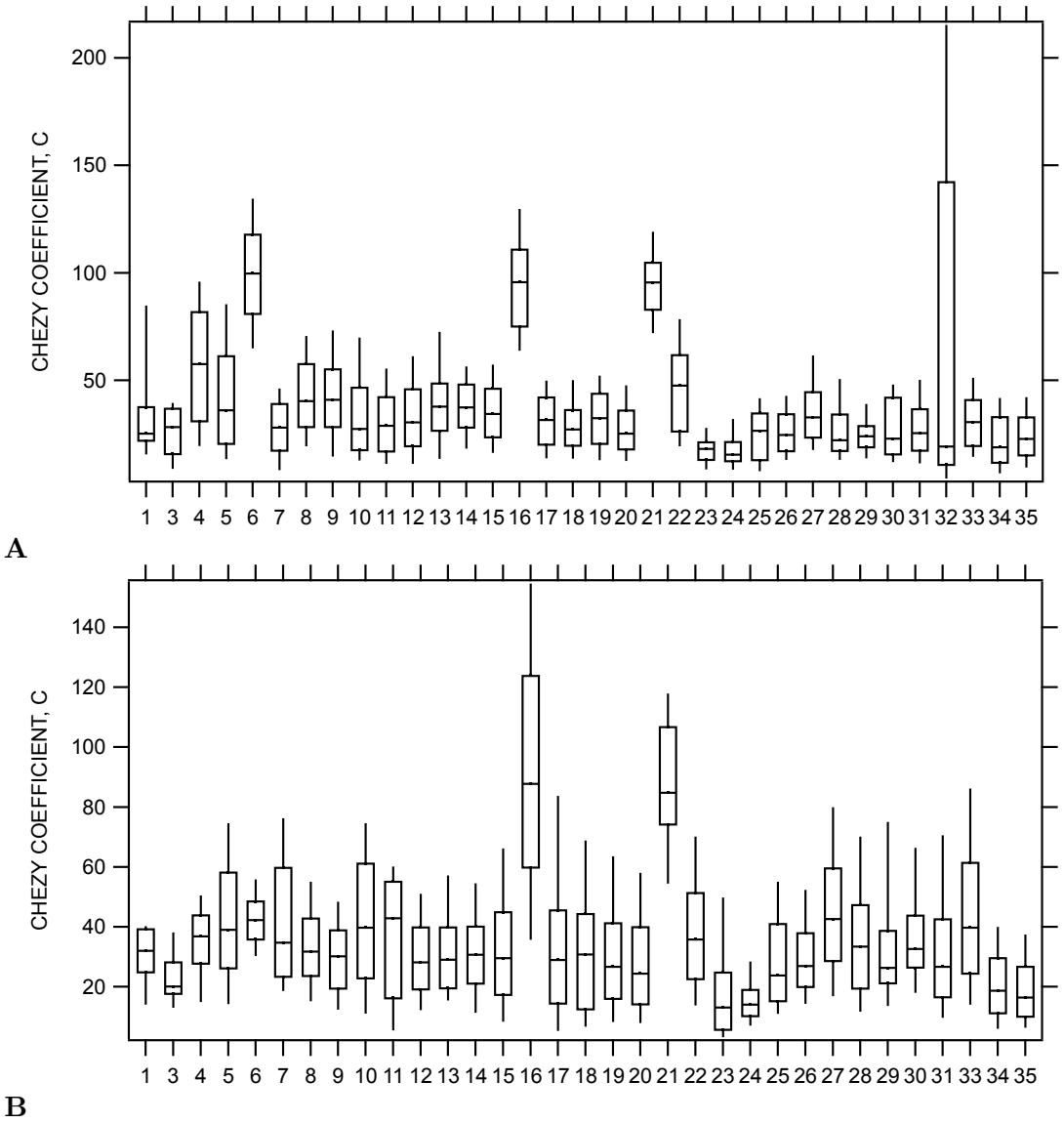

Figure 71: Box plots of the Chezy friction coefficient C broken down by fine TWPD regions (figure 63 and table 10) for (A) flows over the 90th percentiles (storm flow), and (B) flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

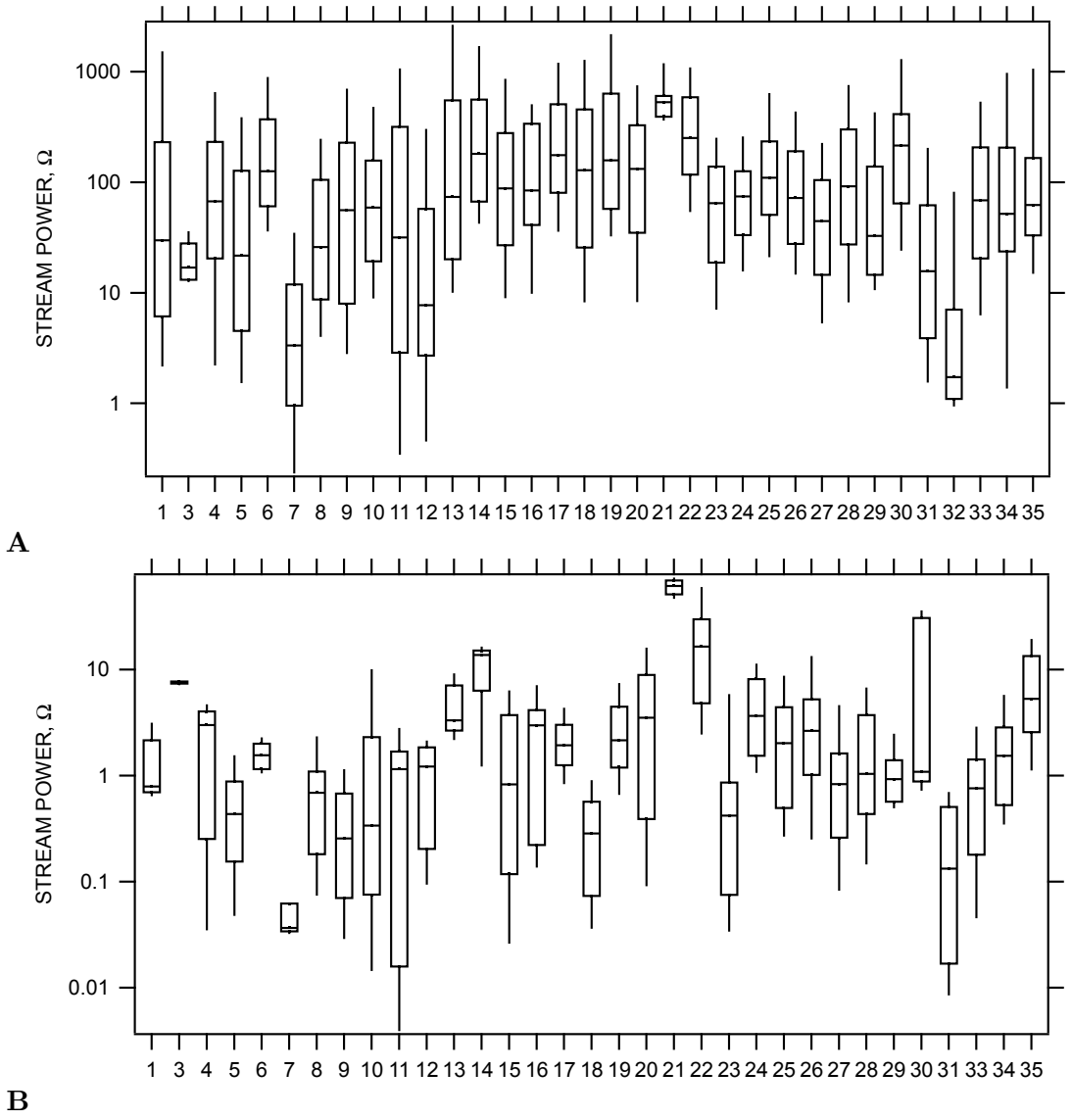

Figure 72: Box plots of stream power Ω broken down by fine TWPD regions (figure 63 and table 10) for (A) flows over the 90th percentiles (storm flow), and (B) flows between the 40th and 60th percentiles (mean flow). Boxes show 25th, 50th, and 75th percentiles; whisker ends are the 10th and 90th percentiles.

C Theoretical $V = V(Q)$ Distributions for Compound Channels

C.1 Introduction

In the course of a project meeting, a question was raised as to whether or not it was physically possible for cross-sectionally averaged velocity, V, to be inversely related to discharge, Q. The question arose because some of the station data in the database exhibited this type of functionality. The propose of the following analysis was to see if it was theoretically possible to produce an inverse relationship between velocity and discharge, and if so, to see what types of conditions produced and strengthen such a relationship.

To do this, the cross-sectionally averaged velocity was computed using the Manning equation (equation 30) in a compound channel that consisted of a main channel with flanking floodplains (figure 73). Discharged was varied so that the flow depth ranged from being confined within the main channel to well out onto the floodplain.

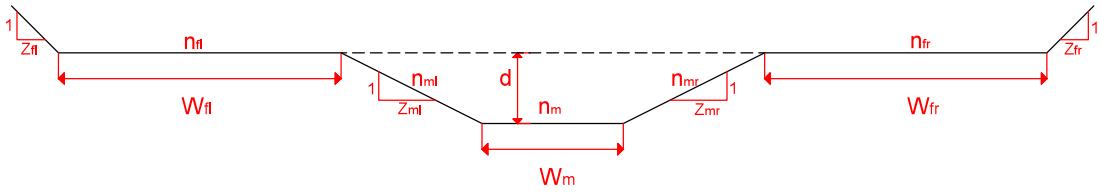

Figure 73: Theoretical channel with definitions for the parameters varied in the analysis.

C.2 Solution method and parameter space

The main channel through which flow was routed is a trapezoidal channel with a main channel, floodplains on either side, and two different cross side slopes (figure 73). The Manning coefficient was given different values for the channel bottom and cross slopes to represent a range of natural roughness settings. The floodplains are located on the right and left of the main channel; each can have a separate length and cross side slope. Similarly, different Manning coefficients could be defined for the right and left floodplain. In figure 73, the subscripts m and f on the parameters refer to the main channel and floodplain respectively, while the l and r subscripts denote left or right. For instance, n_{fr} represents the Manning coefficient of the right side of the floodplain. The bankfull condition is defined as the flow depth that fills the main channel, i.e. d in figure 73.

For the analysis, Q is given and the Manning equation is used to calculate the depth and hence cross-sectional area. From this the velocity at the given discharge is calculated ($V = Q/A$). Default values are given to the parameters (table 11), then each parameter is systematically varied (table 12) and depth is solved for over a range of within and over-bank

discharges. The result of this analysis is a large family of $V = V(Q)$ curves as a function of various boundary conditions. In all cases, river slope was set to $S = 0.0001$.

W_m [m]	n_m	n_{mr}	n_{ml}	Z_{mr}	Z_{ml}	d [m]	W_{fr} [m]	W_{fl} [m]	n_{fr}	n_{fl}	Z_{fr}	Z_{fl}
5	0.025	0.035	0.03	3	2	6	10	10	0.04	0.036	3	3

Table 11: Default parameter values for the channel. Parameters are defined in figure 73.

Q [m^3/s]	W_m [m]	n_m,n_{mr},n_{ml}	Z_{mr},Z_{ml}	d [m]	W_{fr},W_{fl}	n_{fr},n_{fl}	Z_{fr},Z_{fl}
1-300	7-49	0.01-0.07	0-12	1-15	0-30	0.014-0.098	4-16

Table 12: Range of conditions used in the systematic variation. Parameters are defined in figure 73.

For each discharge, depth is calculated from the Manning equation using Newton's method:

$$h_{n+1} = h_n - \frac{f(h_n)}{f'(h_n)} \tag{65}$$

The equivalent Manning coefficient for the section was determined using the composite roughness formula of Motayed and Krishnamurthy (1980):

$$n_c = \frac{PR^{5/3}}{\sum_{i=1}^{N} \frac{P_i {R_i}^{5/3}}{n_i}} \tag{66}$$

in which P_i, R_i, n_i are the wetted perimeter, hydraulic radius, and Manning's n value for any section i, P and R are the wetted perimeter and hydraulic radius of the entire cross section, and N is the number of sections in which the wetted perimeter is divided. Figure 74 illustrates the way in which the channel was divided. The area and wetted perimeters for the main channel and floodplain are split by dashed lines. The Manning equation is solved by the combining the main channel and floodplain channels, i.e. $Q = \Sigma Q_i$.

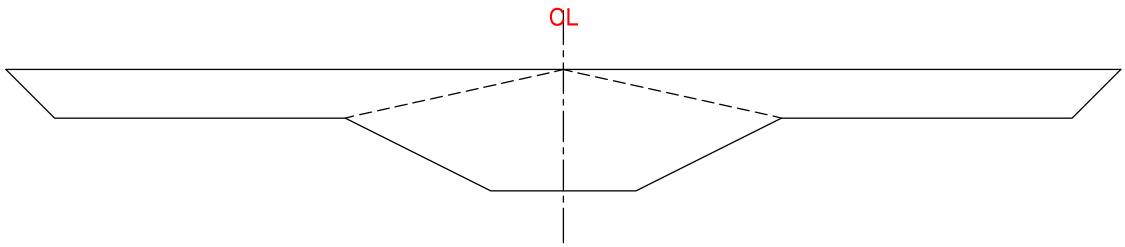

Figure 74: Divided compound channel with diagonal lines

C.3 Results

The two questions this analysis sought to answer were: (1) is it theoretically possible for cross sectionally-averaged velocity, V, and discharge, Q, to be inversely related?; and (2) if so, what channel conditions produce this functionality? From the analysis, it was found that channel velocity can decrease with increasing discharge. A decrease from the maximum velocity at bankfull condition to some reduced velocity can occur once the flow spills out onto the floodplain. This does not occur of all channel configurations, but it does occur for many.

To aid in the interpretation of the results, the bankfull discharge Q_b was defined as the discharge that produced a depth equal to the depth of the main centralized channel, d (figure 73). The velocity at this discharge and depth is defined as the bankfull velocity, V_b. Dips in velocity which produce the inverse relationship between V and Q occur only for discharges greater than Q_b.

It was found that the dip in velocity with increasing discharge once flows spill out onto the floodplain is made stronger by the following:

1. The smaller the main channel is relative to the floodplain or larger compound channel, the stronger the dip in velocity with increasing Q (figure 75)
2. The larger the ratio between the floodplain and main channel Mannings coefficient, the larger the dip in the velocity with increasing Q (figures 76 and 77); note, the channel should have lower n-values and floodplain larger ones.
3. The more rectangular the main channel geometry is, the stronger the dip (figure 78).
4. The narrower and deeper the main channel is relative to the floodplain channel, the stronger and more pronounced the dip (figure 79)

For most conditions, velocity and discharge are negatively correlated in the range of discharges from Q_b to $3Q_b$ but can be as high as $5Q_b$ under extreme conditions (figure 80), and the dip in velocity can range from 0.3-0.9 times the bankfull velocity (max velocity).

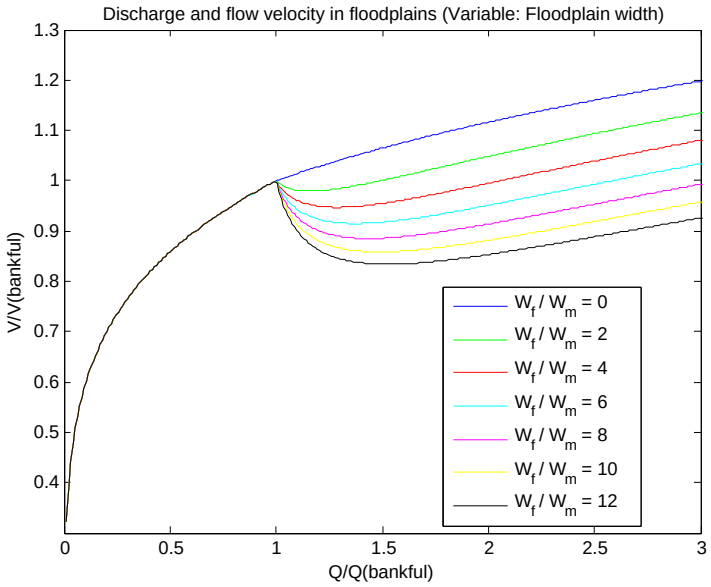

Figure 75: Velocity-discharge relationship as a function of the ratio between the main channel width and the floodplain width.

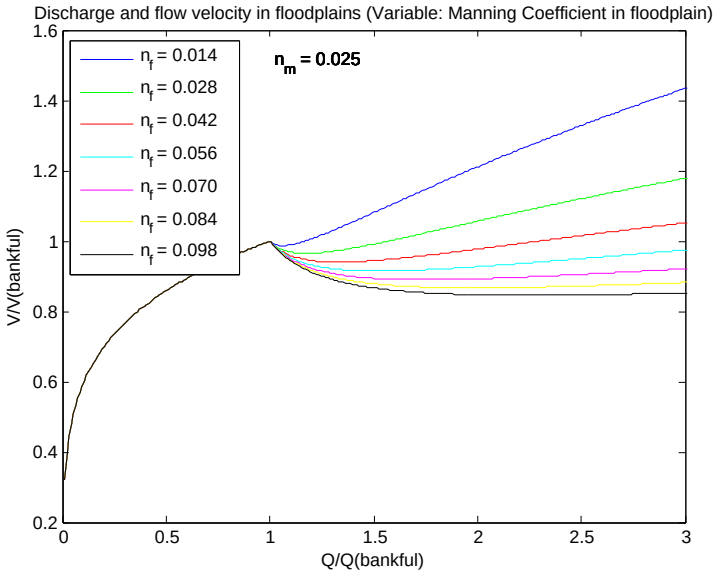

Figure 76: Velocity-discharge relationship as the Manning's n-value on the floodplain changes relative to that of the main channel. All other parameters are held constant.

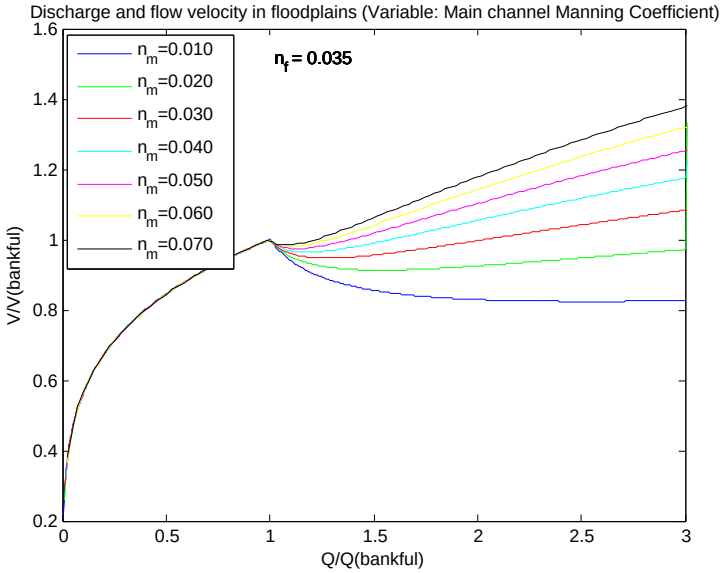

Figure 77: Velocity-discharge relationships for varying main channel Manning's n-value and holding the floodplain value constant. All other parameters are held constant.

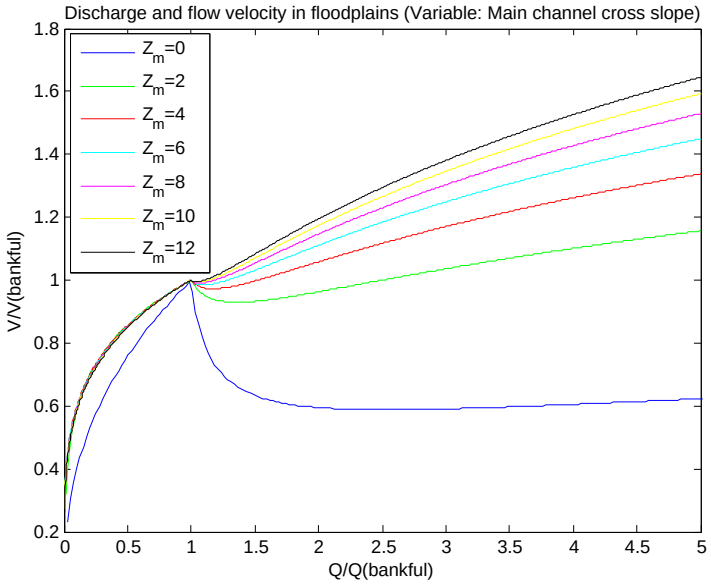

Figure 78: Velocity-discharge relationships for different main channel geometries that are obtained through changing the main channel size slope angle. All other parameters are held constant.

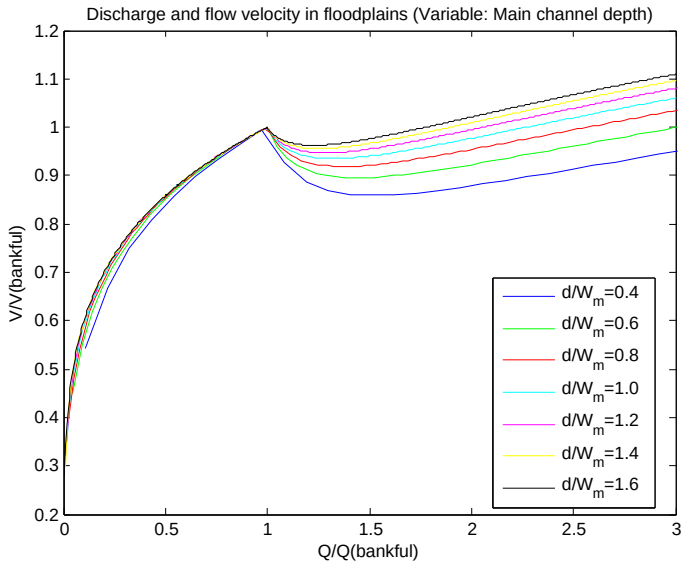

Figure 79: Velocity-discharge relationship for different width-to-depth ratios in the main channel. All other parameters are held constant.

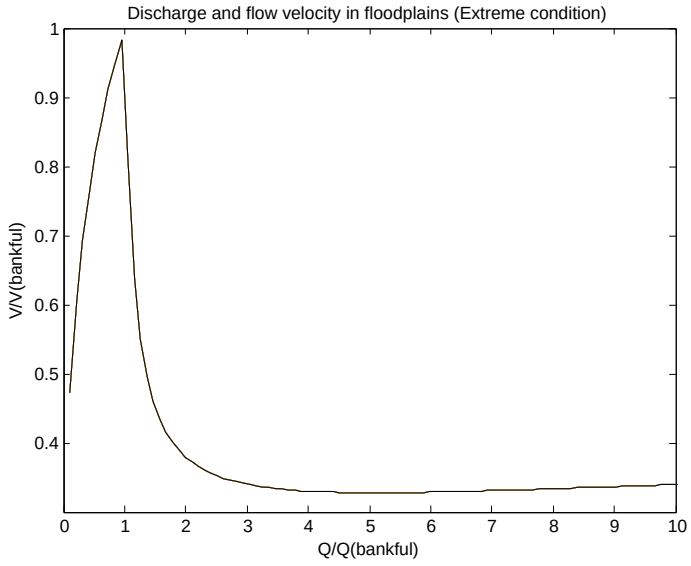

Figure 80: Example of velocity-discharge functionality for an extreme condition. Conditions for the main channel: width = 5 m, depth = 2 m, $n = 0.02$, side slop = 1 (steep); conditions for the floodplain: width = 20 m, Manning coefficient = 0.15, side slope = 10 (very mild).

C.4 Summary

The ancillary parameters calculated from the database included the Froude number, the Manning's n, the Darcy-Weisbach friction factor, the Chezy resistance coefficient, and stream power. The calculated values are presented as state-wide distributions and as distributions broken out by physiographic regions using three different Texas regionalization classifications. The ancillary parameters for the state-wide and regional distributions were calculated for two subsets of the master database associated with this project. For the first subset, ancillary parameters were calculated and presented for measurements whose daily flow discharges were greater than the 90th percentile. This subset roughly corresponds to storm flow conditions. The second subset of data from which the ancillary parameters were calculated was for mean flow conditions with daily discharges between the 40th and 60th percentiles. The distributions are visualized using percent nonexceedance plots and box plots. Little difference in the overall distribution of ancillary parameters Fr, n, f, and C with flow type (storm or mean) was found. However, a larger than would be expected range of values for each parameter overall was present in some of the data. This was especially true in calculated resistance coefficient data (n, f, and C). The very large range of computed values is likely due to some possible errors in the original measurements, the use of the hydraulic depth in place of the hydraulic radius, and the use of the DEM derived main channel slope. The use of the hydraulic depth and main channel slope in the calculations is thought to have produced a bias towards larger Manning n and Darcy-Weisbach friction coefficients and a reduction in the Chezy coefficient. The stream power, Ω, is thought to be less sensitive to errors in the assumptions since it is linearly related to the discharge (a measured value) and does not need the hydraulic radius. The main channel slope however was used in computation of the stream power and $\Omega \propto S_{mc}$ in the calculation.

The analysis of the theoretical functionality of cross-sectionally averaged velocity with discharge in compound channels showed that it is possible to produce an inverse relationship between velocity and discharge. The inverse relation only occurs once what spills out of the main channel and onto the floodplains or larger compound channel. The dip in velocity with increasing discharge is made stronger with an increasing disparity between the main channel and floodplain roughness and with decreasing main channel size and width to depth ratio relative to the floodplains or larger compound channel.

TEXAS TECH UNIVERSITY
Multidisciplinary Research in Transportation

Texas Tech University | Lubbock, Texas 79409
P 806.742.3503 | F 806.742.4168

Disclaimer

The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.