Technical Report

CRWR 270

Suspended Sediment Yield in Texas Watersheds

by

JULIA E. ALLRED COONROD, Ph.D.

EDWARD R. HOLLEY, Ph.D.

and

DAVID R. MAIDMENT, Ph.D.

GEORGE H. WARD, Ph.D.

August 1998

CENTER FOR RESEARCH IN WATER RESOURCES

Bureau of Engineering Research • The University of Texas at Austin

ACKNOWLEDGEMENTS

This project was supported by the Texas Water Development Board (TWDB) under Contract No. 96-483-148. Dr. Bill Longley was the Contract Manager for TWDB. The authors would like to express sincere gratitude to Dr. Longley for his unselfish assistance with both the administrative and the technical aspects of the project. In addition, Roger L. Wolff of the Environmental Section staff spent several days keypunching and checking the daily data for the Lavaca River prior to 1965. Mike Lynn, James Mirabal, Bobby Critendon, and Gary Laneman, all with TWDB, provided valuable insights into TWDB's sediment sampling program.

At the University of Texas at Austin, Dr. Daene McKinney and Dr. Jack Sharp reviewed the report and made beneficial suggestions about the content and presentation.

ABSTRACT

The Texas Water Development Board collected suspended sediment samples across the state of Texas for approximately 60 years. Until this research, no comprehensive analysis of the data had been conducted. This study compiles the suspended sediment data along with corresponding streamflow and rainfall. GIS programs are developed which characterize watersheds corresponding to the sediment gauging stations. The watersheds are characterized according to topography, climate, soils, and land use. All of the data is combined to form several SAS data sets which can subsequently be analyzed using regression.

Annual data for all of the stations across the state are classified temporally and spatially to determine trends in the sediment yield. In general, the suspended sediment load increases with increasing runoff but no correlation exists with rainfall. However, the annual average rainfall can be used to classify the watersheds according to climate, which improves the correlation between sediment load and runoff. The watersheds with no dams have higher sediment loads than watersheds with dams. Dams in the drier parts of Texas reduce the sediment load more than dams in the wetter part of the state. Sediment rating curves are developed separately for each basin in Texas. All but one of the curves fall into a band which varies by about two orders of magnitude.

The study analyzes daily time series data for the Lavaca River near Edna station. USGS data are used to improve the sediment rating curve by the addition of physically related variables and interaction terms. The model can explain an additional 41% of the variability in sediment concentration compared to a simple bivariate regression of sediment load and flow.

The TWDB daily data for the Lavaca River near Edna station are used to quantify temporal trends. There is a high correlation between sediment load and flowrate for the Lavaca River. The correlation can be improved by considering a flow-squared term and by considering seasonal effects. Typically, sediment concentration is the highest during the warmest months. The infrequent high flows carry a large, disproportionate amount of sediment.

Table of Contents

1 INTRODUCTION 1-1
1.1 Purpose Of Research 1-1
1.2 Research Objective 1-2
1.3 Research Contributions 1-4

2 BACKGROUND 2-1
2.1 Sediment Mechanics 2-1
2.2 Historical Work 2-2
2.2.1 Sediment Transport 2-2
2.2.2 Erosion 2-4
2.3 Recent Work 2-6
2.3.1 Sediment Transport 2-6
2.3.2 Erosion 2-10
2.4 Historical Work in Texas 2-11
2.5 TWDB Suspended Sediment Data 2-19
2.6.1 Sampling Method 2-23
2.6.2 Lavaca River Data 2-24
2.6.3 Suspended Sediment Size Analysis 2-27
2.5 Current Work in Texas 2-31

3 RESEARCH METHODS 3-1
3.1 General 3-1
3.2 Collection of temporal data 3-2
3.2.1 Suspended sediment data 3-2
3.2.2 Streamflow data 3-8
3.2.3 Rainfall data 3-10
3.3 Collection of spatial data 3-13
3.3.1 Location of Gauging Stations 3-15
3.3.2 Watershed boundaries 3-19
3.3.3 Watershed Flow Length 3-27
3.3.4 Watershed Soil Types 3-30
3.3.5 Watershed Land Use /Land Cover 3-33
3.3.6 Watershed Rainfall 3-36
3.3.7 Dams and Reservoirs 3-36
3.4 Data Analyses (SAS) 3-37

4 Effects of Gross Watershed Characteristics on Sediment Yield 4-1
4.1 Purpose 4-1
4.2 Data Description and Bivariate Analysis 4-2
  4.2.1 Annual Values 4-4
  4.2.2 Annual Averages 4-5
  4.2.3 Annual Averages with No Dams 4-17
4.3 Data Classifications and Bivariate Analysis 4-20
  4.3.1 Temporal Classifications 4-21
  4.3.2 Spatial Classifications 4-26
4.4 Physical Parameters and Bivariate Analysis 4-38
4.5 Multivariate Analysis 4-47
4.6 Model Selection 4-57
4.7 Model Evaluation 4-61
4.8 Conclusions 4-63

5 Improvement of Sediment Rating Curve using USGS data for the Lavaca River near Edna 5-1
5.1 Purpose 5-1
5.2 Data Description 5-1
5.3 Background and Physical Parameters 5-2
  5.3.1 Sediment Rating Curves - Flow Rate Effects 5-2
  5.3.2 Temperature Effects 5-9
  5.3.3 Size Effects 5-10
  5.3.4 Hydrograph Limb Effects 5-10
5.4 Analysis 5-11
  5.4.1 Multivariate Regression 5-11
  5.4.2 Modeling Software 5-11
  5.4.3 Variable Transformation 5-12
  5.4.4 Omitted data 5-12
  5.4.5 Use of Binary Variables 5-15
  5.4.6 Use of Interaction Terms 5-16
5.5 Model Selection 5-21
5.6 Model Evaluation 5-23
  5.6.1 Warm Effects 5-23
  5.6.2 %fines Effects 5-24
  5.6.3 Flow Effects 5-25
5.7 Conclusions 5-25

6 Daily Sediment Data Analysis for the Lavaca River near Edna 6-1
6.1 Purpose 6-1
6.2 Data Description 6-1
6.3 Watershed Description 6-3
6.4 Bivariate Analysis - "The Sediment Rating Curve" 6-9
6.5 Multivariate Analysis 6-24
  6.5.1 Rainfall Considerations 6-24
  6.5.2 Seasonal Considerations 6-28
  6.5.3 Timing considerations 6-34
6.6 Temporal Analysis 6-35
  6.6.1 Daily graphs 6-36
  6.6.2 Annual Daily Averages 6-37
  6.6.3 Photo review 6-41
6.7 Conclusions 6-43

7 Conclusions 7-1
7.1 Annually Averaged Data for all Stations 7-1
7.2 Annual Data for all Stations 7-1
7.3 Watershed Characteristics 7-2
7.4 Lavaca River near Edna Samples (taken by USGS) 7-3
7.5 Lavaca River near Edna Daily Data (taken by TWDB) 7-3
7.6 Temporal and spatial Implications 7-4
7.7 Future Work 7-5

8 Appendices 8-1
8.1 Appendix A – Fortran programs 8-1
8.2 Appendix B – Arc/Info AML's 8-7
8.3 Appendix C – SAS Programs 8-32
8.4 Appendix D - Values for Watersheds used in SAS 8-65
8.5 Appendix E – Daily Plots for the Lavaca River 8-73

9 References 9-1

Texas Water Development Board Comments

List of Tables

Table 2-1. Sediment Load Data (Table 2, Bulletin 5912) 2-12
Table 2-2. TWDB Suspended Sediment Sampling Stations 2-20
Table 2-3. Comparison of concentrations using two different samplers, Lavaca River near Edna 2-28
Table 2-4. Comparison of percent fines using two different samplers, Lavaca River near Edna 2-29

Table 3-1. Required data 3-1
Table 3-2. Sources of Spatial Data 3-15
Table 3-3. Main Data Sources for Watershed Characteristics 3-16
Table 3-4. Anderson Land Use Codes 3-35
Table 3-5. Level Two Codes for Agricultural Land 3-35

Table 4-1. Study Area Property Variability 4-2
Table 4-2. Basin Property Variability 4-2
Table 4-3. SAS Data Sets 4-3
Table 4-4. Annual Average Values for TWDB Sediment Stations 4-8
Table 4-5. Bivariate regression equations and statistics using average values 4-15
Table 4-6. Bivariate regression equations and statistics using annual values 4-16
Table 4-7. Bivariate regression equations and statistics for stations using average values with no dams 4-20
Table 4-8. Bivariate regression equations and statistics for stations using annual values with no dams 4-20
Table 4-9. Bivariate regression equations and statistics using average values based on at least 5 years of data 4-24
Table 4-10. Bivariate regression equations and statistics using average values based on at least 10 years of data 4-24
Table 4-11. Bivariate regression equations and statistics using annual values for 33 stations with complete records from 1970-1978 4-25
Table 4-12. Bivariate regression equations and statistics using annual values for 33 stations with complete records from 1970-1978 and no dams 4-25
Table 4-13. Temporal Biases 4-26
Table 4-14. Bivariate regression equations and statistics for annual values by river basin 4-29
Table 4-15. Bivariate regression equations and statistics for annual values with no dams, by basin 4-30
Table 4-16. Bivariate regression equations and statistics for annual values by climate 4-37

Table 4-17. Bivariate regression equations and statistics for annual values with no dams, by climate 4-37
Table 4-18. Pearson correlation coefficients for kgload using average values ... 4-42
Table 4-19. Four basic data sets 4-48
Table 4-20. Criteria that reduce the data sets 4-48
Table 4-21. Data Classifications 4-48
Table 4-22. Model comparison using various data sets 4-55
Table 4-23. Correlation coefficients for selected watershed characteristics 4-62

Table 5-1. Summary of Data for the Lavaca River near Edna 5-2
Table 5-2. Regression line for Lavaca River near Edna (USGS Data) 5-4
Table 5-3. Regression line for Lavaca River near Edna (USGS Data) using concentration 5-6
Table 5-4. Regression line for Lavaca River near Edna (USGS Data) using concentration and polynomial term 5-7
Table 5-5. Samples with flow rates less than 1 cfs 5-14
Table 5-6. Comparison of regressions using 76 data points and 72 data points . 5-14
Table 5-7. Categories of Data 5-17
Table 5-8. Regression Results (no interaction terms) 5-19
Table 5-9. Regression Results (interaction terms) 5-20

Table 6-1. Summary of Daily Data for the Lavaca River near Edna Basin 6-3
Table 6-2. Comparison of daily values and annual averages 6-9
Table 6-3. Regression equations and statistics for the Lavaca River near Edna . 6-14
Table 6-4. Ranges of variables for selected data sets 6-15
Table 6-5. Highest values and dates for flow, concentration, and load 6-18
Table 6-6. Regression equations by flow intervals 6-22
Table 6-7. Regression using sine function 6-23
Table 6-8. Highest values and dates for daily rainfall 6-25
Table 6-9. Correlation of flow, concentration, and load 6-26
Table 6-10. Regression equations and statistics with rainfall at Edna 6-27
Table 6-11. Load-flow relationship according to year categories 6-40

List of Figures

Figure 2-1. Relation of average annual rate of sediment production to drainage are size. (Bulletin 5912, Figure 4.) 2-18
Figure 2-2. Locations of sediment gauging stations in relationship to major river basins....................2-22
Figure 2-3. The Lavaca River Basin....................2-26
Figure 2-4. Lavaca Watershed above Edna....................2-27
Figure 2-5. Gradation analysis of Lavaca River Samples, September 13, 1961.2-30
Figure 2-6. Comparison of bed material and suspended sediment, Lavaca River near Edna 2-31

Figure 3-1. Simplified Research Process....................3-2
Figure 3-2. Daily suspended sediment data format....................3-3
Figure 3-3. Daily sediment data in column format....................3-5
Figure 3-4. Monthly suspended sediment data format....................3-6
Figure 3-5. Monthly sediment data in column format....................3-7
Figure 3-6. Daily streamflow format....................3-9
Figure 3-7. Location of rainfall stations in Lavaca basin....................3-11
Figure 3-8. Annual rainfall....................3-12
Figure 3-9. Numerical text file of gauging stations....................3-17
Figure 3-10. Longitude and Latitude of Sediment Gauging Stations....................3-18
Figure 3-11. Eight Direction Pour Point Model in GRID 3-22
Figure 3-12. Sample 4 x 4 Grids 3-24
Figure 3-13. AML Loop....................3-28
Figure 3-14. AML to compute flow length for one watershed....................3-28
Figure 3-15. Flow Length Grid 3-29
Figure 3-16. K-factor AML....................3-32
Figure 3-17. Portion of HSG AML 3-33

Figure 4-1. Extent of Study Area....................4-1
Figure 4-2. Annual Sediment Load for TWDB Stations....................4-5
Figure 4-3. Averaged Annual Sediment Load for TWDB Stations 4-6
Figure 4-4. Sediment Load and Flow for all TWDB Stations....................4-11
Figure 4-5. Sediment Concentration and Flow for all TWDB Stations 4-12
Figure 4-6. Sediment Load and Rainfall for all TWDB Stations....................4-12
Figure 4-7. Sediment Concentration and Rainfall for all TWDB Stations....................4-13
Figure 4-8. Runoff and Rainfall for all TWDB Stations....................4-13
Figure 4-9. Sediment Load and Runoff With and Without Dams....................4-19
Figure 4-10. Concentration and Runoff With and Without Dams 4-19
Figure 4-11. Sediment Load and Runoff by Number of Years....................4-23

Figure 4-12. Sediment Concentration and Runoff By Number of Years 4-23
Figure 4-13. Sediment Load and Runoff for Selected River Basins 4-27
Figure 4-14. Sediment Concentration and Runoff for Selected River Basins 4-28
Figure 4-15. Predicted sediment load based on bivariate regressions equations for different river basins 4-32
Figure 4-16. Predicted sediment load based on bivariate regressions equations for different river basins with no upstream dams 4-32
Figure 4-17. Spatially varied rainfall across Texas 4-34
Figure 4-18. Gauged watersheds according to rainfall category 4-35
Figure 4-19. Sediment Load and Flow for Different Climates 4-36
Figure 4-20. Sediment Concentration and Flow for Different Climates 4-36
Figure 4-21. Compilation of Sediment Yield Curves 4-46
Figure 4-22. Specific sediment yield by climate for average values 4-47

Figure 5-1. Cause and Effect of Sediment Yield 5-3
Figure 5-2. Traditional Sediment Rating Curve for the Lavaca River Near Edna (USGS Data) 5-4
Figure 5-3. Sediment Rating Curve using Concentration for the Lavaca River near Edna (USGS Data) 5-6
Figure 5-4. Modified Sediment Rating Curve using Concentration for the Lavaca River Near Edna (USGS Data) 5-8
Figure 5-5. Multiple Causality Between Variables Effecting Sediment Yield 5-8
Figure 5-6. Binary Variable Model with no Interaction Terms 5-18
Figure 5-7. Binary Variable Model with Interaction Terms 5-21
Figure 5-8. Comparison of predicted and measured values for selected model .. 5-23

Figure 6-1. Frequency of Daily Flows, Lavaca River Near Edna 6-8
Figure 6-2. Sediment Load, Lavaca River near Edna, 1945-1989 6-13
Figure 6-3. Sediment Concentration, Lavaca River near Edna, 1945-1989 6-13
Figure 6-4. Comparison of load equations shown in Table 6-3 6-17
Figure 6-5. Comparison of concentration equations shown in Table 6-3 6-17
Figure 6-6. Relation of total suspended load to flow rate 6-19
Figure 6-7. Frequency of flows with fraction of total suspended sediment load 6-20
Figure 6-8. Comparison of equations expressing load-flow trend 6-23
Figure 6-9. Annual temperature trend 6-29
Figure 6-10. Monthly trends in load, flow, concentration, and rainfall 6-32
Figure 6-11. Typical plot of one year of daily data 6-36
Figure 6-12. Mass curves for flow and load 6-39
Figure 6-13. Double mass curve identifying changes in load-flow relationship. 6-39
Figure 6-14. Lavaca River, 1956 6-42
Figure 6-15. Lavaca River, 1969 6-42

1 INTRODUCTION

1.1 Purpose Of Research

Soil erosion and sedimentation are natural processes that have been accelerated by human activity and have thus become serious problems across the world. Sedimentation, as defined by Vanoni (1975), embodies the processes of erosion, entrainment, transportation, deposition, and compaction of sediment. Increased soil erosion results in poor crop growth and poor economic returns. The resulting sediment can be a major source of pollution. According to the Interagency Sedimentation Work Group's classification (Fan, 1988), there generally exist five locations with sedimentation problems: watershed, stream, reservoir, estuary and coast. Essentially all parts of the environment are affected. Often nutrients and toxicants are attached to sediment, contributing to nonpoint-source pollution. Faye, Carey, Stamer, and Klecker (1980) report for 14 watersheds in Georgia, 60 percent or more of the total annual discharge of trace metals and phosphorus was carried by suspended sediment. The corresponding discharges of nitrogen and organic carbon on suspended sediment ranged from 10 to 70 percent of the total. Also, the tremendous volume and weight of deposited sediment can have a profound environmental and economical impact. According to Wang (1985), two million tons of sediments are discharged into the Gulf of Mexico everyday. Fan and Springer (1990) report that the World Bank estimates the loss of worldwide reservoir storage capacity, due to siltation alone, is approximately 40 million acre-feet or the equivalent of $6 billion in replacement cost every year. In addition to decreased water storage capacity in reservoirs, some of the detrimental effects of sedimentation are the need for dredging of waterways to keep them navigable, the disturbance of biological habitats, and, if sediments travel through reservoirs and reach hydropower plants, the increase in required maintenance of plant structures and machinery. Sediment load is a serious water quality problem in the United States and in Texas. Water quality decreases as sediment concentration increases. Engineers have various methods to estimate the amount of erosion and subsequent sediment loading of waterways. Due to the complexities of sediment detachment, transport, and deposition, no method is completely satisfactory and most methods have large inadequacies. It is necessary to understand the environmental impacts of sedimentation and to quantify these impacts so that proper planning and designing can minimize the detrimental effects of sedimentation.

In recognition of the potential problems associated with sedimentation, the Texas Water Development Board (TWDB) collected daily suspended sediment samples from rivers across the state from the 1920's to 1989. Prior to this study, there has been no comprehensive analysis of the data.

1.2 Research Objective

The primary objective of this research is to present existing suspended sediment data collected from Texas Rivers and to analyze the data using regression techniques combined with knowledge and theory of physical processes to develop models which establish relationships between sediment load and streamflow, precipitation, and watershed parameters to enable the prediction of future suspended sediment loads and to further the understanding of the sedimentation process.

This research objective is carried out in several tasks. Chapter 2 presents the background for the research. Included is a literature review and a description of the suspended sediment sampling program. The next task is the assembling of the data so that it can be used in regression models. This task is non-trivial as described in Chapter 3. Programs are developed to read the data for sediment, streamflow, and rainfall and then write the data to a usable format. GIS is used to determine spatial characteristics of the watersheds and to write the spatial data to a usable format.

Chapter 4 investigates the effects of gross watershed characteristics on sediment yield by considering annual sediment load and average sediment load for 60 stations across the state of Texas. The data are classified according to time, spatial location, and the existence of dams in the basin.

The sediment rating curve for one station, the Lavaca River near Edna, is improved by considering the temperature and the percent fines of each sample (Chapter 5). The analysis uses data that were collected by the USGS.

Finally, in Chapter 6, the daily data for the Lavaca River near Edna are analyzed. The data set includes 45 years of daily samples collected by the TWDB. The sediment data are compared with streamflow, spatially averaged rainfall, and season. Chapter 7 presents conclusions from the research.

1.3 Research Contributions

The contributions made by this research include the following.

- The historical suspended sediment data for Texas are presented and analyzed with respect to streamflow, rainfall, and watershed characteristics.
- GIS methods are used to characterize and describe watersheds according to topography, climate, soil type, land use, and presence of reservoirs.
- New SAS data sets and SAS models are developed to describe the relationship of suspended sediment to streamflow, rainfall, and watershed characteristics.
- Multivariate regression models are developed which quantify spatial and temporal trends of the suspended sediment load across the state. Data from across the state are classified into three climate zones. Regression models for the individual zones are developed. The data are further classified according to the existence of dams upstream from the gauging station. This analysis shows that dams have more impact on the sediment load-streamflow relationship in the drier parts of Texas than the wetter parts of

the state. A reservoir variable is developed and included in the multivariate models to assist in determining the impact the reservoirs have in each basin.

- Techniques to improve sediment rating curves are developed with periodic samples for the Lavaca River near Edna, Texas. Variables relating to the time of year and the origin of the sediment are used along with interaction terms to explain an additional 41% of the variability in sediment concentration in the Lavaca River compared to the simple bivariate regression.
- Long term trends, as well as seasonal trends, are identified for the suspended sediment load for the Lavaca River near Edna, Texas.
- Techniques to improve sediment rating curves are developed with daily samples for the Lavaca River near Edna. Variables relating to the time of year and the limb of the hydrograph are used along with interaction terms to explain an additional 22% of the variability in sediment concentration in the Lavaca River compared to the simple bivariate regression.

2 BACKGROUND

2.1 Sediment Mechanics

Movement of sediment is caused by water or wind, but this discussion will concern sediment motion due to water, only. Detachment of sediment and subsequent erosion occurs due to raindrops or to runoff. Raindrop, or splash, erosion is dependent upon the kinetic energy of rainfall, the shearing resistance of the soil, the grain size of the soil, the ground slope, and the angle at which the rain falls. The impact of a raindrop both compacts the soil surface and disperses soil. Runoff occurs once the infiltration capacity of a soil is exceeded. Erosion occurs once the threshold of movement is reached. For non-cohesive soils, this threshold is dependent upon the shear stress at the ground surface, the sediment grain density and diameter, the fluid density and viscosity, the ground slope, and the acceleration of gravity. A distinction between rill and interill erosion is often made, with rill erosion occurring in small channels where the flow is concentrated, and interill erosion occurring where the runoff is characteristic of true sheet flow. Sediment originating in a watershed and carried to a stream is often termed washload. Sediment is kept in suspension due to turbulence and sometimes by intergranular collisions. The relationship of the sediment fall velocity to the strength of the turbulence then determines how long particles are kept in suspension.

Runoff and streamflow also cause erosion of channel banks and beds. After entrainment, bed material can contribute to the suspended load. Sediment which moves along the river's bottom and is largely composed of material similar to that of the river bed is bedload. Caution should be taken when using terms such as total load, bedload and suspended load, as the terms are not used consistently in the literature. Often these three terms refer to material that has originated in the bed and exclude washload. For the purpose of this research, the loads will include washload unless stated as bed material load.

In general, the amount of suspended sediment per unit area of upstream watershed is inversely proportional to that area of the upstream watershed. This relationship is due to the fact that some of the sediment that erodes is deposited before ever reaching the gauging station. Larger watersheds have more area for sediments to redeposit than smaller watersheds. By far the largest amount of sediment in a river is carried in suspension. Suspended sediment is almost always present in perennial streams, whereas bedload may be present only a small amount of the time.

2.2 Historical Work

2.2.1 Sediment Transport

Interest in predicting bedload sediment transport dates back to 1879 when DuBoys developed the idea that a fluid in motion exerts a shearing force on the stream bed causing sediment to move in the direction of the shear stress. Since that time there have been numerous equations developed to predict bedload, some mechanistic, some empirical, and some probabilistic. The most well known formulas include those derived by Shields (1936), Meyer-Peter and Muller (1948), Einstein (1950), and Bagnold (1966, 1980). Shields' bedload equation is a natural outcome of his well-known incipient motion relation of critical shear stress to particle Reynolds number. Like DuBoys, Shields presented bedload transport rate as a function of excess shear stress in relation to the critical shear stress. Meyer-Peter and Muller's empirically derived bedload equation also implies that bedload transport is a function of excess shear. Einstein departed from the idea that motion begins at "critical" conditions and included the probability of movement in his analysis. Thus, his bedload equation allows for some transport at very small shear stress values. Bagnold introduced the idea that sediment transport is a function of the work done which can be related to the stream power. All of these bedload relationships include coefficients which were derived using laboratory flume experiments. A successful, generally applicable prediction equation of bedload transport has still not been found and little field data are available. Continuous efforts are being devoted to establish the reasons behind the poor performance of predictive bedload equations.

Fewer suspended load transport equations have been developed than bedload equations. Vanoni (1946) and Laursen (1958) show how suspended load transport can be treated theoretically using concepts of continuity, momentum, and turbulent mixing. Einstein (1950), Bagnold (1980), Ackers and White (1973), and Yang (1973) present formulas for total load. They define total load as all sediment (suspended and bedload) that is transported and that originates in the stream bed. The different formulas can produce very different results. When choosing a formula to compute total load, the conditions under which the formula was derived should be considered. One problem is that many factors are either unpredictable or too complex to model without excessive simplification. The bulk properties of the fluid are dependent on the sediment concentration while the sediment concentration is dependent upon the bulk properties of the fluid. Because of these complexities and the fact that wash load is not included in total load formulas, empirical correlations between suspended sediment concentration and streamflow are often used.

2.2.2 Erosion

Most of the suspended sediment load concentration in a river is due to erosion on the watershed and is not made up of bed material. According to the Soil Science Society of America (1979), agricultural lands account for 40 percent of total sediment, streambank erosion accounts for 26 percent, pasture and rangeland account for 12 percent, and forest lands account for 7 percent. The remaining 15 percent is attributable to other federal lands, urban areas, roads, and "other" sediment sources. Methods for predicting erosion were developed as economic losses due to soil erosion affected the agricultural industry. Empirical equations were developed using erosion plots in both laboratory and field conditions. Most equations relate soil loss in depth per year to factors such as rainfall, crop cover, soil characteristics, land use, degree of land slope, and length of slope. The most well known equation in the United States is the Universal Soil Loss Equation (USLE) developed by Wischmeier and Smith (1958). The equation expresses annual soil loss (A) in terms of six factors: rainfall erosivity (R), soil erodibility (K), slope length (L), slope gradient (S), cropping management (C), and erosion control practice (P). According to Wischmeier (1976) the equation may be used to

1. predict average annual loss of soil from a cultivated field with specific land use conditions,
2. guide the selection of cropping and management systems, and conservation practices for specific soils and slopes,
3. predict the change in soil loss that would result from a change in crops or land use on a specific field,
4. determine how conservation practices should be adjusted to allow higher crop yields,
5. estimate soil losses from areas that are not in agricultural use, and
6. provide estimates of soil losses for planners of conservation works.

The USLE is intended to represent a long-term average; it does not accurately estimate erosion for a specific storm event, season, or year, and it does not estimate erosion by concentrated flow.

Sediment yield is a function not only of the amount of upstream erosion, but also of the transport ability of the stream system and includes deposition, scour, and resuspension of the sediments. After computing gross erosion using the USLE, a sediment delivery ratio (SDR) is applied to determine sediment yield. The SDR is defined as the ratio of sediment delivery to gross erosion on the watershed and is generally less than unity due to deposition on land surfaces and in the stream. The SDR can be more than unity for isolated cases since the gross erosion estimate does not always include channel erosion and/or re-suspension of sediment in the stream system. The SDR is usually estimated based on the watershed size and adjusted for the soils, land use, topography, etc.

The Modified Universal Soil Loss Equation (MUSLE) (Williams, 1975) replaces the rainfall energy factor (R) with a runoff factor. MUSLE therefore computes sediment yield from an individual storm. Furthermore delivery ratios are not required because the runoff factor represents energy used in detaching and transporting sediment.

2.3 Recent Work

2.3.1 Sediment Transport

In 1988, the Sedimentation Work Group of the Subcommittee on Sedimentation of the Interagency Advisory Committee on Water Data compiled a review of 12 computer stream sedimentation models. Dr. Shou-shan Fan, Chair of the Work Group, provides a summary report of the 12 models which include 6 privately owned models (CHARIMA, SEDICOUP, FLUVIAL12, HEC2SR, TWODSR, RESSED) and 6 federally owned models (HEC-6, TABS2, IALLUVIAL, STARS, GSTARS, ONED3X). Descriptions of each model are given along with capabilities and equations used. The twelve models that were reviewed share many of the same basic equations: continuity, energy, and momentum. Many of the models allow the user to choose optional sediment transport formulas. The models reviewed calculate the changes of sediment loading in a natural channel. The changes are due to incoming sediment, channel bed scour, bank erosion, settling, and deposition. In several models, the incoming sediment is determined by the user's specified relationship between the amount of sediment and the amount of water entering the study reach. This relationship is often determined using several years of sporadic data from a USGS gauging station. Some of the models use an equilibrium sediment load. The models serve the purpose of generally defining the expected changes in a channel over time. They do not focus on the downstream impacts of the transported sediment or on washload.

Dr. Fan points out that major hindrances to computer sedimentation modeling include that engineers and scientists do not thoroughly understand the physics underlying the sedimentation models, that they also lack appropriate mathematical techniques, and more importantly, that they lack adequate data to calibrate and to verify such heavily data-dependent problems. He finds all the models reviewed to have the following three "drawbacks":

1. They are heavily data dependent; their applicabilities are often limited to the character ranges of the data used to develop the models.
2. Adequate data are critical to model development and implementation. Unfortunately, such data are usually not readily accessible to the public.

3. Many scientists are unable to work on models simply because of the lack of the necessary data.

An evaluation of runoff and erosion models conducted by Wu, Hall, and Bonita (1993) reviewed and tested 3 models: AGNPS, ANSWERS, and CREAMS. Noting that it is sometimes necessary to estimate sediment yield for a specific storm or a series of large storms, they concentrated on large storm events where prediction errors would be most serious. AGNPS uses a modified version of the USLE to compute detachment. ANSWERS and CREAMS also use empirical equations (different from the USLE) to compute detachment. Three watersheds were used to test measured data against computed values. In the computed values, antecedent moisture and crop cover were changed to compute maximum and minimum values of sediment yield. The errors in estimating antecedent moisture and crop cover could not account for the large differences between measured and computed sediment yields. The authors offer no explanation. All three models tend to underestimate sediment yield for larger events.

As part of the HUMUS (Hydrologic Unit Model for the United States) project conducted by the USDA (Arnold, 1995), a Soil and Water Assessment Tool (SWAT) has been developed. The objective of SWAT is to predict the effect of management decisions on water and sediment yields of large watersheds. The program links to GIS, automating inputs and spatially displaying outputs.

Wicks and Bathurst (1996) recently introduced a physically-based, distributed erosion and sediment yield component, SHESED, to their existing hyrological modelling system, SHE. SHESED combines hillslope and channel components of erosion, transport and deposition. The main drawback to the model is the heavy reliance on the calibration of erodibility coefficients. It is interesting to note that in the channel sediment routing procedure, it is assumed that the flow can carry any available load of fine sediments (less than 0.062 mm) but for coarser sediments the load is limited by the calculated capacity transport rate of flow.

ANSWERS-2000 (Bouraoui and Dillaha, 1996) has been developed to simluate long-term average annual runoff and sediment yield from agricultural watersheds. The model can be used without calibration. Predictions of sediment yield for individual storms were within 200% of observed values while predictions of cumulative sediment yield were within 12% and 68% of observed values. Watersheds in Georgia and Virginia were used for the model evaluation.

Kothyari, Tiwari, and Ranvir (1996) analyze the temporal variation of sediment yield carried by the stream during the storm. They combine a time-area curve with sediment delivery to develop a method for prediction of the variation of sediment yield with time. For individual storms, known sediment yield is compared to predicted sediment yield. Sixty-eight percent of the comparisons fall within a ± 40% error band. The authors are pleased with the results. Typically, in sediment studies the standard for “good” agreement tends to be lower than in other fields.

2.3.2 Erosion

The US Department of Agriculture published the results of Water Erosion Prediction Project (WEPP) in August, 1995. WEPP is a major study with the aim of developing a new methodology for erosion prediction based upon fundamental erosion mechanics. WEPP is intended to replace the USLE. According to Dr. John Laflen (1995), the model will be used by all federal agencies. However, no one in the Temple, Texas, office of the Agricultural Research Service expects WEPP to replace the USLE in the near future (i.e. 5-10 years). WEPP can be downloaded from the Internet at

(http://soils.ecn.purdue.edu/~wepphtml/wepp/wepptutpmail.html)

along with a March, 1997 patch program. The model is a DOS-based program and can cause some problems for Windows 95 or Windows NT users. WEPP is a process-oriented, continuous simulation, erosion prediction model. It is applicable to small watersheds (field-sized) and can simulate small profiles (USLE types) up to large fields. It mimics the natural processes that are important in soil erosion. Everyday it updates the soil and crop conditions that affect soil erosion. When rainfall occurs, the plant and soil characteristics are used to determine if surface runoff will occur. If predicted, the program will compute estimated sheet and rill detachment and deposition, and channel detachment and deposition.

2.4 Historical Work in Texas

The Texas Board of Water Engineers was organized in 1914 and began taking suspended sediment samples in Texas rivers as early as 1924. The state agency was reorganized under various names and is currently (1998) the Texas Water Development Board (TWDB). The TWDB discontinued the sediment sampling program in 1989. The TWDB and its predecessor agencies published reports (Texas Water Development Board Reports 306, 184, 106, and 45; Department of Water Resources Report 233; Texas Water Commission Bulletin 6410, and Texas Board of Water Engineers Bulletin 6108) containing monthly data for the stations; however the daily data are available. Original paperwork containing data prior to 1965 is stored in the TWDB's warehouse (Sullivan, 1994). Data from 1965 to 1989 are in digital format and available through the Texas Natural Resources Conservation Commission (TNRCC).

In January, 1959, the Texas Board of Water Engineers published Bulletin 5912, "Inventory and Use of Sedimentation Data in Texas." The report was prepared by the Soil Conservation Service with the main purpose being to "furnish the best possible estimates of average annual sediment production rates for watersheds larger than 100 square miles throughout the State." The report contains 2 tables and 5 figures which summarize the report. Table 1 of Bulletin 5912 (shown below as Table 2-1) provides a summary of sediment measuring stations and data.

Table 2-1. Sediment Load Data (Table 2, Soil Conservation Service, 1959)

Drainage Basin and Stream	Location	Sediment Contributing Area	Length of Record	Adj. Annual Sediment Production Rate		Volume Weight Adj. Factors	Bed Load Adj. Factors	Adj. Annual Sediment Production Rate	Bull. 5912 References
				Per mi^2	Est. Vol. Wt.				
		(mi^2)	(yrs)	(acre-ft)	(lb/ft^3)				
CANADIAN RIVER									
Wolf Creek	Lipscomb	697	6.94	0.531	40	1.00	1.30	0.69	16
RED RIVER									
Pease River	Crowell	2410	5.002	0.412	70	1.40	1.30	0.75	2
Red River	Denison	32840	6.260	0.415	70	1.40	1.30	0.76	2
SABINE RIVER									
Sabine River	Logansport, LA	4858	20.156	0.131	70	1.17	1.30	0.20	2
NECHES RIVER									
Angelina River	Horger-Broaddus	2803	11.817	0.082	70	1.75	1.30	0.187	2,21
Neches River	Rockland								
TRINITY RIVER									
Denton Creek	Roanoke	621	4.62	0.650	60	1.00	1.10	0.71	16
East Fork	Rockwall	840	6.61	0.541	35	1.00	1.10	0.60	16
Trinity River	Rosser	8057	3.181	0.073	70	1.75	1.15	0.15	2
Trinity River	Romayor	17192	21.142	0.198	70	1.75	1.15	0.39	2,21

Table 2-1 (continued)

Drainage Basin and Stream	Location	Sediment Contributing Area	Length of Record	Adj. Annual Sediment Production Rate		Volume Weight Adj. Factors	Bed Load Adj. Factors	Adj. Annual Sediment Production Rate	Bull. 5912 References
				Per mi^2	Est. Vol. Wt.				
		(mi^2)	(yrs)	(acre-ft)	(lb/ft^3)				
SAN JACINTO RIVER									
West Fork	Humble-Conroe	1811	20.753	0.103	70	1.75	1.30	0.23	2,21
East Fork	Cleveland	330	4.833	0.034	70	1.75	1.30	0.078	2,21
San Jacinto	Juffman	2791	6.597	0.182	70	1.75	1.15	0.35	2
Buffalo Bayou	Houston	362	7.65	0.389	60	1.00	1.10	0.43	16
Brays Bayou	Houston	100	7.51	0.208	60	1.00	1.10	0.23	16
White Oak Bayou	Houston	92	7.43	0.580	60	1.00	1.10	0.64	16

Table 2-1 (continued)

Drainage Basin and Stream	Location	Sediment Contributing Area	Length of Record	Adj. Annual Sediment Production Rate		Volume Weight Adj. Factors	Bed Load Adj. Factors	Adj. Annual Sediment Production Rate	Bull. 5912 References
				Per mi^2	Est. Vol. Wt.				
		(mi^2)	(yrs)	(acre-ft)	(lb/ft^3)				
BRAZOS RIVER									
Salt Fork	Aspermont	2216	1.238	1.272	70	0.875	1.30	1.45	2
Salt Fork	Seymour	5250	6.107	1.238	70	0.875	1.30	1.42	2
Double Mountain Fork	Aspermont	1510	9.244	1.765	70	0.875	1.30	2.01	2
Clear Fork	Crystal Falls	4320	3.307	0.131	70	1.0	1.0	0.131	2
Clear Fork	Eliasville	5740	1.244	0.092	70	1.0	1.15	0.105	2
Little River	Little River	5253	4.962	0.143	70	2.0	-	0.29	2
San Gabriel River	Circleville	602	5.403	0.369	70	2.0	-	0.74	2
Leon River	Belton-Gatesville	2313	8.916	0.143	70	1.40	1.30	0.26	2,21
Navasota River	Easterly	949	12.081	0.184	70	1.75	1.15	0.37	2
Brazos River	South Bend	12360	15.710	0.259	70	1.37	1.30	0.46	2,21
Brazos River	Mineral Wells	13910	10.332	0.468	70	1.00	1.15	0.54	2
Brazos River	Glen Rose	15600	4.588	0.537	70	1.00	1.15	0.62	2
Brazos River	Waco	19260	9.254	0.536	70	1.00	1.15	0.62	2
Brazos River	Richmond	34810	33.306	0.538	70	1.40	1.05	0.79	2,21
Big Elm Creek	Buckholtz	166	2.54	2.08	70	1.00	1.05	2.19	16
Big Elm Creek	Temple	68.5	2.29	4.78	50	1.00	1.05	5.02	16
North Elm Creek	Ben Arnold	30.3	2.00	2.00	50	1.00	1.05	2.10	16
Brushy Creek Sub-watersheds									
J	Riesel	9.16	1.30	1.30	40	1.00	1.00	1.30	16
D	Riesel	1.74	0.98	0.98	40	1.00	1.00	0.98	16
Y	Riesel	0.48	1.94	1.94	40	1.00	1.00	1.94	16
W-1	Riesel	0.28	6.60	6.60	40	1.00	1.00	6.60	16
Y-2	Riesel	0.2	1.44	1.44	40	1.00	1.00	1.44	16

Table 2-1 (continued)

Drainage Basin and Stream	Location	Sediment Contributing Area	Length of Record	Adj. Annual Sediment Production Rate Per mi²	Adj. Annual Sediment Production Rate Est. Vol. Wt.	Volume Weight Adj. Factors	Bed Load Adj. Factors	Adj. Annual Sediment Production Rate	Bull. 5912 References
		(mi²)	(yrs)	(acre-ft)	(lb/ft³)				
COLORADO RIVER									
Llano River	Llano	4000	11.167	0.038	70	1.0	1.30	0.049	2
Pedernales River	Johnson City	947	11.167	0.100	70	1.0	1.30	0.130	2
Colorado River	San Saba	18700	27.055	0.161	70	1.40	1.15	0.260	2,21
Colorado River	Tow	19300	5.162	0.174	70	1.40	1.15	0.280	2
Colorado River	Columbus-Eagle	29140	6.997	0.202	70	1.40	1.10	0.310	2
LAVACA RIVER									
Lavaca River	Edna	887	12.083	0.105	70	2.00	1.15	0.241	2,21
GUADALUPE RIVER									
Guadalupe River	Spring Branch	1432	15.748	0.077	70	1.00	1.15	0.088	2,21
Guadalupe River	Victoria	5311	9.083	0.057	70	1.75	1.15	0.113	2
SAN ANTONIO RIVER									
San Antonio	Falls City	2070	5.967	0.069	70	1.75	1.15	0.138	2
San Antonio	Goliad	3918	12.748	0.095	70	1.75	1.15	0.191	2
NUECES RIVER									
Nueces River	Three Rivers	15600	25.583	0.030	70	2.00	1.10	0.066	2
Nueces River	Cotulla	5260	12.748	0.011	70	2.00	1.10	0.024	2

Table 2-1 (continued)

Drainage Basin and Stream	Location	Sediment Contributing Area	Length of Record	Adj. Annual Sediment Production Rate Per mi^2	Adj. Annual Sediment Production Rate Est. Vol. Wt.	Volume Weight Adj. Factors	Bed Load Adj. Factors	Adj. Annual Sediment Production Rate	Bull. 5912 References
		(mi^2)	(yrs)	(acre-ft)	(lb/ft^3)				
RIO GRANDE RIVER									
Rio Grande	El Paso	29271	8.0	0.0067	66.7	0.833	1.20	0.0067	1
Rio Grande	Presidio	66203	8.0	0.0283	66.7	0.833	1.20	0.0283	1
Rio Grande	Johnsons Ranch	70715	8.0	0.0816	66.7	0.833	1.20	0.0816	1
Rio Grande	Aqua Verde	82232	2.0	0.0690	66.7	0.833	1.20	0.0690	1
Rio Grande	Langtry	84795	11.0	0.0686	66.7	0.833	1.20	0.0686	1
Rio Grande	Eagle Pass	130575	21	0.0569	66.7	0.833	1.20	0.0569	2,1
Rio Grande	Laredo	135976	2	0.0258	66.7	0.833	1.20	0.0258	1
Rio Grande	Roma	157204	14.184	0.080	66.7	0.833	1.20	0.0853	2

References:

1. International Boundary and Water Commission, United States and Mexico, Water Bulletin No. 25, "Flow of the Rio Grande and Related Data," 1955.

2. State of Texas, Board of Water Engineers and U.S. Department of Agriculture, Soil Conservation Service - Sixteenth Annual Report of "The Silt Load of Texas Streams 1953-1954."

16. Brune, G.M., Maner, S.B., Renfro, G.W., and Ogle, J.A. "Rates of Sediment Production in the Western Gulf States," U.S. Department of Agriculture, Soil Conservation Service, SCS-TP-127, 21 pp., Illus. (Processed).

21. Unpublished data from files of the Board of Water Engineers, State of Texas, Austin, Texas.

For each station a bedload adjusting factor is used to determine the annual sediment production rate in acre-feet per square mile. Table 2 of Bulletin 5912 summarizes reservoir sedimentation survey data and provides an average annual sediment production rate in acre-feet per square mile. The Soil Conservation Service compiled a map dividing the state of Texas into 14 major land resource areas based on similarity of soils, topography, climate, and vegetation. Essentially all variables of interest, except for drainage area, were lumped together into the major land resource area. Figures 1 through 5 of Bulletin 5912 display curves for the different land resource areas with estimated sediment production rate versus drainage area. Data from Tables 1 and 2 of Bulletin 5912 were used in developing the curves in the figures. Figure 2-1 of this report displays a replication of one of the Bulletin's figures.

The bulletin also contains summaries of specific sedimentation problems within major basins. No further effort was made in correlating sediment production rate with other variables. Also, no investigations of individual storms or seasonal patterns were made; thus, an annual production rate was derived. Reduction in sedimentation rates was proposed for each land resource area according to the projected soil conservation measures.

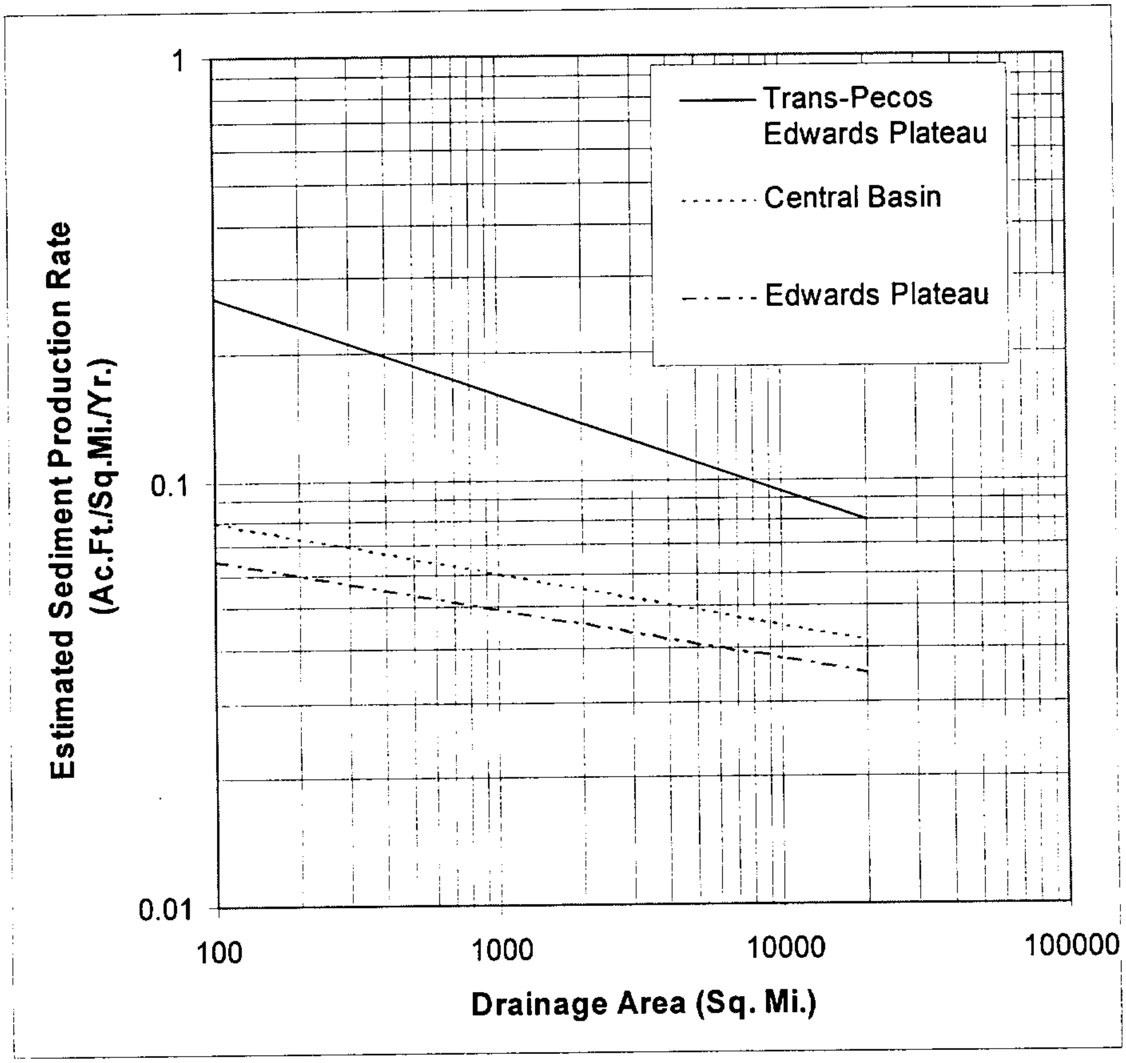


Figure 2-1. Relation of average annual rate of sediment production to drainage area size. (Bulletin 5912, Figure 4.)

2.5 TWDB Suspended Sediment Data

Suspended sediment measurements were made by the TWDB at 60 stations with lengths of record ranging from 3 to 59 years. Daily suspended sediment samples at the 60 stations represent 12 river basins.

Table 2-2 displays stations from which suspended sediment samples were collected under the state's program. The table was compiled using the original data sheets (stored in TWDB's warehouse) and the compilation reports which contain the monthly sediment totals. Some of the data were in order by station or year and some were not. The data tables were neat and ranged from handwritten to typed entries. Every effort was made to place the data in order when going through the boxes. The data have not been presented beyond the publication of monthly values and Bulletin 5912. The sediment stations were located adjacent or near USGS streamflow gauging stations and referenced with the USGS gage number. Figure 2-2 shows the location of the sediment gauging stations.

Table 2-2. TWDB Suspended Sediment Sampling Stations

Sta No.	watercourse	location	area	start	end	total time	average stream-flow	sed load	years used for ave
			mi^2	date	date	yrs	ac ft	tons/mi^2	
Red River Basin									
7299200	Red R. (fork)	Lakeview	2023	05-64	09-77	13.3	57,434	1,114	12
7336820	Red R.	DeKalb	41412	02-69	10-79	10.7	5,529,985	93	10
Sulphur River Basin									
7342500	S. Sulphur R.	Cooper		03-62	10-66				
7343200	Sulphur R.	Talco	1365	11-66	10-89	22.9	1,148,317	1,049	15
7343500	Whiteoak Creek	Talco	494	06-63	10-89	26.4	360,012	68	18
Sabine River Basin									
8022000	Sabine R.	Tatum	3493	06-68	06-89	21.0	1,765,671	38	14
8022500	Sabine R.	Longsport, LA		12-32	03-68	35.3			
Neches River Basin									
8031200	Kickapoo Creek	Brownsboro	232	05-62	10-79	17.4	102,109	13	17
8033000	Neches R.	Diboll	2724	06-66	10-85	19.3	1,051,813	20	16
8033300	Piney Creek	Groveton	79	05-62	10-79	17.4	27,315	42	17
8033500	Neches R.	Rockland		10-30	05-66	35.6			
8037050	Bayou La Nana	Nacogdoches	31.3	06-65	10-85	20.3	22,208	225	17
	Neches R.	3 Rivers		10-27	04-52	24.5			
Trinity River Basin									
8052700	Little Elm Creek	Aubrey		70-63	10-68	5.3			
8062500	Trinity R.	Rosser	8146	11-38	80-89	50.8	1,843,222	117	29
8064500	Chambers Creek	Corsicana	963	06-63	10-79	16.3	319,194	455	16
8065350	Trinity R.	Crockett	13911	05-68	10-89	21.4	4,082,256	115	14
8066200	Long Kning Creek	Livingston	141	06-63	10-79	16.3	112,588	589	16
8066500	Trinity R.	Romayor	17186	08-36	10-89	53.2	5,301,050	200	46
San Jacinto River Basin									
8070000	East Fork S.J. River	Cleveland	325	12-52	10-89	36.9	141,741	80	29
	West Fork S.J. River	Conroe		12-52	04-62	9.3			
	S.J. River	Haffman		09-45	03-52	6.5			
	West Fork S.J. River	Humble		12-32	04-52	19.3			

Table 2-2 (continued)

Sta. No.	watercourse	location	area	start	end	total time	average stream-flow	sed load	years used for ave
			mi^2	date	date	yrs	ac ft	tons/mi^2	
Brazos River Basin									
8084800	California Creek	Stamford	478	07-64	09-79	15.2	21,301	93	15
8085500	Clear Fork	Griffin		10-86	10-89	3.0			
8087300	Clear Fork	Eliasville	5697	05-66	10-82	16.4	241,692	88	16
8088000	Brazos R.	South Bend	13107	01-42	10-89	47.8	552,117	293	39
8088500	Possum Kingdom Res.	Graford	13310	01-42	10-79	37.8	563,424	—	35
8093500	Aquilla Creek	Aquilla	308	06-63	10-89	26.4	92,504	820	19
	Leon R.	Belton		10-45	12-49				
8094800	North Bosque R.	Hico	359	04-62	10-89	27.5	29,630	93	20
8100500	Leon Ri.	Gatesville	2342	03-53	10-89	36.6	175,673	100	28
8109900	Somerville Lake	Somerville	1009	06-62	10-79	17.3	230,956	—	13
8110500	Navosta R.	Easterly	968	01-42	10-89	47.8	303,098	145	37
8114000	Brazos R.	Richmond	35441	06-24	10-79	55.4	5,296,820	651	55
Colorado River Basin									
8146000	San Saba R.	San Saba	3042	04-66	10-89	23.5	158,378	33	16
8147000	Colorado R.	San Saba	17720	09-30	10-89	59.1	825,111	134	52
8148000	Lake Buchanan	Burnet	18370	10-47	10-89	42.0	541,806	—	35
8151500	Llano R.	Llano	4233	08-42	10-89	47.2	257,641	105	38
8153500	Pedernales R.	Johnson City		08-42	10-67	25.2			
8158000	Colorado R.	Austin		08-37	10-89	52.2	1,392,560	4	21
	Colorado R.	Inks Dam		08-42	09-66	24.1			
Lavaca River Basin									
8164000	Lavaca River	Edna	817	09-45	09-89	44.0	235,816	194	37
8164300	Navidad R.	Hallettsville	332	03-62	10-89	27.6	120,057	131	20
8164500	Navidad R.	Ganado	1062	03-76	09-79	3.5	587,741	218	3
Guadalupe River Basin									
8167500	Guadalupe R.	Spring Branch	1315	01-42	10-89	47.8	236,709	136	40
8176500	Guadalupe R.	Victoria	3766	09-45	10-89	44.1	1,682,705	167	17
San Antonio River Basin									
8186000	Cibolo Creek	Falls City	827	06-63	10-89	26.4	108,283	143	19
8188500	San Antonio R.	Goliad	3921	01-42	10-89	47.8	504,608	127	40

Table 2-2 (continued)

Sta. No.	watercourse	location	area	start	end	total time	average stream-flow	sed load	years used for ave
			mi^2	date	date	yrs	ac ft	$tons/mi^2$	
Nueces River Basin									
8194000	Nueces R.	Cotulla	5171	01-42	09-79	37.7	181,665	13	37
8207000	Frio River	Calliham	5491	01-53	09-79	26.7	177,311	23	26
8210500	Lake Corpus Christi	Mathis	16656	02-42	10-89	47.7	636,542	—	20

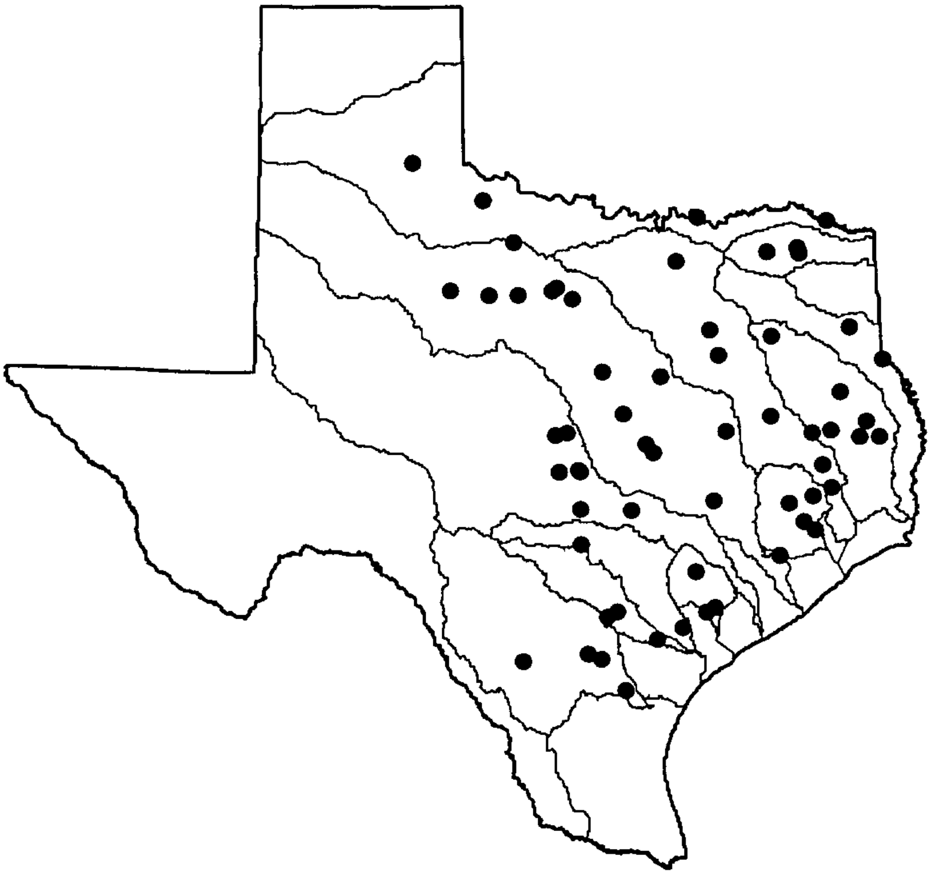


Figure 2-2. Locations of sediment gauging stations in relationship to major river basins

The average length of record was 31 years. The upstream drainage basin areas range from 31 square miles at Bayou La Nana near Nacodoches to 41,412 square miles on the Red River near DeKalb. The rivers chosen for sediment sampling vary greatly in terms of annual average streamflow and annual average suspended sediment load. The Red River at DeKalb averaged the highest annual streamflow at 5.53×10^6 acre-feet with an annual average suspended sediment load of 1,114 tons per square mile. The lowest annual average streamflow was measured at the California Creek station near Stamford where 2.13×10^4 acre-feet and 93 tons per square mile were the measured annual averages. The average annual suspended sediment load ranged from 13 to 1,114 tons per square mile with the exception of those stations located at the outlet of a dam where the average annual load was immeasurable in tons. Because the data cover so many years, a large variation of storms is represented.

2.5.1 Sampling Method

Sediment samples were collected in eight ounce narrow-neck bottles at a position approximately one foot below the water surface near midstream. The percentage of suspended sediment by weight obtained from the sample was multiplied by the factor 1.102 to obtain the mean percentage of suspended sediment in the vertical profile. The average streamflow for the corresponding day is then used in determining the total suspended sediment load.

In the early stages of the program, Farris (1933) studied many samples taken at various depths throughout a cross-section and at different gage heights. He determined that a sample from six tenths depth gave the mean percentage of suspended sediment in the vertical. He further determined that if the sample is taken within the top foot of the stream then the percent of suspended sediment by weight should be multiplied by 1.102 to obtain the value at six tenths depth.

The suspended sediment data are in tons per day. Although the samples were taken at a particular time of day, the suspended sediment loads were computed using the instantaneous bottle sample and the average daily streamflow rather than the actual corresponding streamflow. For most of the analysis in this research the sediment data will be treated as daily values. To convert the data from sediment load to sediment concentration, one acre-foot of streamflow is assumed to weigh 1361.25 tons as outlined in the "Explanation of Data" sections of the TWDB reports. Using this number and unit conversions, the suspended sediment data are converted to a concentration in both percent by weight and mg/l using $(\% \text{ by weight}) * 1.0012 \times 10^6 = \text{mg/l}$.

2.5.2 Lavaca River Data

The station on the Lavaca River near Edna has been chosen for a complete analysis of the daily data. This station has an extensive record with a large variability of storms and there are no major reservoirs upstream of the station to complicate analysis. Furthermore, the Environmental Section of the TWDB recently keyed this station's data prior to 1965 into the computer. Thus, 44 years of daily suspended sediment data for the Lavaca River are available and have been obtained in digital format.

The Lavaca River watershed as defined by TWDB's Plate 1 (1990), and as shown in Figure 2-3, contains all of the Lavaca River, all of the Navidad River, and all of the associated tributaries. The watershed lies between the Colorado River and Guadalupe River watersheds. The Navidad River drains into Lake Texana just prior to discharging into the Lavaca River south of the town of Edna. The Lavaca River drains into Lavaca Bay which is connected to Matagorda Bay and then to the Gulf of Mexico.

Because the Edna sampling station is located upstream of the confluence of the Lavaca and Navidad Rivers, the watershed upstream of the sampling station comprises only a portion of the watershed shown in Figure 2-3. The extent of the watershed which drains to the gauging station located on the US Highway 59 bridge near Edna is shown in Figure 2-4.

The drainage area upstream of the station is 817 square miles. The annual average streamflow for the years of record is 2.29×10^5 acre-feet (316 cfs). The average annual suspended sediment load is 172 tons per square mile (163 mg/l).

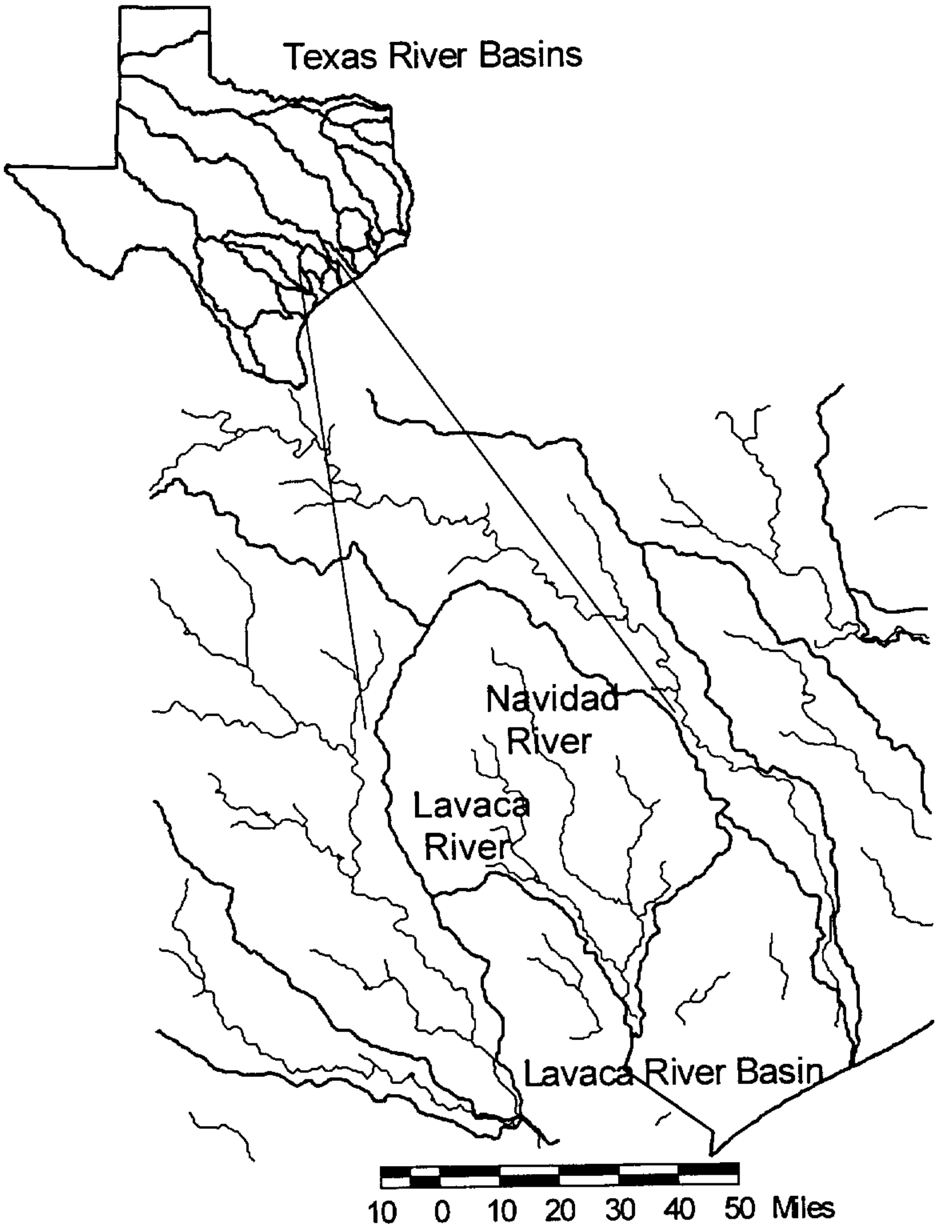


Figure 2-3. The Lavaca River Basin

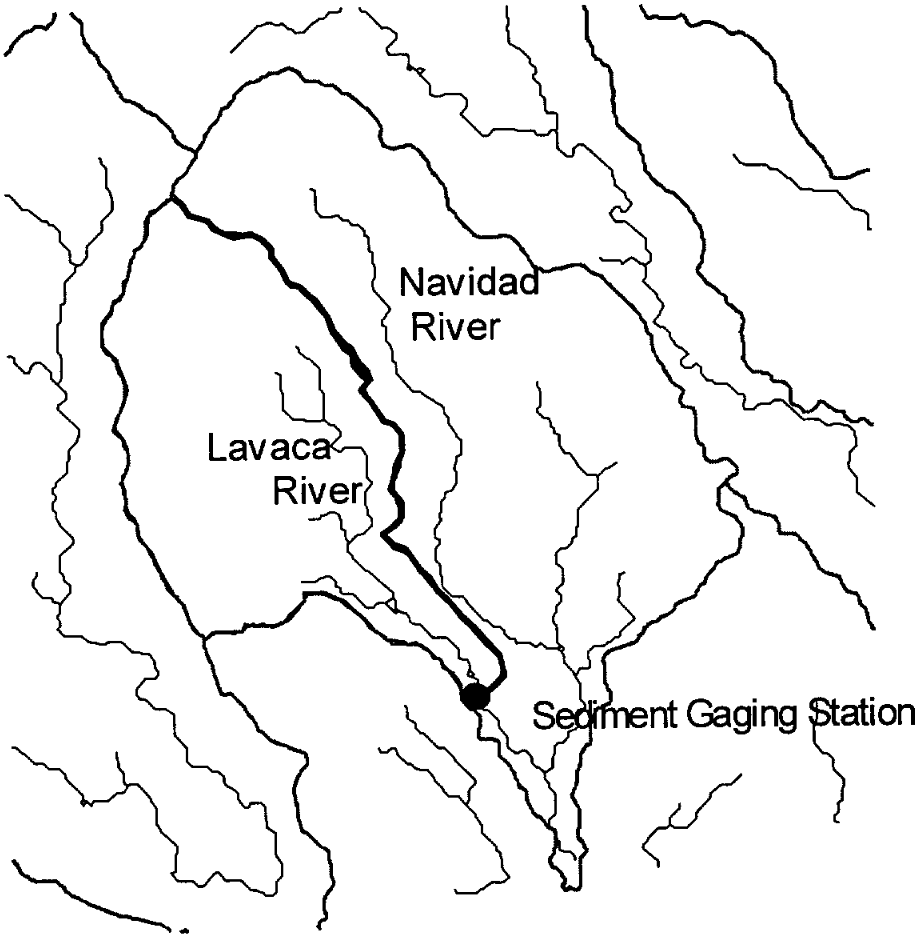


Figure 2-4. Lavaca Watershed above Edna

2.5.3 Suspended Sediment Size Analysis

Welborn (1961) initiated a study to compare results obtained from the Texas Sampler and the USGS depth integrated sampler. He found that no single coefficient can be used for all streams in Texas to determine a depth integrated sample using the Texas Sampler. Welborn found the greatest variation between the results obtained from the two types of samplers in sand-bed streams of southeast Texas. The suspended load of the Sabine, Neches, lower Trinity, and San Jacinto Rivers contain a high percentage of sand and require a multiplier greater that 1.102. For streams carrying higher percentages of silt and clay in suspension, the coefficients for correcting the suspended-sediment concentrations of samples collected with the Texas sampler are nearer unity than those for sandy streams.

Suspended-sediment samples for the Lavaca River near Edna were collected using both samplers in 1961 and 1962 to compare the results. Table 2-3 shows comparisons of suspended sediment concentrations for four sampling dates.

Table 2-3. Comparison of concentrations using two different samplers, Lavaca River near Edna

Sample date	Water temperature (°F)	Discharge (cfs)	Concentration (mg/l)	
			Depth-integrating sampler	Texas sampler
9/13/61	79	13600	244	224
11/14/61	65	7260	335	241
11/15/61	64	10400	103	108
9/19/62	81	920	674	675

Three out of four sampling dates show almost identical concentrations using the two sampling methods. The sample collected on November 14, 1961 using the Texas Sampler underestimates the depth integrated sample by 28 percent. More data would lead to more conclusions. Based on the overall results of the comparison, it is reasonable to believe that there may have been a sampling error for the date in question.

All samples were analyzed using sieves based on the British system with 0.0625 mm being the size able to pass the smallest sieve. The percentage of grains passing the smallest sieve is shown in Table 2-4. The percentages of fines collected using both methods are very close.

Table 2-4. Comparison of percent fines using two different samplers, Lavaca River near Edna

Sample date	% finer than 0.062 mm	
	Depth integrating sampler	Texas sampler
9/13/61	99	100
11/14/61	93	93
11/15/61	97	94
9/19/62	99	99

Welborn used both samples collected on September 13, 1961, as well as the depth integrated sample of November 14, 1961, for a full gradation analyses. Figure 2-5 plots the data obtained from the gradation analysis. It is quickly apparent, that based on this analysis, there is very little discrepancy between that collected from the depth integrating sampler and that collected from the Texas sampler.

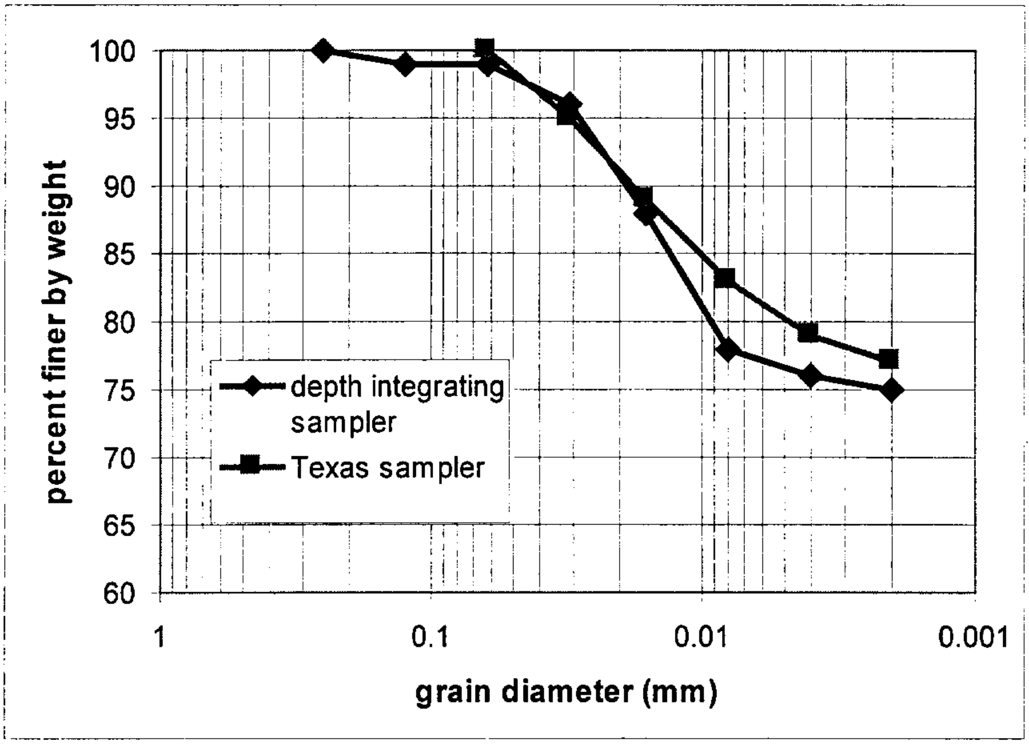


Figure 2-5. Gradation analysis of Lavaca River Samples, September 13, 1961

For his Master's Degree Thesis, James Anderson (1996) analyzed a bed sample of the Lavaca River near Edna. He used a sieve analysis to determine the gradation of the grains finer than 0.105 mm (#200 sieve). Figure 2-6 shows a comparison of Anderson's bed samples to the suspended sediment samples.

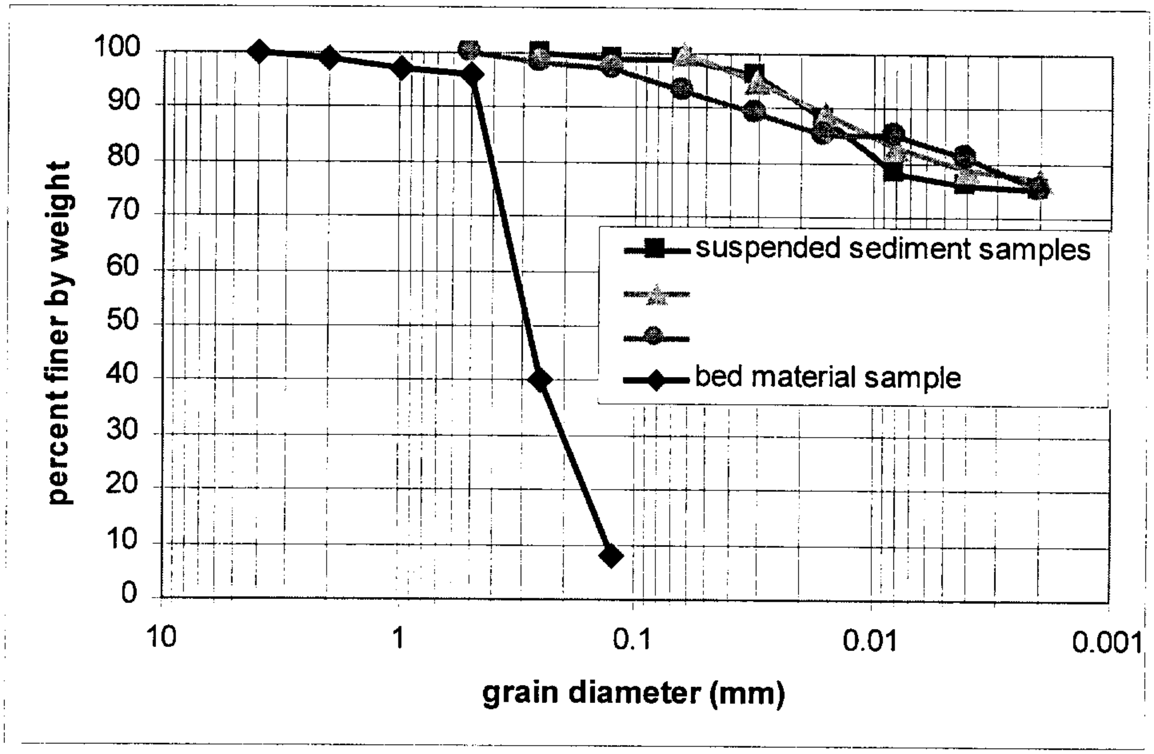


Figure 2-6. Comparison of bed material and suspended sediment, Lavaca River near Edna

Over 90% of the bed sample consisted of grains larger than 0.125 mm diameter. 60% of the grains are larger than 0.25 mm diameter. Anderson's bed samples upstream and downstream of the Edna station are similar. The suspended samples include little to no bed materials.

2.6 Current Work in Texas

Currently (1998) the TWDB is involved in a bathymetric survey program (Sullivan, 1994-1996) in which reservoirs are surveyed in detail using boats equipped with Global Positioning Survey equipment. The bathymetric survey determines the capacity of a reservoir and establishes baseline information for future surveys. Comparison of two surveys of the same reservoir will provide information on the sedimentation rate in the reservoir. It is expected that reservoir owners will have their reservoir surveyed again in 10-15 years or possibly after a large storm event. The surveys will be compared and the results will be used to help water managers better estimate rates at which the reservoir's volume is being depleted. Furthermore, deposition locations can be identified. This program is valuable in that reservoir owners have a more accurate measurement of storage capacity than in the past; however, the program will not aid in determining the capacity depletion of a reservoir after an episodic event unless a survey is conducted just before and just after the event. One rationale behind the program is that this is the most comprehensive way to understand how the total sediment load (suspended load and bedload) is contributing to the accumulated sediment in a reservoir. However, the suspended load is the major contributor to reservoir deposits (Reid and Frostick, 1994). In the meantime, the owners of these reservoirs may be able to benefit from the years of recorded suspended sediment concentrations if correlations can be made with other watershed characteristics.

3 RESEARCH METHODS

3.1 General

The computational tools used for this research include Fortran, Excel, Arc/Info, and SAS. Fortran and Excel were used primarily for data formatting, while Arc/Info and SAS were used for analysis. Collection of both temporal and spatial data, as shown in Table 3-1, was required. Suspended sediment and streamflow data are technically spatial data in that they apply to a certain location. However, the sediment and streamflow data are available only for the gauging station. Land use and land cover are normally considered to be spatial data; however, development, urbanization, crop rotation, and seasonal variation contribute a time variation. Similarly, the location of a dam is spatial data but some of the dams were constructed during the study period, thus adding a temporal component to the dams.

Table 3-1. Required data

Temporal Data	Spatial Data
suspended sediment concentration	[suspended sediment concentration]
streamflow	[streamflow]
rainfall	rainfall (temporally averaged)
	elevation
	soil type
[land use / land cover]	land use / land cover
[dams and reservoirs]	dams and reservoirs

A simplified version of the research process is shown in Figure 3-1.

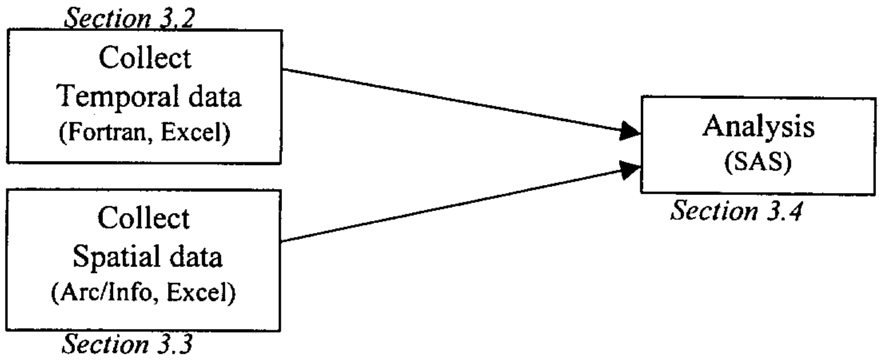
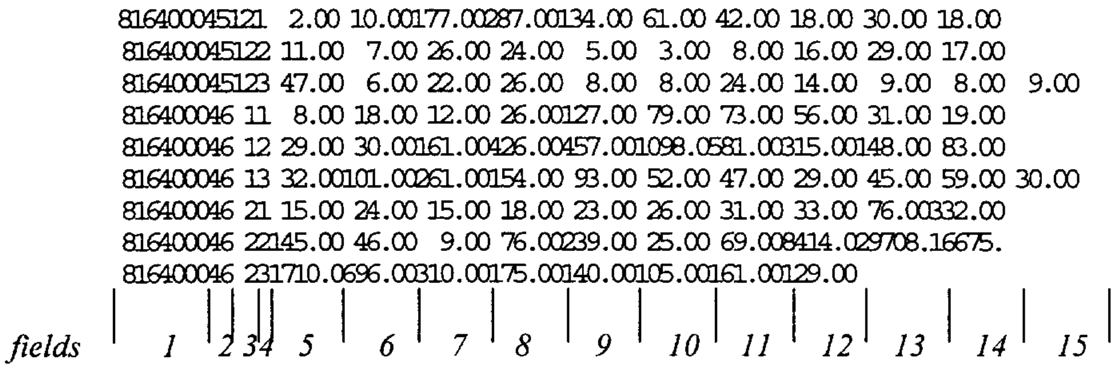


Figure 3-1. Simplified Research Process

3.2 Collection of temporal data

3.2.1 Suspended sediment data

The TWDB suspended sediment data were obtained from the Texas Natural Resource Conservation Commission (TNRCC) - Information Resources (phone no. 512-239-DATA). TNRCC provided both digital and printed data. The digital data were in a Fortran format. Figure 3-2 displays an example for the daily suspended sediment data for station no. 8164000 from January 1, 1945 to February 28, 1945.



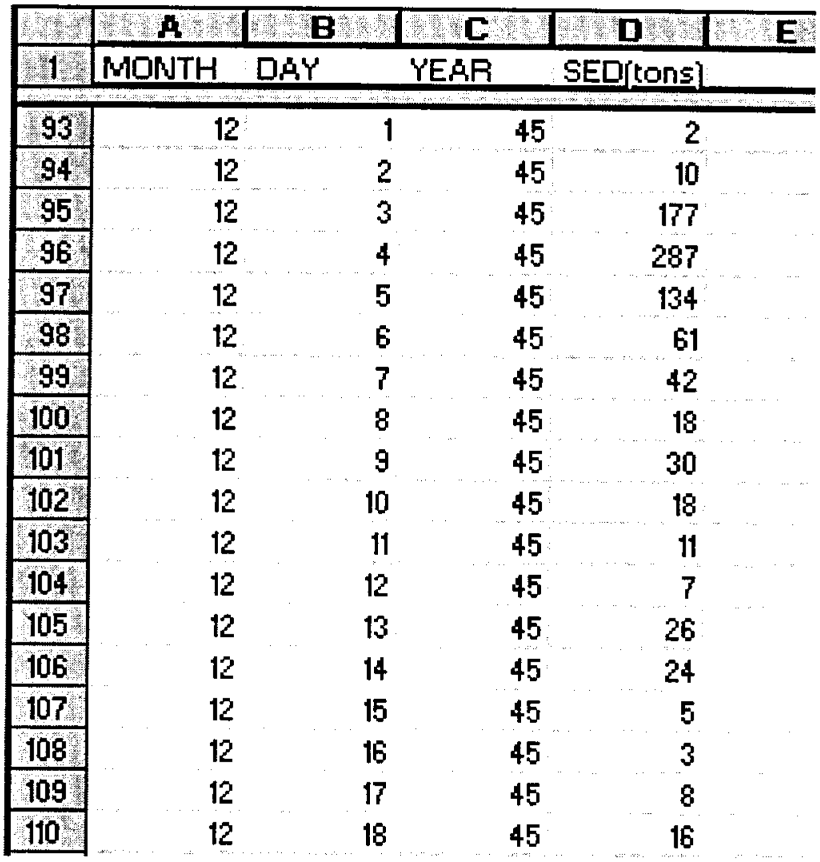
```
816400045121  2.00 10.00177.00287.00134.00 61.00 42.00 18.00 30.00 18.00
816400045122 11.00  7.00 26.00 24.00  5.00  3.00  8.00 16.00 29.00 17.00
816400045123 47.00  6.00 22.00 26.00  8.00  8.00 24.00 14.00  9.00  8.00  9.00
816400046 11  8.00 18.00 12.00 26.00127.00 79.00 73.00 56.00 31.00 19.00
816400046 12 29.00 30.00161.00426.00457.001098.0581.00315.00148.00 83.00
816400046 13 32.00101.00261.00154.00 93.00 52.00 47.00 29.00 45.00 59.00 30.00
816400046 21 15.00 24.00 15.00 18.00 23.00 26.00 31.00 33.00 76.00332.00
816400046 22145.00 46.00  9.00 76.00239.00 25.00 69.008414.029708.16675.
816400046 231710.0696.00310.00175.00140.00105.00161.00129.00
```

Figure 3-2. Daily suspended sediment data format

Field 1 is 8 digits and holds the station number which includes a space in column 1. Field number 2 is 2 digits and holds the last 2 digits of the year. The third field is two digits identifying the month. Field 4 is the card sequence number of either 1, 2, or 3 representing days 1 through 10, days 11 through 20, or day 21 through the end of the month, respectively. The fields on the remaining portion of each line are 6 digit fields that represent the daily sediment amount in tons. For card sequence numbers 1 and 2 there are always 10 remaining fields with sediment data. For card sequence number 3, the number of remaining fields can vary from 8 to 11. Missing data are indicated by -9999. There are a few instances where a daily load exceeded 999,999 tons. In these instances the field is replaced with ******. The printed information from TNRCC was not limited by column width and could be used to determine the values for days where the sediment load value is more than 6 digits. The daily data for every TWDB station for water year 1965 to water year 1989 is contained in a text file named seddata located on the CD for this research.

After reviewing the locations of the stations, the length of record and the variability in the measured sediment load associated with each station, station 8164000 (the Lavaca River near Edna) was chosen for daily data analysis. The data for the Lavaca River near Edna include some very large peaks in sediment load. The sediment load was measured from 1945 to 1989. There are no reservoirs in the watershed above Edna; thus the sediment load is not altered by dams. There are no large areas of urbanization and no major urban development took place during the period of record. The land use did not change dramatically during the period of record. Because the spatial characteristics of the land in the watershed were not subject to much change, the relationships between the temporal data of sediment load, streamflow, and rainfall can be better isolated. After the Lavaca station was chosen for analysis, the TWDB keyed in the 1945 to 1965 data; thus the entire period of record for the Lavaca River near Edna is available in digital format. The data for years 1945 to 1965 are contained in the file 8164000.wy4 on the CD available from this research.

A Fortran program was written to read the data and write the data into a columnar format to be used in a spreadsheet or database program. The result of using the Fortran program with the above data is shown in Figure 3-3, after being read into Excel. There are 16,101 days of data (although, a few of those days have missing data, i.e. -9999) during the period of record. Microsoft Excel 5.0 is limited to 16,384 rows. (Excel 97 now has up to 65,536 rows.)



	A	B	C	D	E
1	MONTH	DAY	YEAR	SED(tons)	
93	12	1	45	2	
94	12	2	45	10	
95	12	3	45	177	
96	12	4	45	287	
97	12	5	45	134	
98	12	6	45	61	
99	12	7	45	42	
100	12	8	45	18	
101	12	9	45	30	
102	12	10	45	18	
103	12	11	45	11	
104	12	12	45	7	
105	12	13	45	26	
106	12	14	45	24	
107	12	15	45	5	
108	12	16	45	3	
109	12	17	45	8	
110	12	18	45	16	

Figure 3-3. Daily sediment data in Excel

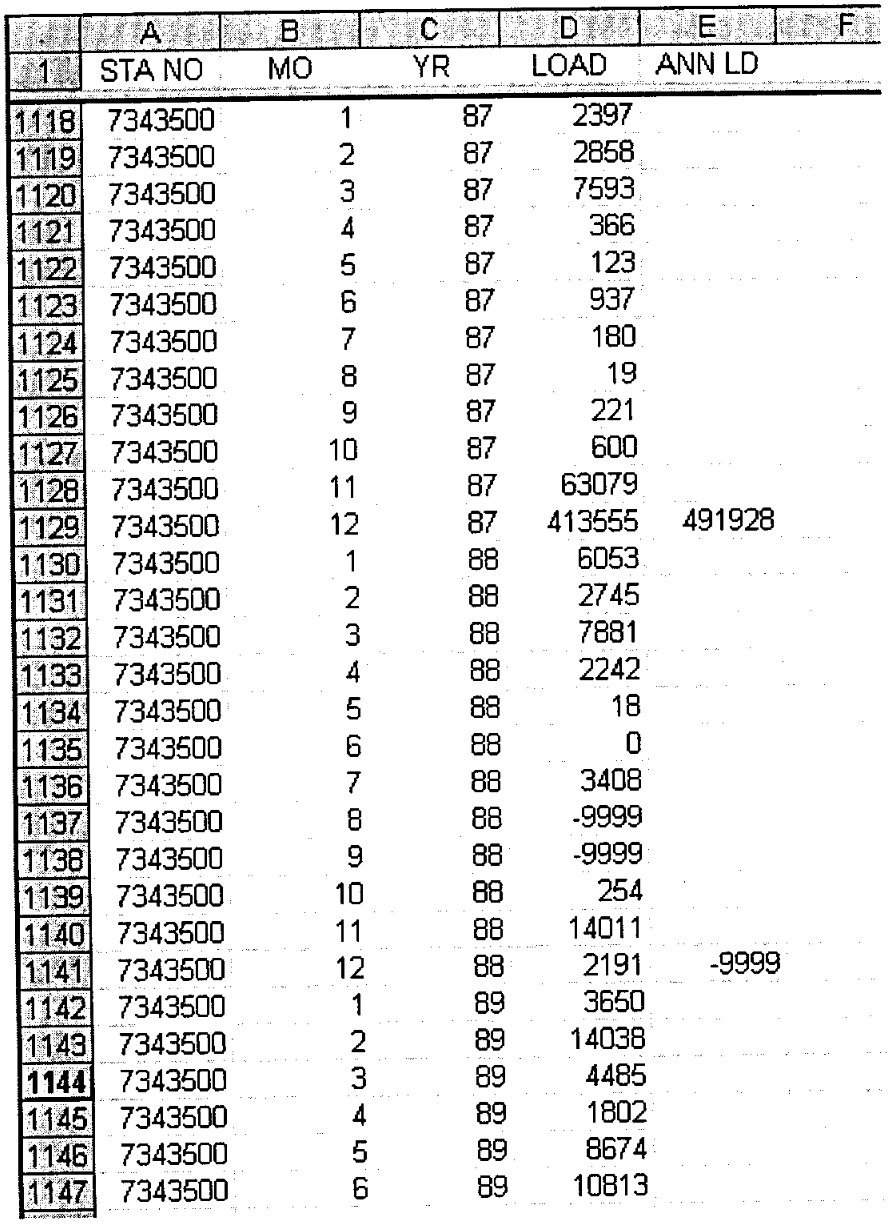
To compare sediment loads from across the state, the annual values were desired. No digital file of annual values was available but the monthly values were available. Figure 3-4 is an example for the monthly suspended sediment data for station 7343500 from January, 1987, to December, 1989, and for station 8022000 from January, 1968, to June, 1971. Refer to Table 2-2 for station numbers and names.

```
       7343500871   2397.00   2858.00   7593.00    366.00    123.00    937.00
       7343500872    180.00     19.00    221.00    600.00  63079.00 413555.00
       7343500881   6053.00   2745.00   7881.00   2242.00     18.00       .00
       7343500882   3408.00  -9999.00  -9999.00    254.00  14011.00   2191.00
       7343500891   3650.00  14038.00   4485.00   1802.00   8674.00  10813.00
       7343500892   2396.00    503.00     60.00  -9999.00  -9999.00  -9999.00
       8022000681  -9999.00  -9999.00  -9999.00  -9999.00  -9999.00  20943.00
       8022000682   3960.00    181.00   3520.00    342.00   7250.00  10870.00
       8022000691   2630.00  17800.00  38770.00  31350.00  38710.00  10730.00
       8022000692    406.00     49.00     73.00     87.00   3020.00   6080.00
       8022000701   9920.00   9210.00  34500.00  17490.00  19360.00   9940.00
       8022000702   1640.00    154.00   1090.00  17043.00   9169.00   2317.00
       8022000711   1350.00   6589.00   5236.00   1205.00   2855.00     56.00
       |       | ||        |         |         |         |         |         |
fields     1    2 3    4       5         6         7         8         9
```

Figure 3-4. Monthly suspended sediment data format

Similar to the daily data, the first field of 8 digits is the station number and the second field of 2 digits is the year. Field number 3 is a card sequence number with a value of 1 if data are from January through June and a value of 2 if data are from July through December. The monthly totals of sediment are in tons in the remaining six fields of 10 digits each. Again, missing data is indicated by -9999.00 and a value with 11 or more digits is indicated by **********. The digital file contains monthly data for every sediment station for the entire period of record. The data are stored in a text file named sedim.txt on the CD for this research. The

Fortran program, used to read the data, writes the data in a columnar format as shown in Figure 3-5, after being read into Excel.



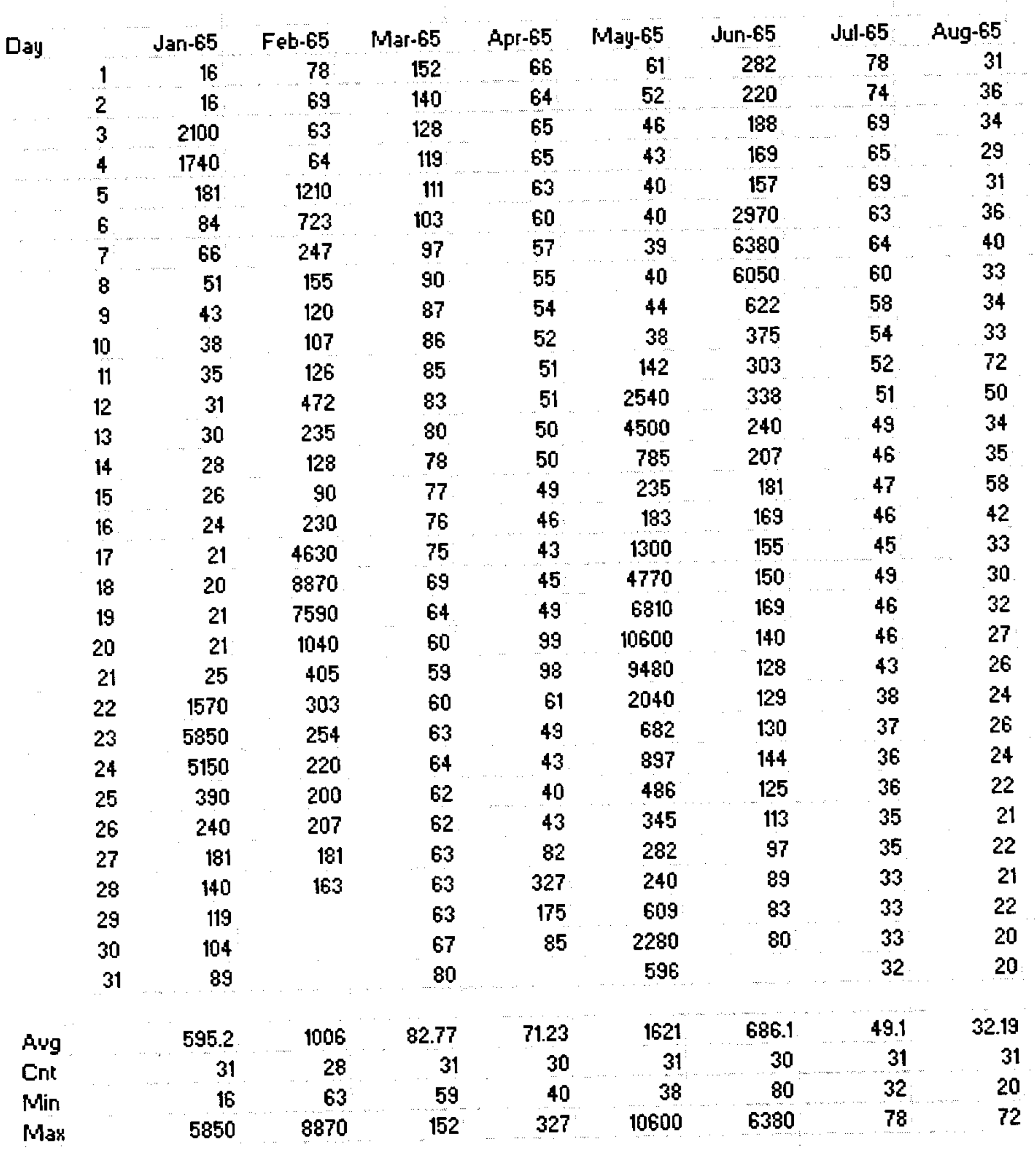
	A	B	C	D	E	F
1	STA NO	MO	YR	LOAD	ANN LD	
1118	7343500	1	87	2397		
1119	7343500	2	87	2858		
1120	7343500	3	87	7593		
1121	7343500	4	87	366		
1122	7343500	5	87	123		
1123	7343500	6	87	937		
1124	7343500	7	87	180		
1125	7343500	8	87	19		
1126	7343500	9	87	221		
1127	7343500	10	87	600		
1128	7343500	11	87	63079		
1129	7343500	12	87	413555	491928	
1130	7343500	1	88	6053		
1131	7343500	2	88	2745		
1132	7343500	3	88	7881		
1133	7343500	4	88	2242		
1134	7343500	5	88	18		
1135	7343500	6	88	0		
1136	7343500	7	88	3408		
1137	7343500	8	88	-9999		
1138	7343500	9	88	-9999		
1139	7343500	10	88	254		
1140	7343500	11	88	14011		
1141	7343500	12	88	2191	-9999	
1142	7343500	1	89	3650		
1143	7343500	2	89	14038		
1144	7343500	3	89	4485		
1145	7343500	4	89	1802		
1146	7343500	5	89	8674		
1147	7343500	6	89	10813		

Figure 3-5. Monthly sediment data in column format

The monthly values were added together (within the spreadsheet mondat.xls) for each calendar year to determine annual values. If any month of the year contained missing data (i.e. -9999.00) then the annual value was also missing and took on the value of -9999. To ensure against errors, the annual values were then compared with the printed reports received from TNRCC. There were a few incidents where a value was keyed in to replace ********** or -9999.00. Because the sediment data since 1982 was not published (TWDB Report 306 included data through September, 1982), it was necessary to use the digital data to compute the average annual sediment load for each station. The averages were computed in Excel. A new worksheet was created with station number, average annual sediment load and the number of years used for the average.

3.2.2 Streamflow data

Two sources were used for obtaining streamflow data: Hydrosphere CD-ROM and the Internet. Both data sources originate from the USGS WATSTORE system. Daily streamflow values for the Lavaca River near Edna, as well as monthly streamflow values for all of the sediment stations were required. Fortran programs were again written and used to change the data format. The format for the daily streamflow values was as shown in Figure 3-6.



Day	Jan-65	Feb-65	Mar-65	Apr-65	May-65	Jun-65	Jul-65	Aug-65
1	16	78	152	66	61	282	78	31
2	16	69	140	64	52	220	74	36
3	2100	63	128	65	46	188	69	34
4	1740	64	119	65	43	169	65	29
5	181	1210	111	63	40	157	69	31
6	84	723	103	60	40	2970	63	36
7	66	247	97	57	39	6380	64	40
8	51	155	90	55	40	6050	60	33
9	43	120	87	54	44	622	58	34
10	38	107	86	52	38	375	54	33
11	35	126	85	51	142	303	52	72
12	31	472	83	51	2540	338	51	50
13	30	235	80	50	4500	240	49	34
14	28	128	78	50	785	207	46	35
15	26	90	77	49	235	181	47	58
16	24	230	76	46	183	169	46	42
17	21	4630	75	43	1300	155	45	33
18	20	8870	69	45	4770	150	49	30
19	21	7590	64	49	6810	169	46	32
20	21	1040	60	99	10600	140	46	27
21	25	405	59	98	9480	128	43	26
22	1570	303	60	61	2040	129	38	24
23	5850	254	63	49	682	130	37	26
24	5150	220	64	43	897	144	36	24
25	390	200	62	40	486	125	36	22
26	240	207	62	43	345	113	35	21
27	181	181	63	82	282	97	35	22
28	140	163	63	327	240	89	33	21
29	119		63	175	609	83	33	22
30	104		67	85	2280	80	33	20
31	89		80		596		32	20
Avg	595.2	1006	82.77	71.23	1621	686.1	49.1	32.19
Cnt	31	28	31	30	31	30	31	31
Min	16	63	59	40	38	80	32	20
Max	5850	8870	152	327	10600	6380	78	72

Figure 3-6. Daily streamflow format

The streamflow data are in cfs. The data were read by a Fortran program and then written in a column format similarly to the sediment data shown in Figure 3-3. To determine concentration, it is necessary to have the streamflow and the load in the same file. The data from one spreadsheet were copied into the other. After insuring the dates matched, the duplicate date columns were deleted.

As displayed in Figure 3-6, the daily streamflow table includes the monthly averages. The monthly averages were used to compute the annual streamflow at each sediment gauging station.

3.2.3 Rainfall data

Rainfall gages are located across the state. The temporal and spatial resolution are both sporadic. The processing of temporally averaged rainfall data is discussed in Section 3.3. The Hydrosphere CD-ROM contains daily rainfall at 17,000 NCDC stations in Texas. As shown in Figure 3-7, there are three rainfall stations located in the Lavaca River basin above Edna: at Yoakum, Hallettesville, and Edna.

The daily rainfall (total rainfall) for each of these three stations was retrieved. The format for the precipitation was identical to that for the streamflow (Figure 3-6) . The data were read using a Fortran program and written to a column format that could be used with the sediment and streamflow data.

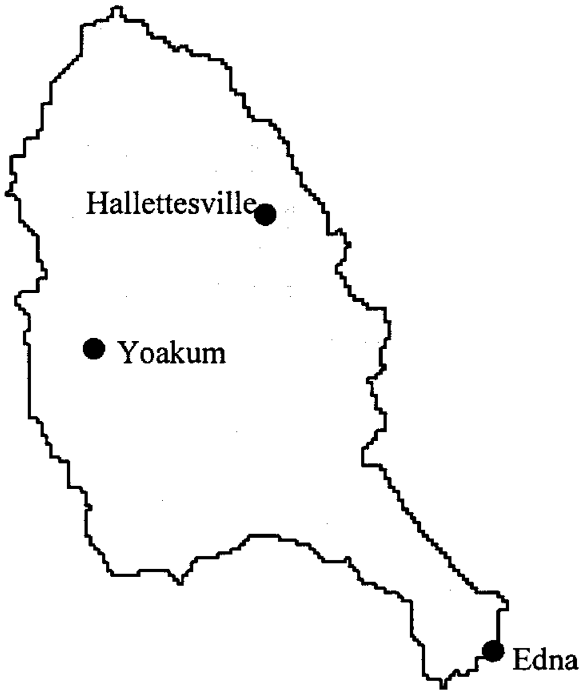
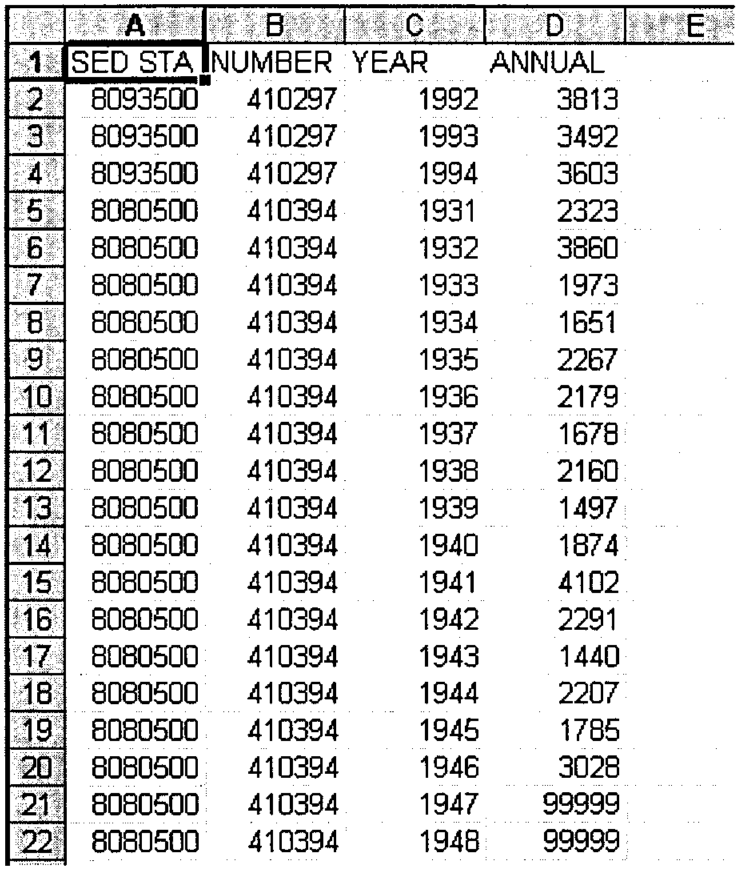


Figure 3-7. Location of rainfall stations in Lavaca basin

Processing of annual rainfall values for the rest of the study area was much more difficult. Spatially distributed annually averaged rainfall is available through the PRISM project (Daly, Neilson, Phillips, 1994), but does not show year-to-year variability. The annual rainfall for the NCDC station closest to each sediment gauging station was determined, then the annual rainfall for those stations was retrieved. The NCDC station number corresponding to the USGS gauging station number was recorded in a spreadsheet along with annual precipitation as shown in Figure 3-8. The heading in column A is the sediment station gage number, in column B is the NCDC gage number which is close to the sediment station, in column C is the year for which the data corresponds, and in column D is the annual precipitation total values in 0.01 inches. The annual rainfalls were copied from the Internet (http://www.ncdc.noaa.gov/coop-precip.html) and pasted into the spreadsheet with the sediment gauging stations. The values were then checked manually to insure everything was copied into the appropriate position.



	A	B	C	D	E
1	SED STA	NUMBER	YEAR	ANNUAL	
2	8093500	410297	1992	3813	
3	8093500	410297	1993	3492	
4	8093500	410297	1994	3603	
5	8080500	410394	1931	2323	
6	8080500	410394	1932	3860	
7	8080500	410394	1933	1973	
8	8080500	410394	1934	1651	
9	8080500	410394	1935	2267	
10	8080500	410394	1936	2179	
11	8080500	410394	1937	1678	
12	8080500	410394	1938	2160	
13	8080500	410394	1939	1497	
14	8080500	410394	1940	1874	
15	8080500	410394	1941	4102	
16	8080500	410394	1942	2291	
17	8080500	410394	1943	1440	
18	8080500	410394	1944	2207	
19	8080500	410394	1945	1785	
20	8080500	410394	1946	3028	
21	8080500	410394	1947	99999	
22	8080500	410394	1948	99999	

Figure 3-8. Annual rainfall

3.3 Collection of spatial data

To correlate loads with watershed characteristics, data on the watershed characteristics had to be obtained. Watershed data could have been obtained from a variety of methods once the appropriate mapping could be obtained. Alternatively, a Geographic Information System (GIS) could be used. Although development of GIS to characterize one watershed may take longer than determining the same parameters using conventional methods, the developed GIS routines can be used to analyze subsequent watersheds with little effort. Thus, GIS was used to characterize watersheds.

The DEC Alpha Workstations, located in the Learning Resource Center in the Department of Civil Engineering at the University of Texas, were used for all of the GIS work. These workstations have an Alpha EV4 processor. The machines are used for many different applications in civil engineering and by many different users; therefore disk space is an issue. Because the GIS work for this research is very space intensive, a temp (temporary) directory was used. Normally, this practice should be avoided because the temp directories are not backed up. This work required approximately 1 GB of disk space.

Arc/Info and ArcView are GIS software packages developed by Environmental Systems Research Institute (ESRI). Arc/Info Version 7 was used for all GIS analyses. ArcView Version 3 was used to display the results of the analyses.

Data are stored in layers in a GIS. If the layer consists of points, arcs, and/or polygons, then the layer is called a cover, or coverage. Alternatively, data can be stored in a raster or square-cell format called a grid. A simple command converts a coverage to a grid and vice versa. Arc/Info provides several hydrologic-analysis tools in the GRID module. Arc/Info is essentially a command line program. The Arc Macro Language (AML) can be used to perform a series of commands. Furthermore, DO loops along with local and global variables can be used in an AML. In addition to GRID and AML, the TABLES module was used frequently.

Table 3-1 lists seven types of spatial data. Those types are repeated in Table 3-2 along with the GIS format of the data. An AML was written for each spatial characteristic to be computed from the above GIS covers and grids. All AML's are included on the CD available for this research. The boundaries of each watershed were determined using the point coverage of the stations and the 500 meter DEM. Once the watershed boundaries were determined they were used in combination with the other covers and grids to define all of the watershed characteristics.

Table 3-2. Sources of Spatial Data

Spatial Data	Source	Class	Attribute
sediment and streamflow gauging stations	latitude and longitude	point cover created	station no.
rainfall	PRISM	grid	average annual rainfall (mm)
elevation	USGS	DEM (grid with elevation)	elevation (m)
soil type	NRCS	STATSGO (polygon cover)	Hydrologic Soil Group K factor
land use / land cover	NRCS	polygon cover	Anderson Land Use Code
dams and reservoirs	TWDB	polygon cover	dam specifications

Table 3-3 lists the main data source used to determine each watershed characteristic. Intermediate covers or grids are not included in this list.

3.3.1 Location of Gauging Stations

The only spatial characteristic associated with the sediment data *per se* is the location where the sediment was measured. The locations of the sediment gauging stations correspond with the location of the stream gauging station. The gauging stations were numbered consecutively from 1 to 61 in the order of increasing number associated with the USGS gauging no. Figure 3-9 displays a portion of the text file with this numbering system.

Table 3-3. Main Data Sources for Watershed Characteristics

Watershed Characteristic	Required Grid or Cover
Extent of watershed (watershed boundaries, area, and perimeter)	stations, DEM, streams
length	stations, DEM
soilA, soilB, soilC, soilD	STATSGO
k_avg, k_max	STATSGO
streams	streams
elev_ave, elev_min, elev_max	DEM
slope_ave, slope_max	DEM
rain_sta, rain_avg, rain_max, rain_min	rainfall
%LU1, %LU2, %LU3, %LU4, %LU5, %LU6	land use
river slope	DEM
fracarea	dams, DEM
resleng	dams, DEM

length = hydrologic length
soilA, soilB, soilC, soilD = fraction of NRCS hyrologic soil groups A, B, C, D
k_avg, k_max = average soil erodibility, maximum soil erodibility
elev_ave, elev_min, elev_max = average, minimum, and maximum elevation
slope_ave,slope_max = average and maximum slope from one grid cell to the next
rain_sta, rain_max, rain_min = temporally averaged value of rainfall at the station, maximum, and minimum temporally averaged rainfall
%LU1, %LU2, %LU3, %LU4, %LU5, %LU6 = the percentage of the watershed with Anderson Land Use Codes 1, 2, 3, 4, 5, 6, respectively.
fracarea = fraction of area that feeds upstream reservoirs
resleng = distance from gauging station to a reservoir in the basin

```
1     7299200
2     7308000
3     7331600
4     7336820
5     7342500
6     7343200
7     7343500
8     8022000
         .
         .
         .
55    8183500
56    8186000
57    8188500
58    8194000
59    8207000
60    8210000
61    8210500
```

Figure 3-9. Numerical Text File of Gauging Stations

A text file with the latitude and longitude of each sediment gauging station was also created as well as a text file with the consecutive number and the name of the station. A portion of the file with the latitude and longitude of each station is shown in Figure 3-10. Note that the longitude (horizontal coordinate) is given first. The latitude and longitude of each station are available in the printed reports received from TNRCC as well as the TWDB published reports. Furthermore, since the sediment stations correspond with USGS streamflow stations, the locations can be obtained from the USGS internet site (http://txwww.cr.usgs.gov/cgi-bin/txnwis) or the USGS Water Resources Data publications.

1	-100.745	34.57305
2	-99.6833	34.1
3	-96.5631	33.81889
4	-94.6942	33.6875
5	-95.5942	33.35555
6	-95.1325	33.38638
7	-95.0925	33.32222
8	-94.4578	32.36972
	⋮	
55	-98.0639	28.95139
56	-97.93	29.01388
57	-97.3844	28.64944
58	-99.2406	28.42555
59	-98.3464	28.49194
60	-98.185	28.43611
61	-97.86	28.03805

Figure 3-10. Longitude and Latitude of Sediment Gauging Stations

A point coverage of the 61 stations was built in Arc/Info and then projected into Texas Albers. The Texas Albers projection preserves area. Area is the most important map measurement to preserve because it affects the volume of rainfall and the amount of sediment available for supply. The Albers projection is conical, in which the meridians are straight lines meeting in a common point beyond the limits of the map. The parallels are concentric circles, the center of which is at the point of intersection of the meridians. The meridians and the parallels intersect at right angles and the arcs of longitude along any given parallel are of equal length. The spheroid is intersected by a cone at two parallels know as the standard parallels for the area to be represented. On the two standard parallels, arcs of longitude are represented by their true lengths, or at an exact scale. Between the standard parallels, the scale along the meridians is too large and beyond them too small (Deetz and Adams, 1969). The Albers projection is constructed in such a way that the area of the earth's surface between any pair of parallels and meridians is correctly preserved in the flat map representation. The projection parameters for the Texas Albers projection are as follows:

Units: meters
Datum: nad83
First standard parallel: 27 25 00
Second standard parallel: 34 55 00
Longitude of central meridian: -100 00 00
Latitude of projection's origin: 31 10 00
False easting: 1000000.0
False northing: 1000000.0

After the point coverage of the stations was created, the coverage could be used in conjunction with the DEM to determine the watershed boundaries.

3.3.2 Watershed boundaries

Digital Elevation Models (DEMs) can be used to define topography within a GIS. DEMs are available from the USGS via the Internet:

(http://edcwww.cr.usgs.gov/glis/hyper/guide/usgs_dem).

The models were developed by digitizing quad sheets at various scales. Because the watersheds associated with the sediment stations cover almost all of Texas and small portions of neighboring states, a small scale DEM was chosen.

A 15-arc-second DEM corresponds to 1:2,000,000 USGS quad sheets. When projected to spheroid coordinates, a 500 meter grid cell resolution is comparable to 15 seconds. Because the DEM was used for watershed delineation, it was necessary to adjust the DEM by filling in the spurious pits and by "burning in" the streams, as explained below. Both of these procedures were used with a 500 meter DEM of Texas for the "Spatial Water Balance of Texas" (Reed, Maidment, and Patoux, 1997). The same processed DEM was used for this research.

The FILL command in the GRID module of Arc/Info fills sinks or levels peaks in a continuous grid to remove small imperfections in the data. Sinks and peaks are often errors in data due to resolution of the data or rounding of elevations to the nearest integer value. Sinks should be filled to ensure proper delineation of basins and streams. If the sinks are not filled, a derived drainage network may be discontinuous. The DEM was also modified so that streams delineated from the DEM are consistent with digitized streams in EPA's River Reach File 1 (RF1). This process of DEM modification is called "burning in the streams."

The simplest stream burn-in procedure involves (1) creating a gridded representation of the digitized stream network (RF1) and identifying cells as being either stream cells or land surface cells, (2) raising the elevation of land surface cells relative to stream cells, and (3) deriving the drainage network based upon flow direction values defined by the burned DEM. Because many arcs in RF1 are not connected to the major river systems, burning in these arcs creates inland drainage basins or pits. Some of these disconnected arcs may represent real inland drainage basins or playas; however, information as to where inland drainage occurs was not readily available, so the DEM was filled a second time to eliminate pits created by the burning procedure. One drawback of the DEM analysis used here is that non-contributing drainage areas are considered as contributing to downstream runoff. This situation may cause the estimates of runoff per unit area in some watersheds to be to low; however, the runoff total at watershed outlets is still consistent with measured flows. A grid of stream cells contains those cells with a flow accumulation greater than 1000 cells or 250 km^2. A 1000 cell threshold was chosen because watersheds delineated with fewer than 1000 cells tend to be poorly defined (Reed, Maidment, and Patoux, 1997).

After the DEM was processed (i.e., the pits were filled and the streams were dug), then flow direction and flow accumulation grids were created. A flow direction grid simply defines which way water flows out of each grid cell. A grid is nothing more than a group of cells, located spatially, with each cell having either a numerical value or "NO DATA." Thus, a flow direction grid has numerical values corresponding to direction of flow. The eight-direction model is shown in Figure 3-11.

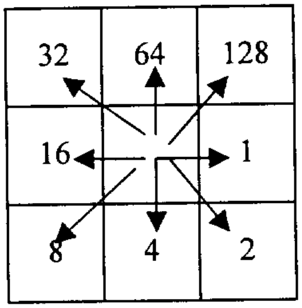


Figure 3-11. Eight Direction Pour Point Model in GRID

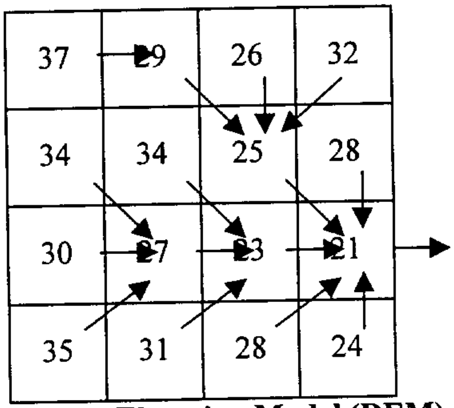
A sample 4 cell by 4 cell grid is shown in Figure 3-12 to illustrate GRID functions and commands. When applying the flowdirection command to the DEM, a new grid is created as shown by (A) and (B) of Figure 3-12. A flow accumulation grid can be created from the flow direction grid. The accumulated flow is based upon the number of cells flowing into each cell in the output grid. The current processing cell is not considered in this accumulation. Output cells with a high flow accumulation are areas of concentrated flow and may be used to identify stream channels. Figure 3-12(C) exhibits an example of a flow accumulation grid. Output cells with a flow accumulation of zero are local topographic highs and may be used to identify ridges. The results of FLOWACCUMULATION can be used to create a stream network by applying a threshold value to select cells with a high accumulated flow. An example is shown in Figure 3-12(D) where the expression to create a grid with the value 1 represents the stream network on a background of

NODATA. A flow direction grid and a flow accumulation grid were necessarily created so that watersheds for each station could be delineated.

The 500-meter DEM for Texas requires 9 MB of disk storage space, while the flow direction grid takes 14 MB and the flow accumulation grid requires 40 MB. Because the flow direction and flow accumulation grids are easily replicated, once they were used, they were deleted.

The point coverage of the sediment gages was converted to a grid coverage. The watershed function in GRID determines the contributing area above a set of cells in a grid. The flow direction grid and an outlet grid are required. Figure 3-12 (E) and (F) display an example of an outlet grid and a watershed grid. The numbers 12, 20, and 31 are simply identity numbers. Note that three adjacent outlet cells are shown with outlet 12 being downstream of 20 and 31.

The area of each watershed can be computed in at least two ways. The value of the flow accumulation grid, corresponding to the cell containing the gauging station, can be multiplied by the area of the grid cell. Alternatively, the grid can be converted to a polygon coverage. For this work, the grids were converted to polygon coverages. The polygon attribute table (PAT) contains the area and perimeter for each polygon in the coverage. Consider the watershed grid shown in Figure 3-12(F). When this sample grid is converted to a polygon coverage, three polygons will be created. The area represented by polygon 12



37	29	26	32
34	34	25	28
30	27	23	21
35	31	28	24

(A) Digital Elevation Model (DEM)

1	2	4	8
2	2	2	4
1	1	1	1
128	128	128	64

(B) Flow Direction Grid

0	1	0	0
0	0	4	0
0	3	6	15
0	0	0	0

(C) Flow Accumulation Grid

*	*	*	*
*	*	1	*
*	1	1	1
*	*	*	*

* = NO DATA

(D) Stream Grid

*	*	*	*
*	*	20	*
*	*	31	12
*	*	*	*

* = NO DATA

(E) Outlet Grid

20	20	20	20
31	31	20	12
31	31	31	12
31	31	12	12

(F) Watershed Grid

Figure 3-12. Sample 4 x 4 Grids

will include the area of the four grid cells with value 12. The actual contributing area includes the nested watersheds with identities 20 and 31.

For this research, it was desired to characterize the watershed for each gage separately. After creating a watershed coverage with all of the stations, it was determined which watersheds were nested in other watersheds. In the example above, watersheds 20 and 31 are nested in watershed 12. Six separate text files were created so that there were no nested watersheds in any one file. The procedure of creating a point coverage, an outlet grid, and subsequent watershed grid was repeated for each file. Then when the watershed grids were converted to polygon coverages, the areas in the PAT were for the entire watershed.

The area of the watershed contributing to each gauging station was the first parameter to compute. If the area computed for a station compared reasonably with the USGS published area for a station then it was assumed that the station was located correctly. Conversely, if the area did not compare, then the location of the station was incorrect. The most common reason for a station to have an incorrect area was the occurrence of the cell containing the gage not falling on the stream. The point coverage of the stations was edited over and over until the areas of the corresponding watersheds were considered reasonable.

The location of each sediment gage was reviewed using an AML developed by Seann Reed (Ph.D. student at the University of Texas at Austin). The AML works in GRID. A view zoomed in on the gauging station is created and the station and streams are drawn. The macro asks the user if the station should be moved. If the user responds 'yes' then the user is prompted to use the mouse to choose the cell on the stream where the gage should be located. As part of this work, the AML was modified to also draw major roads and reservoirs. These steps are necessary to insure the correct watershed for each gage can be identified. If a gage did not fall on the stream then an incorrect watershed was delineated. The roads and reservoirs were also useful in determining locations as many of the stations were on bridges or dam outlets. In several instances, it was difficult to determine from visual inspection of the computer drawing whether a gage should be located upstream or downstream of a tributary junction. To remedy this dilemma, two approaches were taken. First, the USGS home page on the Internet allows the user to see a map of the gage. The map can be magnified, printed, then compared with the screen image produced by the AML. Second, a watershed can be delineated for several points along a stream. The watershed area for those points can be compared with the published USGS gauging station area.

With accurate representation of the contributing area in GIS, it became possible to calculate parameters that would characterize each watershed. For some parameters it was easier to compute the parameter for one individual watershed at a time. To allow for this procedure, a separate polygon coverage was created for each station. To create separate polygon coverages for each station, separate grids for each station had to first be created. Then the grid could be converted to a polygon coverage. Basically two GRID commands were required for each watershed:

```
Grid:  shed1 = con(shedcov1 == 1,1)
Grid:  gridpoly shed1 poly1
```

Five AMLs were written so that individual polygon coverages could be obtained. A different AML was required for five of the six grids that contained no nested watersheds. The AMLs are indshed1.aml, indshed2.aml, indshed3.aml, indshed4.aml, and indshed56.aml. After confidence had been established in the areas of each watershed and the individual coverages had been created, it became possible to characterize the watersheds.

3.3.3 Watershed Flow Length

The function FLOWLENGTH calculates upstream or downstream distance along a flow path for each cell. A primary use of the FLOWLENGTH function is to calculate the length of the longest flow path within a given basin. This measure is often used to calculate the time of concentration of a basin. An AML was written to determine the flow length of a basin and to write the length to a file. A loop, shown in Figure 3-13, is created so that the length can be computed for all 61 watersheds. Notice the third line says which AML to run. This loop can be used with other AMLs by simply specifying a different AML to run.

```
&sv i = 1
&do while(%i%<62)
&r length.aml %i%
&sv i = %i% + 1
&end
```

Figure 3-13. AML Loop

A flow direction grid is required to compute the flow length. Since, the flow length for each individual watershed was required, a flow direction grid for each watershed was needed. There was not enough disk space to store the 61 individual flow direction grids so the AML creates and kills the flow direction grids as well as the flow length grids. Length.aml is listed in Figure 3-14.

```
&arg 1
grid
kill flowdir all
mape poly%1%
setwindow poly%1%
flowdir = txmfd
kill downgrid all
downgrid = flowlength(flowdir, #, DOWNSTREAM)
length%1% = zonalstats(shed%1%, downgrid, max)
quit
tables
select length%1%
alter max, hydleng, 16,,,,
unload length.txt
quit
&return
```

Figure 3-14. AML to compute flow length for one watershed

After a flow direction grid for the individual watershed is computed, the flow length grid can be created. Figure 3-15 shows a flowlength grid created for watershed 12 of the 4 x 4 cell grid that was presented in Figure 3-12.

0	500	0	0
0	0	1207	0
0	707	1207	1914
0	0	0	0

Figure 3-15. Flow Length Grid

The zonalstats function records in an output INFO table the specified statistics of the values of all cells in the value grid that belong to the same zone. A zonal grid is an integer grid that identifies the zone for each cell. A value grid is an integer or floating-point grid that defines the values of the cells that are to be used in the zonal calculations. To compute the flow length, the watershed grid is the zonal grid while the flowlength grid was the value grid. The maximum value of the flowlength grid is the flow length of the watershed. This value was written to a text file. Because the flow length was computed for the watershed in numerical order, the text file was a column of lengths that could easily by added to the PAT.

3.3.4 Watershed Soil Types

All soils information was obtained from the State Soil Geographic (STATSGO) Data Base available from the Natural Resource Conservation Service (NRCS) (www.ftw.nrcs.usda.gov/nsdi_node.html). Soil maps for the STATSGO data base were made by the NRCS by generalizing the detailed soil survey data. The mapping scale for STATSGO map is 1:250,000. The level of mapping is designed to be used for broad planning and management uses covering state, regional, and multi-state areas. The number of soil polygons per quadrangle map is between 100 and 400. The minimum area mapped is about 1,544 acres.

Each STATSGO map is linked to the Soil Interpretations Record attribute data base. The attribute data base gives the proportionate extent of the component soils and their properties for each map unit. The STATSGO map units consist of 1 to 21 components each. The Soil Interpretations Record data base includes over 25 physical and chemical soil properties, interpretations, and productivity. Examples of information that can be queried from the data base are available water capacity, soil reaction, salinity, flooding, water table, bedrock, and interpretations for engineering uses, cropland, woodland, rangeland, pastureland, wildlife, and recreation development. STATSGO data are available as an ArcInfo 7.0 coverage and can be downloaded from the NRCS ftp site (ftp.ftw.nrcs.usda.gov/put/statsgo/unix).

When using STATSGO, the first step is choosing which parameters to extract from the INFO files and add as attributes to the coverage. For this research two soil properties were chosen, soil erodibility and hydrologic soil group. Soil erodibility, or K factor, is used in the Universal Soil Loss Equation (USLE). The K factor is a function of the percent of silt, the percent of coarse sand, the soil structure, the permeability of the soil, and the percent of organic matter (Wischmeier, Johnson, and Cross, 1971).

Hydrologic Soil Group (HSG) is an indicator of the perviousness of the soils. HSG A is the most pervious while D is the least pervious. Hydrologic soil groups are used frequently in rainfall/runoff programs. Because D soils provide little infiltration, they are dominated with clays. Similarly, A soils are dominated by sands that have a low runoff potential. An AML was used to create a K-factor grid and to create a soils coverage containing attributes of the percent of each hydrologic soil group. Figure 3-16 lists the aml used to create a K-factor grid for the research area.

After the K-factor grid was computed, a simple aml was used to determine the averages, maximums, and minimums of K-factor for each watershed. The ZONALSTATS function was used with the watershed grid being the zone grid and the K-factor grid was the value grid.

```
/* this aml creates a kfactor grid coverage
/* from statsgo
copy /res2/maidment/statsgo/statsgo
copyinfo /res2/maidment/statsgo/layer
copyinfo /res2/maidment/statsgo/comp
tables
additem comp kfact 6 6 n 3 seqnum
sel comp
res layernum = 1
relate add
laycomp
comp
info
museq
museq
ordered
rw

calc laycomp//kfact = laycomp//kfact + kfact
relate drop
laycomp

additem comp fkfact 6 6 n 3 kfact
statistics muid kfact.sta
sum fkfact
end
q
joinitem statsgo.pat kfact.sta statsgo.pat muid muid ordered
```

Figure 3-16. K-factor AML

The procedure to compute the percentage of each hydrologic soil group for each watershed is slightly more difficult; however, the procedure is well documented in Dr. David Maidment's Web Site (http://www.ce.utexas.edu/prof/maidment/CE397/statsgo/viewstat.htm). Figure 3-17 displays a portion of the aml used to compute the percent hydrologic soil group for one watershed.

Because the STATSGO coverage has many map units, it is necessary to compute the area of each HSG for each map unit and then add those areas for each watershed. The procedure is repeated for each of the four HSGs, then the area of each HSG is divided by the total area to determine the percentage of each HSG in each watershed.

```
clip statsgo polyl soilclip
tables
additem soilclip.pat A_area 8 18 f 5
additem soilclip.pat B_area 8 18 f 5
additem soilclip.pat C_area 8 18 f 5
additem soilclip.pat D_area 8 18 f 5
select soilclip.pat
calculate A_area = area * A-pct / 100
calculate B_area = area * B-pct / 100
calculate C_area = area * C-pct / 100
calculate D_area = area * D-pct / 100
statistics soilstatl
sum A_area
sum B_area
sum C_area
sum D_area
end
no
no
quit
&return
```

Figure 3-17. Portion of HSG AML

3.3.5 Watershed Land Use /Land Cover

The Land Use and Land Cover (LULC) data files describe the vegetation, water, natural surface, and cultural features on the land surface. The USGS provides these data sets and associated maps as a part of its National Mapping Program. The LULC mapping program is designed so that standard topographic maps of a scale of 1:250,000 can be used for compilation and organization of the land use and land cover data. The data are available via the Internet at (http://edcwww.cr.usgs.gov/glis/hyper/guide/1_250_lulc).

Manual interpretation of aerial photographs acquired from NASA high-altitude missions and other sources was first used to compile the land use land cover maps. Secondary sources from earlier land use maps and field surveys were also incorporated into the LULC maps as needed. At a later time, the LULC maps were digitized to create a national digital LULC database. The polygons have a minimum size of 10 acres. Each polygon represents a homogeneous element in the mapping scheme that is labeled with an Anderson Land Use Code. Land use coverages are available on the Internet at

(http://edcwww.cr.usgs.gov/glis/hyper/guide/1_250_lulc).

Anderson Land Use Codes have one-digit (Level One) and two digit (Level Two) codes. The one-digit codes are shown in Table 3-4. An example of the two-digit codes is shown in Table 3-5.

Table 3-4. Anderson Land Use Codes

Level One Code	Land Use
1	Urban or Built-up Land
2	Agricultural Land
3	Rangeland
4	Forest Land
5	Water
6	Wetland
7	Barren Land
8	Tundra
9	Perennial Snow or Ice

Table 3-5. Level Two Codes for Agricultural Land

Level Two Code	Agricultural Use
21	Cropland and Pasture
22	Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas
23	Confined Feeding Operations
24	Other Agricultural Land

Level One codes were used for this research. The polygon coverage was converted to a grid with the value of the one-digit code. Again, the ZONALSTATS function in GRID was used to obtain the percent of the area of each watershed in each land use. No tundra (code=8) or perennial ice (code = 9) exists in the study area.

3.3.6 Watershed Rainfall

Annually averaged rainfall for the United States was generated using the PRISM model (Daly, Neilson, and Phillips, 1994). Data input to the model consisted of 1961-1990 monthly average precipitation totals from over 8000 CLIM81, SNOTEL, and selected state network stations. In addition, some data-sparse areas were supplemented by a total of about 500 shorter-term stations. A station was included in the data set if it had at least 20 years of valid data. The annual rainfall grid was resampled and projected to be consistent with the other grids for this project. The ZONALSTATS function was used to compute the average, maximum, and minimum average annual rainfall for each watershed. The same function (but with the outlet grid serving as the zone grid) was used to determine the average annual rainfall at each station.

3.3.7 Dams and Reservoirs

The TWDB provided a coverage with polygons representing the extents of the reservoirs in the state. For this work, the locations of the dams, corresponding drainage areas, and corresponding hydrologic lengths to the sediment gages are desired. To determine an outlet for each reservoir, flow accumulation grids were created from the DEM for each reservoir. With each reservoir representing a different zone, the ZONALMAX function was used to determine the outlets for each reservoir. Similarly to the sediment gauges, the outlets do not necessarily correspond correctly to the streams coverage. To adjust the locations of the outlets, the same procedure that was used for the gauging stations was employed. The location of the outlet was used to determine the hydrologic distance from the reservoir to the sediment gauging station. Rather that using the outlet to find the contributing areas of the reservoirs, the contributing area listed in the attribute table was assumed to be correct.

The variable *resvar* is computed by dividing the fractional area by the weighted fractional length. The fractional area (*fracarea*) is the total area that accumulates in dams divided by the total area of the watershed. The weighted fractional length is the weighted length divided by the total length. The distance from the dam to the station is multiplied by the contributing area to the dam. These values are accumulated for all of the furthest downstream dams in the basin. The cumulative total of these values is divided by the cumulative total area contributing to dams to determine the weighted length.

3.4 Data Analyses (SAS)

SAS was used for all statistical analyses including regression. The SAS System is an integrated system of software providing complete control over data management and analysis. SAS is a registered trademark, not an acronym. Because SAS's capabilities are so broad and powerful, it is sometimes called a programming language.

All of the SAS work was done by remote login to the UNIX Timesharing Service (UTS): http://www.utexas.edu/cc/services/unix/. A cluster of computer systems known as UNIX Timesharing Services (UTS) provides general-access interactive UNIX timesharing. The UTS cluster consists of two Digital Equipment Corporation AlphaServers running Digital UNIX. Machines in the cluster have identical resources and software and share 40 GB of disk space. Because of its speed and 64-bit architecture, UTS is particularly useful for data-intensive applications.

SAS programs communicate with the computer by SAS "statements." There are two kinds of SAS statements: DATA statements and PROC statements. DATA statements are used to indicate where the variables are on data lines, the names of the variables, and how to create new variables from existing variables. The PROC (short for procedure) statements indicate what kind of analyses to perform and provide specifications for those analyses.

The SAS data sets were created from EXCEL spreadsheets which were created as explained in Sections 3.2 and 3.3. The procedures that were used will be discussed along with the analyses in the following chapters.

4 Effects of Gross Watershed Characteristics on Sediment Yield

4.1 Purpose

The primary purpose of analyzing the annual sediment data from across Texas is to determine the effects of gross watershed characteristics on sediment yield. Figure 4-1 displays the extent of the study area in relationship to the Texas boundary.

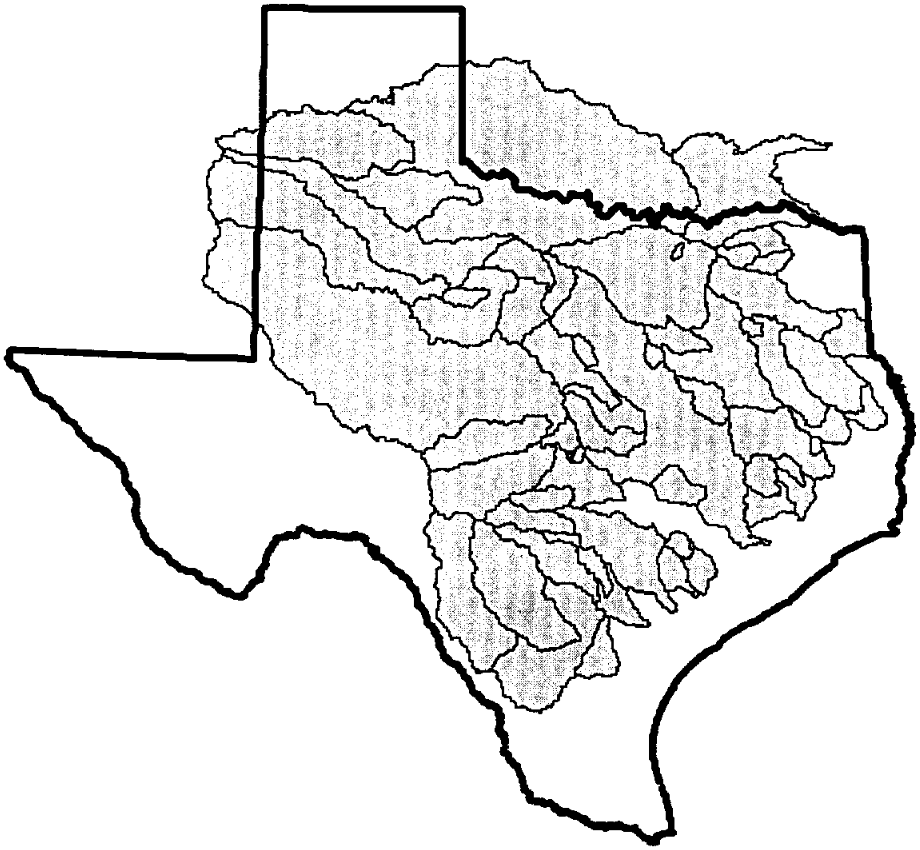


Figure 4-1. Extent of Study Area

As displayed in Table 4-1, the sampled basins contain a wide range of soil types, topography, and rainfall.

Table 4-1. Study Area Property Variability

Study Area Property	Maximum Value	Minimum Value
soil erodibility	0.50	0.00
elevation	1508 m	5 m
ground slope	33.7%	0%
avg. annual rainfall	1462 mm	338 mm

In addition, the individual basins above each gauging station vary by size, shape, and average watershed properties. Some of these variations are presented in Table 4-2.

Table 4-2. Basin Property Variability

Basin Property	Maximum Value	Minimum Value
area	12,300 km^2	81.1 km^2
length	1670 km	31.3 km
river slope	2.96%	0.33%

4.2 Data Description and Bivariate Analysis

Three SAS data sets were created: one with annual values for all the stations for all of the years the stations were in operation, one with averaged annual values for all the stations, and one with averaged annual values for all the stations with no upstream dams. In each case, an ASCII file containing the data and the watershed characteristics was read by the SAS program, new variables were formed, and a SAS data set resulted. Table 4-3 shows the names of the SAS data sets. Appendix C includes the SAS programs used to create the SAS data sets. Appendix D also lists values for the variables of the different watersheds.

Table 4-3. SAS Data Sets

	Annual values	**Annual averages**
All stations	Alldata.annual	Avedata.averages
Stations with no upstream reservoirs		Avedata.avenodam

There are several reasons that data sets with both individual annual values for particular years and annual average values were used. The data set using individual annual values is significantly larger than the set using average values. Using the large data set provides more degrees of freedom allowing the multivariate regression model to contain more variables. Furthermore, the year to year variability of sediment load, stream flow, and rainfall can be considered in the analysis of the annual value data set. It is unnecessary to create a fourth data set with annual values for those stations without upstream dams. These data can be analyzed using the annual value data set.

The new variables formed in the SAS programs are essentially of three types. The first type of new variable is simply the same variable but in different units. For example, *annload* is the sediment load given in tons and *kgload* is the sediment load given in kilograms. Because the different data sources vary on the chosen system of measurement, the SAS data set was created so that all variables would have values in both English and metric units. The second type of new variable is formed by combining two or more variables. The variable *sedconc* is created by dividing *annload* by *annflow*, where *annflow* is the annual flow. The last type of new variable consists of using a different functional form of the variable. As indicated by the name, *lnsedconc* is the natural logarithm of *sedconc*.

4.2.1 Annual Values

The annual values for sixty stations comprise a total of 1360 data points. Figure 4-2 shows the amount of annual data available for analysis. There are sixty stations and up to sixty-one years of data for each station.

The alldata.annual data set contains year-to-year variability in sediment load, stream flow, and rainfall near the station. Also, the presence of reservoirs in the basin changes from year to year. All of the watershed characteristics for individual basins were assumed to have the same value for each year. It should be noted that streamflow and annual rainfall near the station are unavailable for several stations.

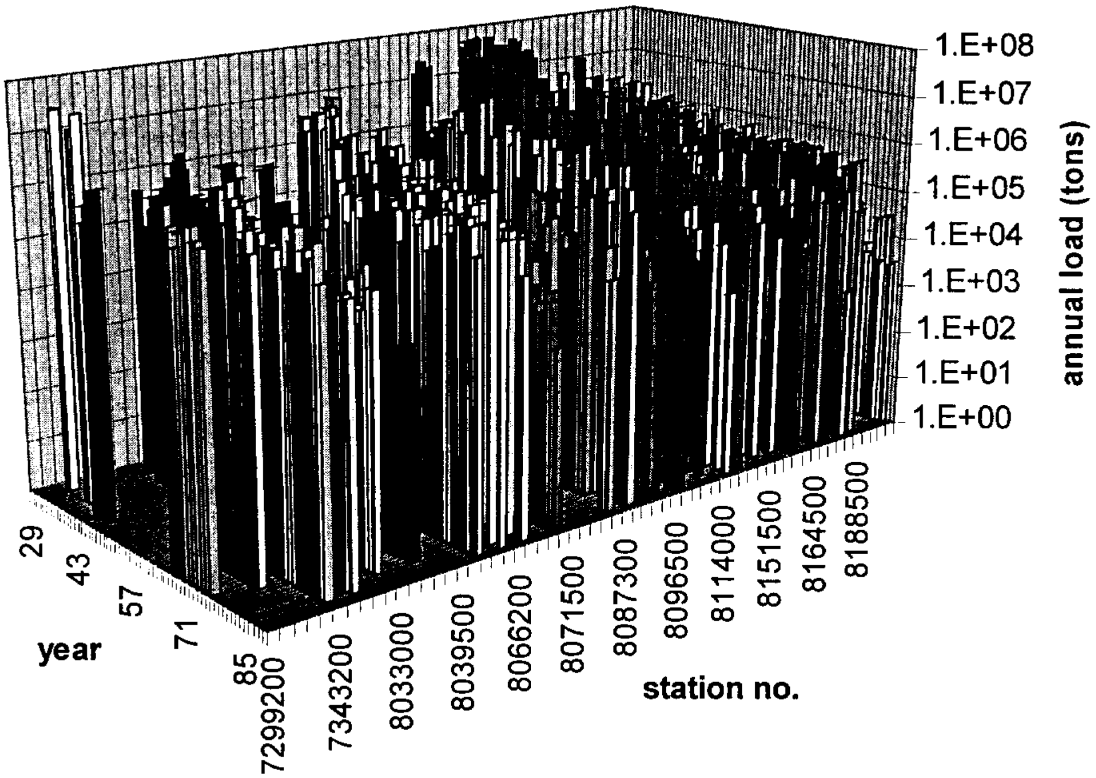


Figure 4-2. Annual Sediment Load for TWDB Stations

4.2.2 Annual Averages

The average sediment load, stream flow, and rainfall near the station were computed in SAS for each station. This information was combined with the watershed characteristics to form an ASCII file named averages.dat. The SAS program was used to form the SAS data set called avedata.averages. Figure 4-3 displays the data set the same way that Figure 4-2 displays the annual data set. Note that time is important in the average data set because each average was computed over a different number of years and over different years.

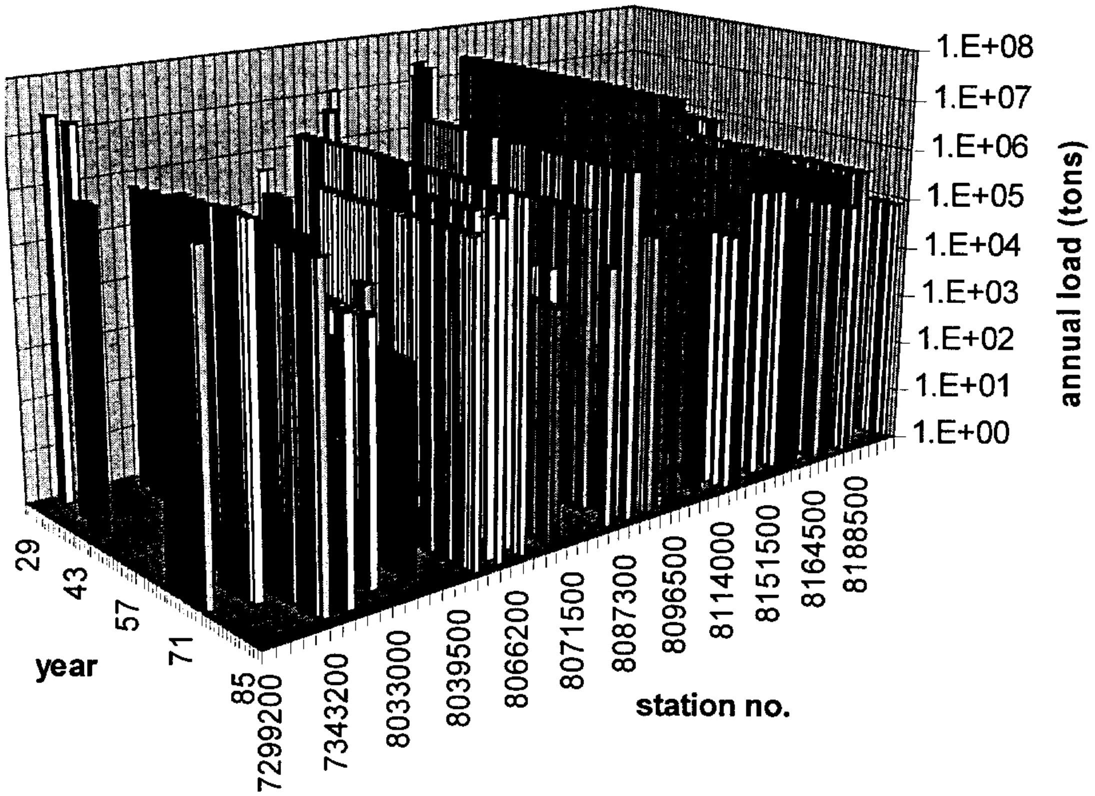


Figure 4-3. Averaged Annual Sediment Load for TWDB Stations

Table 4-4 displays the annual averages of sediment load, stream flow, and rainfall as used in the multivariate regression models. Consecutive numbers 38 and 39 are not included in the table. Consecutive number 38 corresponds with USGS gauging station number 8104500 - Little River Station on the Little River. The station's operation was discontinued in 1929. Consecutive numbers 39 and 40 correspond with USGS gauging stations 8109900 - Somerville Lake near Somerville and 8110000 - Yegua Creek near Somerville. These numbers refer to the same station. Apparently after Somerville Lake was built, the location of the station was adjusted slightly and the gauge was renumbered. The data table is not completely consistent with Table 2-2 because the tables were compiled differently for different purposes. Table 2-2 was compiled using the published sediment reports. Table 4-4 was compiled using the monthly data. There are six stations with no values for average annual rainfall (indicated by "." in the table). These stations had no close rainfall stations during the time period that sediment samples were collected. As discussed in Chapter 3, the annual rainfall values from NCDC weather stations near sediment stations were retrieved for each year that sediment samples were collected. Average annual values of rainfall from across the study area were also used in the analysis. However, the average annual values of rainfall were obtained by the PRISM study (Daly, Neilson, and Phillips, 1994) and represent a 30-year average from 1961-1990.

Table 4-4. Annual Average Values for TWDB Sediment Stations

Consecutive No.	Station No.	No. of Yrs	Average Annual Sediment Load	Average Annual Runoff	Average Sediment Concentration	Average Annual Rainfall
			(kg)	(m^3)	(mg/l)	(mm)
1	7299200	10	1.64E+09	6.98E+02	26,586	453
2	7308000	4	1.11E+09	1.24E+03	9,079	596
3	7331600	4	1.94E+10	5.04E+04	3,959	.
4	7336820	9	3.65E+09	1.22E+05	289	1,109
5	7342500	3	1.29E+08	2.22E+03	754	924
6	7343200	21	1.17E+09	1.51E+04	777	1,211
7	7343500	23	4.67E+07	4.95E+03	99	1,181
8	8022000	19	1.24E+08	2.61E+04	53	.
9	8022500	32	6.60E+08	3.58E+04	229	1,262
10	8031200	16	2.81E+06	1.48E+03	20	1,094
11	8033000	17	6.12E+07	1.58E+04	43	1,128
12	8033300	8	2.46E+06	2.56E+02	103	1,089
13	8033500	35	2.97E+08	2.37E+04	113	1,247
14	8037050	9	8.72E+06	3.75E+02	253	1,122
15	8038500	8	1.24E+08	2.03E+04	74	1,130
16	8039500	6	4.64E+08	3.88E+04	128	1,444
17	8052700	3	3.51E+07	3.36E+02	1,018	.
18	8062500	36	8.53E+08	2.56E+04	375	896
19	8064500	15	3.22E+08	4.57E+03	799	941
20	8065350	17	1.44E+09	5.75E+04	266	1,096
21	8066200	15	6.88E+07	9.33E+02	658	1,189
22	8066500	51	2.83E+09	7.42E+04	387	1,165
23	8068000	7	1.20E+08	3.49E+03	284	1,081
24	8069500	15	3.77E+08	1.20E+04	300	1,477

Table 4-4 (continued)

Consecutive No.	Station No.	No. of Yrs	Average Annual Sediment Load	Average Annual Runoff	Average Sediment Concentration	Average Annual Rainfall
25	8070000	36	2.28E+07	2.16E+03	106	1,315
26	8071500	6	7.26E+08	2.10E+04	268	1,972
27	8080500	3	4.09E+09	1.66E+03	23,156	590
28	8082500	3	9.38E+09	6.44E+03	15,805	.
29	8084800	14	4.05E+07	3.12E+02	1,072	641
30	8085500	2	7.27E+07	2.20E+03	390	.
31	8087300	15	4.59E+08	3.14E+03	1,008	629
32	8088000	43	3.18E+09	7.81E+03	4,199	736
33	8088600	36	8.79E+07	8.84E+03	118	663
34	8093500	23	1.96E+08	1.25E+03	1,685	876
35	8094800	23	3.01E+07	4.52E+02	576	798
36	8100500	33	1.94E+08	2.50E+03	770	814
37	8102500	4	4.84E+08	4.86E+03	963	626
40	8110000	16	9.75E+06	2.98E+03	58	1,004
41	8110500	41	1.21E+08	4.34E+03	280	907
42	8114000	54	2.07E+10	7.49E+04	2,544	1,127
43	8146000	20	7.95E+07	1.97E+03	383	683
44	8147000	61	2.25E+09	1.12E+04	1,807	678
45	8148000	35	1.50E+07	2.80E+05	1	759
46	8148090	12	5.00E+07	8.19E+03	55	776
47	8151500	43	3.69E+08	3.75E+03	869	688
48	8153500	23	6.21E+08	1.49E+03	1,682	874
49	8158000	46	3.31E+08	2.06E+04	109	839
50	8164000	43	1.33E+08	3.33E+03	499	1,103
51	8164300	22	3.71E+07	1.69E+03	262	1,026
52	8164500	2	1.12E+08	5.18E+03	230	.

Table 4-4 (continued)

Consecutive No.	Station No.	No. of Yrs	Average Annual Sediment Load	Average Annual Runoff	Average Sediment Concentration	Average Annual Rainfall
53	8167500	46	1.58E+08	3.66E+03	445	850
54	8176500	43	4.68E+08	1.85E+04	258	928
55	8183500	5	2.04E+08	1.85E+03	1,068	811
56	8186000	25	9.32E+07	1.45E+03	563	758
57	8188500	47	4.39E+08	7.46E+03	651	930
58	8194000	36	5.94E+07	2.54E+03	220	576
59	8207000	25	1.14E+08	2.49E+03	607	638
60	8210000	25	6.61E+08	9.13E+03	921	712
61	8211000	41	6.95E+07	7.55E+03	78	859

Figures 4-4 through 4-8 show relationships between annually averaged values. Regression equations summarized in Table 4-5 describe these bivariate relationships. The relationships between these variables are not very strong; however, some general trends can be noted. As flow increases, velocity and shear increase both overland and in-stream. In turn, the soil lost increases and the capacity of the stream to transport the sediment also increases. Figure 4-4 exhibits the expected increase in load for stations with higher annual runoff. Figure 4-5 however, tells a different story. The average concentration is simply the load divided by the flow. It is difficult by visual inspection of the figure to detect a relationship between concentration and flow for the various stations. The regression equation however, indicates a slight negative trend.

Figures 4-6 and 4-7 compare the same two dependent variables, average load and average concentration, with rainfall. These figures demonstrate no relationship between the load and the rainfall; however, they do illustrate a negative trend between concentration and rainfall. This negative trend is not surprising when considering the watershed properties that the annual rainfall represents. Annual rainfall indicates the type of climate, which in turn, indicates the amount and type of vegetation. The negative trend reiterates the fact that rivers tend to be muddier in arid climates and clearer in humid climates. The negative trend is not apparent when considering annual load (runoff times concentration) because the annual runoff is lower in arid climates.

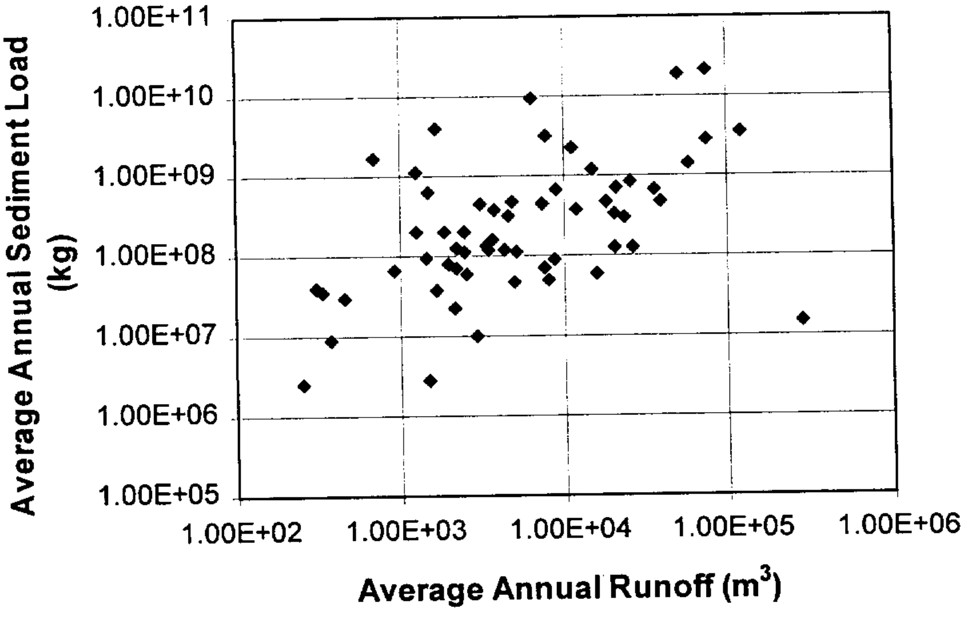


Figure 4-4. Sediment Load and Runoff for all TWDB Stations

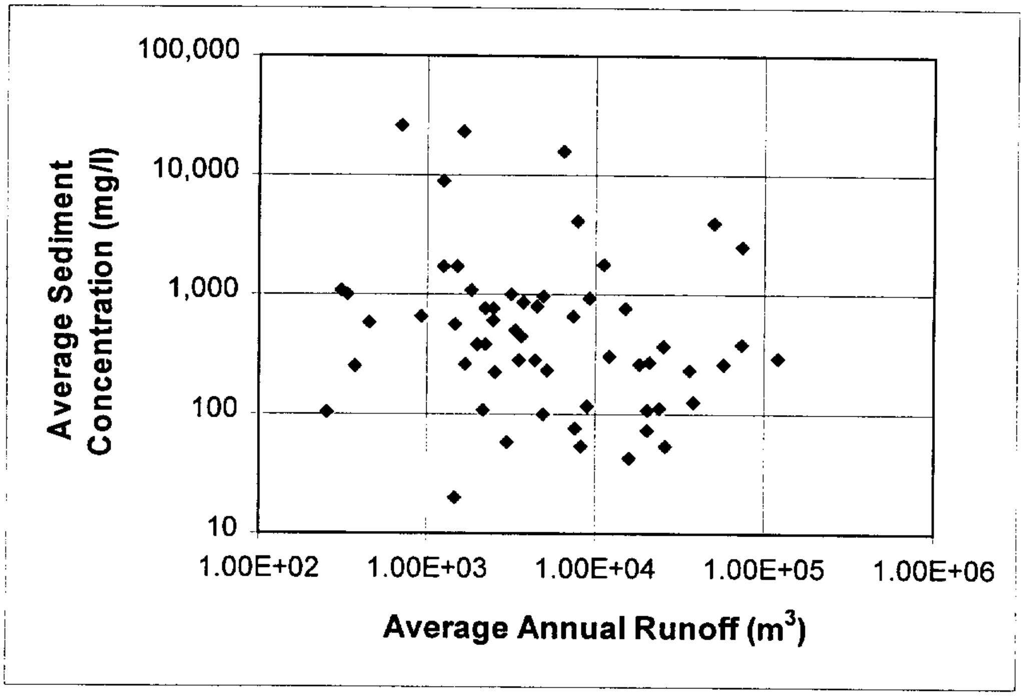


Figure 4-5. Sediment Concentration and Runoff for all TWDB Stations

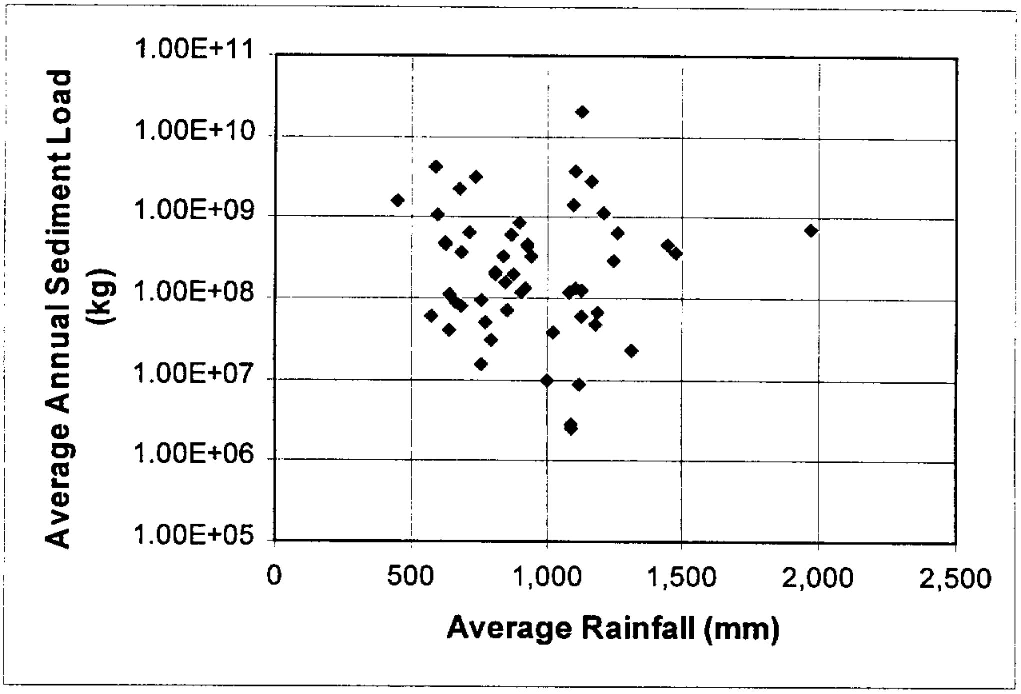


Figure 4-6. Sediment Load and Rainfall for all TWDB Stations

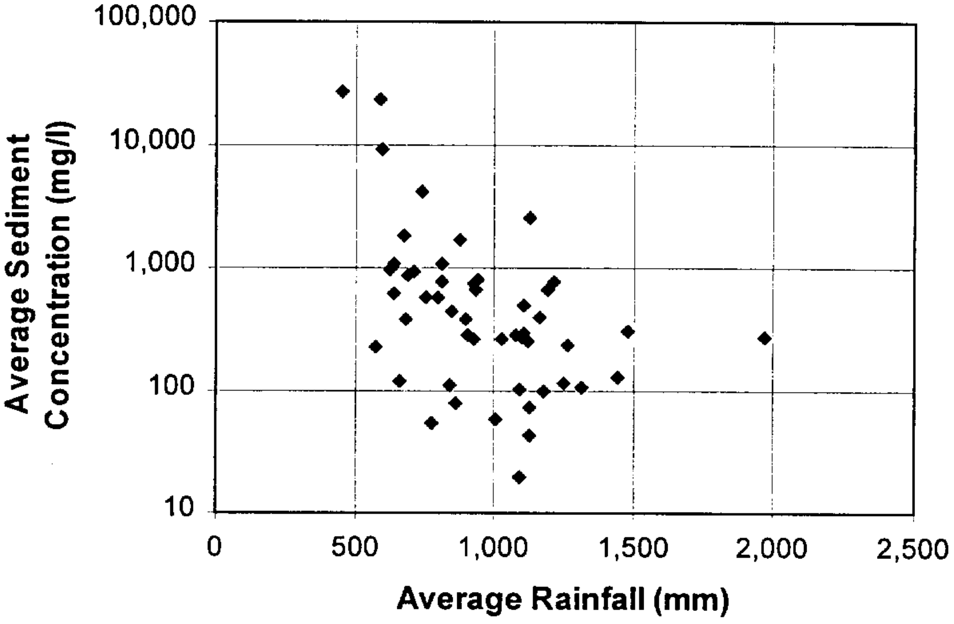


Figure 4-7. Sediment Concentration and Rainfall for all TWDB Stations

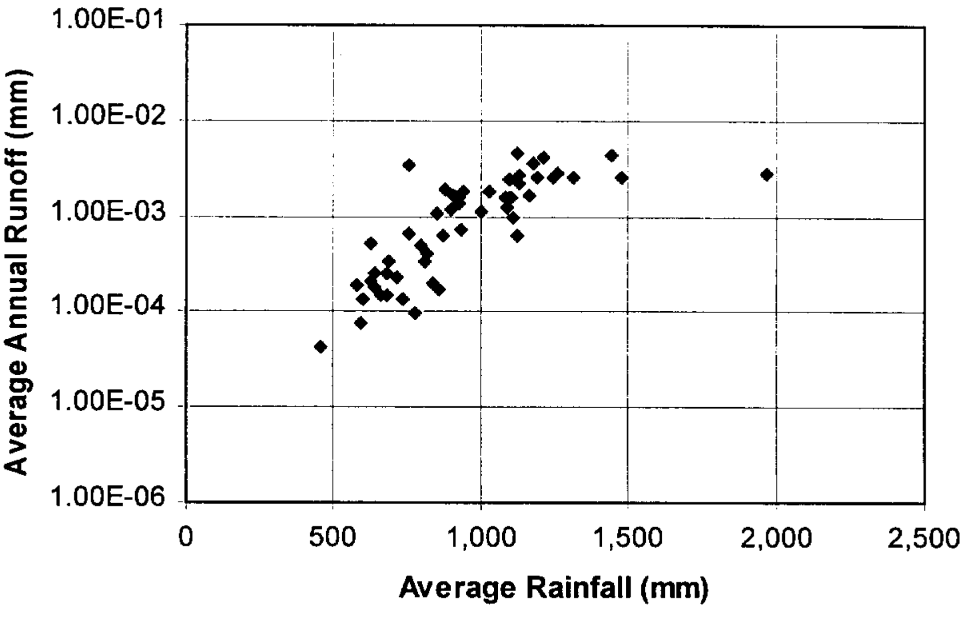


Figure 4-8. Runoff and Rainfall for all TWDB Stations

Figure 4-8 displays the positive trend between rainfall and runoff. In general, stations receiving greater amounts of rain have higher annual runoff. As in Figures 4-4 through 4-7, a large amount of scatter is present. Numerous watershed parameters and physical processes determine the dependent variables in each case. This research attempts to determine the watershed parameters and variables that reduce and explain the scatter.

Tables 4-5 and 4-6 quantify trends (or lack of trends) displayed in the figures. In almost every case, the t-statistics of the independent variables indicate a high probability that the variable is significant. Similar tables are presented throughout this chapter. The tables require some explanation. The variable names are the same as those used in the SAS programs. The table caption indicates which of the three data sets was used for analysis. The degrees of freedom (DF), plus the number of independent variables used in the regression equation, indicate the number of data points in the analysis. The degrees of freedom can change for the same data set if there are some missing variables in the data set. For example, equations (1) and (2) of Table 4-5 have 58 degrees of freedom while the remaining equations possess only 52. Six of the fifty-nine stations have no values for annual rainfall near the stations. The regression analysis, shown by equations (1) and (2) is repeated in equations (3) and (4), using only those stations with values for annual rainfall.

Table 4-5. Bivariate regression equations and statistics using average values

Equation *(t-statistics)*	R^2	DF	SE	F
(1) $\mathit{lnloadkg} = 13.89 + 0.62(\mathit{lnflowm3})$ *(11.22) (4.37)*	0.25	58	1.69	19.1
(2) $\mathit{lnconc} = 9.08 - 0.35\ (\mathit{lnflowm3})$ *(7.63) (-2.60)*	0.11	58	1.62	6.7
(3) $\mathit{lnloadkg} = 14.23 + 0.57(\mathit{lnflowm3})$ *(11.21) (3.93)*	0.23	52	1.63	15.4
(4) $\mathit{lnconc} = 9.40 - 0.40\ (\mathit{lnflowm3})$ *(7.76) (-2.88)*	0.14	52	1.55	8.3
(5) $\mathit{lnloadkg} = 19.38 - 0.00025(\mathit{arain\_mm})$ *(21.27) (-0.27)*	0.00	52	1.86	0.1
(6) $\mathit{lnconc} = 8.01 - 0.0022(\mathit{arain\_mm})$ *(10.45) (2.79)*	0.13	52	1.56	7.7
(7) $\mathit{lnflowm3} = 6.77 + 0.0019(\mathit{arain\_mm})$ *(9.39) (2.62)*	0.12	52	1.47	6.8

R^2 = coefficient of determination
DF = degrees of freedom
SE = Standard error
F = F-Value

Variables:
lnloadkg = log of load in kg
lnconc = log of concentration (mg/l)
lnflowm3 = total runoff (m^3)
arain_mm = annual rainfall near the station (mm)

Table 4-6 shows that the relationships between variables are similar for the annual data set and average data set in Table 4-5. The coefficient of determination (R^2) is slightly higher when using the annual data set for the dependent variable *lnloadkg*, but is lower when using the dependent variable *lnconc*. Note that equation (6) in Table 3-6 shows no correlation between the sediment concentration and the annual rainfall. In addition, the t-statistic expresses a low probability that that the annual rainfall has any statistical significance in the variation of concentration. This lack of correlation confirms that annual rainfall describes climate and does not explain annual variability. Individual storms affect sediment concentration and load. It is difficult to describe the intensity or frequency of individual storms in an annual variable.

Table 4-6. Bivariate regression equations and statistics using annual values

Equation *(t-statistics)*	R^2	DF	SE	F
(1) $\mathit{lnloadkg} = 12.00 + 0.78(\mathit{lnflowm3})$ *(47.63) (27.23)*	0.36	1342	1.87	742
(2) $\mathit{lnconc} = 7.44 - 0.22\ (\mathit{lnflowm3})$ *(7.63) (-2.60)*	0.04	1342	1.87	58.7
(3) $\mathit{lnloadkg} = 12.70 + 0.68(\mathit{lnflowm3})$ *(42.36) (19.77)*	0.29	960	1.86	391
(4) $\mathit{lnconc} = 8.15 - 0.32\ (\mathit{lnflowm3})$ *(27.17) (-9.36)*	0.08	960	1.86	87.6
(5) $\mathit{lnloadkg} = 17.01 + 0.0016(\mathit{arain\_mm})$ *(79.22) (7.41)*	0.05	960	2.15	54.8
(6) $\mathit{lnconc} = 5.58 - 0.00020(\mathit{arain\_mm})$ *(28.73) (-0.99)*	0.00	960	1.94	1
(7) $\mathit{lnflowm3} = 6.87 + 0.0018(\mathit{arain\_mm})$ *(41.58) (10.78)*	0.11	960	1.65	116

4.2.3 Annual Averages with No Dams

It is well known that dams trap sediment. It is also well known that sediment-free water, released from a dam, scours surrounding channel banks and the stream bed to obtain an equilibrium sediment load. To determine the amount of impact dams have on sediment load, it is necessary to know sediment loads when there are no dams. Some gauged watersheds have never had reservoirs. Other stations had dams come on line during sampling years. The data set for no dams includes averages for thirty-one stations. Figures 4-9 and 4-10 show the relationship between load and runoff and concentration and runoff as illustrated in preceding Figures 4-4 and 4-5. Different symbols represent those stations without dams. In general, sediment load and sediment concentration are lower for those stations with dams in the watershed.

In addition to creating the annual averages with the no dams data set, a variable was added to the annual value data set. This variable takes the value of "yes" or "no" according to whether reservoirs were present in the watershed during the sampling year. This variable allows the annual value data set to be classified according to whether dams are present or not, thus precluding the need to create a fourth data set of annual values with no dams. Tables 4-7 and 4-8 quantify the relationship between sediment load and runoff and between sediment concentration and runoff for those basins with no dams. The correlation between load and runoff is markedly improved ($R^2 = 0.36$ to $R^2 = 0.55$) when considering the annual data without dams. Essentially no improvement in the correlation exists when considering the average data with no dams. For both average and annual data sets, the intercept decreased and the slope increased when considering only those stations without dams.

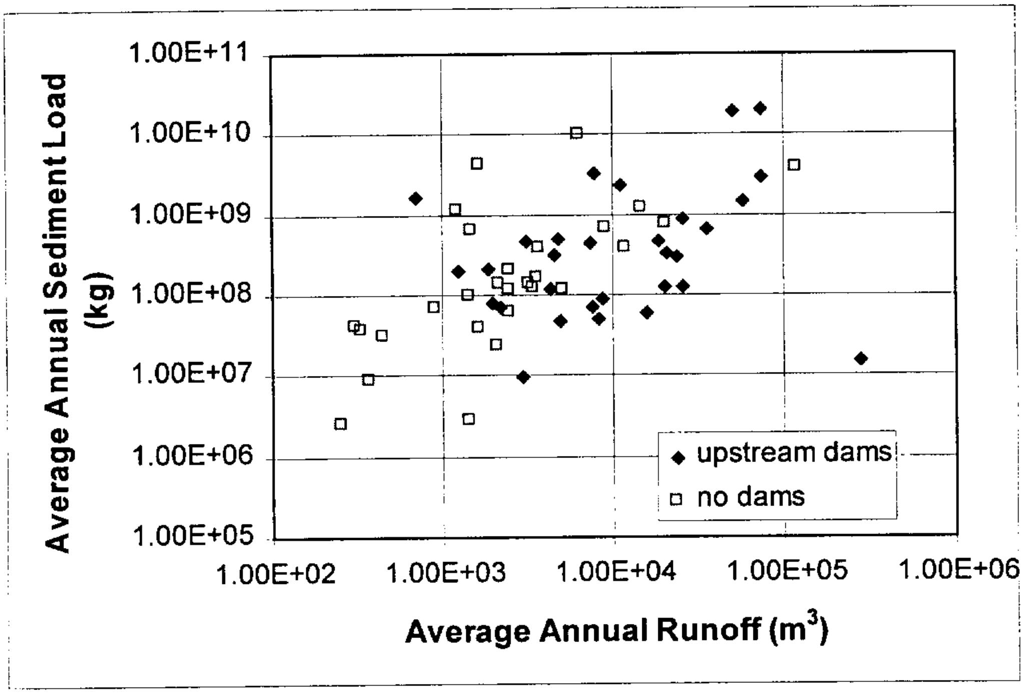


Figure 4-9. Sediment Load and Runoff With and Without Dams

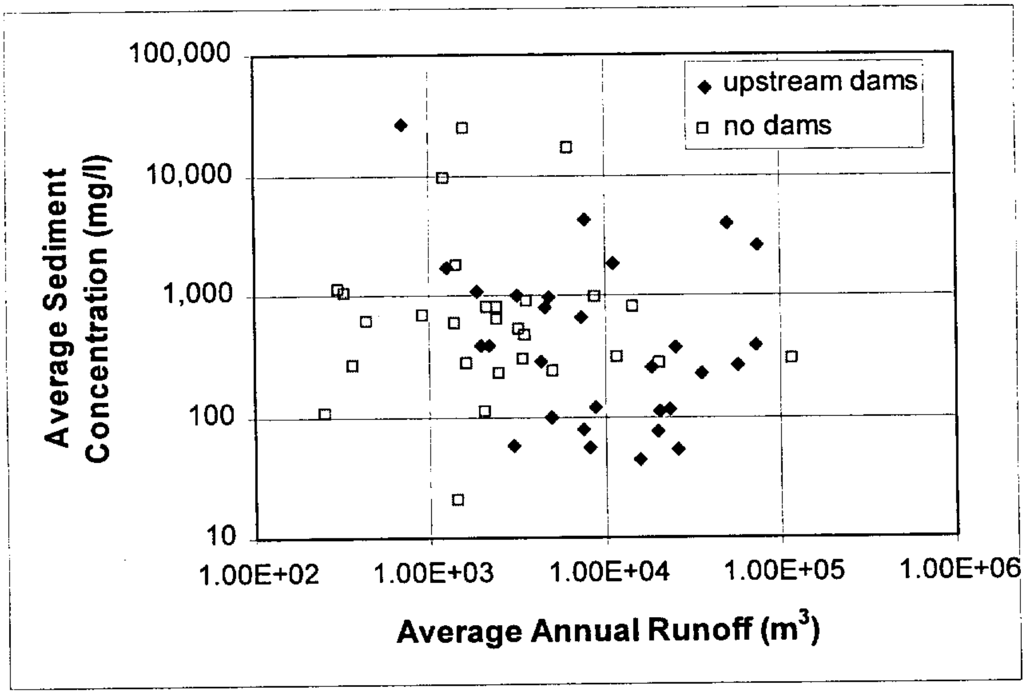


Figure 4-10. Concentration and Runoff With and Without Dams

Table 4-7. Bivariate regression equations and statistics for stations using average values with no dams

Equation	R^2	DF	SE	F
(1) $lnloadkg = 13.24 + 0.73(lnflowm3)$ $(7.22)\quad(3.26)$	0.27	30	1.73	10.6
(2) $lnconc = 8.46 - 0.25\ (lnflowm3)$ $(4.70)\quad(-1.16)$	0.04	30	1.70	1.3

Table 4-8. Bivariate regression equations and statistics for stations using annual values with no dams

Equation	R^2	DF	SE	F
(1) $lnloadkg = 11.03 + 0.98\ (lnflowm3)$ $(33.64)\quad(27.39)$	0.55	611	1.43	750
(2) $lnconc = 5.96 - 0.017\ (lnflowm3)$ $(21.02)\quad(-0.47)$	0.00	611	1.43	0.22

4.3 Data Classifications and Bivariate Analysis

The large amount of data and the large variation in the data reveal inherent biases, in both temporal and spatial contexts. Classifying the data can assist in removing biases.

4.3.1 Temporal Classifications

The fact that the sampling for each station did not cover the same time span causes temporal biases. More data points exist for those stations that were sampled for longer periods of time. In addition, some stations were sampled during wet years while other stations were sampled during dry years. Several approaches can be taken to remove temporal biases.

The first approach is to use the annual data set and run the SAS program for each individual year in order to compute different model coefficients for each year. The next approach uses the average data set and inversely weights the averages according to the number of years of data used for the average. The assumption that the more years used for the average, the more reliable the average, requires determining the number of years required for a reliable average. A value such as 5 or 10 years can be chosen and all data with fewer sampled years are not used. Figures 4-11 and 4-12 repeat the comparison between sediment volume and runoff while differentiating between stations that use more years to compute annual averages.

The highest sediment loads and concentrations occur at stations with fewer than five years of accumulated data. A few outliers have more than 10 years of data. The equations in Tables 4-9 and 4-10 are developed omitting stations with fewer than 5 years and 10 years of data respectively. Note that the load-runoff model improves when dropping the stations with fewer than 5 years of data, but the model worsens when dropping the stations with fewer than 10 years of data. Four stations with high annual load have fewer than 5 years of data. These stations strongly influence the model equation. Eliminating stations strongly influencing the model improves the correlation between variables. When removing additional stations with records of fewer than 10 years, the correlation improves. The stations with 5 to 9 years of record have less influence on the model. By eliminating station with fewer than two years of data, several stations with at least 10 years of data become outliers and have more influence on the model than when all the stations are used.

It is interesting to note that the stations with the most influence in one model are not necessarily the stations with the most influence in another model. Two stations with low annual flow volumes have 5 to 9 years of data. Those two data points influence the concentration-flow model more than in the load-flow model. The concentration-flow model improves when considering the stations with at least 10 years of data. When the model improves, in each instance it improves because stations with varying amounts of influence are removed, not because the temporal bias is removed. The same analysis was done removing the stations with upstream dams. Almost no effect was noted in the models by removing the stations with upstream dams.

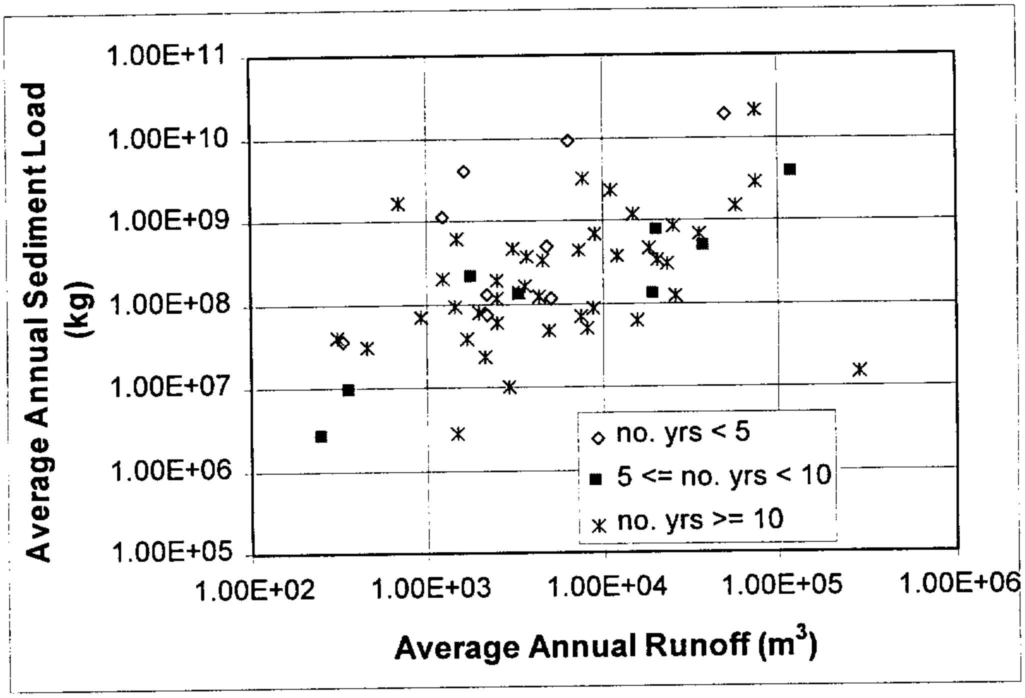


Figure 4-11. Sediment Load and Runoff by Number of Years

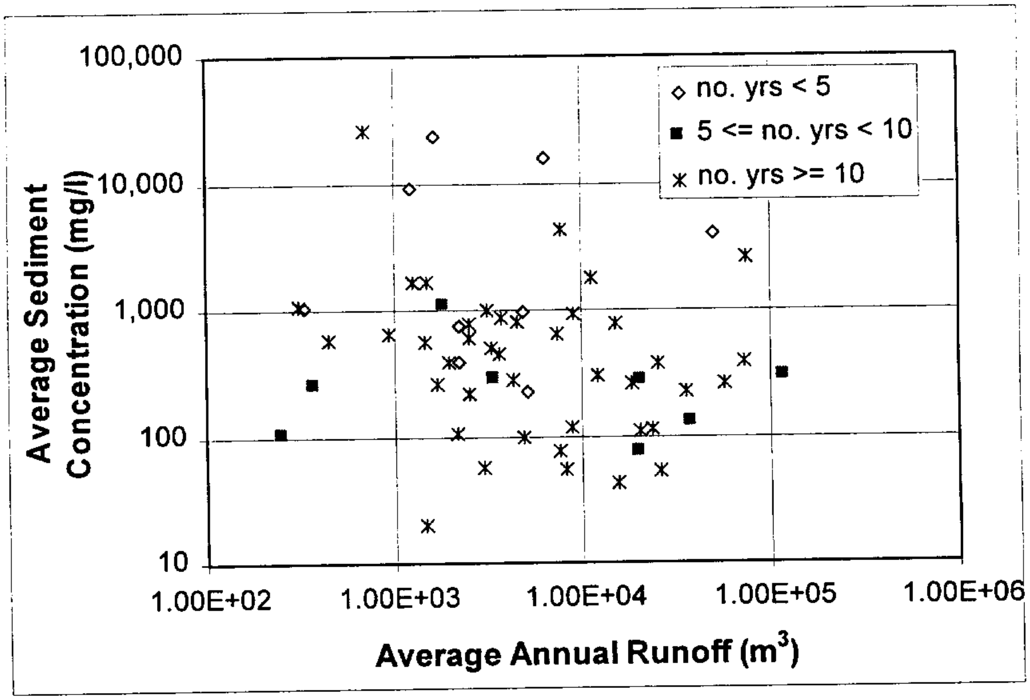


Figure 4-12. Sediment Concentration and Runoff By Number of Years

Table 4-9. Bivariate regression equations and statistics using average values based on at least 5 years of data

Equation *(t-statistics)*	R^2	DF	SE	F
(1) *lnloadkg* = 13.58 + 0.63(*lnflowm3*) *(11.02) (4.49)*	0.30	49	1.55	20.1
(2) *lnconc* = 8.72 - 0.34 (*lnflowm3*) *(7.49) (-2.60)*	0.12	49	1.46	6.77

Table 4-10. Bivariate regression equations and statistics using average values based on at least 10 years of data

Equation *(t-statistics)*	R^2	DF	SE	F
(1) *lnloadkg* = 14.86 + 0.49(*lnflowm3*) *(9.89) (2.86)*	0.17	41	1.60	8.18
(2) *lnconc* = 9.85 - 0.47 (*lnflowm3*) *(6.91) (-2.88)*	0.17	41	1.52	8.29

The latter two approaches do not address the problem that some data may have been taken during wet years while other data were collected during drier years. The data were reviewed to select a time interval for which most of the stations have a complete record. The nine year period from 1970 to 1978 has thirty-three stations with complete records of data. Tables 4-11 and 4-12 show the load-flow and concentration-flow relationships for these thirty-three stations during the 1970-1978 time period. The equations in Table 4-12 were developed by omitting stations with upstream dams. As with previous data sets, the intercept decreases but the slope increases when omitting stations with dams from the load-flow equation. The coefficient of determination is relatively high for the load-flow equation omitting data with dams.

Table 4-11. Bivariate regression equations and statistics using annual values for 33 stations with complete records from 1970-1978

Equation *(t-statistics)*	R^2	DF	SE	F
(1) $lnloadkg = 12.84 + 0.65(lnflowm3)$ *(26.45) (11.82)*	0.32	296	1.72	140
(2) $lnconc = 8.28 - 0.35\ (lnflowm3)$ *(17.07) (-6.30)*	0.12	296	1.72	39.6

Table 4-12. Bivariate regression equations and statistics using annual values for 33 stations with complete records from 1970-1978 and no dams

Equation *(t-statistics)*	R^2	DF	SE	F
(1) $lnloadkg = 11.17 + 0.89(lnflowm3)$ *(19.43) (11.84)*	0.55	116	1.07	140
(2) $lnconc = 6.62 - 0.11(lnflowm3)$ *(11.51) (-1.45)*	0.02	116	1.07	2.10

Table 4-13 summarizes the above methods to remove temporal biases. As noted previously, it is difficult to determine if a model changes because a bias is removed or because some stations are having an extreme influence. As more variables are used in the models to explain variability in the dependent variables of sediment load and sediment concentration, it is necessary to continue to address the biases that may exist.

Table 4-13. Temporal Biases

	Number of years of data for a station	Wet years vs. dry years
Annual - by year	✓	✓
Average - weighted by year	✓	
Average - number of years > 5 or 10	✓	
9 year time frame	✓	✓

4.3.2 Spatial Classifications

Selectivity bias also exists in the spatial domain. There are two sampling stations in the Guadalupe River Basin while there are 14 stations in the Brazos River Basin. The locations of stations were not controlled in a statistical fashion to achieve a representative sample population, nor were the stations located in a river basin used to provide data for hydrologic modeling. Convenience dictated the locations of stations. Approaches taken to remove the biases associated with station location include classifying the data by basin and by climate.

Modeling the basins separately still means more data for the wetter areas of the state. Modeling by different climates results in more data for some basins than for others. Because all basins are different, individual parameters cannot be controlled, but the data can be classified by similar parameters. To model by basin, the stations had to be separated into 12 major river basins as defined by the TWDB (1990). Figure 4-13 illustrates the relationship between load and runoff separately for three basins. The load-flow relationship for the individual basins considerably improves. The Neches, Trinity, and San Jacinto river basins show a high correlation between load and runoff even when the stations with upstream dams remain in the analysis. These three basins are all located in the eastern portion of the state.

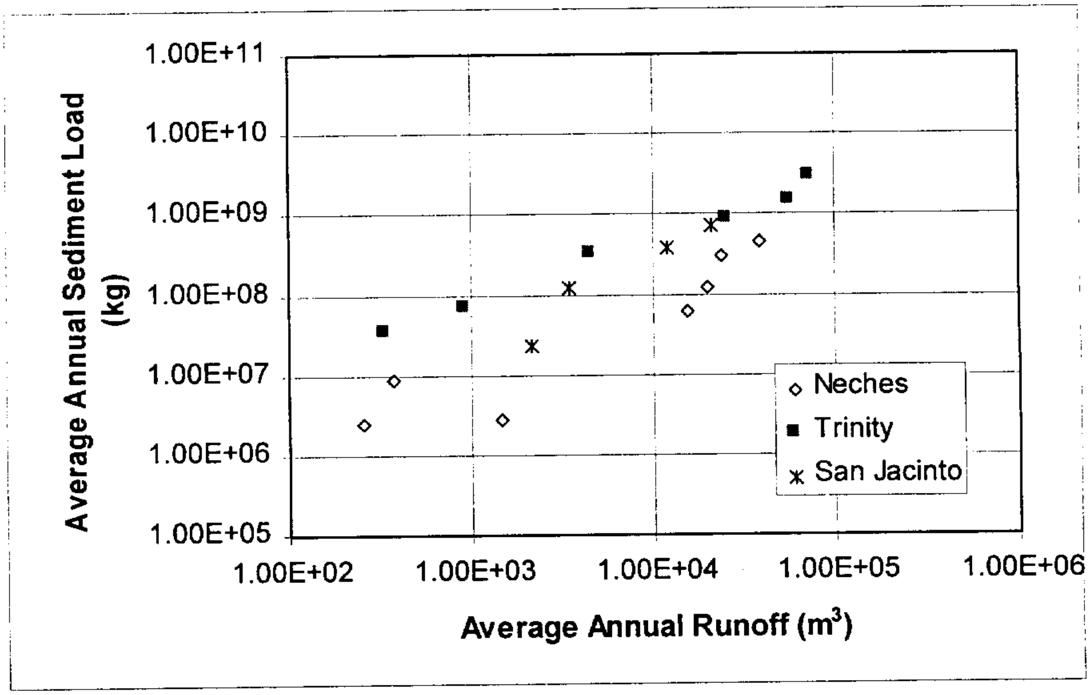


Figure 4-13. Sediment Load and Runoff for Selected River Basins

Figure 4-14 exhibits the concentration-flow relationship for three different basins. The Red River and Colorado River exhibit a negative trend while the San

Jacinto shows a positive trend between variables. Both the Red and Colorado River Basins extend into the arid climate of New Mexico. Tables 4-14 and 4-15 include the regression equations for load and flow and concentration and flow for each individual basin. The equations developed in Table 4-14 include basins with upstream dams while equations developed in Table 4-15 omit basins with upstream dams. The Sulphur, San Jacinto, and Lavaca basins have no dams upstream of the sampling stations during the sampling years.

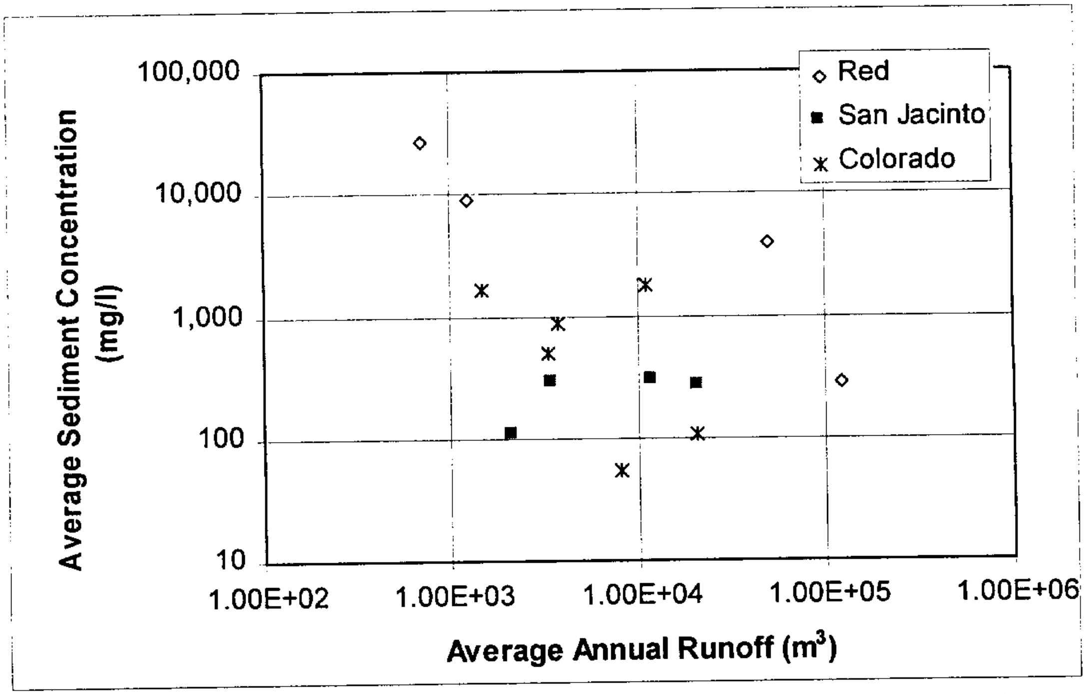


Figure 4-14. Sediment Concentration and Runoff for Selected River Basins

Table 4-14. Bivariate regression equations and statistics for annual values by river basin

River Basin	**Dependent Variable**	**Intercept** ***(t-statistic)***	**Lnflowm3 Coefficient** ***(t-statistic)***	**R^2**	**DF**	**SE**	**F**
Red	*lnloadkg*	19.17 *(29.47)*	0.28 *(3.90)*	0.38	26	0.90	15.1
	lnconc	14.62 *(22.46)*	-0.72 *(-10.25)*	0.81	26	0.90	105
Sulphur	*lnloadkg*	3.35 *(1.78)*	1.75 *(8.27)*	0.60	46	1.17	68.3
	lnconc	-1.20 *(-0.64)*	0.75 *(3.54)*	0.22	46	1.17	12.6
Sabine	*lnloadkg*	11.39 *(4.77)*	0.78 *(3.35)*	0.21	44	0.97	11.2
	lnconc	6.83 *(2.86)*	-0.22 *(-0.94)*	0.02	44	0.97	0.89
Neches	*lnloadkg*	9.18 *(31.99)*	0.96 *(27.94)*	0.87	115	0.82	781
	lnconc	4.63 *(16.12)*	-0.043 *(-1.25)*	0.01	115	0.82	1.56
Trinity	*lnloadkg*	12.34 *(36.58)*	0.81 *(23.73)*	0.81	136	0.70	563
	lnconc	7.78 *(23.08)*	-0.19 *(-5.74)*	0.20	136	0.7	33
San Jacinto	*lnloadkg*	6.70 *(14.00)*	1.35 *(23.00)*	0.90	63	0.55	529
	lnconc	2.15 *(4.48)*	0.35 *(5.98)*	0.37	63	0.55	35.7
Brazos	*lnloadkg*	9.09 *(21.11)*	1.24 *(24.33)*	0.66	306	1.73	592
	lnconc	4.54 *(10.54)*	0.24 *(4.67)*	0.07	306	1.73	22
Colorado	*lnloadkg*	19.68 *(26.08)*	-0.14 *(-1.75)*	0.01	239	2.25	3.1
	lnconc	15.13 *(20.05)*	-1.14 *(-14.12)*	0.46	239	2.25	199
Lavaca	*lnloadkg*	11.65 *(18.22)*	0.83 *(9.92)*	0.60	66	0.67	98.5
	lnconc	7.10 *(11.10)*	-0.17 *(-2.05)*	0.06	66	0.67	4.21
Guadalupe	*lnloadkg*	10.05 *(18.81)*	1.01 *(16.63)*	0.76	88	0.71	277
	lnconc	5.49 *(10.28)*	0.012 *(0.20)*	0.00	88	0.71	0.04
San Antonio	*lnloadkg*	9.11 *(16.46)*	1.22 *(17.88)*	0.81	76	0.63	320
	lnconc	4.55 *(8.23)*	0.22 *(3.18)*	0.12	76	0.63	10.1
Nueces	*lnloadkg*	10.41 *(12.02)*	0.92 *(8.65)*	0.37	126	1.35	74.8
	lnconc	5.86 *(6.76)*	-0.077 *(-0.73)*	0.00	126	1.35	0.53

Table 4-15. Bivariate regression equations and statistics for annual values with no dams, by basin

River Basin	Dependent Variable	Intercept *(t statistic)*	Lnflowm3 Coefficient *(t statistic)*	R^2	DF	SE	F
Red	*lnloadkg*	17.66 *(12.43)*	0.50 *(2.34)*	0.31	13	0.51	5.47
	lnconc	13.11 *(9.22)*	-0.50 *(-2.31)*	0.31	13	0.51	5.33
Sulphur	*lnloadkg*	3.35 *(1.78)*	1.75 *(8.27)*	0.60	46	1.17	68.3
	lnconc	-1.20 *(-0.64)*	0.75 *(3.54)*	0.22	46	1.17	12.6
Sabine	*lnloadkg*	13.88 *(2.82)*	0.61 *(1.32)*	0.14	12	1.01	1.74
	lnconc	9.32 *(1.90)*	-0.39 *(-0.83)*	0.06	12	1.01	0.68
Neches	*lnloadkg*	9.04 *(29.79)*	0.99 *(26.46)*	0.88	94	0.85	700
	lnconc	4.48 *(14.77)*	-0.015 *(-0.31)*	0.00	94	0.85	0.09
Trinity	*lnloadkg*	9.39 *(10.43)*	1.27 *(9.00)*	0.83	18	0.48	81
	lnconc	4.84 *(5.37)*	0.27 *(1.90)*	0.17	18	0.48	3.57
San Jacinto	*lnloadkg*	6.70 *(14.00)*	1.35 *(23.00)*	0.90	63	0.55	529
	lnconc	2.15 *(4.48)*	0.35 *(5.98)*	0.37	63	0.55	35.7
Brazos	*lnloadkg*	9.12 *(20.17)*	1.33 *(21.48)*	0.83	97	1.16	461
	lnconc	4.56 *(10.09)*	0.33 *(5.34)*	0.23	97	1.16	28.5
Colorado	*lnloadkg*	8.65 *(6.97)*	1.27 *(7.87)*	0.48	67	1.47	62.0
	lnconc	4.09 *(3.30)*	0.27 *(1.67)*	0.04	67	1.47	2.77
Lavaca	*lnloadkg*	11.65 *(18.22)*	0.83 *(9.92)*	0.60	66	0.67	98.5
	lnconc	7.10 *(11.10)*	-0.17 *(-2.05)*	0.06	66	0.67	4.21
Guadalupe	*lnloadkg*	8.74 *(8.54)*	1.19 *(9.23)*	0.66	45	0.86	85.3
	lnconc	4.19 *(4.09)*	0.19 *(1.51)*	0.05	45	0.86	2.27
San Antonio	*lnloadkg*	4.92 *(4.32)*	1.81 *(11.19)*	0.84	24	0.63	125
	lnconc	0.37 *(0.33)*	0.81 *(5.01)*	0.52	24	0.63	25.1
Nueces	*lnloadkg*	11.18 *(15.48)*	0.95 *(10.83)*	0.68	55	0.74	117
	lnconc	6.62 *(9.17)*	-0.047 *(-0.54)*	0.01	55	0.74	0.29

The Brazos and Nueces basins show significant improvement in the load-runoff correlation when omitting the stations with upstream dams. When omitting stations with dams from both of these basins, the intercept and the slope increase. When omitting stations with dams from every other basin, at least one of the coefficients is decreased.

Examining the regression equations by basin lends a few insights to the nature of the data. Figures 4-15 and 4-16 show plots of the bivariate regression equations for sediment load. On these plots, regression lines are shown for the extent of the plot no matter the extent of data. These plots are not shown for the purpose of predicting load but for the purpose of analyzing the data. The Lavaca River Basin is shown with a black bold line for comparisons with Chapters 5 and 6.

Both plots show that the majority of the regression lines fall in a band that varies about two orders of magnitude. This band has a steeper slope when no upstream dams are present. When including basins with dams in the analysis, the regression lines for the Red River and Colorado River Basins do not fall on the band. However, when removing stations with upstream dams, the Colorado Basin regression line falls in the band. The Colorado River Basin has a large series of dams, the Highland Lakes. There is essentially no correlation between load and runoff in this basin without removing those stations that are affected by upstream dams. When these stations are removed from the analysis, the regression equation coefficients are more in line with the other basins.

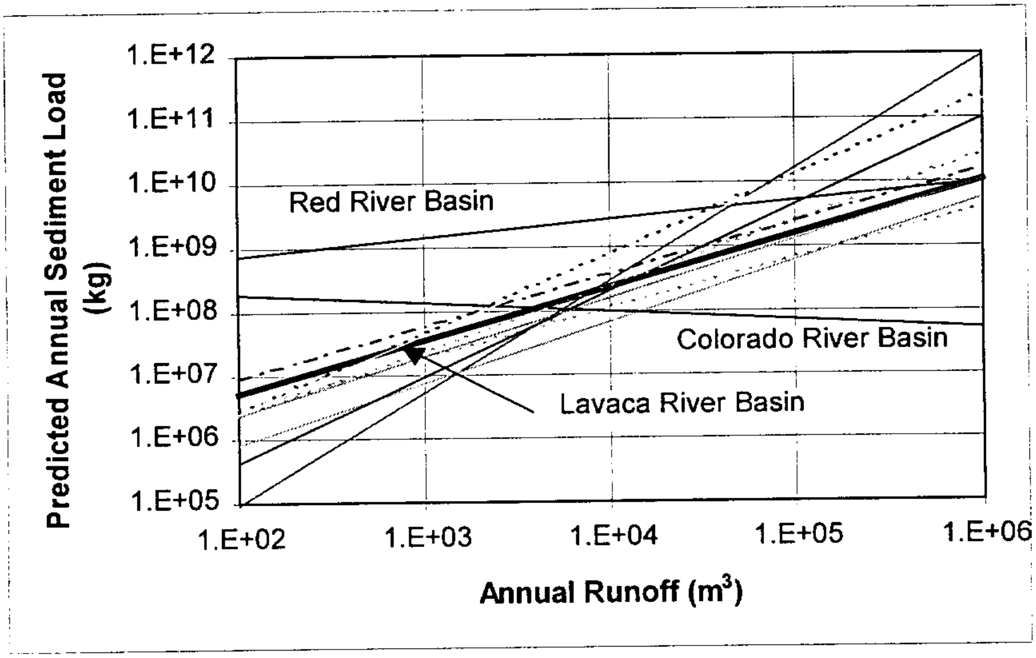


Figure 4-15. Predicted sediment load based on bivariate regressions equations for different river basins

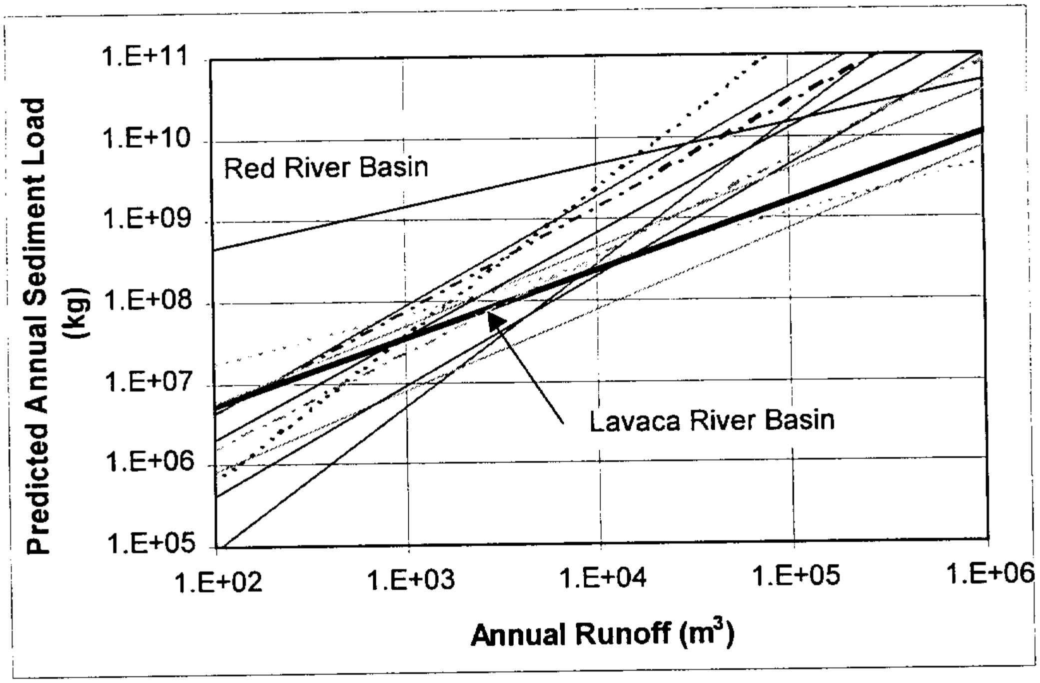


Figure 4-16. Predicted sediment load based on bivariate regressions equations for different river basins with no upstream dams

Similar plots for sediment concentration show an almost horizontal band. The Red River Basin does not fall in the band. Although for a few basins a large amount of variation in concentration can be attributed to flow, little relationship is seen in general between sediment concentration and flow. This fact would lead one to believe that the best method of prediction is to use the historically measured average concentration with hydrologically predicted (or known) flow to predict sediment load.

To model by climate, three classifications of rainfall for Texas are defined: high, medium, and low. The climate types are defined according to the average (temporal and spatial) annual rainfall over the basin. The spatially averaged annual rainfall for a gauged watershed ranges from 475 mm for Station no. 8080500 - Double Mountain Fork near Aspermont (in the Brazos River Basin) to 1212 mm for Station No. 8066200 - Long King Creek at Livingston (in the Trinity River Basin). To determine the dividing values for climate type, the minimum average rainfall was subtracted from the maximum average rainfall. The value was then divided by three. The resulting definition for low rainfall climate basins is those basins with an average annual rainfall of less than 720 mm. High rainfall basins average at least 966 mm rainfall. The medium rainfall climate basins have rainfall averages between the low and high rainfall climates.

Figure 4-17 displays the spatial variation of annual rainfall across the state according to the high, medium, and low rainfall categories.

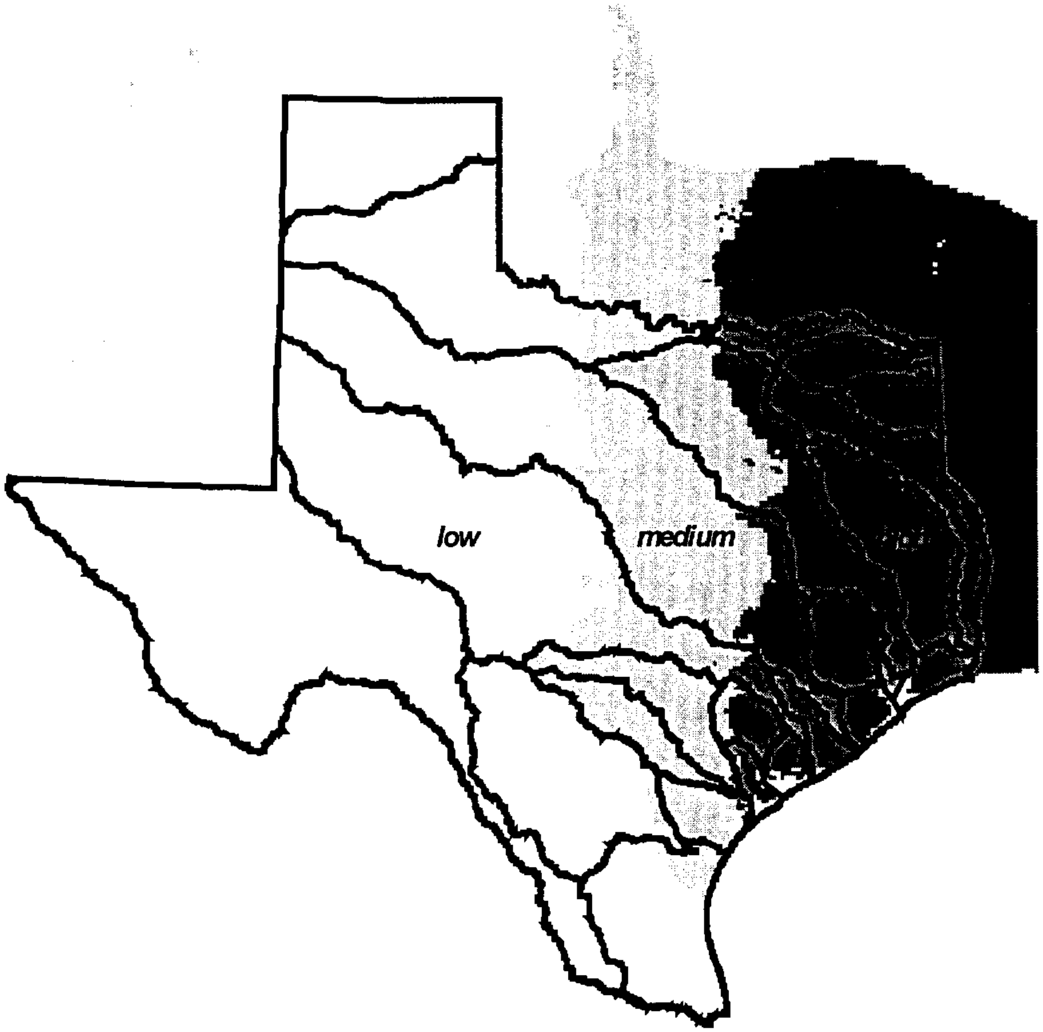


Figure 4-17. Spatially varied rainfall across Texas

Figure 4-18 emphasizes that a basin is classified according to spatially average rainfall. Large basins that extend across the state may be classified as low rainfall even though the actual station is located in a high rainfall portion of the state.

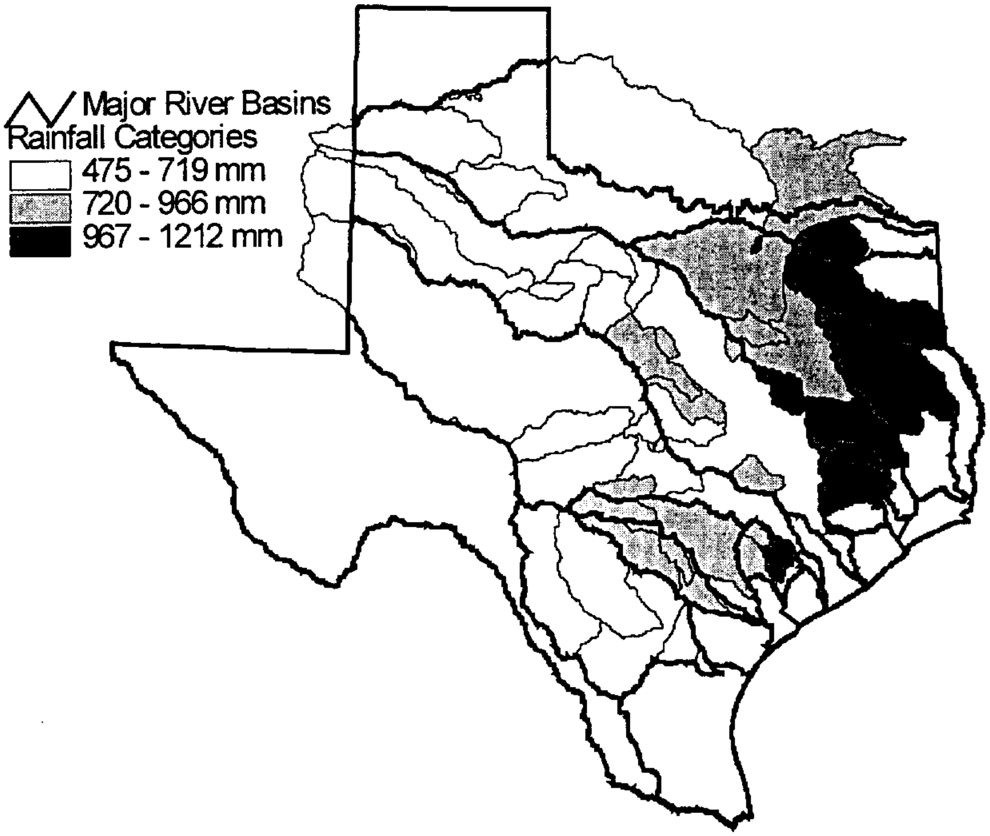


Figure 4-18. Gauged watersheds according to rainfall category

The load-flow and concentration-flow relationships, according to climate are shown in Figures 4-19 and 4-20. Corresponding regression equations are presented in Tables 4-16 and 4-17.

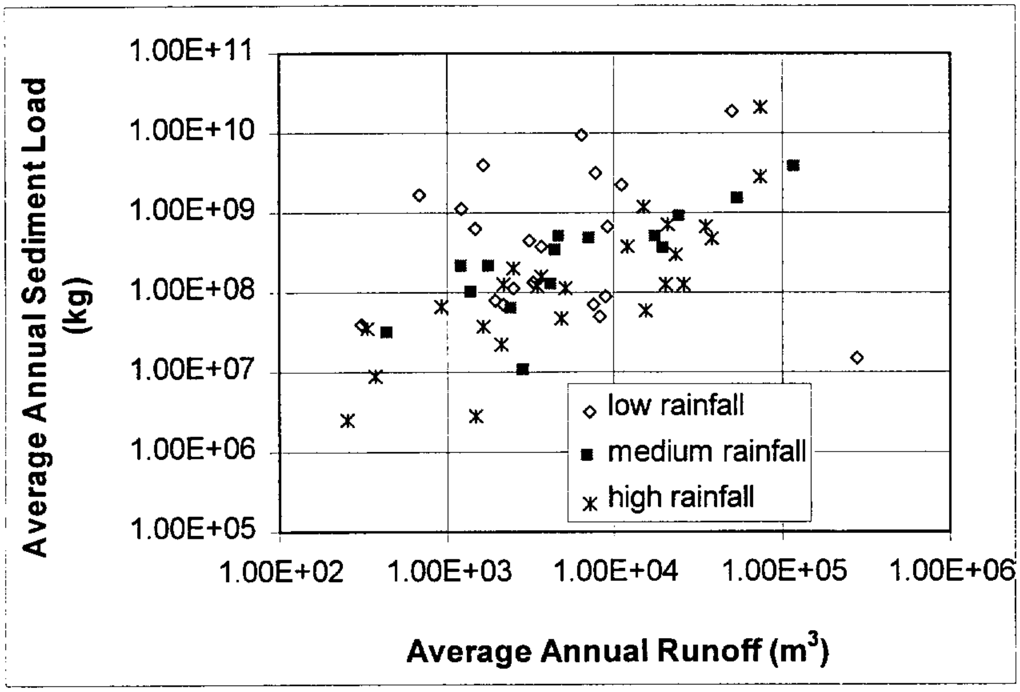


Figure 4-19. Sediment Load and Runoff for Different Climates

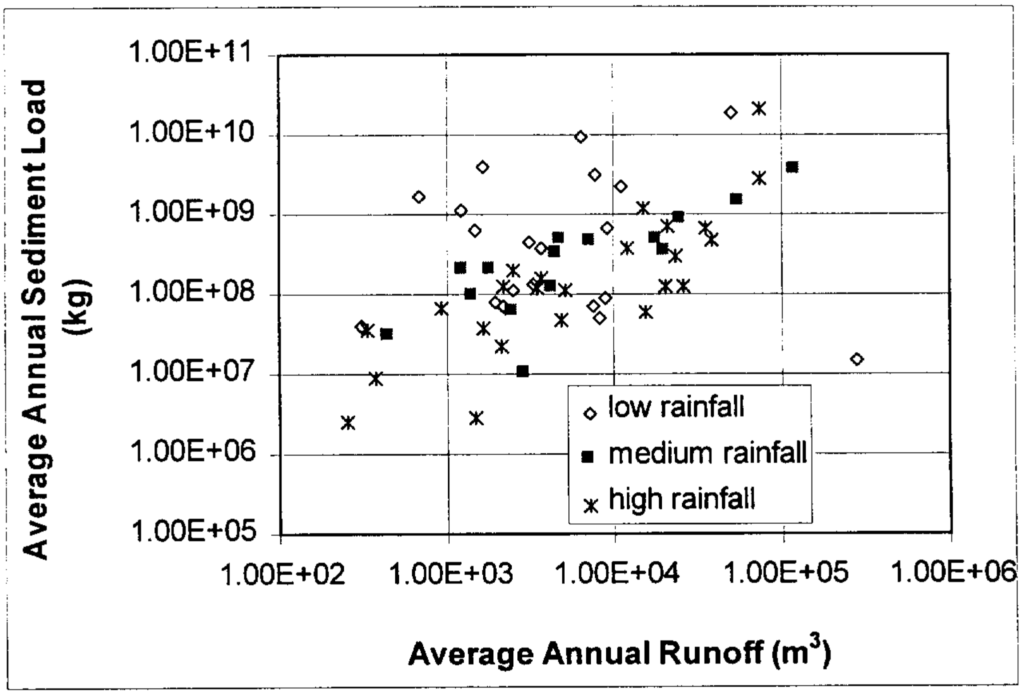


Figure 4-20. Sediment Concentration and Runoff for Different Climates

Table 4-16 Bivariate regression equations and statistics for annual values by climate

Climate by rain	Dependent variable	intercept	Lnflowm3 coefficient	R^2	DF	SE	F
low	*lnloadkg*	13.25 *(26.42)*	0.64 *(11.42)*	0.19	564	2.54	130
	lnconc	8.70 *(17.34)*	-0.36 *(-6.34)*	0.07	564	2.54	40.2
medium	*lnloadkg*	11.00 *(43.99)*	0.95 *(31.7)*	0.74	363	0.99	1005
	lnconc	6.45 *(25.78)*	-0.054 *(-1.83)*	0.01	363	0.99	3.34
high	*lnloadkg*	10.05 *(32.98)*	0.96 *(28.18)*	0.66	413	1.10	794
	lnconc	5.49 *(18.02)*	-0.043 *(-1.27)*	0.00	413	1.10	1.62

Table 4-17 Bivariate regression equations and statistics for annual values with no dams, by climate

Climate by rain	Dependent variable	intercept	Lnflowm3 coefficient	R^2	DF	SE	F
low	*lnloadkg*	9.09 *(17.16)*	1.26 *(18.07)*	0.65	180	1.66	327
	lnconc	4.53 *(8.56)*	0.26 *(3.77)*	0.07	180	1.66	14.2
medium	*lnloadkg*	9.90 *(19.97)*	1.12 *(16.10)*	0.64	145	1.09	259
	lnconc	5.35 *(10.78)*	0.12 *(1.70)*	0.02	145	1.09	2.91
high	*lnloadkg*	10.15 *(26.20)*	0.96 *(20.84)*	0.61	284	1.10	434
	lnconc	5.60 *(14.44)*	-0.044 *(0.95)*	0.00	284	1.10	0.90

Examination of the tables leads to the conclusion that dams have more impact on sediment load in low rainfall climates than in medium or high rainfall climates. The equations for data with dams show decreasing intercept and slope for both load and concentration as climate becomes more humid. The analysis without dams shows increasing intercepts but decreasing coefficients for the more humid (high rainfall) climates. The more humid a climate, the more stable the environment. The higher rainfalls result in more land cover. Humid climates produce high runoff with correspondingly high loads but the concentration of sediment in the rivers is lower than in more arid (low rainfall) climates.

4.4 Physical Parameters and Bivariate Analysis

Chapter 3 - Research Methods discusses how GIS was used to determine watershed characteristics. A number of independent variables must be included to adequately represent storm and basin characteristics. Topography, land use, soil type, and rainfall are the most important factors. Some of the watershed characteristics can be used directly as variables in the SAS models and some of the watershed characteristics are combined to create new variables. Appendix D shows the values of the variables for the individual watersheds. Following is a description of some of the important watershed characteristics for the erosion/transport/deposition process.

basin area (*area*)

The watershed area is indicative of the supply of sediment available. The area multiplied by the rainfall is the potential runoff (*potq_m3*) which can cause erosion and transport sediment. The gauged flow divided by the area (*runoffmm*) is the actual runoff.

basin perimeter (*perimetr*)

The basin perimeter is indicative of the size and shape of the watershed. The variable is included because the perimeter has been shown to be a significant variable in rainfall-runoff processes (Haan, 1977), and it is easy to determine using GIS.

basin length (*length*)

The length is also indicative of the basin area. However, the longer the basin the more opportunities there are for sediments to redeposit. It is hypothesized that when length and area are both included in a model one of the variables will represent supply while the other variable represents opportunities for deposition.

river slope (*USGSslp*)

Steeper river slopes carry faster flows and have higher values of bed shear than flatter slopes. This variable represents both the supply and transport process. The shear stress developed in steeper rivers is likely to sweep bed material into suspension. Higher stream power is capable of sustaining the particles in suspension.

travel time (*time*)

The travel time is computed as the length divided by the square root of the river slope. The length and the river slope together are indicative of the travel time in the watershed. Because the velocity of the stream flow is often not readily available, several travel time equations are based on the slope. The actual travel time is equal to the length of the basin divided by the average velocity of the flow. According to Manning's equation the velocity of the flow is proportional to the square root of the slope. Thus, the travel time should be inversely proportional to the square root of the slope. A related variable is *leng32*. This variable is the length raised to the 3/2 power. The variable is created to correspond with the velocity of flow in the basin. The average velocity of flow through a watershed is equal to the length of the watershed divided by the time of travel.

average overland slope (*slop_avg*)

On a sloping surface, more water droplets and soil particles are splashed downslope than upslope. Selby (1994) points out the relationship between splash detachment and hillslope is usually a curvilinear one and is usually influenced by particle size. This relationship provides reason to interact slope and soil variables, as well as including a polynomial term to express the curvilinear relationship. Similar variables are the maximum and minimum overland slope (*slop_max, slop_min*).

stream frequency (*strmfreq*)

The stream frequency is a ratio of stream plan area to basin area. The stream frequency indicates whether sediment that erodes from hillslopes will travel far before entering a channel where the capacity for transport is greater than on the hillslope.

land use/land cover (*lu1-lu7*)

The land use determines the sediment available for erosion. Paved areas and well seeded areas are not a major source of sediment. The vegetative cover protects soils from erosion. The one-digit Anderson land use codes include seven types of land use in the study area. Table 3-4 presents the one-digit Anderson land use codes. The land use variables are equal to the fractional area of the basin that is of a particular land use. Variables *lu1* through *lu7* are the fraction of the area that falls in land use code 1 through land use code 7. The variables are then renamed and combined as follows: *urb=lu1*, *ag=lu2*, *range=lu3*, *forest=lu4*, *veg=lu3+lu4*, *water=lu5+lu6*, and *bare=lu7*.

hydrologic soil groups (*soilA, soilB, soilC, soilD*)

There are too many soil types to include all of them. The fractional area of each basin in each hydrologic soils group is determined. The less pervious soils allow the most runoff. However, the less pervious soils are also those soils that are more resistant to erosion. An erosion resistant soils variable (*ers*) is combined by adding *soilC* to *soilD*.

average soil erodibility (*k_avg*)

Soil erodibility is the K factor included in the Universal Soil Loss Equation. The soil erodibility factor represents the average soil loss per unit of rainfall factor. The average, maximum, and minimum soil erodibilities are computed for each basin.

fraction of area that feeds upstream reservoirs (*fracarea*)

Areas that contribute runoff and sediments to reservoirs are unlikely to contribute to the overall sediment downstream of the reservoir unless during large flows the sediments are able to pass through the reservoir. Because there is not the same sediment contribution in the river as there would have been before the reservoir was built, the river will erode downstream of the dam to obtain an equilibrium sediment load.

reservoir variable (*resvar*)

The reservoir variable is a function of the distance from the gauging station to the reservoirs in the basin and the contributing drainage area for the reservoirs. This variable is included because reservoirs far upstream in a watershed do not have the same impact as reservoirs near the gauging station.

annual average rainfall for each basin (*rain_avg*)

Rainfall and runoff cause erosion. Arid climates have such low rainfall amounts and corresponding runoff amounts that the amount of erosion is less. Climates with very high rainfall are well forested and the land cover intercepts the rain and causes less erosion. By considering both rainfall and land cover in the model, it will be clear why semi-arid terrain often has the highest sediment yields. Similar variables are the maximum and minimum rainfalls (*rain_max, rain_min*).

annual average rainfall at each sediment station (*rain_sta*)

It is suspected that the watershed area nearest the station is the largest contributor of sediments because the sediment originating near the station has not had the same chance to redeposit as sediment from far upstream in the watershed. Sediment yield correlating as well or better with the station rainfall as with the average basin rainfall indicates that this hypothesis is correct.

rainfall variability range across each basin (*rainvar*)

The rainfall variability is computed by subtracting the minimum average annual rainfall from the maximum average annual rainfall. The average annual rainfall varies dramatically from west Texas to east Texas. A river basin like the Colorado extends into arid climates of New Mexico where the average annual rainfall is on the order of 475 mm. Smaller basins, such as the Lavaca, have much less variability.

Table 4-18 presents correlation coefficients of some watershed variables to sediment load and sediment concentration. Each correlation coefficient which is shown is associated with a p-value less than 0.1 indicating that it is unlikely that this correlation is due to chance. If a correlation coefficient is not shown in the table, then the associated p-value is greater than 0.1. This table is presented to give general information and does not imply causality.

Table 4-18. Pearson correlation coefficients for *kgload* using average values

independent variable	dependent variable			
	kgload	*lnloadkg*	*sedconc*	*lnconc*
(1) *area*	0.60	0.49		
(2) *lnarea*	0.43	0.66		
(3) *length*	0.65	0.59	0.23	
(4) *lnlength*	0.46	0.62	0.27	
(5) *leng32*	0.70	0.55		
(6) *shape1*	0.41	0.45		
(7) *lnshape1*	0.37	0.51		
(8) *shape2*	0.49	0.50		
(9) *lnshape2*	0.42	0.56		
(10) *potq_m3*	0.63	0.53		
(11) *lnpotq*	0.44	0.60		
(12) *runoffmm*		-0.28	-0.32	-0.49
(13) *lnromm*		-0.30	-0.52	-0.53
(14) *k_max*				-0.23
(15) *slopmax*	0.35	0.46	0.29	
(16) *range*		0.22		
(17) *USGSslop*				0.30
(18) *relief_m*	0.56	0.53	0.34	0.24
(19) *soilA*		-0.24		-0.34
(20) *soilB*	0.27	0.29	0.41	
(21) *soilC*			-0.23	
(22) *ers*	-0.24		-0.34	
(23) *elev_avg*	0.30	0.38	0.64	0.44

Table 4-18 (continued)

independent variable	dependent variable			
	kgload	*lnloadkg*	*sedconc*	*lnconc*
(24) *lnelev*	0.30	0.35	0.47	0.41
(25) *elev_max*	0.49	0.50	0.48	0.35
(26) *rain_sta*			-0.43	-0.40
(27) *rain_avg*	-0.24	-0.28	-0.44	-0.38
(28) *lnrain*	-0.24	-0.29	-0.49	-0.38
(29) *strmfreq*	0.25	0.41	0.25	
(30) *bare*				-0.30
(31) *rainvar*	0.59	0.59		
(32) *time*	0.61	0.59	0.26	
(33) *lntime*	0.43	0.60	0.25	

A large difference in the correlation between non-transformed and log transformed variables is due to the range of values which variables cover. If variables do not cover more than an order of magnitude then there is no point in considering a log transformed variable.

Variables related to the area of the gauged watershed have the highest correlation with the sediment load. In fact the correlation between sediment load and area is greater than the correlation between sediment load and flow. The first thirteen variables in Table 4-18 are directly related to area. *Length* and *area* have a 0.95 correlation coefficient. This high correlation indicates that the shapes of the basins do not vary dramatically. Two different shape factors are used. *Shape1* is equal to the area divided by the length of each basin. *Shape2* is equal to the area divided by the perimeter of each basin.

Runoff potential has a high correlation with sediment load. The actual runoff has a relatively high (but negative) correlation with sediment concentration. The negative correlation is not necessarily expected and requires some explanation. The correlation coefficient between the potential runoff and the actual runoff is -0.22. As potential runoff increases, actual runoff decreases. This coefficient has a 90% probability that the relationship is not due to chance; however, the correlation is low. This effect is explained by the nature of different climates. The watersheds with high runoff potential are the watersheds in the more humid climates. The vegetation in these areas will use a significant portion of the rainfall and runoff. As shown in Table 4-17, the more humid climates will have lower concentrations of sediment in the streams.

Other variables showing relatively high correlation with sediment concentration include those variables associated with topography, soil type, land use, and rainfall. Again, a high correlation does not imply causality. There is a high correlation between elevation and sediment concentration. High elevations do not cause high sediment concentrations. The high elevations of Texas occur in the drier climates that have higher sediment concentrations. The elevation correlation is stronger than the rainfall correlation. This stronger correlation may have to do with the resolution of the data. The elevation grid has a finer resolution than the rainfall grid.

For the most part, the sign of the correlation with concentration is as expected. Steeper slopes are more susceptible to erosion and produce higher concentrations. The negative correlation with rainfall is due to climate type. The erosion resistant soils have lower sediment concentrations. The only land use variable that had a significant correlation at the 90% level is *bare*. As the amount of bare lands increase, the concentration decreases. This effect is not readily explained but the correlation coefficient is low. The same variable in a multivariate regression model can have an opposite effect. The only size variable with significant correlation is the length of the watershed. The length may correlate better than the area with concentration because it may be associated with whether the gauged stream is in equilibrium.

Specific sediment yield - the efflux of sediment from unit area per unit time - varies widely from one environment to another. If the hyper-arid and arid deserts are ignored because runoff is extremely rare, there is a strong inverse relationship between sediment yield and area (Langbein & Schumm, 1958). According to Reid and Frostick (1994), the greatest yields are derived from semi-arid terrain despite the infrequency of flash floods, while temperate forests produce only about one third as much sediment despite the much higher frequency of storm events and the fact that flow is perennial. The specific sediment yield relationship for many different basins is shown by Reid and Frostick (1994) and repeated in Figure 4-21.

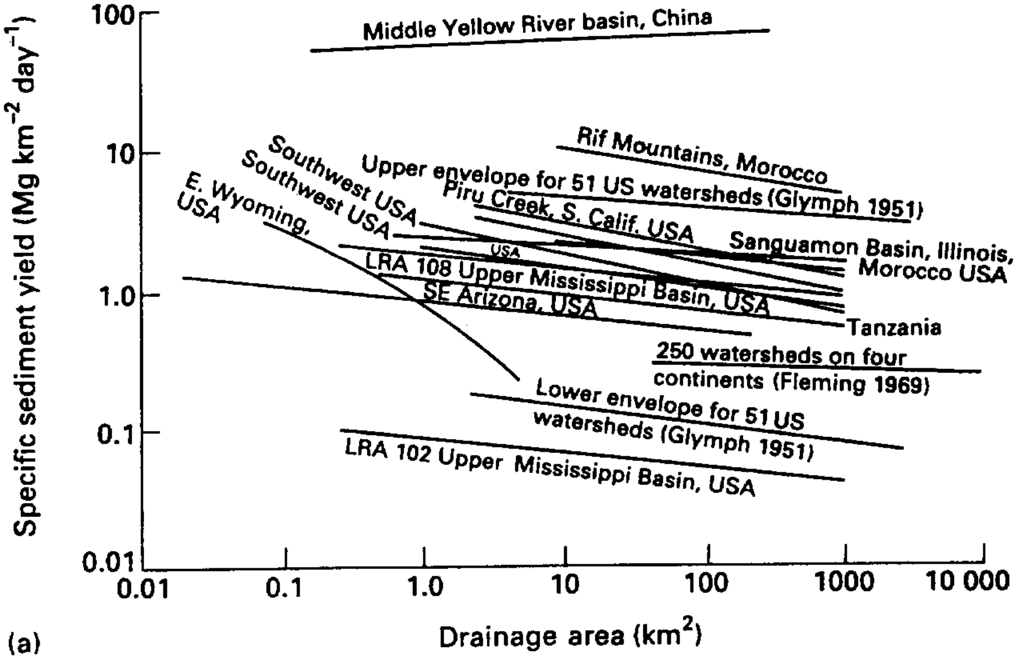


Figure 4-21. Compilation of Sediment Yield Curves (Reid and Frostick, 1994)

Figure 4-22 shows the same relationship with the average values of this research. Although the inverse trend is apparent, the correlation is very poor. No conclusions can be drawn from the large area - low specific sediment yield data points as these five data points all represent basins with upstream dams. The bivariate regression equations are not shown because there is no statistical significance in the relationship regardless of whether dams are considered. Figure 4-21 does not include basins as large as some of the TWDB stations. The basins with areas less than 10,000 km^2 fall in the vicinity of Glymphs's lower envelope for 51 US watersheds.

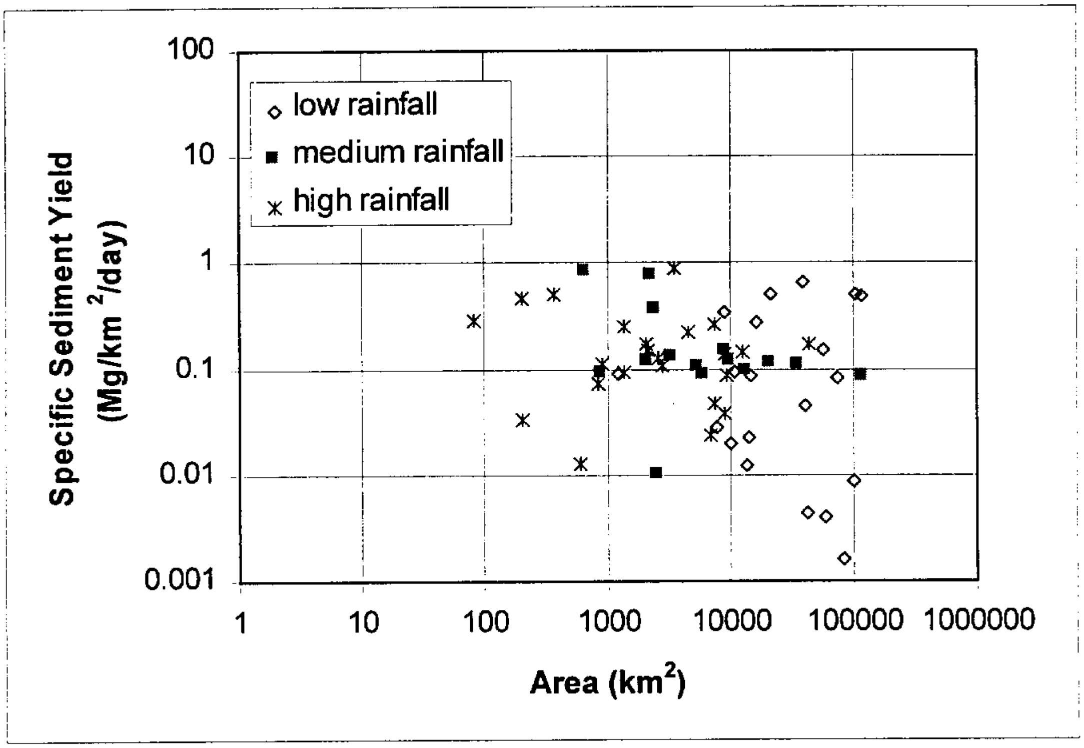


Figure 4-22. Specific sediment yield by climate for average values

4.5 Multivariate Analysis

Sediment load and concentration are dependent upon many variables; thus, it is reasonable that the bivariate analyses do not yield high correlations. Most of the additional variables are related to the watershed, not the stream which carries the sediment. Correlating suspended sediment with watershed variables assumes that the suspended sediment is largely attributable to washload or that the stream characteristics are strongly correlated with the watershed. Both assumptions are reasonable.

With so much data and so many variables, countless regression models can be analyzed. As Section 4.3 demonstrates, there are many ways to analyze the data. Although there are three SAS data sets, there are numerous data sets formed from those three. The four basic data sets are shown in Table 4-19. Each of the four data sets can be reduced by setting a criteria or by classifying the data as shown in Tables 4-20 and 4-21.

Table 4-19. Four basic data sets

average values		annual values	
with dams	without dams	with dams	without dams

Table 4-20. Criteria that reduce the data sets

No. yrs >= 5 (averages)
No. yrs >=10 (averages)
Use of variable not available for each station
Use data with complete data set from 1970-1978

Table 4-21. Data Classifications (number of classifications)

Years (61)
Basins (12)
Climates (3)

An example is used to help explain the complexities involved with multivariate regression for the data. Table 4-18 can serve as a starting point for choosing variables for a model. SAS is used to perform a stepwise regression for the dependent variable *lnloadkg* and the independent variables: *lnarea, lnleng, lnpotq, lnromm, slop_max, range, relief_m, soilB, lnelev, lnrain, strmfreq, lntime,* and *lnresvar*. These variables are not available for stations 8033300 and 8037050; thus, those 2 stations are removed from the analyses. The same stepwise regression is used for annual and average data. Retaining all variables that are significant at the 0.15 level, two very different models result for the average and annual data sets. The models respectively are

$$\begin{aligned}\mathit{lnloadkg} &= -16.63 + 1.64(\mathit{lnleng}) + 0.80(\mathit{lnelev}) + 3.29\,(\mathit{lnrain}) - 0.25(\mathit{lnresvar})\\ &\quad R^2 = 0.54 \text{ Adj. } R^2=0.51 \text{ DF}=56 \text{ SE}=1.27 \text{ F}=15.5 \qquad \textbf{(4-1)}\end{aligned}$$

$$\begin{aligned}\mathit{lnloadkg} &= 10.75 + 2.13(\mathit{lnleng}) + 0.50(\mathit{lnromm}) - 0.076(\mathit{slop\_max})\\ &\quad - 0.012(\mathit{soilB}) - 0.32(\mathit{lnresvar})\\ &\quad R^2 = 0.46 \text{ Adj. } R^2=0.46 \text{ DF}=1300 \text{ SE}=1.65 \text{ F}=219 \qquad \textbf{(4-2).}\end{aligned}$$

Only two variables, *lnleng* and *lnresvar*, are common to both models. In both models, the coefficient for *lnleng* is positive and the coefficient for *lnresvar* is negative. The coefficients have different values but are of similar magnitude. The *lnleng* variable represents the size of the watershed. Larger watersheds have larger sediment loads. The negative coefficient for the reservoir variable simply means that the larger the reservoir variable, the less sediment is received at the sampling station. The reservoir variable increases as the percentage of the watershed which is dammed increases.

Equation 4-1 shows positive coefficients for *lnelev* and *lnrain*. The positive coefficient for *lnelev* implies that if all other variables in the model are held constant, the station with the highest elevation will have the highest sediment load. In an area like Texas, the number of meters a station is above sea level has nothing to do directly with sediment load. The elevation of a Texas station is however, a good indicator of the station's location. The higher elevations are in the western portion of the state where the climate is more arid. This area of the state is also more prone to thunderstorms which are highly erosive.

Note the coefficients in the model for annual values shown by Equation 4-2. The positive coefficients for *lnleng* and *lnromm* are expected. In this case, *lnleng* corresponds with the area of the basin. Larger basins create more sediment yield. A positive coefficient for *lnromm* simply implies a cause and effect relationship between runoff and sediment yield. If additional variables related to the size of the watershed are included, then the coefficient for all of them is not expected to be positive. The negative coefficient for *slop_max* does not correctly reflect the impact steep slopes have on erosion. Because of the lower stability of a steep slope, it is expected that more erosion occurs on a steep slope than a flat slope. The impact on sediment concentration from a high percentage of hydrologic soil group

B soils is uncertain. Because the coefficients for *slop_max* and *soilB* are not easily explained, it is better to create a model without these coefficients.

The stepwise procedure for the annual values requires seven steps. The third step includes three variables: *lnromm*, *lntime*, and *lnresvar*. The coefficient for *lntime* is +1.55. Step four adds *lnleng* to the regression. R^2 increases from 0.42 to 0.45. The coefficient for *lnleng* is +1.96 and the coefficient for *lntime* is -0.11. When *lntime* is the only variable related to the size of the basin, it takes on a positive coefficient. However when several variables relate to the size of the basin, the variable coefficients will take on differing signs representing both supply and opportunities for deposition.

The model shown by Equation 4-1 (and repeated in Equation 4-3) is also used to develop coefficients for the variables using the annual data set. Models using the average and annual data sets with the same four variables are

$$\mathit{lnloadkg} = -16.63 + 1.64\,(\mathit{lnleng}) + 0.80\,(\mathit{lnelev}) + 3.29(\mathit{lnrain}) - 0.25\,(\mathit{lnresvar})$$
$$R^2 = 0.54 \quad \text{Adj. } R^2=0.51 \quad DF=56 \quad SE=1.27 \quad F=15.5 \qquad \mathbf{(4\text{-}3)}$$

$$\mathit{lnloadkg} = -26.20 + 2.03\,(\mathit{lnleng}) + 0.64(\mathit{lnelev}) + 4.40(\mathit{lnrain}) - 0.32\,(\mathit{lnresvar})$$
$$R^2 = 0.41 \quad \text{Adj. } R^2=0.41 \quad DF=1300 \quad SE=1.72 \quad F=229 \qquad \mathbf{(4\text{-}4).}$$

All of the variables in Equation 4-4 are significant at the 0.0001 level. The model for the average values has stronger correlation because the average data set has fewer outliers since the years with extreme values get averaged in with the less-extreme years. The magnitude of each coefficient is similar but if the equations are used for prediction, the results will be different.

Using average values for the variables (*lnleng*=5.47, *lnelev*=5.58, *lnrain*=6.69 and *lnresvar*=-2.19) results in *lnloadkg*=19.36 using Equation 4-3, and *lnloadkg*=18.61 using Equation 4-4. The corresponding loads are $2.56\text{x}10^8$ kg/year and $1.21\text{x}10^8$ kg/year; thus, the model developed using average values can result in predicting twice the annual load of the model developed using the annual values.

The model shown by Equation 4-2 is revised by omitting variables *slop_max* and *soilB*. Equations 4-5 and 4-6 result for the average and annual data sets, respectively as

$$lnloadkg = 11.29 + 1.55\,(lnleng) + 0.13\,(lnromm) - 0.26\,(lnresvar)$$
$$R^2 = 0.51 \quad \text{Adj. } R^2=0.48 \quad DF=56 \quad SE=1.31 \quad F=18.7 \qquad \textbf{(4-5)}$$

$$lnloadkg = 11.97 + 1.84\,(lnleng) + 0.56\,(lnromm) - 0.31\,(lnresvar)$$
$$R^2 = 0.45 \quad \text{Adj. } R^2=0.44 \quad DF=1300 \quad SE=1.67 \quad F=348 \qquad \textbf{(4-6).}$$

All of the variables in Equation 4-5 are significant at the 0.15 level. All of the variables in Equation 4-6 are significant at the 0.0001 level. Not much is lost by omitting the variables *slop_max* and *soilB*. The values for R^2 and the standard error (SE) are similar for Equation 4-2 and 4-6. In fact, Equation 4-6 is preferable because the coefficients of all of the variables are physically legitimate.

The correlation is again stronger for the model developed using the average data set. If the equations are used for predicting purposes, the results vary. Using average values for the variables (*lnleng*=5.47, *lnromm*=-7.16, and *lnresvar*=-2.19) results in *lnloadkg*=19.41 ($2.68\text{x}10^8$ kg/year) using Equation 4-5, and *lnloadkg*=18.70 ($1.33\text{x}10^8$ kg/year) using Equation 4-6. Considering Equations 4-3, 4-4, 4-5, and 4-6, the data set chosen can affect the model as much as the variables chosen.

The coefficients for the same model for the annual data set can be determined using a weighting factor equal to the inverse of the number of years of data for each station. By using the weighting factor, stations with more years of data do not influence the model more that those stations with fewer years of data. The stepwise regression result is

$$\begin{aligned} lnloadkg &= 10.20 + 3.01(lnleng) + 0.60(lnromm) - 0.01(range) \\ &\quad + 0.21(lnelev) - 0.95(lntime) - 0.35(lnresvar) \\ &R^2 = 0.52 \quad \text{Adj. } R^2 = 0.52 \quad DF=1300 \quad SE=0.32 \quad F=350 \qquad \textbf{(4-7).} \end{aligned}$$

Variables *lnleng*, *lnromm*, and *lnresvar* are common to both the weighted (Equation 4-7) and un-weighted (Equation 4-2) versions of the stepwise model. If the weighted model is reduced to just using these common variables, the result is

$$\begin{aligned} lnloadkg &= 11.29 + 1.93\,(lnleng) + 0.52\,(lnromm) - 0.36\,(lnresvar) \\ &R^2 = 0.50 \quad \text{Adj. } R^2=0.50 \quad DF=1300 \quad SE=0.33 \quad F=437 \qquad \textbf{(4-8)} \end{aligned}$$

which is similar to Equation 4-6. The coefficient of determination, R^2, improves slightly, but the standard error, SE, is reduced by eighty percent, from 1.67 to 0.33. The weighted model with the same selection of variables as in Equation 4-4 results in

$$\begin{aligned} lnloadkg &= -21.42 + 1.96\,(lnleng) + 0.71(lnelev) + 3.69(lnrain) - 0.37\,(lnresvar) \\ &R^2 = 0.45 \quad \text{Adj. } R^2=0.45 \quad DF=1300 \quad SE=0.35 \quad F=262 \qquad \textbf{(4-9).} \end{aligned}$$

The coefficients are similar to those in Equation 4-4, but the standard error is reduced from 1.72 to 0.35.

The weighted models are similar to the models without the weight; however the weighted models have slightly higher agreement. In general, the *lnloadkg* = *f(lnleng, lnromm, lnresvar)* model has a higher R^2 than the *lnloadkg = f(lnleng, lnelev, lnrain, lnresvar)* model. The *lnloadkg = f(lnleng, lnromm, lnresvar)* model will be the focus of the remaining discussion. The coefficients and statistics for this model are shown for numerous data sets in Table 4-22. Models (1) through (9) are developed using the annual values while models (10) through (14) are developed using the average values. A few of the models are repeated from equations presented earlier in this section.

Table 4-22. Model comparison using various data sets

(dependent variable = *lnloadkg*)

	coefficients							
data description	intercept	*lnleng*	*lnromm*	*lnresvar*	R^2	DF	SE	F
(1) annual	11.97	1.84	0.56	-0.31	0.45	1300	1.67	348
(2) weighted annual	11.29	1.93	0.52	-0.36	0.50	1300	0.33	437
(3) arid, weighted annual	4.23	2.46	0.14	-0.59	0.52	522	0.36	190
(4) moderate, weighted annual	18.66	1.51	1.11	-0.12	0.78	363	0.20	422
(5) humid, weighted annual	14.75	1.60	0.80	-0.26	0.59	413	0.25	198
(6) annual, no dams	14.19	1.88	0.72		0.55	569	1.29	343
(7) no dams, weighted annual	12.84	2.09	0.64		0.65	569	0.34	525
(8) all dams, annual	12.80	1.51	0.37	-0.57	0.42	730	1.82	174
(9) all dams, weighted annual	14.69	1.33	0.47	-0.47	0.45	730	0.34	200
(10) averages	11.29	1.55	0.13	-0.26	0.51	56	1.31	18.7
(11) arid, averages	0.62	2.39	-0.45*	-0.48	0.63	20	1.42	9.71
(12) moderate, averages	19.09	1.30	0.97	-0.15*	0.65	14	0.97	6.94
(13) humid, averages	8.84	1.71	-0.13*	-0.25*	0.57	20	1.09	7.37
(14) averages, no dams	11.71	1.76	0.20		0.65	29	1.18	25.4

* variable is not significant at the 0.15 level

Models (6), (7), and (14) in Table 4-22 are created with data sets without dams. These three models have higher values for the intercept and for the *lnleng* and *lnromm* coefficients. The reservoir variable does not completely account for the effects of reservoirs. When data include stations with reservoirs, then the other variables in the model are also affected. Models (8) and (9) were developed by including only stations with upstream reservoirs. Both models have a higher negative coefficient for the reservoir variable than when stations without dams are included in the analyses. The models which exclude those basins with dams [(6), (7), and (14)] have a higher R^2 then the counterparts for all the stations [(1), (2), and (10)] or for those stations with dams [(8) and (9)]. The reservoir variable is effective in explaining a portion of the variability in sediment load due to reservoir construction, but not all of the variability.

When the *lnloadkg = f(lnleng, lnromm, lnresvar)* model is determined separately for the three climates, the results improve. For both the annual and average values, the stations classified as having a moderate climate result in the model with the highest correlation and lowest standard error. According to the models, reservoirs in the arid climates have more impact on the load than those in the moderate or humid climates. The coefficients for the variables do not consistently increase or decrease from one climate to the next. There is no reasonable physical explanation for the reservoirs to have the least impact in the moderate climates. Because there are missing variables, the values of the coefficients may be biased.

The number of data points used to determine the models for the separate climates using average values is much lower than for any of the other models. The smaller number of data points results in decreasing the significance of some of the variables. The coefficients in equations (11), (12), and (13) of Table 4-22 that are not significant at the 0.15 level are marked. Of these four coefficients, only *lnromm* in equation (13) is not significant at the 0.25 level. The models improve for specific data sets, allowing for fewer generalizations about the nature of sediment yield.

4.6 Model Discussion

Choosing the model that gives the highest R^2 may not give the best model. Gujurati (1995) states that the objective of multivariate regression is to obtain dependable estimates of the true regression coefficients and draw statistical inferences about them. It is important to evaluate the logical or theoretical relevance of the explanatory variables to the dependent variable and to evaluate the statistical significance of the coefficients.

Although the variables were chosen to represent physical processes, the physical processes are not always evident in the model with the highest correlations. Several steps are taken to narrow the number of possible models. The physical processes are considered to insure the theoretical relevance of the explanatory variables.

The "best" model for one data set is not the "best" model for another data set. Those stations that fall in the arid category have better correlation when the dependent variable is concentration rather than load. Stations in the moderate and humid categories have better correlation using sediment load. Two models selected as an example for the arid climate stations are

$$\begin{aligned} lnconc &= -20.97 + 2.89\,(lnelev) - 0.25\,(bare) - 0.37\,(USGSslp) \\ &\quad + 23.73\,(k\_max) - 0.46\,(lnresvar) \\ &\quad R^2 = 0.62\ \ \text{Adj. } R^2{=}0.61\ \ DF{=}522\ \ SE{=}0.35\ \ F{=}165 \end{aligned} \qquad \textbf{(4-10)}$$

$$\begin{aligned} lnloadkg &= 10.35 - 1.07\,(lnarea) + 3.09\,(lnleng) + 0.22\,(lnromm) \\ &\quad + 0.0019\,(relief\_m) - 0.53\,(lnresvar) \\ &\quad R^2 = 0.54\ \ \text{Adj. } R^2{=}0.54\ \ DF{=}522\ \ SE{=}0.36\ \ F{=}123 \end{aligned} \qquad \textbf{(4-11).}$$

The models were developed using the weighted annual data set. The values of the coefficients for *bare* and *USGSslp* in Equation 4-10 are not explained well physically. As the percentage of bare soils increases from one watershed to another, the amount of soils more susceptible to erosion increases. As the slope of a river increases from watershed to another, more particles are kept in suspension. The model however does not express these principles based on the coefficients of the variables. On the other hand, if these two variables are removed from the model, R^2 is reduced from 0.62 to 0.49. The interdependence of the variables makes it difficult to define causality.

Equation 4-11 has coefficients that are reasonable. Two variables, *lnarea* and *lnleng*, represent the size of the basin. The positive coefficient for *lnleng* indicates the potential supply of sediment. The negative coefficient for *lnarea* indicates the opportunities for deposition of sediment. As runoff and the relief increase, the sediment yield increases. The reservoir variable has a larger negative coefficient than for the other climates. Possibly, it is more difficult for rivers to stay in equilibrium in climates with lower rainfalls. A similar model to Equation 4-11 uses *lnelev* rather than *relief_m*. The models are very similar but a positive coefficient for *relief_m* has more physical basis than a positive coefficient for *lnelev*.

Equations 4-12 and 4-13 present two selected models for the moderate climate. As with the arid climate, the models were developed using the weighted annual data set.

$$\begin{aligned} \mathit{lnloadkg} = {} & 16.74 + 0.84\,(\mathit{lnarea}) + 0.27(\mathit{lnleng}) + 1.09\,(\mathit{lnromm}) \\ & + 0.19(\mathit{lnelev}) - 0.16\,(\mathit{lnresvar}) \\ & R^2 = 0.79 \quad \text{Adj. } R^2=0.79 \quad DF=363 \quad SE=0.19 \quad F=275 \end{aligned} \qquad \textbf{(4-12)}$$

$$\begin{aligned} \mathit{kgload} = {} & 1960000000 + 35700\,(\mathit{flowm3}) - 2050000\,(\mathit{rain\_avg}) \\ & - 876000\,(\mathit{leng\_km}) \\ & R^2 = 0.70 \quad \text{Adj. } R^2=0.70 \quad DF=363 \quad SE=146000000 \quad F=279 \end{aligned} \qquad \textbf{(4-13)}$$

Lnleng and *lnelev* are significant at the 0.28 and 0.20 levels in Equation 4-12. In contrast to Equation 4-11, *lnarea*, *lnleng*, and *lnromm* all have positive coefficients. The reservoir variable consistently has a negative coefficient in all models. The reservoir variable is more independent than the other variables related to watershed characteristics.

Because there is not a great amount of variability in the magnitude of the variables, the model (Equation 4-13) using variables which are not transformed shows good correlation. In this model, the length variable has a negative coefficient which would indicate the longer basins are depositing more sediment and not keeping the sediment in suspension. The negative coefficient for rainfall may well be related to the land cover that exists in the areas of higher rainfall. Adding the reservoir variable to Equation 4-13 does not improve the model.

Equations 4-14 and 4-15 present two selected models for the humid climate. Again, the models were developed using the weighted annual data set.

$$\begin{aligned}\mathit{lnloadkg} = {} & 15.51 + 0.51\,(\mathit{lnarea}) + 0.99\,(\mathit{lnleng}) + 0.94\,(\mathit{lnromm}) - 0.49\,(\mathit{slop\_max}) \\ & + 0.0067\,(\mathit{relief\_m}) - 0.24\,(\mathit{lnresvar}) \\ & R^2 = 0.73 \quad \text{Adj. } R^2 = 0.73 \quad DF = 413 \quad SE = 0.20 \quad F = 185 \qquad \mathbf{(4\text{-}14)}\end{aligned}$$

$$\begin{aligned}\mathit{kgload} = {} & 239000000 + 32200\,(\mathit{flowm3}) + 0.116\,(\mathit{area}) \\ & - 5740\,(\mathit{length}) - 729000000\,(\mathit{resvar}) \\ & R^2 = 0.66 \quad \text{Adj. } R^2 = 0.65 \quad DF = 413 \quad SE = 127000000 \quad F = 195 \qquad \mathbf{(4\text{-}15)}\end{aligned}$$

Although steeper slopes might be an indication of more exposure and more erosion, the model shown by Equation 4-14 has a negative coefficient associated with the variable *slop_max*. When the variable *slop_max* is dropped from the model, the R^2 drops from 0.73 to 0.63. Because this model is developed for the more humid regions of the state, perhaps the *slop_max* variable is related to limestone formations which contribute very little to sediment yield.

As in the moderate climate, the model using variables which are not transformed shows good correlation. In this model, the two variables relating to size (*area* and *length*) have opposite signs. The reservoir variable improves the model for the humid climates.

A model developed using the annual weighted data set but including the stations from all of the classifications results in

$$lnloadkg = 12.72 - 0.25\ (lnarea) + 2.08\ (lnleng) + 0.57\ (lnromm) + 0.00071\ (relief\_m) - 0.35\ (lnresvar)$$
$$R^2 = 0.51 \quad \text{Adj. } R^2=0.51 \quad DF=1300 \quad SE=0.33 \quad F=267 \qquad \textbf{(4-16).}$$

Equation 4-16 can be used to make some generalizations on how sediment yield varies across the state of Texas. The runoff and relief variables indicate the type of climate of the watershed. Watersheds with a high value for relief are in the steeper terrain areas of the state. Watersheds with higher runoff depths will be in the more eastern parts of the state.

4.7 Model Evaluation

Correlation among the independent variables causes the regression estimates to change depending on which independent variables are being used. For example, the impact of rainfall on sediment load depends on whether elevation is in the equation or not. Furthermore, the effect on a variable coefficient is different depending on the data set.

Variables relating to watershed size, runoff, and reservoirs are the most effective in describing the variation of sediment yield at a station. Of course, the variables relating to the specific watershed characteristics have a direct impact on the actual runoff.

In using the annual data set instead of the average data set, the degrees of freedom are increased so that more variables can be included in a model. As more variables are included in the models, R^2 will necessarily increase. However, the small incremental increases in R^2 are at the expense of reasonable coefficient interpretation. Although including more variables increases R^2 to values closer to 1, the coefficients for the variables may be difficult to explain physically. One reason the variable coefficients take unexpected values is multicollinearity. Many watershed characteristics are related to each other. Table 4-23 presents the Pearson correlation coefficients for several of the watershed variables. For the sediment gauging stations, elevation and area even show a considerable correlation.

Table 4-23. Correlation coefficients for selected watershed characteristics

	lnarea	*lnleng*	*lnromm*	*lnelev*	*relief_m*
lnarea	1.00	0.98	-0.49	0.55	0.83
lnleng	0.98	1.00	-0.52	0.62	0.87
lnromm	-0.49	-0.52	1.00	-0.81	-0.64
lnelev	0.55	0.62	-0.81	1.00	0.77
relief_m	0.83	0.87	-0.64	0.77	1.00

The models presented in Equations 4-10 to 4-16 are developed using stepwise regression, followed by an evaluation of model coefficients. When the average data set is used to determine variable coefficients for the same models, the models have similar values of R^2. However, because of the low degrees of freedom, the adjusted R^2 is much lower and the individual variables are less significant.

4.8 Conclusions

The models developed to help analyze the effects of gross watershed characteristics are not necessarily intended to be accurate prediction tools. The models certainly provide guidance as to how watershed characteristics affect sediment load. Furthermore, the reservoir variable (*resvar*) can be a reasonable indicator of how a new dam may affect sediment load.

Because the data are so complex, it is recommended that the data be used according to the purpose of the analysis. The equations presented in Table 4-22, as well as equations 4-10 through 4-16 can be used to generally estimate sediment loads and how those loads are affected by the addition or deletion of dams. To quantify the potential effects of building a new dam in the Guadalupe River Basin (for example), equations (4) and (12) of Table 4-22 and equation 4-12 can be used. If a more precise estimation is required, the data set should be narrowed so that more variables can be held constant. The SAS models and SAS data sets, that are included on the CD available for this research, can easily be modified to produce very specific models for specific situations. It is best to choose a model created from similar data. If estimates of sediment load are required in west Texas, it is best to use an analysis including only west Texas stations and similarly for any area of the state.

The data are enormous and extremely useful. Although general trends can be noted from the analyses, a model accurately describing how sediment yield varies for all conditions in Texas cannot be created. Equation 4-16 generalizes the trends of sediment load for the entire state. Chapters 5 and 6 present further analysis of the sediment yield for the Lavaca River.

5 Improvement of Sediment Rating Curve using USGS data for the Lavaca River near Edna

5.1 Purpose

The primary purpose of analyzing the USGS sediment concentration data for the Lavaca River is to determine the effect of water temperature and grain size distribution on sediment concentration. The data are sporadic in the temporal sense but the knowledge gained from analysis can be applied to the study of daily data that have no information on temperature or grain size distribution.

5.2 Data Description

The USGS takes periodic water quality measurements at various stations. For the Lavaca River near Edna, these measurements include 76 observations at intervals of approximately 1 to 5 months during the period from 1972 to 1992. The data were obtained from a Hydrosphere CD-ROM. The collected data include stream flow rate, the percent of sampled grains that is finer than 0.062 mm, the water temperature, and the sediment concentration. A summary of the data is included in Table 5-1.

Table 5-1. Summary of Data for the Lavaca River near Edna

measured quantity	average	minimum	maximum
flow rate (cfs)	597	0.01	26900
sediment concentration (mg/l)	114	3	745
sediment load (tons/day)	215	0	10801
% grains < 0.062 mm	63	4	100
temperature (°C)	21	6.5	31.5

There is a large variation in the 76 observations. The average sediment concentration is rather low, requiring very large samples of sediment-water mixture to allow enough soil for sieve and gradation tests. On average, 63% of the grains pass the smallest sieve, implying that the sediment samples are mostly fine-grained. The average water temperature is moderately warm, characteristic of the Texas coast.

5.3 Background and Physical Parameters

5.3.1 Sediment Rating Curves - Flow Rate Effects

The stream flow rate in an unsteady hydrograph is a function of temporal and spatial characteristics of precipitation, watershed characteristics, and channel storage and resistance characteristics. Sediment yield is also a function of precipitation, watershed characteristics, and channel characteristics as shown in Figure 5-1.

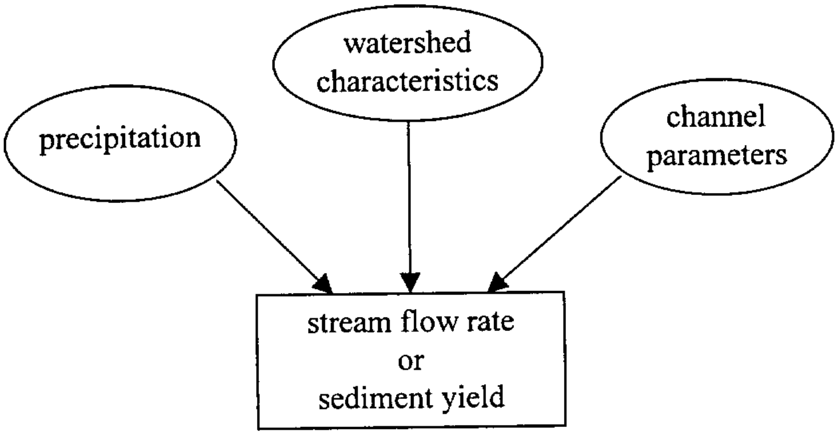


Figure 5-1. Cause and Effect of Sediment Yield

Thus, it is reasonable to assume a high degree of correlation between sediment load and stream flow rate. This assumption results in the traditional sediment rating curve shown in Figure 5-2 and Table 5-2. The regression line, in the often used form, is

$$load = 0.90\ (flow)^{1.21} \qquad \textbf{(5-1)}$$

where load is in tons/day and flow is in cfs.

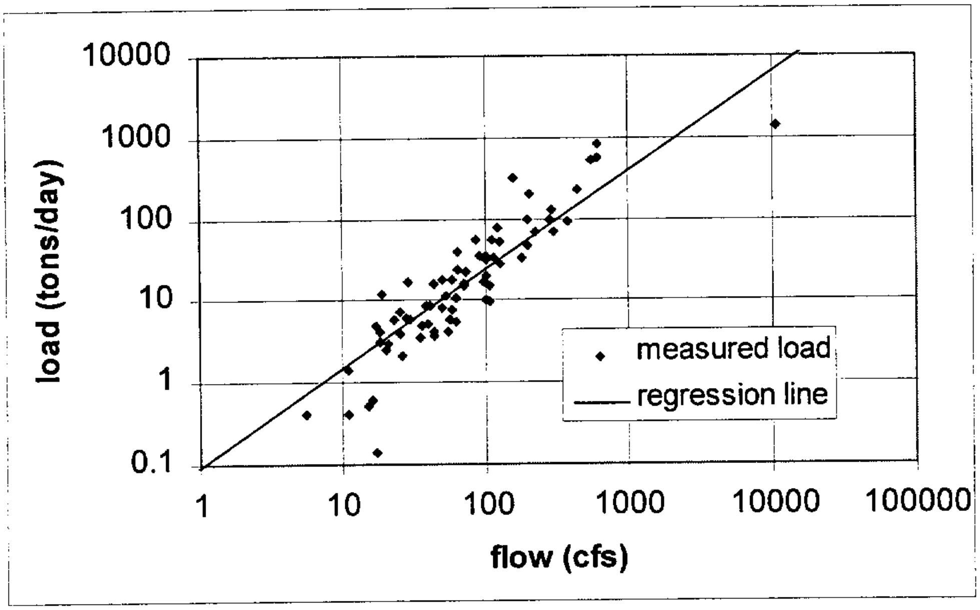


Figure 5-2. Traditional Sediment Rating Curve for the Lavaca River Near Edna (USGS Data)

Table 5-2. Regression line for Lavaca River near Edna (USGS Data)

Equation *(t-statistics)*	R^2	DF	SE	F
ln load = -2.41 + 1.21 (*ln flow*) *(-12.0) (26.8)*	0.91	75	0.87	719

Table 5-2 is similar to the tables shown in Chapter 4, with R^2 = coefficient of determination, DF = degrees of freedom, SE = standard error, and F = F-value. The high degree of correlation is not typical of sediment rating curves. Table 4-15 shows sediment rating curves based on annual values for the different basins in Texas (without dams). The coefficient of determination varies from 0.14 to 0.90.

Although a high degree of correlation exists, the stream flow does not fully explain the variability in sediment load. Furthermore, the relationship between sediment load and stream flow rate is biased because the sediment load is a product of sediment concentration and stream flow rate; thus the stream flow rate is included in both the abscissa and ordinate of the sediment rating curve.

Another form of a sediment rating curve can include the sediment concentration rather than the sediment load so that the bias in the traditional sediment rating curve is removed. As expected, the coefficient of determination (R^2) is dramatically reduced, from $R^2 = 0.91$ to $R^2 = 0.22$. As shown in Figure 5-3, the samples can deviate by an order of magnitude from the regression line. The regression line, which can be written as

$$conc = 33.4\ (flow)^{0.21} \qquad \textbf{(5-2)}$$

is shown in Table 5-3 along with the statistics of the regression.

Knowledge of how flow rate physically affects the sediment concentration should help explain some of the scatter. The stream flow rate is limited by the amount of rainfall, whereas the sediment yield is limited by sediment available for supply and by capability of transporting sediment. Therefore, it is expected that as flow rate increases, the rate of change of sediment concentration with flow rate will decrease. To account for this limiting phenomena, it is necessary to include a polynomial term in the regression (i.e., $(\ln flow)^2$ in addition to $\ln flow$). The regression statistics for this curve are shown in Table 5-4.

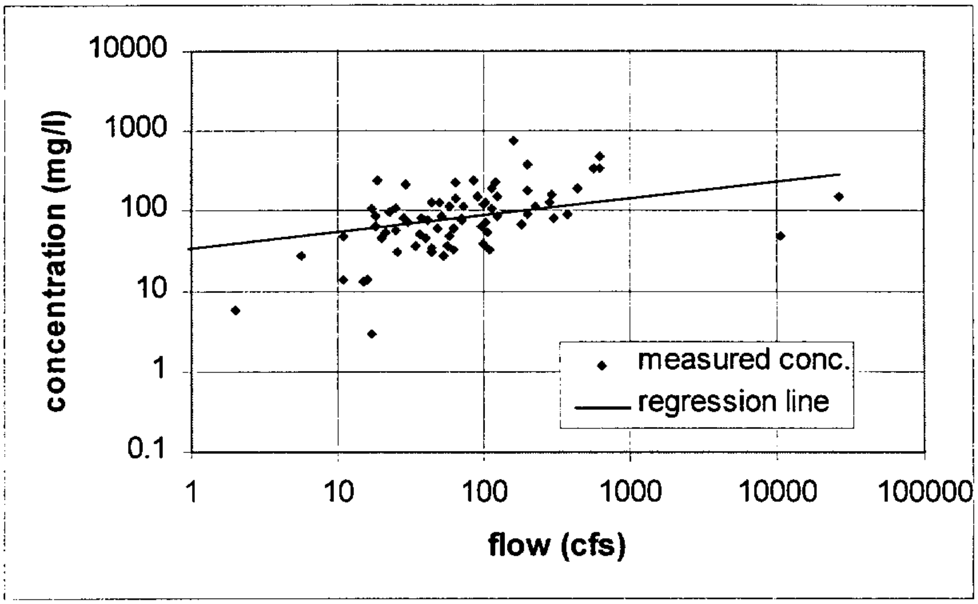


Figure 5-3. Sediment Rating Curve using Concentration for the Lavaca River near Edna (USGS Data)

Table 5-3. Regression line for Lavaca River near Edna (USGS Data) using concentration

Equation *(t-statistics)*	R^2	DF	SE	F
$\ln conc = 3.51 + 0.21 (\ln flow)$ *(17.5) (4.63)*	0.22	75	0.87	21.4

Table 5-4. Regression line for Lavaca River near Edna (USGS Data) using concentration and polynomial term

Equation *(t-statistics)*	R^2	DF	SE	F
$\mathit{ln\ conc} = 1.02 + 1.16\,(\mathit{ln\ flow}) - 0.0799\,(\mathit{ln\ flow})^2$ $(1.84)\quad(5.35)\qquad(-4.00)$	0.39	71	0.75	22.4

This polynomial relationship with flow rate is not typically included in sediment rating curves. The flow rate explains the general trend of sediment concentration, as shown in Figure 5-4, when the polynomial term is included. Even though the correlation is improved over that shown in Figure 5-3, the data continue to show a wide range of concentration values for a given flow rate.

The regression equation shown in Figure 5-4 shows a decrease in concentration with flow when the flow exceeds approximately 1400 cfs. Furthermore, a maximum concentration of 190 mg/l is achieved. Eleven of the 76 samples exceed 190 mg/l. It is evident from the figure that the two samples with flow rates exceeding 10,000 cfs have considerable influence in determining the shape of the regression line. To further answer the questions raised by this analysis, the addition of other variables is explored.

Assuming that the temperature of the water is a proxy variable for viscosity and the percentage of fines is a proxy variable for particle fall velocity, a model showing causality between variables is shown in Figure 5-5. The multiple causality between variables is one of the reasons sediment concentration is so hard to predict.

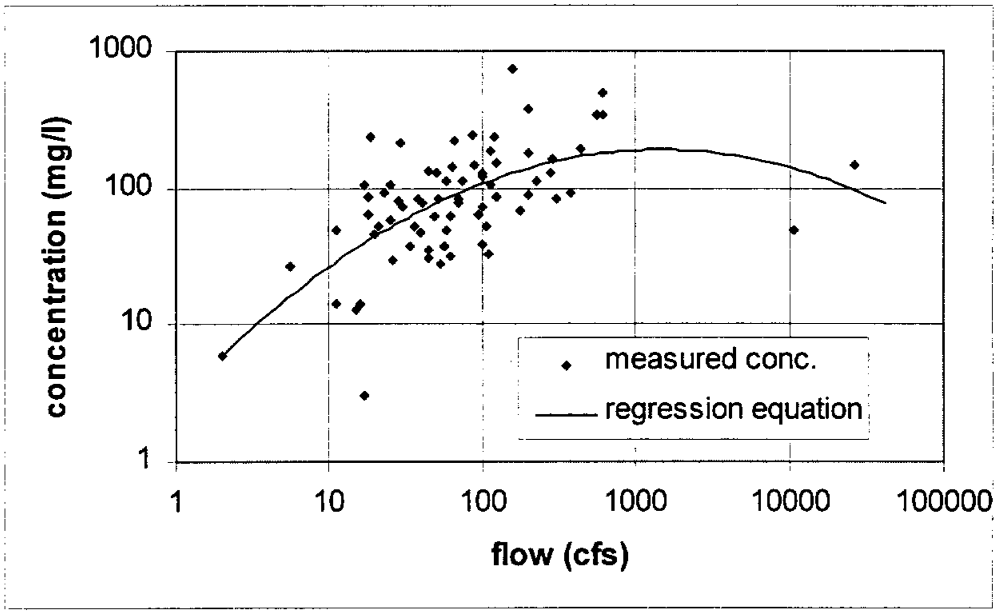


Figure 5-4. Modified Sediment Rating Curve using Concentration for the Lavaca River Near Edna (USGS Data)

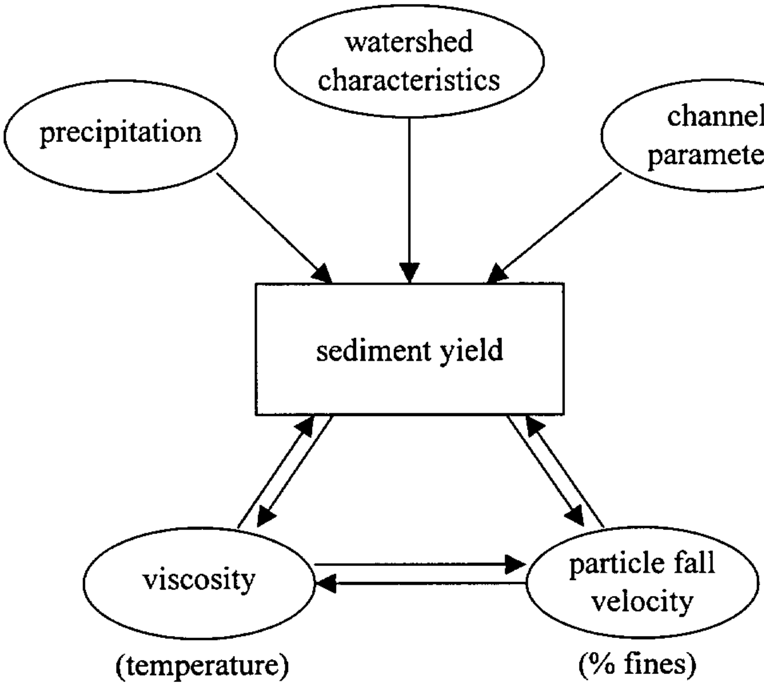


Figure 5-5. Multiple Causality Between Variables Affecting Sediment Yield

5.3.2 Temperature Effects

Water temperature can explain some of the variability in sediment concentration for a given flow rate. A temperature increase results in lower fluid viscosity and thereby higher fall velocities. Based on this fact alone, one would expect sediment concentration to decrease with increased temperature. These changes in fluid characteristics also cause a change in the bed form. Vanoni (1975) reports studies by Colby and Scott (1965) that show a ripple and dune bed at 83°F and a relatively flat bed at 39°F for the same discharge in the same river. In this case, an increase in temperature is responsible for an increase in roughness which would result in a decrease in velocity and a corresponding decrease in stream power. Therefore temperature is a factor in the local hydraulics and in the transporting capability of the river. Higher water temperatures result in less transporting capacity of suspended sediment, but the bed formation action may contribute to bed load transport.

Perhaps more important is the time of year which is reflected by the temperature. Colder months will have less ground cover and thus less protection from erosion. Also the nature of storms may be different from summer to winter months.

5.3.3 Size Effects

The size information serves several purposes. The size of the grains determines the particle fall velocity. The stream power must be greater to transport particles of fast fall velocities than for those with slower fall velocities. Although the information does not present the gradation of the samples, knowledge of the percent fines can be useful. The larger the percent fines, the more uniform the sediment concentration is vertically. If there is a very small percent of fine particles then bed material may be a significant contributor to the suspended load.

5.3.4 Hydrograph Limb Effects

Sediment concentrations are frequently higher on the rising limb of the hydrograph (Reid and Frostick, 1994; Holley, 1992). The topsoil is loose and more prone to erosion before and during the beginning of a rainstorm. After some time passes, the rain compacts the soil and limits the additional erosion which can occur. If the time base of the hydrograph is long enough that the daily stream flows can be properly classified as to the limb of the hydrograph on the days when the sediment samples were collected, then a variable describing hydrograph limb should further improve the model. Some hydrographs will have a time base shorter than a day. For those days, trying to define the limb of the hydrograph has no meaning when using daily data, as in this research.

5.4 Analysis

5.4.1 Multivariate Regression

Multivariate regression can be used as an exploratory tool in understanding possible causative factors of a physical process such as sediment yield. The goal of the regression model is to explain as much as possible of the variation observed in the response variable, leaving as little variation as possible to unexplained "noise" (Helsel and Hirsch, 1992). The data contain several variables affecting sediment concentration, making multivariate regression a reasonable analysis tool.

5.4.2 Modeling Software

SAS was used for determining the coefficients for the statistical models. This size of data set, however, is easily manageable in a spreadsheet such as Excel. Excel is capable of simple bivariate and multivariate regression. In addition, regression diagnostics such as residual plots are available. The disadvantage of using Excel lies in variable transformation. Every time a transformed variable is used, a new column of data must be inserted. Performing numerous operations on the spreadsheet with the original data leaves opportunities for error production. Furthermore, when developing a model with transformed variables, the model is not using the original data. On the other hand, using a statistical package such as SAS requires a data file to be read but not manipulated. In addition, review of the

SAS data file provides information on all data transformations. SAS was used for the analyses but some analyses were replicated using Excel.

5.4.3 Variable Transformation

Helsel and Hirsch (1992) note the primary reason to transform the response variable is that the data are heteroscedastic, that is, the variance of the residuals is a function of an independent variable. They also note one helpful generalization, namely that any y variable that covers more than an order of magnitude of values in the data set, as sediment loads typically do, probably needs to be transformed. The traditional and modified sediment rating curves previously presented use transformed variables for flow, load, and concentration. All three of these parameters cover more than an order of magnitude of values. Conversely, temperature and % fines do not warrant transformation. To confirm these assumptions, regressions with and without transformed variables were performed in addition to plotting the residuals.

5.4.4 Omitted data

The data contain 76 values. When the data are log transformed, 4 of the 76 values for flow rate are negative (i.e., 4 of the samples were taken with flow rates less than 1 cfs.) The negative values can cause an incorrect function shape for some models. For example, consider a model of the form

$$ln\ conc = b_0 + b_1\,(ln\ flow) + b_2\,(ln\ flow)^2 \qquad \textbf{(5-3)}$$

with a positive value for b_1 and a negative value for b_2, like the equation in Table 5-4. For a large range of values for *ln flow*, the second term will be a large positive number and the third term will be a smaller negative number. For this range of *ln flow* values, sediment concentration increases with flow rate, but as flow increases the rate at which the sediment increases begins to decrease. However, if a negative value for *ln flow* is used in the equation, the second and third terms of Equation 5-3 will both be negative so that sediment concentration decreases with flow rate. On the other extreme, a large value of flow rate exists where the sediment concentration starts to decrease with increased flow rate. The requirement that one should not apply a statistical model to values outside the range of data for which the model was developed is reiterated!

To correct the problems of the negative values of (*ln flow*), two approaches can be taken:

a) the units can be changed such that the values will be greater than 1, or
b) the four values with flow rates less than 1 cfs can be deleted from the analysis.

The latter approach was used largely because the very small values of flow are associated with small values of sediment concentration and contribute very little to the overall load. Table 5-5 shows the four samples with flow rates less than 1 cfs.

Table 5-5. Samples with flow rates less than 1 cfs

date	flow (cfs)	conc. (mg/l)
08/16/89	0.17	16
10/17/89	0.05	198
09/05/90	0.01	16
10/30/90	0.02	16

Note the sample taken on 10/17/89 with a concentration of 198 mg/l. With a flow rate of only 0.05 cfs, it is very difficult to take a water sample without collecting some bed material. This sample causes all of the models to have poor correlations. When these four samples are removed from the data set, the concentration models improve. Table 5-6 compares the models using 76 variables (DF=75) and 72 variables (DF=71). Most significantly improved is the model with the polynomial term (the last row in the table).

Table 5-6. Comparison of regressions using 76 data points and 72 data points

Equation *(t-statistics)*	R^2	DF	SE	F
$\ln load = -2.41 + 1.21 (\ln flow)$ *(-12.0) (26.8)*	0.91	75	0.87	719
$\ln load = -2.96 + 1.33(\ln flow)$ *(-9.69) (19.6)*	0.85	71	0.82	386
$\ln conc = 3.51 + 0.21 (\ln flow)$ *(17.5) (4.63)*	0.22	75	0.87	21.4
$\ln conc = 2.96 + 0.33 (\ln flow)$ *(9.70) (4.86)*	0.25	71	0.82	23.7
$\ln conc = 3.50 + 0.183 (\ln flow) - 0.00552 (\ln flow)^2$ *(17.3) (2.96) (-0.62)*	0.23	75	0.87	10.8
$\ln conc = 1.02 + 1.16 (\ln flow) - 0.0799 (\ln flow)^2$ *(1.84) (5.35) (-4.00)*	0.39	71	0.75	22.4

5.4.5 Use of Binary Variables

As implied by the name, binary variables simply take the value of 0 or 1. An advantage of using binary variables is the ability to plot results of multivariate regressions in two dimensions. To transform the variables *temperature* (temperature of the water) and *% fines* (the percent grains finer than 0.062 mm) from continuous variables to binary variables requires selecting a value that can be used to separate the data into two groups for each variable. For preliminary analysis, temperatures greater than 20° C are considered warm, giving the binary variable *warm* a value of 1. Temperatures less than or equal to 20° C are considered cold, or "not" *warm*, and the binary variable *warm* is given a value of 0. In a similar manner, large grain samples are defined as those samples with 50% or less of the sediment finer than 0.062 mm. Small grain samples have more than 50% fines and the variable *large* was given a value of 0. The binary variables for *temperature* and *% fines* are defined in the 'data' statement of the SAS file as follows:

```
if temperature <=20 then warm=0;
  else warm = 1;
if %fines <=50 then large = 1;
  else large = 0;
```

The hydrograph limb is a categorical variable. Two binary variables (*rise* and *fall*) are required to express three categories. Each sample in the data set was categorized as occurring on the rising limb, the falling limb, or base flow. To determine this categorization, the flow rates preceding and following the sample had to be known. Daily average flow rates, obtained from the USGS web site, were used. Using daily average flow rates can not take into account storms, located close to the gauging station, which produce hydrographs with a base time shorter than a day. If the data point is on the rising limb of the hydrograph then *rise*=1 and *fall*=0. Similarly, if the data point is on the falling limb of the hydrograph then *rise*=0 and *fall*=1. If the flow rate appears steady during the several days around which the sample was taken, then base flow is assumed and *rise* and *fall* both take values of 0.

5.4.6 Use of Interaction Terms

Interaction terms are a function of one variable and at least one other variable. The purpose of using interaction terms is both physical and mathematical. In some instances, two variables may be multiplied together and included in a multivariate regression at the exclusion of the individual variables. For example, to test the hypotheses that sediment load is dependent on the travel time of a watershed, it is appropriate to use a variable with the length divided by the slope.

If interaction terms are not used in a multivariate regression with binary variables, the model may be biased. Consider the model given by

$$\mathit{ln\ conc} = \underset{(5.42)}{2.05} + \underset{(6.26)}{0.43}\,(\mathit{ln\ flow}) + \underset{(2.29)}{0.44}\,(\mathit{warm}) + \underset{(3.27)}{0.61}\,(\mathit{large}) \qquad \textbf{(5-4)}$$

Equation	R^2	DF	SE	F
(5-4)	0.39	71	0.76	14.2

which was developed for the Lavaca River near Edna data. Table 5-7 shows the four categories of data, which can be plotted. The model is an improvement over the bivariate regression, but if the model is plotted as shown in Figure 5-6, it is quickly apparent that the slope is constant for all categories of data. Quick inspection of the measured samples in the same figure show concentration increasing at a higher rate with flow for the "large, warm" samples than for the "small, cold" samples. Because no interaction terms are included in the model, the rate of change of ln concentration with ln flow is required (mathematically) to be the same for all categories of data.

Table 5-7. Categories of Data

	"not" warm 0	**warm 1**
"not" large 0	cold, small 0 , 0	warm, small 1 , 0
large 1	cold, large 0 , 1	warm, large 1 , 1

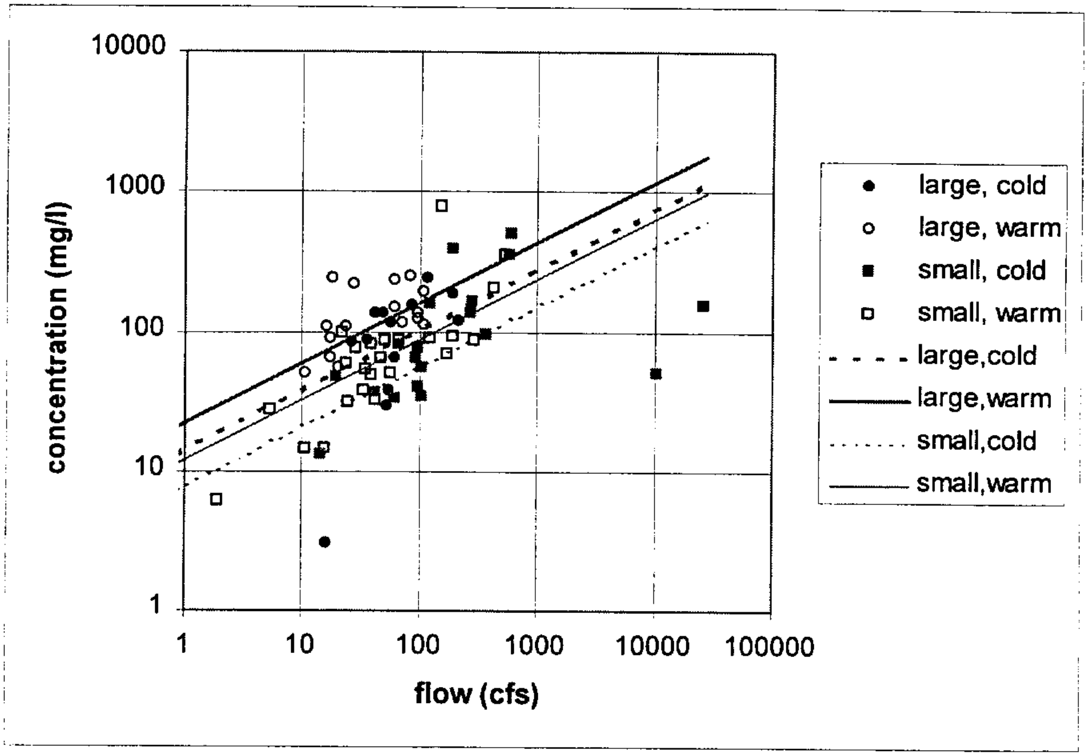


Figure 5-6. Binary Variable Model with no Interaction Terms

Expressions for each line on the graph can be obtained by placing the different values of the binary variables into the regression model for each category of data. The results are summarized in Table 5-8. This model, with no interaction terms, says that temperature and size effects are the same for all ranges of flow rates. Furthermore, flow effects are the same no matter what temperature the water or size of grains of the sample.

Table 5-8. Regression Results (no interaction terms)

Category	Slope	Intercept
cold, small	0.43	2.05
warm, small	0.43	2.49
cold, large	0.43	2.66
warm, large	0.43	3.10

The addition of interaction terms to the model produces the same results as four separate regressions, one for each category of data. The model expressed by

$$\begin{aligned} \mathit{ln\ conc} = \underset{(5.22)}{2.46} + \underset{(3.89)}{0.34}\,(\mathit{ln\ flow}) - \underset{(-0.53)}{0.33}\,(\mathit{warm}) + \underset{(-0.41)}{0.32}\,(\mathit{large}) \\ + \underset{(1.31)}{0.183}\,(\mathit{ln\ flow})(\mathit{warm}) + \underset{(0.34)}{0.066}\,(\mathit{ln\ flow})(\mathit{large}) \end{aligned} \qquad \textbf{(5-5)}$$

Equation	R^2	DF	SE	F
(5-5)	0.41	71	0.76	8.98

includes interaction terms for the Lavaca River near Edna.

Physically, this model states that temperature and size have a different effect for lower flows than for higher flows. An additional interaction term (*warm*)(*large*) would be included to account for size effects varying at different temperatures, not necessarily because the variable (*warm*)(*large*) represents something physically. Although the model expressed in Equation 5-5 has only a slightly better coefficient of determination than that of Equation 5-4 and some of the t-statistics are poor, the model is included to help explain the importance of interaction terms. The model coefficients for the four groups of data are summarized in Table 5-9. The results of Equation 5-5 are plotted in Figure 5-7. Comparison of Tables 5-8 and 5-9 along with Figures 5-6 and 5-7 emphasizes the mathematical importance of interaction terms.

Table 5-9. Regression Results (interaction terms)

Category	Slope	Intercept
cold, small	0.34	2.46
warm, small	0.52	2.13
cold, large	0.41	2.78
warm, large	0.59	2.45

The warm temperature lines have steeper slopes implying that sediment concentration increases more rapidly with flow during the warmer months. Also evident is that the samples that have larger grains have higher concentrations for the same flow rate. Presumably more bed material is contributing to these samples.

Similar to *temperature* and *% fines*, the effect that hydrograph limb has on sediment concentration is different for higher and lower flow rates; thus, interaction terms such as (*ln flow*)(*rise*) should be used rather than just *rise*.

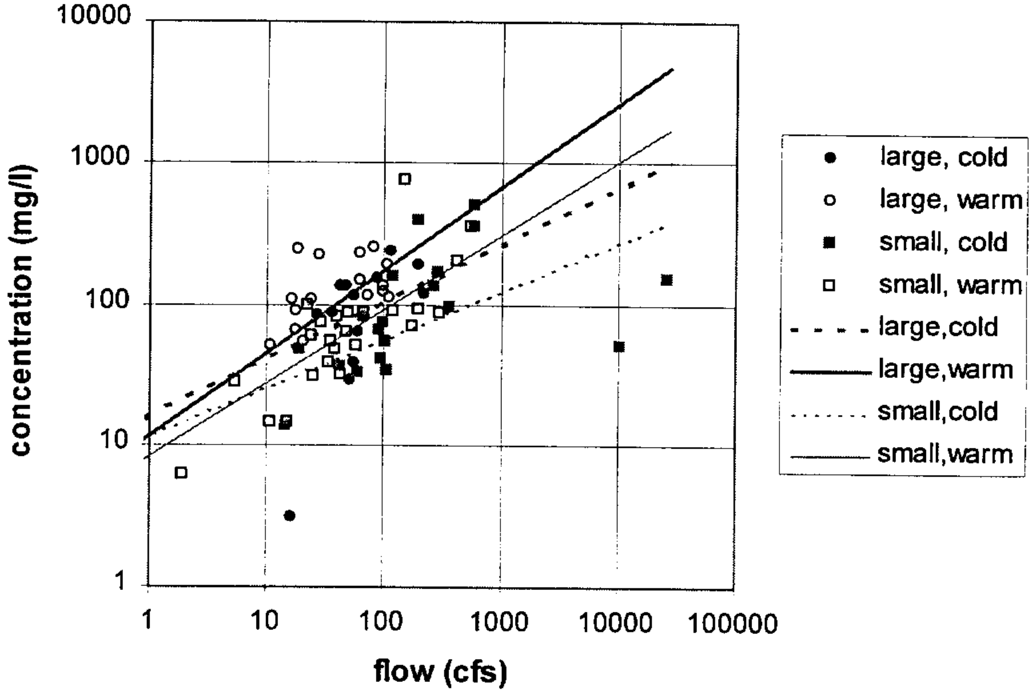


Figure 5-7. Binary Variable Model with Interaction Terms

5.5 Model Selection

After determining the best model using the binary variables *warm* and *large* as first defined, several new models were developed using redefined binary variables *warm* and *large*, and using the continuous variables *temperature* and *% fines*. As expected, using continuous variables improved the original model slightly. Numerous values defining the difference between *warm* and not *warm* and *large* and not *large* were used in the multivariate regressions because the original dividing values were chosen arbitrarily. The end result was to define *warm* as those temperatures that were greater than 17° C. This newly defined binary variable *warm* resulted in a better model than using the continuous variable *temperature*. This result is expected when considering that the variable *temperature* is actually a proxy variable for the time of year. The binary variable *warm* represents two different times of year and is therefore more meaningful than the continuous variable *temperature* which takes into account daily fluctuations of temperature independent of the season.

The model expressed by

$$\mathit{ln\ conc} = \underset{(-0.12)}{-\,0.12} + \underset{(6.05)}{1.65}\,(\mathit{ln\ flow}) - \underset{(-6.95)}{0.135}\,(\mathit{ln\ flow})^2 - \underset{(-3.53)}{0.037}\,(\%\mathit{fines})$$

$$+\,\underset{(5.20)}{3.07}\,(\mathit{warm}) + \underset{(2.62)}{0.0066}\,(\%\mathit{fines})(\mathit{ln\ flow}) - \underset{(-4.54)}{0.55}\,(\mathit{warm})(\mathit{ln\ flow}) \qquad \textbf{(5-6)}$$

Equation	R^2	DF	SE	F
(5-6)	0.66	71	0.57	21.4

is significantly improved over the traditional sediment rating curve using concentration. The coefficient of determination is improved from $R^2 = 0.25$ to $R^2 = 0.66$. Notice the addition of the $(\mathit{ln\ flow})^2$ term which was discussed earlier. A significant amount of scatter still exists in the data. Several new variables were introduced to help reduce the scatter. Variables that could be deduced from the data set included those associated with the hydrograph limb and those further defining the time of year. The model was not successfully improved.

To further analyze the effectiveness of the model expressed by Equation 5-6, a regression of predicted and measured values was conducted resulting in an R^2 of 0.66. Values are shown in Figure 5-8.

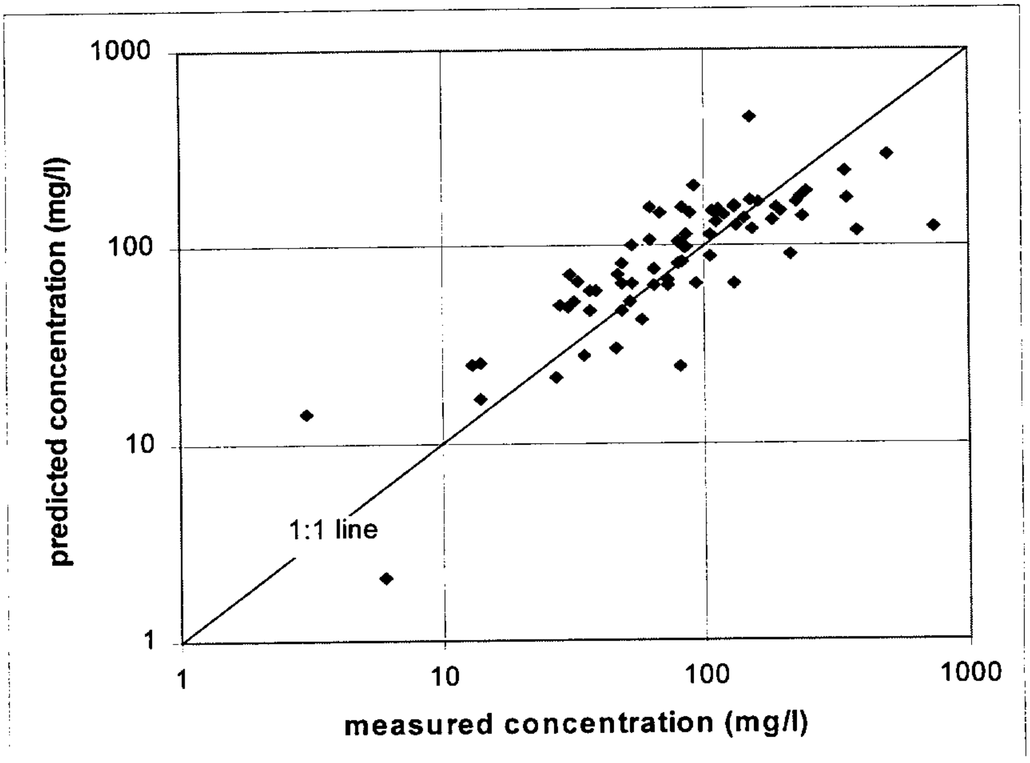


Figure 5-8. Comparison of predicted and measured values for selected model

5.6 Model Evaluation

5.6.1 Warm Effects

By evaluating the selected model expressed by Equation 5-6, the relative effects of the variables can be determined. Evaluation of the terms that contain the variable warm shows that during warm seasons, for most flows, the sediment concentration will be higher than for cold seasons. Rainfall patterns vary seasonally. Warm summer months are more likely to have intense rainstorms. Thunderstorms with intense rainfall cause much more erosion than hours or days of light drizzle. Large amounts of rainfall over short periods of time supply high kinetic energy for high rates of erosion. This seasonal effect becomes less apparent as the flow increases. In fact, when the flow is greater than 260 cfs (*ln flow* = 5.56) the sediment concentration is higher for a cold season. Caution should be used in interpreting this phenomenon of the model. The negative coefficient for the (*ln flow*)(*warm*) term may simply represent more of a limiting effect that season has on sediment concentration rather than a physical parameter causing warm seasons to have lower sediment concentration for high flow rates than cold seasons. On the other hand, during cold seasons, vegetation is sparser resulting in more bare ground subject to erosion. The sparse vegetation explanation is not very likely for a basin such as the Lavaca River where the climate is warm and cover changes little.

5.6.2 %fines Effects

Analysis of the terms containing the *%fines* variable shows that for most flows (less than 272 cfs), the higher sediment concentrations contain the larger particles (smaller percent fines). Larger particles are likely to have originated in the channel bed. When bed material is contributing to the suspended load, the concentration is higher. For flows greater than 272 cfs, higher percentages of fines are found in the greater sediment concentrations. This difference becomes more pronounced the greater the flow. Perhaps at larger flows, the wash load becomes an even larger contributor to the suspended load. The flow rate where the season and size effects change is approximately the same. For the data sample, 88% of the flows were less than 260 cfs.

5.6.3 Flow Effects

In general, sediment concentration increases with flow similar to the relationship expressed by a bivariate regression. However, the model shows that as flow rates get large, the concentrations do not continue to increase at the same rate.

The model limits concentration to 632 mg/l at a flow rate of 7300 cfs. One of the samples was measured at 745 mg/l; however, the next largest measured concentration was 491 mg/l.

5.7 Conclusions

The selected model, expressed by equation 5-6, shows reduced variability in sediment concentration by including the additional variables *warm* and *%fines,* with the appropriate interaction terms. The model can explain an additional 41% of the variability in sediment concentration compared to the simple bivariate regression which is commonly used for a sediment rating curve.

As is apparent from the model presented, it is necessary that correct interaction terms and correct functional form of variables be used in the model to accurately represent the involved physics. Furthermore, if there are omitted variables, the coefficients will be biased. Coefficients for similar models for other gauging stations can readily be determined using a program such as SAS or EXCEL. More accurate predictions of sediment load result in improved knowledge of storage capacity in reservoirs and better estimates of non-point source pollution.

6 Daily Sediment Data Analysis for the Lavaca River near Edna

6.1 Purpose

Analyzing the daily data from one station allows several investigations that are not possible with the annual value analysis of all of the stations. The purpose of analyzing the daily sediment data for one station is to improve the sediment rating curve, evaluate seasonal changes and temporal changes, and to relate sediment load to spatial varied rainfall and watershed characteristics. By better quantifying the sediment concentration in the Lavaca River, studies of the delta formation as well as studies of the water quality and quantity can be improved. Additionally, similar types of analyses can be used in other areas and basins.

6.2 Data Description

The Lavaca River near Edna station has an extensive record with a large variability of storms and there are no major reservoirs upstream of the station to complicate analysis. Furthermore, in 1994, the Environmental Section of the TWDB keyed this station's data prior to 1965 into the computer. Thus, 44 years of daily suspended sediment data for the Lavaca River are available and have been obtained in digital format.

The Lavaca River watershed as defined by TWDB's Plate 1 (1990), and as shown in Figure 2-3, contains all of the Lavaca River, all of the Navidad River, and all of the associated tributaries. The watershed lies between the Colorado River and Guadalupe River watersheds. The Navidad River drains into Lake Texana just prior to discharging into the Lavaca River south of the town of Edna. The Lavaca River drains into Lavaca Bay which is connected to Matagorda Bay and then to the Gulf of Mexico.

Because Edna is located upstream of the confluence of the Lavaca and Navidad Rivers, the watershed upstream of the sampling station comprises only a portion of the watershed shown in Figure 2-3. The extent of the watershed that drains to the gauging station located at the US Highway 59 bridge near Edna is shown in Figure 2-4.

Suspended sediment samples were taken from the Lavaca River near Edna everyday from September, 1945, to September, 1989. The drainage area upstream of the station is 817 square miles. The annual average streamflow for the years of record is 2.29×10^5 acre-feet (316 cfs). The average annual suspended sediment load is 172 tons per square mile (164 mg/l). Daily rainfall was recorded during the same period at three stations in the basin: Edna, Hallettesville, and Yoakum. Table 6-1 summarizes the daily data for flow, sediment, and rainfall. The number of values for *rain* is lower than for *eppt*, *hppt*, and *yppt* because there had to be rainfall data at each location on a given day for the spatially averaged rainfall to be computed.

Table 6-1. Summary of Daily Data for the Lavaca River near Edna Basin

Variable	Number of values	Mean	Standard Deviation	Mini-mum	Maximum	Skew-ness
flow	16101	316	1396	0.04	53000	12.48
sedconc	16058	163	375	0.01	15896	12.09
sedld	16058	386	2925	0.10	243000	47.30
rain	13676	0.10	0.34	0.00	9.98	7.31
eppt	15025	0.11	0.42	0.00	7.34	7.35
hppt	15645	0.10	0.40	0.00	12.50	8.31
yppt	15645	0.10	0.38	0.00	10.42	7.53

flow = daily flowrate (cfs)
sedconc = daily average sediment concentration (mg/l)
sedld = daily total sediment load (tons)
rain = spatially averaged rainfall (in.)
eppt = daily rainfall at Edna (in.)
hppt = daily rainfall at Hallettesville (in.)
yppt = daily rainfall at Yoakum (in.)

6.3 Watershed Description

The watershed description provided herein uses information from an Open-File Report (USGS 1973), Report 268 (Texas Department of Water Resources), and Texas Water Development Board Report 92 (Kunze, 1969).

The Lavaca River basin is in the central part of the gulf Coastal Plain of Texas. The fan-shaped basin, drained by the Lavaca River in the west and the Navidad River in the east, is about 80 miles long and about 55 miles across at its widest point. The basin is bounded on the southeast by the Colorado-Lavaca coastal basin, on the northeast by the Colorado River basin, and on the Northwest by the Guadalupe coastal basin. The drainage area for the entire basin includes all or part of eight counties, is about 2,314 square miles, and approximately 0.9 percent of the area of the State of Texas. The Lavaca River above Edna includes portions of Jackson, Lavaca, De Witt, and Fayette counties. The Lavaca River rises in southern Fayette County at an elevation of about 470 feet and flows south southeastward through Lavaca and Jackson Counties to Lavaca Bay.

The terrain of the northernmost area of the Lavaca River basin is rolling to level and is moderately wooded with hardwood and pecan trees. The drainage pattern in this area is fairly well defined and surface water runs off quickly. In the middle section of the basin, the topography changes to slightly rolling or level prairie covered with native grasses and groves of hardwood. Pecan trees grow profusely along the streams. In the southernmost part of the basin, the terrain becomes a flat, grassy prairie with live oaks, mesquite, and huisache. Because the slope of the streams in this area is very low, surface water runoff is slow.

There are about 66,500 acres (269 km^2) in the flood plains of the Lavaca and Navidad Rivers. About 24,000 acres (97 km^2) have been cleared and are being used for pasture and cultivated crops, including rice. The Texas Blackland Prairie land-resource area is the major source of sediment in the basin. The area occupies the upper 32 percent of the basin and furnishes large quantities of fine sediment to the streams. Sediment damage is extensive in the floodplains of the upper portions of the Lavaca and Navidad Rivers and their tributaries. The Intracoastal Waterway is not suffering any sediment damage by the Lavaca River since most of the sediment entering Lavaca Bay is deposited immediately. The delta built by the Lavaca River in Lavaca Bay covers more than 1,000 acres (4.05 km^2).

Oil is produced in the central and southern parts of the basin, and irrigation is practiced extensively in the southern half. Surface streams are probably degraded from time to time by oil-field brine and by return flow from irrigation. Municipal wastes may also affect water quality in some streams during extreme low flow. Low flow in some streams in the basin may be maintained for indefinite periods by return flow from irrigation, local wastewater, seepage from bank storage, and seepage into streams that have cut below the water table. However, there are periods of no flow. Most of the flow in streams in the basin is surface runoff, which is dependent on the quantity and intensity of local precipitation. Generally, the Lavaca River basin has an abundant supply of surface water of very good quality.

The surface streams probably obtain most of their chemical characteristics from the geologic formations that crop out within the basin. The exposed rocks range in age from Miocene to Holocene and crop out in bands nearly parallel to the coast. The main constituents are limy clay, clay, sandstone, and limy sand. Both the Lavaca River and its principal tributary, the Navidad River, traverse all of the formations; therefore, the streams contain chemical constituents dissolved from each formation.

In 1960, Yoakum (population of 5,761) and Edna (population of 5,038) were the only two communities with a population of more than 5,000. In 1990, Yoakum (5,611) and Edna (5,343) remained the only two communities with more than 5,000. The basin of the Lavaca River above Edna has no major reservoirs. Lake Texana is created on the Navidad River by the Palmetto Bend Reservoir.

The economy of the Lavaca River basin is based chiefly on agriculture and livestock. Corn and cotton are major crops in the northern half of the basin, and rice, cotton, truck produce, and grain sorghum are the major crops in the southern half. Oil production and oil field supply are the major non-agrarian sources of income. The greatest concentration of oil fields is in the central and southern parts of the basin. Natural gas and other minerals also contribute to the local economy.

The basin receives an average of about 38 inches of rainfall per year, of which about 5 inches enter Lavaca Bay as runoff. Moderate summers and mild winters are characteristic of the area. Rainfall is fairly evenly distributed throughout the year. In the northern half of the basin, average monthly rainfall is usually at a peak in May and again in September. In the southern half of the basin, the rainfall generally peaks during the summer season. However, rainfall throughout the basin is subject to much greater variations than indicated by the annual and monthly averages. For example, during the 1931-65 period, precipitation at Hallettesville ranged from a low of 0.00 inches in October, 1934 to a high of 24.68 inches in July, 1936. Similarly, precipitation at Edna ranged from a low of 0.00 inches during several months to a high of 14.38 inches in June 1960. Precipitation so unevenly distributed in time does not sustain streamflow; therefore, flow in most tributaries in the basin is intermittent, and periods of no flow have occurred in both the Lavaca and Navidad Rivers.

Runoff, like rainfall, in the Lavaca River basin is highly variable. Discharge of the Lavaca River near Edna has ranged from no flow to 73,000 cfs. The magnitude and frequency of high and low flows can best be shown by Figure 6-1. A curve with a steep slope throughout indicates a highly variable stream flow is largely from direct runoff, whereas a curve with a flat slope shows surface or groundwater storage. The steep slope of the curve further supports the fact that flow in the streams of the Lavaca River mostly comes from surface runoff (Kunze, 1969).

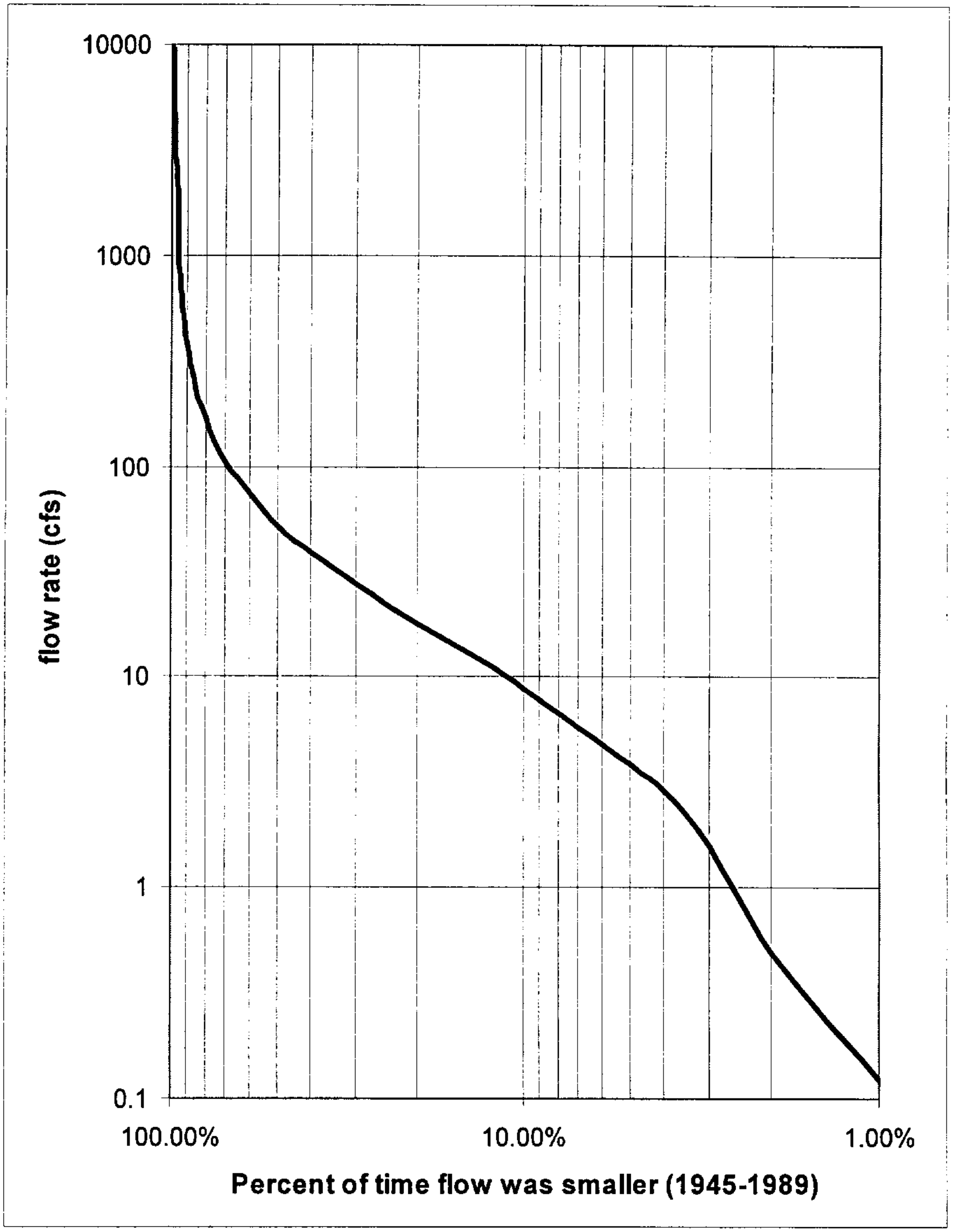


Figure 6-1. Frequency of Daily Flows, Lavaca River Near Edna

6.4 *Bivariate Analysis - "The Sediment Rating Curve"*

Sediment rating curves are often developed with annual averaged or monthly averaged values for sediment load and flow rate. Annual and monthly averages dampen the effect of extreme events. Table 6-2 compares the mean and variance for daily values and annual means. The slight discrepancies in the means are due to the beginning and ending years of sampling which have incomplete records and were not included in the annual means. Using annual averages of daily data results in more than 95% reduction in the variance of all three variables.

Table 6-2. Comparison of daily values and annual averages

	daily values		annual averages	
	mean	variance	mean	variance
flow (cfs)	316	1,949,162	322	51,697
conc. (mg/l)	163	140,404	163	6,003
load (tons)	386	8,557,071	400	115,571

The distinction between annual values (totals), annual average values, and annual means can be confusing. An annual average for year j is computed using

$\overline{Q_j} = \frac{1}{365}\sum_{i=1}^{365} Q_i$ where Q_i is the daily value and $\sum_{i=1}^{365} Q_i$ is the annual value or total.

The long term average over N years of daily values is $\overline{Q_N} = \frac{1}{365 \times N}\sum_{i=1}^{365\times N} Q_i$. The mean of the of the annual averages is computed using

$\overline{\overline{Q}} = \frac{1}{N}\sum_{j=1}^{N}\overline{Q_j} = \frac{1}{N}\left(\sum_{j=1}^{N}\frac{1}{365}\sum_{i=1}^{365}Q_{i,j}\right) = \frac{1}{365\times N}\sum_{i=1}^{365\times N}Q_{i,j}$, yielding the same results as the mean of the daily values. The average of the annual values is computed using $\overline{Q}_{annual\ total} = \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{i=1}^{365}Q_i\right) = \frac{1}{N}\sum_{i=1}^{365\times N}Q_{i,j} = 365\times\overline{\overline{Q}}$. The concentration is simply the load divided by the flow, $C = \frac{L}{Q}$. However, the average concentration for year j, $\overline{C_j} = \sum_{i=1}^{365}\frac{L_i}{Q_i} \neq \frac{\overline{L_j}}{\overline{Q_j}}$ is considerably smaller than the average load divided by the average flow.

Data from the year 1950 will be used as an example. There are 365 daily values of flow in 1950. All 365 values are used in the daily analysis. The total volume of flow for 1950 is 19736 cfs while the average daily flow for the year is 54 cfs. Similarly, the total load for 1950 is 47815 tons while the average daily load is 131 tons. The concentration is not as straightforward. The daily concentration is computed using daily load and daily flow. Taking the average of these 365 daily concentration values results in 197 mg/l. The annual concentration is computed using the annual load and annual flow. The resulting annual concentration value for 1950 is 816 mg/l. This value might be thought of as an annual "flow-weighted" concentration. Analyses in Chapter 4 use annual values (totals) and the average of all of the annual (total) values. The current analysis uses annual average values (based on daily data) and daily values.

Using the annual averages results in using 43 data points to compute relationships of sediment load with other variables. In comparison, using the daily values results in using over 16,000 data points. Figure 6-2 displays all of the daily data points. Although the data follow a positive trend, there is a large spread. The values for sediment load are recorded as integers; thus, the straight lines of data for the low values of sediment load are seen in the figure.

The straight lines caused by the integer values can also be seen in the concentration values shown in Figure 6-3. The concentration trend is positive for daily flowrates less than about 1000 cfs. When the flowrate is on the order of 1000 cfs and the maximum concentration is on the order of 1000 mg/l, the trend actually begins a downward decline. This reverse in trend indicates that the river has a maximum sediment concentration. On average, a much higher than average flow is required to obtain the maximum concentration; however as the flow increases even more, the concentration is diluted. In other words, the volume of sediment in the river is supply limited. This supply limit is also shown in the load trend where the maximum load is approximately 10,000 tons/day.

A flow rate of 1000 cfs is more than three times the mean daily flow for the Lavaca River near Edna. According to the frequency analysis for 1945-1989, flows equal or exceed 1000 cfs only five percent of the time. These rarely-occurring events cause the depth of flow to increase and reach areas of the channel banks that are much wider and more vegetated than those areas of the channel that typically have continual flow. The larger cross section and increased roughness of the river channel during large storms result in slower velocities, thus decreased sediment capacity.

Regression equations were developed to quantify the relationships shown in Figures 6-2 and 6-3. The regression equations using the SAS models are shown in Table 6-3. The "*ln*" is included in variable names if the natural logarithm of the value is used. *Lnconc* is the natural log of daily concentration of sediment in mg/l. *Lnflow* is the natural log of daily flowrate in cfs. *Lnflowsq* is equal to the square of *lnflow*. *Lnload* is the natural log of daily load in tons.

Table 6-3 is similar to tables presented in Chapters 4 and 5. The degrees of freedom (DF), plus the number of independent variables used in the regression equation, indicate the number of data points in the analysis. Equations (1), (2), (3), and (4) with 16,057 degrees of freedom use daily values, while equation (5), (6), (7), and (8) with 42 degrees of freedom use annual averages (of daily values). Note that for the annual mean equations, the dependent variables *lnload* and *lnconc* do not have the same relationship as in the other sets of equations because $\overline{C_j} = \sum_{i=1}^{365} \frac{L_i}{Q_i} \neq \frac{\overline{L_j}}{\overline{Q_j}}$. Equations (9), (10), (11), and (12) use annual (total) values like the analyses in Chapter 4.

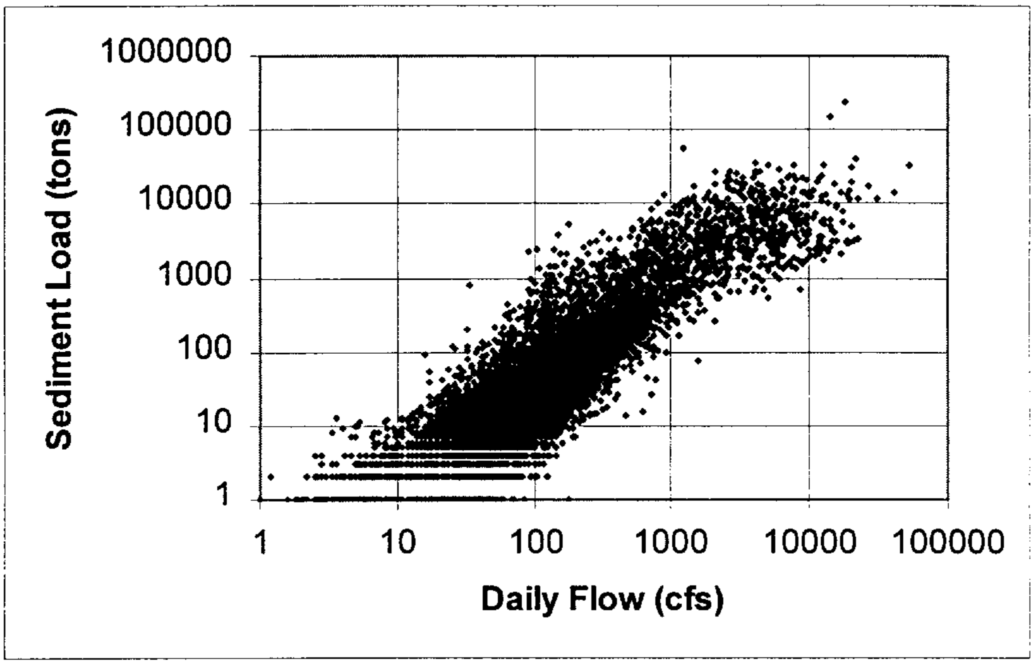


Figure 6-2. Sediment Load, Lavaca River near Edna, 1945-1989

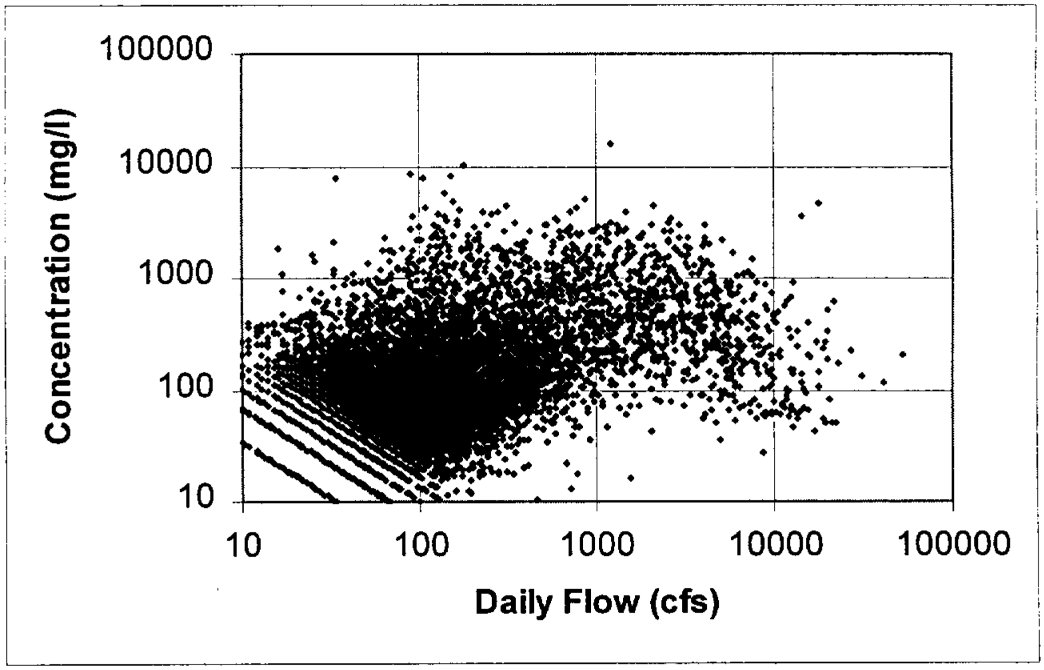


Figure 6-3. Sediment Concentration, Lavaca River near Edna, 1945-1989

Table 6-3. Regression equations and statistics for the Lavaca River near Edna

	Equation *(t-statistics)*	R^2	DF	SE	F
Daily values	(1) *lnconc* = 3.12 + 0.285*(lnflow)* *(135) (54)*	0.15	16057	1.17	2890
	(2) *lnconc* = 3.79 - 0.168*(lnflow)* + 0.0598*(lnflowsq)* *(141) (-14) (43)*	0.24	16057	1.11	2536
	(3) *lnload* = -2.70 + 1.285*(lnflow)* *(-117) (242)*	0.79	16057	1.17	58714
	(4) *lnload* = -2.03 + 0.832*(lnflow)* + 0.0598*(lnflowsq)* *(-75) (71) (43)*	0.81	16057	1.11	33660
Annual means	(5) *lnconc* = 4.64 + 0.068*(lnflow)* *(14) (1.1)*	0.03	42	0.42	1.2
	(6) *lnconc* = 6.35 - 0.70*(lnflow)* + 0.080*(lnflowsq)* *(6.5) (-1.68) (1.9)*	0.11	42	0.40	2.4
	(7) *lnload* = 1.70 + 0.73*(lnflow)* *(3.7) (8.8)*	0.65	42	0.55	77
	(8) *lnload* = 2.25 + 0.48*(lnflow)* + 0.026*(lnflowsq)* *(1.7) (0.83) (0.43)*	0.65	42	0.56	38
Annual values	(9) *lnconc* = 8.91 - 0.25*(lnflow)* *(9.2) (-3.0)*	0.18	42	0.57	9.0
	(10) *lnconc* = 8.46 - 0.17*(lnflow)* – 0.0040*(lnflowsq)* *(1.21) (-0.13) (-0.07)*	0.18	42	0.57	4.4
	(11) *lnload* = 3.09 + 0.75*(lnflow)* *(3.2) (8.8)*	0.65	42	0.57	77
	(12) *lnload* = 2.64 + 0.83*(lnflow)* - 0.0040*(lnflowsq)* *(0.4) (0.64) (-0.07)*	0.65	42	0.57	37

R^2 = coefficient of determination
SE = standard error
DF = degrees of freedom
F = F-value

As previously discussed, the load is the product of flow and concentration necessitating the correlation between load and flow to be better than the correlation between concentration and flow. The coefficient of determination is considerably higher for the equations developed with daily values than for the annual mean equations. This higher correlation is probably due to the amount of data. There are so many data points in the daily data set that an extreme value is less likely to have undue influence in the model. Such is not the case for the annual means.

To reiterate the difference between the daily values, the annual means (based on daily values), and annual values, the ranges of the values are shown in Table 6-4. It is imperative that anyone employing such equations understand

Table 6-4. Ranges of variables for selected data sets

Data set		variable		
		sedld (tons)	*flow* (cfs)	*sedconc* (mg/l)
Daily values	average	386	316	163
	maximum	243,000	53,000	0.01
	minimum	0.10	0.04	15896
Annual means (based on daily values)	average	400	322	164
	maximum	1542	1028	431
	minimum	14	8	56
		annload (tons)	*annflow* (cfs)	*sedconc* (mg/l)
Annual values (totals)	average	146,512	117,756	499
	maximum	611,558	375,066	1267
	minimum	5123	2,906	108

whether daily values or annual values should be used. Selected equations from Table 6-3 are shown graphically in Figures 6-4 and 6-5 for the appropriate range of data.

The sediment rating curve developed using over 16,000 data points is almost identical to the sediment rating curve developed in Chapter 5 using 72 data points. The USGS spot sampling program for the Lavaca near Edna collected enough high flows that the sporadic samples are adequate to produce a sediment rating curve. The equations developed using annual means overestimate the load compared with the equations developed using daily values. If one extrapolates from the annual means equations for large flows then the sediment load will be underestimated. The equations for annual values (totals) cover a different range of data.

As expected from viewing Figures 6-2 and 6-3, the daily regression models improve when including the lnflowsq term. However, the regression model does not exhibit the same trend that the human eye detects. For the model to produce a concave down curve, the sign of the coefficient for *lnflow* should be positive while the sign of the coefficient for *lnflowsq* should be of a smaller magnitude but negative. Opposite signs are shown in equations (2) and (4) of Table 6-3. Figures 6-2 and 6-3 show over 16,000 data points. There are large areas in both figures where the individual points cannot be distinguished; thus the density of those points cannot be used when one visually tries to identify a trend. The points with flow

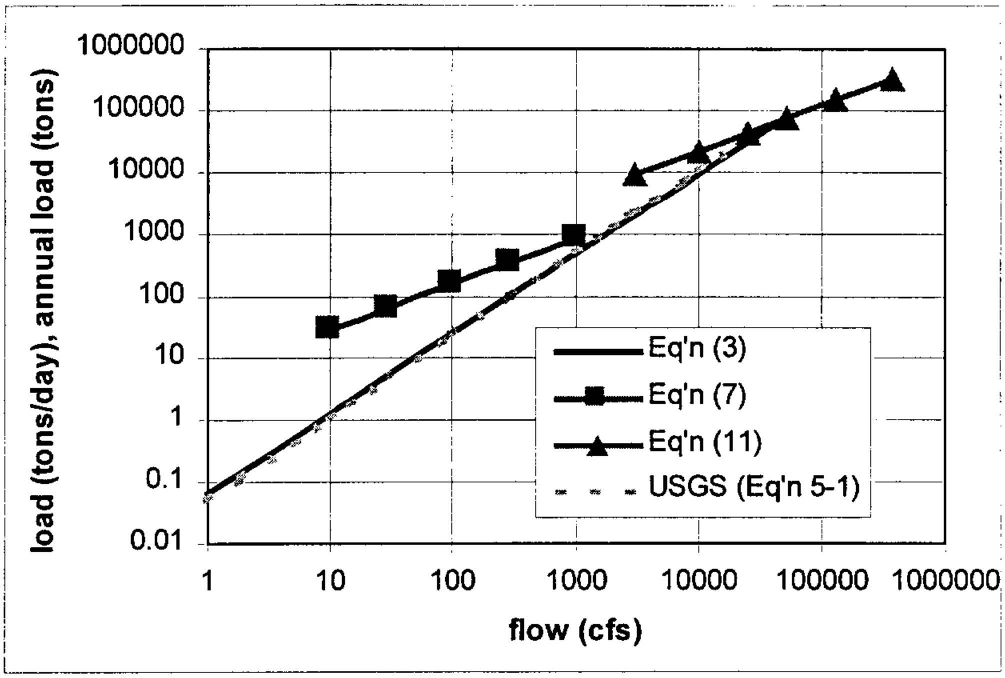


Figure 6-4. Comparison of load equations shown in Table 6-3

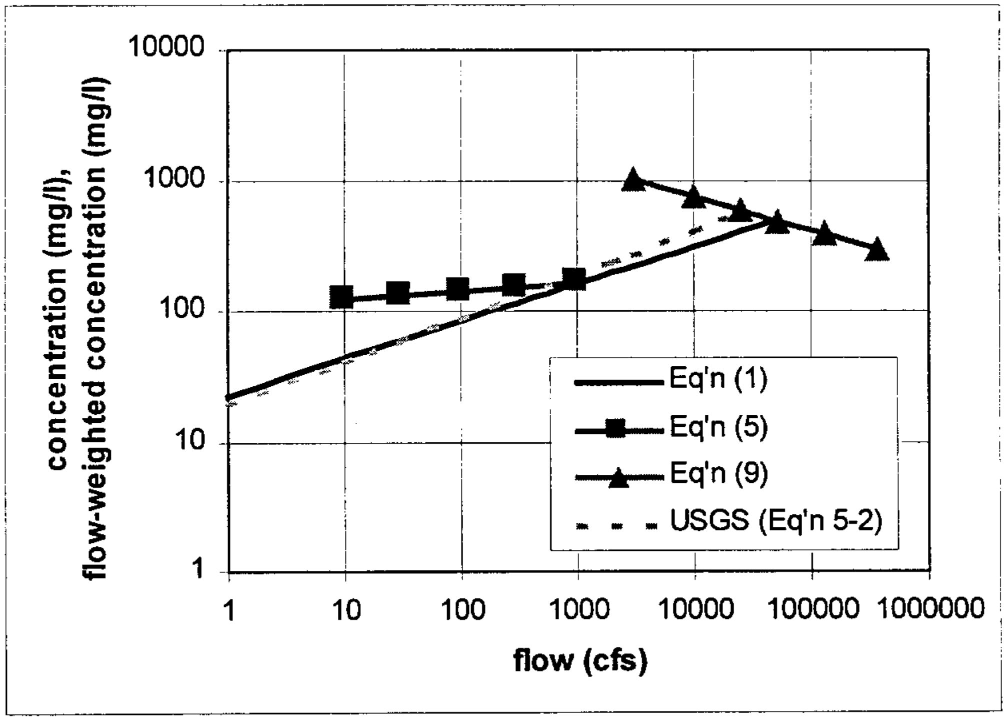


Figure 6-5. Comparison of concentration equations shown in Table 6-3

exceeding 1000 cfs represent only five percent of the total data. The regression equations do not adequately represent the effect of large flows.

Table 6-5 lists the five highest values for flow (*flow*), concentration (*sedconc*), and load (*sedld*). The values are listed in chronological order for each variable. Note that the highest flow does not result in the highest load, etc. The sum of the five highest loads (530,760 tons) is 8.6% of the total load (6,194,837 tons) but only 0.03% of the total observations. Thus, the reliability of overall predictions of sediment load depends heavily upon how accurately the high loads can be predicted. The amount of total sediment that is carried by the large storms is quickly evident from inspection of Figure 6-6. To prepare this figure, the flows are rank ordered by magnitude from the smallest to the largest. The loads are summed from the smallest flow rate to each ranked flow rate to give cumulative amounts of sediment load. Each cumulative amount of sediment load is divided by the total sediment load to give a fraction of the total sediment carried by flows smaller than or equal to each ranked flow rate. Days where the Lavaca River carries the average flow rate of 316 cfs essentially do not contribute to the total load over the 45 year

Table 6-5. Highest values and dates for flow, concentration, and load

flow (cfs)	*sedconc* (mg/l)	*sedld* (tons)
22,800 (10/20/60)	7,884 (09/16/55)	40,023 (10/17/57)
27,000 (04/18/73)	10,243 (03/20/57)	36,180 (04/10/59)
53,000 (06/14/73)	8,297 (03/15/61)	57,117 (02/05/65)
41,200 (06/15/73)	15,896 (02/05/65)	243,000 (06/13/81)
30,800 (09/02/81)	8,628 (09/19/70)	154,440 (06/14/81)

period. Figure 6-7 combines the results of Figures 6-1 and 6-6. It is quickly apparent that 10 percent of the flows carry about 90 percent of the load.

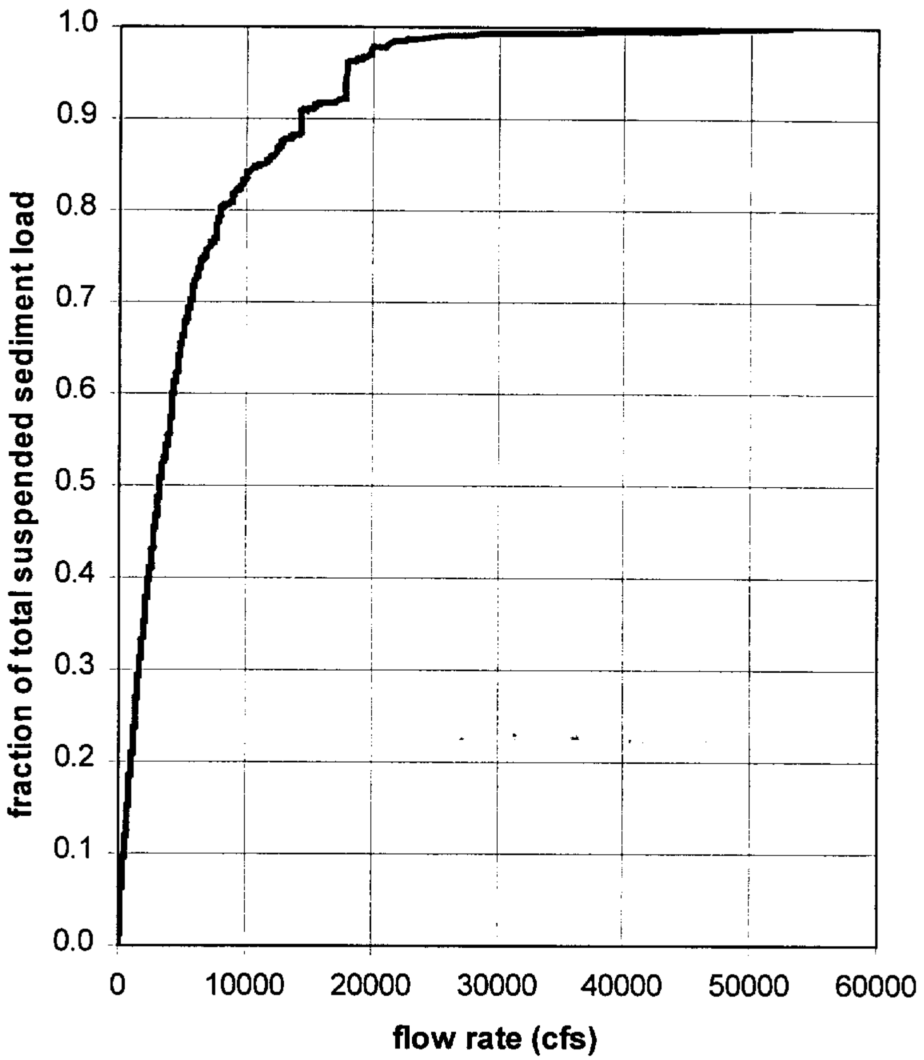


Figure 6-6. Relation of total suspended load to flow rate

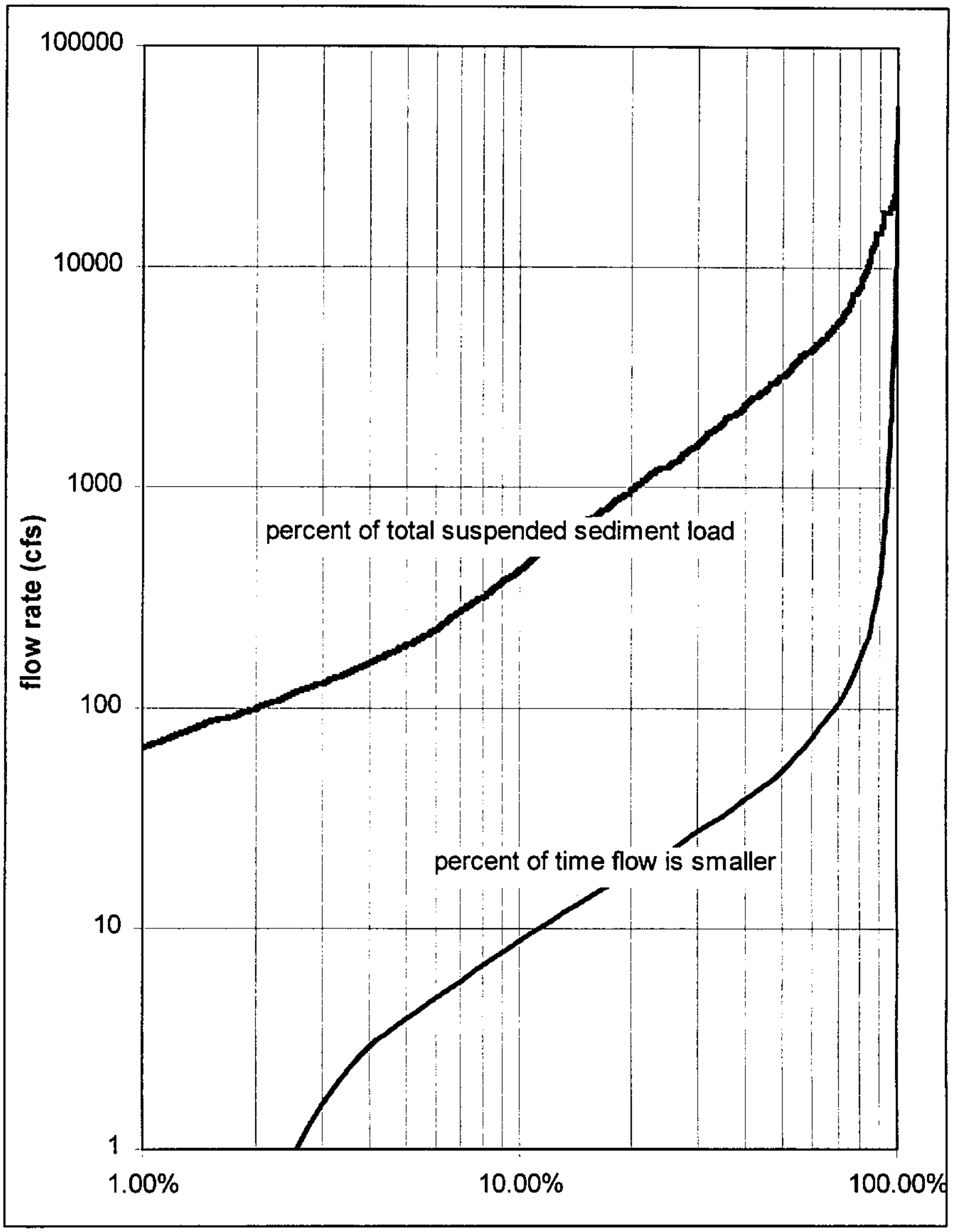


Figure 6-7. Frequency of flows with fraction of total suspended sediment load

The average values of flow, load, and concentration were used as a guide to set threshold values. If data that is less than 10% of the mean is not included, then 14595 data points result. If data that is less than the mean is not included then 3644 data points result. Finally, data that is more than 50% greater than the mean results in a data set of 2343 points. In each case, the correlation between load and flow is increased but the correlation between concentration and flow is not increased. As an alternative, six flow categories were determined. The categories and bivariate regression models are presented in Table 6-6. In every case, except for the lowest flows, the concentration-flow correlation is very poor and thus not presented. The load-flow correlations are, in general, poor; however this poor correlation is not surprising since the variation for the dependent variable is small when considering a small range of flows. In spite of the poor correlations, the models have significant variables. For the lower and higher flows, the coefficient for *lnflow* is smaller than for the flows between 10 and 1000 cfs. These individual lines better represent the trend that is seen when one inspects the scatter plot of the data.

Table 6-6. Regression equations by flow intervals

flow "category"	Equation (*t-statistics*)	R^2	DF	SE	F
$1 \le \mathit{flow} < 10$	$\mathit{lnconc} = 5.68 + -0.72(\mathit{lnflow})$ *(62) (-14)*	0.24	655	0.56	208
$1 \le \mathit{flow} < 10$	$\mathit{lnload} = -0.14 + 0.28(\mathit{lnflow})$ *(-1.53) (5.6)*	0.05	655	0.56	31
$10 \le \mathit{flow} < 100$	$\mathit{lnload} = -2.82 + 1.29(\mathit{lnflow})$ *(-51) (85)*	0.45	8962	0.87	7250
$100 \le \mathit{flow} < 500$	$\mathit{lnload} = -4.28 + 1.62(\mathit{lnflow})$ *(-22) (43)*	0.34	3742	1.00	1890
$500 \le \mathit{flow} < 1000$	$\mathit{lnload} = -5.37 + 1.81(\mathit{lnflow})$ *(-3.5) (7.7)*	0.11	484	1.02	59
$1000 \le \mathit{flow} < 5000$	$\mathit{lnload} = 0.80 + 0.94(\mathit{lnflow})$ *(1.3) (12)*	0.19	601	0.91	143
$\mathit{flow} \ge 5000$	$\mathit{lnload} = 4.44 + 0.46(\mathit{lnflow})$ *(3.7) (3.5)*	0.65	42	0.56	38

To better express the trend shown in the scatter plot, a sine function can be used as shown in Table 6-7. The degrees of freedom are reduced from earlier models because flows less than 1 cfs were not considered. The coefficient of determination is the same for load-flow relationship using the sine function model and the simple bivariate model. The coefficient of determination for the concentration-flow relationship is doubled using the sine function instead of the simple bivariate model. In both cases, the standard error is considerably reduced.

Table 6-7. Regression using sine function

Equation *(t-statistics)*	R^2	DF	SE	F
$lnload = 4.63 - 3.59(sin(2\pi lnflow/10.8))$ *(413) (-233)*	0.79	14677	1.00	54383
$lnconc = 4.95 - 1.11(sin(2\pi lnflow/10.8))$ *(488) (-79)*	0.30	14677	0.91	6285

Figure 6-8 shows plots of equations from Tables 6-3, 6-6, and 6-7. The sine function equation may express the trend the best for flows of about 10 to 5000 cfs. To predict the sediment load of extreme flows, the equations in Table 6-6 should

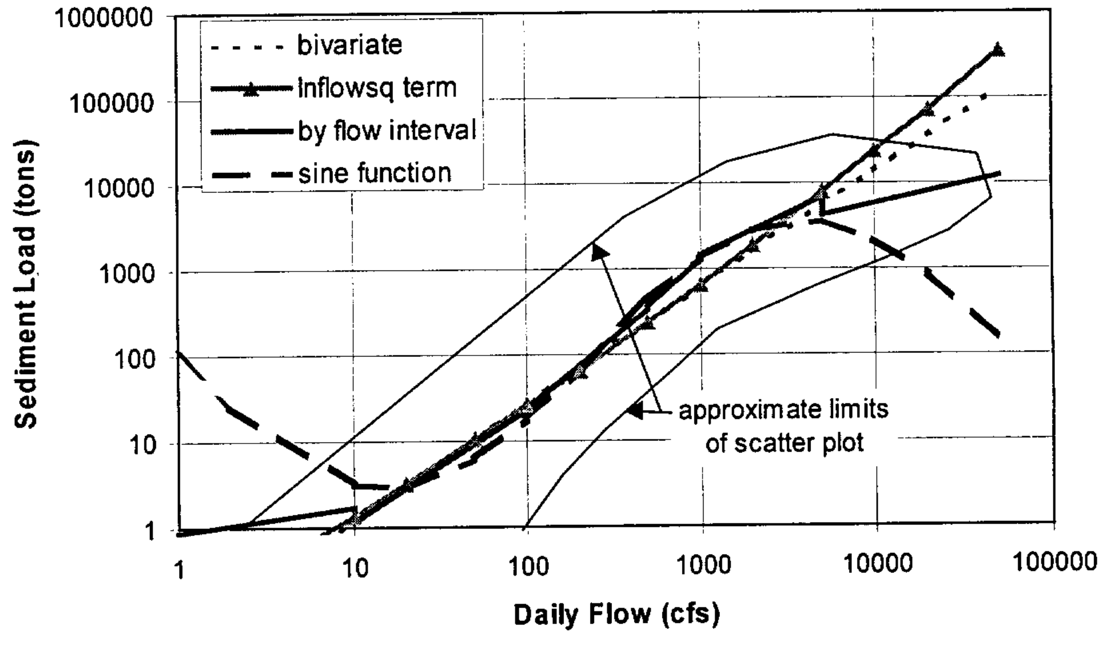


Figure 6-8. Comparison of equations expressing load-flow trend

be used and compared to assuming a maximum concentration of 1000 mg/l. The simplest bivariate model will over predict the load for extreme events causing unnecessary conservatism.

6.5 *Multivariate Analysis*

Bivariate analyses of water quality parameters are not uncommon. Both flow and constituents carried by the flow are dependent upon many of the same physical parameters, i.e. watershed characteristics and rainfall. Because the rainfall combined with the watershed characteristics is responsible for both runoff and soil loss, it is reasonable to believe that the sediment load in the river correlates with the rainfall in the watershed.

6.5.1 Rainfall Considerations

To determine the spatially averaged rainfall, three rainfall gauging stations were used: Edna, Hallettesville, and Yoakum. Arc/Info was used to create Theissen polygons so that the rainfall could be spatially distributed. In some cases, there were days that did not have rainfall at all three stations. In the SAS models, if a day had a missing value for rainfall at one location, then the spatially averaged rainfall was not computed. For purposes of plotting and visually analyzing trends, the missing values were replaced. If rainfall was available at only one station, then that was the assumed rainfall over the entire basin. If rainfall was available at two of the three stations, then the average of the two stations was assumed to be the rainfall at the third gage.

Similar to Table 6-5, Table 6-8 lists the five highest values for rainfall at the three gauging locations as well as the spatially averaged rainfall. Although, the flow in the river is high when there is a lot of rain, the highest rainfall dates are not always the same as the highest flow dates, indicating that there are days when the three gages do not adequately detect rainfall over the basin. Also evident from the table is the spatial variability of rainfall. On June 26, 1973, two more inches of rain fell at Hallettesville than Yoakum, and less than 0.3 inches of rainfall was detected at Edna. Tables 6-5 and 6-8 also help identify some of the more extreme years. As mentioned, 1973 was a wet year, as were 1960 and 1981.

Table 6-8. Highest values and dates for daily rainfall

eppt (in)	*hppt* (in)	*yppt* (in)	*rain* (in)
7.25 (06/25/60)	6.75 (09/12/61)	5.65 (09/13/51)	4.79 (08/29/46)
7.00 (09/11/61)	6.75 (11/04/65)	6.34 (08/04/71)	5.10 (09/21/67)
7.34 (05/07/72)	7.05 (09/21/67)	6.68 (05/07/72)	5.43 (05/07/72)
6.98 (11/19/82)	12.5 (06/26/73)	10.42 (06/26/73)	9.98 (06/26/73)
6.79 (04/11/85)	6.40 (08/31/81)	6.25 (08/31/81)	5.56 (08/31/81)

Table 6-9 shows the correlation coefficients for flow, load concentration, and rain. *Lnrain* correlates better with *lnload* than *lnflow*. *Lnflow* is more correlated with *lnhppt* than any other rainfall variable. *Lnload* is more correlated with *lnhppt* than any other rainfall variable but the difference is not as pronounced as for *lnflow*. *Lnconc* correlates more with *lneppt* than any other rainfall variable.

Lnrain correlated the most with *lnhppt* because Hallettesville has the largest area of the watershed. When the rainfall variables are included in the models, the models are almost unaffected as far as improved correlation; however, the standard error of the model is reduced. The rainfall at Edna proved to be as effective at improving the model as any of the other rainfall variables. Table 6-10 shows regression equations for *lnconc* (not *lnload*) that include the rainfall that fell at Edna.

Table 6-9. Correlation of flow, concentration, and load

	lnflow	*lnload*	*lnconc*	*lnrain*	*lneppt*	*lnhppt*	*lnyppt*
lnflow	1.00	0.89	0.39	0.17	0.16	0.19	0.16
lnload	0.89	1.00	0.77	0.19	0.19	0.19	0.18
lnconc	0.39	0.77	1.00	0.15	0.17	0.14	0.14
lnrain	0.17	0.19	0.15	1.00	0.69	0.89	0.87
lneppt	0.16	0.19	0.17	0.69	1.00	0.44	0.38
lnhppt	0.19	0.19	0.14	0.89	0.44	1.00	0.69
lnyppt	0.16	0.18	0.14	0.87	0.38	0.69	1.00

Table 6-10. Regression equations and statistics with rainfall at Edna

Equation *(t-statistics)*	R^2	DF	SE	F
(1) $\mathit{lnconc} = 2.88 + 0.357(\mathit{lnflow})$ *(106)* *(60)*	0.22	12405	0.96	3555
(2) $\mathit{lnconc} = 3.62 -0.0223(\mathit{lnflow}) + 0.0338(\mathit{lnflowsq})$ *(59)* *(0.87)* *(13)*	0.23	12405	0.95	1893
(3) $\mathit{lnconc} = 2.91 +0.35\ (\mathit{lnflow}) + 0.18(\mathit{eppt})$ *(107)* *(56)* *(9)*	0.23	12405	0.96	1829
(4) $\mathit{lnconc} = 3.60 +0.035(\mathit{lnflow}) + 0.0315(\mathit{lnflowsq}) +$ *(58)* *(1.4)* *(12)* $0.15\ (\mathit{eppt})$ *(7)*	0.24	12405	0.95	1286

A first glance at the models including the *eppt* variable shows a reduced standard error in comparison with those models presented in Table 6-3. The degrees of freedom are also reduced because rainfall data did not exist at Edna everyday. Further investigation shows that it is the omission of some data that reduces the standard error rather than the effect of an additional variable. The bivariate model, which excludes data points when rainfall was not recorded at Edna, has a reduced standard error (0.96) from the bivariate model that includes all of the data (SE=1.17). Although the analysis shows that rainfall does not adequately account for the scatter in the load-flow relationship, the positive coefficient of *eppt* confirms that rainfall near the gauging station will cause increased sediment load.

When transforming the rainfall variables using the natural logarithm, the data set is greatly reduced because of the number of days that have a zero value for rainfall. The omission of data points results in improved correlation between concentration and flow, but the rainfall variables do not improve the models.

Variables were also formed that considered the proximity and travel time from each rainfall gauge to the streamflow gauge. The variables considered various combinations of hydraulic length and slope. The additions of such variables to the model were unsuccessful in explaining any additional variability in sediment load and sediment concentration.

6.5.2 Seasonal Considerations

The analysis of the periodic samples (Chapter 5), that included the temperature of each sample, showed that the time of year was an important factor in determining sediment yield. To do a similar analysis with the daily data, temperature data are required. Average daily temperature data from November, 1977 to September, 1981 were obtained from the National Weather Service. Because there were no data on some days, 698 data points resulted. A periodic function was fit to the data to determine an average temperature for each day of the year. The data and the function are shown in Figure 6-9.

The day of the year was determined using the date functions in Excel. The serial number of each date was computed, then the serial number from the last day of the previous year was subtracted. No correction was made for leap year. The

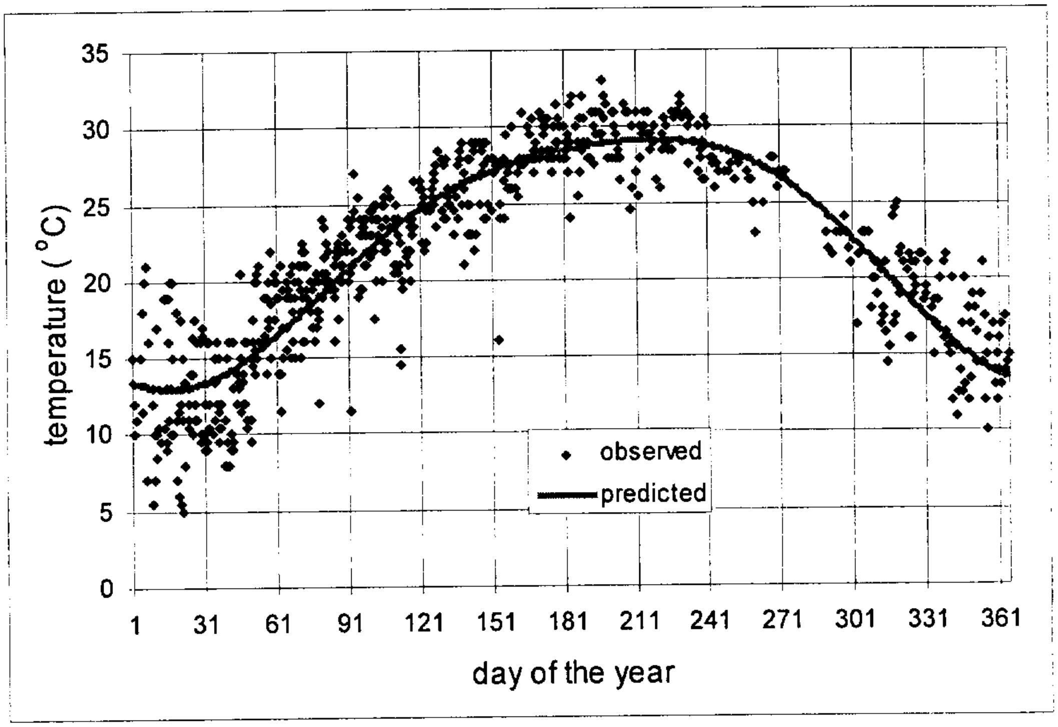


Figure 6-9. Annual temperature trend

predicted temperature for any day for the year can be approximated by

$$\begin{aligned} Temperature\ (^{o}C) &= 22.3 - 2.76\ \sin(2\pi d) - 7.52\ \cos(2\pi d) \\ &\quad - 0.48\ \sin(4\pi d) - 1.30\ \cos(4\pi d) \end{aligned} \qquad \textbf{(6-1)}$$

where d = (day of the year)/365. The equation results in an $R^2 = 0.83$ and a standard error of 2.63 °C.

The periodic function was included in the SAS model to predict temperature for each day of the year. Temperature could then be included as a variable (*temp*) in the model. The inclusion of the variable temp results in

$$\mathit{lnconc} = \underset{(70)}{3.14} - \underset{(-13)}{0.156}\mathit{(lnflow)} + \underset{(43)}{0.0591}\mathit{(lnflowsq)} + \underset{(18)}{0.0277}\mathit{(temp)} \qquad \textbf{(6-2)}$$

Equation	R^2	DF	SE	F
(6-2)	0.26	16057	1.10	1837

Equation 6-2 shows a slight improvement in the model and confirms that the concentration is higher during warmer times of year.

As discussed in Chapter 5, the average daily temperature is representative of the time of year. The time of year indicates the type of cover over the watershed as well as indicating the nature of rainfall storms. The vegetation cover is the greatest during the time of year with the highest temperature. More vegetation reduces erosion. The highest temperatures will also be when the most intense rainstorms fall.

Summer and winter variables were defined to represent the seasonal variation. Selecting months that are classified as summer or winter is a somewhat arbitrary process. Figure 6-9 can assist in the process. If the sediment data was collected in January, February, or March (the three months with temperatures less than 15 °C) then the binary variable *winter* was assigned a value of 1. If the sediment data was collected in May through September (the five months with temperatures greater than 25 °C), then the binary variable *summer* was assigned a value of 1. Furthermore, to account for different seasonal effects during high flows versus low flows, interaction terms were formed. *Winflow* is the interaction term formed by multiplying *winter* by *flow*. Likewise, *sumflow* is formed by multiplying *summer* by *flow*. The interaction terms were formed with *flow* instead of *lnflow* simply because the correlation proved to be better when using the former. The resulting model

$$\mathit{lnconc} = \underset{(127)}{3.67} - \underset{(-22)}{0.256}(\mathit{lnflow}) + \underset{(51)}{0.0811}(\mathit{lnflowsq}) - \underset{(-3.4)}{0.0781}(\mathit{winter})$$

$$- \underset{(-11)}{0.000253}(\mathit{winflow}) + \underset{(17)}{0.342}(\mathit{summer}) - \underset{(-27)}{0.000249}(\mathit{sumflow}) \qquad \textbf{(6-3)}$$

Equation	R^2	DF	SE	F
(6-3)	0.29	16057	1.07	1091

is somewhat satisfactory. The coefficient of determination is improved from the original bivariate model from 0.15 to 0.29. Most of the improvement is from the inclusion of the *lnflowsq* term, but the seasonal variables are significant. For a given flowrate, the positive coefficient for *summer* and the negative coefficient for *winter* indicate that the sediment concentration will be greater in the summer months than in the winter months. During the summer, as the flow increases, the seasonal effect is diminished because of the negative coefficient for the interaction term *sumflow*. The opposite is true during the winter -- as the flow increases, the seasonal effect is amplified because both *winter* and *winflow* have negative coefficients.

To further identify seasonal trends, a daily average for each month for load, flow, concentration, and rainfall (spatially averaged) was computed. The averages were plotted against the month of the year, as shown in Figure 6-10.

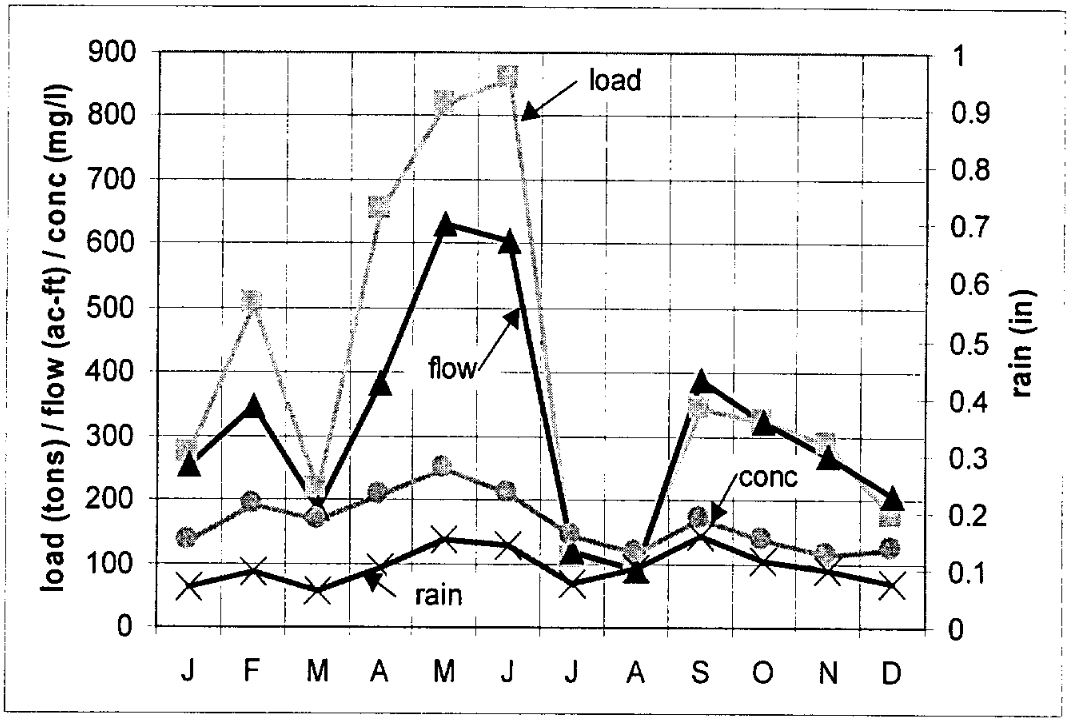


Figure 6-10. Monthly trends in load, flow, concentration, and rainfall

Seasonal trends can quickly be identified. The highest rainfall averages are in May and June, then again in September. The flow in the Lavaca River corresponds with the rainfall over the basin; although the high rainfall in September produces sixty-seven percent less flow per depth of rain than the high rainfall in

May. This phenomenon may be explained by the low moisture antecedent conditions prior to the September rains. The sediment concentration is at its highest in May. Perhaps, the cultivating of farm lands during this time is responsible. The dip in the curves in March is peculiar. Perhaps the March dip is simply due to the lower rainfall in March.

The percent change in each variable from month to month is computed to determine an average percent change from month to month. The average percent change from month to month for load, flow, concentration, and rain is 90, 62, 20, and 31, respectively. The average daily rainfall varies less from month to month than load and flow.

Comparison of the average daily temperature (Figure 6-9) with the averages for sediment load and flow (Figure 6-10) shows that while the average daily temperature is increasing, the load to flow ratio is the greatest. During the first half of the year, additional soil is lost from cultivating practices and contributes to the sediment load carried by the river. During the second half of the year, the ground is stabilized in addition to being protected by a vegetative canopy.

The trends shown in Figure 6-10 lead to a different formulation of seasonal variables than those that were developed based on the temperature trend. when cultivation is taking place. Based on Figure 6-7, a binary variable was formed to identify months when the load-flow ratio is the highest (i.e., February, April, May and June.) The inclusion of the variable in the model was less effective than the summer and winter variables. According to the models, the temperature effect on suspended sediment is due to the nature of summer storms and not due to land cover or cultivation practices.

6.5.3 Timing considerations

Additional variables considered in the multivariate analysis include whether the sample was taken while the river was rising or falling as well as correlating the sediment load with lagged flow. The binary variable rising was assigned a 1 if the flow was 5% greater than the previous day's flow. Similarly, falling was assigned a 1 if the flow was 5% less than the previous day's flow. This definition of the rising and falling limb of the hydrograph proved to be effective. The resulting model is given by

$$\mathit{lnconc} = \underset{(125)}{3.04} + \underset{(48)}{0.261}(\mathit{lnflow}) + \underset{(17)}{0.460}(\mathit{rising}) + \underset{(8.9)}{0.187}(\mathit{falling}) \tag{6-4}$$

Equation	R^2	DF	SE	F
(6-4)	0.17	16057	1.16	1082

The larger positive coefficient for *rising* than for *falling* shows that sediment concentration is greater on the rising limb of the hydrograph than on the falling limb. Furthermore, if the flow is quasi-steady (not rising or falling), then the sediment concentration is the least. In other words, when the flow is storm induced, then concentrations will be higher. Similar results are shown when *lnflowsq* is also included in the model:

$$\mathit{lnconc} = \underset{(131)}{3.70} - \underset{(-14)}{0.164}(\mathit{lnflow}) + \underset{(41)}{0.0573}(\mathit{lnflowsq}) + \underset{(10)}{0.263}(\mathit{rising}) + \underset{(7.7)}{0.154}(\mathit{falling}) \quad \textbf{(6-5)}$$

Equation	R^2	DF	SE	F
(6-5)	0.25	16057	1.11	1306

Interaction terms, *flowfall* and *flowrise*, were formed with the product of *lnflow* and *falling* and *rising*, respectively. The interaction terms do not assist in further accounting for the variability in the data. The interaction terms demonstrate that the effect of the hydrograph limb is most important for larger flows.

Including a variable for the previous day's flow was also investigated. Because of the small basin, the concentration correlates better with the same day's flow than with the previous day's flow. The previous day's flow is a significant variable but is not effective at improving the model.

6.6 Temporal Analysis

Several different methods were used to investigate the temporal nature of the data. Some of these methods were presented in 6.5.2 Seasonal Considerations. In addition to those methods presented in 6.5.2, daily graphs, annual daily averages, and aerial photos were reviewed.

6.6.1 Daily graphs

The daily concentration, flowrate, and spatially averaged rain were plotted against time for the entire sequence of data. Each year of data is on a separate plot. A typical annual plot for 1970 is shown in Figure 6-11. Each figure is included in Appendix E. Several things can be observed from these plots. The sediment concentration is highly variable. Even during times of relatively steady flow, the concentration fluctuates regularly. The daily plots also assist in interpreting the changes in annual daily averages.

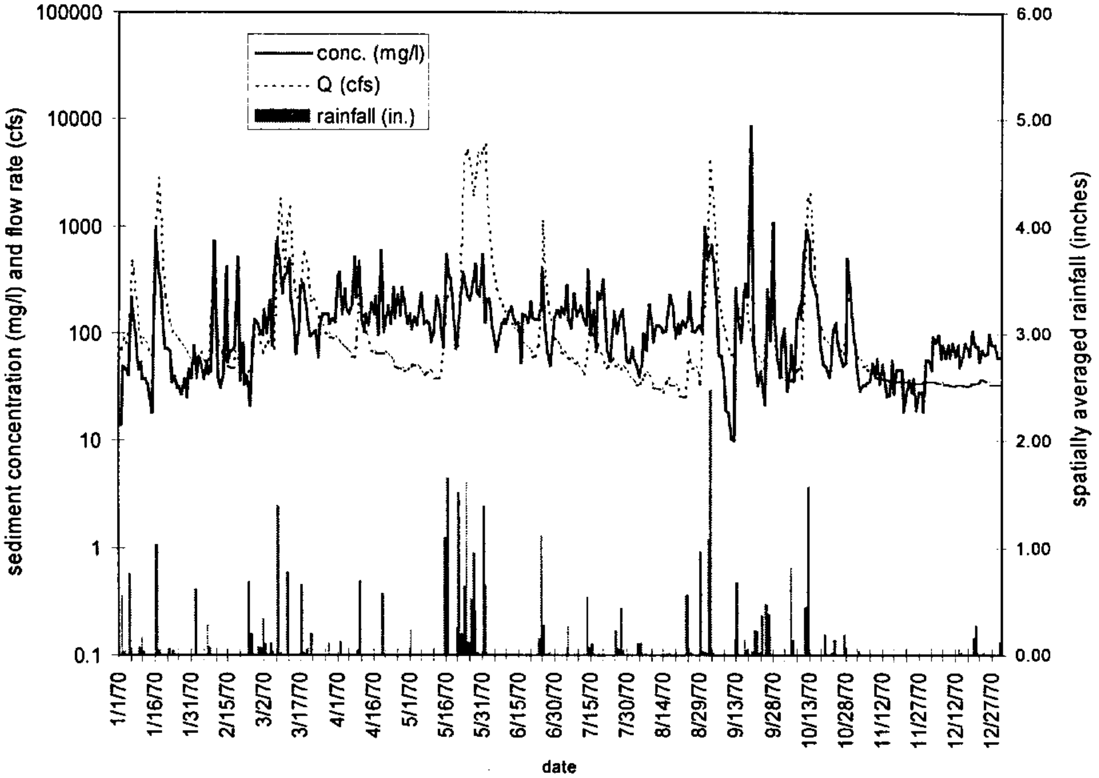


Figure 6-11. Typical plot of one year of daily data

6.6.2 Annual Daily Averages

The annual average values were plotted against time. The averages vary from year to year and the extreme years can quickly be identified. However, no trend can be detected from the plot. To assist in identifying a temporal change in the flow and load, a mass curve technique was employed. The cumulative load and the cumulative flow were plotted against time as shown in Figure 6-12. The slope of the cumulative load curve is relatively constant compared to the slope of the cumulative flow curve which changes dramatically in the late 1950's. The cumulative load was then plotted against the cumulative flow as shown in Figure 6-13. Each point on the figure represents the end of the year. Changes between the load-flow relationship are quickly apparent. The period from 1945-1959 has the steepest slope, indicating more load was produced for a given flow during that time. Perhaps, poorer land management practices were in effect during that time. 1960 was a wet year with a lower average daily concentration. The load-flow relationship appears relatively constant from 1965 to 1980 in spite of the fact that 1973 and 1979 had the two highest flows for this period of record. A break in the double mass curve is apparent at 1981 when the highest sediment load for this period of record occurred.

Investigation of the daily plots for 1960 may help explain the first break in the double mass curve. 1960 was a very wet year with a total rainfall of 54 inches.

The warm months of July, August, and September are full of consecutive days of rain. In fact, there was only one day during the period of August 7, 1960 to September 9, 1960 in which rain was not detected at one of the three stations. The consecutive days of rain result in compacted soil and reduced erosion. 1981 was definitely an extreme year producing more sediment load than any other year of the record. 1981 had a total rainfall of 55 inches. The daily plot for 1981 shows extreme rainfall events. On August 31, 1981, the spatially averaged rainfall was 5.6 inches. On October 31, 1981, the spatially averaged rainfall was 4.2 inches. There were 17 days during 1981 where the daily rainfall exceeded 1 inch.

To further account for the annual trends, three time categories were created: one for 1945-1959, 1966-1980, and 1982-1989. The load-flow and concentrations-flow relationships for each category are shown in Table 6-11.

Based on the double mass curve, the load was expected to be the highest in the first time category. The first time category regression equation produces greater loads than the other two equations if the flows are low. If the flows are high, the higher coefficients for *lnflow* in the second two equations will cause higher sediment loads.

The temporal analysis thus far does not account for any land use changes that may have occurred during the period of record. To account for land use changes, aerial photos were reviewed.

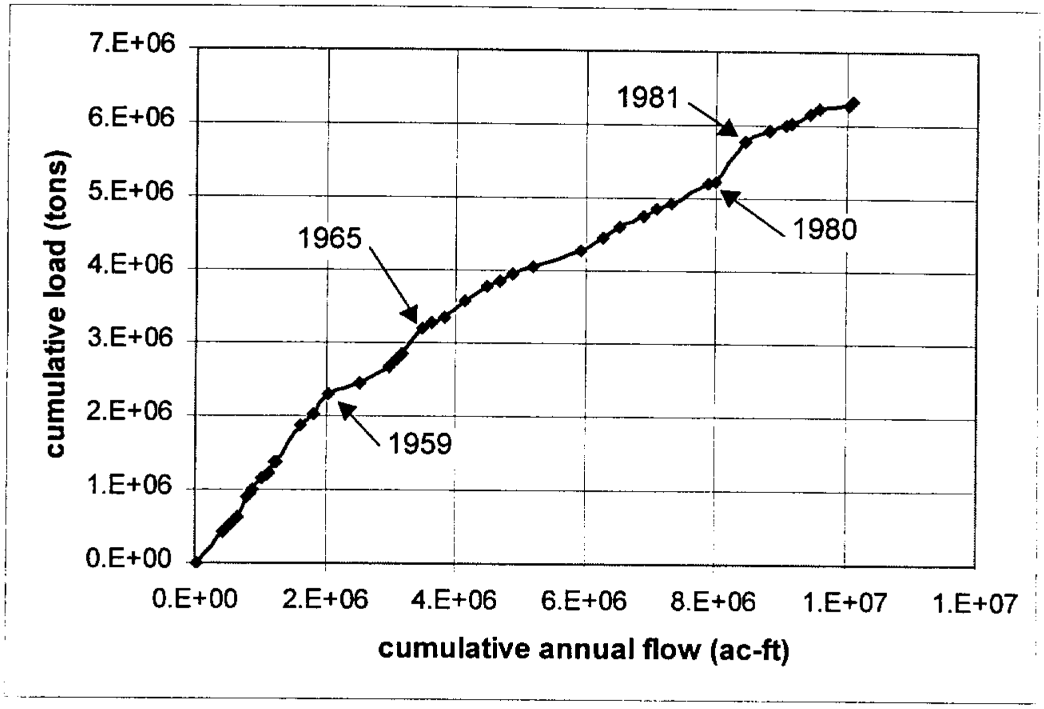


Figure 6-12. Mass curves for flow and load

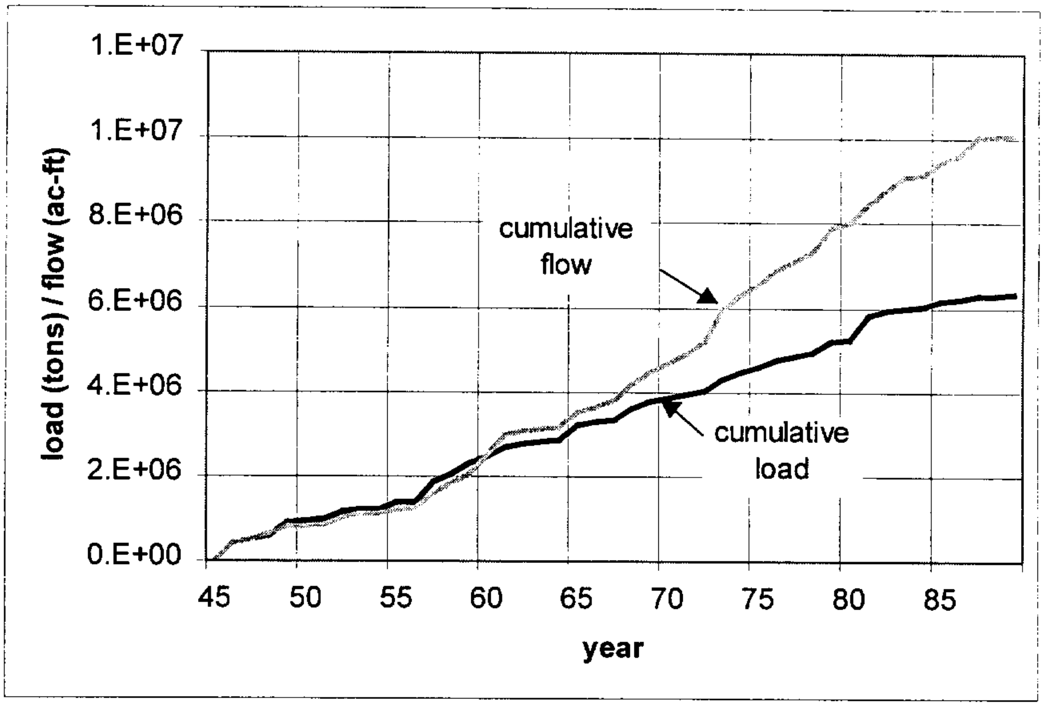


Figure 6-13. Double mass curve identifying changes in load-flow relationship

Table 6-11. Load-flow relationship according to year categories

	Equation *(t-statistics)*	R^2	DF	SE	F
(1946-1959)	$lnload = -2.27 + 1.21(lnflow)$ *(-59) (119)*	0.73	5234	1.49	14213
	$lnload = -1.91 + 0.65(lnflow) + 0.10(lnflowsq)$ *(-55) (40) (41)*	0.80	5234	1.29	10204
	$lnconc = 3.55 + 0.21(lnflow)$ *(92) (21)*	0.08	5234	1.49	443
	$lnconc = 3.91 - 0.35(lnflow) + 0.10(lnflowsq)$ *(112) (-21) (41)*	0.30	1126	1.29	1126
(1966-1980)	$lnload = -3.00 + 1.35(lnflow)$ *(-74) (159)*	0.82	5478	0.87	25134
	$lnload = -2.68 + 1.21(lnflow) + 0.013(lnflowsq)$ *(-30) (35) (4)*	0.82	5478	0.87	12611
	$lnconc = 2.82 + 0.35(lnflow)$ *(69) (41)*	0.23	5478	0.87	1664
	$lnconc = 3.14 + 0.21(lnflow) + 0.013(lnflowsq)$ *(35) (6) (4)*	0.24	5478	0.87	843
(1982-1989)	$lnload = -3.57 + 1.40(lnflow)$ *(-69) (117)*	0.83	2786	0.95	13716
	$lnload = -3.55 + 1.40(lnflow) + 0.00077(lnflowsq)$ *(-37) (33) (41)*	0.83	2786	0.95	6855
	$lnconc = 2.25 + 0.41(lnflow)$ *(43) (34)*	0.29	2786	0.95	1136
	$lnconc = 2.27 + 0.40(lnflow) + 0.00077(lnflowsq)$ *(24) (9) (0.17)*	0.29	2786	0.95	568

6.6.3 Photo review

Archived aerial photos were ordered from the USDA. The cost was prohibitive to order all of the photos. Instead, the photo index was ordered for each available year as well as the actual photos in the area near the station. By reviewing the photo indexes, general basin characteristics can be compared from year to year. The photos near the station can reveal any dramatic changes in the river upstream of the station. The scale on the photo index is approximately 1 inch = 1 mile. The photo indexes exist for March, 1956, February, 1962, and March, 1968. The frames obtained in the area of the Lavaca River are dated 2/11/56, 1/17/61, and 11/15/69. These frames are 10 inches by 10 inches at a 1:20,000 scale so that each frame covers approximately 8 square miles. The index and flight lines are not consistent between the three, making it slightly more difficult to compare between years. In addition, the contrast on the photos varies dramatically. There is a much larger contrast in the 1968 index than in the other two. Figures 6-14 and 6-15 both show a portion of the 1:20000 scale maps for 1956 and 1969, respectively.

Figure 6-14. Lavaca River, 1956

Figure 6-15. Lavaca River, 1969

In general, the area along the river meanders and is well forested. The 1961 photos show a small area near the river where the trees are gone. The 1969 photo shows more trees gone from the same area and some additional area was cleared. Essentially no difference in urban development from 1956 to 1961 to 1969 is evident. The Lavaca River follows the same course upstream and downstream of the Edna station. The 10-inch frames clearly show the progression of the highway construction. In 1956, U.S. Highway 59 is under construction and the south side of the bridge is in place. In 1961, the highway is in place. In 1969, the interchange was added. The interchange area is outside of the study area.

From 1956 to 1961, development of Edna increased to the North and East. By 1968, the Edna bypass was under construction and development continued to increase to the North and slightly to the south. From 1956 to 1968, development increased in the Hallettesville area but the area just became denser and did not spread much. The tree band along the river is visibly narrower in places in 1968.

The area in general has small roads, small-farmed areas, and moderate cover. A few more plowed fields exist in 1970 but general use remains the same in the basin. There is no change in land use apparent from the photos that would be responsible for the break in the double mass curve from 1959 to 1960.

6.7 Conclusions

The Lavaca River near Edna exhibits a high correlation between sediment load and flowrate. The correlation can be improved by considering a flow-squared term and by considering seasonal effects. The infrequent high flows carry a large, disproportionate amount of sediment. To predict future sediment loads, it is best to use a model developed with the data that is post-1965. The following model was developed using post-1965 data and excluding flows and concentrations that are less than 10% of the mean flow and mean concentration:

$$\mathit{lnconc} = \underset{(-2.0)}{-0.410} + \underset{(22)}{1.23}(\mathit{lnflow}) - \underset{(-16)}{0.0771}(\mathit{lnflowsq}) + \underset{(9.8)}{0.0658}(\mathit{temp})$$

$$- \underset{(-5.3)}{0.00707}(\mathit{flowtemp}) - \underset{(-4.4)}{0.614}(\mathit{rising}) + \underset{(7.7)}{0.225}(\mathit{flowrise})$$

$$- \underset{(-4.3)}{0.548}(\mathit{falling}) + \underset{(4.0)}{0.110}(\mathit{flowfall}) \qquad \mathbf{(6\text{-}6)}$$

Equation	R^2	DF	SE	F
(6-6)	0.37	6561	0.73	485

Equation 6-6 can be used to predict the average daily concentration if the daily flowrate is known. If the daily flowrate is known, all other variables should be known. The load can then be calculated based on the estimated concentration. The simpler bivariate relationships obtained from the same data are

$$\mathit{lnload} = \underset{(-80)}{-3.21} + \underset{(175)}{1.38}(\mathit{lnflow}) \qquad \mathbf{(6\text{-}7)}$$

Equation	R^2	DF	SE	F
(6-7)	0.82	6561	0.79	30642

$$\mathit{lnconc} = 2.61 + 0.383(\mathit{lnflow}) \quad \textbf{(6-8)}$$
$$\quad (65) \qquad (48)$$

Equation	R^2	DF	SE	F
(6-8)	0.26	6561	0.79	2349

To determine loads from unusual years or events, lessons can be learned from 1960 and 1981. As shown in 1960, if there are many days of consecutive rain, the flow will continue to increase but the load will not. On the other hand, if extreme isolated storms occur, such as those in 1981, then the load will also be extreme.

7 Conclusions

The Texas Water Development Board's suspended sediment sampling program has resulted in a tremendous data base that can be used to help quantify spatial and temporal trends of sediment yield in Texas. Also useful in defining spatial and temporal trends are the streamflow, the rainfall, and the watershed characteristics. Watershed characteristics from each watershed corresponding to a sediment gauging station can be obtained using GIS with the macros written for this research. All of the data can be incorporated into a SAS data set to enable creation of new variables and statistical analyses.

7.1 Annually Averaged Data for all Stations in Texas

In general, stations with higher runoff produce more sediment load, whereas stations with higher rainfall have lower average sediment concentrations. No correlation exists between sediment concentration and annual runoff from station to station. The stations that have upstream dams have lower sediment loads than stations with no dams; however, no difference in sediment concentration can be detected between the basins that have dams and the basins with no dams.

7.2 Annual Data for all Stations

Sediment rating curves are developed for each basin using the annual values. Although the range of values for the coefficients of the model appear to vary somewhat dramatically, when the sediment rating curves are plotted, they generally fall in the same band that varies about two orders of magnitude. The Red River Basin and the Colorado River Basin are the two curves which fall outside the band. When the sediment rating curves are re-computed using those stations with no dams, the Colorado River Basin curve falls within the band. The Red River Basin rating curve is largely influenced by two stations which are located in the Texas panhandle and which have relatively low annual runoff but high sediment yields.

The watersheds can be classified according to the spatially averaged rainfall over the basin. The low rainfall climate streams carry the largest sediment concentrations and thus have higher sediment loads for corresponding streamflows than do medium or high rainfall climates. A similar climate analysis considering the stations with no dams shows that the dams have the most impact on sediment load in low rainfall climates.

7.3 Watershed Characteristics

The sediment rating curves can be improved by considering various watershed characteristics as additional variables. One variable of particular interest relates to reservoirs in the watershed. The variable *resvar* is a function of the area in the watershed that contributes to reservoirs and the distances these reservoirs are from the sediment gauging station. The variable *resvar* increases with larger reservoir contributing areas and with shorter distances to the sediment gauging station. When this variable is included in a multivariate regression model, the coefficient is negative. Regression models that include the reservoir variable can be used as a tool to predict changes in sediment yield due to the addition of dams in the basin.

When using a regression model to predict future sediment loads, it is best to use a model developed for watersheds with similar amounts of annual rainfall.

7.4 Lavaca River near Edna Samples (taken by USGS)

Periodic suspended sediment samples collected by the USGS include concentration, percent fines, and water temperature. The additional information can be used to form interaction terms improving the coefficient of determination from 0.25 to 0.61 for the concentration-flow relationship. A seasonal effect can be detected. For most flows, the sediment concentration is higher during the warmer months. As the flow increases, the seasonal effect becomes less apparent. The higher sediment concentrations correspond with smaller percent fines implying that when larger particles contribute to the load, the concentration will be higher. The model also shows that a linear relationship does not exist between concentration and flow. As flow increases concentration increases; however concentration does not continue to increase at the same rate as flow.

7.5 Lavaca River near Edna Daily Data (taken by TWDB)

A frequency analysis of forty-five years of daily flows shows that about ninety percent of the flows are less than 385 cfs. Observations of plots of load-flow and concentration-flow imply non-linear relationships. Over the forty-five year period of record, only 10% of the total sediment load was produced by flow rates less than 385 cfs. Hence, 90% of the sediment load is produced by the highest 10% of the flows. Because the load-flow relationship is non-linear, different sediment rating curves are developed for different flow rates. Including variables accounting for seasonal effects improves the model more than including rainfall variables. A double mass diagram helps identify periods where the load-flow relationship changes. Models developed with post-1965 data can be used to predict suspended sediment concentration. The variables in the model include flow, temperature, whether the river is rising or falling, and interaction terms of the variables.

7.6 Temporal and Spatial Implications

Spatial and temporal trends in sediment yield are identified using the TWDB data. Substitution of time for space and vice versa is an interesting notion that may be able to be pursued further with the data. For example, there are differences in sediment response due to the aridity of the climate as you go from one location to another across Texas. If a site has highly variable rainfall, it may be possible to use models developed from nearby spatial locations to predict the pattern of temporal loads at an ungauged site. Daily data from several stations would have to be considered to see if the time-space inference is feasible. Furthermore, it may be possible to predict seasonal trends in sediment load in an ungauged area using the

Lavaca River as an index and a ratio of average annual load for the station to average annual load for the Lavaca River.

7.7 Future Work

There are several opportunities for future work.

(1) Analyze periodic USGS samples for more stations to see if the seasonal and grain size effects are similar between stations.

(2) Use a depth integrated sampler at several locations across the river cross section, for a wide range of flow rates, for each station, to determine a more definitive relationship between the Texas sampler and the cross section sediment yield.

(3) Analyze the daily data from several nearby locations to further explore the time-space implications.

(4) Compare predicted load with bathymetric survey results.

8 Appendices

8.1 Appendix A - Fortran programs

The following Fortran programs were used to change the format of the data so that it could be used in Excel or SAS.

****

```
C      PROGRAM STATLOAD.FOR
C
C      THIS PROGRAM READS AND WRITES SEDIMENT LOAD FOR ONE STATION
C
C
       DIMENSION NDAYS(12),VAL(31)
C
C      OPEN(1,FILE="D1.DAT")
       OPEN(2,FILE="STATLOAD.OUT")
C
       NDAYS(1)=31
       NDAYS(2)=28
       NDAYS(3)=31
       NDAYS(4)=30
       NDAYS(5)=31
       NDAYS(6)=30
       NDAYS(7)=31
       NDAYS(8)=31
       NDAYS(9)=30
       NDAYS(10)=31
       NDAYS(11)=30
       NDAYS(12)=31
C
5      CONTINUE
C
       READ(1,10,END=900)ISTA1,IYR1,IMO1,ICARD1,(VAL(I),I=1,10)
10     FORMAT(I8,I2,I2,I1,10F6.0)
       IF(ICARD1.NE.1)GO TO 5
       READ(1,10,END=900)ISTA2,IYR2,IMO2,ICARD2,(VAL(I),I=11,20)
       IF(ICARD2.NE.2)GO TO 5
       READ(1,20,END=900)ISTA3,IYR3,IMO3,ICARD3,(VAL(I),I=21,31)
20     FORMAT(I8,I2,I2,I1,11F6.0)
       IF(ICARD3.NE.3)GO TO 5
C
       ND=NDAYS(IMO1)
       IF(IMO1.EQ.2.AND.AMOD(IYR1,4).EQ.0)ND=29
```

```
C
      DO 100 I=1,ND
      WRITE(2,30)IMO1,I,IYR1,VAL(I)
      WRITE(*,30)IMO1,I,IYR1,VAL(I)
30    FORMAT(I5,I5,I5,F10.2)
100   CONTINUE
C
      GO TO 5
C
900   CONTINUE
      CLOSE(1)
      CLOSE(2)
      END
```

****

```
C     PROGRAM STATFLOW.FOR
C
C     THIS PROGRAM READS AND WRITES STREAMFLOW FOR ONE STATION.
C
C
      CHARACTER*128 ALINE
      CHARACTER*8 FIELD
      DIMENSION FIELD(31,12),NDAYS(12)
C
C     OPEN(1,FILE="D1.DAT")
      OPEN(2,FILE="STATFLOW.OUT")
C
      NDAYS(1)=31
      NDAYS(2)=28
      NDAYS(3)=31
      NDAYS(4)=30
      NDAYS(5)=31
      NDAYS(6)=30
      NDAYS(7)=31
      NDAYS(8)=31
      NDAYS(9)=30
      NDAYS(10)=31
      NDAYS(11)=30
      NDAYS(12)=31
C
      READ(1,11,END=900) ALINE
      READ(1,11,END=900) ALINE
C      READ(1,11,END=900) ALINE
11    FORMAT(A128)
5     CONTINUE
C
      READ(1,10,END=900)IYR
10    FORMAT(14X,I2)
      DO 50 I=1,31
```

```
      READ(1,20,END=900)IDAY,(FIELD(I,J),J=1,12)
20    FORMAT(I8,12A8)
50    CONTINUE
      DO 60 L=1,6
      READ(1,30,END=900)ALINE
30    FORMAT(A128)
60    CONTINUE
C
      DO 200 J=1,12
      ND=NDAYS(J)
      IF(J.EQ.2.AND.AMOD(IYR,4).EQ.0)ND=29
C
      DO 100 I=1,ND
      IF(FIELD(I,J).EQ.'        ')GO TO 100
      WRITE(2,40)J,I,IYR,FIELD(I,J)
      WRITE(*,40)J,I,IYR,FIELD(I,J)
40    FORMAT(I5,I5,I5,2X,A8)
100   CONTINUE
200   CONTINUE
C
      GO TO 5
C
900   CONTINUE
      CLOSE(1)
      CLOSE(2)
      END
```

****

```
C     This program reads rainfall data from the format created
C     using the Hydrosphere CD ROMs.  The data is written to a
C     column format which can be used in SAS or EXCEL.
C
C     PROGRAM STATRAIN.FOR
      INTEGER STARTYR,ENDYR,TOTYRS
      CHARACTER*128 ALINE
      CHARACTER*8 FIELD,TEXT,DATE
      DIMENSION FIELD(366,50),IYR(50),DATE(366)
C
C     OPEN(1,FILE="D4.DAT")
      OPEN(2,FILE="STATRAIN.OUT")
C
C
      DO 10 M=1,11
      READ(1,11,END=900) ALINE
11    FORMAT(A128)
10    CONTINUE
C
      READ(1,12,END=900)TEXT,STARTYR
```

```
12     FORMAT(A8,I8)
       READ(1,13,END=900)TEXT,ENDYR
13     FORMAT(A8,I8)
       READ(1,14,END=900)TEXT,TOTYRS
14     FORMAT(A8,I8)
C
       DO 20 M=1,9
       READ(1,15,END=900) ALINE
15     FORMAT(A128)
20     CONTINUE

C
       READ(1,16,END=900)TEXT,(IYR(K),K=1,TOTYRS)
16     FORMAT(A8,50I8)
       DO 50 I=1,366
       READ(1,17,END=900)DATE(I),(FIELD(I,K),K=1,TOTYRS)
17     FORMAT(A8,50A8)
50     CONTINUE
C
C       WRITE(*,40)DATE(3),IYR(3),FIELD(3,3)
C       WRITE(*,40)DATE(64),IYR(8),FIELD(64,8)
C       40     FORMAT(A8,I5,2X,A8)
C
       DO 100 K=1,TOTYRS
       DO 200 I=1,366
C       IF(FIELD(I,K).EQ.'        ')GO TO 200
       WRITE(2,40)DATE(I),IYR(K),FIELD(I,K)
       WRITE(*,40)DATE(I),IYR(K),FIELD(I,K)
40     FORMAT(A8,I5,2X,A8)
200    CONTINUE
100    CONTINUE
C
C
900    CONTINUE
       CLOSE(1)
       CLOSE(2)
       END

                   ********************************


C      PROGRAM MONLOAD.FOR
C
C      This program reads and writes monthly sediment data.
C
       DIMENSION DATA(12)
C
C      OPEN(1,FILE="D1.DAT")
       OPEN(2,FILE="MONLOAD.OUT")
C
C
```

```
5     CONTINUE
C
      READ(1,10,END=900)ISTA1,IYR1,ICARD1,(DATA(I),I=1,6)
10    FORMAT(I8,I2,I1,6F10.0)
      IF(ICARD1.NE.1)GO TO 5
      READ(1,10,END=900)ISTA2,IYR2,ICARD2,(DATA(I),I=7,12)
      IF(ICARD2.NE.2)GO TO 5
C
C
      DO 100 I=1,12
      IMO=I
      WRITE(2,30)ISTA1,IMO,IYR1,DATA(I)
      WRITE(*,30)ISTA1,IMO,IYR1,DATA(I)
30    FORMAT(I8,I5,I5,F12.2)
100   CONTINUE
C
      GO TO 5
C
900   CONTINUE
      CLOSE(1)
      CLOSE(2)
      END
```

* * *

```
C     PROGRAM ANNLOAD.FOR
C
C     This program reads writes monthly sediment data and
C     write annual sediment data.
C
      REAL LD(12)
C
C     OPEN(1,FILE="D1.DAT")
      OPEN(2,FILE="ANNLOAD.OUT")
C
5     CONTINUE
C
      DO 100 I=1,12
      READ(1,10,END=900)ISTA,IMO,IYR,LD(I)
10    FORMAT(I8,I5,I5,F12.2)
100   CONTINUE
C
      ANNLD=LD(1)+LD(2)+LD(3)+LD(4)+LD(5)+LD(6)
     1      +LD(7)+LD(8)+LD(9)+LD(10)+LD(11)+LD(12)
      DO 200, I=1,12
      IF(LD(I).EQ.-9999)ANNLD=-9999
200   CONTINUE
      WRITE(2,30)ISTA,IYR,ANNLD
      WRITE(*,30)ISTA,IYR,ANNLD
30    FORMAT(I8,I5,F12.2)
C
```

```
      GO TO 5
C
900   CONTINUE
      CLOSE(1)
      CLOSE(2)
      END
```

8.2 Appendix B – Arc/Info AML's

The following are text files and macros that were developed and used for the GIS portion of this research. They are listed alphabetically by name.

Albprj

```
input
projection geographic
datum NAD27
units dd
parameters
output
projection albers
datum NAD83
units meters
parameters
27 25 00
34 55 00
-100 00 00
31 10 00
1000000.0
1000000.0
end
```

Cov1.txt

```
1,931204.3793928,1378828.350427
2,1026882.566392,1325379.350659
5,1409597.149572,1251327.683163
7,1454652.183163,1249389.932932
8,1520101.191392,1146860.016342
10,1413700.316392,1135367.141342
12,1467323.816392,1005240.191392
14,1508188.816392,1061093.183163
17,1288493.074622,1239199.058163
19,1342655.256764,1110407.75767
21,1481642.291705,961422.9372098
23,1435880.641342,906852.1209325
25,1468690.826645,917610.3671174
27,980769.7957608,1203822.439928
29,1033178.616886,1196088.566392
34,1263687.729715,1083348.558713
35,1186076.706945,1091328.908707
36,1212654.566161,1032428.003487
38,1252930.064475,979937.2283572
39,1334692.767248,911346.1584759
40,1334715.650477,911346.1584759
```

```
41,1351591.642248,1006437.587572
43,1122150.352701,1005874.750336
44,1136635.241211,1006937.874209
47,1126820.208969,954851.1495721
48,1153132.283707,903858.0746221
50,1321633.868235,759818.8410806
51,1308199.478482,816033.8341898
53,1154044.878045,854895.322925
55,1188113.024572,754877.9685571
56,1201476.107982,762848.8876221
58,1073645.325971,695832.7214759
59,1161695.808163,704436.2043429
end
```

Cov2.txt

```
3,1318141.150246,1299343.733657
6,1452070.683163,1256835.79103
9,1564651.775477,1105147.066161
11,1494198.806814,1007797.332851
15,1542182.650246,1019315.699391
18,1332490.982751,1145486.624441
24,1456068.067067,882899.0301986
28,1069197.399572,1267843.049934
30,1072547.241211,1196870.63267
37,1243146.634019,991230.8764647
45,1151292.071461,955829.7917051
52,1334228.394175,767332.2297158
54,1291073.883788,739943.363609
57,1255797.340849,722892.2804297
60,1178268.806814,697303.959875
end
```

Cov3.txt

```
4,1490764.441392,1292310.999209
13,1533357.283476,997842.4829825
16,1557123.301514,997845.9085808
20,1411653.467429,1027053.50767
26,1469133.731402,870320.0888257
31,1118912.658707,1201366.058163
46,1154622.424259,952845.3293429
61,1210350.349753,654378.5670674
end
```

Cov4.txt

```
22,1493638.732982,928907.6881155
32,1126631.583969,1206360.292379
49,1221939.74098,899899.2626221
end
```

Cov5.txt

```
33,1147918.403522,1189850.076645
end
```

Cov6.txt

```
42,1421629.649572,836416.8972005
end
```

Indsheds.aml

```
/* create individual polygons of watersheds
grid
shed1 = con(shedcov4 == 1,1)
shed2 = con(shedcov4 == 2,1)
shed5 = con(shedcov4 == 3,1)
shed7 = con(shedcov4 == 4,1)
shed8 = con(shedcov4 == 5,1)
shed10 = con(shedcov4 == 6,1)
shed12 = con(shedcov4 == 7,1)
shed14 = con(shedcov4 == 8,1)
shed17 = con(shedcov4 == 9,1)
shed19 = con(shedcov4 == 10,1)
shed21 = con(shedcov4 == 11,1)
shed23 = con(shedcov4 == 12,1)
shed25 = con(shedcov4 == 13,1)
shed27 = con(shedcov4 == 14,1)
shed29 = con(shedcov4 == 15,1)
shed34 = con(shedcov4 == 16,1)
shed35 = con(shedcov4 == 17,1)
shed36 = con(shedcov4 == 18,1)
shed38 = con(shedcov4 == 19,1)
shed39 = con(shedcov4 == 20,1)
shed40 = con(shedcov4 == 21,1)
shed41 = con(shedcov4 == 22,1)
shed43 = con(shedcov4 == 23,1)
shed44 = con(shedcov4 == 24,1)
shed47 = con(shedcov4 == 25,1)
shed48 = con(shedcov4 == 26,1)
shed50 = con(shedcov4 == 27,1)
shed51 = con(shedcov4 == 28,1)
shed53 = con(shedcov4 == 29,1)
shed55 = con(shedcov4 == 30,1)
shed56 = con(shedcov4 == 31,1)
shed58 = con(shedcov4 == 32,1)
shed59 = con(shedcov4 == 33,1)
q
gridpoly shed1 poly1
gridpoly shed2 poly2
gridpoly shed3 poly3
gridpoly shed4 poly4
```

```
gridpoly shed5 poly5
gridpoly shed6 poly6
gridpoly shed7 poly7
gridpoly shed8 poly8
gridpoly shed9 poly9
gridpoly shed10 poly10
gridpoly shed11 poly11
gridpoly shed12 poly12
gridpoly shed13 poly13
gridpoly shed14 poly14
gridpoly shed15 poly15
gridpoly shed16 poly16
gridpoly shed17 poly17
gridpoly shed18 poly18
gridpoly shed19 poly19
gridpoly shed20 poly20
gridpoly shed21 poly21
gridpoly shed22 poly22
gridpoly shed23 poly23
gridpoly shed24 poly24
gridpoly shed25 poly25
gridpoly shed26 poly26
gridpoly shed27 poly27
gridpoly shed28 poly28
gridpoly shed29 poly29
gridpoly shed30 poly30
gridpoly shed31 poly31
gridpoly shed32 poly32
gridpoly shed33 poly33
gridpoly shed34 poly34
gridpoly shed35 poly35
gridpoly shed36 poly36
gridpoly shed37 poly37
gridpoly shed38 poly38
gridpoly shed39 poly39
gridpoly shed40 poly40
gridpoly shed41 poly41
gridpoly shed42 poly42
gridpoly shed43 poly43
gridpoly shed44 poly44
gridpoly shed45 poly45
gridpoly shed46 poly46
gridpoly shed47 poly47
gridpoly shed48 poly48
gridpoly shed49 poly49
gridpoly shed50 poly50
gridpoly shed51 poly51
gridpoly shed52 poly52
gridpoly shed53 poly53
gridpoly shed54 poly54
gridpoly shed55 poly55
```

```
gridpoly shed56 poly56
gridpoly shed57 poly57
gridpoly shed58 poly58
gridpoly shed59 poly59
&return
```

kfact.aml

```
/* this aml creates a kfactor grid coverage
/* from statsgo
copy /res2/maidment/statsgo/statsgo
copyinfo /res2/maidment/statsgo/layer
copyinfo /res2/maidment/statsgo/comp
tables
additem comp kfact 6 6 n 3 seqnum
sel comp
res layernum = 1
relate add
laycomp
comp
info
museq
museq
ordered
rw

calc laycomp//kfact = laycomp//kfact + kfact
relate drop
laycomp

additem comp fkfact 6 6 n 3 kfact
statistics muid kfact.sta
sum fkfact
end
q
joinitem statsgo.pat kfact.sta statsgo.pat muid muid ordered
&return
```

Lonlat.dat

```
1     -100.745    34.57305
2     -99.6833    34.1
3     -96.5631    33.81889
4     -94.6942    33.6875
5     -95.5942    33.35555
6     -95.1325    33.38638
7     -95.0925    33.32222
8     -94.4578    32.36972
9     -94.0061    31.97222
10    -95.6053    32.30944
11    -94.8097    31.13277
12    -95.0864    31.14027
13    -94.3986    31.02472
```

```
14     -94.6411     31.61611
15     -94.2944     31.21139
16     -94.13       31.03556
17     -96.8925     33.28333
18     -96.4622     32.42666
19     -96.3706     32.10805
20     -95.6575     31.33555
21     -94.9586     30.71611
22     -94.8506     30.425
23     -95.4569     30.24444
24     -95.2578     30.02694
25     -95.1039     30.33638
26     -95.1333     29.99444
27     -100.18      33.00806
28     -99.2672     33.58083
29     -99.6422     32.93083
30     -99.2242     32.93444
31     -98.7664     32.96
32     -98.6436     33.02416
33     -98.4333     32.86666
34     -97.2011     31.84444
35     -98.0347     31.9775
36     -97.7617     31.43277
37     -97.4411     31.07
38     -97.3458     30.96639
39     -96.5072     30.32166
40     -96.5072     30.32167
41     -96.2983     31.16944
42     -95.7575     29.58222
43     -98.7192     31.21305
44     -98.5642     31.21777
45     -98.4183     30.75111
46     -98.385      30.73
47     -98.6694     30.75111
48     -98.4003     30.29083
49     -97.6942     30.24444
50     -96.6861     28.95972
51     -96.8125     29.46666
52     -96.5522     29.02555
53     -98.3828     29.86055
54     -97.0128     28.79277
55     -98.0639     28.95139
56     -97.93       29.01388
57     -97.3844     28.64944
58     -99.2406     28.42555
59     -98.3464     28.49194
60     -98.185      28.43611
61     -97.86       28.03805
end
```

Lonlat2.dat

1	-100.745	34.5731
2	-99.6833	34.1000
3	-96.5631	33.8189
4	-94.6942	33.6875
5	-95.5942	33.3556
6	-95.1325	33.3864
7	-95.0925	33.3222
8	-94.4578	32.3697
9	-94.0061	31.9722
10	-95.6053	32.3094
11	-94.8097	31.1328
12	-95.0864	31.1403
13	-94.3986	31.0247
14	-94.6411	31.6161
15	-94.2944	31.2114
16	-94.1300	31.0356
17	-96.8925	33.2833
18	-96.4622	32.4267
19	-96.3706	32.1081
20	-95.6575	31.3356
21	-94.9586	30.7161
22	-94.8506	30.4250
23	-95.4569	30.2444
24	-95.2578	30.0269
25	-95.1039	30.3364
26	-95.1333	29.9944
27	-100.180	33.0081
28	-99.2672	33.5808
29	-99.6422	32.9308
30	-99.2242	32.9344
31	-98.7664	32.9600
32	-98.6436	33.0242
33	-98.4333	32.8667
34	-97.2011	31.8444
35	-98.0347	31.9775
36	-97.7617	31.4328
37	-97.4411	31.0700
38	-97.3458	30.9664
39	-96.5072	30.3217
40	-96.5072	30.3217
41	-96.2983	31.1694
42	-95.7575	29.5822
43	-98.7192	31.2131
44	-98.5642	31.2178
45	-98.4183	30.7511
46	-98.3850	30.7300
47	-98.6694	30.7511
48	-98.4003	30.2908
49	-97.6942	30.2444
50	-96.6861	28.9597

```
      51   -96.8125 29.4667
      52   -96.5522 29.0256
      53   -98.3828 29.8606
      54   -97.0128 28.7928
      55   -98.0639 28.9514
      56   -97.9300 29.0139
      57   -97.3844 28.6494
      58   -99.2406 28.4256
      59   -98.3464 28.4919
      60   -98.1850 28.4361
      61   -97.8600 28.0381
end
```

Nattotx.prj

```
input
projection albers
units meters
datum nad27
parameters
29 30 00
45 30 00
-96 00 00
23 00 00
0.0
0.0
output
projection albers
units meters
datum nad83
parameters
27 25 00
34 55 00
-100 00 00
31 10 00
1000000.0
1000000.0
end
```

Newout.txt

```
1,931204.3793928,1378828.350427
2,1026882.566392,1325379.350659
3,1318141.150246,1299343.733657
4,1490764.441392,1292310.999209
5,1409597.149572,1251327.683163
6,1452070.683163,1256835.79103
7,1454652.183163,1249389.932932
8,1520101.191392,1146860.016342
9,1564651.775477,1105147.066161
10,1413700.316392,1135367.141342
11,1494198.806814,1007797.332851
12,1467323.816392,1005240.191392
```

```
13,1533357.283476,997842.4829825
14,1508188.816392,1061093.183163
15,1542182.650246,1019315.699391
16,1557123.301514,997845.9085808
17,1288493.074622,1239199.058163
18,1332490.982751,1145486.624441
19,1342655.256764,1110407.75767
20,1411653.467429,1027053.50767
21,1481642.291705,961422.9372098
22,1493638.732982,928907.6881155
23,1435880.641342,906852.1209325
24,1456068.067067,882899.0301986
25,1468690.826645,917610.3671174
26,1469133.731402,870320.0888257
27,980769.7957608,1203822.439928
28,1069197.399572,1267843.049934
29,1033178.616886,1196088.566392
30,1072547.241211,1196870.63267
31,1118912.658707,1201366.058163
32,1126631.583969,1206360.292379
33,1147918.403522,1189850.076645
34,1263687.729715,1083348.558713
35,1186076.706945,1091328.908707
36,1212654.566161,1032428.003487
37,1243146.634019,991230.8764647
38,1252930.064475,979937.2283572
39,1334692.767248,911346.1584759
40,1334715.650477,911346.1584759
41,1351591.642248,1006437.587572
42,1421629.649572,836416.8972005
43,1122150.352701,1005874.750336
44,1136635.241211,1006937.874209
45,1151292.071461,955829.7917051
46,1154622.424259,952845.3293429
47,1126820.208969,954851.1495721
48,1153132.283707,903858.0746221
49,1221939.74098,899899.2626221
50,1321633.868235,759818.8410806
51,1308199.478482,816033.8341898
52,1334228.394175,767332.2297158
53,1154044.878045,854895.322925
54,1291073.883788,739943.363609
55,1188113.024572,754877.9685571
56,1201476.107982,762848.8876221
57,1255797.340849,722892.2804297
58,1073645.325971,695832.7214759
59,1161695.808163,704436.2043429
60,1178268.806814,697303.959875
61,1210350.349753,654378.5670674
end
```

Outcov.aml

```
kill sedsta all
generate sedsta  (cov1,cov2,...)
input newout.txt
points
q
build sedsta points
addxy sedsta
kill outgrid all
pointgrid sedsta outgrid (cov1grid,cov2grid,...)
500
y
nodata
&return
```

Rain.aml

```
/* This AML creates an INFO table with the
/* minimum, maximum and mean rainfall over
/* a watershed.  The values are annual
/* averages in mm based on a coverage
/* created by Oregon State University for
/* 1961-1990.
/*
/* Create a rainfall grid just larger than
/* the watershed.
grid
mape lavgrd
setwindow lavgrd
rain = annrain
kill annrain all
/*
/* Convert the rainfall grid into a polygon
/* coverage for the purpose of plotting and
/* to insure all cells in the watershed will
/* have a rainfall value.
rainpoly = gridpoly(rain)
quit
/*
/* Convert back into a grid coverage and
/* compute statistics.
clip rainpoly txtmpa rainclip
polygrid rainclip raingrid grid-code
500
y
grid
rainstat = zonalstats(areagrid,raingrid)
kill rain all
kill rainpoly all
kill raingrid all
quit
&return
```

Rainll.dat

```
1 -96.67 28.98
2 -96.94 29.43
3 -97.14 29.29
4 -98.30 30.00
5 -96.00 28.88
end
```

Rainname.dat

```
1 EDNA
2 HALLETTESVILLE
3 YOAKUM
end
```

Resleng.dat

```
/* on June 24, 1996, aml's were accidentally deleted
/*
/* the following code is used to determined the
/* flowlength from a dam to a sediment gauging station
/*
grid
flowdir%i% = selectmask(txmfd,shed%i%)
downgrd%i% = flowlength(flowdir%i%,#,DOWNSTREAM)
/*
/*
mape poly%i%
arcs poly%i%
arcs resnar
/*
/* zoom in on reservoirs of interest
/*
gridpaint dams
/* if some cells are colored black, then
/*  points damptspr
cellvalue dams *
cellvalue downgrd%i%
/*
/* copy results to text file and bring into a spreadsheet
&return
```

Resshed.aml

```
/*
/* this aml locates outlets and watersheds
/*  for all of the reservoirs in Texas
/*
polygrid resrvoir damgrid res_num
500
y
grid
```

```
outgrid = zonalmax(damgrid,flowacc)
damout = con(outgrid == flowacc,damgrid)
/* convert to point coverage to move to streams
dampts = gridpoint(damout,res_num)
/* once reservoir outlets are moved to the stream
/* then the corresponding watersheds can be computed
damshed = watershed(txmfd,damout)
kill damgrid all
dampoly = gridpoly(damshed)
quit
&return
```

Sastable.aml

```
/* this aml calculates values for input
/* into SAS multivariate regression model
grid
rainstat5 = zonalstats(shedcov5,raingrid)
rainsta5 = zonalstats(cov5grid,raingrid,max)
slopstat5 = zonalstats(shedcov5,slopgrid)
elevstat5 = zonalstats(shedcov5,dem)
strmfreq5 = zonalstats(shedcov5,txst,sum)
quit
tables
select rainstat5
alter mean,rain_avg,,,,,
alter min,rain_min,,,,,
alter max,rain_max,,,,,
alter value,grid-code,,,,,
select rainsta5
alter max,rain_sta,,,,,
alter value,grid-code,,,,,
select slopstat5
alter mean,slop_avg,,,,,
alter min,slop_min,,,,,
alter max,slop_max,,,,,
alter value,grid-code,,,,,
select elevstat5
alter mean,elev_avg,,,,,
alter min,elev_min,,,,,
alter max,elev_max,,,,,
alter value,grid-code,,,,,
select strmfreq5
alter sum,streams,,,,,
alter value,grid-code,,,,,
quit
joinitem cov5shed.pat rainstat5 sasstat5 grid-code grid-code
joinitem sasstat5 rainsta5 sasstat5 grid-code grid-code
joinitem sasstat5 slopstat5 sasstat5 grid-code grid-code
joinitem sasstat5 elevstat5 sasstat5 grid-code grid-code
joinitem sasstat5 strmfreq5 sasstat5 grid-code grid-code
&return
```

Sedshed.aml

```
kill outgrid all
copy /res3/jcoonrod/task3/outgrid (cov1grid,cov2grid,...)
kill shedgrid all
grid
shedgrid = watershed(txmfd,outgrid)
/*shedcov1                    cov1grid
kill sedshed all
sedshed = gridpoly(shedgrid)
/*cov1shed = gridpoly(shedcov1)
/*
quit
&return
```

Selout.aml

```
/*** This AML (written by Seann Reid, UT Ph.D. candidate, 1994,
edited by author)
/*** paints the vicinity of outlet locations in a point
/*** coverage so that the user can select the outlet cell from the
/*** streamlink grid which is closest to that point as a watershed
outlet.
/*** Several new graphics windows are created.  The number of
outlet
/*** locations that can be selected in one execution is influenced
by the
/*** number of new windows that can fit on the screen.
/*** Modified to write a point file of coordinates to be used in
/*** watershed delineation.

/* Initiate AML from GRID

/*&args linkgrid outlets outfile
&sv linkgrid = txstrms
&sv outlets = cov1grid
&sv outfile = outlet2.txt
&type running selout.aml

/** INITIALIZE MARKER PARAMETERS
markersize 3
markercolor 2
markersymbol 2
/*&messages &on
/*&messages &off &all
&if [iteminfo %outlets% -point X-COORD -exists] = .FALSE. &then
  &do
  &sys arc addxy %outlets% point
  &end

/*grid
```

```
&if [extract 1 [show display]] ne 9999 &then
  &do
   display 9999
  &end
/*&type got here0
mape %linkgrid%
/*describe %linkgrid%
/*&type got herel
/*&sv cellsize = %grd$dx%
&sv cellsize = 500
mapunits meters
units map
&sv cellrange = 10.0

&type got here2
&sv startrec = [ RESPONSE 'Type a record number to start with ?' ]
&sv cntl = 1

/*&sv idstring = [ QUOTE %outlets%# ]
&sv end_of_points = .FALSE.
cursor out_cur declare %outlets%.pat info ro

/*** OPEN THE CURSOR THAT IS POINTING TO THE FIRST RECORD
cursor out_cur open
&do &while %cnt1% lt %startrec%
  cursor out_cur next
  &sv cntl = %cnt1% + 1
&end

&sv count = %startrec%

/*** Open the output file for writing.

&sv wfunit = [open %outfile% openstat -append]
  &if %openstat% ne 0 &then
     &do
        &type openstat = %openstat%
        &stop Cannot open the output file %outfile%
     &end
  &else &type File %outfile% opened succussfully for writing.

/*** Processing loop ***
&do &until %end_of_points% = .TRUE.
   /*&sv count = %:out_cur.%idstring%%
   &sv x = %:out_cur.X-COORD%
   &sv y = %:out_cur.Y-COORD%

   &sv xmin = ( %x% - %cellrange% * %cellsize% )
   &sv xmax = ( %x% + %cellrange% * %cellsize% )
   &sv ymin = ( %y% - %cellrange% * %cellsize% )
   &sv ymax = ( %y% + %cellrange% * %cellsize% )
```

```
   mape %xmin% %ymin% %xmax% %ymax%
   &type %xmin% %xmax% %ymin% %ymax%
   &type %x% %y%
   gridshades %linkgrid%
   /*arcs %linkgrid%
   points %outlets%
   /*arcs %outlets%
   arcs roads
   arcs resrvoir
   /*&sv mcolor = [ UNQUOTE [RESPONSE 'Enter markercolor #' ] ]
   /*markercolor %mcolor%
   points %outlets%
   &sv selyn = [ RESPONSE 'Select an outlet cell? (y/n)' ]
   &if ( %selyn% = 'y' ) &then
     &do
       &type Select an outlet cell.
       where
       &sv xloc = [ extract 1 [ show where ] ]
       &sv yloc = [ extract 2 [ show where ] ]
       &sv comma = ,
       &type %count%%comma%%xloc%%comma%%yloc%
       &sv outline = %count%%comma%%xloc%%comma%%yloc%

/* Write the number of sub-watersheds being processed to the output
file.
       &if [ write %wfunit% %outline% ] ne 0 &then
          &do
            &type Error in writing to output file.  Exiting AML.
            &return
          &end
        /*  &sv count = %count% + 1
      &end
  &sv count = %count% + 1
  &sv yesno = [ RESPONSE 'Do you wish to continue (y/n)?' ]
  &if ( %yesno% = 'n' ) &then
    &do
      &type Stopping execution.
      /*&sv end_of_points = .TRUE.
      /*close the write file unit
        &if [close %wfunit%] = 0 &then
          &type %outfile% closed successfully
        cursor out_cur remove
        &messages &on
      &return
    &end
/* In previous version, I created a new graphics window for each
/* outlet but of course this won't work with 185 potential
/* watersheds.

/*   &type %xmin% %ymin% %xmax% %ymax%
```

```
&if [extract 1 [show display]] ne 9999 &then
  &do
   display 9999
  &end
/*&type got here0
mape %linkgrid%
/*describe %linkgrid%
/*&type got here1
/*&sv cellsize = %grd$dx%
&sv cellsize = 500
mapunits meters
units map
&sv cellrange = 10.0

&type got here2
&sv startrec = [ RESPONSE 'Type a record number to start with ?' ]
&sv cntl = 1

/*&sv idstring = [ QUOTE %outlets%# ]
&sv end_of_points = .FALSE.
cursor out_cur declare %outlets%.pat info ro

/*** OPEN THE CURSOR THAT IS POINTING TO THE FIRST RECORD
cursor out_cur open
&do &while %cntl% lt %startrec%
  cursor out_cur next
  &sv cntl = %cntl% + 1
&end

&sv count = %startrec%

/*** Open the output file for writing.

&sv wfunit = [open %outfile% openstat -append]
  &if %openstat% ne 0 &then
     &do
        &type openstat = %openstat%
        &stop Cannot open the output file %outfile%
     &end
  &else &type File %outfile% opened succussfully for writing.

/*** Processing loop ***
&do &until %end_of_points% = .TRUE.
   /*&sv count = %:out_cur.%idstring%%
   &sv x = %:out_cur.X-COORD%
   &sv y = %:out_cur.Y-COORD%

   &sv xmin = ( %x% - %cellrange% * %cellsize% )
   &sv xmax = ( %x% + %cellrange% * %cellsize% )
   &sv ymin = ( %y% - %cellrange% * %cellsize% )
   &sv ymax = ( %y% + %cellrange% * %cellsize% )
```

```
   mape %xmin% %ymin% %xmax% %ymax%
   &type %xmin% %xmax% %ymin% %ymax%
   &type %x% %y%
   gridshades %linkgrid%
   /*arcs %linkgrid%
   points %outlets%
   /*arcs %outlets%
   arcs roads
   arcs resrvoir
   /*&sv mcolor = [ UNQUOTE [RESPONSE 'Enter markercolor #' ] ]
   /*markercolor %mcolor%
   points %outlets%
   &sv selyn = [ RESPONSE 'Select an outlet cell? (y/n)' ]
   &if ( %selyn% = 'y' ) &then
     &do
       &type Select an outlet cell.
       where
       &sv xloc = [ extract 1 [ show where ] ]
       &sv yloc = [ extract 2 [ show where ] ]
       &sv comma = ,
       &type %count%%comma%%xloc%%comma%%yloc%
       &sv outline = %count%%comma%%xloc%%comma%%yloc%

/* Write the number of sub-watersheds being processed to the output
file.
       &if [ write %wfunit% %outline% ] ne 0 &then
          &do
            &type Error in writing to output file.  Exiting AML.
            &return
          &end
         /*  &sv count = %count% + 1
      &end
  &sv count = %count% + 1
  &sv yesno = [ RESPONSE 'Do you wish to continue (y/n)?' ]
  &if ( %yesno% = 'n' ) &then
    &do
      &type Stopping execution.
      /*&sv end_of_points = .TRUE.
      /*close the write file unit
        &if [close %wfunit%] = 0 &then
          &type %outfile% closed successfully
        cursor out_cur remove
        &messages &on
      &return
    &end
/* In previous version, I created a new graphics window for each
/* outlet but of course this won't work with 185 potential
/* watersheds.

/*   &type %xmin% %ymin% %xmax% %ymax%
```

```
/*   &if %count% eq 1 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL DISPLAY UR
/*   &if %count% eq 2 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL WINDOW win1 LL
/*   &if %count% eq 3 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL DISPLAY LL
/*   &if %count% eq 4 &then
/*       windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL WINDOW win3 UR

cursor out_cur next
&if %:out_cur.AML$NEXT% = .FALSE. &then
  &do
    &sv end_of_points = .TRUE.
    cursor out_cur remove
/*    &type %end_of_points%
  &end

/*Clear screen before moving to next point.
clear
/*&sv count = %count% + 1
/*End of Main Processing Loop
&end

&if [close %wfunit%] = 0 &then
        &type %outfile% closed successfully
&messages &on
/*&type quitting grid
/*q

&return
```

Selout1.aml

```
/*** This AML (written by Seann Reid, UT Ph.D. candidate, 1994,
edited by author)
/*** paints the vicinity of outlet locations in a point
/*** coverage so that the user can select the outlet cell from the
/*** streamlink grid which is closest to that point as a watershed
outlet.
/*** Several new graphics windows are created.  The number of
outlet
/*** locations that can be selected in one execution is influenced
by the
/*** number of new windows that can fit on the screen.
/*** Modified to write a point file of coordinates to be used in
/*** watershed delineation.

/* Initiate AML from GRID
```

```
/*&args linkgrid outlets outfile
&sv linkgrid = txstlm
&sv outlets = resrvoir
&sv outfile = outlets.txt
&type running selout.aml

/** INITIALIZE MARKER PARAMETERS
markersize 3
markercolor 2
markersymbol 2
/*&messages &on
/*&messages &off &all
&if [iteminfo %outlets% -point X-COORD -exists] = .FALSE. &then
  &do
  &sys arc addxy %outlets% point
  &end

/*grid

&if [extract 1 [show display]] ne 9999 &then
  &do
   display 9999
  &end
/*&type got here0
mape %linkgrid%
/*describe %linkgrid%
/*&type got here1
/*&sv cellsize = %grd$dx%
&sv cellsize = 500
mapunits meters
units map
&sv cellrange = 10.0

&type got here2
&sv startrec = [ RESPONSE 'Type a record number to start with ?' ]
&sv cntl = 1

/*&sv idstring = [ QUOTE %outlets%# ]
&sv end_of_points = .FALSE.
cursor out_cur declare %outlets%.pat info ro

/*** OPEN THE CURSOR THAT IS POINTING TO THE FIRST RECORD
cursor out_cur open
&do &while %cntl% lt %startrec%
  cursor out_cur next
  &sv cntl = %cntl% + 1
&end

&sv count = %startrec%
```

```
/*** Open the output file for writing.

&sv wfunit = [open %outfile% openstat -append]
  &if %openstat% ne 0 &then
     &do
        &type openstat = %openstat%
        &stop Cannot open the output file %outfile%
     &end
  &else &type File %outfile% opened succussfully for writing.

/*** Processing loop ***
&do &until %end_of_points% = .TRUE.
   /*&sv count = %:out_cur.%idstring%%
   &sv x = %:out_cur.X-COORD%
   &sv y = %:out_cur.Y-COORD%

   &sv xmin = ( %x% - %cellrange% * %cellsize% )
   &sv xmax = ( %x% + %cellrange% * %cellsize% )
   &sv ymin = ( %y% - %cellrange% * %cellsize% )
   &sv ymax = ( %y% + %cellrange% * %cellsize% )

   mape %xmin% %ymin% %xmax% %ymax%
   &type %xmin% %xmax% %ymin% %ymax%
   &type %x% %y%
   gridshades %linkgrid%
   /*arcs %linkgrid%
   points %outlets%
   &sv mcolor = [ UNQUOTE [RESPONSE 'Enter markercolor #' ] ]
   markercolor %mcolor%
   points %outlets%
   &sv selyn = [ RESPONSE 'Select an outlet cell? (y/n)' ]
   &if ( %selyn% = 'y' ) &then
     &do
       &type Select an outlet cell.
       where
       &sv xloc = [ extract 1 [ show where ] ]
       &sv yloc = [ extract 2 [ show where ] ]
       &sv comma = ,
       &type %count%%comma%%xloc%%comma%%yloc%
       &sv outline = %count%%comma%%xloc%%comma%%yloc%

/* Write the number of sub-watersheds being processed to the output
file.
        &if [ write %wfunit% %outline% ] ne 0 &then
           &do
             &type Error in writing to output file.  Exiting AML.
             &return
           &end
         /*  &sv count = %count% + 1
       &end
   &sv count = %count% + 1
```

```
 &sv yesno = [ RESPONSE 'Do you wish to continue (y/n)?' ]
 &if ( %yesno% = 'n' ) &then
  &do
   &type Stopping execution.
   /*&sv end_of_points = .TRUE.
   /*close the write file unit
    &if [close %wfunit%] = 0 &then
     &type %outfile% closed successfully
    cursor out_cur remove
    &messages &on
   &return
  &end
/* In previous version, I created a new graphics window for each
/* outlet but of course this won't work with 185 potential
/* watersheds.

/*   &type %xmin% %ymin% %xmax% %ymax%
/*   &if %count% eq 1 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL DISPLAY UR
/*   &if %count% eq 2 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL WINDOW win1 LL
/*   &if %count% eq 3 &then
/*     windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL DISPLAY LL
/*   &if %count% eq 4 &then
/*      windows create win%count% %xmin% %ymin% %xmax% %ymax% ~
/*        SIZE 350 350 POS UL WINDOW win3 UR

cursor out_cur next
&if %:out_cur.AML$NEXT% = .FALSE. &then
 &do
  &sv end_of_points = .TRUE.
  cursor out_cur remove
/*   &type %end_of_points%
 &end

/*Clear screen before moving to next point.
clear
/*&sv count = %count% + 1
/*End of Main Processing Loop
&end

&if [close %wfunit%] = 0 &then
        &type %outfile% closed successfully
&messages &on
/*&type quitting grid
/*q

&return
```

soils.aml

```
/*clip statsgo poly1 soilclip
tables
additem soilclip.pat A_area 8 18 f 5
additem soilclip.pat B_area 8 18 f 5
additem soilclip.pat C_area 8 18 f 5
additem soilclip.pat D_area 8 18 f 5
select soilclip.pat
calculate A_area = area * A-pct / 100
calculate B_area = area * B-pct / 100
calculate C_area = area * C-pct / 100
calculate D_area = area * D-pct / 100
/*statistics soilstat1
/*sum A_area
/*sum B_area
/*sum C_area
/*sum D_area
/*end
/*no
/*no
quit
&return

/* copy and clip covarages for individual watershed
/* soils
copy /mnt/thesis/tx/statsgo
clip statsgo <watershed polygon coverage> soilclip
kill statsgo

tables
additem soilclip.pat A_area 8 18 f 5
additem soilclip.pat B_area 8 18 f 5
additem soilclip.pat C_area 8 18 f 5
additem soilclip.pat D_area 8 18 f 5
select soilclip.pat
calculate A_area = area * A-pct / 100
calculate B_area = area * B-pct / 100
calculate C_area = area * C-pct / 100
calculate D_area = area * D-pct / 100
statistics
sum A_area
sum B_area
sum C_area
sum D_area

no
no
quit
&return
```

Task3.aml

```
/* LAND USE
copy /mnt/thesis/tx/txlus
clip txlus txtmpa luclip
kill txlus

polgrid txtmpa lavgrd
500
y
polygrid luclip lugrid level1
500
y

grid
areagrid = lavgrd / 2
lusum = areagrid * lugrid

kill luclip all
kill lugrid all

/* look at land use VAT (LUSUM.VAT) for totals in each land use
code

/* FLOW PATH
copy /mnt/thesis/tx/txdir
grid
mape lavgrd
setwindow lavgrd
flowdir = txdir
kill txdir all
upgrid = flowlength (flowdir, #, UPSTREAM)
downgrid = flowlength (flowdir, #, DOWNSTREAM)
flowlen = upgrid + downgrid
hydlen = zonalmax(lavgrd, flowlen)
&return
```

Task5.aml

```
/* Lavaca watershed only
/*
/* generate point coverage of rainfall stations
generate stations
input rainll.dat
points
quit
build stations points
addxy stations
/* project cover stations albstats albprj
project cover stations fortheis albprj
/* several bogus points were added to stations
/* so that theissen polygons would extend over
/* the watershed
```

```
kill stations all
thiessen fortheis theis
clip theis txtmpa thclip
kill theis all
kill fortheis all
/* create grid for each theis polygon
polygrid thclip thgrid thclip-id
500
y
/* locate centroids and zone grids for each polygon
grid
thcen = zonalcentroid(thgrid)
zone1 = con(thgrid == 1, 1)
zone2 = con(thgrid == 2, 1)
zone3 = con(thgrid == 3, 1)
/* should thgrid be thcen?
/* determine land uses in each zone
luedna = zone1 * lugrid
luhall = zone2 * lugrid
luyoak = zone3 * lugrid
/* look at VAT and count cells
/*
/* determine soils
tables
select soilclip.pat
statistic thclip-id thsoilstat
 sum A_area
 sum B_area
 sum C_area
 sum D_area
/*statistics in file thsoilstat
/*
/* hyd length from centroid to sediment station
grid
 cen1 = zone1 * upgrid
 cen2 = zone2 * upgrid
 cen3 = zone3 * upgrid
&return
```

Usgs_no.txt

```
1      7299200     Red_R.      Lakeview
2      7308000     Pease_R.    Crowell
3      7331600     Red_R.      Denison
4      7336820     Red_R.      DeKalb
5      7342500     S._Sulphur_R.     Cooper
6      7343200     Sulphur_R.  Talco
7      7343500     Whiteoak_Creek    Talco
8      8022000     Sabine_R.   Tatum
9      8022500     Sabine_R.   Longsport,LA
10     8031200     Kickapoo_Creek    Brownsboro
11     8033000     Neches_R.   Diboll
```

```
12      8033300      Piney_Creek Groveton
13      8033500      Neches_R.   Rockland
14      8037050      Bayou_La_Nana     Nacogdoches
15      8038500      Angelina_R. Zavalla
16      8039500      Angelina_R. Horger
17      8052700      Little_Elm_Cr     Aubrey
18      8062500      Trinity_R.  Rosser
19      8064500      Chambers_Creek    Corsicana
20      8065350      Trinity_R.  Crockett
21      8066200      Long_King_Creek   Livingston
22      8066500      Trinity_R.  Romayor
23      8068000      West_Fork_S.J.    Conroe
24      8069500      West_Fork_S.J.    Humble
25      8070000      East_Fork_S.J.    Cleveland
26      8071500      San_Jacinto_R.    Huffman
27      8080500      Dbl_Mtn_Fork      Aspermont
28      8082500      Brazos_R.   Seymour
29      8084800      California_Cr     Stamford
30      8085500      Clear_Fork  Griffin
31      8087300      Clear_Fork  Eliasville
32      8088000      Brazos_R.   South_Bend
33      8088600      Possum_King_Res   Graford
34      8093500      Aquilla_Creek     Aquilla
35      8094800      North_Bosque_R.   Hico
36      8100500      Leon_Ri.    Gatesville
37      8102500      Leon_Ri.    Belton
38      8104500      Little_Ri.  Little_River
39      8109900      Somerville_Lake   Somerville
40      8110000      Yegua_Creek Somerville
41      8110500      Navosta_R.  Easterly
42      8114000      Brazos_R.   Richmond
43      8146000      San_Saba_R. San_Saba
44      8147000      Colorado_R. San_Saba
45      8148000      Lake_Buchanan     Burnet
46      8148090      Colorado_R. Inks_Dam
47      8151500      Llano_R.    Llano
48      8153500      Pedernales_R.     Johnson_City
49      8158000      Colorado_R. Austin
50      8164000      Lavaca_River      Edna
51      8164300      Navidad_R.  Hallettsville
52      8164500      Navidad_R.  Ganado
53      8167500      Guadalupe_R.      Spring_Branch
54      8176500      Guadalupe_R.      Victoria
55      8183500      San_Antonio_R.    Falls_City
56      8186000      Cibolo_Creek      Falls_City
57      8188500      San_Antonio_R.    Goliad
58      8194000      Nueces_R.   Cotulla
59      8207000      Frio_River  Calliham
60      8210000      Nueces_River      Three_Rivers
61      8210500      Lake_Corp_Chr     Mathis
end
```

Usgsno.txt

```
1       7299200
2       7308000
3       7331600
4       7336820
5       7342500
6       7343200
7       7343500
8       8022000
9       8022500
10      8031200
11      8033000
12      8033300
13      8033500
14      8037050
15      8038500
16      8039500
17      8052700
18      8062500
19      8064500
20      8065350
21      8066200
22      8066500
23      8068000
24      8069500
25      8070000
26      8071500
27      8080500
28      8082500
29      8084800
30      8085500
31      8087300
32      8088000
33      8088600
34      8093500
35      8094800
36      8100500
37      8102500
38      8104500
39      8109900
40      8110000
41      8110500
42      8114000
43      8146000
44      8147000
45      8148000
46      8148090
47      8151500
48      8153500
49      8158000
```

```
50     8164000
51     8164300
52     8164500
53     8167500
54     8176500
55     8183500
56     8186000
57     8188500
58     8194000
59     8207000
60     8210000
61     8210500
end
```

8.3 *Appendix C – SAS Programs*

The CD available for this research contains all of the SAS programs written for this research. The programs are in three different sub-directories: Texas, LavUSGS, and Lavaca corresponding with the work presented in Chapters 4, 5, and 6, respectively. This appendix contains respresentative programs written for this research.

****

```
*Filename: /Texas/best.sas;
*Author:  Julie Coonrod;
*Date:  July, 1997;

*This model compares the "best" model between
sets of data;

options pagesize=60 linesize=80;
 libname avedata '.';

data a;
 set avedata.averages;
* set alldata.annual;
*set avedata.avenodam;
 wtfac=1/no_yrs;
 lnrelf = log(relief_m);
*if noyrs < 10 then kgload='.';
*if noyrs < 10 then sedconc='.';
*exclude stations with reservoirs;
*if dams='yes' then kgload='.';
*if dams='yes' then sedconc='.';
*if dams='yes' then lnloadkg='.';
*if dams='yes' then lnconc='.';
*to remove bias of stations with more years;
*if year<70 then kgload='.';
*if year<70 then sedconc='.';
*if year<70 then lnloadkg='.';
*if year<70 then lnconc='.';
*if year>78 then kgload='.';
*if year>78 then sedconc='.';
*if year>78 then lnloadkg='.';
*if year>78 then lnconc='.';
```

```
* the following stations do not have complete records from 70-78;
*if sta_no = 8033300 then kgload='.';
*if sta_no = 8033300 then sedconc='.';
*if sta_no = 8033300 then lnloadkg='.';
*if sta_no = 8033300 then lnconc='.';
*if sta_no = 8037050 then kgload='.';
*if sta_no = 8037050 then sedconc='.';
*if sta_no = 8037050 then lnloadkg='.';
*if sta_no = 8037050 then lnconc='.';
*if sta_no = 8031200 then kgload='.';
*if sta_no = 8031200 then sedconc='.';
*if sta_no = 8031200 then lnloadkg='.';
*if sta_no = 8031200 then lnconc='.';
*if sta_no = 8164500 then kgload='.';
*if sta_no = 8164500 then sedconc='.';
*if sta_no = 8164500 then lnloadkg='.';
*if sta_no = 8164500 then lnconc='.';
run;

*proc sort;
* by climate;
*run;

proc reg;
* weight wtfac;
* by climate;
 model lnloadkg = lnarea lnleng lnromm
                  lnelev lnresvar;
 model lnloadkg = lnarea lnleng lnromm slop_max
                  relief_m lnresvar;
 model lnloadkg = lnarea lnleng lnromm
                  relief_m lnresvar;
 model lnloadkg = lnarea lnleng lnromm
                  lnrelf lnresvar;
 model lnconc = k_max
                soilA lnelev lnrain
                lnresvar;
 model lnconc = lnresvar lnelev bare
                USGSslp k_max;
 model lnconc = lnresvar lnelev k_max;
 model sedconc = relief_m soilB elev_avg
                 rain_avg time
                 resvar;
 model kgload=flowm3 rain_avg leng_km;
 model kgload = flowm3 area length resvar;
run;
```

****

```
*Filename: /Texas/alldata.sas;
*Author:  Julie Coonrod;
*Date:  May, 1996;

*This model considers all of the data for each sediment station in
the state of Texas;
*The SAS program is used to narrow data ranges according to that
of interest;

options pagesize=60 linesize=80;
 libname alldata '.';

data a;
 set alldata.annual;
*exclude stations with reservoirs;
*if dams='yes' then annload='.';
* define variables;
* invres=1/resvar;
run;

proc sort;
 by sta_no;
proc chart;
  hbar sta_no / type = mean sumvar=annload;
  hbar year / type = freq sumvar=year;
  hbar sta_no / group=year;
run;

proc means maxdec=0;
  by sta_no;
run;

proc summary;
  class sta_no;
  var annload annflow;
run;

*proc reg;
* determine individual relationships;
* flowrate;
 *model annload=annflow;
 *model lnload=lnflow;
 *model lnconc=lnflow;
* *model lnconc=lnflow lnresvar;
 *model normload=flowm3;
 *model lnnormld=lnflowm3;
 *model annload=runoffin;
 *model normload=runoffmm;
* area;
 *model annload=area_mi;
 *model normload=area_km;
```

```
 *model sedconc=area_km;
 *model annload=rain_avg rain_sta rainvar annrain
                        / selection = stepwise;
 *model sedconc=rain_avg rain_sta rainvar annrain
                       / selection = stepwise;
* misc and multivariate;
 *model kgload=area length;
* *model kgload=resvar;
* *model kgload=resarea;
* *model kgload=invres;
* *model kgload=area resarea;
* *model kgload=area length resvar resarea invres
             / selection = stepwise;
 *model kgload=area time;
 *model normload=area rain_avg;
 *model annload=soilD urb;
 *model sedconc=ers rainers2 annrain rain_avg;
 *model kgload=lengflow flowm3 length;
 *model kgload=flowm3 resvar;
 *model annload=area resvar;
* *model sedconc=slop_avg slope rain_avg ers urb ag flowm3
               resvar invres / selection = stepwise;
* *model annload=area length time rain_avg rainvar ers
               slop_avg slope ag resvar invres
              / selection = stepwise;
* *model lnload=lnflow lnresvar;
 *model lnload=lnflow;
 *model lnconc=lnflow;
* *model lnconc=lnflow lnresvar;
* *model lnconc=lnflow lnresvar lnflowsq;
* *model lnconc=lnflow lnresvar ers;
* *model sedconc=annflow resvar ers;
 *model kgload=flowm3 area length strmfreq
              slop_max annrain rainvar ers
              ag urb
              / selection = stepwise;
 *model sedconc=flowm3 area length strmfreq
               slop_max annrain rainvar ers
               ag urb
               / selection = stepwise;
 *model kgload=flowm3 area length streams
              slop_avg annrain rain_avg ers
              ag urb
              / selection = stepwise;
 *model sedconc=flowm3 strmfreq slope
               slop_max annrain rainvar ers
               ag urb
               / selection = stepwise;
 *model sedconc=ers;
run;
```

```
proc reg;
* determine individual relationships;
* flowrate;
 *model annload=annflow;
 *model lnload=lnflow;
 *model annload=annflow area length ers ag rain_avg
            USGSslp / selection = stepwise;
 *model lnload=lnflow lnarea lnleng lners lnag lnrain
             lnslope / selection = stepwise;
 *model sedconc=flowm3 area length strmfreq
              slop_max annrain rainvar ers
              ag urb
              / selection = stepwise;
 *model kgload=flowm3 area length streams
             slop_avg annrain rain_avg ers
             ag urb
             / selection = stepwise;
 *model sedconc=flowm3 strmfreq USGSslp
              slop_max annrain rainvar ers
              ag urb
              / selection = stepwise;
**models with interaction terms;
 *model sedconc = rain_avg ers rainers2;
 *model sedconc = flow length lengflow;
 *model sedconc = flow USGSslp slopflow;
*run;
```

****

```
*Filename: /Texas/alldata1.sas;
*Author:  Julie Coonrod;
*Date:  May, 1996;

*This model considers all of the data for each sediment station in
the
state of Texas;
*The SAS program is used to narrow data ranges according to that
of interest;

options pagesize=60 linesize=80;

data a;
 infile 'alldata.dat';
 input sta_no 1-8 year 9-11 annload 12-20 annflow 21-29
       ncdc_no 30-36 annrain 37-42 area 43-55 perimetr 56-63
       length 64-71 k_avg 72-76 k_min 77-81 k_max 82-86
       streams 87-93
       elev_avg 94-98 elev_min 99-103 elev_max 104-108
       slop_avg 109-113 slop_min 114-118 slop_max 119-123
       rain_sta 124-128 rain_avg 129-133
       rain_min 134-138 rain_max 139-143
```

```
       lu0 144-147 lu1 148-151 lu2 152-155 lu3 156-159
       lu4 160-163 lu5 164-167 lu6 168-171 lu7 172-175
       soilA 176-179 soilB 180-183
       soilC 184-187 soilD 188-191 dams $ 192-195
       fracarea 196-201 resvar 202-207 USGSslp 208-213;
* later, may want to add resvar and resarea;
* missing values;
 if annload=-9999 then annload='.';
 if annflow=-9999 then annflow='.';
 if annrain=99999 then annrain='.';
*exclude stations with reservoirs;
 if dams='yes' then annload='.';
* change units;
 flowaf=annflow*86400/43560;
 sedpct=annload/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 kgload=annload*2000*0.4536;
 area_km=area/1000/1000;
 area_mi=area*3.28084*3.28084/5280/5280;
 leng_km=length/1000;
 flowm3=annflow*0.02832;
 peri_km=perimetr/1000;
* define variables;
 slope=(elev_max-elev_min)/length;
 rainvar=rain_max-rain_min;
 potq=area*rain_avg/1000;
 runoffin=flowaf*43560/5280/5280/area_mi*12;
 runoffmm=flowm3/area*1000;
 leng32=leng_km**1.5;
 relief=elev_max-elev_min;
 normload=kgload/area_km;
 time1=length/USGSslop;
 time2=length/USGSslop**0.5;
 shape1=area/length;
 shape2=area/perimetr;
 strmfreq=streams/area;
 urb=lu1;
 ag=lu2;
 veg=lu3+lu4;
 ers=soilC+soilD;
* interaction terms;
 rainers2=rain_avg*(1-ers);
 lengflow=length*flowm3;
* resflow=resvar*flowm3;
* interaction terms;
 rainers2=rain_avg*ers;
 lengflow=length*flowm3;
* take logarithms;
 lnflow=log(annflow);
```

```
 lnload=log(annload);
 lnconc=log(sedconc);
 lnnormld=log(normload);
 lnflowm3=log(flowm3);
 lnareami=log(area_mi);
 lnarea=log(area);
 lnleng=log(length);
 lntime=log(time2);
 lners=log(ers);
 lnag=log(ag);
 lnromm=log(runoffmm);
 lnelev=log(runoffmm);
 lnrain=log(annrain);
* lnresvar=log(resvar);
 lnflowsq=lnflow*lnflow;
run;

proc corr;
 var area perimetr length;
 var k_avg k_max ers soilC soilD;
 var k_max elev_max slop_avg slop_max USGSslop;
 var annrain rain_sta rain_avg rainvar;
 var urb ag veg ers slop_avg rain_avg;
 var rain_avg strmfreq;
 var annload annflow area length USGSslp;
 var annload annflow annrain rain_avg;
 var annload annflow ers urb ag strmfreq;
 var annload annflow slop_avg slope USGSslop time2;
run;
```

****

```
*Filename: /Texas/alldata.sas;
*Author:  Julie Coonrod;
*Date:  May, 1996;

*This *model considers all of the data for each sediment station in
the
state of Texas;
*The SAS program is used to narrow data ranges according to that
of interest;

options pagesize=60 linesize=80;
 libname alldata '.';

data a;
 set alldata.annual;
*exclude stations with reservoirs;
if dams='yes' then kgload='.';
if dams='yes' then sedconc='.';
```

```
if dams='yes' then normload='.';
if dams='yes' then lnloadkg='.';
if dams='yes' then lnconc='.';
if dams='yes' then lnnormld='.';
*to remove bias of stations with more years;
if year<70 then kgload='.';
if year<70 then sedconc='.';
if year<70 then normload='.';
if year<70 then lnloadkg='.';
if year<70 then lnconc='.';
if year<70 then lnnormld='.';
if year>78 then kgload='.';
if year>78 then sedconc='.';
if year>78 then normload='.';
if year>78 then lnloadkg='.';
if year>78 then lnconc='.';
if year>78 then lnnormld='.';
* the following stations do not have complete records from 70-78;
if sta_no = 8033300 then kgload='.';
if sta_no = 8033300 then sedconc='.';
if sta_no = 8033300 then normload='.';
if sta_no = 8033300 then lnloadkg='.';
if sta_no = 8033300 then lnconc='.';
if sta_no = 8033300 then lnnormld='.';
if sta_no = 8037050 then kgload='.';
if sta_no = 8037050 then sedconc='.';
if sta_no = 8037050 then normload='.';
if sta_no = 8037050 then lnloadkg='.';
if sta_no = 8037050 then lnconc='.';
if sta_no = 8037050 then lnnormld='.';
if sta_no = 8031200 then kgload='.';
if sta_no = 8031200 then sedconc='.';
if sta_no = 8031200 then normload='.';
if sta_no = 8031200 then lnloadkg='.';
if sta_no = 8031200 then lnconc='.';
if sta_no = 8031200 then lnnormld='.';
if sta_no = 8164500 then kgload='.';
if sta_no = 8164500 then sedconc='.';
if sta_no = 8164500 then normload='.';
if sta_no = 8164500 then lnloadkg='.';
if sta_no = 8164500 then lnconc='.';
if sta_no = 8164500 then lnnormld='.';
* categorize by size and rainfall;
if area<40972795270 then size='small';
if area>=40972795270 and area<81864523459
               then size='medium';
if area>=81864523459 then size='large';
if rain_avg<720 then climate='arid';
if rain_avg>=720 and rain_avg<966
            then climate='mod';
if rain_avg>=966 then climate='humid';
```

```
* define variables;
* invres=1/resvar;
run;

*proc sort;
* by climate;
* by year;
*run;

*proc print;
* by year;
* id sta_no;
* var basin area kgload arain_mm rain_sta;
*run;

proc reg;
* by year;
* by basin;
* by categories;
* by size;
* by climate;
* determine individual relationships;
* flowrate;
 model kgload=flowm3;
 model sedconc=flowm3;
 model normload=flowm3;
 model lnloadkg=lnflowm3;
 model lnconc=lnflowm3;
 model lnnormld=lnflowm3;
run;

proc reg;
* rainfall;
 model kgload=arain_mm;
 model kgload=rain_sta;
 model kgload=rain_avg;
 model sedconc=arain_mm;
 model sedconc=rain_sta;
 model sedconc=rain_avg;
 model normload=arain_mm;
 model normload=rain_sta;
 model normload=rain_avg;
 model lnloadkg=lnarain;
 model lnloadkg=lnrain;
 model lnconc=lnarain;
 model lnconc=lnrain;
 model lnnormld=lnarain;
 model lnnormld=lnrain;
* runoff;
 model sedconc=runoffmm;
 model sedconc=potq_m3;
```

```
 model kgload=runoffmm;
 model kgload=potq_m3;
 model normload=runoffmm;
 model normload=potq_m3;
 model lnconc=lnromm;
 model lnconc=lnpotq;
 model lnloadkg=lnromm;
 model lnloadkg=lnpotq;
 model lnnormld=lnromm;
 model lnnormld=lnpotq;
* area;
 model sedconc=area_km;
 model kgload=area_km;
 model normload=area_km;
 model lnconc=lnarea;
 model lnloadkg=lnarea;
 model lnnormld=lnarea;
* length;
 model sedconc=leng_km;
 model kgload=leng_km;
 model normload=leng_km;
 model lnconc=lnleng;
 model lnloadkg=lnleng;
 model lnnormld=lnleng;
* concentration variations;
 model sedconc=strmfreq;
 model sedconc=urb;
 model sedconc=ag;
 model sedconc=range;
 model sedconc=forest;
 model sedconc=veg;
 model sedconc=water;
 model sedconc=bare;
 model sedconc=soilA;
 model sedconc=soilB;
 model sedconc=soilC;
 model sedconc=soilD;
 model sedconc=ers;
 model sedconc=k_avg;
 model sedconc=k_max;
 model sedconc=USGSslp;
 model sedconc=slop_avg;
 model sedconc=slop_max;
 model sedconc=relief_m;
 model lnconc=lnag;
 model lnconc=lnveg;
 model lnconc=lners;
 model lnconc=lnslop;
 *model lnconc=lnfor;
 *model lnconc=lnkavg;
 *model lnconc=lnslopa;
```

```
 model kgload=flowm3 potq_m3 leng_km
              / selection = stepwise;
 model kgload=flowm3 rain_avg area_km leng_km
              / selection = stepwise;
 model kgload=flowm3 arain_mm area_km leng_km
              / selection = stepwise;
 model lnloadkg=lnflowm3 lnarain lnarea lnleng
                / selection = stepwise;
 model lnloadkg=lnflowm3 lnrain lnarea lnleng
                / selection = stepwise;
 model lnloadkg=lnflowm3 lnpotq lnarea lnleng
                / selection = stepwise;
 model lnnormld=lnflowm3 lnarain lnarea lnleng
                / selection = stepwise;
 model lnnormld=lnromm lnarain lnarea lnleng
                / selection = stepwise;
 model sedconc=range forest USGSslp relief_m
               / selection = stepwise;
 model lnconc= lnpotq lnag lnveg lners lnslop;
* (kitchen sink approach);
 model kgload = area_km peri_km leng_km
                          shapel shape2 strmfreq leng32
                          time k_avg elev_avg relief_m
                          slop_avg slop_max rain_sta
                 arain_mm rain_avg rainvar USGSslp urb
                          ag range forest veg bare ers
                          / selection = stepwise;
model sedconc = flowm3 area_km peri_km leng_km
                          shapel shape2 strmfreq leng32
                          time k_avg elev_avg relief_m
                          slop_avg slop_max rain_sta
                   arain_mm rain_avg rainvar USGSslp urb
                          ag range forest veg bare ers
                          / selection = stepwise;
model normload = flowm3 area_km peri_km leng_km
                          shapel shape2 strmfreq leng32
                          time k_avg elev_avg relief_m
                          slop_avg slop_max rain_sta
                          rain_avg rainvar USGSslp urb
                          ag range forest veg bare ers
                          / selection = stepwise;
model lnloadkg =  lnarea lnperi lnleng
                          lnshapel lnshape2
                          lntime  lnelev lnrelief
                          lnslop
                   lnarain lnrain lnrainv
                          lnag lnveg lners
                          / selection = stepwise;
model lnconc = lnflowm3 lnarea lnperi lnleng
                          lnshapel lnshape2
                          lntime  lnelev lnrelief
```

```
                           lnslop
                     lnarain lnrain lnrainv
                           lnag lnveg lners
                           / selection = stepwise;
model lnnormld =  lnarea lnperi lnleng
                           lnshape1 lnshape2
                           lntime  lnelev lnrelief
                           lnslop
                           lnrain lnrainv
                           lnag lnveg lners
                           / selection = stepwise;
* interaction terms;
model kgload = flowm3 flowsq lengflow slopflow
               length USGSslp /selection = stepwise;
model normload = flowm3 flowsq lengflow slopflow
                 length USGSslp / selection = stepwise;
model sedconc = flowm3 flowsq lengflow slopflow
                length USGSslp / selection = stepwise;
model lnloadkg = lnflowm3 lnflsq lnslflow lnleng
                 lnslop / selection = stepwise;
model lnconc = lnflowm3 lnflsq lnslflow lnleng
               lnslop / selection = stepwise;
model lnnormld = lnflowm3 lnflsq lnslflow lnleng
                 lnslop / selection = stepwise;
* ;
model kgload = rain_avg ers rainers rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rainers rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model kgload = rain_avg ers rainers2 rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rainers2 rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model kgload = rain_avg ers rain_k rainslp rainelev k_avg
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rain_k rainslp rainelev k_avg
```

```
                              rainurb rainag rainwet rainbare ersveg
                              ersslop
                              slop_avg elev_avg urb ag water bare veg
                  / selection = stepwise;
* model kgload=resvar;
* model kgload=resarea;
* model kgload=invres;
* model kgload=area resarea;
* model kgload=area length resvar resarea invres
             / selection = stepwise;
* model annload=area length time rain_avg rainvar ers
                slop_avg slope ag resvar invres
               / selection = stepwise;
* model lnload=lnflow lnresvar;
* model lnconc=lnflow lnresvar;
* model lnconc=lnflow lnresvar lnflowsq;
* model lnconc=lnflow lnresvar ers;
* model sedconc=annflow resvar ers;
run;
```

****

```
*Filename: /Texas/aaveone.sas;
*Author:  Julie Coonrod;
*Date:  May, 1996;

*This model considers all of the data for each sediment station in
the state of Texas;
*The SAS program is used to narrow data ranges according to that
of interest;

options pagesize=60 linesize=80;
 libname avedata '.';

data a;
 set avedata.average;
*exclude stations with reservoirs;
if dams='yes' then kgload='.';
if dams='yes' then sedconc='.';
if dams='yes' then normload='.';
if dams='yes' then lnloadkg='.';
if dams='yes' then lnconc='.';
if dams='yes' then lnnormld='.';
* categorize by size and rainfall;
if area<40972795270 then size='small';
if area>=40972795270 and area<81864523459
                    then size='medium';
if area>=81864523459 then size='large';
if rain_avg<720 then climate='arid';
if rain_avg>=720 and rain_avg<966
```

```
                then climate='mod';
if rain_avg>=966 then climate='humid';
* define variables;
* invres=1/resvar;
run;

proc sort;
 by climate;
run;

*proc print;
* by year;
* id sta_no;
* var basin area kgload arain_mm rain_sta;
*run;

proc reg;
* by basin;
* by categories;
* by size;
 by climate;
* determine individual relationships;
* flowrate;
 model kgload=flowm3;
 model sedconc=flowm3;
 model normload=flowm3;
 model lnloadkg=lnflowm3;
 model lnconc=lnflowm3;
 model lnnormld=lnflowm3;
* rainfall;
 model kgload=arain_mm;
 model kgload=rain_sta;
 model kgload=rain_avg;
 model sedconc=arain_mm;
 model sedconc=rain_sta;
 model sedconc=rain_avg;
 model normload=arain_mm;
 model normload=rain_sta;
 model normload=rain_avg;
 model lnloadkg=lnarain;
 model lnloadkg=lnrain;
 model lnconc=lnarain;
 model lnconc=lnrain;
 model lnnormld=lnarain;
 model lnnormld=lnrain;
* runoff;
 model sedconc=runoffmm;
 model sedconc=potq_m3;
 model kgload=runoffmm;
 model kgload=potq_m3;
 model normload=runoffmm;
```

```
 model normload=potq_m3;
 model lnconc=lnromm;
 model lnconc=lnpotq;
 model lnloadkg=lnromm;
 model lnloadkg=lnpotq;
 model lnnormld=lnromm;
 model lnnormld=lnpotq;
* area;
 model sedconc=area_km;
 model kgload=area_km;
 model normload=area_km;
 model lnconc=lnarea;
 model lnloadkg=lnarea;
 model lnnormld=lnarea;
* length;
 model sedconc=leng_km;
 model kgload=leng_km;
 model normload=leng_km;
 model lnconc=lnleng;
 model lnloadkg=lnleng;
 model lnnormld=lnleng;
* concentration variations;
 model sedconc=strmfreq;
 model sedconc=urb;
 model sedconc=ag;
 model sedconc=range;
 model sedconc=forest;
 model sedconc=veg;
 model sedconc=water;
 model sedconc=bare;
 model sedconc=soilA;
 model sedconc=soilB;
 model sedconc=soilC;
 model sedconc=soilD;
 model sedconc=ers;
 model sedconc=k_avg;
 model sedconc=k_max;
 model sedconc=USGSslp;
 model sedconc=slop_avg;
 model sedconc=slop_max;
 model sedconc=relief_m;
 model lnconc=lnag;
 model lnconc=lnveg;
 model lnconc=lners;
 model lnconc=lnslop;
 *model lnconc=lnfor;
 *model lnconc=lnkavg;
 *model lnconc=lnslopa;
 model kgload=flowm3 potq_m3 leng_km
              / selection = stepwise;
 model kgload=flowm3 rain_avg area_km leng_km
```

```
             / selection = stepwise;
model kgload=flowm3 arain_mm area_km leng_km
             / selection = stepwise;
model lnloadkg=lnflowm3 lnarain lnarea lnleng
               / selection = stepwise;
model lnloadkg=lnflowm3 lnrain lnarea lnleng
               / selection = stepwise;
model lnloadkg=lnflowm3 lnpotq lnarea lnleng
               / selection = stepwise;
model lnnormld=lnflowm3 lnarain lnarea lnleng
               / selection = stepwise;
model lnnormld=lnromm lnarain lnarea lnleng
               / selection = stepwise;
model sedconc=range forest USGSslp relief_m
              / selection = stepwise;
model lnconc= lnpotq lnag lnveg lners lnslop;
* (kitchen sink approach);
model kgload = area_km peri_km leng_km
                           shape1 shape2 strmfreq leng32
                           time k_avg elev_avg relief_m
                           slop_avg slop_max rain_sta
                  arain_mm rain_avg rainvar USGSslp urb
                           ag range forest veg bare ers
                           / selection = stepwise;
model sedconc = flowm3 area_km peri_km leng_km
                            shape1 shape2 strmfreq leng32
                            time k_avg elev_avg relief_m
                            slop_avg slop_max rain_sta
                     arain_mm rain_avg rainvar USGSslp urb
                            ag range forest veg bare ers
                            / selection = stepwise;
model normload = flowm3 area_km peri_km leng_km
                            shape1 shape2 strmfreq leng32
                            time k_avg elev_avg relief_m
                            slop_avg slop_max rain_sta
                            rain_avg rainvar USGSslp urb
                            ag range forest veg bare ers
                            / selection = stepwise;
model lnloadkg =  lnarea lnperi lnleng
                            lnshape1 lnshape2
                            lntime  lnelev lnrelief
                            lnslop
                     lnarain lnrain lnrainv
                            lnag lnveg lners
                            / selection = stepwise;
model lnconc = lnflowm3 lnarea lnperi lnleng
                            lnshape1 lnshape2
                            lntime  lnelev lnrelief
                            lnslop
                      lnarain lnrain lnrainv
                            lnag lnveg lners
```

```
                           / selection = stepwise;
model lnnormld =  lnarea lnperi lnleng
                           lnshape1 lnshape2
                           lntime  lnelev lnrelief
                           lnslop
                           lnrain lnrainv
                           lnag lnveg lners
                           / selection = stepwise;
* interaction terms;
model kgload = flowm3 flowsq lengflow slopflow
               length USGSslp /selection = stepwise;
model normload = flowm3 flowsq lengflow slopflow
                 length USGSslp / selection = stepwise;
model sedconc = flowm3 flowsq lengflow slopflow
                length USGSslp / selection = stepwise;
model lnloadkg = lnflowm3 lnflsq lnslflow lnleng
                 lnslop / selection = stepwise;
model lnconc = lnflowm3 lnflsq lnslflow lnleng
               lnslop / selection = stepwise;
model lnnormld = lnflowm3 lnflsq lnslflow lnleng
                 lnslop / selection = stepwise;
* ;
model kgload = rain_avg ers rainers rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rainers rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model kgload = rain_avg ers rainers2 rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rainers2 rainslp rainelev
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model kgload = rain_avg ers rain_k rainslp rainelev k_avg
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
               / selection = stepwise;
model sedconc= rain_avg ers rain_k rainslp rainelev k_avg
                          rainurb rainag rainwet rainbare ersveg
                          ersslop
                          slop_avg elev_avg urb ag water bare veg
```

```
                / selection = stepwise;
* model kgload=resvar;
* model kgload=resarea;
* model kgload=invres;
* model kgload=area resarea;
* model kgload=area length resvar resarea invres
             / selection = stepwise;
* model aveload=area length time rain_avg rainvar ers
                slop_avg slope ag resvar invres
               / selection = stepwise;
* model lnload=lnflow lnresvar;
* model lnconc=lnflow lnresvar;
* model lnconc=lnflow lnresvar lnflowsq;
* model lnconc=lnflow lnresvar ers;
* model sedconc=aveflow resvar ers;
run;



*Filename: /Texas/alldata.sas;
*Author:  Julie Coonrod;
*Date:  May, 1996;

*This *model considers all of the data for each sediment station in
the
state of Texas;
*The SAS program is used to narrow data ranges according to that
of interest;

options pagesize=60 linesize=80;
 libname alldata '.';

data a;
 set alldata.averages;
 if resvar=0 then resvar=0.001;
 invres=1/resvar;
run;

proc sort;
 by climate;
run;

proc print;
 by climate;
 id sta_no;
 var basin area kgload rain_avg rain_sta;
run;

proc reg;
* weight by noyears;
* by basin;
```

```
* by categories;
* by size;
 by climate;
* determine individual relationships;
* flowrate;
* model kgload=flowm3;
* model sedconc=flowm3;
* model normload=flowm3;
* model lnloadkg=lnflowm3;
* model lnconc=lnflowm3;
* model lnnormld=lnflowm3;
* rainfall;
* model kgload=arain_mm;
* model kgload=rain_sta;
* model kgload=rain_avg;
* model sedconc=arain_mm;
* model sedconc=rain_sta;
* model sedconc=rain_avg;
* model normload=arain_mm;
* model normload=rain_sta;
* model normload=rain_avg;
* model lnloadkg=lnarain;
* model lnloadkg=lnrain;
* model lnconc=lnarain;
* model lnconc=lnrain;
* model lnnormld=lnarain;
* model lnnormld=lnrain;
* runoff;
* model sedconc=runoffmm;
* model sedconc=potq_m3;
* model kgload=runoffmm;
* model kgload=potq_m3;
* model normload=runoffmm;
* model normload=potq_m3;
* model lnconc=lnromm;
* model lnconc=lnpotq;
* model lnloadkg=lnromm;
* model lnloadkg=lnpotq;
* model lnnormld=lnromm;
* model lnnormld=lnpotq;
* area;
* model sedconc=area_km;
* model kgload=area_km;
* model normload=area_km;
* model lnconc=lnarea;
* model lnloadkg=lnarea;
* model lnnormld=lnarea;
* length;
* model sedconc=leng_km;
* model kgload=leng_km;
* model normload=leng_km;
```

```
* model lnconc=lnleng;
* model lnloadkg=lnleng;
* model lnnormld=lnleng;
* concentration variations;
* model sedconc=strmfreq;
* model sedconc=urb;
* model sedconc=ag;
* model sedconc=range;
* model sedconc=forest;
* model sedconc=veg;
* model sedconc=water;
* model sedconc=bare;
* model sedconc=soilA;
* model sedconc=soilB;
* model sedconc=soilC;
* model sedconc=soilD;
* model sedconc=ers;
* model sedconc=k_avg;
* model sedconc=k_max;
* model sedconc=USGSslp;
* model sedconc=slop_avg;
* model sedconc=slop_max;
* model sedconc=relief_m;
* model lnconc=lnag;
* model lnconc=lnveg;
* model lnconc=lners;
* model lnconc=lnslop;
 *model lnconc=lnfor;
 *model lnconc=lnkavg;
 *model lnconc=lnslopa;
 model kgload=flowm3 potq_m3 leng_km
             / selection = stepwise;
 model kgload=flowm3 rain_avg area_km leng_km
             / selection = stepwise;
 model kgload=flowm3 arain_mm area_km leng_km
             / selection = stepwise;
 model lnloadkg=lnflowm3 lnarain lnarea lnleng
               / selection = stepwise;
 model lnloadkg=lnflowm3 lnrain lnarea lnleng
               / selection = stepwise;
 model lnloadkg=lnflowm3 lnpotq lnarea lnleng
               / selection = stepwise;
 model lnnormld=lnflowm3 lnarain lnarea lnleng
               / selection = stepwise;
 model lnnormld=lnromm lnarain lnarea lnleng
               / selection = stepwise;
 model sedconc=range forest USGSslp relief_m
              / selection = stepwise;
 model lnconc= lnpotq lnag lnveg lners lnslop;

run;
```

```
                  ********************************
*Filename: /Texas/avevals.sas;
*Author:  Julie Coonrod;
*Date:  April, 1996;

*This model considers average annual values for each sediment
station in the
state of Texas disregarding reservoirs in place during sampling
years;

options pagesize=60 linesize=80;

data a;
 infile 'avevals.dat';
 input con_no 1-6 sta_no 7-16 noyrs 17-22 aveload 23-32 aveflow 33-
42
       gridcode 43-48 area 49-60 perimetr 61-70
       length 71-80 k_avg 81-86 k_min 87-92 k_max 93-98
       streams 99-104
       elev_avg 105-110 elev_min 111-116 elev_max 117-122
       slop_avg 123-128 slop_min 129-134 slop_max 135-140
       rain_sta 141-146 rain_avg 147-152
       rain_min 153-158 rain_max 159-164
       lu0 165-170 lu1 171-176 lu2 177-182 lu3 183-188
       lu4 189-194 lu5 195-200 lu6 201-206 lu7 207-212
       soilA 213-218 soilB 219-224 soilC 225-230 soilD 231-236;
* remove averages that are not for at least a 10-year period;
 if noyrs<10 then aveload='.';
* change units;
 flowaf=aveflow*86400/43560;
 flowm3=flowaf*1233;
 sedpct=aveload/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 kgload=aveload*2000*0.4536;
 area_km=area/1000/1000;
 area_mi=area*3.28084*3.28084/5280/5280;
 leng_km=length/1000;
 peri_km=perimetr/1000;
 elevft=elev_avg*3.28084;
* define variables;
 slope=(elev_max-elev_min)/length;
 rainvar=rain_max-rain_min;
 potq=area*rain_avg/1000;
 rain_in=rain_avg/10/2.54;
 runoffin=flowaf*43560/5280/5280/area_mi*12;
 runoffmm=flowm3/area*1000;
 leng32=leng_km**1.5;
```

```
 relief=elev_max-elev_min;
 normload=kgload/area_km;
 normconc=sedconc/area_km;
 time1=length/slope;
 time2=length/slope**0.5;
 shape1=area/length;
 shape2=area/perimetr;
 strmfreq=streams/area;
 urb=lu1;
 ag=lu2;
 veg=lu3+lu4;
 ers=soilC+soilD;
* interaction terms;
 rainers2=rain_avg*ers;
 lengflow=length*flowm3;
* take logarithms;
 lnflow=log(aveflow);
 lnload=log(aveload);
 lnconc=log(sedconc);
 lnnormld=log(normload);
 lnflowm3=log(flowm3);
 lnareami=log(area_mi);
 lnareakm=log(area_km);
run;

proc reg;
* determine individual relationships;
* flowrate;
 model aveload=aveflow;
 model lnload=lnflow;
 model lnconc=lnflow;
* model normload=flowm3;
* model lnnormld=lnflowm3;
* model aveload=runoffin;
* model normload=runoffmm;
* area;
 model aveload=area_mi;
 model normload=area_km;
 model sedconc=area_km;
* landuse;
*model sedconc=urb veg ag lu7 / selection = stepwise;
* soil type;
* model sedconc=k_avg ers soilD soilC
                         / selection = stepwise;
 model sedconc=ers;
 model sedconc=rain_avg;
 model sedconc=k_avg;
 model sedconc=slope;
* model sedconc=slope slop_avg slop_max
                        / selection = stepwise;
 model sedconc=slop_avg;
```

```
* model aveload=rain_avg rain_sta rainvar
                        / selection = stepwise;
* model sedconc=rain_avg rain_sta rainvar
                       / selection = stepwise;
* misc and multivariate;
 model kgload=area length;
 model kgload=area time2;
 model normload=area rain_avg;
 model aveload=soilD urb;
 model sedconc=ers rainers2 annrain rain_avg;
 model kgload=lengflow flowm3 length;
* model sedconc=slop_avg slope rain_avg ers urb ag flowm3
              / selection = stepwise;
* model aveload=area length time2 rain_avg rainvar ers
               slop_avg slope ag / selection = stepwise;
run;
                  ********************************


* FILENAME:  LavUSGS/72chap5.SAS;
* AUTHOR:    JULIE COONROD;
* DATE:      FEBRUARY, 1996;

* DESCRIPTION:  This is a multivariate regression model relating
    sediment concentration to flow, size, and temperature
variables.
    As part of the USGS water quality sampling program, periodic
    water samples are taken.  The suspended sediment concentration
    is measured along with flow rate, %fines, and water
temperature.;

* SPECIFICS:  This model uses binary variables to describe size and
    temperature.  The use of binary variables enables plotting in
two dimensions and helps determine controlling processes.;

options pagesize=60 linesize=80;

data a;
 infile 'task4.dat';
 input date $ 1-9 time $ 10-18 flow 19-27 sedconc 28-36 sedld 37-45
                  size 46-54 temp 55-63 kvisc 64-74
                  fall 75-82 rise 83-90 base 91-98;
 load=flow*sedconc*86400/0.001/35.31/1000/1000/0.454/2000;
 lnload=log(load);
 lnflow=log(flow);
 lnconc=log(sedconc);
 if lnflow < 0 then lnflow = '.';
* If lnflow < 0 then lnflow + lnflowsq produces an incorrect
function
  shape.  Note that lnflow < 0 when flow < 1 cfs;
 if temp <= 20 then warm = 0;
   else warm = 1;
```

```
 if size <= 50 then large = 1;
   else large = 0;
 lnflowsq=lnflow*lnflow;
 largwarm=large*warm;
 flowwarm=lnflow*warm;
 flowlarg=lnflow*large;
 flowlgwm=lnflow*warm*large;
run;

proc reg;
 model lnload = lnflow;
 model lnconc = lnflow;
 model lnconc = lnflow lnflowsq;
 model lnconc = lnflow large;
 model lnconc = lnflow warm;
 model lnconc = lnflow warm large;
 model lnconc = lnflow warm large flowwarm flowlarg;
 model lnconc = lnflow lnflowsq large warm;
 model lnconc = lnflow large warm largwarm;
 model lnconc = lnflow lnflowsq large warm largwarm;
 model lnconc = lnflow large flowlarg;
 model lnconc = lnflow lnflowsq large flowlarg;
 model lnconc = lnflow warm flowwarm;
 model lnconc = lnflow lnflowsq warm flowwarm;
 model lnconc = lnflow large warm largwarm flowlarg flowwarm
flowlgwm;
 model lnconc = lnflow lnflowsq large warm largwarm flowlarg
flowwarm
                flowlgwm;
 model lnconc = lnflow lnflowsq large warm flowlarg flowwarm;
 model lnconc = lnflow lnflowsq large warm largwarm flowlarg
flowwarm
                flowlgwm;
 run;
                  ********************************

* FILENAME:  LavUSGS/bstchap5.SAS;
* AUTHOR:    JULIE COONROD;
* DATE:      FEBRUARY, 1996;

* DESCRIPTION:  This is a multivariate regression model relating
    sediment concentration to flow, size, and temperature
variables.
    As part of the USGS water quality sampling program, periodic
    water samples are taken.  The suspended sediment concentration
    is measured along with flow rate, %fines, and water
temperature.;

* SPECIFICS:  This model uses binary variables to describe size and
    temperature.  The use of binary variables enables plotting in
two
```

```
    dimensions and helps determine controlling processes.;

options pagesize=60 linesize=80;

data a;
 infile 'task4.dat';
 input date $ 1-9 time $ 10-18 flow 19-27 sedconc 28-36 sedld 37-45
                  size 46-54 temp 55-63 kvisc 64-74
                  fall 75-82 rise 83-90 base 91-98;
 lnflow=log(flow);
 lnconc=log(sedconc);
 if lnflow < 0 then lnflow = '.';
* If lnflow < 0 then lnflow + lnflowsq produces an incorrect
function
  shape.  Note that lnflow < 0 when flow < 1 cfs;
 if temp <= 17 then warm = 0;
   else warm = 1;
 if size <= 50 then large = 1;
   else large = 0;
 lnflowsq=lnflow*lnflow;
 sizeflow=size*lnflow;
 largwarm=large*warm;
 flowwarm=lnflow*warm;
 flowlarg=lnflow*large;
 flowlgwm=lnflow*warm*large;
run;

proc reg;
 model lnconc = lnflow;
 model lnconc = lnflow lnflowsq;
 model lnconc = lnflow large;
 model lnconc = lnflow warm;
 model lnconc = lnflow large warm;
 model lnconc = lnflow lnflowsq large warm;
 model lnconc = lnflow large warm largwarm;
 model lnconc = lnflow lnflowsq large warm largwarm;
 model lnconc = lnflow large flowlarg;
 model lnconc = lnflow lnflowsq large flowlarg;
 model lnconc = lnflow warm flowwarm;
 model lnconc = lnflow lnflowsq warm flowwarm;
 model lnconc = lnflow large warm largwarm flowlarg flowwarm
flowlgwm;
 model lnconc = lnflow lnflowsq large warm largwarm flowlarg
flowwarm
                flowlgwm;
 model lnconc = lnflow lnflowsq large warm flowlarg flowwarm;
 model lnconc = lnflow lnflowsq large warm largwarm flowlarg
flowwarm
                flowlgwm;
 model lnconc = lnflow lnflowsq size warm sizeflow flowwarm;
 run;
```

* *

```
*filename:  Lavaca/bivar.sas;
options pagesize=60 linesize=80;

data a;

 infile 'COL_DATA.prn';

 input month 1-8 day 9-16 year 17-24 sedld 25-32 flow 33-40

       eppt 41-48 hppt 49-56 yppt 57-64;

 date = mdy(month,day,year);
 format date mmddyy8.;
 if sedld = -9999 then sedld = '.';
 if flow >= 1 and flow < 10 then flowcat=1;
 if flow >= 10 and flow < 100 then flowcat=2;
 if flow >= 100 and flow < 500 then flowcat=3;
 if flow >= 500 and flow < 1000 then flowcat=4;
 if flow >= 1000 and flow < 5000 then flowcat=5;
 if flow >= 5000 then flowcat=6;
 *if sedld < 1 then sedld = '.';
 flowaf=flow*86400/43560;
 sedpct=sedld/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 lnflow=log(flow);
 lnconc=log(sedconc);
 lnload=log(sedld);
 lnflowsq = lnflow*lnflow;
run;

proc sort;
by flowcat;
run;


proc reg;
by flowcat;
 model lnconc = lnflow;
 model lnconc = lnflow lnflowsq;
 model lnload = lnflow;
 model lnload = lnflow lnflowsq;
run;
```

* *

```
Lavaca/Chap6gen.sas;
```

```
options pagesize=60 linesize=80;


data a;

 infile 'COL_DATA.prn';

 input month 1-8 day 9-16 year 17-24 sedld 25-32 flow 33-40
       eppt 41-48 hppt 49-56 yppt 57-64;

 date = mdy(month,day,year);
 format date mmddyy8.;
 if sedld = -9999 then sedld = '.';
 if eppt = -9999 then eppt = '.';
 if hppt = -9999 then hppt = '.';
 if yppt = -9999 then yppt = '.';
 if flow = 0 then flow = 0.1;
 if sedld = 0 then sedld = 0.1;
 flowaf=flow*86400/43560;
 sedpct=sedld/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 lnflow=log(flow);
 lnconc=log(sedconc);
 lnload=log(sedld);
 lneppt=log(eppt);
 lnhppt=log(hppt);
 lnyppt=log(yppt);
 rain=0.13*eppt+0.42*hppt+0.45*yppt;
 lnrain=log(rain);
proc means maxdec=2 n nmiss mean std stderr min max sum var cv
skewness
kurtosis;
 var flow sedconc sedld rain eppt hppt yppt;
run;

proc univariate normal plot;
 var flow sedconc sedld lnflow lnconc lnload eppt hppt yppt rain;
 id date;
run;

proc chart;
 vbar lnload / levels = 20;
 vbar sedld / midpoints = 0 to 50 by 5;
 vbar sedld / midpoints = 0 to 1000 by 50;
run;

proc corr;
 var lnflow lnload lnconc lnrain lneppt lnhppt lnyppt;
run;
```

```
proc corr;
 var flow sedconc sedld rain eppt hppt yppt;
run;

proc plot;
 plot sedconc*flow;
 plot lnconc*lnflow;
 plot lnload*lnflow;
run;

proc reg;
 model lnconc = lnflow;
 plot predicted.*lnconc;
 plot lnconc*predicted.;
 plot predicted.*lnflow = 'P' lnconc*lnflow='*' / overlay;
 model lnload = lnflow;
 plot predicted.*lnload;
 plot predicted.*lnflow = 'P' lnload*lnflow='*' / overlay;
run;
```

****

```
* file name:  Lavaca\chap6tmp.sas;
options pagesize=60 linesize=80;

data a;
 infile 'COL_DATA.prn';
 input month 1-8 day 9-16 year 17-24 sedld 25-32 flow 33-40
       eppt 41-48 hppt 49-56 yppt 57-64;
 date = mdy(month,day,year);
 format date mmddyy8.;
 if sedld = -9999 then sedld = '.';
 if eppt = -9999 then eppt = '.';
 if hppt = -9999 then hppt = '.';
 if yppt = -9999 then yppt = '.';
 if flow = 0 then flow = 0.1;
 if sedld = 0 then sedld = 0.1;
 flowaf=flow*86400/43560;
 sedpct=sedld/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 lnflow=log(flow);
 lnconc=log(sedconc);
 lnload=log(sedld);
* the following deals with low values;
* if flow < 315 then flow = '.';
* if lnflow < 100 then lnflow = '.';
* if sedconc < 163 then lnconc = '.';
* if lnload < 100 then lnload = '.';
*  if sedld < 386 then sedld = '.';
```

```
 lneppt=log(eppt);
 lnhppt=log(hppt);
 lnyppt=log(yppt);
 rain=0.13*eppt+0.42*hppt+0.45*yppt;
 lnrain=log(rain);
 lnflowsq=lnflow*lnflow;
* the following variables are rainfall*area/length;
 e=1;
 edna=eppt*106/14.5**e;
 hall=hppt*343/50.6**e;
 yoak=yppt*368/43.1**e;
dayofyr=(month(date)-1)*30.44+day(date);
 temp=22.3-2.76*sin(2*3.1416*dayofyr/365)-
7.52*cos(2*3.1416*dayofyr/365)
          -0.48*sin(4*3.1416*dayofyr/365)-
1.3*cos(4*3.1416*dayofyr/365);
flowtemp=lnflow*temp;
lagflow = lag1(flow);
 if flow >1.05*lagflow then rising=1;
   else rising=0;
 if flow <0.95*lagflow then falling=1;
   else falling=0;
 flowfall=lnflow*falling;
 flowrise=lnflow*rising;
* determine seasonal variables;
 if month(date) = 12 or month(date) = 1 or month(date) = 2 then
winter=1;
     else winter=0;
 if month(date) > 4 and month(date) < 10 then summer =1;
     else summer=0;
 winflow = winter*lnflow;
 sumflow = summer*lnflow;
run;

proc reg;
 model lnload = lnflow;
 model lnconc = lnflow;
 model lnload = lnflow lnflowsq;
 model lnconc = lnflow lnflowsq;
 model lnconc = lnflow temp;
 model lnconc = lnflow lnflowsq temp;
 model lnconc = lnflow lnflowsq temp flowtemp;
 model lnconc = lnflow rising falling;
 model lnconc = lnflow lnflowsq rising falling;
 model lnconc = lnflow rising flowrise;
 model lnconc = lnflow falling flowfall;
 model lnconc = lnflow lnflowsq rising flowrise falling flowfall;
 model lnconc = lnflow lnflowsq temp flowtemp rising flowrise
 falling
                 flowfall;
  model lnload = lnflow winter summer;
```

```
 model lnconc = lnflow winter summer;
 model lnconc = lnflow lnflowsq winter winflow sumflow summer;
 run;
```

* *

```
file name:  Lavaca\chap6yr.sas

options pagesize=60 linesize=80;

data a;

 infile 'COL_DATA.prn';

 input month 1-8 day 9-16 year 17-24 sedld 25-32 flow 33-40
       eppt 41-48 hppt 49-56 yppt 57-64;
 date = mdy(month,day,year);
 format date mmddyy8.;

 if sedld = -9999 then sedld = '.';
 if eppt = -9999 then eppt = '.';
 if hppt = -9999 then hppt = '.';
 if yppt = -9999 then yppt = '.';
 if flow = 0 then flow = 0.1;
 if sedld = 0 then sedld = 0.1;
 flowaf=flow*86400/43560;
 sedpct=sedld/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 lnflow=log(flow);
 lnconc=log(sedconc);
 lnload=log(sedld);
* the following deals with low values;
* if lnflow < 0 then lnflow = '.';
* if lnconc < 0 then lnconc = '.';
* if lnload < 0 then lnload = '.';
 lneppt=log(eppt);
 lnhppt=log(hppt);
 lnyppt=log(yppt);
 rain=0.13*eppt+0.42*hppt+0.45*yppt;
 lnrain=log(rain);
 lnflowsq=lnflow*lnflow;
* the following variables are rainfall*area/length;
 e=1;
 edna=eppt*106/14.5**e;
 hall=hppt*343/50.6**e;
 yoak=yppt*368/43.1**e;
dayofyr=(month(date)-1)*30.44+day(date);
```

```
 temp=22.3-2.76*sin(2*3.1416*dayofyr/365)-
7.52*cos(2*3.1416*dayofyr/365)
          -0.48*sin(4*3.1416*dayofyr/365)-
1.3*cos(4*3.1416*dayofyr/365);
flowtemp=lnflow*temp;
lagflow = lag1(flow);
 if flow >1.05*lagflow then rising=1;
   else rising=0;
 if flow <0.95*lagflow then falling=1;
   else falling=0;
 flowfall=lnflow*falling;
 flowrise=lnflow*rising;
* determine seasonal variables;
 if month(date) = 12 or month(date) = 1 or month(date) = 2 then
winter=1;
     else winter=0;
 if month(date) > 4 and month(date) < 10 then summer =1;
     else summer=0;
 winflow = winter*flow;
 sumflow = summer*flow;
run;

proc sort;
  by year;
run;

proc means;
  by year;
run;

proc reg;
 by month;
 model lnload = lnflow;
 model lnconc = lnflow;
 model lnload = lnflow lnflowsq;
 model lnconc = lnflow lnflowsq;
 model lnconc = lnflow temp;
 model lnconc = lnflow lnflowsq temp;
 model lnconc = lnflow lnflowsq temp flowtemp;
 model lnconc = lnflow rising falling;
 model lnconc = lnflow lnflowsq rising falling;
 model lnconc = lnflow rising flowrise;
 model lnconc = lnflow falling flowfall;
 model lnconc = lnflow lnflowsq rising flowrise falling flowfall;
 model lnconc = lnflow lnflowsq temp flowtemp rising flowrise
falling flowfall;
* model lnload = lnflow winter summer;
* model lnconc = lnflow winter summer;
* model lnconc = lnflow lnflowsq winter winflow sumflow summer;
 run;
```

```
                  *******************************

* file name:  Lavaca\chap6spa.sas
options pagesize=60 linesize=80;
data a;
 infile 'COL_DATA.prn';
 input month 1-8 day 9-16 year 17-24 sedld 25-32 flow 33-40
       eppt 41-48 hppt 49-56 yppt 57-64;
 date = mdy(month,day,year);
 format date mmddyy8.;
 if sedld = -9999 then sedld = '.';
 if eppt = -9999 then eppt = '.';
 if hppt = -9999 then hppt = '.';
 if yppt = -9999 then yppt = '.';
 if flow = 0 then flow = 0.1;
 if sedld = 0 then sedld = 0.1;
 flowaf=flow*86400/43560;
 sedpct=sedld/(flowaf*1361.25);

sedconc=sedpct/1.102*2000*0.454*1000000*1361.25/43560/0.02832*0.001
;
 lnflow=log(flow);
 lnconc=log(sedconc);
 lnload=log(sedld);
* the following deals with low values;
* if lnflow < 0 then lnflow = '.';
* if lnconc < 0 then lnconc = '.';
* if lnload < 0 then lnload = '.';
 lneppt=log(eppt);
 lnhppt=log(hppt);
 lnyppt=log(yppt);
 rain=0.13*eppt+0.42*hppt+0.45*yppt;
 lnrain=log(rain);
 lnflowsq=lnflow*lnflow;
* the following variables are rainfall*area/length;
 e=2;
 edna=eppt*106/(14.5**e);
 hall=hppt*343/(50.6**e);
 yoak=yppt*368/(43.1**e);
dayofyr=(month(date)-1)*30.44+day(date);
 temp=22.3-2.76*sin(2*3.1416*dayofyr/365)-
7.52*cos(2*3.1416*dayofyr/365)
          -0.48*sin(4*3.1416*dayofyr/365)-
1.3*cos(4*3.1416*dayofyr/365);
flowtemp=lnflow*temp;
lagflow = lag1(flow);
 if flow >1.05*lagflow then rising=1;
   else rising=0;
 if flow <0.95*lagflow then falling=1;
   else falling=0;
 flowfall=lnflow*falling;
```

```
 flowrise=lnflow*rising;
run;

proc reg;
 model lnload = lnflow;
 model lnconc =lnflow;
 model lnload = lnflow lnflowsq;
 model lnconc = lnflow lnflowsq;
 model lnload = edna hall yoak;
 model lnconc = edna hall yoak;
 model sedld = edna hall yoak;
 model sedconc = edna hall yoak;
 model lnload = lnflow edna hall yoak;
 model lnconc = lnflow edna hall yoak;
 model lnload = lnflow eppt hppt yttp;
 model lnconc = lnflow eppt hppt yppt;
 model lnconc = lnflow eppt;
 model lnconc = lnflow lnflowsq eppt;
 model lnconc = lnflow rain;
 model lnconc = rain;
 model lnconc = lnflow lnflowsq rain;
 model lnconc = lnflow lnflowsq eppt;
 model lnconc = lnflow lnflowsq hppt;
 model lnconc = lnflow lnflowsq yppt;
 model lnconc = lnflow temp;
 model lnconc = lnflow lnflowsq temp;
 model lnconc = lnflow lnflowsq temp flowtemp;
 model lnconc = lnflow rising falling;
 model lnconc = lnflow lnflowsq rising falling;
 model lnconc = lnflow rising flowrise;
 model lnconc = lnflow falling flowfall;
 model lnconc = lnflow lnflowsq rising flowrise falling flowfall;
 model lnconc = lnflow lnflowsq temp flowtemp rising flowrise
falling flowfall;
 run;
```

8.4 Appendix D – Values for Watersheds used in SAS

No.	Station No.	No. Yrs.	Kgload	Flowm3	Area (m^2)	Perimeter (m)	Length (m)
1	7299200	10	1.639E+09	698	1.63E+10	9.82E+05	7.27E+05
2	7308000	4	1.109E+09	1,245	8.94E+09	7.70E+05	3.09E+05
3	7331600	4	1.938E+10	50,435	1.03E+11	2.50E+06	1.04E+06
4	7336820	9	3.651E+09	121,740	1.23E+11	3.33E+06	1.31E+06
5	7342500	3	1.294E+08	2,220	1.37E+09	2.28E+05	8.43E+04
6	7343200	21	1.170E+09	15,055	3.54E+09	3.64E+05	1.33E+05
7	7343500	23	4.666E+07	4,949	1.37E+09	2.33E+05	7.84E+04
8	8022000	19	1.242E+08	26,078	9.08E+09	7.76E+05	2.54E+05
9	8022500	32	6.603E+08	35,830	1.26E+10	9.59E+05	3.32E+05
10	8031200	16	2.814E+06	1,479	5.99E+08	1.41E+05	4.34E+04
11	8033000	17	6.125E+07	15,760	7.03E+09	6.86E+05	2.42E+05
12	8033300	8	2.463E+06	256	2.05E+08	.	.
13	8033500	35	2.972E+08	23,738	9.34E+09	8.33E+05	2.91E+05
14	8037050	9	8.721E+06	375	8.11E+07	.	.
15	8038500	8	1.241E+08	20,292	7.37E+09	5.76E+05	1.94E+05
16	8039500	6	4.638E+08	38,785	9.08E+09	6.59E+05	2.33E+05
17	8052700	3	3.512E+07	336	2.04E+08	9.00E+04	3.13E+04
18	8062500	36	8.527E+08	25,641	2.11E+10	9.65E+05	3.24E+05
19	8064500	15	3.219E+08	4,573	2.50E+09	3.65E+05	1.16E+05
20	8065350	17	1.439E+09	57,531	3.61E+10	1.45E+06	5.30E+05
21	8066200	15	6.877E+07	933	3.70E+08	1.01E+05	3.13E+04
22	8066500	51	2.826E+09	74,212	4.46E+10	1.88E+06	7.27E+05
23	8068000	7	1.200E+08	3,488	2.18E+09	2.65E+05	7.97E+04
24	8069500	15	3.774E+08	12,046	4.66E+09	4.37E+05	1.15E+05
25	8070000	36	2.281E+07	2,160	8.48E+08	1.98E+05	6.30E+04
26	8071500	6	7.263E+08	21,007	7.38E+09	5.42E+05	1.40E+05
27	8080500	3	4.090E+09	1,658	2.18E+10	1.37E+06	5.43E+05
28	8082500	3	9.379E+09	6,443	3.90E+10	1.62E+06	7.34E+05
29	8084800	14	4.045E+07	312	1.23E+09	2.16E+05	8.23E+04
30	8085500	2	7.266E+07	2,201	1.03E+10	6.51E+05	2.75E+05

No.	Station No.	No. Yrs.	Kgload	Flowm3	Area (m^2)	Perimeter (m)	Length (m)
31	8087300	15	4.594E+08	3,142	1.49E+10	8.15E+05	3.46E+05
32	8088000	43	3.182E+09	7,808	5.75E+10	1.95E+06	8.60E+05
33	8088600	36	8.788E+07	8,844	5.98E+10	2.02E+06	9.10E+05
34	8093500	23	1.961E+08	1,253	6.64E+08	1.45E+05	4.44E+04
35	8094800	23	3.010E+07	452	9.24E+08	1.75E+05	6.14E+04
36	8100500	33	1.945E+08	2,496	6.12E+09	6.26E+05	2.37E+05
37	8102500	4	4.841E+08	4,861	9.30E+09	7.75E+05	3.21E+05
40	8110000	16	9.754E+06	2,976	2.65E+09	3.04E+05	9.13E+04
41	8110500	41	1.214E+08	4,336	2.60E+09	3.51E+05	1.08E+05
42	8114000	54	2.073E+10	74,865	1.15E+11	3.37E+06	1.67E+06
43	8146000	20	7.952E+07	1,968	7.75E+09	6.33E+05	2.26E+05
44	8147000	61	2.248E+09	11,200	7.39E+10	2.21E+06	8.32E+05
45	8148000	35	1.503E+07	280,352	8.32E+10	2.29E+06	9.13E+05
46	8148090	12	5.003E+07	8,189	8.35E+10	2.30E+06	9.19E+05
47	8151500	43	3.689E+08	3,752	1.09E+10	7.05E+05	2.33E+05
48	8153500	23	6.214E+08	1,493	2.30E+09	3.14E+05	1.09E+05
49	8158000	46	3.305E+08	20,590	1.02E+11	2.58E+06	1.06E+06
50	8164000	43	1.329E+08	3,335	2.11E+09	3.09E+05	1.12E+05
51	8164300	22	3.715E+07	1,686	8.98E+08	1.62E+05	5.15E+04
52	8164500	2	1.118E+08	5,185	2.84E+09	3.58E+05	1.23E+05
53	8167500	46	1.581E+08	3,662	3.43E+09	4.24E+05	1.80E+05
54	8176500	43	4.677E+08	18,489	1.36E+10	1.08E+06	4.57E+05
55	8183500	5	2.044E+08	1,851	5.48E+09	6.26E+05	2.47E+05
56	8186000	25	9.316E+07	1,449	2.14E+09	4.73E+05	1.76E+05
57	8188500	47	4.389E+08	7,463	1.02E+10	8.51E+05	3.45E+05
58	8194000	36	5.941E+07	2,543	1.35E+10	8.95E+05	3.20E+05
59	8207000	25	1.144E+08	2,493	1.40E+10	7.94E+05	3.13E+05
60	8210000	25	6.607E+08	9,128	4.01E+10	1.29E+06	4.70E+05
61	8211000	41	6.952E+07	7,548	4.30E+10	1.37E+06	5.42E+05

No.	k_avg	k_max	streams	elev_ avg	elev_ min	elev_ max	slop_ avg	slop_ max	rain_ sta
1	0.32	0.36	1833	1105	591	1508	1.05	18.20	502
2	0.30	0.40	991	751	402	1095	1.14	9.79	616
3	0.31	0.41	10994	590	164	1508	1.02	18.20	1002
4	0.31	0.41	12887	527	99	1508	1.09	33.73	1209
5	0.37	0.41	106	162	116	220	0.56	3.10	1104
6	0.36	0.41	306	152	91	220	0.59	3.10	1132
7	0.37	0.41	79	134	85	184	0.57	2.50	1150
8	0.31	0.45	757	135	65	215	0.77	5.32	1183
9	0.33	0.45	1046	124	55	215	0.77	5.32	1251
10	0.29	0.36	34	148	116	188	0.80	4.60	1062
11	0.30	0.50	510	118	42	221	1.13	6.70	1126
12	.	.	.	.	.	.	.	.	1137
13	0.33	0.50	700	108	35	221	1.04	6.70	1257
14	.	.	.	.	.	.	.	.	1151
15	0.29	0.45	586	107	36	221	1.13	6.74	1247
16	0.29	0.50	751	101	31	221	1.10	6.74	1317
17	0.35	0.41	0	203	166	244	0.75	1.99	983
18	0.30	0.42	1699	230	91	426	1.01	5.67	912
19	0.33	0.41	197	173	95	273	0.85	3.43	941
20	0.31	0.42	2968	189	51	426	0.92	5.67	1038
21	0.35	0.38	12	82	33	121	1.03	2.95	1234
22	0.32	0.50	3706	167	15	426	0.89	5.67	1262
23	0.26	0.36	170	87	34	140	0.79	3.16	1191
24	0.28	0.42	412	71	16	140	0.55	3.16	1221
25	0.29	0.38	94	87	39	136	0.85	2.82	1288
26	0.28	0.42	804	67	5	140	0.54	3.16	1268
27	0.27	0.39	2056	1058	512	1431	0.68	10.59	599
28	0.29	0.40	4145	952	383	1449	0.75	10.59	691
29	0.29	0.39	105	519	453	609	0.58	2.25	638
30	0.28	0.39	1192	561	363	808	0.99	9.24	679
31	0.27	0.39	1691	523	325	808	1.05	9.24	755
32	0.28	0.40	6186	807	325	1449	0.83	10.59	766

No.	k_avg	k_max	streams	elev_ avg	elev_ min	elev_ max	slop_ avg	slop_ max	rain_ sta
33	0.28	0.40	6421	790	284	1449	0.84	10.59	787
34	0.32	0.41	37	205	150	260	0.82	2.74	879
35	0.29	0.40	70	401	299	479	1.04	3.16	804
36	0.25	0.40	510	411	228	606	1.14	6.94	837
37	0.25	0.40	823	373	151	606	1.29	6.94	859
40	0.30	0.42	202	123	59	195	0.84	3.04	1002
41	0.32	0.41	163	142	86	212	0.72	3.34	985
42	0.28	0.42	11452	520	14	1449	0.96	10.59	1159
43	0.20	0.34	592	601	360	760	1.07	5.45	680
44	0.24	0.38	6954	808	355	1401	0.77	11.72	695
45	0.24	0.38	7718	782	308	1401	0.81	11.72	734
46	0.23	0.38	7727	780	268	1401	0.81	11.72	746
47	0.19	0.29	778	593	304	761	1.59	11.06	683
48	0.21	0.29	162	538	336	681	1.78	7.28	794
49	0.23	0.38	9169	733	134	1401	0.99	13.97	814
50	0.29	0.41	202	74	9	166	0.66	2.60	1046
51	0.30	0.39	51	99	54	163	0.87	2.29	982
52	0.30	0.41	263	65	10	163	0.50	2.58	1052
53	0.19	0.30	279	539	298	730	2.08	8.61	849
54	0.26	0.41	1348	263	12	730	1.34	8.61	953
55	0.23	0.35	528	312	95	704	1.67	11.19	726
56	0.26	0.35	246	245	88	607	1.19	5.87	737
57	0.25	0.36	1020	245	29	704	1.36	11.19	887
58	0.24	0.41	1254	336	117	734	1.23	10.73	577
59	0.25	0.41	1445	271	52	727	1.25	12.58	639
60	0.26	0.41	3739	250	32	734	1.06	12.58	700
61	0.26	0.41	3979	239	15	734	1.05	12.58	800

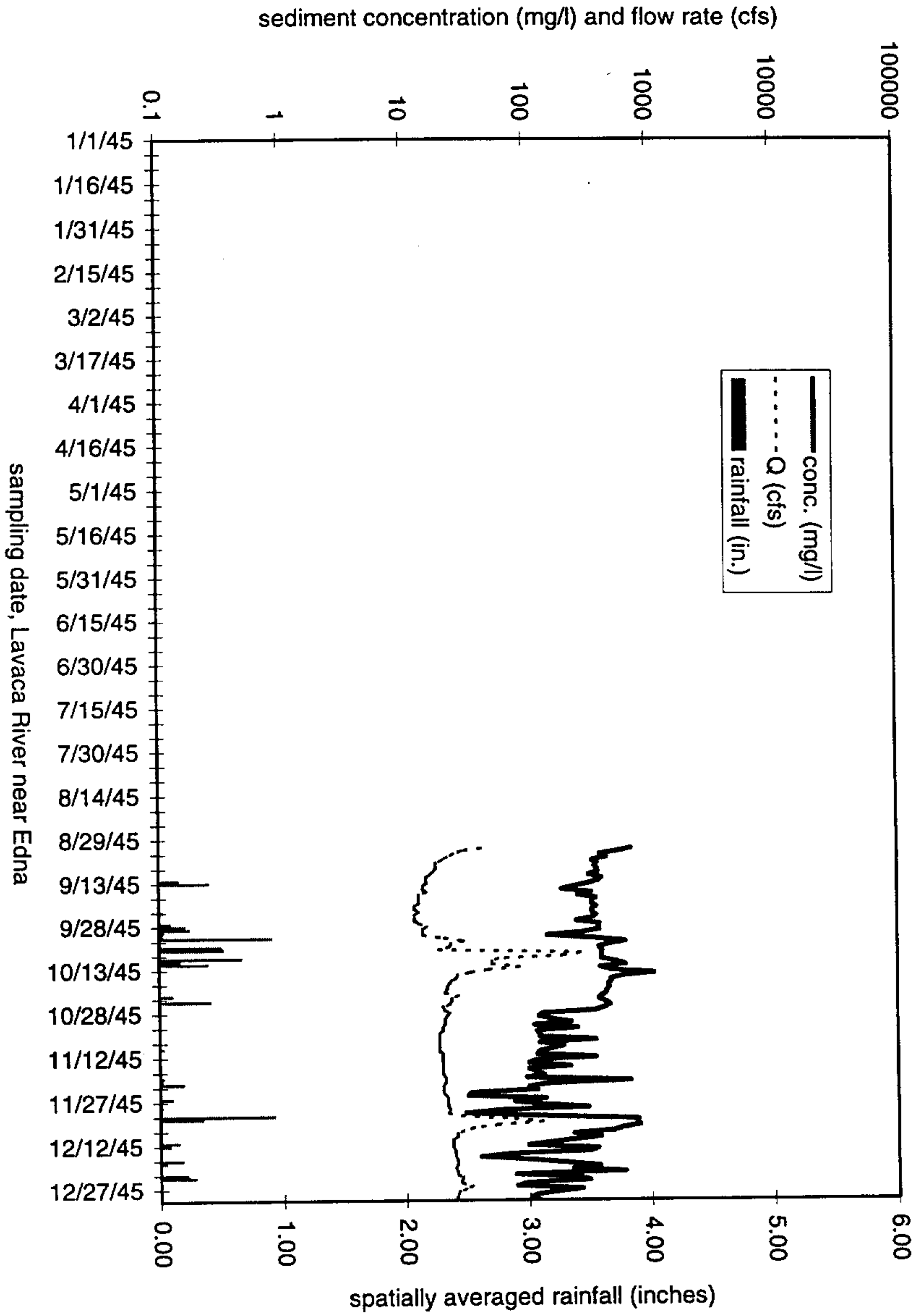
No.	rain_ avg	rain_ min	rain_ max	lu0	lu1	lu2	lu3	lu4	lu5	lu6	lu7
1	478	417	562	0	6	55	37	0	0	1	0
2	537	463	634	0	0	44	55	0	0	1	0
3	663	417	1067	1	3	50	40	4	1	0	0
4	737	417	1462	1	2	52	35	7	1	0	0
5	1090	1057	1150	0	1	91	0	8	0	0	0
6	1111	1057	1185	12	2	39	38	8	1	0	0
7	1132	1104	1163	5	23	60	0	12	1	0	0
8	1091	1007	1198	1	3	57	0	36	2	0	0
9	1124	1007	1272	1	3	50	0	45	2	0	0
10	1054	1030	1089	0	1	72	0	27	0	0	0
11	1090	1020	1185	1	2	34	0	61	2	0	0
12	.	.	.	.	.	.	.	.	.	.	.
13	1111	1020	1295	1	2	28	0	67	1	0	0
14	.	.	.	.	.	.	.	.	.	.	.
15	1162	1080	1278	1	2	29	0	65	2	1	0
16	1186	1080	1394	1	2	25	0	67	5	1	1
17	992	969	1022	0	1	98	0	1	0	0	0
18	903	735	1092	3	10	58	16	9	2	0	1
19	917	852	975	0	1	94	2	2	1	0	0
20	938	735	1101	2	7	63	11	14	2	1	1
21	1212	1174	1236	1	1	11	0	88	0	0	1
22	972	735	1320	2	6	57	9	23	2	0	1
23	1135	1093	1195	4	11	26	0	53	4	0	2
24	1142	1064	1233	8	11	26	0	52	2	0	2
25	1202	1141	1288	0	2	8	0	89	0	0	0
26	1174	1064	1289	6	8	19	0	64	2	0	2
27	475	371	615	0	2	60	36	0	0	0	1
28	503	371	694	0	1	60	37	0	0	1	1
29	628	603	647	50	3	37	9	0	0	0	0
30	636	569	731	17	3	33	42	4	0	0	0
31	655	569	755	12	2	27	54	4	1	0	0
32	555	371	768	3	1	50	43	1	0	0	0

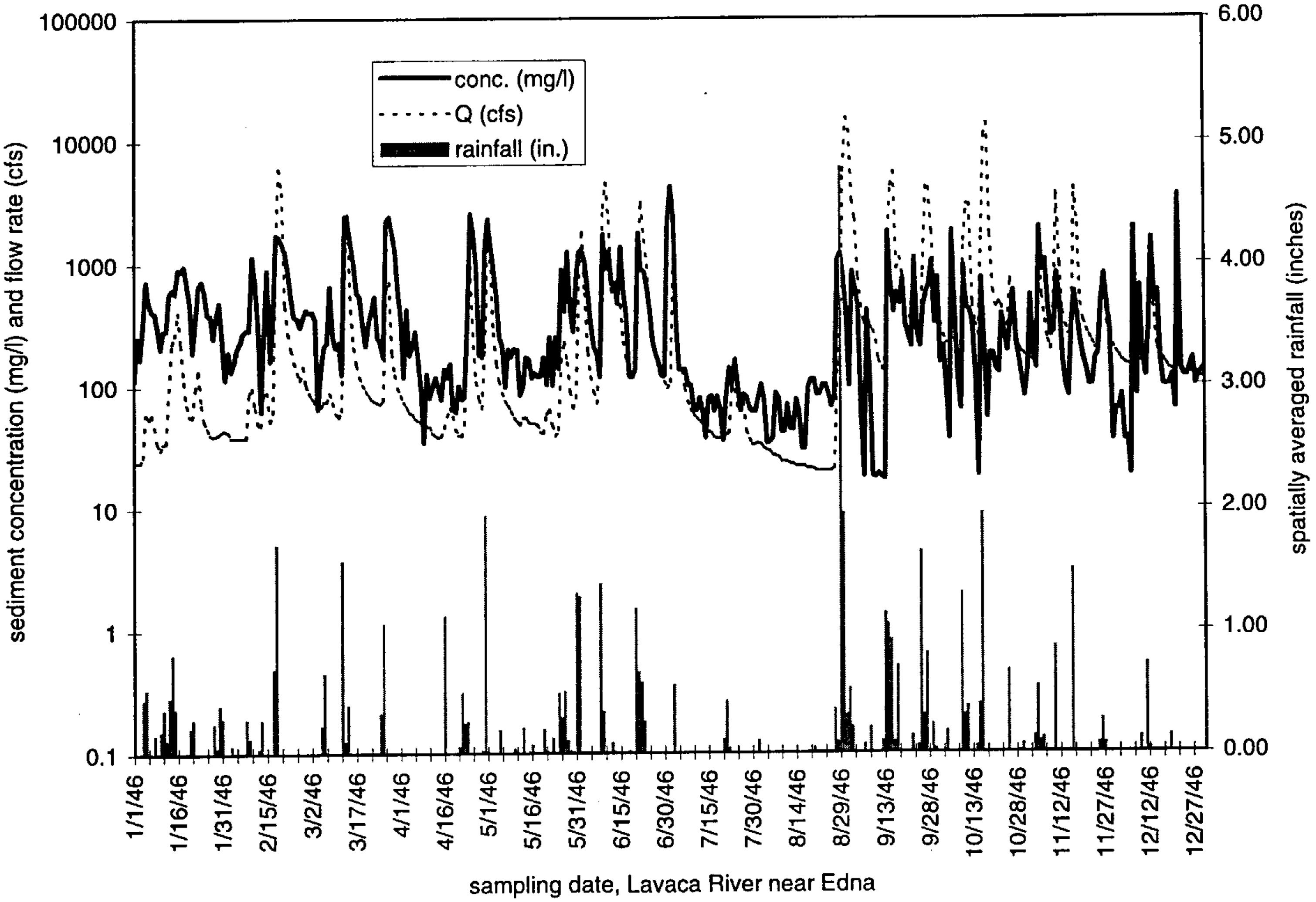
No.	rain_ avg	rain_ min	rain_ max	lu0	lu1	lu2	lu3	lu4	lu5	lu6	lu7
33	563	371	793	3	1	48	44	2	0	0	0
34	879	855	902	2	2	69	23	4	0	0	0
35	788	763	816	7	2	54	33	4	1	0	0
36	751	655	849	5	1	40	41	8	5	0	0
37	773	655	879	6	2	33	41	13	4	0	0
40	937	873	1029	4	1	41	22	28	2	1	2
41	975	926	1020	63	1	3	10	22	0	0	0
42	711	371	1159	12	2	39	38	8	1	0	0
43	640	564	727	0	0	12	48	39	0	0	0
44	516	338	761	8	4	16	55	3	0	0	14
45	532	338	780	7	3	15	54	7	0	0	12
46	533	338	787	7	3	15	54	8	1	0	12
47	658	572	773	0	11	2	34	52	0	0	0
48	765	697	825	0	1	17	10	71	0	0	0
49	564	338	862	6	4	13	49	16	1	0	10
50	966	896	1051	0	1	62	6	27	4	0	0
51	966	927	993	0	1	90	2	8	0	0	0
52	1011	927	1079	0	0	62	4	31	3	0	0
53	774	705	871	0	38	7	12	42	0	0	1
54	847	705	960	0	11	35	10	42	0	0	1
55	757	693	866	0	11	36	9	41	1	0	2
56	808	729	871	2	2	75	4	15	0	0	2
57	775	693	896	1	7	51	7	32	1	0	1
58	585	531	709	0	1	7	67	25	0	0	0
59	647	551	764	0	1	21	56	22	0	0	0
60	618	531	764	1	1	15	67	17	0	0	0
61	624	531	810	1	1	16	66	17	0	0	0

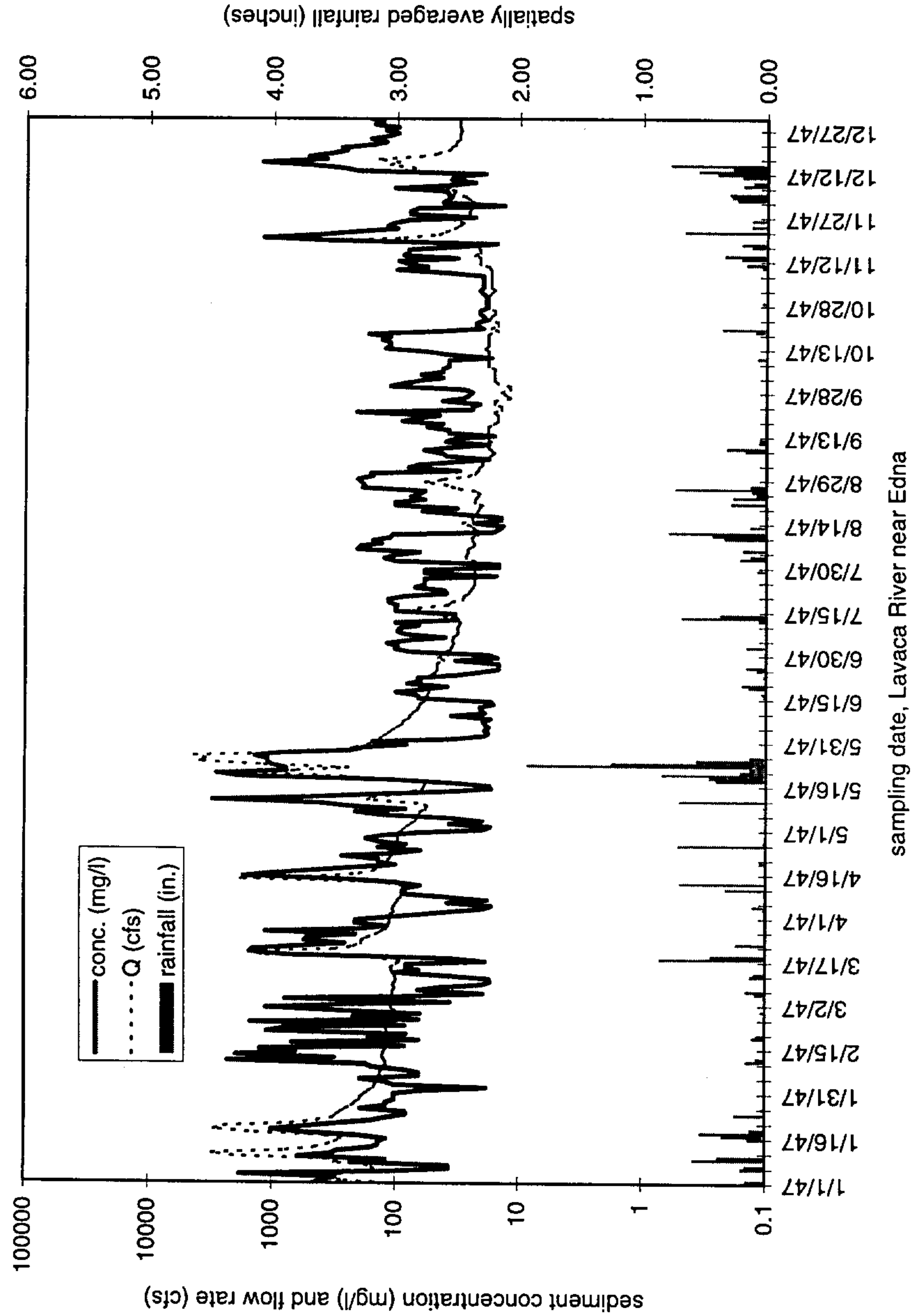
No.	soil A	soil B	soil C	soil D	USGS slp	normload	sed conc	frac area	resvar
1	1	25	12	63	0.384	100282	26586	0	0
2	6	37	12	46	2.383	123987	9079	0	0
3	5	35	21	39	1.243	188188	3959	0.06	0.15
4	5	35	20	41	1.020	29740	289	0.84	4.17
5	0	5	2	94	0.965	94787	754	0	0
6	0	5	3	92	0.795	330428	777	0	0
7	5	6	22	67	0.902	34018	99	0	0
8	9	21	28	41	0.367	13690	53	0.33	0.57
9	8	20	34	38	0.330	52551	229	0.06	0.13
10	36	7	28	29	0.994	4697	20	0	0
11	17	35	31	16	0.376	8714	43	0.31	0.49
12	.	.	.	.	.	12016	103	0	0
13	13	32	28	27	0.352	31817	113	0.03	0.04
14	.	.	.	.	.	107581	253	0	0
15	9	44	38	9	0.434	16844	74	0	0
16	9	43	35	13	1.845	51051	128	0	0
17	0	6	9	85	1.698	172138	1018	0	0
18	0	15	34	51	0.795	40371	375	0.76	2.23
19	0	5	22	73	1.149	128745	799	0.17	0.6
20	5	13	26	56	0.534	39912	266	0.58	1.09
21	0	20	21	59	2.960	186111	658	0	0
22	5	14	25	55	0.384	63413	387	0.53	1.35
23	0	34	23	42	0.970	55107	284	0	0
24	0	33	34	33	0.781	81040	300	0	0
25	0	40	30	29	1.136	26902	106	0	0
26	0	38	30	32	0.671	98481	268	0	0
27	4	66	20	10	1.616	187323	23156	0	0
28	5	50	26	19	1.398	240469	15805	0	0
29	3	24	64	8	1.346	32889	1072	0	0
30	3	25	54	18	1.121	7086	390	0.24	0.6
31	2	21	52	26	1.077	30922	1008	0.35	1.18
32	4	39	35	22	1.285	55379	4199	0.09	0.58

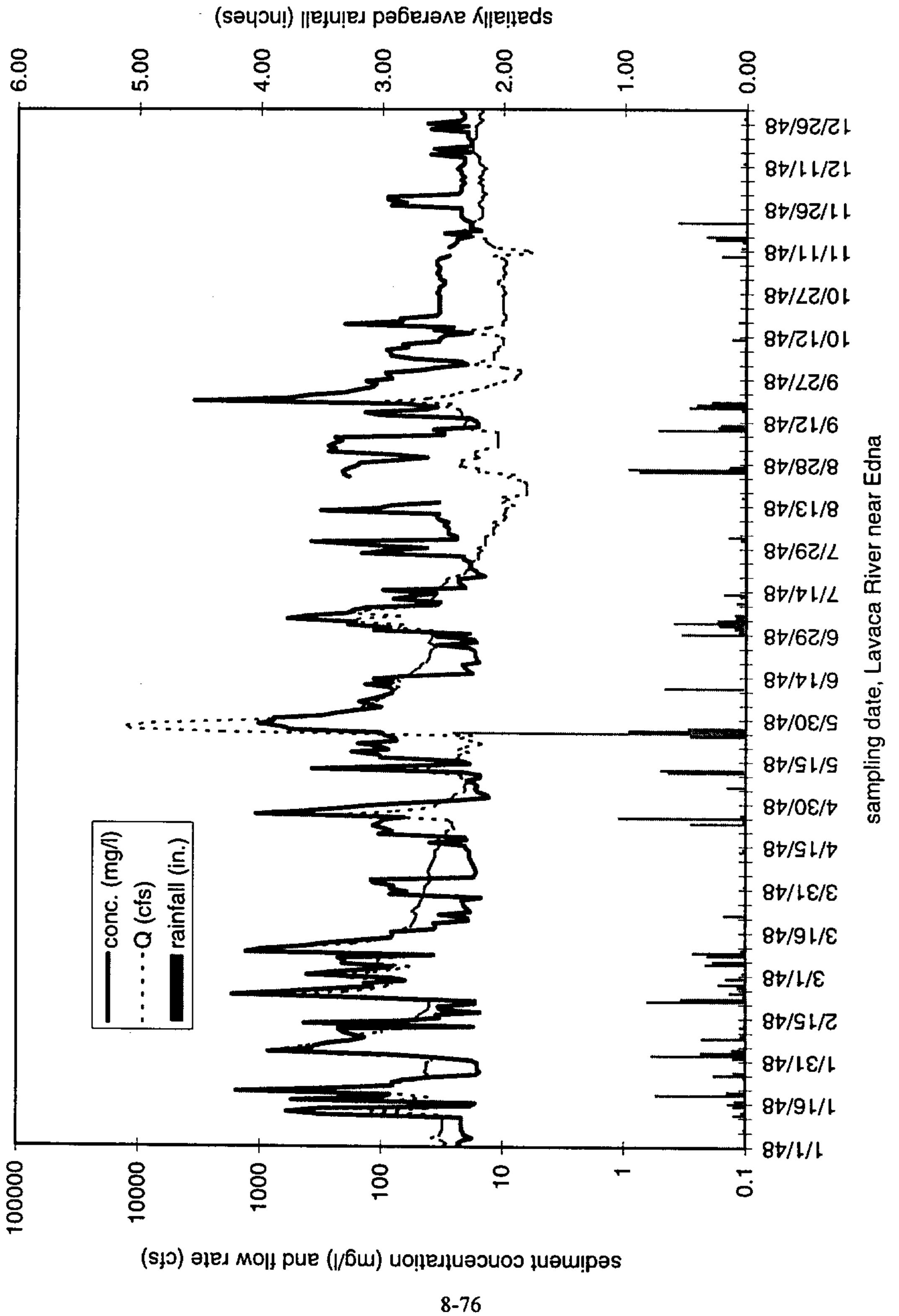
No.	soil A	soil B	soil C	soil D	USGS slp	normload	sed conc	frac area	resvar
33	4	38	36	23	1.244	1470	118	0.61	323.9
34	0	20	42	38	1.845	295589	1685	0.17	2.23
35	0	15	56	29	2.127	32597	576	0	0
36	0	16	22	62	1.140	31769	770	0.44	0.74
37	0	13	48	38	1.127	52039	963	0	0
40	10	29	13	49	1.041	3681	58	0.74	134.87
41	9	22	8	62	0.867	46774	280	0.21	0.55
42	3	27	32	38	0.803	179966	2544	0.32	0.81
43	1	9	20	70	1.700	10256	383	0.17	0.48
44	4	40	32	24	1.182	30419	1807	0.32	0.92
45	3	36	31	29	1.099	181	1	0.35	0.85
46	3	36	31	29	1.086	599	55	0.97	17.9
47	2	9	17	73	1.774	33950	869	0	0
48	4	8	32	56	2.924	270301	1682	0	0
49	3	31	30	36	1.034	3247	109	0.99	8.4
50	2	12	39	47	0.834	62852	499	0	0
51	0	18	32	49	1.772	41354	262	0	0
52	2	9	39	50	0.772	39329	230	0	0
53	0	6	20	73	1.979	46076	445	0	0
54	4	14	26	56	1.384	34393	258	0.39	0.9
55	5	12	23	59	1.888	37295	1068	0.3	0.52
56	9	11	28	51	2.285	43465	563	0	0
57	5	15	30	50	1.356	43184	651	0.16	0.23
58	1	19	30	50	1.682	4401	220	0.34	1.18
59	1	23	38	38	1.625	8154	607	0	0
60	2	22	35	41	1.187	16461	921	0.03	0.05
61	2	23	36	40	1.068	1619	78	0.67	8.15

8.5 *Appendix E – Daily Plots for the Lavaca River*

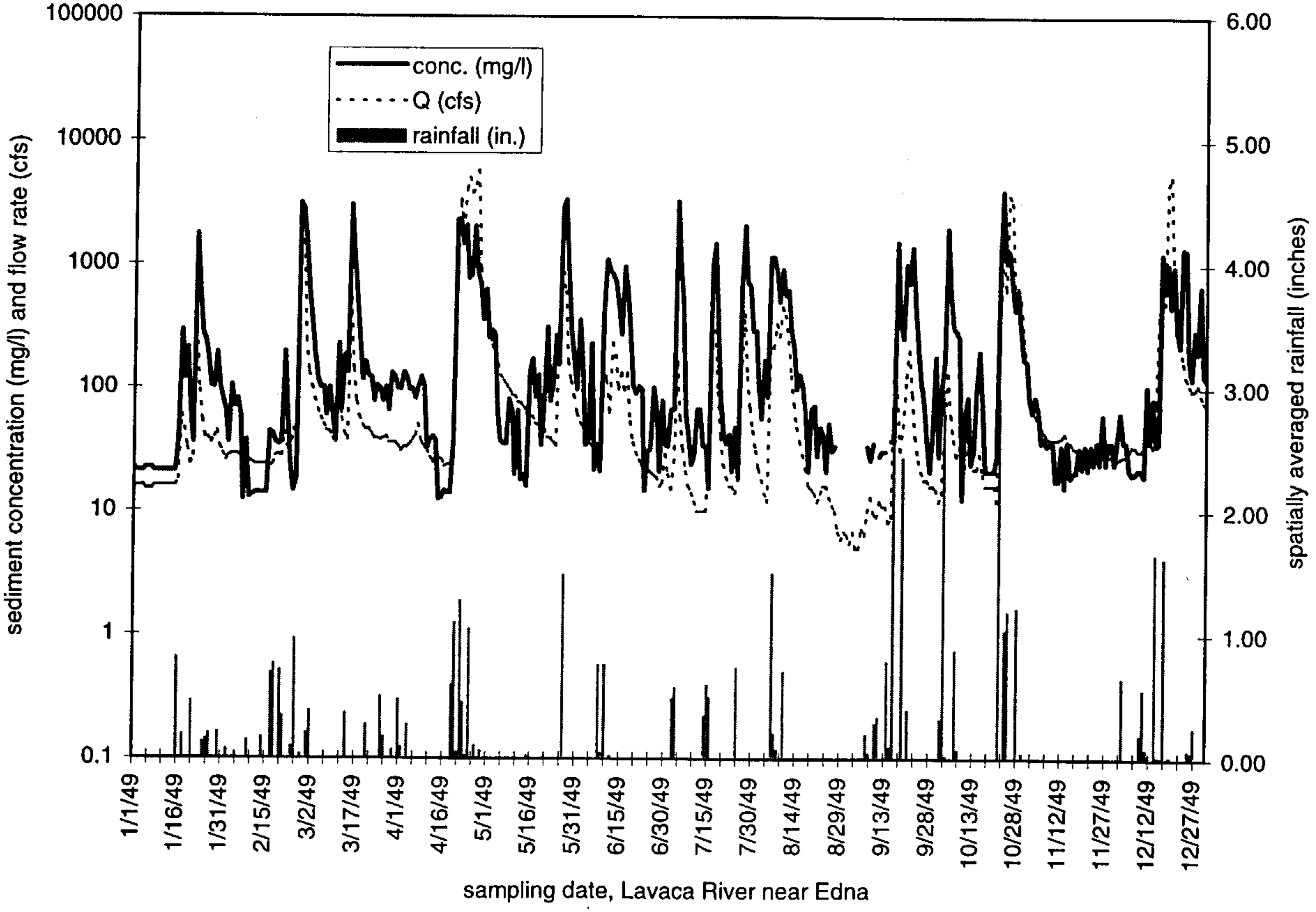



conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/46
1/16/46
1/31/46
2/15/46
3/2/46
3/17/46
4/1/46
4/16/46
5/1/46
5/16/46
5/31/46
6/15/46
6/30/46
7/15/46
7/30/46
8/14/46
8/29/46
9/13/46
9/28/46
10/13/46
10/28/46
11/12/46
11/27/46
12/12/46
12/27/46
sampling date, Lavaca River near Edna

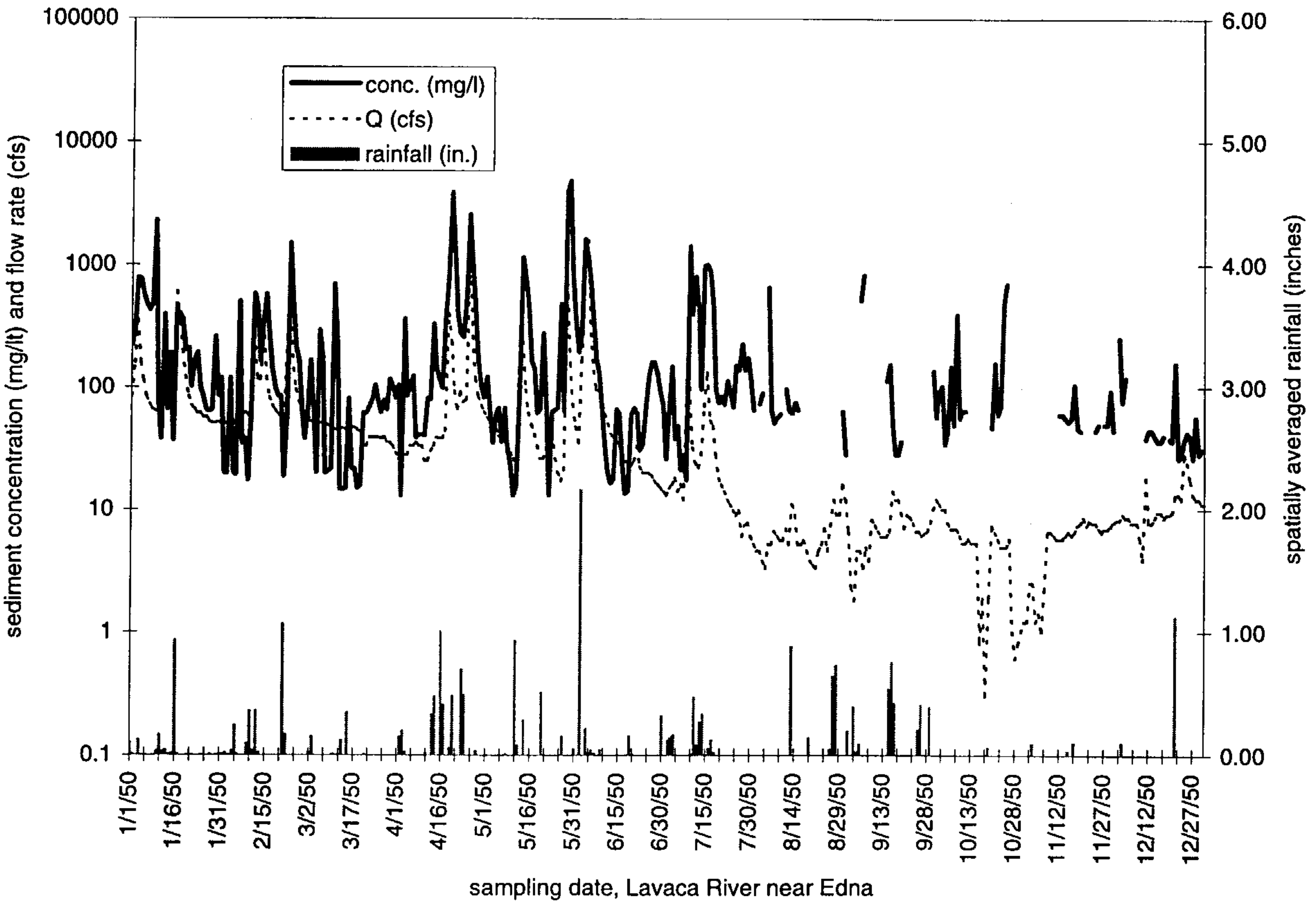

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/47
1/16/47
1/31/47
2/15/47
3/2/47
3/17/47
4/1/47
4/16/47
5/1/47
5/16/47
5/31/47
6/15/47
6/30/47
7/15/47
7/30/47
8/14/47
8/29/47
9/13/47
9/28/47
10/13/47
10/28/47
11/12/47
11/27/47
12/12/47
12/27/47
sampling date, Lavaca River near Edna



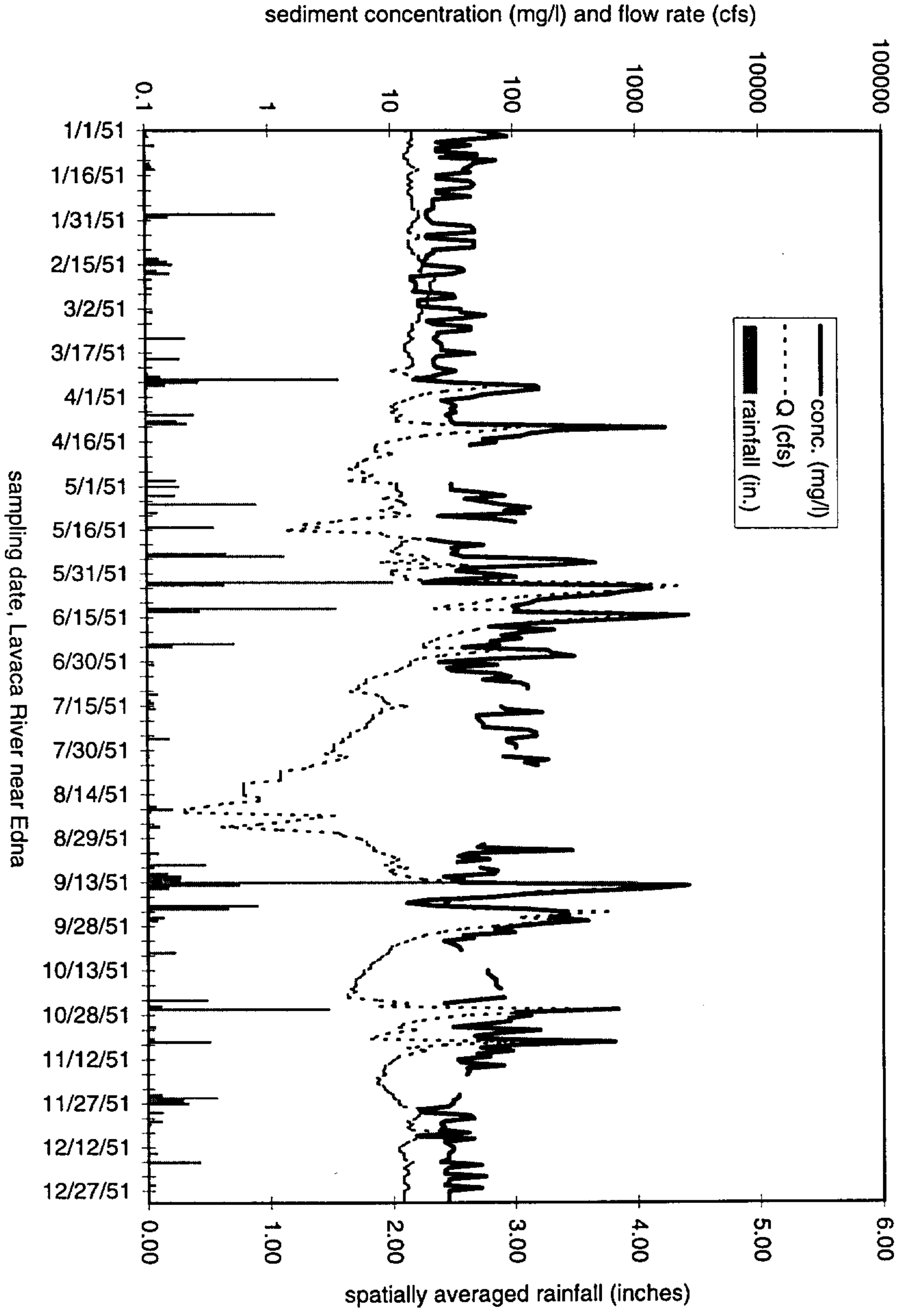
conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
sampling date, Lavaca River near Edna
1/1/48
1/16/48
1/31/48
2/15/48
3/1/48
3/16/48
3/31/48
4/15/48
4/30/48
5/15/48
5/30/48
6/14/48
6/29/48
7/14/48
7/29/48
8/13/48
8/28/48
9/12/48
9/27/48
10/12/48
10/27/48
11/11/48
11/26/48
12/11/48
12/26/48

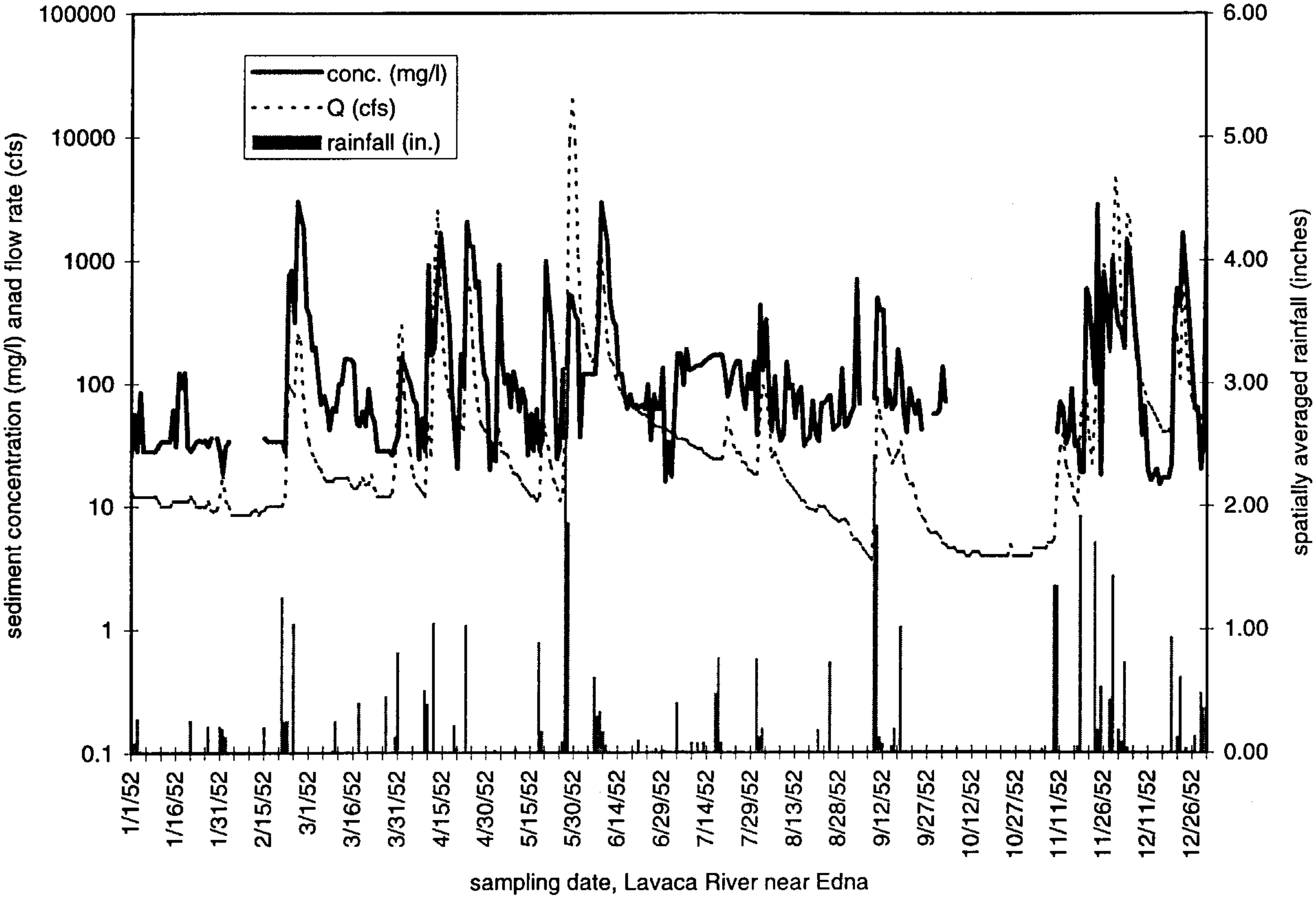


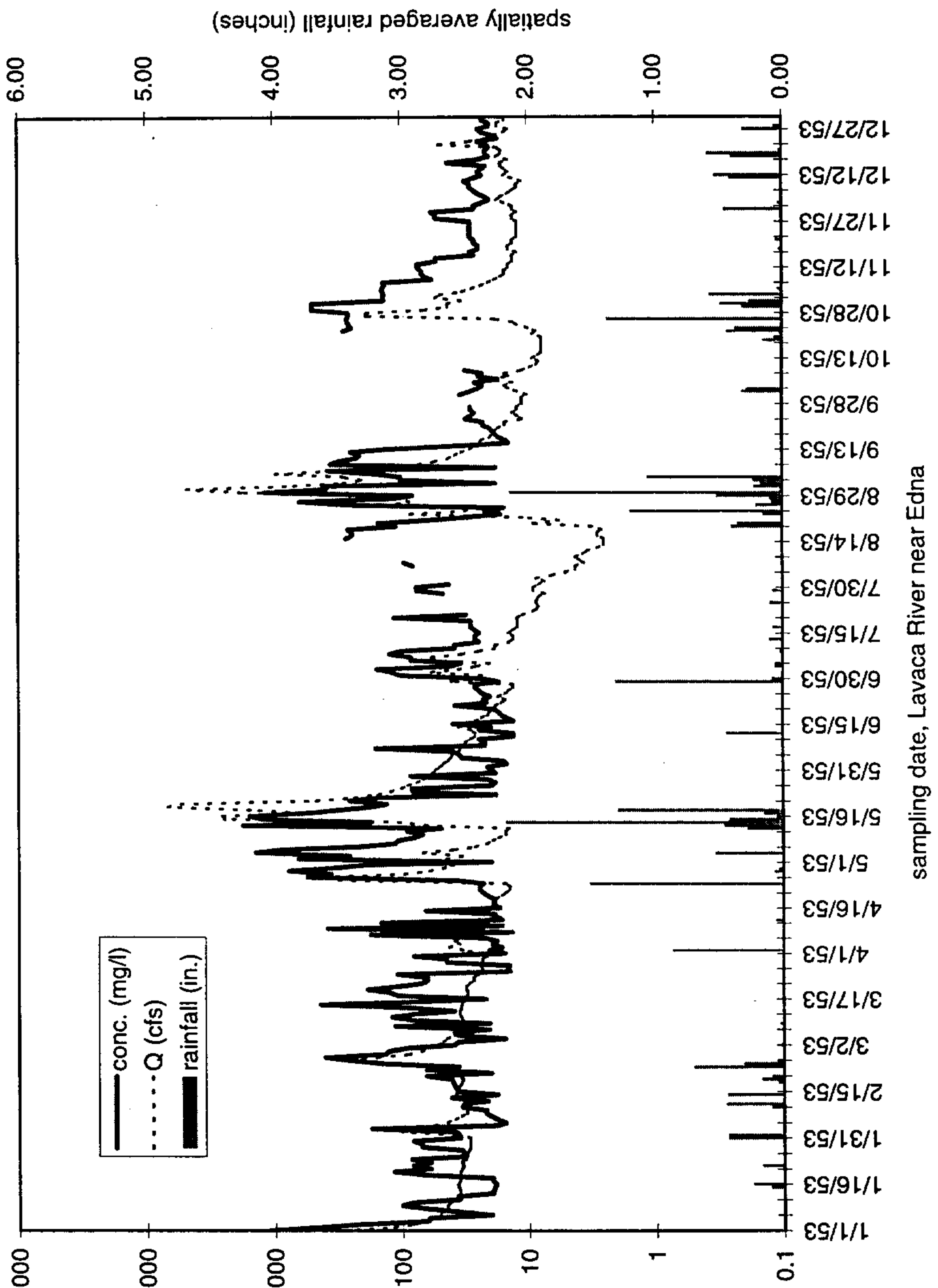
conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/49
1/16/49
1/31/49
2/15/49
3/2/49
3/17/49
4/1/49
4/16/49
5/1/49
5/16/49
5/31/49
6/15/49
6/30/49
7/15/49
7/30/49
8/14/49
8/29/49
9/13/49
9/28/49
10/13/49
10/28/49
11/12/49
11/27/49
12/12/49
12/27/49
sampling date, Lavaca River near Edna

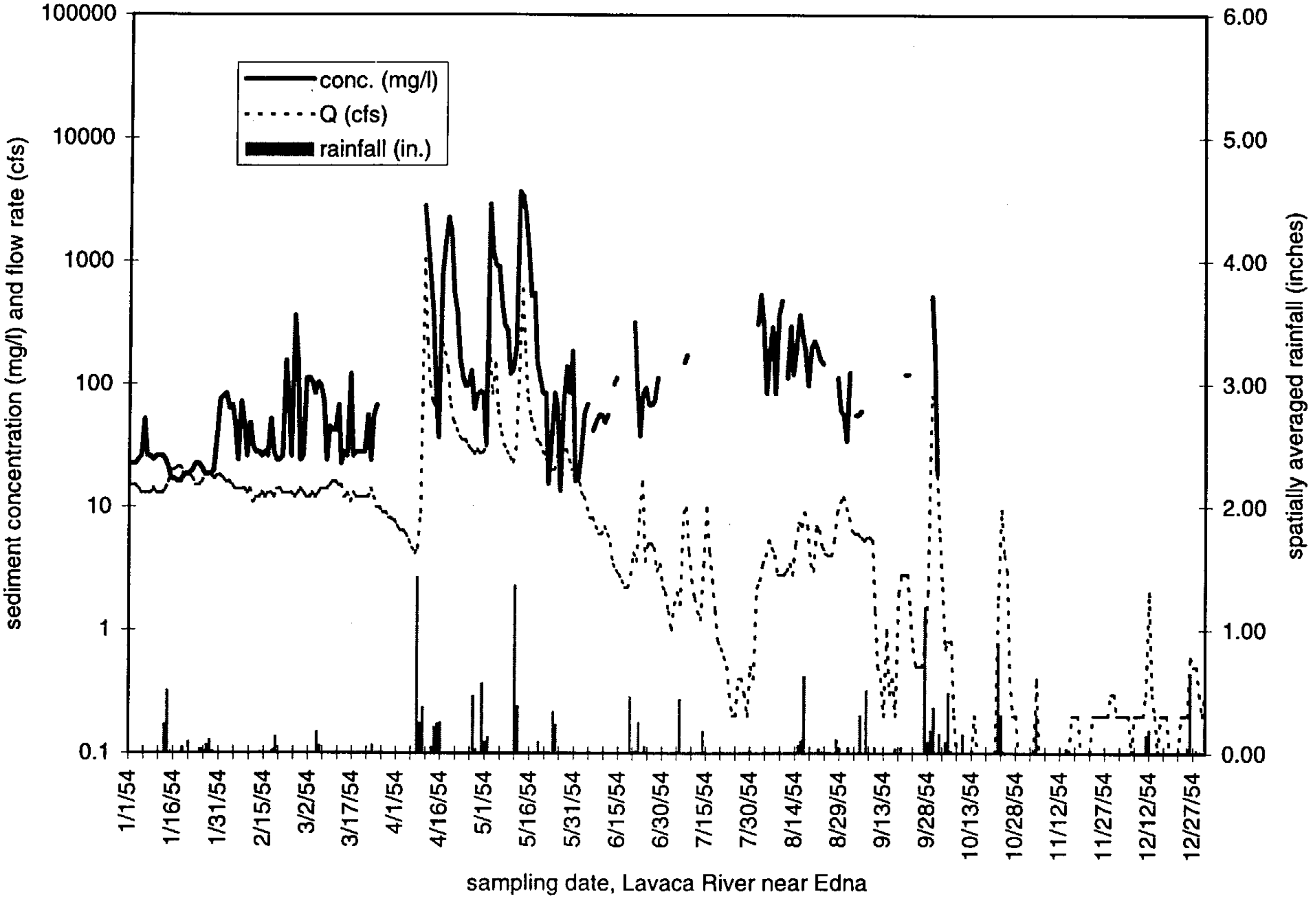


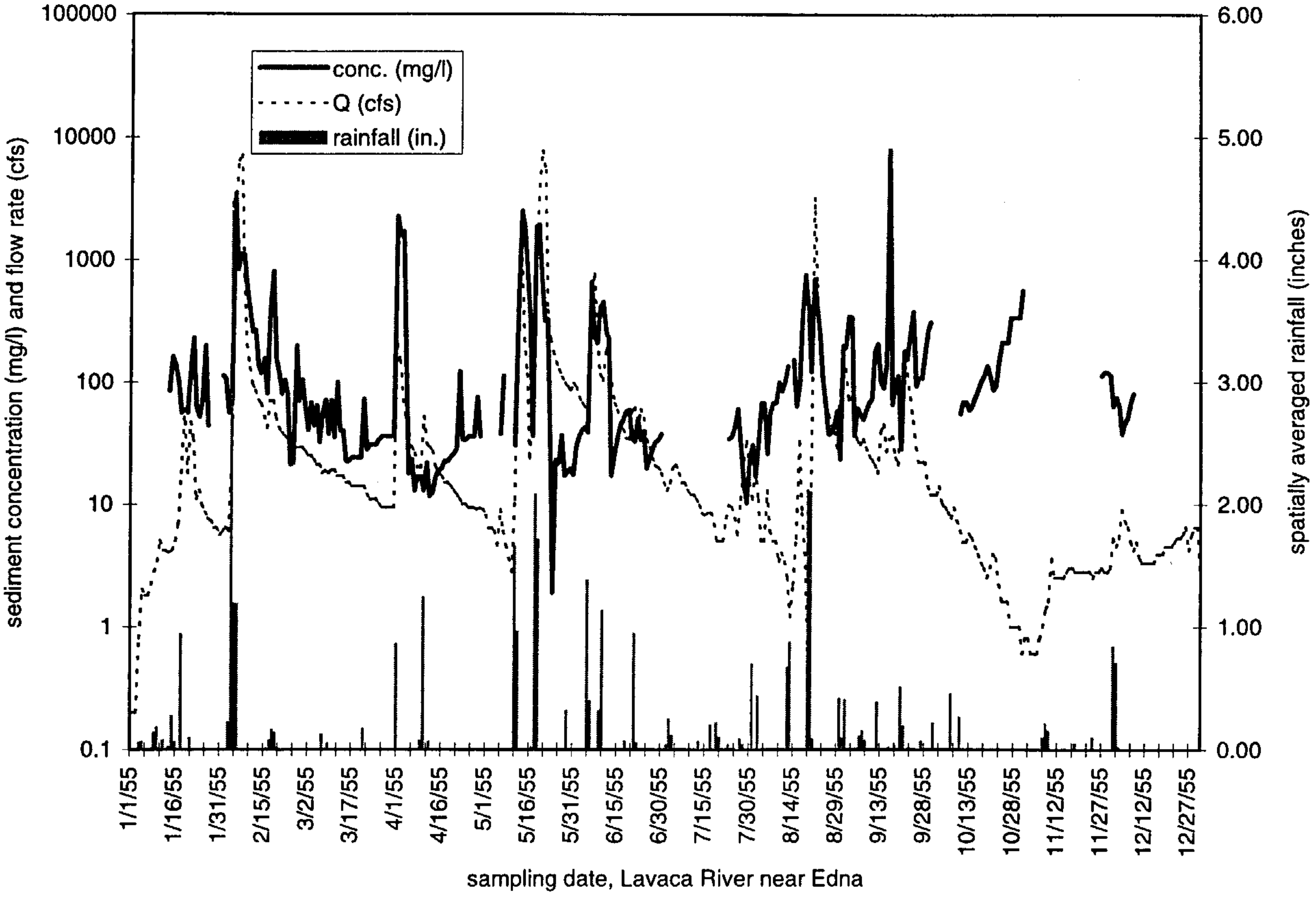
100000
10000
1000
100
10
1
0.1
sediment concentration (mg/lt) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/50
1/16/50
1/31/50
2/15/50
3/2/50
3/17/50
4/1/50
4/16/50
5/1/50
5/16/50
5/31/50
6/15/50
6/30/50
7/15/50
7/30/50
8/14/50
8/29/50
9/13/50
9/28/50
10/13/50
10/28/50
11/12/50
11/27/50
12/12/50
12/27/50
sampling date, Lavaca River near Edna

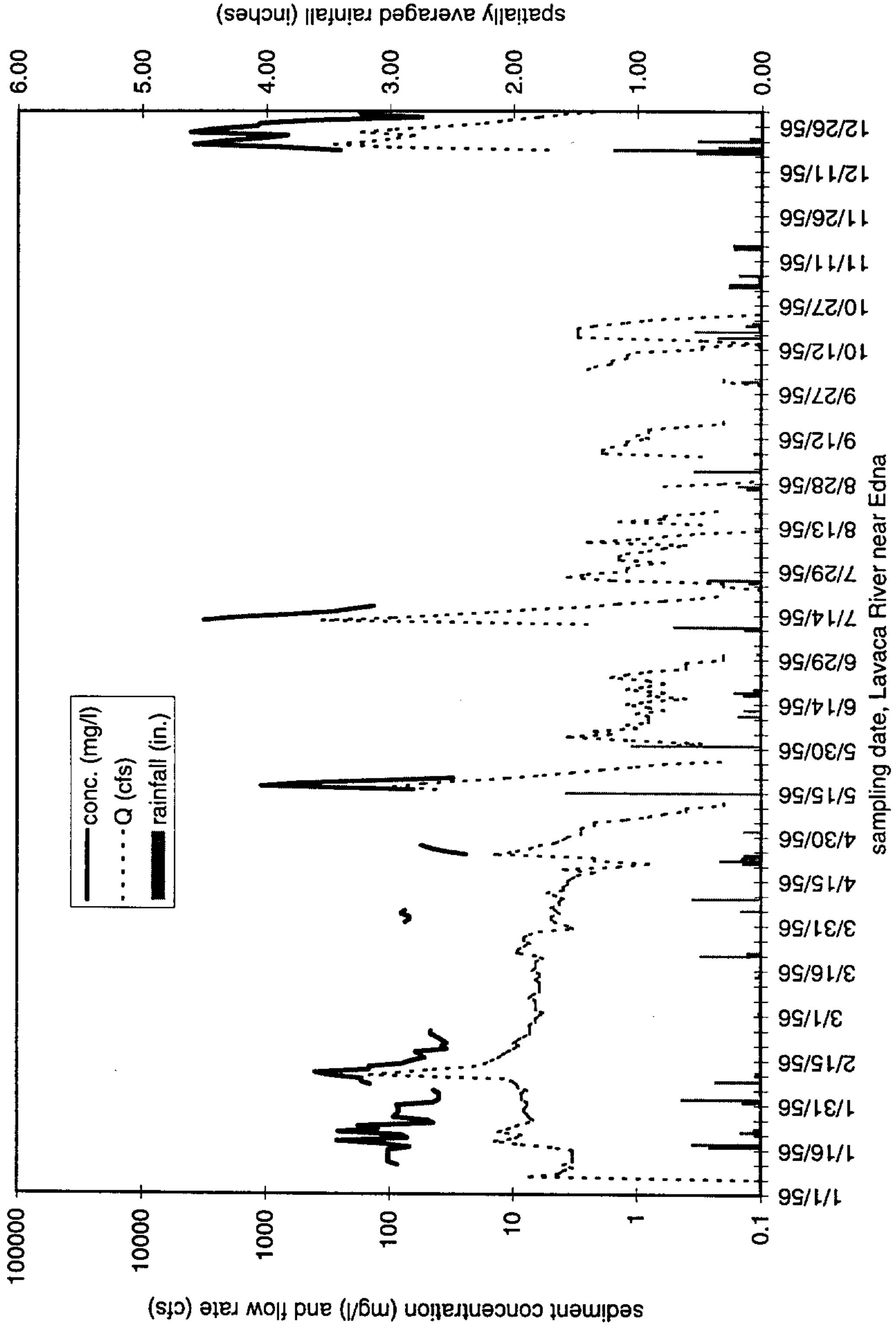

sediment concentration (mg/l) and flow rate (cfs)
0.1
1
10
100
1000
10000
100000
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/51
1/16/51
1/31/51
2/15/51
3/2/51
3/17/51
4/1/51
4/16/51
5/1/51
5/16/51
5/31/51
6/15/51
6/30/51
7/15/51
7/30/51
8/14/51
8/29/51
9/13/51
9/28/51
10/13/51
10/28/51
11/12/51
11/27/51
12/12/51
12/27/51
sampling date, Lavaca River near Edna
0.00
1.00
2.00
3.00
4.00
5.00
6.00
spatially averaged rainfall (inches)

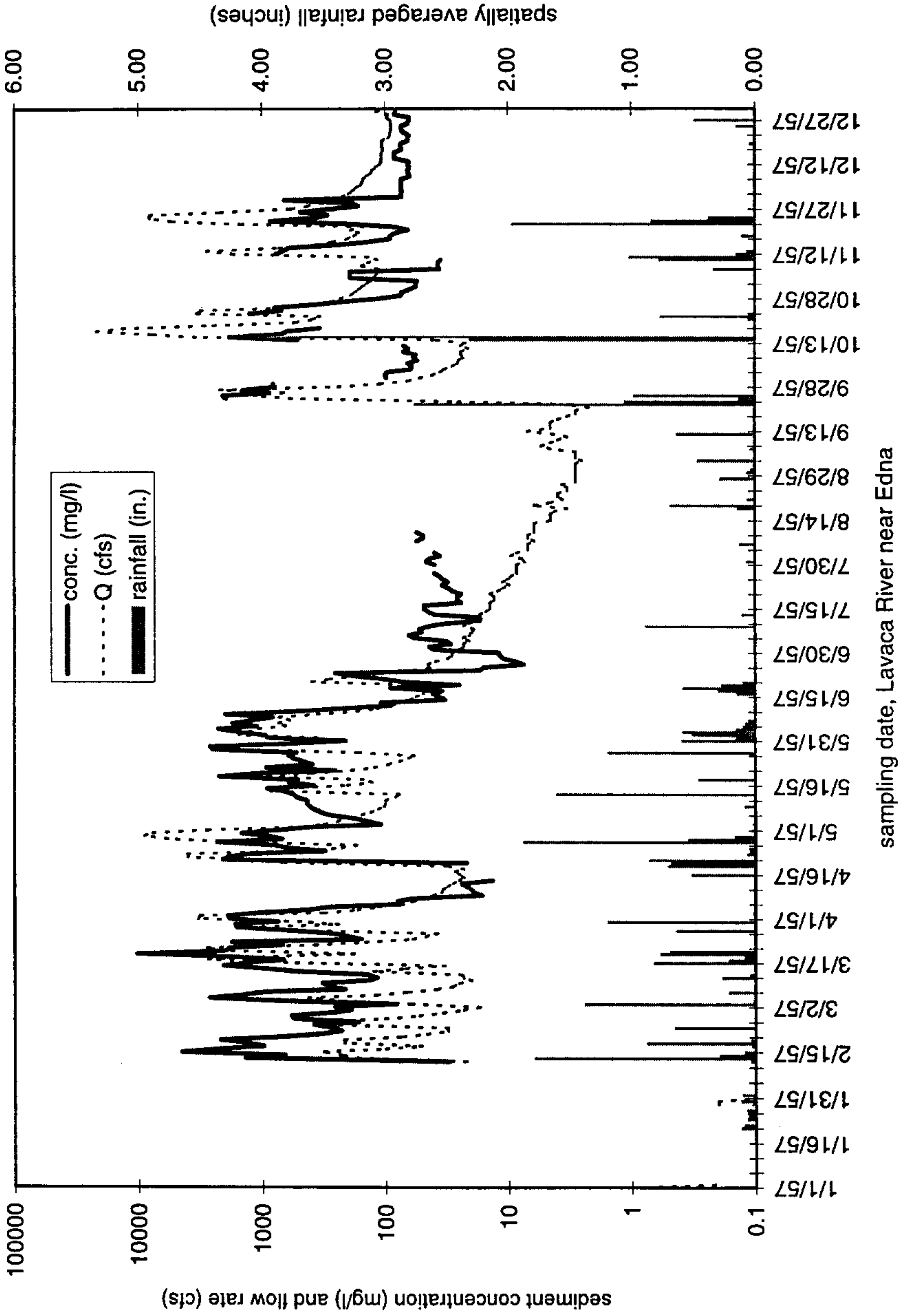

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) anad flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/52
1/16/52
1/31/52
2/15/52
3/1/52
3/16/52
3/31/52
4/15/52
4/30/52
5/15/52
5/30/52
6/14/52
6/29/52
7/14/52
7/29/52
8/13/52
8/28/52
9/12/52
9/27/52
10/12/52
10/27/52
11/11/52
11/26/52
12/11/52
12/26/52
sampling date, Lavaca River near Edna

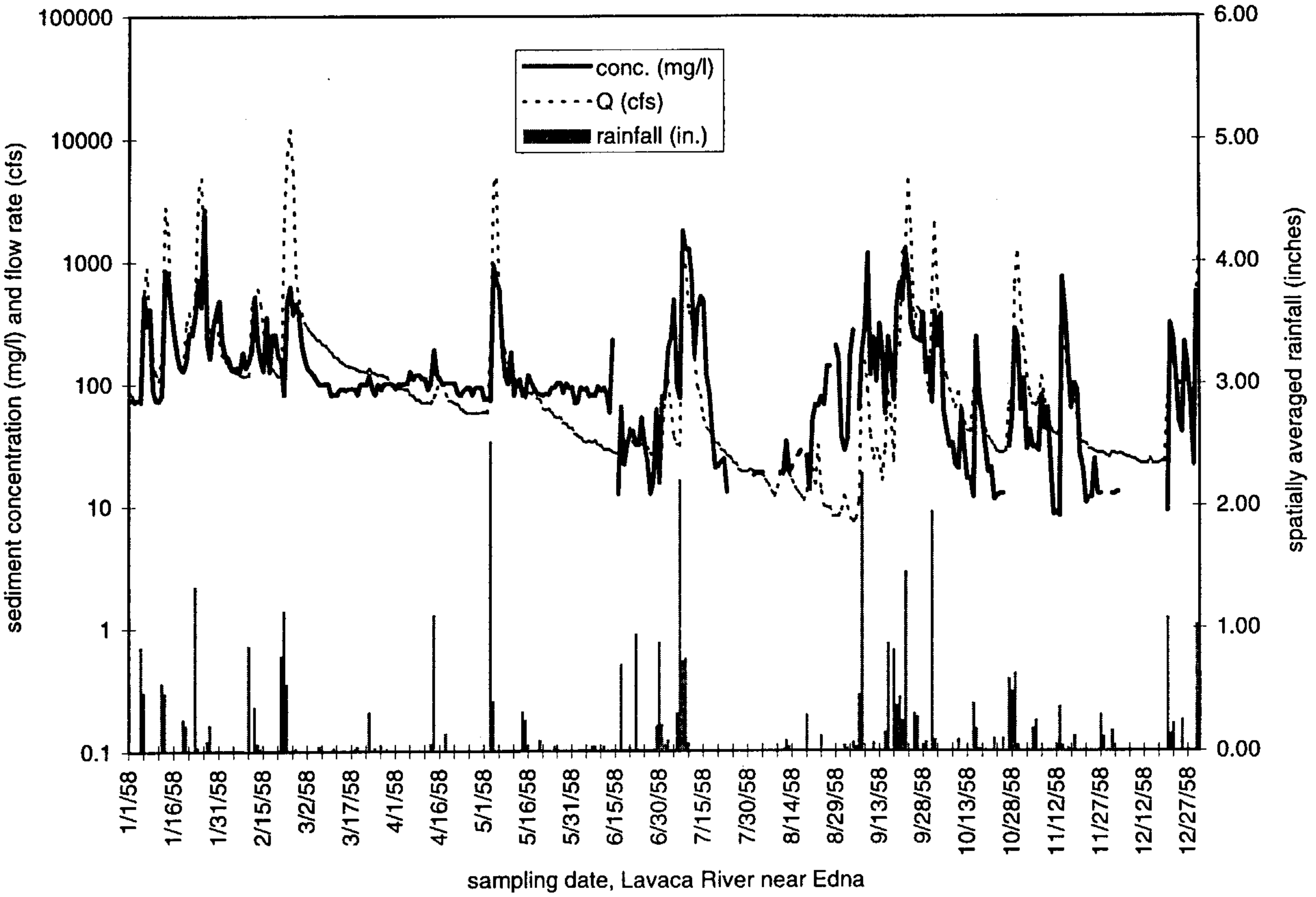

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
sampling date, Lavaca River near Edna
1/1/53
1/16/53
1/31/53
2/15/53
3/2/53
3/17/53
4/1/53
4/16/53
5/1/53
5/16/53
5/31/53
6/15/53
6/30/53
7/15/53
7/30/53
8/14/53
8/29/53
9/13/53
9/28/53
10/13/53
10/28/53
11/12/53
11/27/53
12/12/53
12/27/53

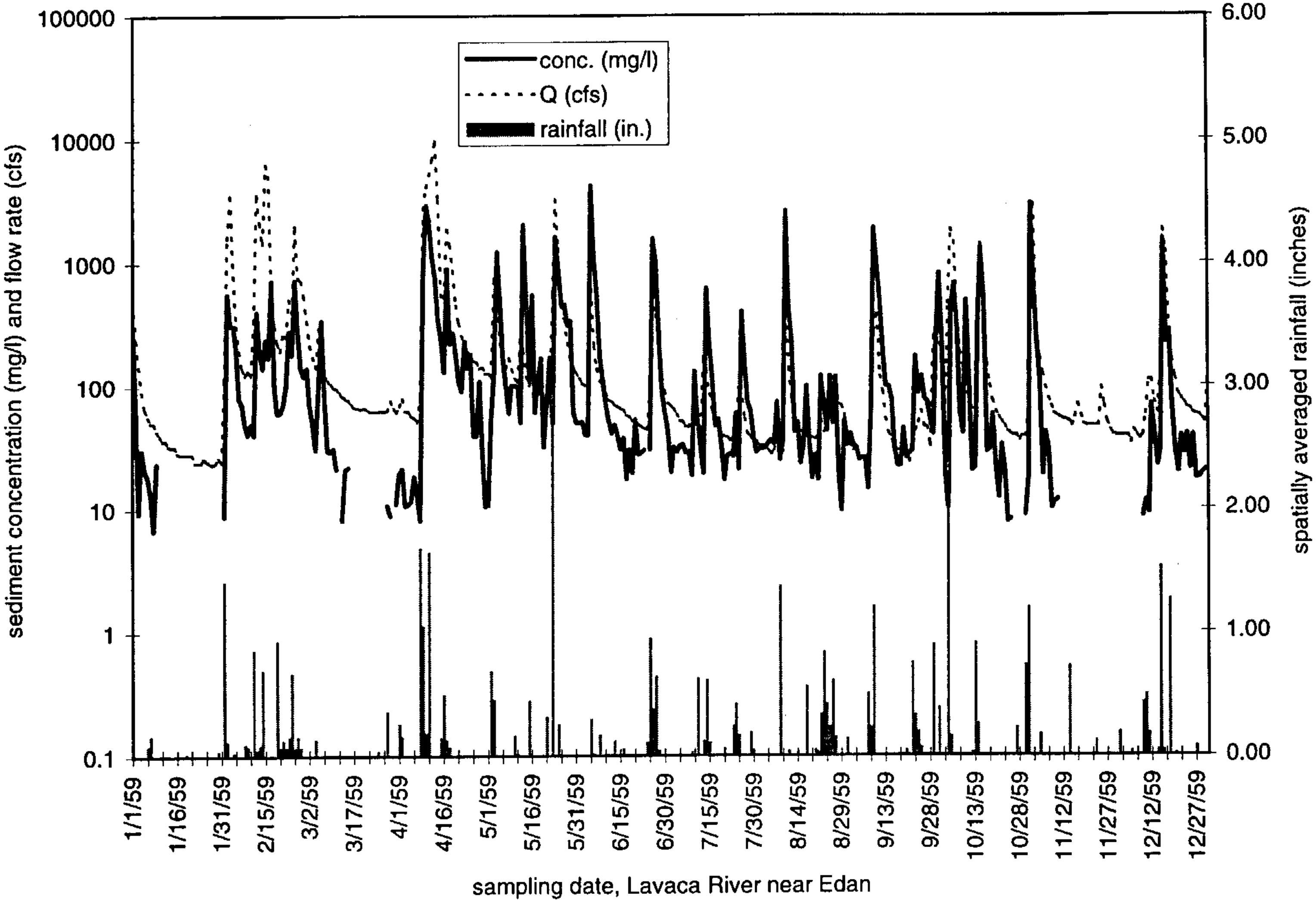

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/54
1/16/54
1/31/54
2/15/54
3/2/54
3/17/54
4/1/54
4/16/54
5/1/54
5/16/54
5/31/54
6/15/54
6/30/54
7/15/54
7/30/54
8/14/54
8/29/54
9/13/54
9/28/54
10/13/54
10/28/54
11/12/54
11/27/54
12/12/54
12/27/54
sampling date, Lavaca River near Edna

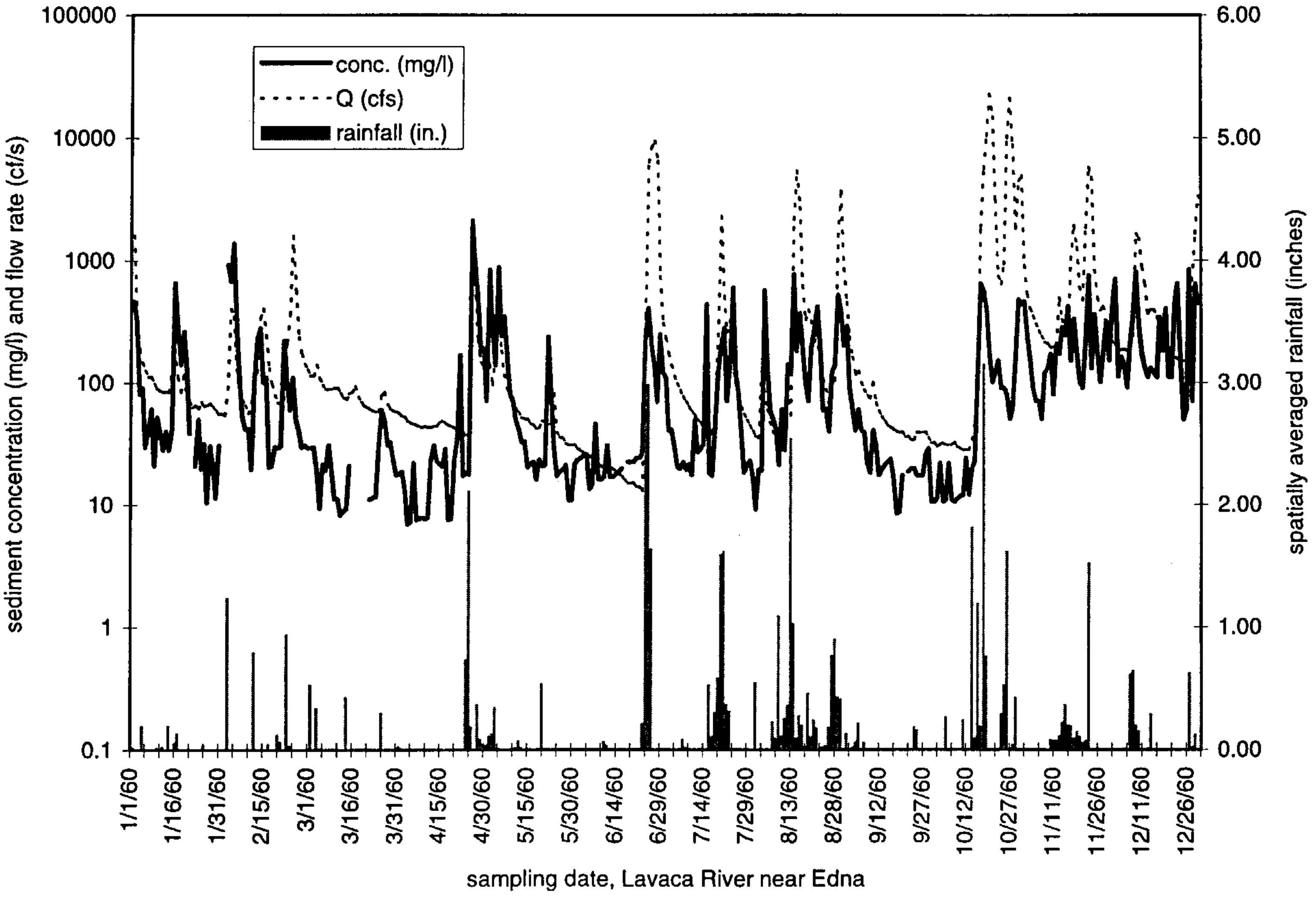

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/55
1/16/55
1/31/55
2/15/55
3/2/55
3/17/55
4/1/55
4/16/55
5/1/55
5/16/55
5/31/55
6/15/55
6/30/55
7/15/55
7/30/55
8/14/55
8/29/55
9/13/55
9/28/55
10/13/55
10/28/55
11/12/55
11/27/55
12/12/55
12/27/55
sampling date, Lavaca River near Edna

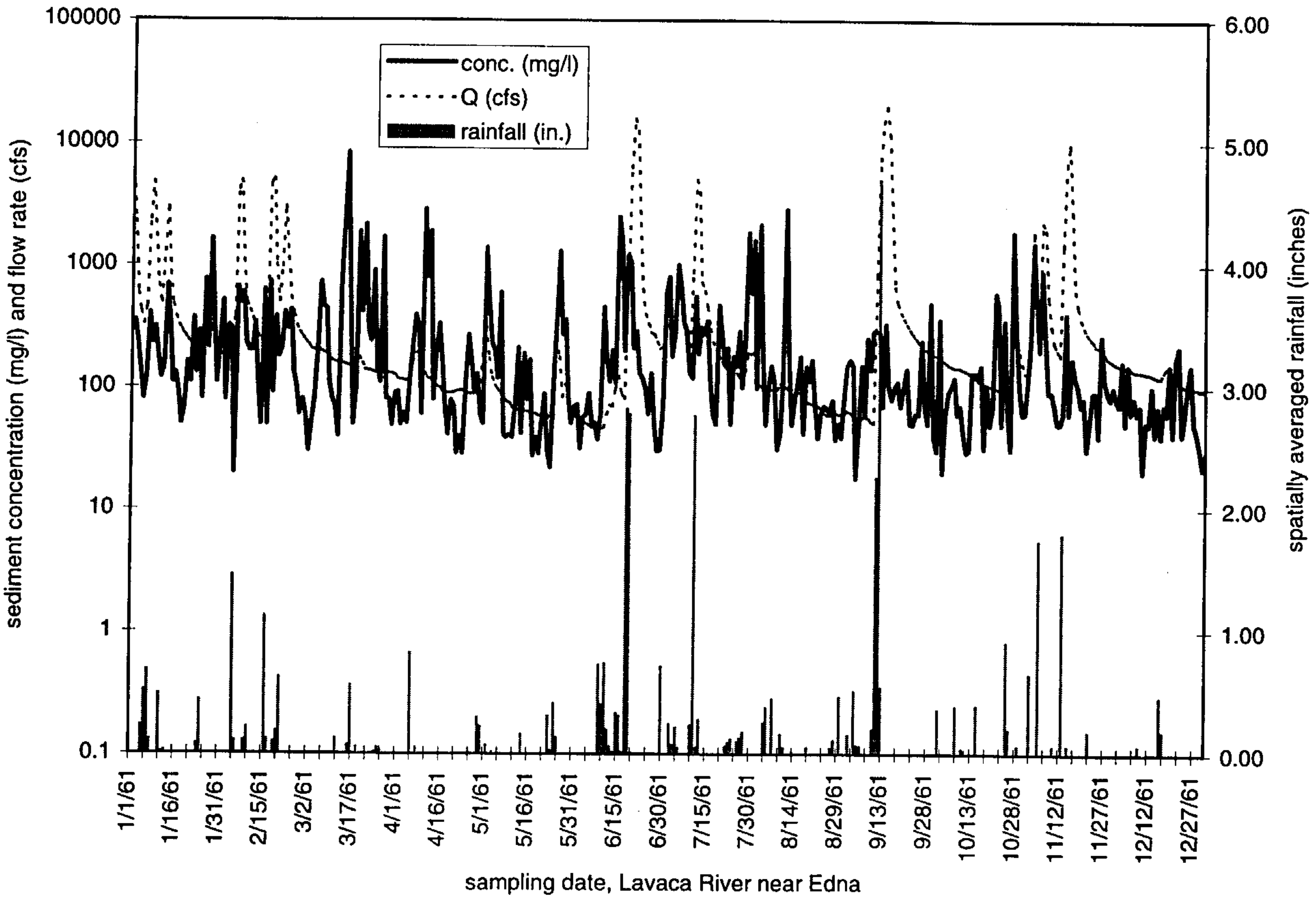

sediment concentration (mg/l) and flow rate (cfs)
0.1
1
10
100
1000
10000
100000
spatially averaged rainfall (inches)
0.00
1.00
2.00
3.00
4.00
5.00
6.00
sampling date, Lavaca River near Edna
1/1/56
1/16/56
1/31/56
2/15/56
3/1/56
3/16/56
3/31/56
4/15/56
4/30/56
5/15/56
5/30/56
6/14/56
6/29/56
7/14/56
7/29/56
8/13/56
8/28/56
9/12/56
9/27/56
10/12/56
10/27/56
11/11/56
11/26/56
12/11/56
12/26/56
conc. (mg/l)
Q (cfs)
rainfall (in.)

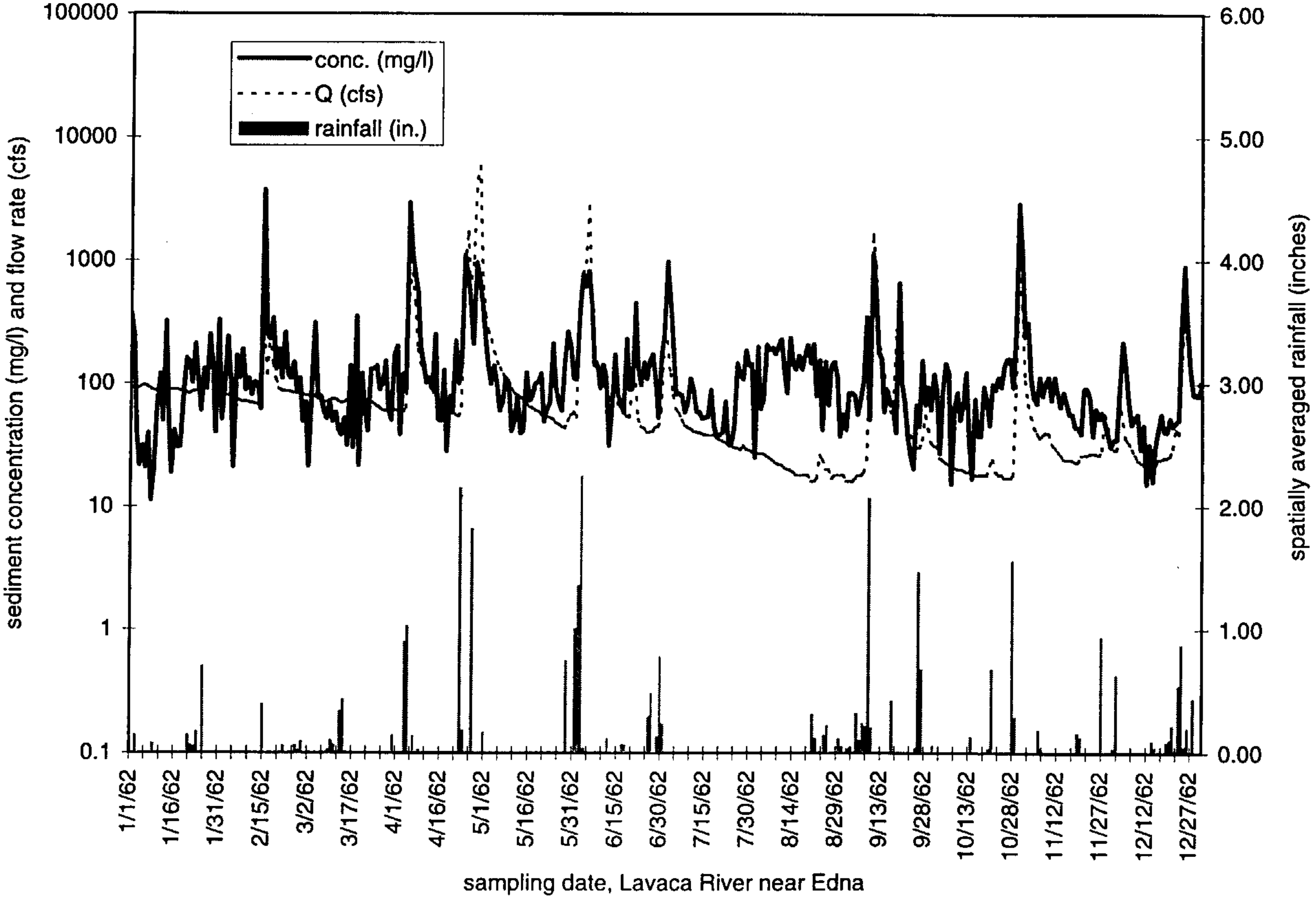

sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/57
1/16/57
1/31/57
2/15/57
3/2/57
3/17/57
4/1/57
4/16/57
5/1/57
5/16/57
5/31/57
6/15/57
6/30/57
7/15/57
7/30/57
8/14/57
8/29/57
9/13/57
9/28/57
10/13/57
10/28/57
11/12/57
11/27/57
12/12/57
12/27/57
sampling date, Lavaca River near Edna
conc. (mg/l)
Q (cfs)
rainfall (in.)

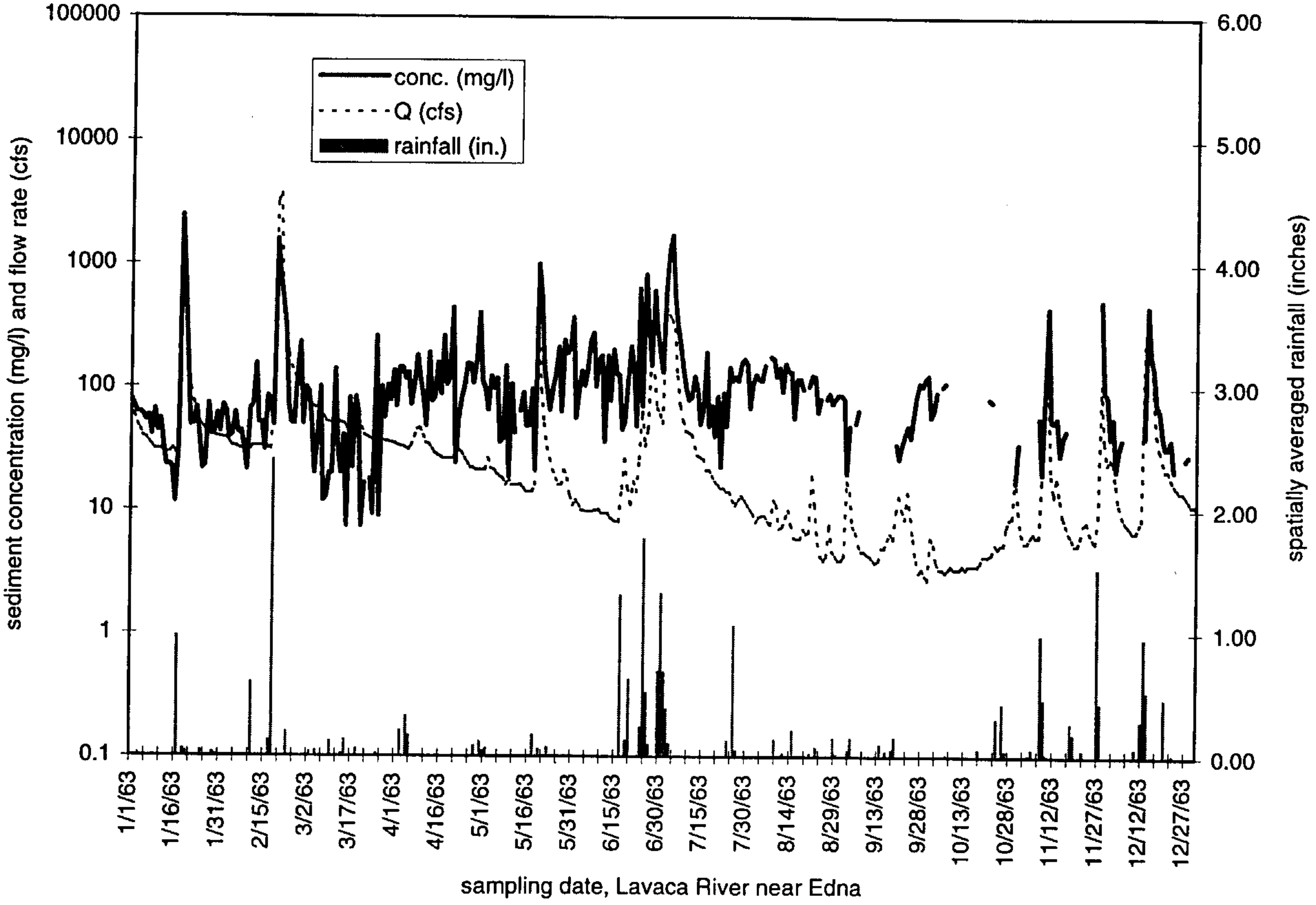

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/58
1/16/58
1/31/58
2/15/58
3/2/58
3/17/58
4/1/58
4/16/58
5/1/58
5/16/58
5/31/58
6/15/58
6/30/58
7/15/58
7/30/58
8/14/58
8/29/58
9/13/58
9/28/58
10/13/58
10/28/58
11/12/58
11/27/58
12/12/58
12/27/58
sampling date, Lavaca River near Edna

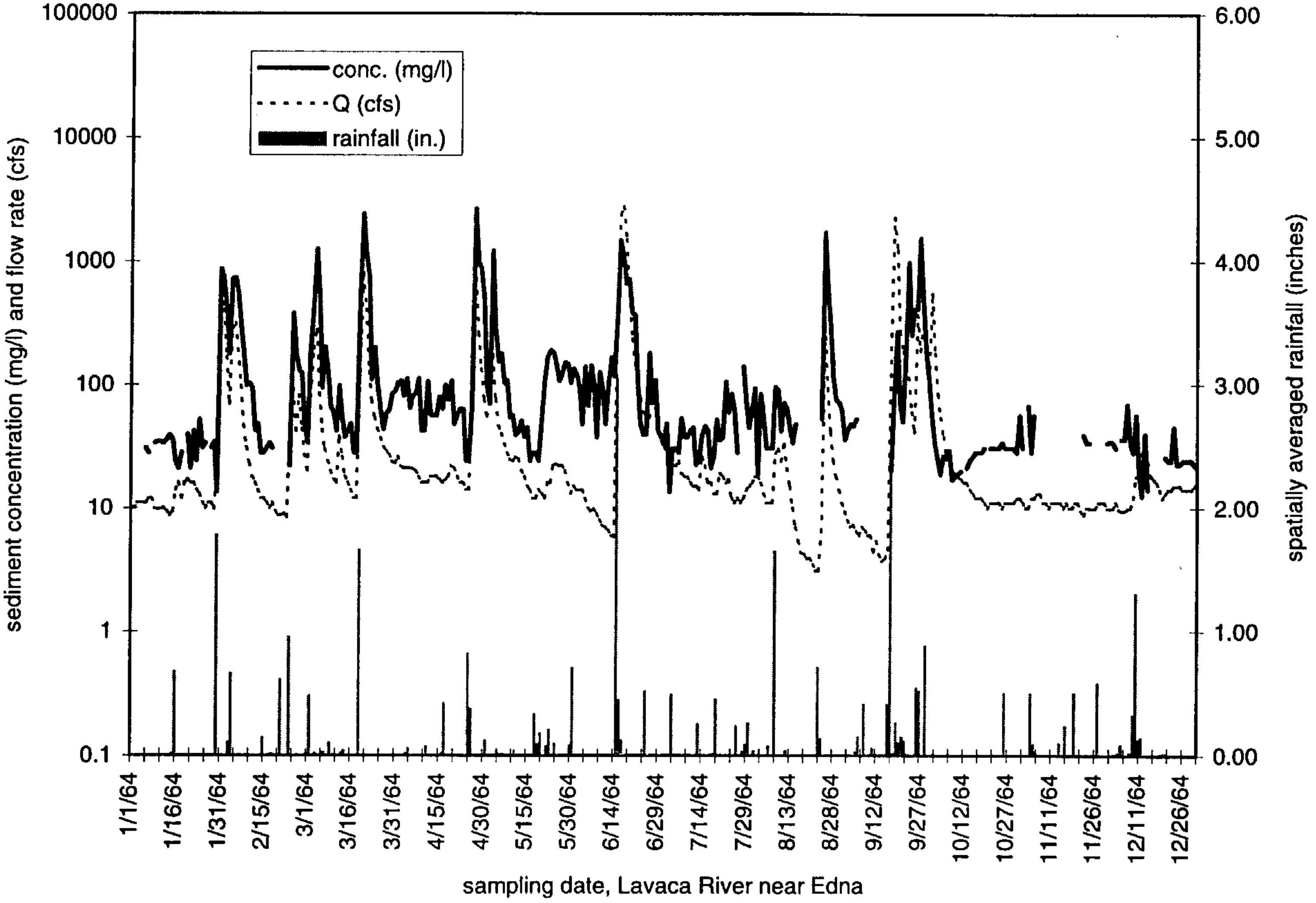

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/59
1/16/59
1/31/59
2/15/59
3/2/59
3/17/59
4/1/59
4/16/59
5/1/59
5/16/59
5/31/59
6/15/59
6/30/59
7/15/59
7/30/59
8/14/59
8/29/59
9/13/59
9/28/59
10/13/59
10/28/59
11/12/59
11/27/59
12/12/59
12/27/59
sampling date, Lavaca River near Edan

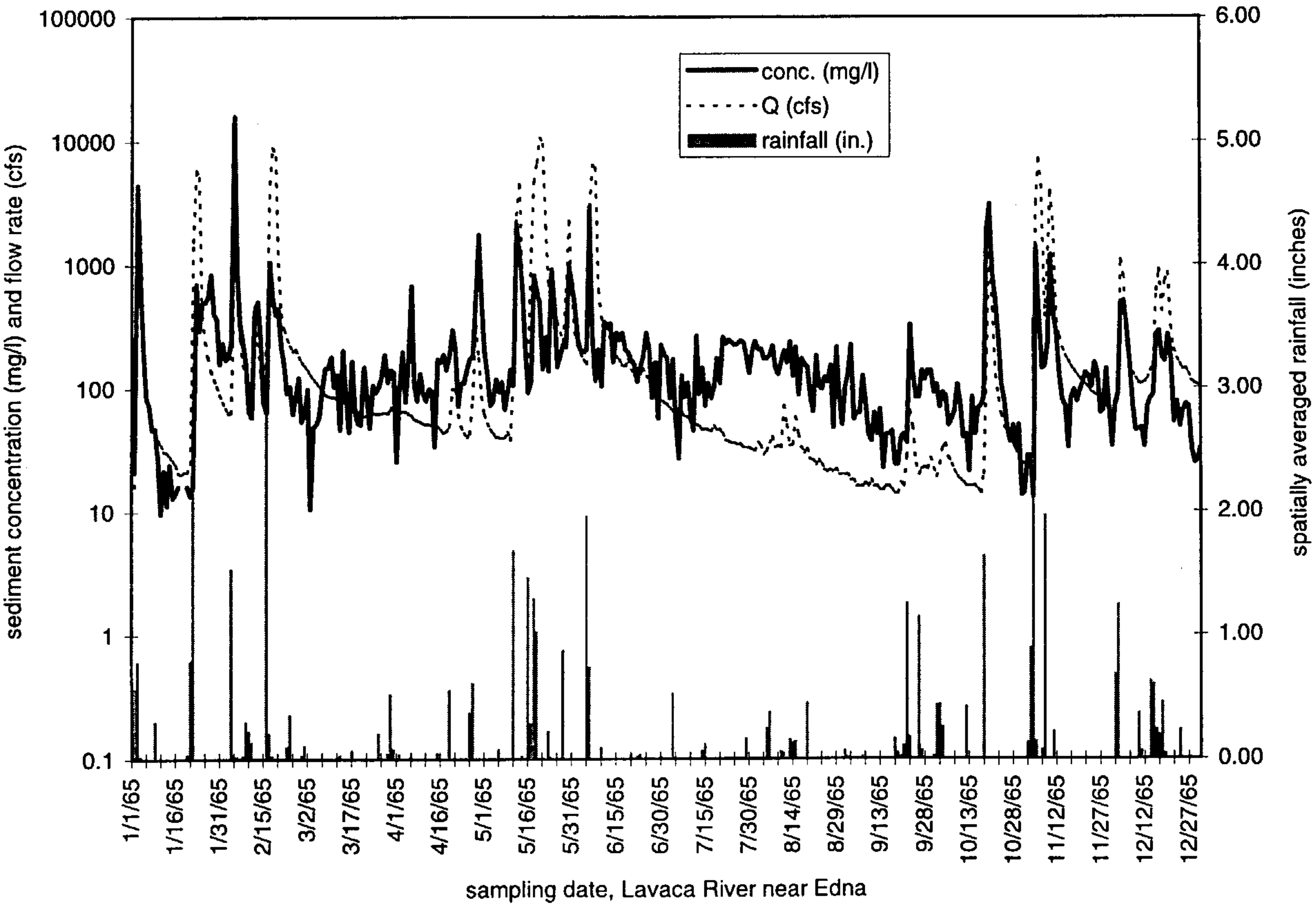

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cf/s)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/60
1/16/60
1/31/60
2/15/60
3/1/60
3/16/60
3/31/60
4/15/60
4/30/60
5/15/60
5/30/60
6/14/60
6/29/60
7/14/60
7/29/60
8/13/60
8/28/60
9/12/60
9/27/60
10/12/60
10/27/60
11/11/60
11/26/60
12/11/60
12/26/60
sampling date, Lavaca River near Edna

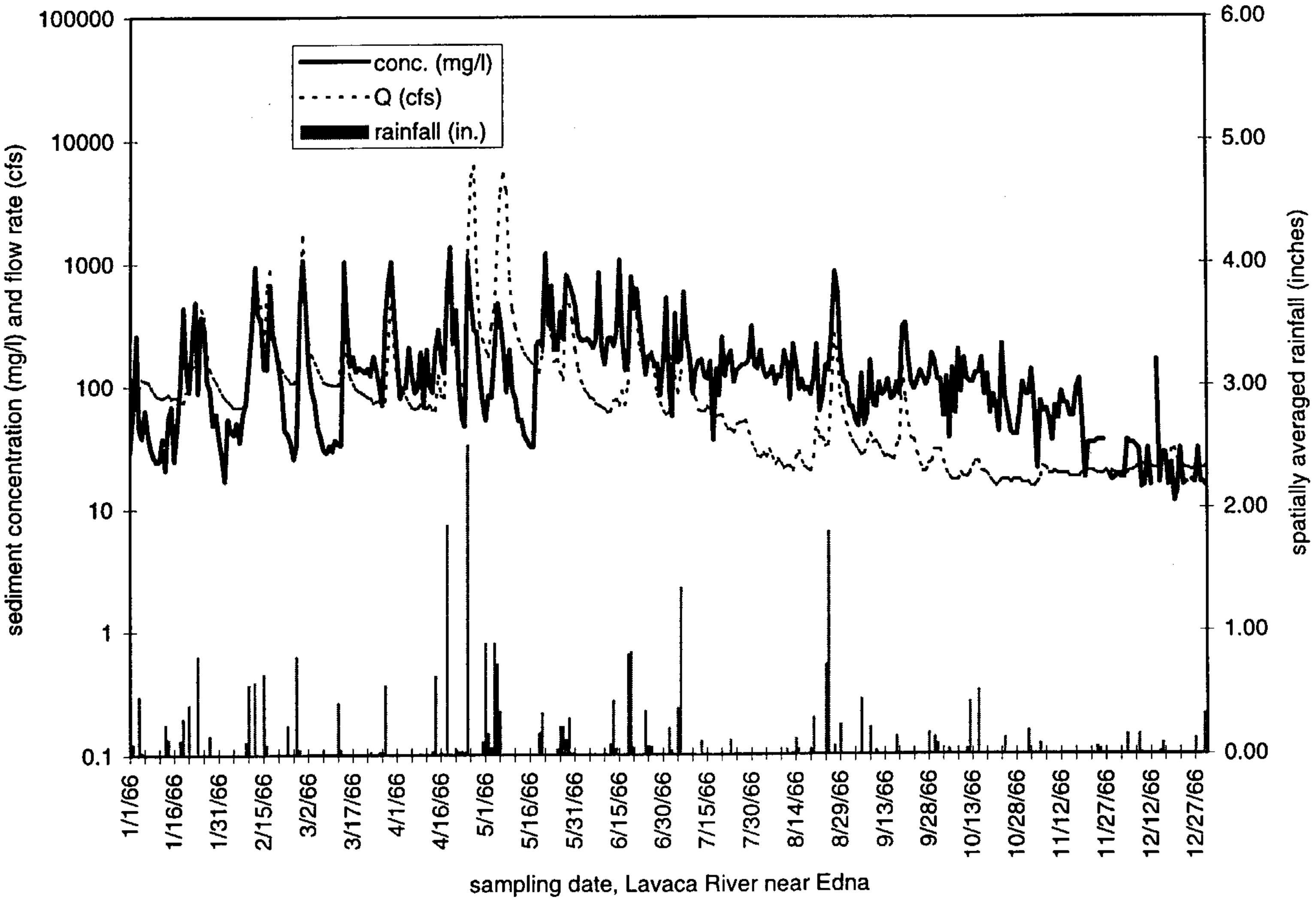

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/61
1/16/61
1/31/61
2/15/61
3/2/61
3/17/61
4/1/61
4/16/61
5/1/61
5/16/61
5/31/61
6/15/61
6/30/61
7/15/61
7/30/61
8/14/61
8/29/61
9/13/61
9/28/61
10/13/61
10/28/61
11/12/61
11/27/61
12/12/61
12/27/61
sampling date, Lavaca River near Edna

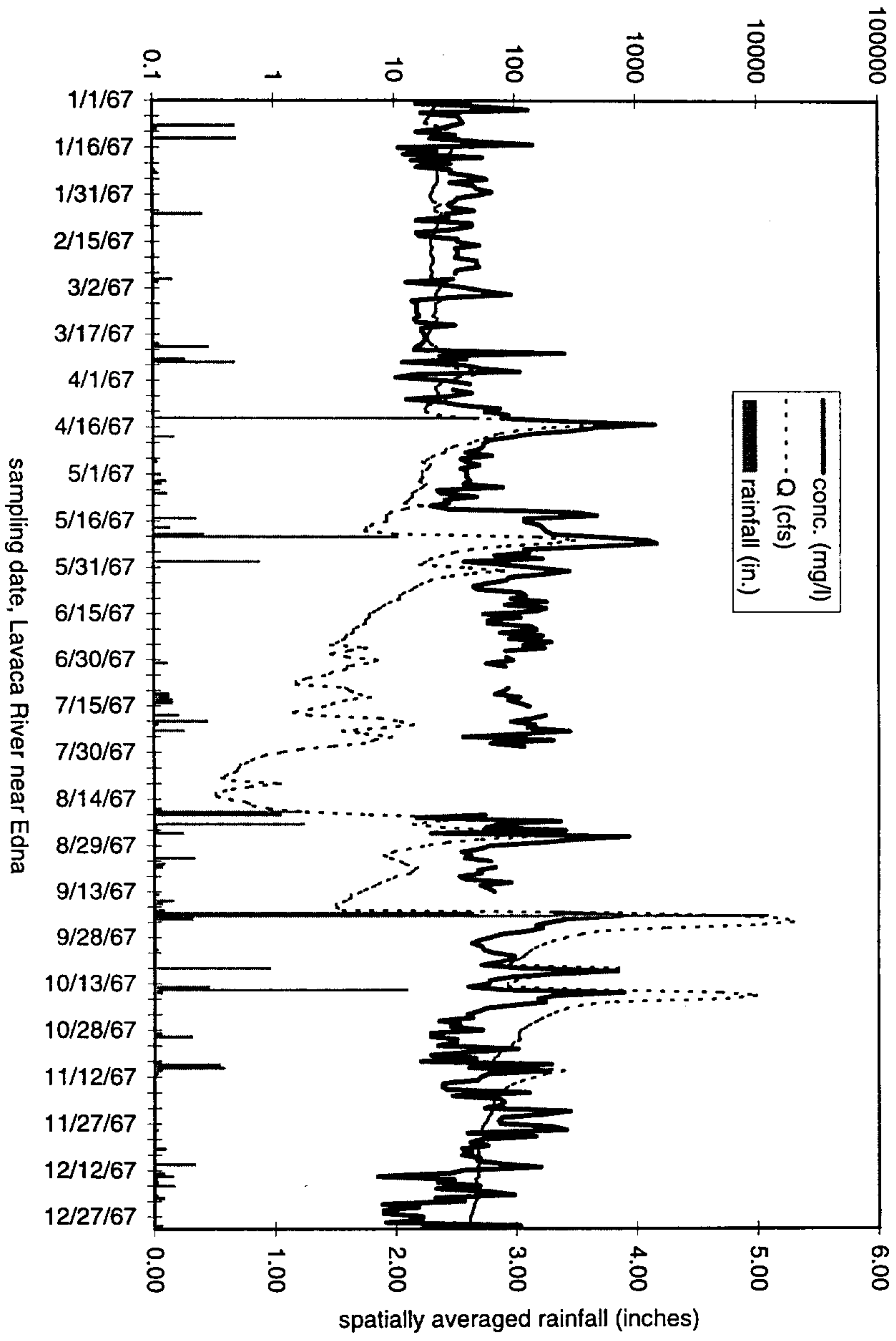

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/62
1/16/62
1/31/62
2/15/62
3/2/62
3/17/62
4/1/62
4/16/62
5/1/62
5/16/62
5/31/62
6/15/62
6/30/62
7/15/62
7/30/62
8/14/62
8/29/62
9/13/62
9/28/62
10/13/62
10/28/62
11/12/62
11/27/62
12/12/62
12/27/62
sampling date, Lavaca River near Edna

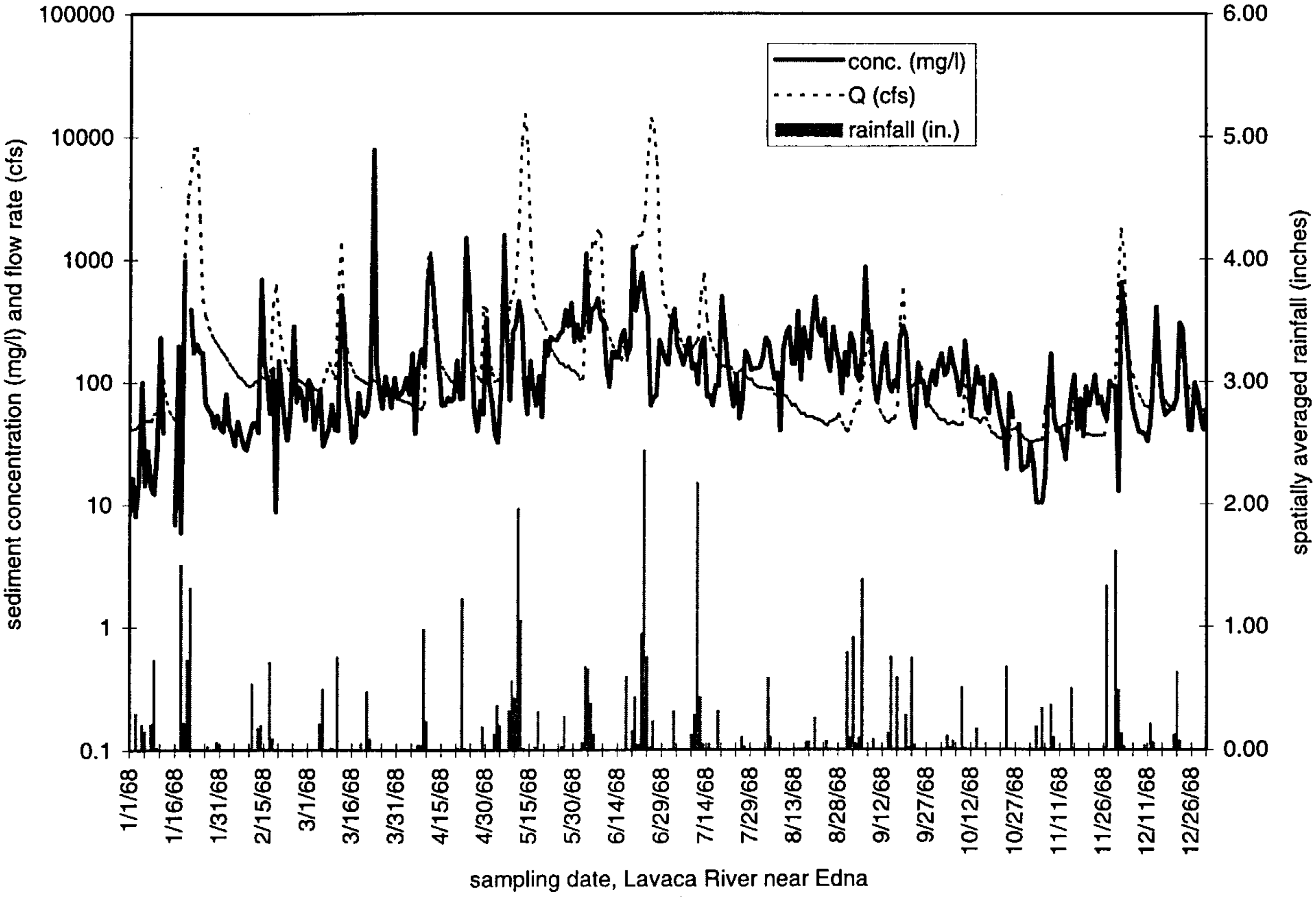

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/63
1/16/63
1/31/63
2/15/63
3/2/63
3/17/63
4/1/63
4/16/63
5/1/63
5/16/63
5/31/63
6/15/63
6/30/63
7/15/63
7/30/63
8/14/63
8/29/63
9/13/63
9/28/63
10/13/63
10/28/63
11/12/63
11/27/63
12/12/63
12/27/63
sampling date, Lavaca River near Edna

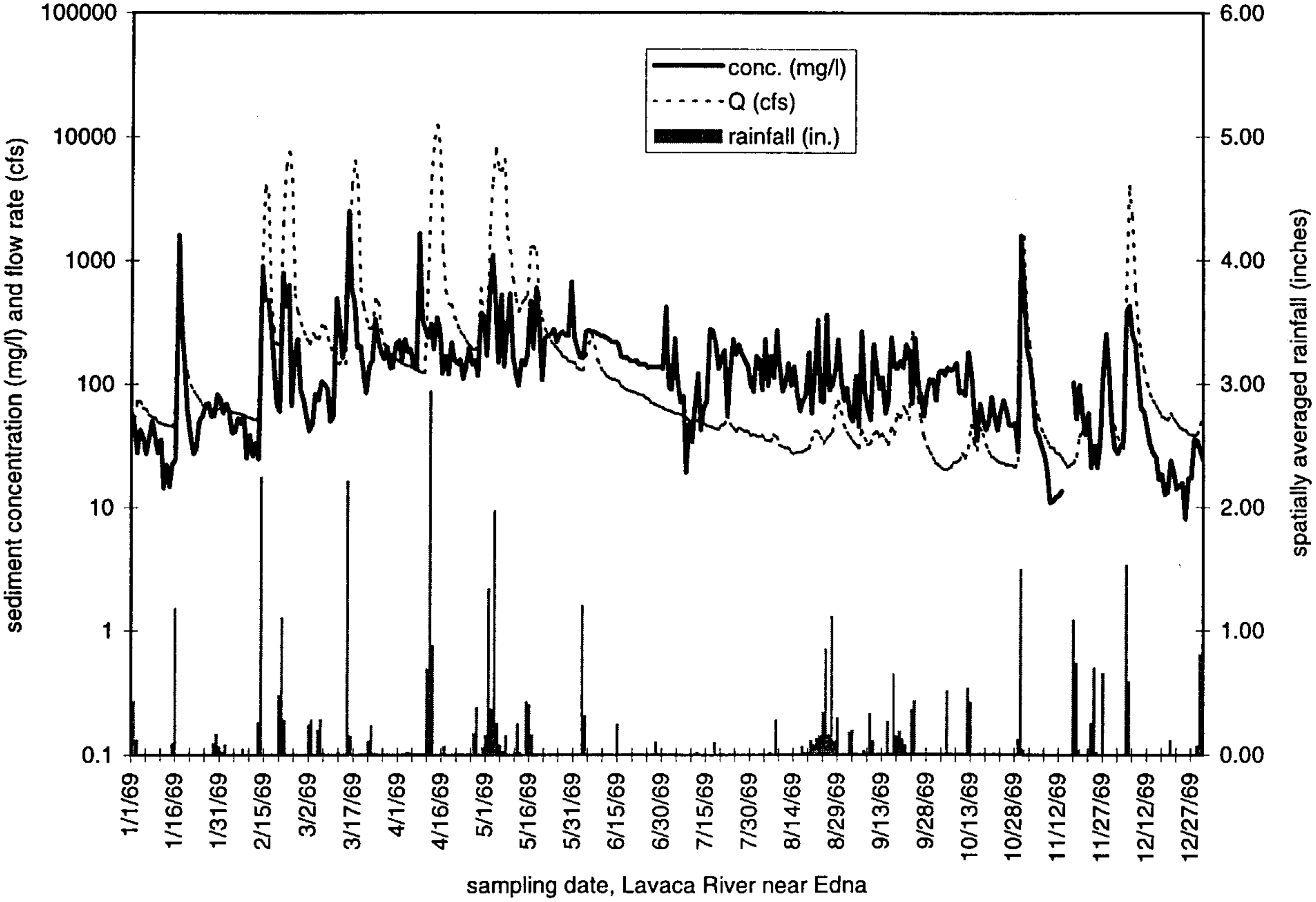

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/64
1/16/64
1/31/64
2/15/64
3/1/64
3/16/64
3/31/64
4/15/64
4/30/64
5/15/64
5/30/64
6/14/64
6/29/64
7/14/64
7/29/64
8/13/64
8/28/64
9/12/64
9/27/64
10/12/64
10/27/64
11/11/64
11/26/64
12/11/64
12/26/64
sampling date, Lavaca River near Edna

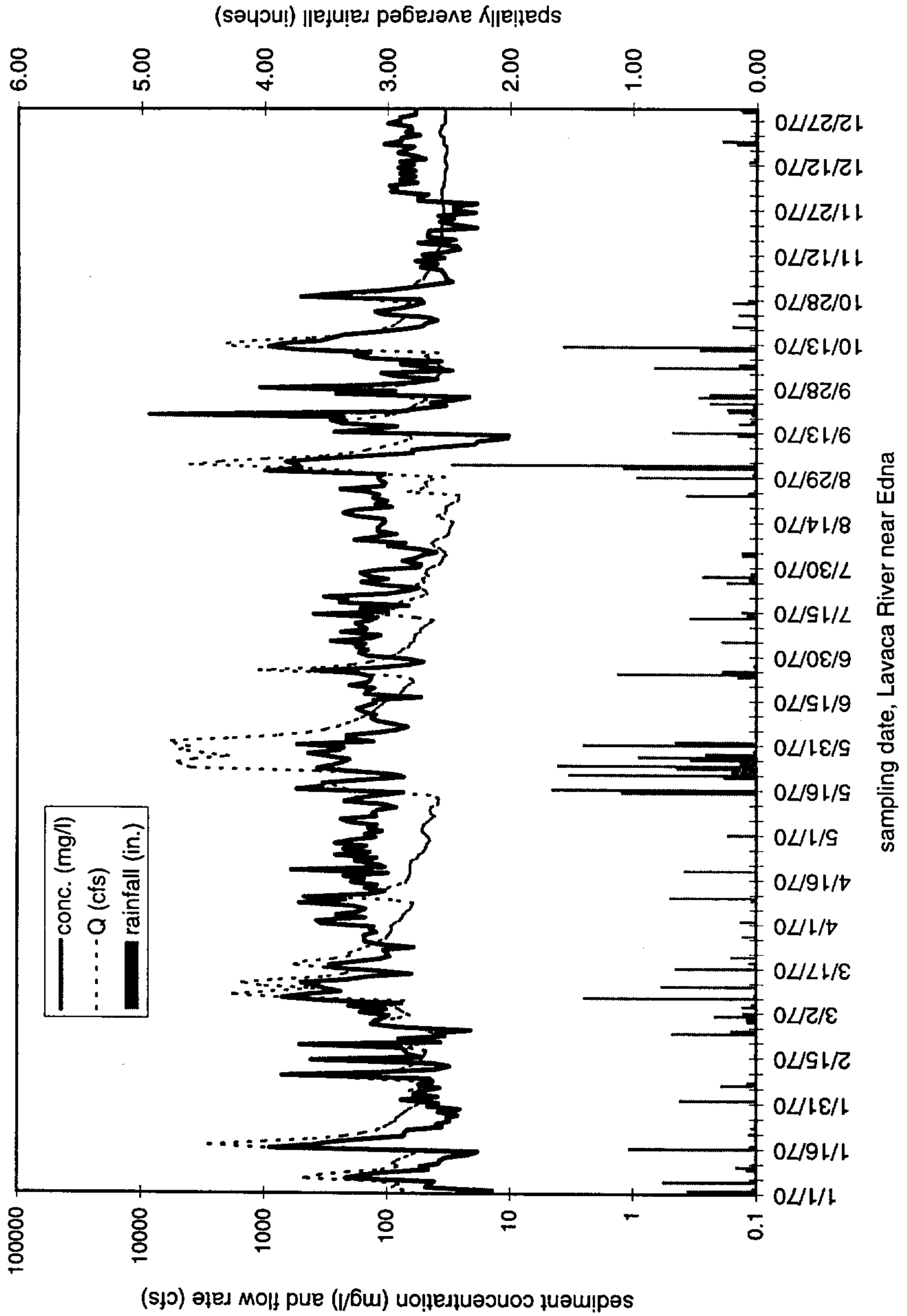

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/65
1/16/65
1/31/65
2/15/65
3/2/65
3/17/65
4/1/65
4/16/65
5/1/65
5/16/65
5/31/65
6/15/65
6/30/65
7/15/65
7/30/65
8/14/65
8/29/65
9/13/65
9/28/65
10/13/65
10/28/65
11/12/65
11/27/65
12/12/65
12/27/65
sampling date, Lavaca River near Edna

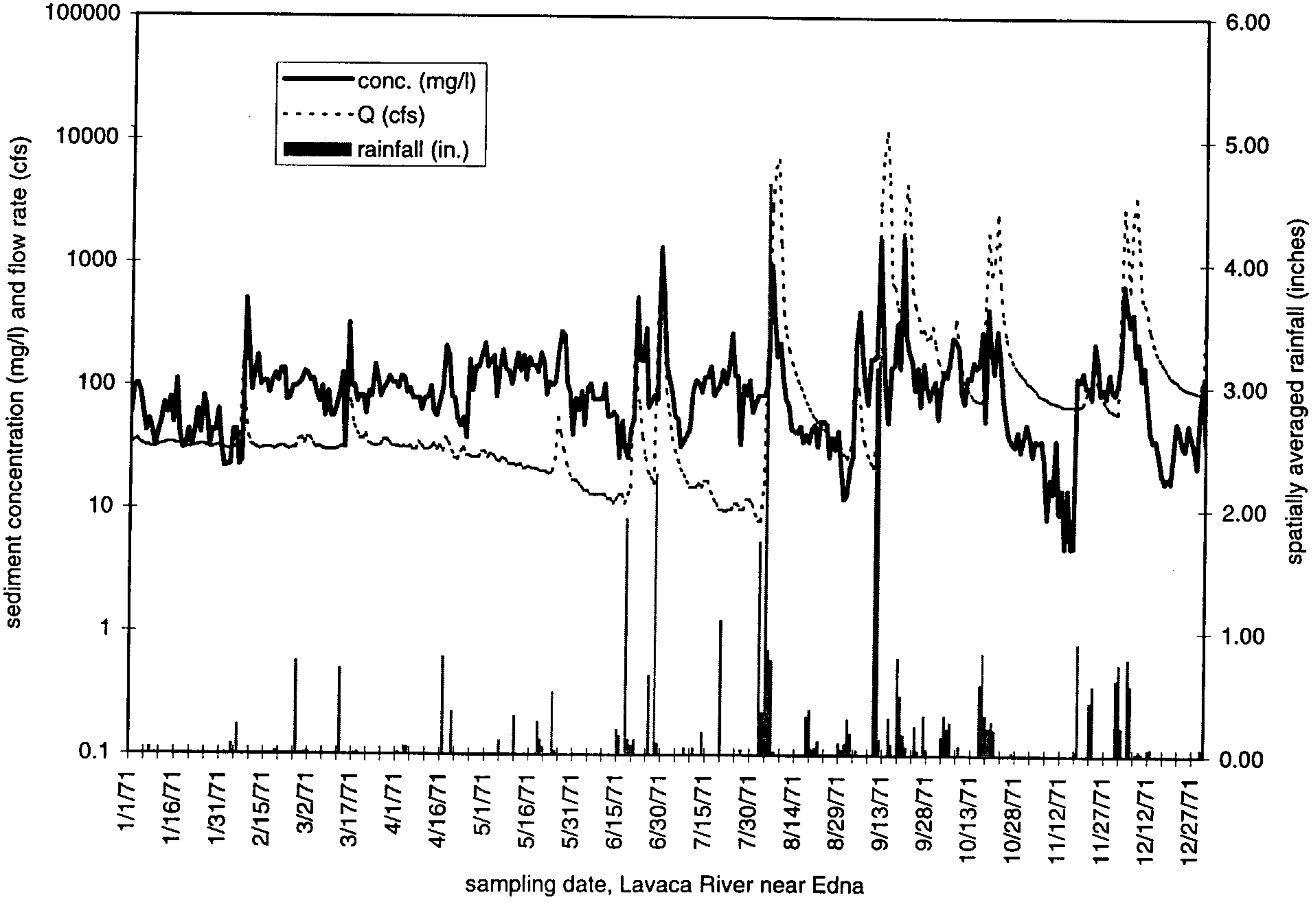

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/66
1/16/66
1/31/66
2/15/66
3/2/66
3/17/66
4/1/66
4/16/66
5/1/66
5/16/66
5/31/66
6/15/66
6/30/66
7/15/66
7/30/66
8/14/66
8/29/66
9/13/66
9/28/66
10/13/66
10/28/66
11/12/66
11/27/66
12/12/66
12/27/66
sampling date, Lavaca River near Edna


sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/67
1/16/67
1/31/67
2/15/67
3/2/67
3/17/67
4/1/67
4/16/67
5/1/67
5/16/67
5/31/67
6/15/67
6/30/67
7/15/67
7/30/67
8/14/67
8/29/67
9/13/67
9/28/67
10/13/67
10/28/67
11/12/67
11/27/67
12/12/67
12/27/67
sampling date, Lavaca River near Edna
0.00
1.00
2.00
3.00
4.00
5.00
6.00
spatially averaged rainfall (inches)

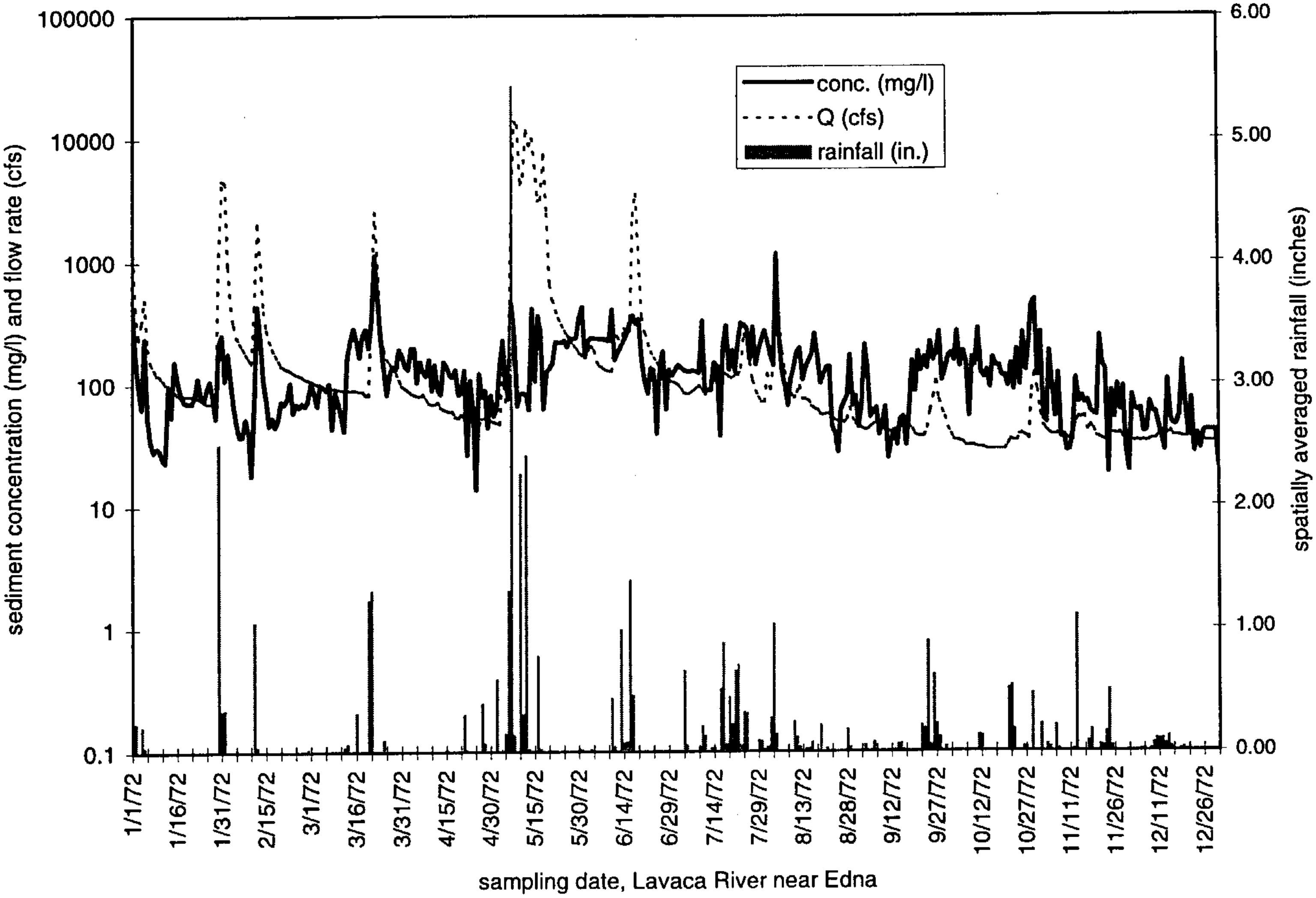

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/68
1/16/68
1/31/68
2/15/68
3/1/68
3/16/68
3/31/68
4/15/68
4/30/68
5/15/68
5/30/68
6/14/68
6/29/68
7/14/68
7/29/68
8/13/68
8/28/68
9/12/68
9/27/68
10/12/68
10/27/68
11/11/68
11/26/68
12/11/68
12/26/68
sampling date, Lavaca River near Edna

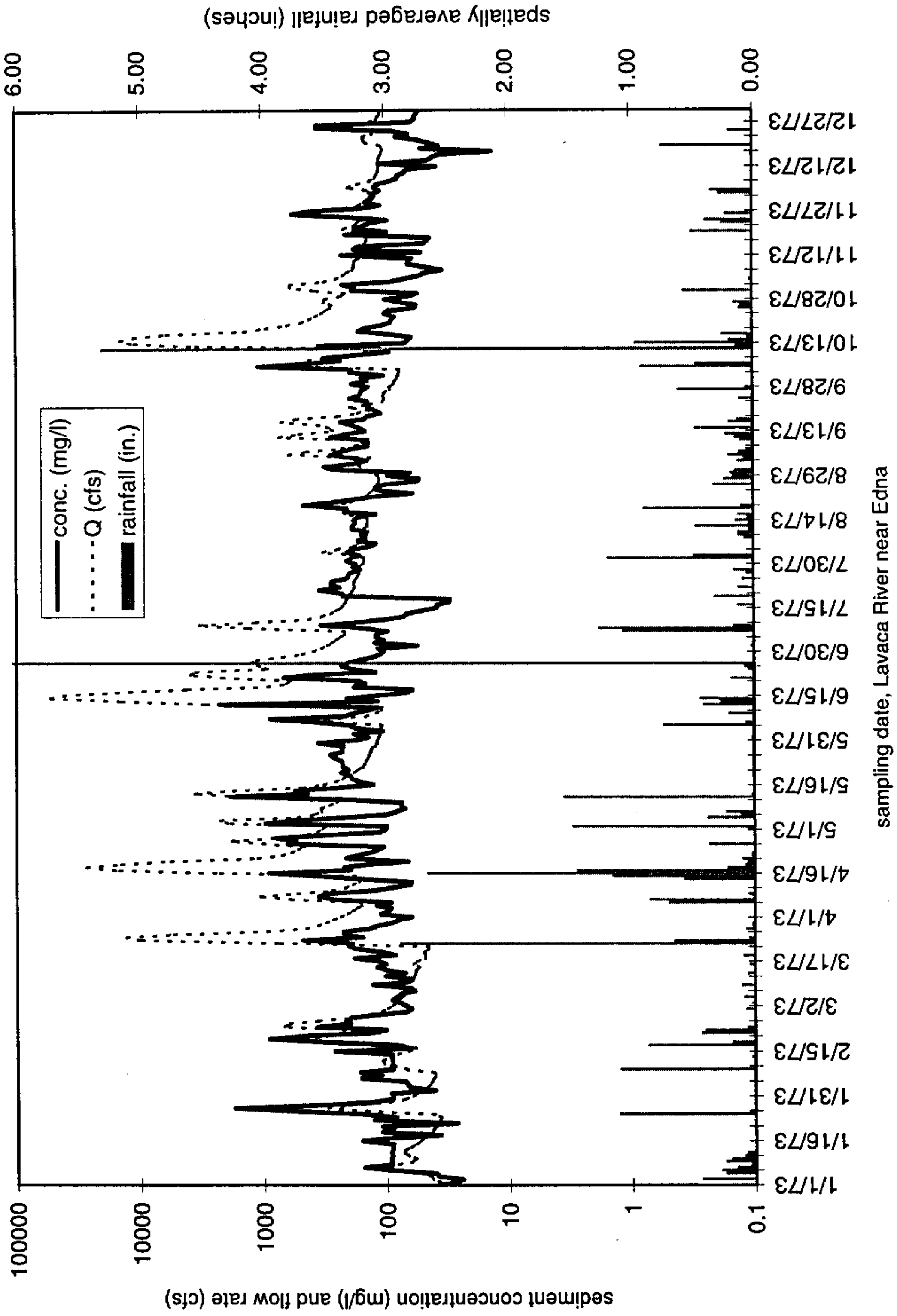

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/69
1/16/69
1/31/69
2/15/69
3/2/69
3/17/69
4/1/69
4/16/69
5/1/69
5/16/69
5/31/69
6/15/69
6/30/69
7/15/69
7/30/69
8/14/69
8/29/69
9/13/69
9/28/69
10/13/69
10/28/69
11/12/69
11/27/69
12/12/69
12/27/69
sampling date, Lavaca River near Edna

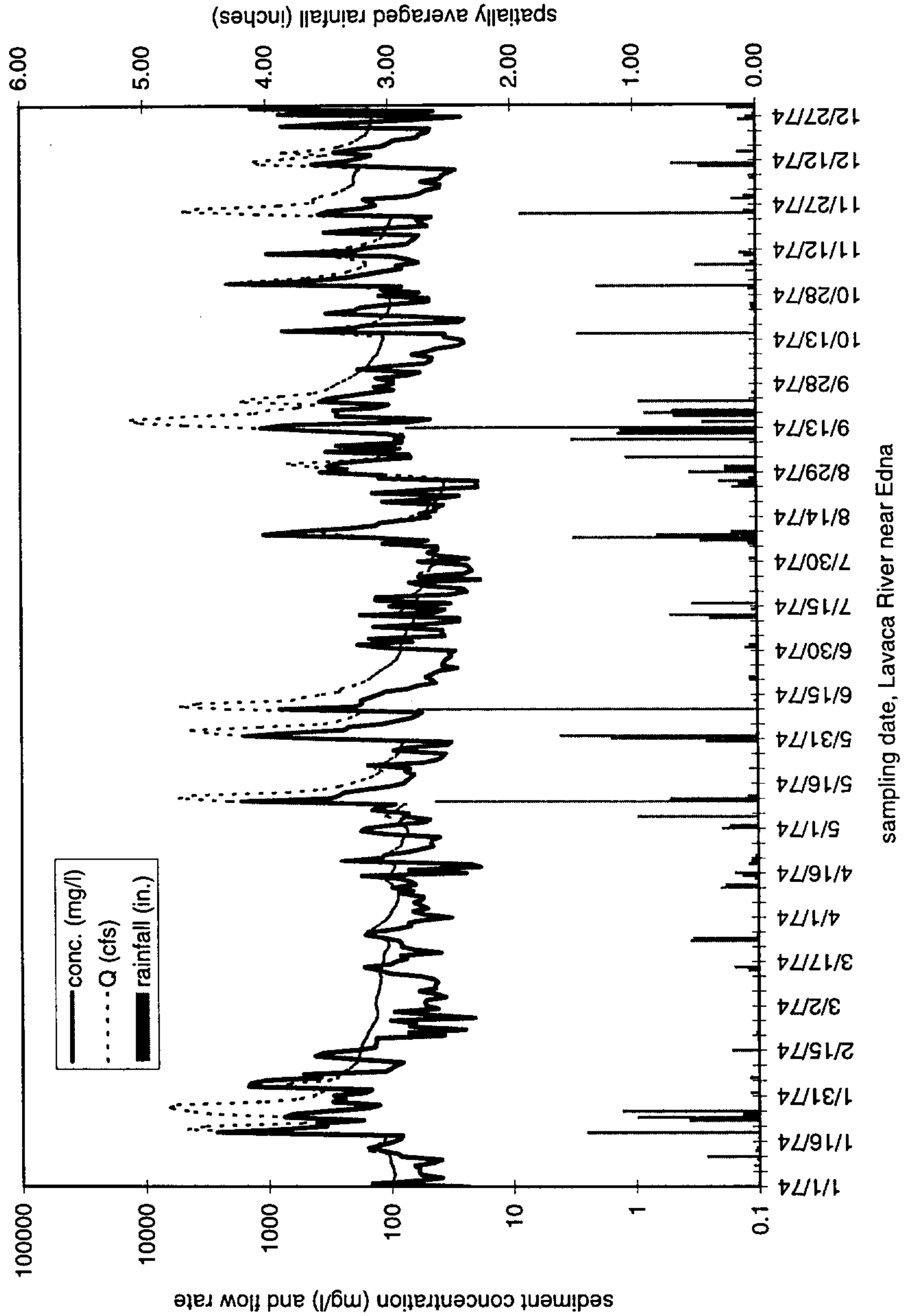

sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/70
1/16/70
1/31/70
2/15/70
3/2/70
3/17/70
4/1/70
4/16/70
5/1/70
5/16/70
5/31/70
6/15/70
6/30/70
7/15/70
7/30/70
8/14/70
8/29/70
9/13/70
9/28/70
10/13/70
10/28/70
11/12/70
11/27/70
12/12/70
12/27/70
sampling date, Lavaca River near Edna



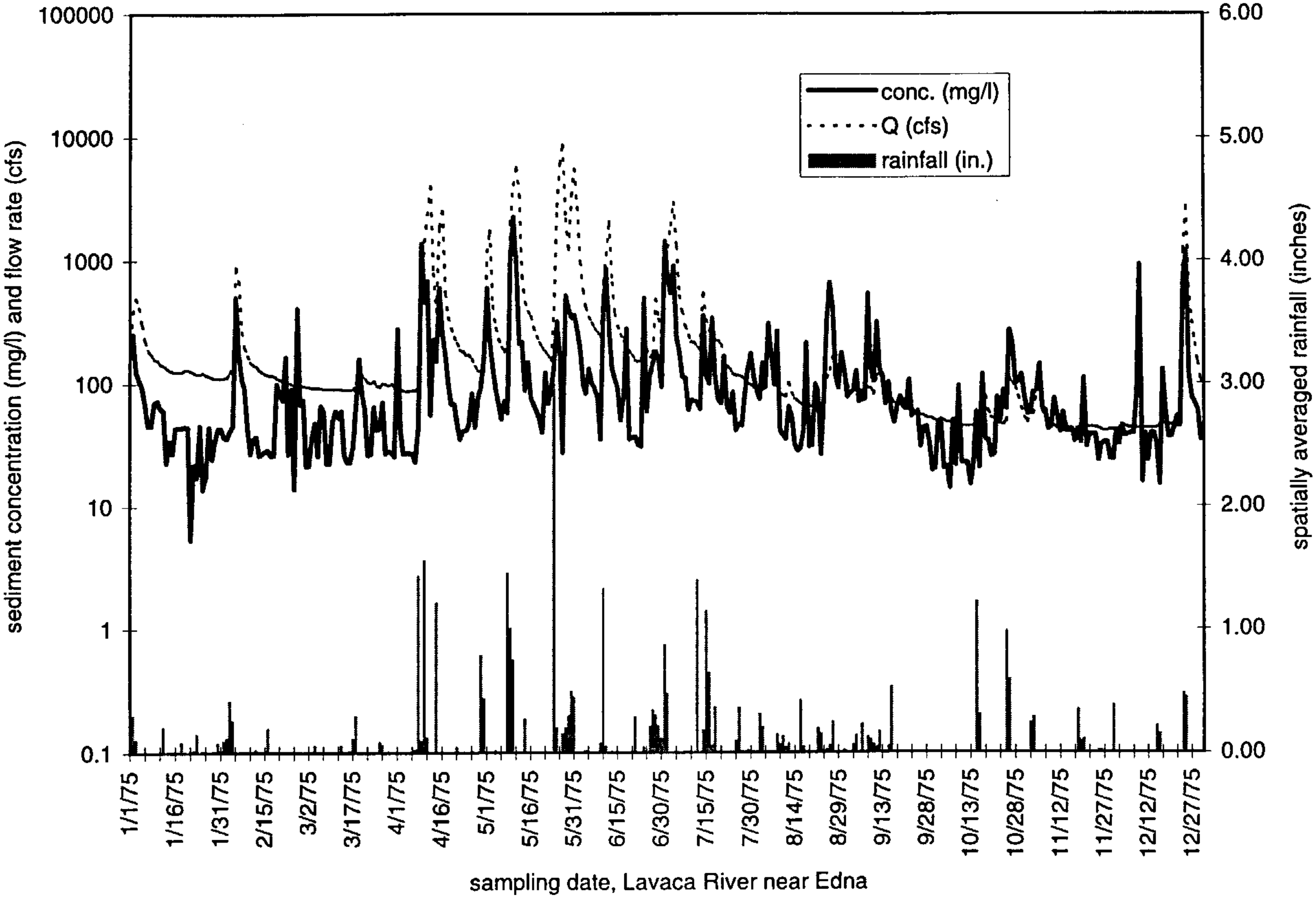
conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/71
1/16/71
1/31/71
2/15/71
3/2/71
3/17/71
4/1/71
4/16/71
5/1/71
5/16/71
5/31/71
6/15/71
6/30/71
7/15/71
7/30/71
8/14/71
8/29/71
9/13/71
9/28/71
10/13/71
10/28/71
11/12/71
11/27/71
12/12/71
12/27/71
sampling date, Lavaca River near Edna

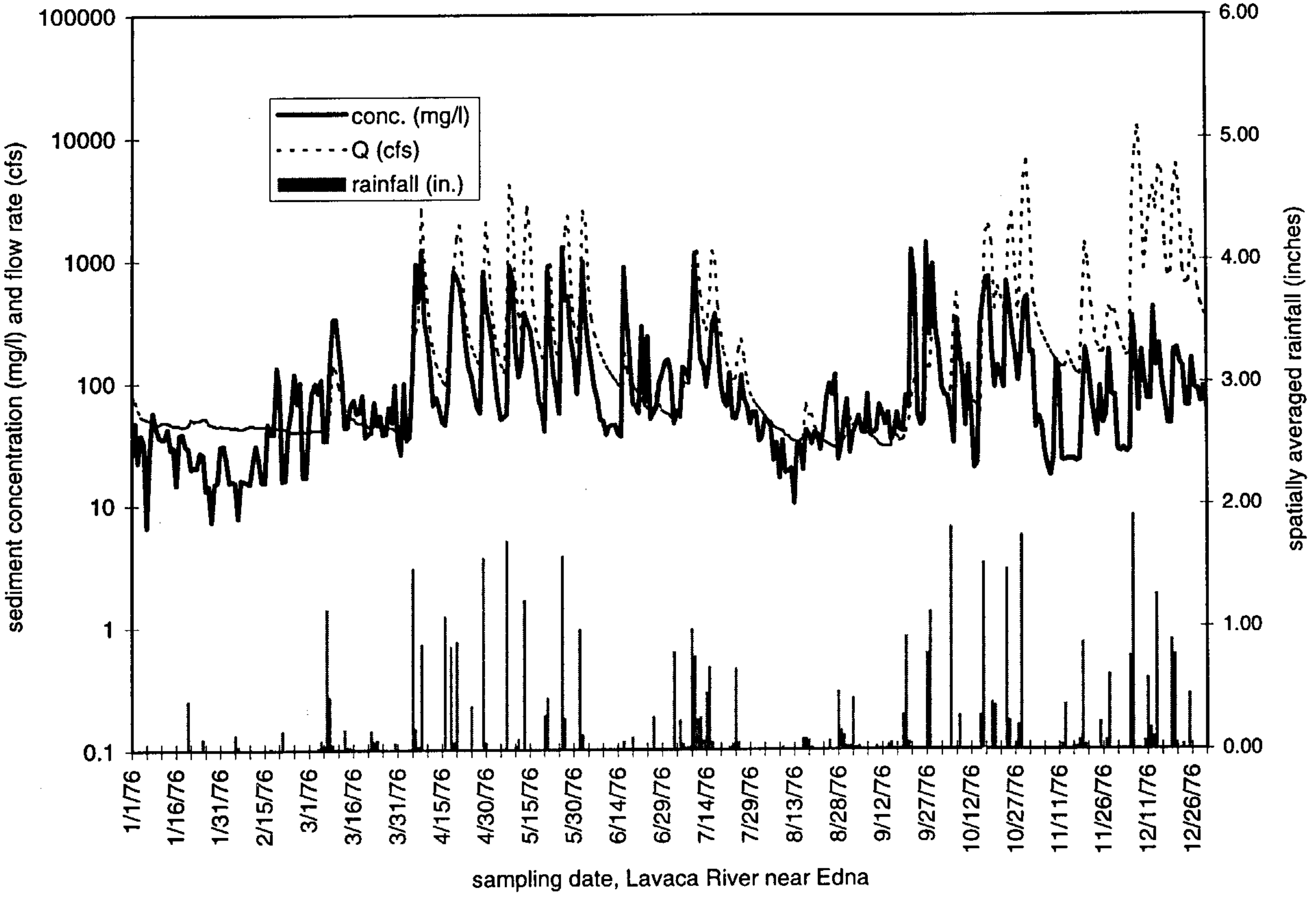

sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/72
1/16/72
1/31/72
2/15/72
3/1/72
3/16/72
3/31/72
4/15/72
4/30/72
5/15/72
5/30/72
6/14/72
6/29/72
7/14/72
7/29/72
8/13/72
8/28/72
9/12/72
9/27/72
10/12/72
10/27/72
11/11/72
11/26/72
12/11/72
12/26/72
sampling date, Lavaca River near Edna

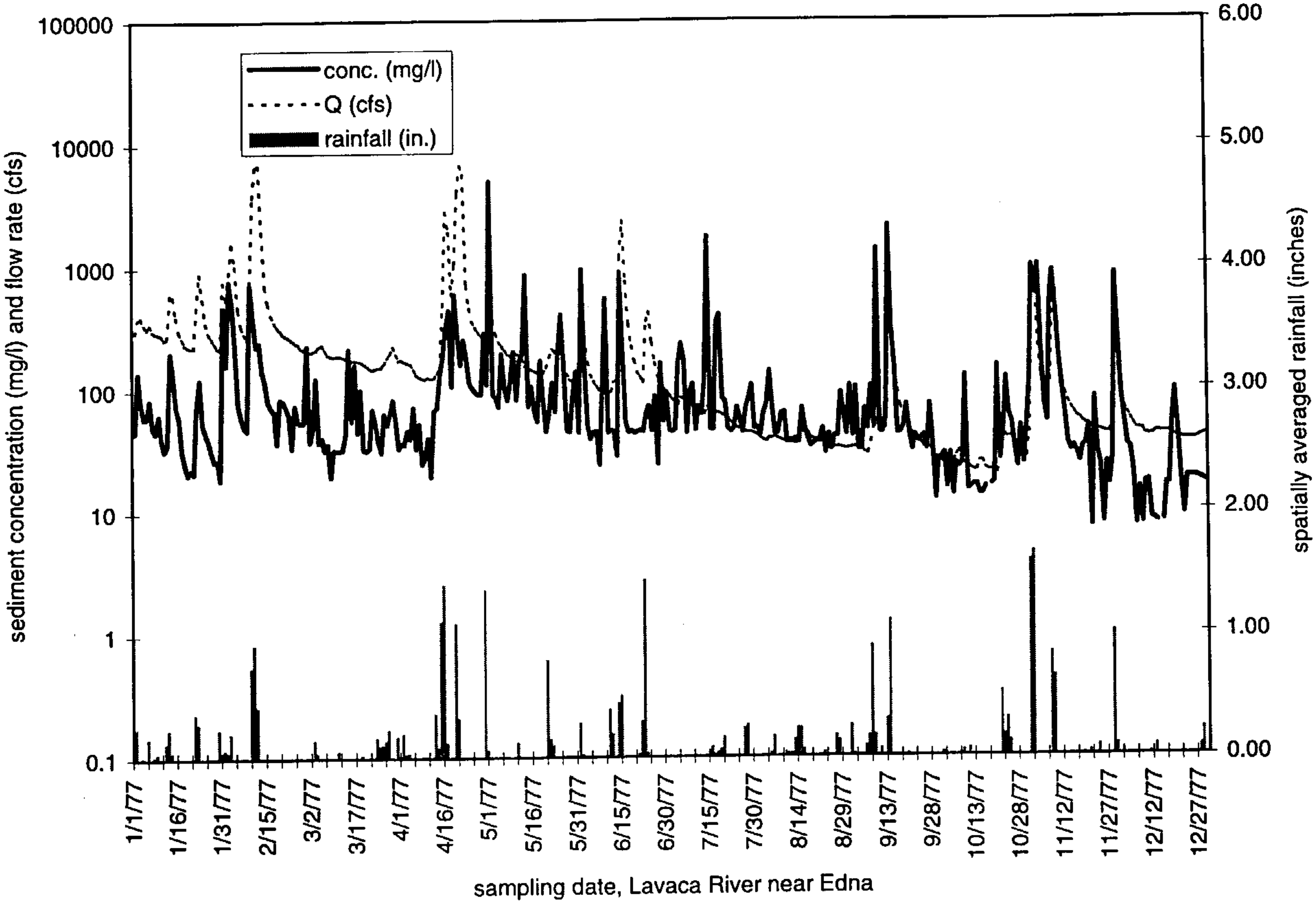

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
0.1
1
10
100
1000
10000
100000
spatially averaged rainfall (inches)
0.00
1.00
2.00
3.00
4.00
5.00
6.00
sampling date, Lavaca River near Edna
1/1/73
1/16/73
1/31/73
2/15/73
3/2/73
3/17/73
4/1/73
4/16/73
5/1/73
5/16/73
5/31/73
6/15/73
6/30/73
7/15/73
7/30/73
8/14/73
8/29/73
9/13/73
9/28/73
10/13/73
10/28/73
11/12/73
11/27/73
12/12/73
12/27/73



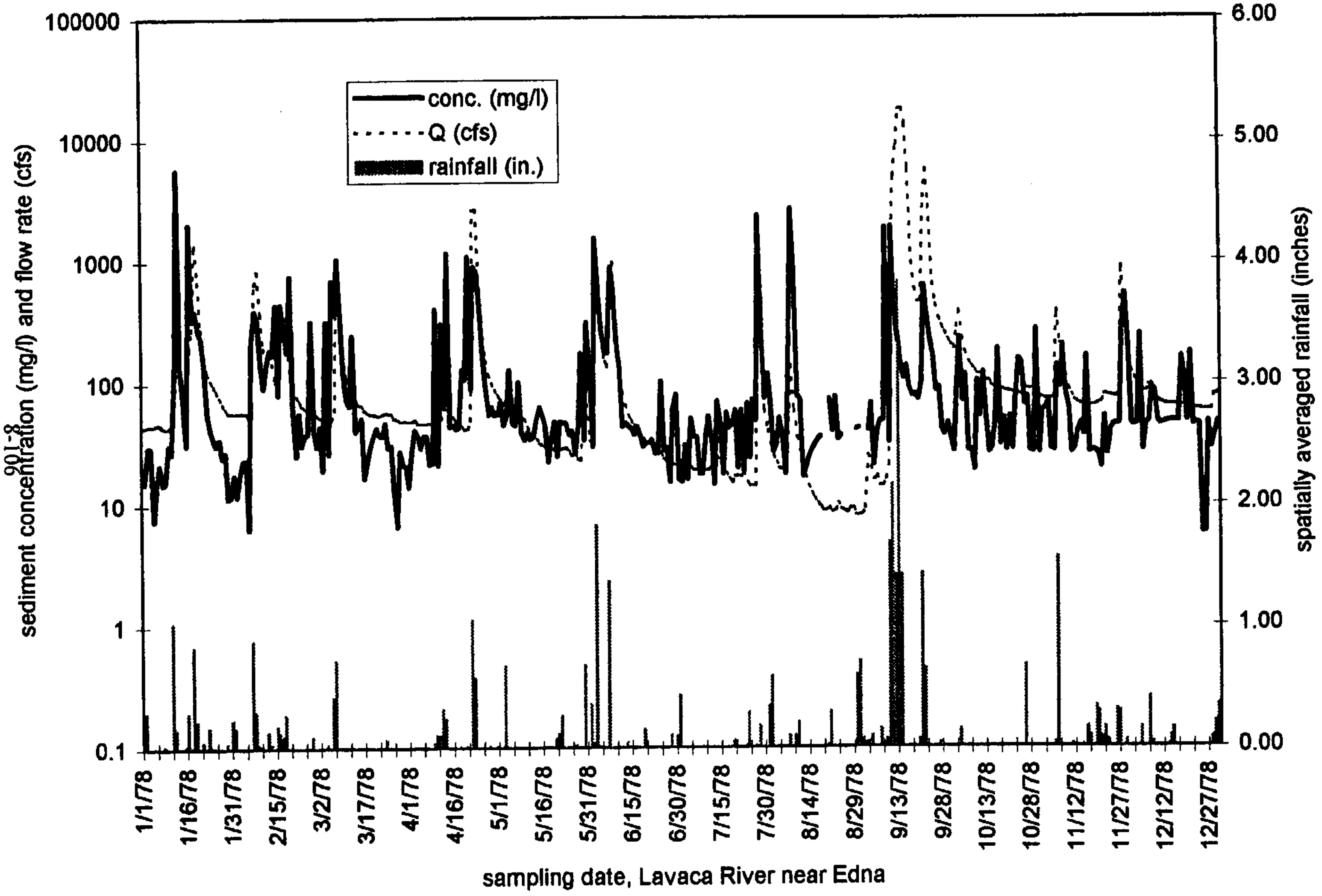
sediment concentration (mg/l) and flow rate
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
sampling date, Lavaca River near Edna
1/1/74
1/16/74
1/31/74
2/15/74
3/2/74
3/17/74
4/1/74
4/16/74
5/1/74
5/16/74
5/31/74
6/15/74
6/30/74
7/15/74
7/30/74
8/14/74
8/29/74
9/13/74
9/28/74
10/13/74
10/28/74
11/12/74
11/27/74
12/12/74
12/27/74
conc. (mg/l)
Q (cfs)
rainfall (in.)

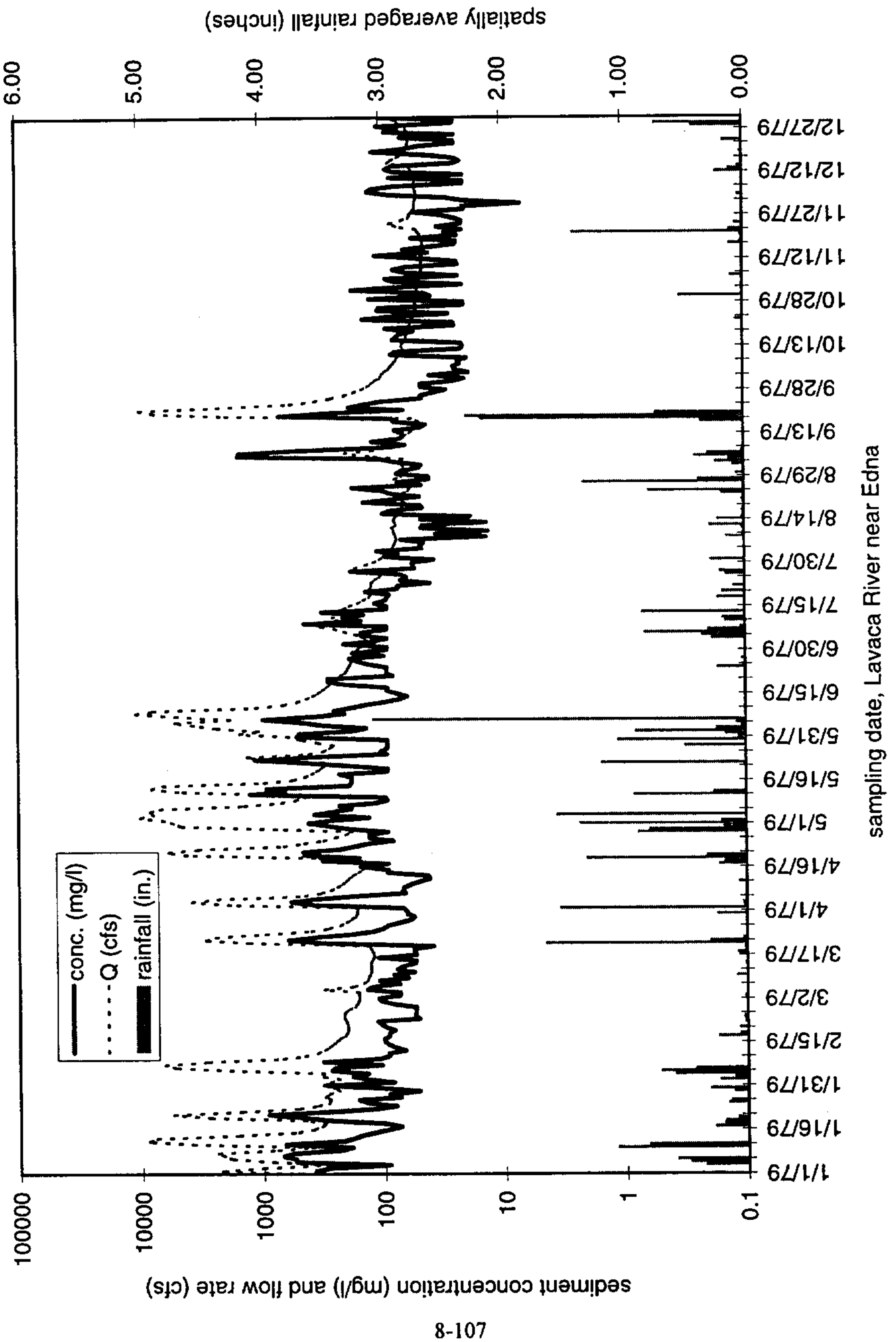

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/75
1/16/75
1/31/75
2/15/75
3/2/75
3/17/75
4/1/75
4/16/75
5/1/75
5/16/75
5/31/75
6/15/75
6/30/75
7/15/75
7/30/75
8/14/75
8/29/75
9/13/75
9/28/75
10/13/75
10/28/75
11/12/75
11/27/75
12/12/75
12/27/75
sampling date, Lavaca River near Edna


conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/76
1/16/76
1/31/76
2/15/76
3/1/76
3/16/76
3/31/76
4/15/76
4/30/76
5/15/76
5/30/76
6/14/76
6/29/76
7/14/76
7/29/76
8/13/76
8/28/76
9/12/76
9/27/76
10/12/76
10/27/76
11/11/76
11/26/76
12/11/76
12/26/76
sampling date, Lavaca River near Edna

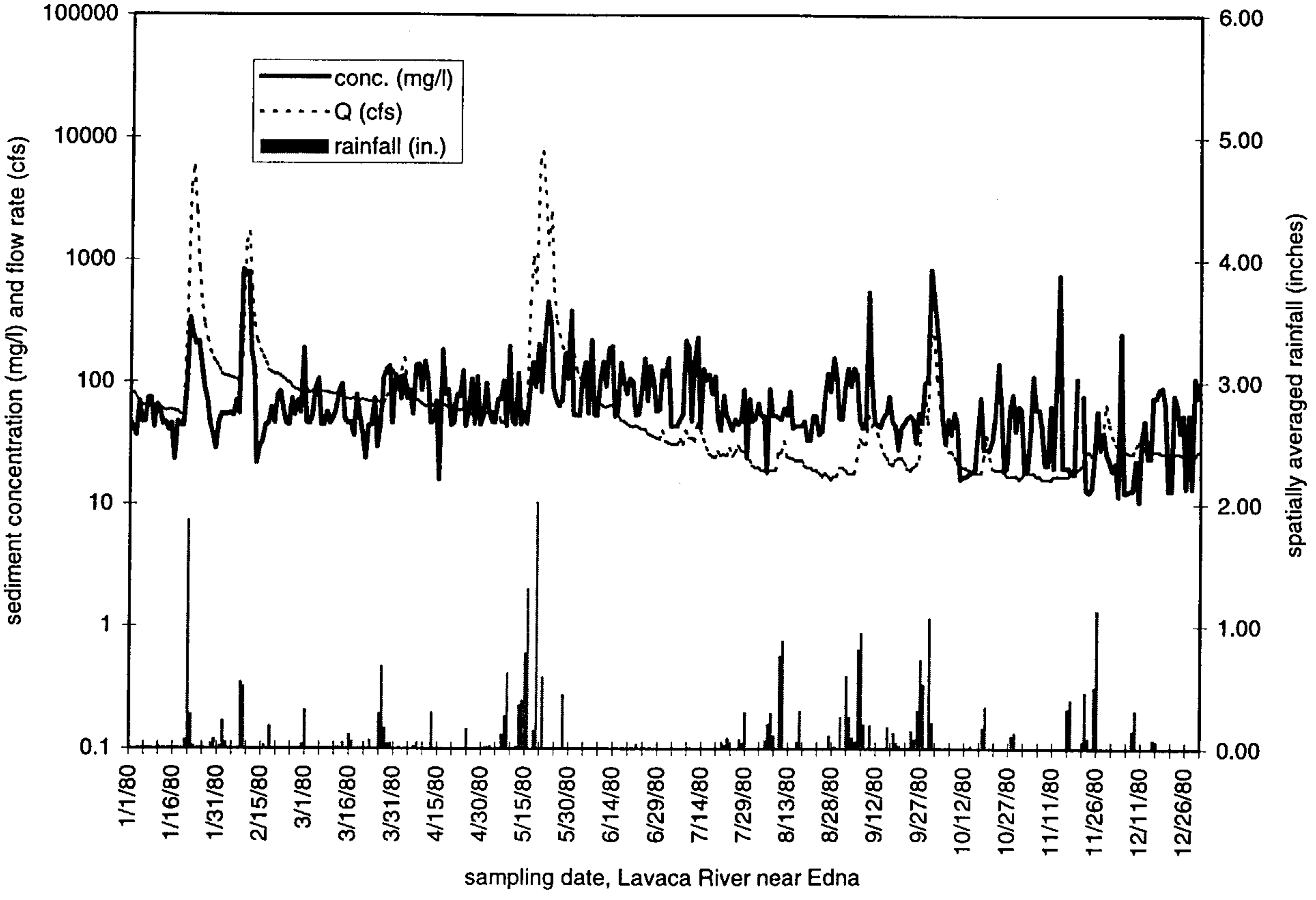


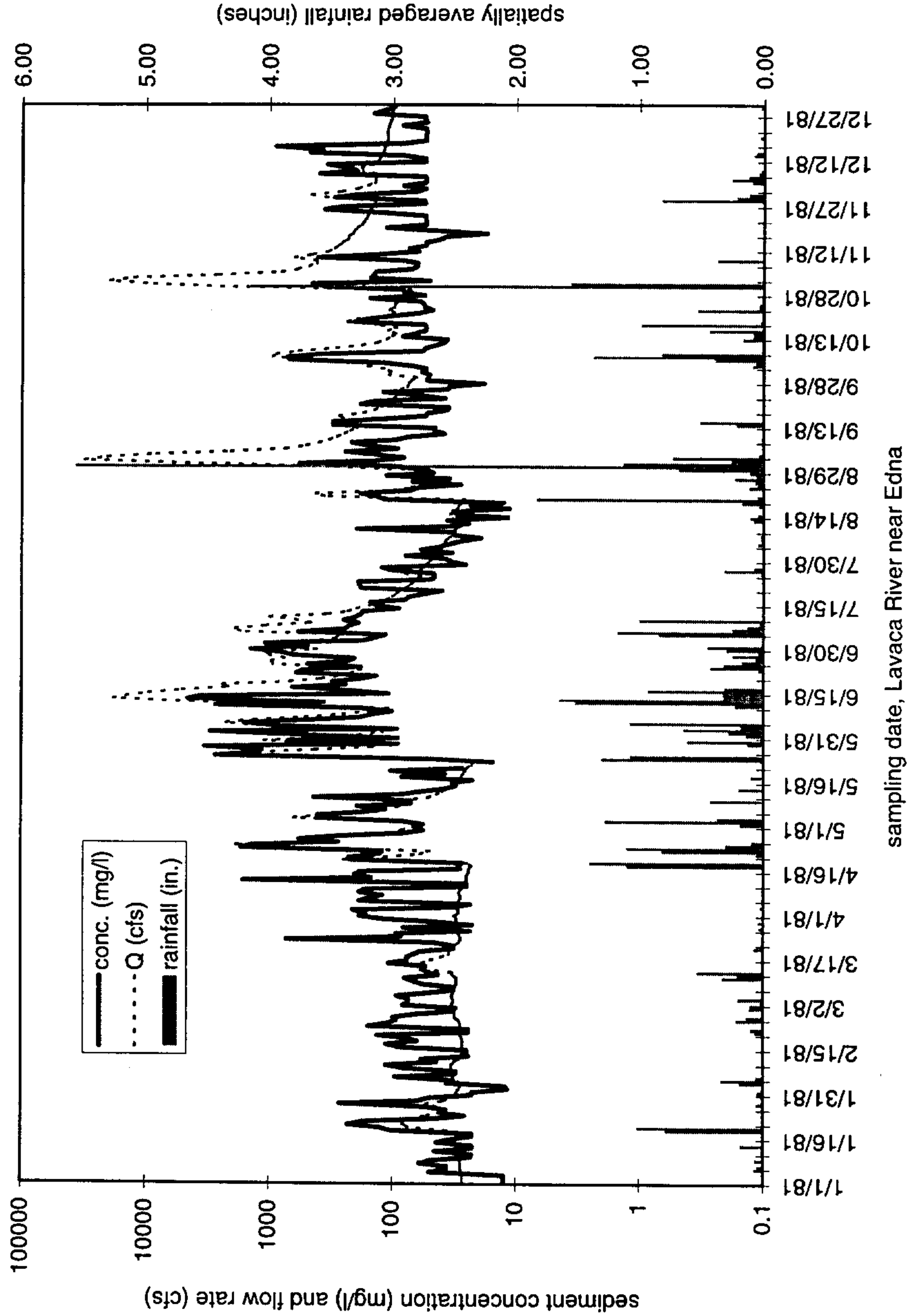
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/77
1/16/77
1/31/77
2/15/77
3/2/77
3/17/77
4/1/77
4/16/77
5/1/77
5/16/77
5/31/77
6/15/77
6/30/77
7/15/77
7/30/77
8/14/77
8/29/77
9/13/77
9/28/77
10/13/77
10/28/77
11/12/77
11/27/77
12/12/77
12/27/77
sampling date, Lavaca River near Edna

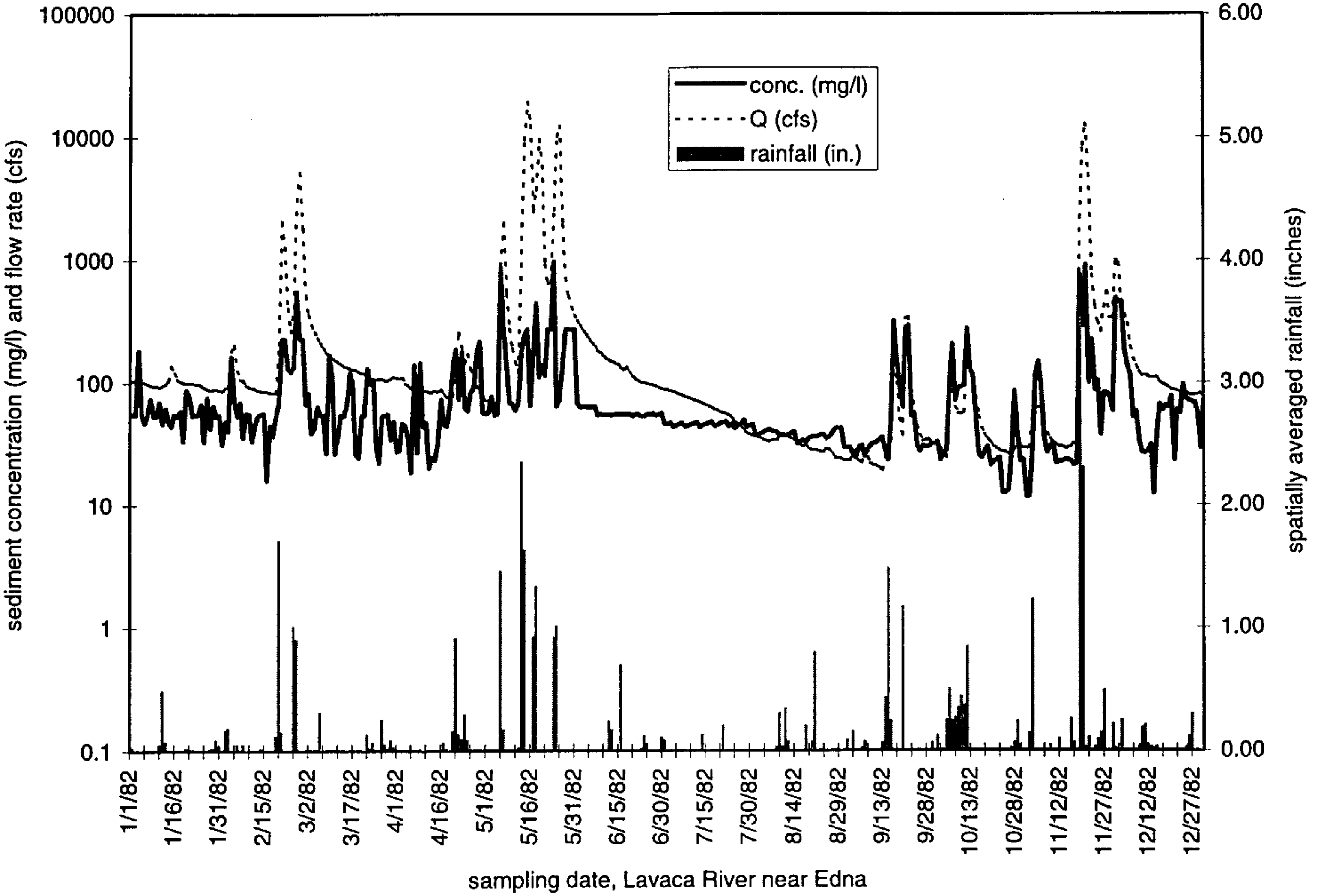

conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/78
1/16/78
1/31/78
2/15/78
3/2/78
3/17/78
4/1/78
4/16/78
5/1/78
5/16/78
5/31/78
6/15/78
6/30/78
7/15/78
7/30/78
8/14/78
8/29/78
9/13/78
9/28/78
10/13/78
10/28/78
11/12/78
11/27/78
12/12/78
12/27/78
sampling date, Lavaca River near Edna

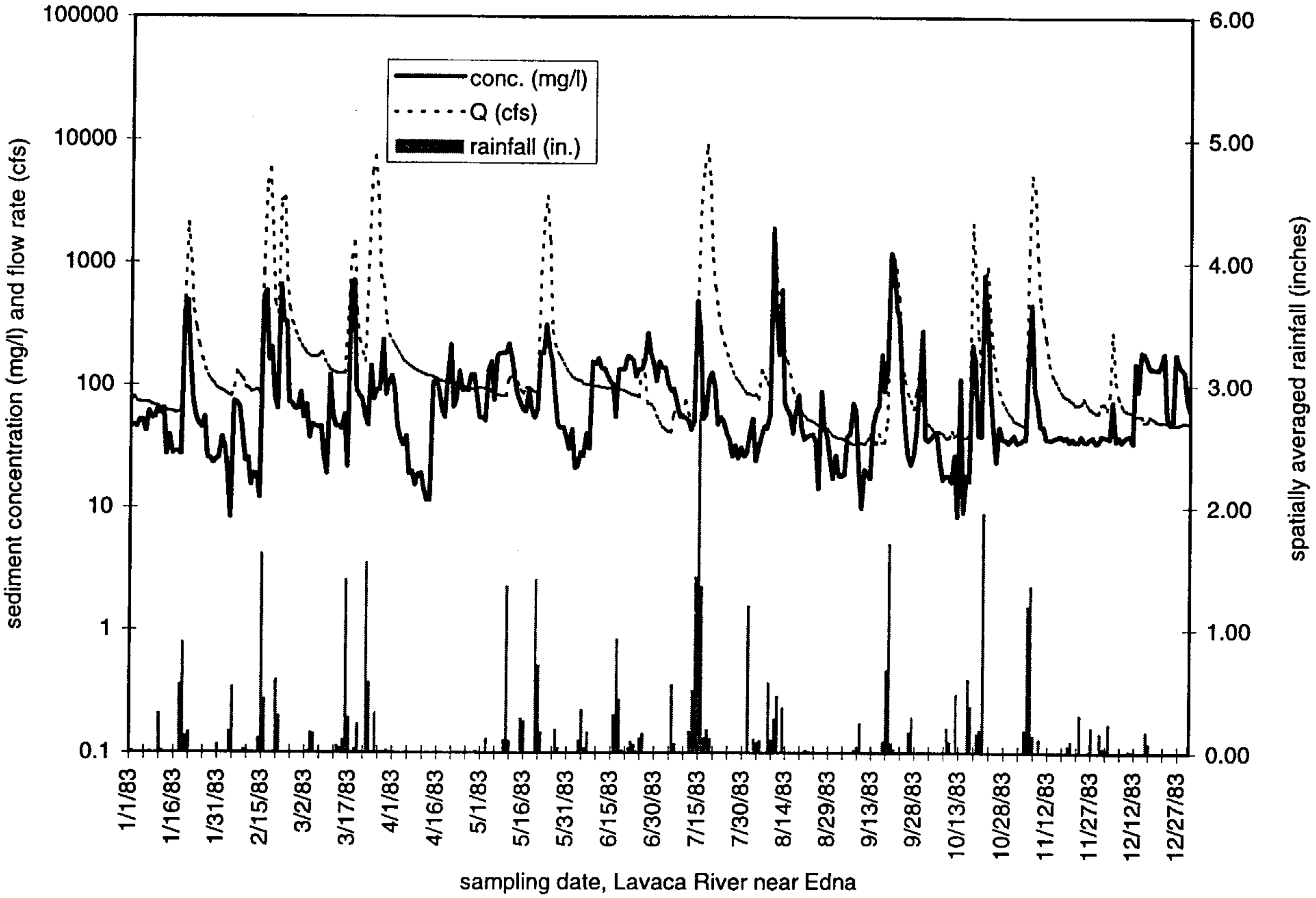


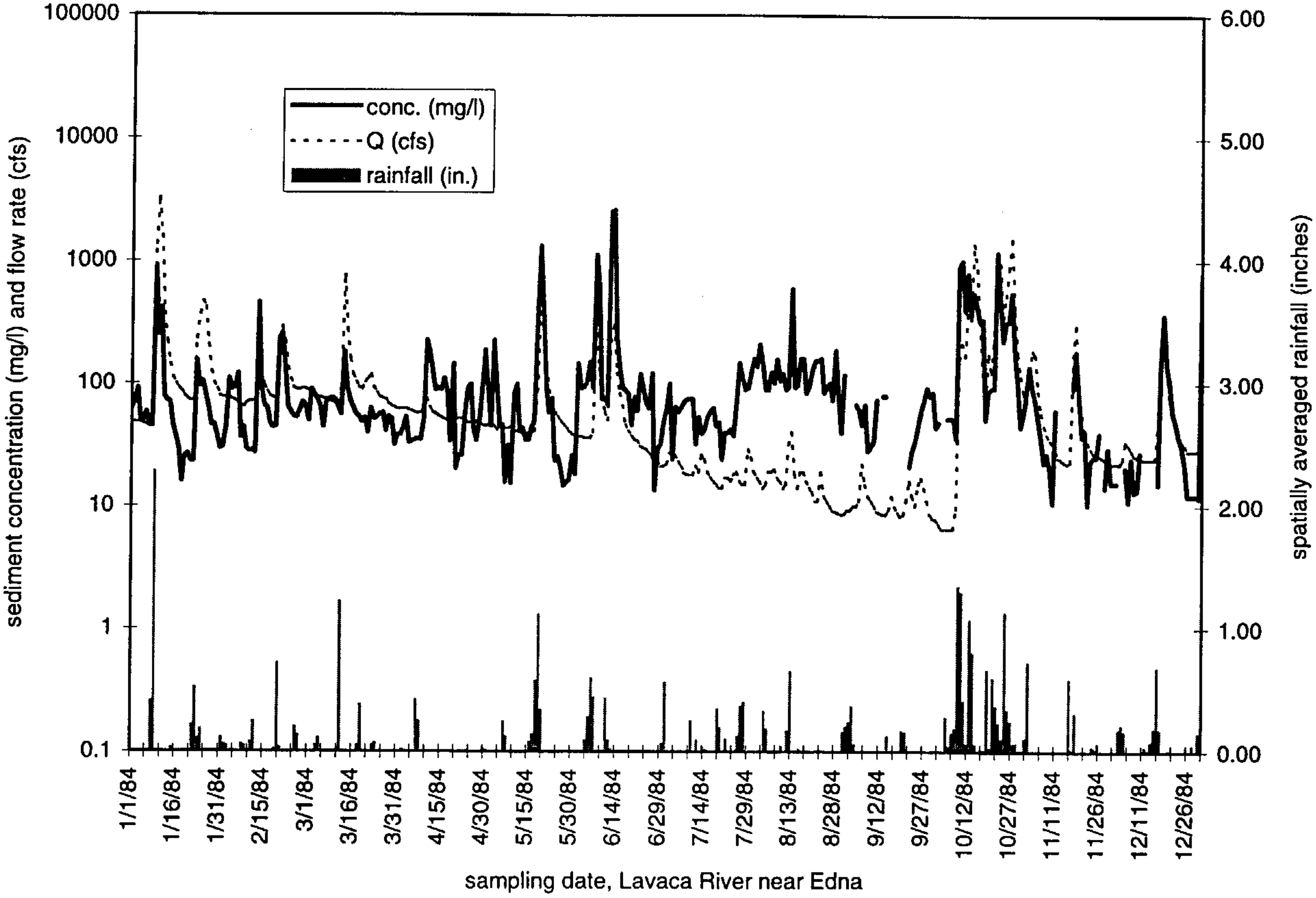
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
conc. (mg/l)
Q (cfs)
rainfall (in.)
100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
1/1/79
1/16/79
1/31/79
2/15/79
3/2/79
3/17/79
4/1/79
4/16/79
5/1/79
5/16/79
5/31/79
6/15/79
6/30/79
7/15/79
7/30/79
8/14/79
8/29/79
9/13/79
9/28/79
10/13/79
10/28/79
11/12/79
11/27/79
12/12/79
12/27/79
sampling date, Lavaca River near Edna


conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/80
1/16/80
1/31/80
2/15/80
3/1/80
3/16/80
3/31/80
4/15/80
4/30/80
5/15/80
5/30/80
6/14/80
6/29/80
7/14/80
7/29/80
8/13/80
8/28/80
9/12/80
9/27/80
10/12/80
10/27/80
11/11/80
11/26/80
12/11/80
12/26/80
sampling date, Lavaca River near Edna

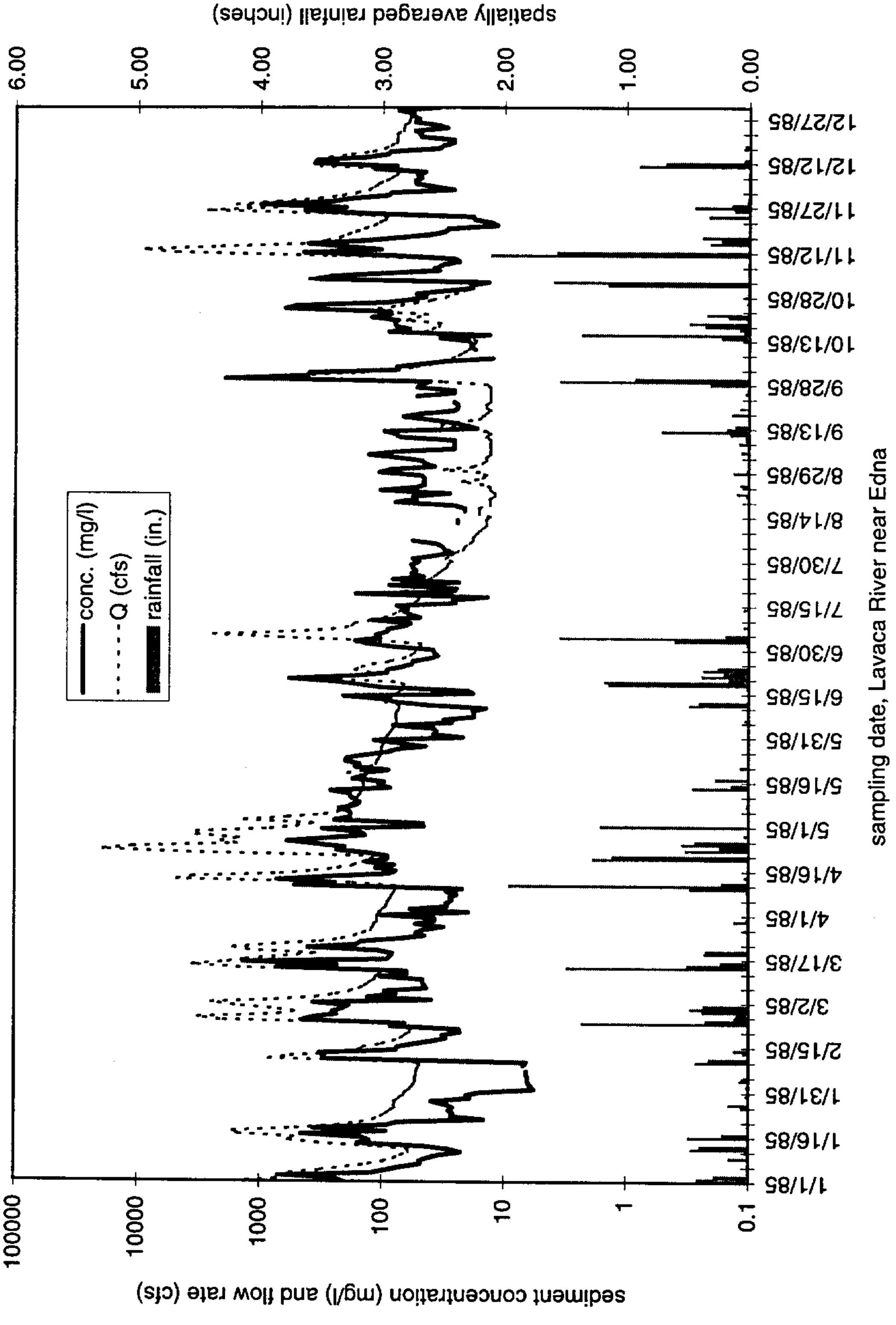

sediment concentration (mg/l) and flow rate (cfs)
0.1
1
10
100
1000
10000
100000
spatially averaged rainfall (inches)
0.00
1.00
2.00
3.00
4.00
5.00
6.00
sampling date, Lavaca River near Edna
1/1/81
1/16/81
1/31/81
2/15/81
3/2/81
3/17/81
4/1/81
4/16/81
5/1/81
5/16/81
5/31/81
6/15/81
6/30/81
7/15/81
7/30/81
8/14/81
8/29/81
9/13/81
9/28/81
10/13/81
10/28/81
11/12/81
11/27/81
12/12/81
12/27/81
conc. (mg/l)
Q (cfs)
rainfall (in.)

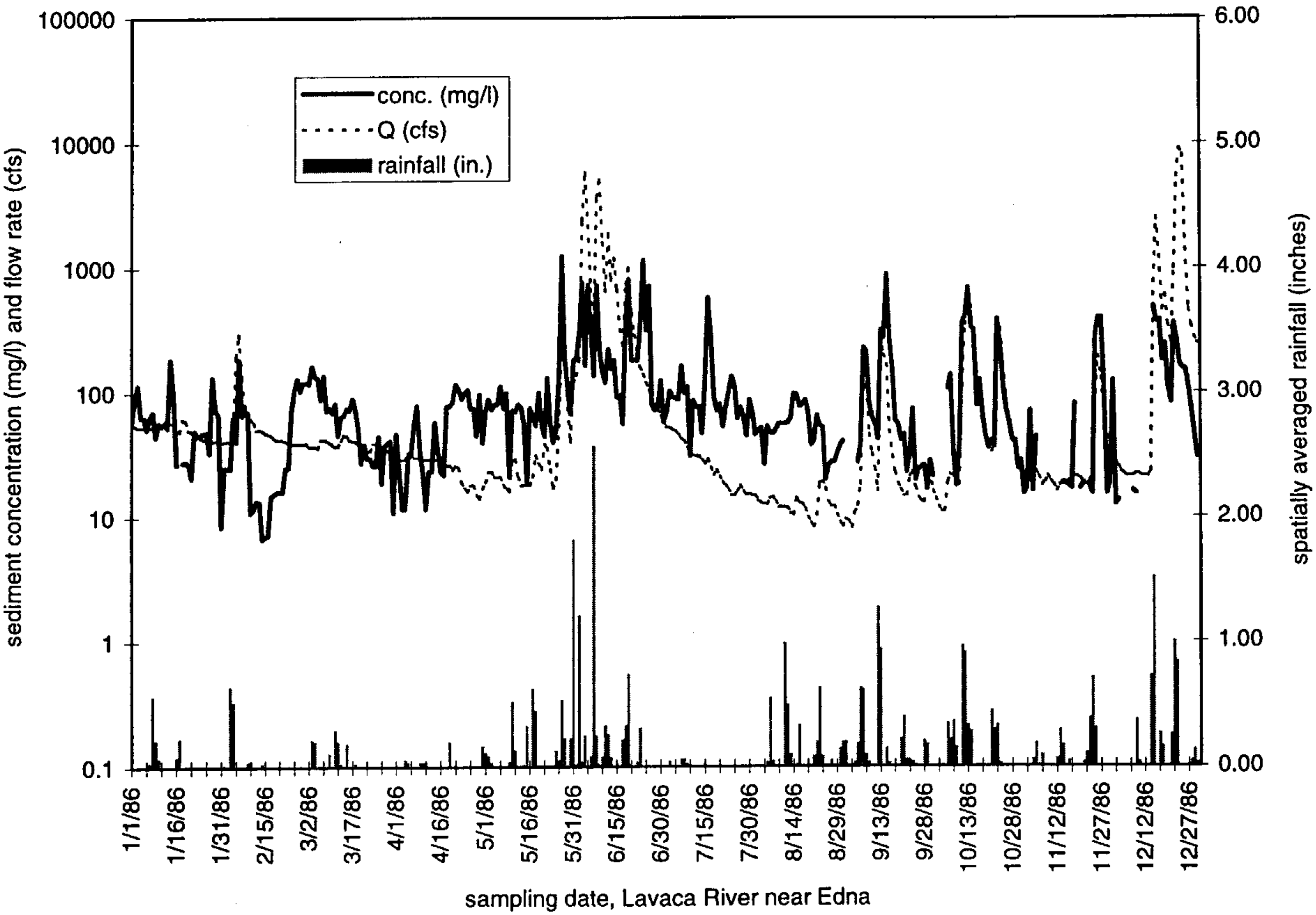

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/82
1/16/82
1/31/82
2/15/82
3/2/82
3/17/82
4/1/82
4/16/82
5/1/82
5/16/82
5/31/82
6/15/82
6/30/82
7/15/82
7/30/82
8/14/82
8/29/82
9/13/82
9/28/82
10/13/82
10/28/82
11/12/82
11/27/82
12/12/82
12/27/82
sampling date, Lavaca River near Edna

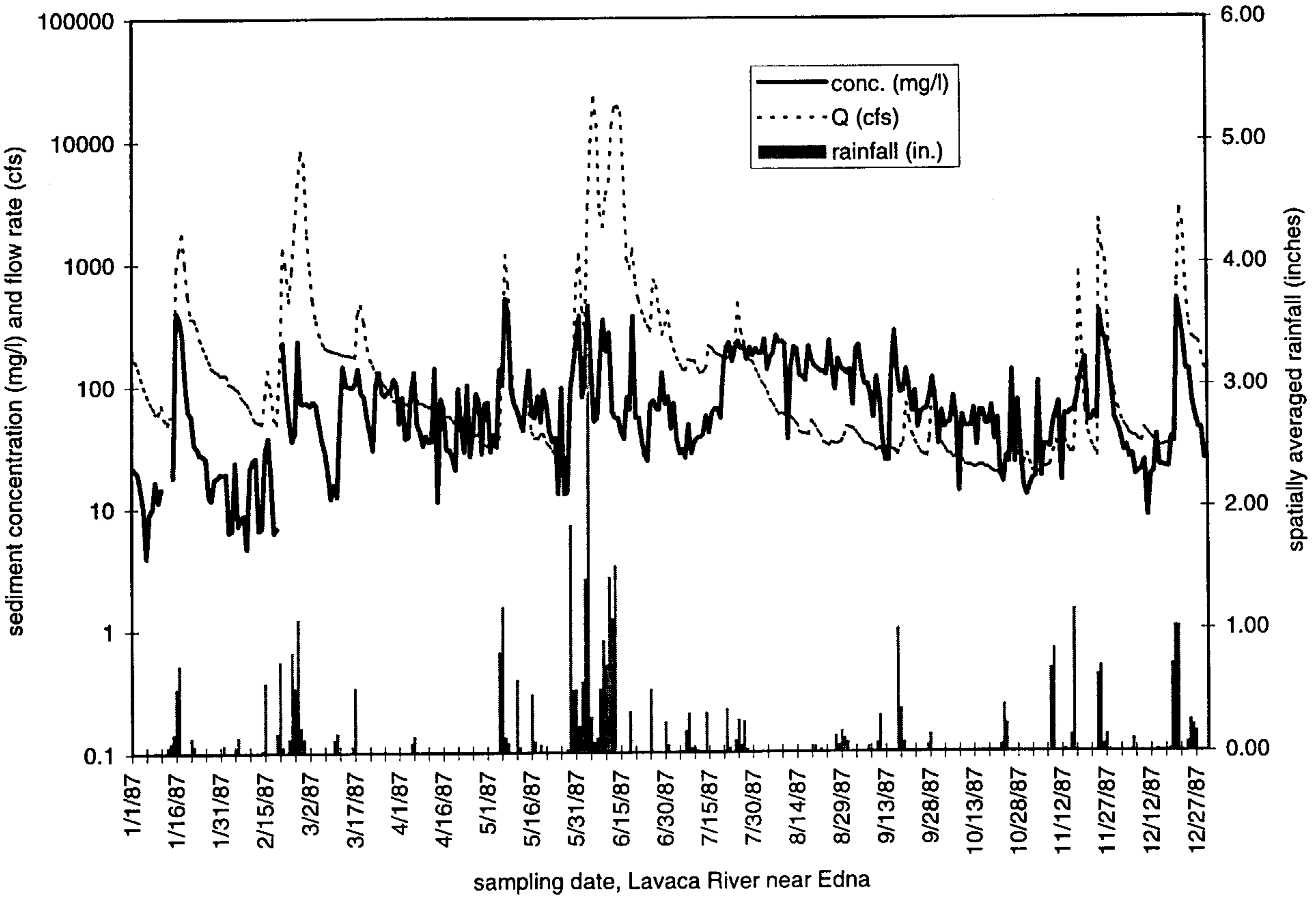

100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/83
1/16/83
1/31/83
2/15/83
3/2/83
3/17/83
4/1/83
4/16/83
5/1/83
5/16/83
5/31/83
6/15/83
6/30/83
7/15/83
7/30/83
8/14/83
8/29/83
9/13/83
9/28/83
10/13/83
10/28/83
11/12/83
11/27/83
12/12/83
12/27/83
sampling date, Lavaca River near Edna

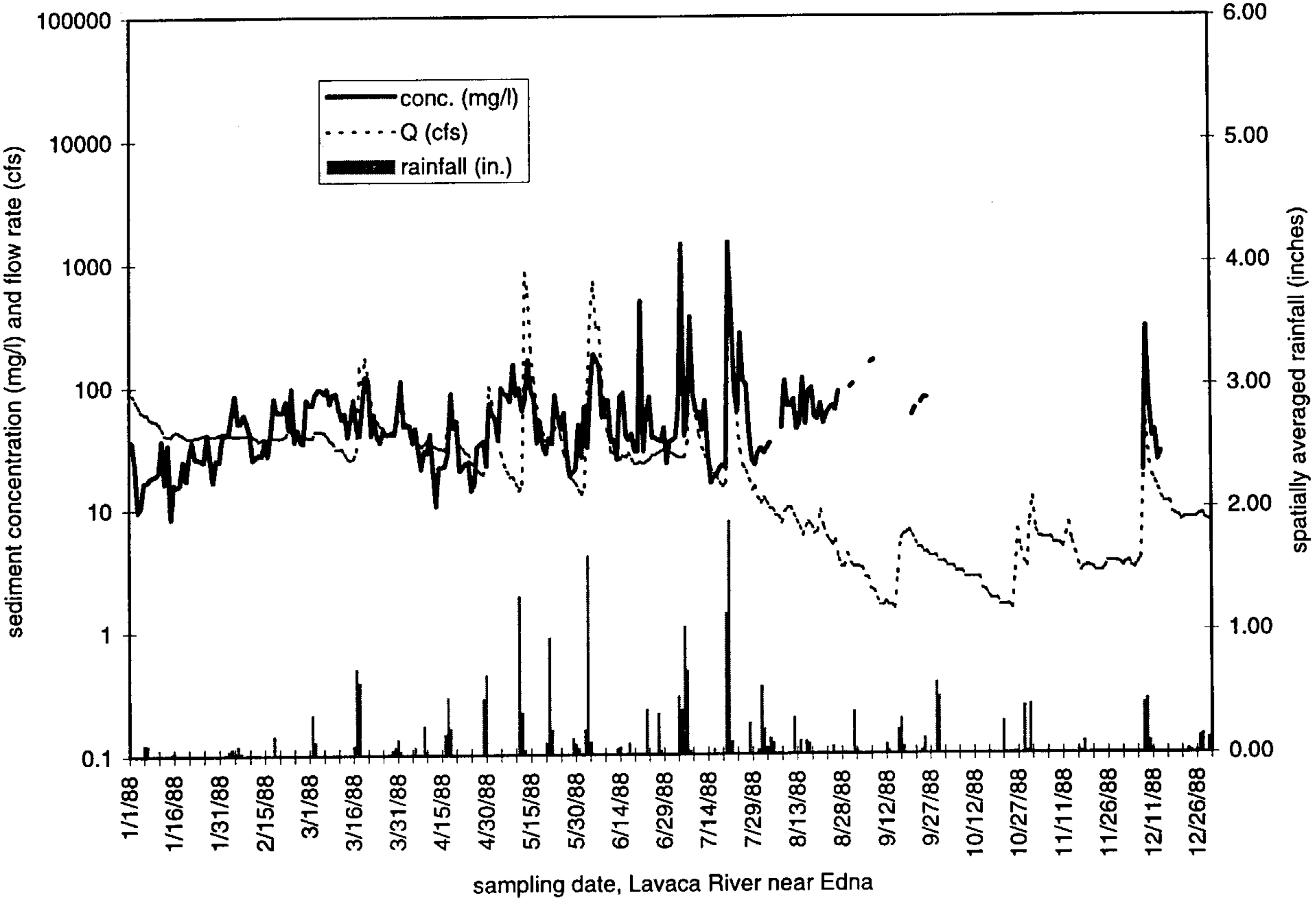


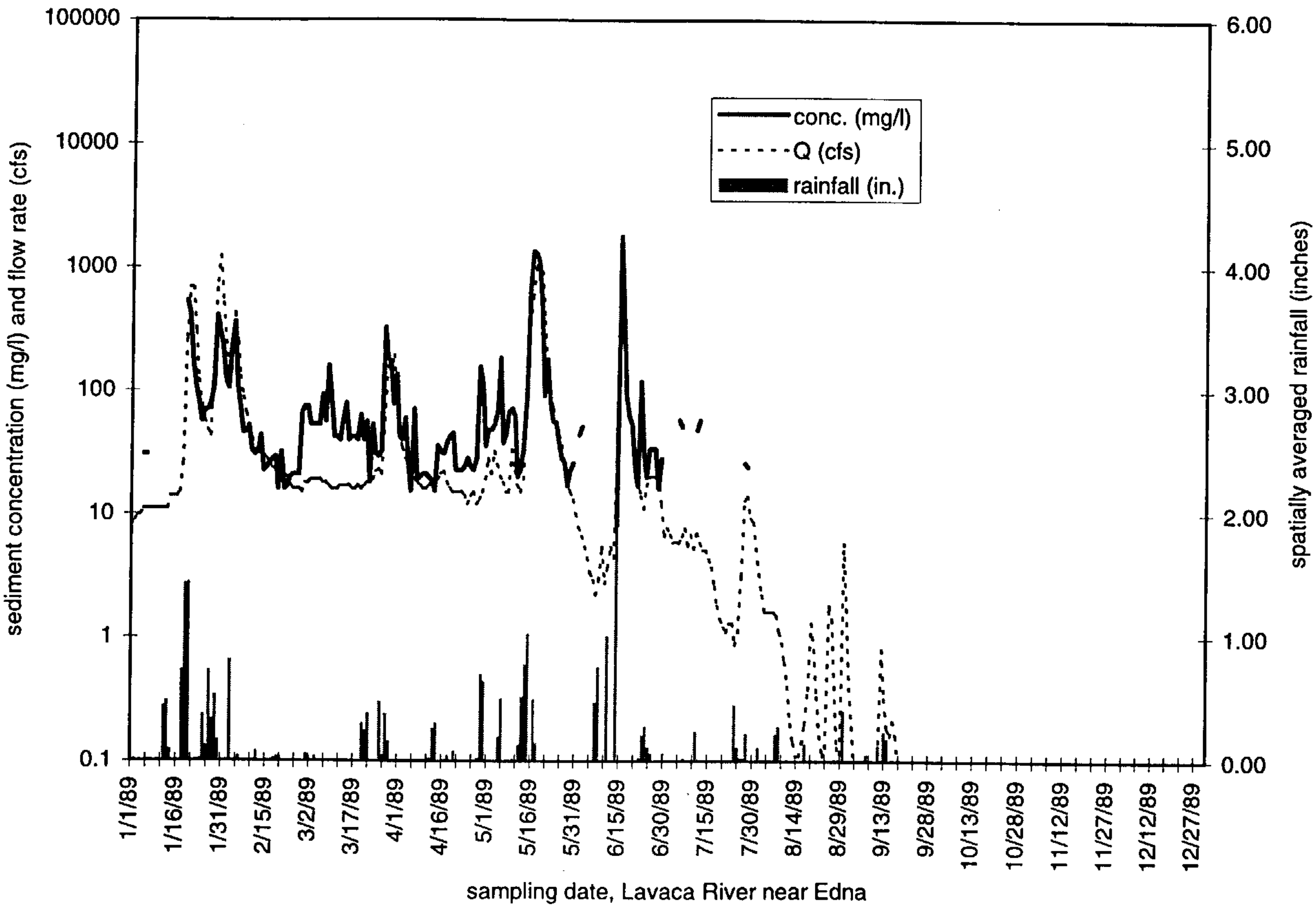
conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/84
1/16/84
1/31/84
2/15/84
3/1/84
3/16/84
3/31/84
4/15/84
4/30/84
5/15/84
5/30/84
6/14/84
6/29/84
7/14/84
7/29/84
8/13/84
8/28/84
9/12/84
9/27/84
10/12/84
10/27/84
11/11/84
11/26/84
12/11/84
12/26/84
sampling date, Lavaca River near Edna


sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/85
1/16/85
1/31/85
2/15/85
3/2/85
3/17/85
4/1/85
4/16/85
5/1/85
5/16/85
5/31/85
6/15/85
6/30/85
7/15/85
7/30/85
8/14/85
8/29/85
9/13/85
9/28/85
10/13/85
10/28/85
11/12/85
11/27/85
12/12/85
12/27/85
sampling date, Lavaca River near Edna


100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/86
1/16/86
1/31/86
2/15/86
3/2/86
3/17/86
4/1/86
4/16/86
5/1/86
5/16/86
5/31/86
6/15/86
6/30/86
7/15/86
7/30/86
8/14/86
8/29/86
9/13/86
9/28/86
10/13/86
10/28/86
11/12/86
11/27/86
12/12/86
12/27/86
sampling date, Lavaca River near Edna


100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/87
1/16/87
1/31/87
2/15/87
3/2/87
3/17/87
4/1/87
4/16/87
5/1/87
5/16/87
5/31/87
6/15/87
6/30/87
7/15/87
7/30/87
8/14/87
8/29/87
9/13/87
9/28/87
10/13/87
10/28/87
11/12/87
11/27/87
12/12/87
12/27/87
sampling date, Lavaca River near Edna


conc. (mg/l)
Q (cfs)
rainfall (in.)
sediment concentration (mg/l) and flow rate (cfs)
100000
10000
1000
100
10
1
0.1
spatially averaged rainfall (inches)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1/1/88
1/16/88
1/31/88
2/15/88
3/1/88
3/16/88
3/31/88
4/15/88
4/30/88
5/15/88
5/30/88
6/14/88
6/29/88
7/14/88
7/29/88
8/13/88
8/28/88
9/12/88
9/27/88
10/12/88
10/27/88
11/11/88
11/26/88
12/11/88
12/26/88
sampling date, Lavaca River near Edna


100000
10000
1000
100
10
1
0.1
sediment concentration (mg/l) and flow rate (cfs)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
spatially averaged rainfall (inches)
conc. (mg/l)
Q (cfs)
rainfall (in.)
1/1/89
1/16/89
1/31/89
2/15/89
3/2/89
3/17/89
4/1/89
4/16/89
5/1/89
5/16/89
5/31/89
6/15/89
6/30/89
7/15/89
7/30/89
8/14/89
8/29/89
9/13/89
9/28/89
10/13/89
10/28/89
11/12/89
11/27/89
12/12/89
12/27/89
sampling date, Lavaca River near Edna

9 REFERENCES

Ackers, P., and White, W.R., 1973. *Sediment transport: New approach and analysis.* Journal of Hydraulics, Vol. 99, no. HY11, pp 2041-60.

Adey, E.A. and Cook, H.M., 1964. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1959- September 1961, Texas Water Commission, Bulletin 6410.

Anderson, James R. 1996. *Nonpoint Source Pollution by Organocholrine Pesticides in the Lavaca-Navidad Watershed, Texas.* Thesis. The University of Texas at Austin.

Arnold, Jeff , 1995. Agricultural Research Service in Temple, Texas. Interview.

Arnold, J.G., Srinivasan, R., and Muttiah, R.S., 1994. *Large-Scale Hydrologic Modeling and Assessment*, Effects of Human-Induced Changes On Hydrologic Systems, Proceedings Annual Summer Symposium of the American Water Resources Association.

Bagnold, R.A., 1966. *An Approach to the Sediment Transport Problem from General Physics.* U.S. Geological Survey Professional Paper, 422-I.

Bagnold, R.A., 1980. *An empirical correlation of bedload transport rates in flumes and natural rivers.* Proceedings of the Royal Society of London, Vol. 372, pp.453-473.

Bouraoui, Faycal and Dillaha, Theo, 1996. *ANSWERS-2000: runoff and sediment transport model.* Journal of Environmental Engineering, Volume 122, no. 6, American Society of Civil Engineers, New York, New York.

Cody, Ronald P. and Smith, Jeffrey K., 1991. *Applied Statistics and the SAS Programming Language*, Third Edition, Prentice Hall.

Colby, B.R. and Scott, C.H. 1965. *Effects of Water Temperature on the Discharge of Bed Material.* U.S. Geological Survey Professional Paper, 462-G.

Cook, H.M., 1967. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1961- September 1963, Texas Water Development Board, Report 45.

Cook, H.M., 1970. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1963- September 1965, Texas Water Development Board, Report 106.

Daly, C., Neilson, R.P., and Phillips, D.L., 1994. *A statistical-topographic model for mapping climatological precipitation over mountainous terrain.* Journal of Applied Meteorology 33, 140-158.

Deetz, H.C., and Adams, O.S., 1969. *Elements of map projection with applications to map and chart construction*, Greenwood, New York, 226.

Dougherty, John P., 1979. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1971- September 1975, Texas Department of Water Resources, Report 233.

DuBoys P. 1879. *Le Rhone et les rivieres a lit affouillable.* Ann. Ponts et Chauss., 18, 141-195.

Einstein, H.A. 1950. *The Bed-load Function for Sediment Transportation in Open Channel Flows.* U.S. Department of Agriculture Technical Bulletin 1026.

Fan, Shou-Shan, 1988. *Twelve Selected Computer Stream Sedimentation Models Developed in the United States.* Interagency Ad Hoc Sedimentation Work Group, Published by Federal Energy Regulatory Commission.

Fan and Springer, 1990. *Major Sedimentation Issues and Ongoing Investigations at the Federal Energy Regulatory Commission.* Hydraulic Engineering, Volume 2, Proceedings of the 1990 National Conference. 1015-1020.

Farris, Orville A. 1933. *The Silt Load of Texas Streams.* United States Department of Agriculture Technical Bulletin No. 382.

Faye, R.E., Carey, W.P., Stamer, J.K., and Kleckner, R.L. 1980. *Erosion, Sediment Discharge, and Channel Morphology in the Upper Chattahoochee River Basin, Georgia.* Geological Survey Professional Paper 1107.

Graf, Walter Hans, 1971. *Hydraulics of Sediment Transport*, McGraw-Hill.

Gujarati, Damodar N., 1995. *Basic Econometrics*, McGraw-Hill, Inc.

Haan, Charles T., 1977. *Statistical Methods in Hydrology*, The Iowa State University Press.

Helsel, D.R. and Hirsch, R.M. *Statistical Methods in Water Resources.* Elsevier, 1992.

Holley, E.R. 1992. "Sediment transport in the lower Guadalupe and San Antonio Rivers," Center of Research in Water Resources, University of Texas at Austin, Tech. Memo. 92-1.

Julien, Pierre Y. *Erosion and Sedimentation*, Cambridge University Press, 1995.

Kothyari, U.C., Tiwari, A.K., Singh, Ranvir. 1996. *Temporal Variation of Sediment Yield.* Journal of Hydrologic Engineering, Volume 1, No.4, p169-176. American Society of Civil Engineers, New York, New York.

Kunze, H.L. 1969. *Reconnaissance of the Chemical Quality of Surface Waters of the Lavaca River Basin, Texas.* TWDB Report 92, prepared by United States Gelogical Survey in cooperation with the Texas Water Development Board.

Laflen, John, 1995. Purdue University. Telephone interview.

Laursen, E.M. 1958. *The Total Sediment Load of Streams.* Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Paper 1530.

Lanbein, W.B. and Schumm, S.A.,1958. *Yield of sediment in relation to mean annual precipitation.* Transactions of the American Geophysical Union, p. 1076-1084.

Leopold, L.B. and Maddock, T., 1953. *The Hydraulic Geometry of Stream Channels and Some Physiographic Implications.* U.S. Geological Survey, Prof. Paper, 252.

Longley, W.L., ed. 1994. *Freshwater inflows to Texas Bays and Estuaries: ecological relationships and methods for determination of needs.* Texas Water Development Board and Texas Parks and Wildlife Department, Austin, TX. pp.56-71.

Mirabal, James, 1974. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1965- September 1971, Texas Water Development Board, Report 184.

Nordin, C.F., 1963. *A preliminary study of sediment transport parameters, Rio Puerco near Bernardo, New Mexico.* U.S. Geological Survey Professional Paper, 462-C.

Meyer-Peter E. & Muller R. 1948. *Formulas of Bed-load Transport.* International Association for Hydraulic Structures Research, Report of Second Meeting, Stockholm, 39-64.

Pye, Kenneth, Editor, 1994. *Sediment Transport and Depositional Processes,* Blackwell Scientific Publications.

Quincy, Roger M., 1988. *Suspended-Sediment Load of Texas Streams*, Compilation Report, October 1975- September 1982, Texas Water Development Board, Report 306.

Raudkivi, A.J., 1990. *Loose Boundary Hydraulics*, Pargamon Press.

Reed, S., Maidment, D., and Patoux, J. "Spatial Water Balance of Texas" (http://www.ce.utexas.edu/prof/maidment/GISHydro/docs/wbtexas/wbtexas.htm) , Center for Research and Water Resources, The University of Texas at Austin, 2/23/97.

Reid, I. and Frostick, L.E., 1994. "Fluvial Sediment Transport and Deposition, " Chap. 4, *Sediment Transport and Depositional Processes*, (ed. Kenneth Pye), Blackwell Scientific Publications. pp. 91,113.

Selby, M.J., 1994. "Hillslope Sediment Transport, " Chap. 3, *Sediment Transport and Depositional Processes*, (ed. Kenneth Pye), Blackwell Scientific Publications. p. 67.

Soil Science Society of America, 1979. *Universal Soil Loss Equation: Past, Present, and Future.* SSSA Special Publication Number 8.

Shields, A. 1936. *Anwendun der Aehnlichkeils mechanik, und der Turbulenzforschung auf die Geschiebebewegung.* Mitteilungen der Preussischen Versuchsanstalt fur Wasserbau und Schiffbau, 26, 98-109.

Soil Conservation Service, U.S. Department of Agriculture, 1959. *Inventory and Use of Sedimentation Data in Texas*, Texas Board of Water Engineers, Bulletin 5912.

Srinivasan, R. and Arnold, J.G., 1994. *Integration of a Basin-Scale Water Quality Model with GIS*, Water Resources Bulletin, American Water Resources Association, Vol. 30, No.3.

Sullivan, Scott. 1994-1996. Texas Water Development Board. Interviews.

Stout, I.M., Bentz, L.C., and Ingram, H.W., 1961. *Silt Load of Texas Streams*, Compilation Report, June 1889- September 1959, Texas Board of Water Engineers, Bulletin 6108.

Texas Department of Water Resources, February, 1982. *Erosion and Sedimentation by Water in Texas*, Report 268, Average Annual Rates Estimate in 1979.

Texas Water Development Board, December, 1990. *Plate 1, Surface Water Development, Existing Reservoirs and Recommended Projects.*

USGS Texas District Open-File Report, September, 1973.

Vanoni, V.A., 1946. *Transportation of Suspended Sediment by Water.* American Society of Civil Engineers, Volume 111, Paper No. 2267, 67-133.

Vanoni, V.A., 1975. *Sedimentation Engineering.* ASCE Manual 25, p.1.

Wang, Sam S.Y., 1985. *Mathematical Models of Sediment Effects on Water Resources Systems*, Interim Report, Project Number G915-05, Water Resources Research Institute, Mississippi State University.

Welborn, Clarence T., 1967. *Comparative Results of Sediment Sampling with the Texas Sampler and the Depth-Integrating Samplers and Specific Weight of Fluvial Sediment Deposits in Texas.* Texas Water Development Board Report 37 in cooperation with the United States Geological Survey.

Wicks, J.M. and Bathurst, J.C., 1996. *SHESED: a physically based, distributed erosion and sediment yield component for the SHE hydrological modelling system.* Journal of Hydrology, Volume 175, No.1. Elsevier Science.

Williams, J.R., 1969. *Flood Routing with Variable Travel Time or Variable Storage Coefficients.* Transactions of the ASAE 12(1):100-103.

Williams, J.R., 1975. *Sediment-yield prediction with universal equation using runoff energy factor.* pp.244-252. Present and prospective technology for predicting sediment yield and sources: Proceedings of the sediment-yield workshop, USDA Sedimentation Lab, Oxford, MS, November, 1972.

Wischmeier, W.H. and Smith, D.D., 1958. *Rainfall Energy and its Relationship to Soil Loss.* Transactions of the American Geophysical Union, Volume 39, 285-291.

Wischmeier, W.H., Johnson, C.B., and Cross, B.V., 1971. *A Soil-Erodibility Nomograph for Farmland and Construction Sites*, Journal of Soil and Water Conservation, Volume 26, No. 5, 189-193.

Wischmeier, W.H., 1976. *Use and Misuse of the Universal Soil Loss Equation.* Journal of Soil Water Conservation, Volume 31, 5-9.

Wu, Hall, and Bonta, 1993. *Evaluation of Runoff and Erosion Models*, Journal of Irrigation and Drainage Engineering, Vol. 119, no. 4, March/April, 1993, ASCE, pp. 364-382.

Yang, C.T., 1973. *Incipient motion and sediment transport.* Journal of Hydraulics, Vol. 99, no HY10, ASCE, pp. 1679-1704.

Yang, Chih Ted, 1996. *Sediment Transport*, , McGraw-Hill.

ATTACHMENT 1

TEXAS WATER DEVELOPMENT BOARD

COMMENTS FOR UNIVERSITY OF TEXAS AT AUSTIN
RESEARCH PLANNING GRANT CONTRACT
CONTRACT NO. 96-483-148

The contractors diligently followed the scope of work contained in the contract and included every analysis that had been planned, and a few that had not originally been considered. The work was comprehensive, thorough, and of high quality. The draft report contained a complete description of the methods used, data collected, procedures followed to analyze the data, analysis of the results, conclusions, recommendations, references, and copies of programs and Arc Macro Language (AML) routines developed for the study. The draft final report meets or exceeds all requirements for the report listed in Article IV, Paragraph 3 of the contract, and contains all information that was outlined in the Scope of Work. The report is well written and understandable.