Analysis of Suspended Sediment Loads in Streams and Rivers using Linear Regression and Pearson Correlation

By Chantal Wase Sakwe

Submitted in partial fulfillment of the requirements
for the Degree of
Masters in Science
in the
Environmental Science Program

Youngstown State University
Fall 2015

Analysis of Suspended sediment Loads in Streams and Rivers using Linear Regression and Pearson Correlation

Chantal W. Sakwe

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Chantal W. Sakwe, Student Date

Approvals:

Dr. Isam E. Amin, Thesis Advisor Date

Dr. Alan M. Jacobs, Committee Member Date

Dr. Carl G. Johnston, Committee Member Date

Mr. Lawrence P. Gurlea, Committee Member Date

Dr. Gloria P. Johnston, Committee Member Date

Dr. Salvatore A. Sanders, Dean of Graduate Studies Date

Abstract

Sediment is the number one pollutant in the United States. Sediment in streams and rivers also carries along other pollutants such as nutrients (phosphorus and nitrates) and harmful bacteria. This study employs linear regression to predict the suspended sediment load, a dependent variable, as a function of the stream water discharge, an independent variable in seven U.S. Rivers and streams. The major objectives of the study are to investigate the effect of the sediment record length on the accuracy of the prediction, and to investigate the correlation of the suspended sediment load with nutrients and fecal coliform. The linear regression results showed that sediment sources/sinks and hydrologic variations that take place throughout the year play a vital role in the regression analysis of suspended sediment data. Five of the seven investigated rivers produced accurate predictions of the suspended sediment load. The sediment record length can affect the value of the correlation coefficient between the streamflow and suspended sediment rate. The percent deviation between the predicted and suspended sediment is less likely affected by the sediment record length and most likely affected by the hydrologic variations and sediment sources/sinks. Two of the investigated rivers show that significant hydrologic variations and sediment sources/sinks can increase the percent deviation, regardless of the sediment record length. This study also includes a Pearson correlation between the suspended sediment load, nutrients and fecal coliform. The Pearson correlation results showed no correlation between the suspended sediment loads and nutrients but high correlation was seen between the fecal coliform and the suspended sediment loads in the investigated rivers. The two investigated rivers for Pearson correlation are influenced by flow, which greatly determines the other parameters, even though the two rivers have different physical conditions.

Acknowledgements

I would like to express my heartfelt appreciation to my advisor, Dr. Isam E. Amin for his guidance, help, valuable remarks, close follow-up and support during this project. I would also like to thank Dr. Gloria Patricia Johnston for the SPSS lessons, always being ready to give me a helping hand and for letting me use her data.

I appreciate the helpful comments of my committee members: Dr. Carl G. Johnston, Dr. Alan M. Jacobs and Mr. Lawrence P. Gurlea and I thank them for being a part of this work.

My sincere gratitude goes to my family, especially to my husband Sona Ngoh for the encouraging words, helpful ideas and financial support. Thanks to my parents, Pierre and Joana Sakwe, my brother Louis Sakwe and sisters Nobelle Sakwe and Joy Sakwe, your love and support brought me this far.

Table of Contents

List of Figures

List of Tables

CHAPTER 1. Introduction

1.1 Introduction

Agricultural activities have generally been noted as the greatest contributor to water quality deterioration because it releases different types of material into water: sediments, pesticides, animal manures, fertilizers and other sources of inorganic and organic matter. Runoff and percolation are the major ways by which these pollutants are transported into surface water bodies and groundwater and are therefore known as non-point sources of pollution (Ongley 1996a).

Sediment is considered the greatest water pollutant in the United States mainly because sediment by itself is a pollutant, and other pollutants in the river could become either physically or chemically attached to the sediment and either carried along with it or stored at the bottom of the river (United States Army Corps of Engineers 2006; Laubel et al 1999; Ongley 1996b). About 25% of the stream length (268760 meters) in the United States suffers from the destructive effects related to excessive sediment loads (United States Environmental Protection agency 2006). Depending on the size of the particles in the river and flow pattern, the sediments can be transported as either bed load or suspended load (Van Rijn 1984). Bed load is stony material, such as gravel and cobbles that rolls along a river bed which cannot be suspended by the water current of the river due to its heavy weight, while suspended load is made up of sand, silt, clay-sized particles that are suspended in water due to the turbulence of the water (Ongley 1996b). Suspended sediment makes up the majority of sediment in a waterbody (Gray and Simŏes 2008; Ongley 1996b).

Most often, environmental problems in waterbodies including rivers, lakes, ponds, streams, canals, reservoirs and ditch systems are related to sediment accumulation (Botkin and Keller 2005). The critical role played by the different sources in river pollution has led to the growing concern on sediment transport as the major method by which non-point source pollutants from land are carried to the river (Russell et al. 2001). Other pollutants such as phosphorus, nitrogen, heavy metals, organic matter, pesticides and bacteria could be carried and stored by the sediment as sediment most often comes from forest areas, agricultural fields, impervious surface runoff, construction and mining sites (Amin and Jacobs 2007; Le et al. 2010). For example; transporting pollutants such as nitrogen, phosphorus and heavy metals is linked with the transportation of contaminated sediment downstream. Phosphorus and nitrogen are significant contributors to eutrophication and growth of toxic algae which results in drastic decrease in the dissolved oxygen in water bodies causing aquatic life to suffer and in most cases die. Moreover, nitrogen can be converted to nitrate which is of great concern in drinking water because it causes an illness known as methemoglobinemia commonly known as blue baby syndrome, which causes oxygen stress in the body (Amin and Jacobs 2007). Coarser sediments such as sand and gravel being transported in the water can also cause damaging effects to the environment (Laubel et al. 1999). For example, species population can be greatly affected when the pores of spawning grounds and substratum are covered with sediment (Madsen 1995; Lamba et al 2015). Erosion of river and stream banks often is an important source of sediment in the water, and accounts for up to about 50% of the transported sediment (Svendsen et al. 1995). Generally, channel processes are determined by sediment load in rivers (Berkun et al. 2015). Sediment build up raises water levels and occupies space that would have otherwise been occupied by water thereby reducing its total storage capacity and also increases flood episodes. Reservoirs filled with sediment are no longer useful for water storage. Sediment blocks sunlight from penetrating the water thereby inhibiting efficient primary production in the aquatic environment. (Troeh et al. 2004).

Sediment control is therefore crucial to mitigating the environmental effects of excess sediment in a water body (Amin and Jacobs 2007) and prediction of the suspended sediment load, the focus of this study, is essential for sediment control and mitigation of the pollution.

1.2 Problem statement

An excess sediment concentration in streams has detrimental effects on aquatic ecosystems, for example it covers vegetation and spawning grounds, blocks sunlight penetration and degrades water quality as explained above. Suspended sediment load prediction, therefore is important in the design of effective sediment control strategies and future water resources management. Therefore, awareness and knowledge of the prediction of suspended sediment load in streams is important for the protection of water resources and aquatic ecosystems.

1.3 Objectives

The objectives of the study are:

1. To investigate the effect of the sediment record length on the accuracy of predicting suspended sediment in streams and rivers using linear regression analysis.
2. To predict the annual suspended sediment load as the sum of the monthly loads and compare this prediction with the annual observed of sediment load.
3. To investigate the correlation between suspended sediment in streams, nutrients and fecal coliform.

CHAPTER 2. Literature Review

2.1 Literature Review

Sediment transport in rivers is associated with the transportation and storage of other pollutants. These pollutants can be attached to the sediment particles either physically or chemically. Sediment load in rivers is related to erosion of sediment particles (Lamba et al 2015). There are four types of erosion processes which can lead to the sediment accumulation in rivers: sheet, rill, gully and in-stream erosion. Sheet, rill and gully erosion are caused by overland flow for example on agricultural fields. Sheet and gully erosion are influenced mainly by rainfall (Merrit et al. 2003). The removal of sediment from the stream bank and bed is known as in-stream erosion. Severe in-stream erosion can cause the collapse of the stream bank, which leads to an increase in the amount of sediment in the stream. The majority of the sediment that is transported during high flow periods in a river mostly originates from the river channel (Merritt et al. 2003). Sources, sinks and mobilization of sediment are subject to change depending on the time and location. Suspended sediment input from the various sources can change seasonally. For instance in the spring, bare soils erode easily whereas the erosion rate decreases during the crop growing season (Trimble 1999).

The two main factors that influence suspended sediment load are rainfall and streamflow (Jie and Yu 2011). Generally, the natural river flows usually change according to the season; high flows are attributed to spring rains and snowmelt whereas low flows are related to warm, dry summer (American Rivers 2002). As the sediments are transported, three types of particles can be identified: fine particles that makeup the suspended material which includes silt, clay, and sand; the bedload and the saltation load, which are made up of coarse particles. Saltation load is material that is between bedload and suspended sediments, it is the material that can be lifted off the river bed due to its light weight but cannot remain long enough in suspension because it is heavier than the suspended particles (Ongley 1996). The movement of all three particle types in the water can be distinguished into three categories: suspended particle motion, rolling and sliding motion or both and the saltation motion which is the hopping of the particles down the stream bed.

The suspended sediment load of a stream is usually determined by direct measurement of the concentrations of sediment or by using sediment transport equations, with direct measurement being the more reliable method. However, using the direct measurement method, which includes setting up gaging stations at particular locations, over a long period of time is neither economical nor feasible. There are three common groups of sediment transport equations which include: physically based, empirical, and regression-based. The physically based models require huge amounts of data and involve a lot of parameter estimates (Tayfur 2003). Practically, empirical models cannot be applied to general cases and work only for situations and locations for which they have been developed (Yang 1996; Tayfur 2003). Regression based models have commonly been used to predict sediment load in rivers owing to the fact that they are simple and easily applicable (Jain 2001; Walling 1977). Furthermore, they are free from the limitations of the physically based models (large data requirement) and empirical equations (specific situation and location requirement).

Ferguson (1987) has shown a possibility of underestimating bias which is inevitable when using log-transformed regression analysis to derive relationships between water discharge and sediment load. His idea is that the bias is the main factor related to errors in this approach and can be corrected by using a simple correction factor which is based on the standard error of estimate of the logarithmic regression. Ferguson in his study tested the accuracy of bias corrected sediment load against the observed sediment loads in streams in an attempt to rate the application of the bias correction factor. His results showed that the bias correction factor reduced the degree of underestimation linked with the standard rating curve estimates yielding predicted values that were within 10 percent of the observed values. Hansen and Bray (1987) applied the smearing estimate correction factor to estimate the suspended sediment yield of the Kennebecasis River in New Brunswick, Canada, and got similar results as Ferguson (1987) with suspended sediment loads within 10% of the observed values. Some authors are still hesitant about using the bias correction factor for fear that it would produce unreliable results of the suspended sediment loads (Hansen and Bray 1987; Ferguson 1987). Conversely, other authors have been inspired by their findings and have been encouraged to apply the bias correction factor assuming that, this method will produce reliable values of sediment yield which can be used in establishing sediment budgets (Stott et al. 1986).

Leopold and Maddock (1953) were able use the linear regression method to relate suspended sediment load and discharge of the Powder River at Arvnda, Wyoming. Their results showed that the suspended sediment load per unit of discharge increased as discharge increased. Further investigation showed that suspended sediment load per unit volume of water decreased slightly downstream.

Leopold and Miller (1956) in their study went a step further to differentiate the effect of discharge on suspended sediment load in various perennial and ephemeral streams. They also applied the linear regression model to establish a relationship between the suspended sediment load and discharge. Their work showed that the suspended sediment load of the arroyo increased downstream more rapidly than discharge, therefore producing higher loads of suspended sediment downstream. It was the reverse for the perennial streams, as the suspended sediment load was observed to decrease downstream as determined by Leopold and Maddock (1953).

Bhowmik et al. (1980) in their study investigated the sediment transport and water discharge for two separate years, 1967-1968 and 1977-1978 in the Kankakee River in Illinois. From the data collected, regression equations were used to establish relationships between water discharge and suspended sediment discharge. They observed from the daily data that a peak water discharge did not always coincide with a peak sediment discharge and at many stations, the sediment load per square mile changed significantly for the same water discharge, making the establishment of direct relationships between water discharge and sediment load for every gauging station challenging.

Amin and Jacobs (2007) in their study showed how linear regression analysis can be used to account for sediment sources and sinks in the prediction of suspended sediment load in streams. In this study, regression equations were used to predict the monthly and annual suspended sediment load in Rio Puerco River in central New Mexico, considering the effects of sources and sinks as opposed to the traditional linear regression approach which does not take the effects of sources and sinks into consideration. Applying the traditional linear regression approach, the monthly regression relationship was used to examine the effects of sources and sinks because it produced the highest correlation coefficient compared to the daily and annual relationships. Three regression lines were fitted through the monthly suspended sediment data, whose equations were used to determine how much sediment was either gained from sources or lost to sinks. The correlation coefficient increased from 0.93 to 0.98 when the effects of sources and sinks were included in the simulation, leading to considerable improvement in the accuracy of the prediction of the suspended sediment load using linear regression. This was reflected in the absolute value of the percent deviation of the predicted load from the observed values. The percent deviation is defined as the difference between the observed and predicted sediment divided by the observed value and multiplied by 100%. Considering sediment sources and sinks has lowered the median and mean value of the percent deviation, i.e., increased the accuracy of the prediction. The median and mean values of the percent deviation were 20.5% and 45.7% for the traditional approach and 12.9% and 38.5% for the proposal approach.

Past studies have mentioned that in order to get good results using regression analysis, sediment data should preferably be collected on a daily or weekly basis over a period between 10 to 20 years (Bhomik et al. 1980; Amin and Jacobs 2007).

2.2 Advantages and Disadvantages of the Linear Regression Model

Using linear regression for sediment load prediction is useful and because of the erratic nature of sediment in streams, this method can be applied to sediment load studies for an extended period of time at a particular gauging station. This approach is also preferred as it usually produces good results for monthly and annual sediment load predictions in large drainage basin streams. This is because when the daily loads are added to give the monthly loads, the daily variations are wiped out making the monthly relationship more accurate. In general, the annual relationship is considered less accurate compared to the monthly regression relationship, as the annual regression relationship experiences more hydrologic variations during the year. The monthly regression relationship is, therefore, a more rational choice for predicting the suspended sediment load using linear regression (Amin and Jacobs 2007).

Even though the linear regression approach has advantages, it also has drawbacks. Predictability problems could be encountered, for instance, sediment supply and discharge are major contributors in determining sediment load. Therefore, a change in sediment supply can offset the relationship outcome between discharge and suspended sediment (Araujo et al. 2012). The daily peak water discharge may not always coincide with the daily peak sediment discharge and the sediment load can change significantly for the same water discharge (Bhowmik et al. 1980). Data associated with water quality conditions can be challenging and inaccuracy and discontinuous measurements can be problematic (Araujo et al. 2012).

CHAPTER 3. Materials and Methods

3.1 Study sites

Seven rivers were investigated in this study:

1. Maumee River at Waterville, Ohio [Sediment record = 54 years]
2. Sacramento River, near Red Bluff, California [Sediment record = 54years]
3. Cuyahoga River at Independence, Ohio [Sediment record = 37years]
4. Mississippi River at McGregor, Iowa [Sediment record =30years]
5. Kankakee River at Momence, Illinois [Sediment record =17years]
6. Smoky Hill River at Lindsborg, Kansas [Sediment record =18years]
7. Cuyahoga River at Old Portage, Ohio [Sediment record =10years]

3.1.1 Maumee River

The Maumee River is located in northwestern Ohio. The river begins at Fort Wayne, Indiana, at the confluence of the St. Mary's and the St. Joseph Rivers and it is about 130 miles long. It covers approximately 5,024 square miles in Ohio and flows through about 18 counties. The Maumee River flows into Lake Erie at Toledo, Ohio. It has about four thousand miles of streams, creeks, and rivers emptying into it, making it the largest watershed of any river flowing into a Great Lake. The watershed is mostly made up of agricultural lands with urban development to an extent, hay and pasture lands, and forest (Ohio History Central – Maumee River, n.d; Ohio Environmental Protection Agency- Maumee River Watershed, n.d).

The Maumee River is a perennial river that flows continuously throughout the year. The Maumee River watershed contributes about 50% of sediment to the Lake Erie as well as about 40% of the phosphorus load into this great lake which has been noted to be mainly from agriculture, including manure (Maumee River Watershed, n.d.). During the period 1950 - 2003 the Waterville gaging station has an annual peak streamflow of approximately $3.1x10^6 ft^3/s$ and a low of $7.2x10^5 ft^3/s$ with an average annual suspended sediment load record of $1.3x10^6$ tons/year (United States Geological Survey Sediment Portal). The annual mean precipitation in the area is about 33.21 inches. Rainfall in the area is evenly distributed throughout the year. June has the highest rainfall recorded in the year with an average of 3.80 inches (Waterville Ohio weather 2015).

3.1.2 Sacramento River

This is the largest river and watershed system in California (by discharge, it is the second largest U.S. river draining into the Pacific, after the Columbia River). It covers about 27,000 square miles and drains the eastern slopes of the Coast Range, Mount Shasta, the western slopes of the southernmost region of the Cascades, and the northern portion of the Sierra Nevada. It is about 447 km long and approximately one third of the state's total surface runoff ends up in this river. Primary tributaries to the Sacramento River are the Pit, Feather, and American Rivers. The Sacramento River Basin provides drinking water for residents of northern and southern California, supplies water for agriculture and is a home to hundreds of wildlife species, including four separate runs of Chinook salmon (Mount, n.d.).

The Sacramento River is an ephemeral river that flows only after a rainfall event. Annual precipitation varies and over 80% of the precipitation in this area falls within the 6 months from November to April and originates mostly from region wide storms. The mean annual precipitation can be as high as or even above 100 inches and as low as and less than 10 inches (Jones et al. 1972), specifically 24.52 inches at Red Bluff (US Climate data 2015). The annual sediment load in the river varies greatly due to the variations in streamflow (Jones et al. 1972). During the period 1960 – 2013, the annual peak streamflow at the Red Bluff gaging station was $1.8x10^7 ft^3/s$ and a low of $2.8x10^6 ft^3/s$, with an average annual suspended sediment load record of $1.8x10^6$ tons/year between these years (United States Geological Survey Sediment Portal).

3.1.3 Cuyahoga River

The Cuyahoga River is located in Northeast Ohio. The river begins officially in Geauga County and from this point, is about 85miles long and flows south to Cuyahoga Falls, where it turns sharply north until it empties into Lake Erie. The river drains 813 square miles of land traveling through at least 6 counties. Today this river is a home to many different aquatic species, as it is recovering from past pollution. Sewage from cities and waste from industrial activities was dumped in the river especially in the section between Akron and Cleveland, and in the summer of 1969, the river caught fire and is now referred to as the river that burned. The river even though not fully recovered now inhabits about 40 species of fish, sensitive aquatic bugs and even fish that survive only in clean waters such as the steelhead trout and northern pike (Environmental Protection Agency 2013).

The Cuyahoga River is a perennial river. It receives an annual average precipitation of about 35 in. including snowfall, and approximately 20 inches out of the 35 inches normally fall from April to September. The river contributes minimal sediment load to Lake Erie (Cuyahoga Valley National Park Ohio 2013). Between 1972 and 2009, the peak annual flow of the gaging station at Independence was $5.3x10^5 ft^3/s$, and a low flow of $1.3x10^4 ft^3/s$ has been recorded with an average annual suspended sediment load of $2.4x10^5$ tons/year. At the Old Portage gaging station, an annual peak flow of $2.0x10^5 ft^3/s$ and a low flow of $9.2x10^4 ft^3/s$ have been recorded between 1972 and 1981, with an average annual suspended sediment load of $3.2x10^4$ tons/year (United States Geological Survey Sediment Portal).

3.1.4 Mississippi River

This is the second longest river in the United States after the Missouri River. The Mississippi River is about 2,340 miles long and flows through ten states: Minnesota, Wisconsin, Iowa, Illinois, Missouri, Kentucky, Tennessee, Arkansas, Mississippi, and Louisiana. It discharges into the Gulf of Mexico. The Mississippi River system drains the agricultural plains between the Appalachian Mountains to the east and the Rocky Mountains to the west, with its drainage basin (approximately 1,234,700 square miles) occupying up to 40 percent of the United States, making it the fifth largest in the world. The river serves as a primary drinking water source for millions of Americans and also supports a wide variety of fish and wildlife (American Rivers 2014).

The Mississippi River is a perennial river. The study area has an average annual precipitation of 33.43inches. Rainfall is equally distributed throughout the year. August being the wettest month of the year has an average rainfall of 4.6inches. (McGregor

Weather 2015). A large amount of sediment of about 2.1×10^8 tons/year is discharged by this river into the Gulf of Mexico (Myint and Walker 2002). The peak discharge at McGregor gaging station between the years 1975 and 2004 was $2.4x10^7ft^3/s$, with a low flow of $5.7x10^6ft^3/s$ and average annual suspended sediment of $1.6x10^6$ tons/year (United States Geological Survey Sediment Portal).

3.1.5 Kankakee River

This River flows westward from Indiana to Illinois. The Kankakee River has twelve larger tributary streams, including the Iroquois River, and cuts across at least thirteen northwestern Indiana Counties and rises from the springs and swamplands of Northwest Indiana. Its total basin area is approximately 5165 square miles, 2169 miles in Illinois and 2996 miles in Indiana with a total length of about 150 miles, of which 91 miles is in Indiana. It joins the Des Plaines River and becomes the Illinois River. (Ivens et al. 1981). The Kankakee River is a very high quality system, supporting a high diversity of fishes and mussels. It is well known as an excellent sport fishery for smallmouth bass, walleye, channel catfish, rock bass, and northern pike. It provides water predominantly to the population surrounding it (The Kankakee River 2010).

The Kankakee is an ephemeral river. The study area has an average annual precipitation of 39.17in. It is characterized by high rainfall and the wettest month is June, with an average precipitation of 4.21inches. (City of Momence 2015). The peak annual discharge at this gaging station between the years 1979 and 1995 was $2.0x10^6ft^3/s$, with a low flow of $7.1x10^5ft^3/$ and an annual suspended sediment loads of $2.6x10^5$ tons/year (United States Geological Survey Sediment Portal).

3.1.6 Smoky Hill River

This river is about 575 miles long and cuts across a few counties and eventually unites with the Republican River to form the Kansas River. The Smoky Hill River covers an area of approximately 8,810 square miles. In some areas of the western half of the Smoky Hill-Saline basin is the Ogallala high plains aquifer which accounts for the majority of water use, with irrigation using the most water (Kansas Water Office 2011).

The Smoky Hill River is a perennial river. The study area receives an annual precipitation of approximately 16 inches in the extreme west and 30 inches in the east. There are large variations in the annual precipitation and most of the annual rainfall results from thunderstorms. Most of the precipitation is seen between April and September which accounts for up to 75% of the annual rainfall (Kansas Water Office 2009). The annual peak discharge at Lindsborg gaging station between the years 1958 to 1975 was $2.0x10^6 ft^3/s$, with a low flow of $8.3x10^4 ft^3/s$ and an average annual suspended sediment load of $2.1x10^6$ tons/year (United States Geological Survey Sediment Portal).

3.2 Sediment data

The sediment data for the seven investigated streams was downloaded from the United States Geological Survey (USGS) sediment portal website (http://cida.usgs.gov/sediment/). The data is recorded as daily sediment load and streamflow of the rivers. The daily data consist of mean daily streamflow and mean daily suspended sediment data.

3.3 Nutrients and Fecal coliform data

The nutrient data for the Maumee River at Waterville was obtained from the USGS website. The data were recorded as annual loads of suspended sediment concentration, total phosphorus, nitrate and total nitrate. The data was collected in 2002.

Zoar Valley of Cattaraugus, NY data was used because of available data for comparison purposes. The data for this stream was provided by Dr. Gloria Patricia Johnston, which was used for her Masters thesis – "Assessment of water quality, benthic community structure and microbial indicators in Cattaraugus Creek, Zoar Valley, NY" (Basto Salgado 2005). The data was collected over a two year period; 2004 and 2005.

Zoar Valley of Cattaraugus creek is found in western New York State and is a tributary of the Lake Erie, with unique natural ecosystems in Northeastern United States. It possibly surrounds the premier old-growth broadleaf forest in the area The Cattaraugus creek watershed covers approximately 552 square miles, and includes forests, wetlands, agricultural lands and small villages. The Cattaraugus creek is approximately 50 miles in length. The creek originates at the Java Lake in Wyoming County and flows west into Lake Erie. Flow discharge is typically $<177 ft^3/s$ (Basto Salgado 2005).

3.4 Data Analysis

Data analysis was done using excel and SPSS. Four different lengths of sediment record were considered for comparison purposes: 50 years+, 30 years+, 15 years+ and 10 years. The monthly and annual data for the sediment load and flow rates were determined from the daily data. The daily values for each month of each year were summed to get the monthly values for that year, which were then be added up to get the annual data for each year.

3.4.1 Linear Regression analysis

Linear regression is commonly used in the study of sediment discharge to show the relationship between suspended sediment and water discharge. It investigates the long term variability of the suspended sediment load at particular gaging stations (Amin and Jacobs 2007). This study will predict the suspended sediment load, a dependent variable, as a function of stream water discharge, an independent variable, at all investigated rivers. The regression equation usually takes the form:

$$\mathbf{Y = aX^m} \qquad (1)$$

where Y= Suspended sediment load
X= Stream water discharge
a= a constant
m= slope of the regression line

In logarithmic form the regression equation reads:

$$\mathbf{\log Y = m \log X + \log a} \qquad (2) \qquad \text{which is in the form } y = mx + b,$$

where b = log a = intercept and
m = slope of the regression line.

The regression equation will therefore be used to predict the suspended sediment load on a daily, monthly and annual basis.

The regression analysis involves the transformation of the suspended sediment load data into logarithmic functions (equation 2) which will have to be retransformed to the original engineering units. This retransformation involves a bias that needs to be corrected (Amin and Jacobs 2007). This will be accomplished by applying the smearing estimator as proposed by Cohn and Gilroy 1991 using the following equation:

$$\mathbf{Y_{SE} = Y\ [\ \sum 10^{res} / \mathit{n}]}$$

where Y_{SE} = the corrected predicted sediment load using the smearing factor
Y = the predicted sediment load
n = number of the predicted sediment loads
res = residuals = [log observed sediment load] – [log predicted sediment load] = correction factor.

The percent deviation of the annual predicted sediment from the observed sediment was determined for each river and used to define the accuracy of the prediction. The percent deviation is the difference between the observed sediment and the predicted sediment as shown above. The lower the percent deviation, the better the accuracy of the prediction (Palmer and O'Connell 2009).

3.4.2 Pearson correlation analysis

Pearson correlation analysis using SPSS is the most commonly used measurement between data sets to see how well the data is related. It measures the strength and direction of association that exists between two variables on at least an interval scale. The Pearson correlation coefficient denoted r, can be any value within the range of +1 and -1. If the value is zero, then no relationship exist between the two variables. A positive relationship is seen if the value is between +1 and 0, which can be interpreted as an increase in the value of one variable results to an increase in the value of the other variable. Likewise, a negative association is seen when the value is less than 0 that is an increase in the value of one variable produces the reverse in the value of the other variable. (Laerd Statistics 2013).

Nevertheless, getting a Pearson correlation coefficient of +1 does not always mean a unit increase in one variable will lead to a unit increase in the other. It basically means that the data points fall along the line of best fit. Therefore the strength of the relationship can be determined by how close the Pearson correlation coefficient is to +1 or -1, depending on if the association is positive or negative respectively.

The Pearson correlation coefficient can be interpreted as follows:

	Coefficient, *r*	
Strength of Association	Positive	Negative
Small	.1 to .3	-0.1 to -0.3
Medium	.3 to .5	-0.3 to -0.5
Large	.5 to 1.0	-0.5 to -1.0

(Laerd Statistics 2013

CHAPTER 4. Results and Discussion

4.1 Results of Linear Regression

Regression equations were used to build relationships between the suspended sediment loads and the streamflow of the seven rivers investigated. The suspended sediment load is measured in tons per day for the daily analysis, tons per month for the monthly analysis and tons per year for the annual analysis. Streamflow is expressed in cubic feet per second (ft^3/s).

4.1.1 Sediment Record length: 50 Years and above

4.1.1.1 Maumee River at Waterville, Ohio

Figure 1 shows that the daily suspended sediment and streamflow correlate very well, with a correlation coefficient r = 0.96. The monthly and annual relationship of the suspended sediment and streamflow are illustrated in figures 2 and 3, respectively. The monthly suspended sediment also correlates well with the streamflow and has a correlation coefficient r = 0.97. The annual correlation coefficient r = 0.86. The monthly correlation coefficient is higher than the daily because the monthly suspended sediment load is obtained by adding the daily loads, and this clears out the daily variations. The annual processes produced are also prone to various levels of hydrologic changes. The relatively low correlation coefficient of the annual relationship is due to the immense hydrologic changes that take place throughout the year.

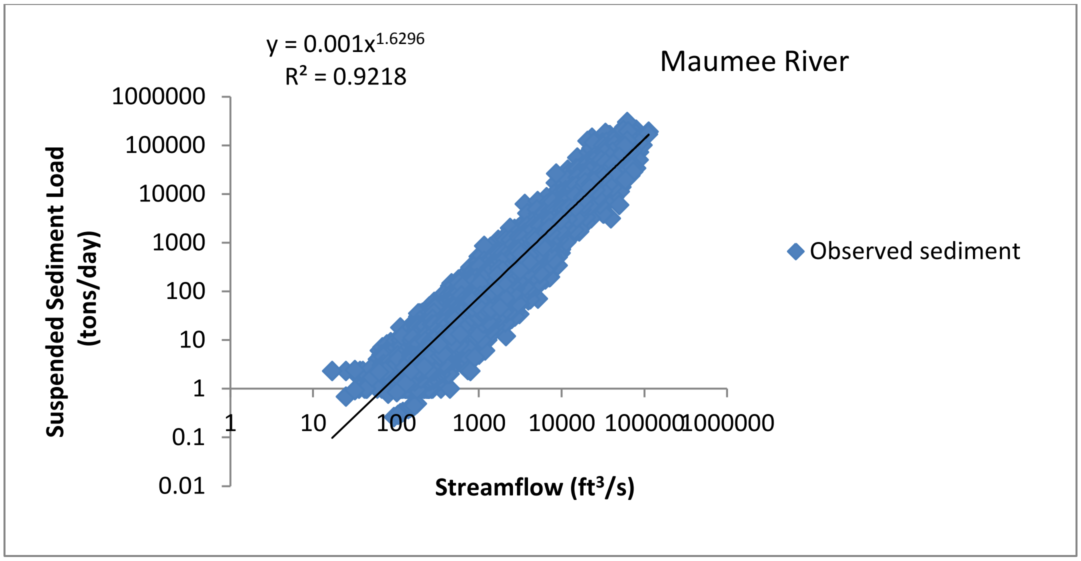

Figure 1: Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Maumee River.
The regression equation for this relationship is $Y = 0.001X^{1.6296}$

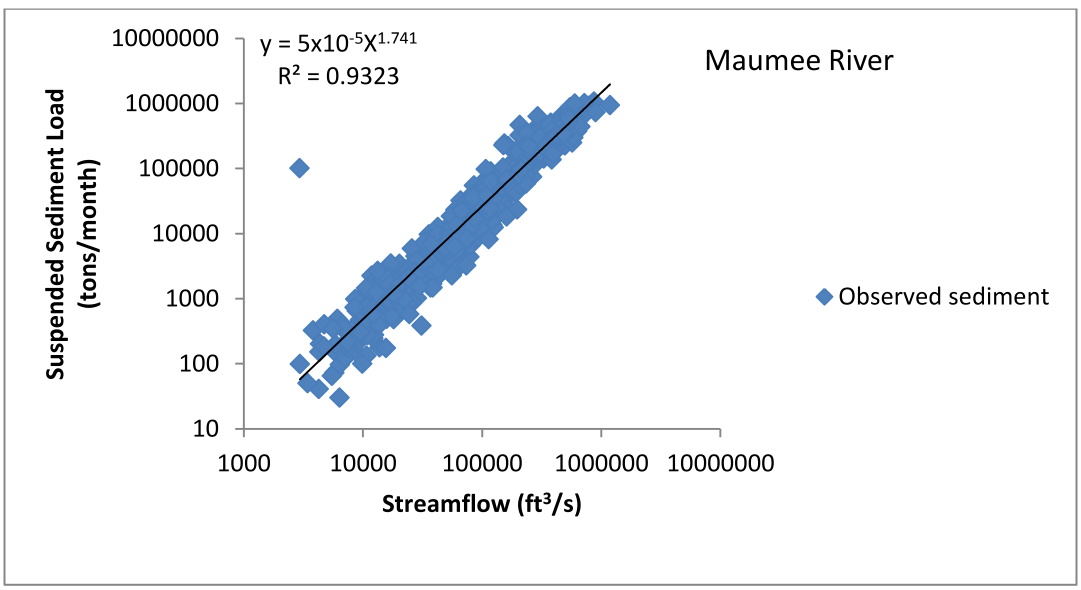

Figure 2. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft3/s) of the Maumee River.
The regression equation for this relationship is $Y = 5x10^{-5}X^{1.741}$

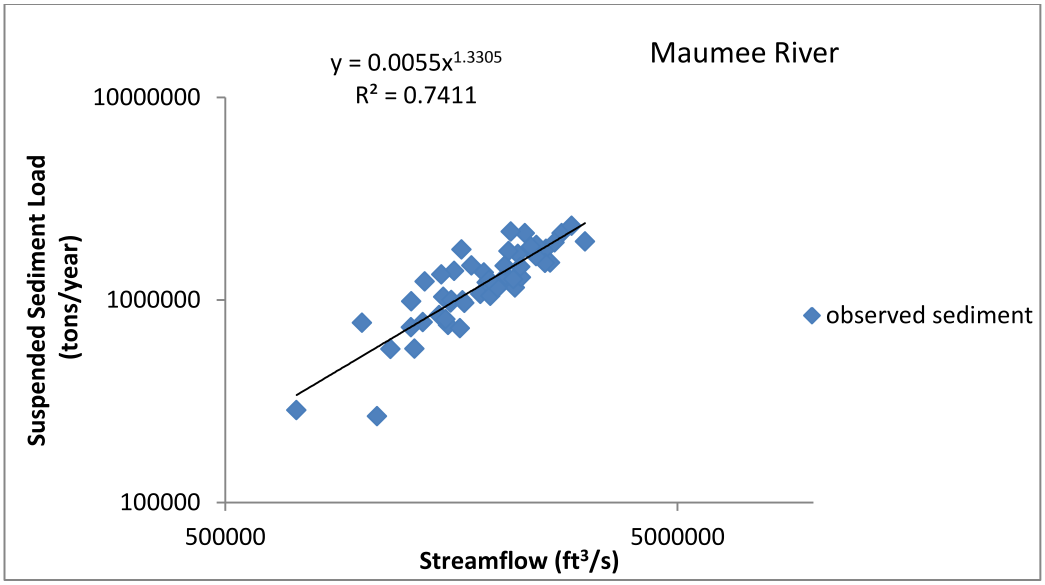

Figure 3. Regression relationship between the annual suspended sediment load (tons/year) and streamflow (ft3/s) of the Maumee River.
The regression equation for this relationship is $Y = 0.0055X^{1.3305}$

4.1.1.2 Sacramento River near Red Bluff, California

Figures 4, 5 and 6 show the regression relationships between the suspended sediment load and streamflow of the gaging station near Red Bluff on the Sacramento River. Figure 4 shows a good correlation between the daily suspended sediment load and streamflow, $r = 0.89$. The monthly and annual relationships also show a good correlation between the suspended sediment and streamflow, $r = 0.87$ and 0.83, respectively. The correlation coefficients of the daily and monthly relationships indicate practically similar variations in the hydrologic conditions and the daily and monthly sediment loads. As explained in the last section, the annual correlation coefficient is lower than the daily and monthly correlation coefficients because of the vast hydrologic changes that take place throughout the year.

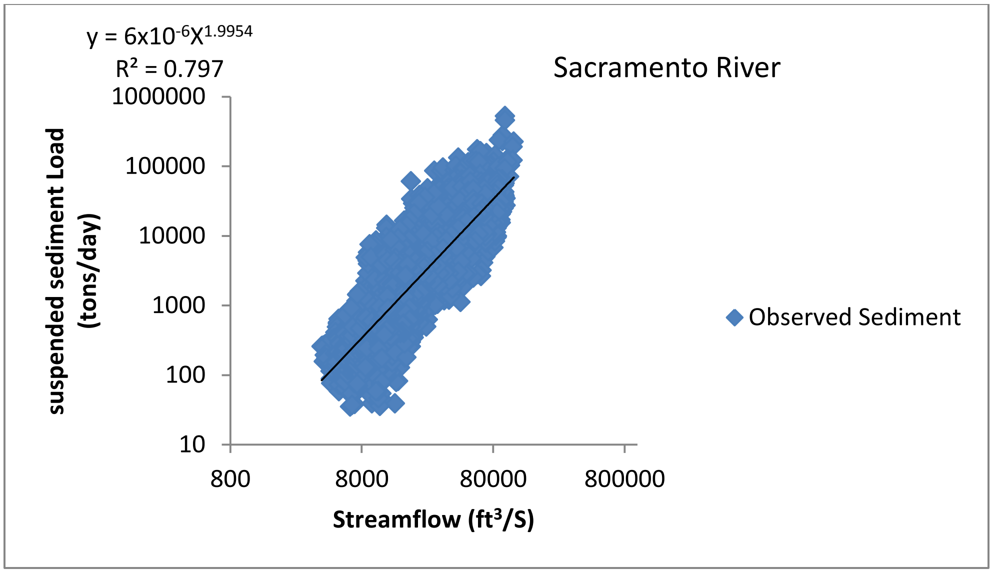

Figure 4. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Sacramento River.
The regression equation for this relationship is $Y = 6x10^{-6}X^{1.9954}$

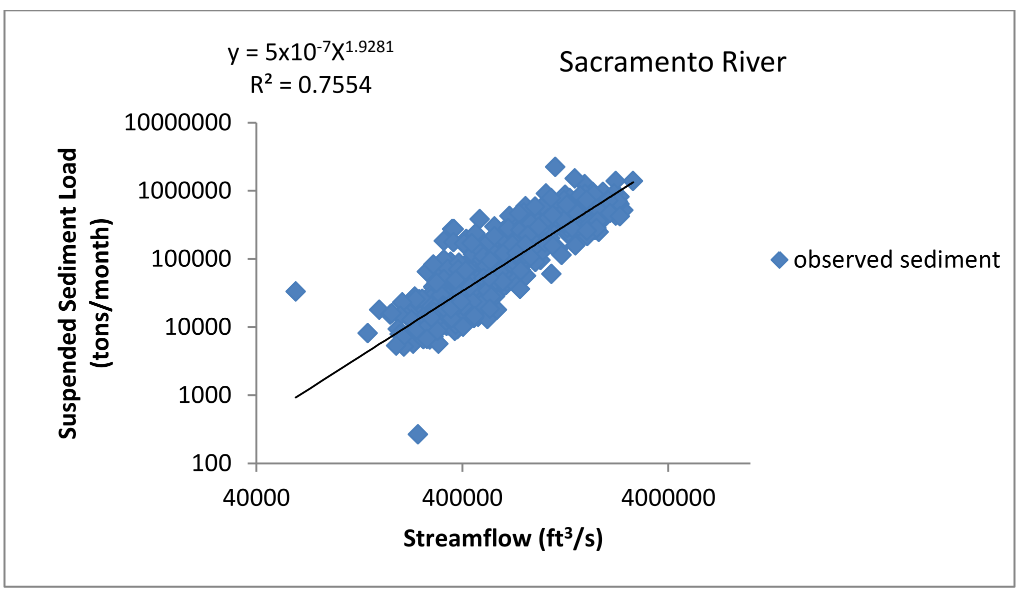

Figure 5. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft^3/s).
The regression equation for this relationship is $Y = 5x10^{-7}X^{1.9281}$

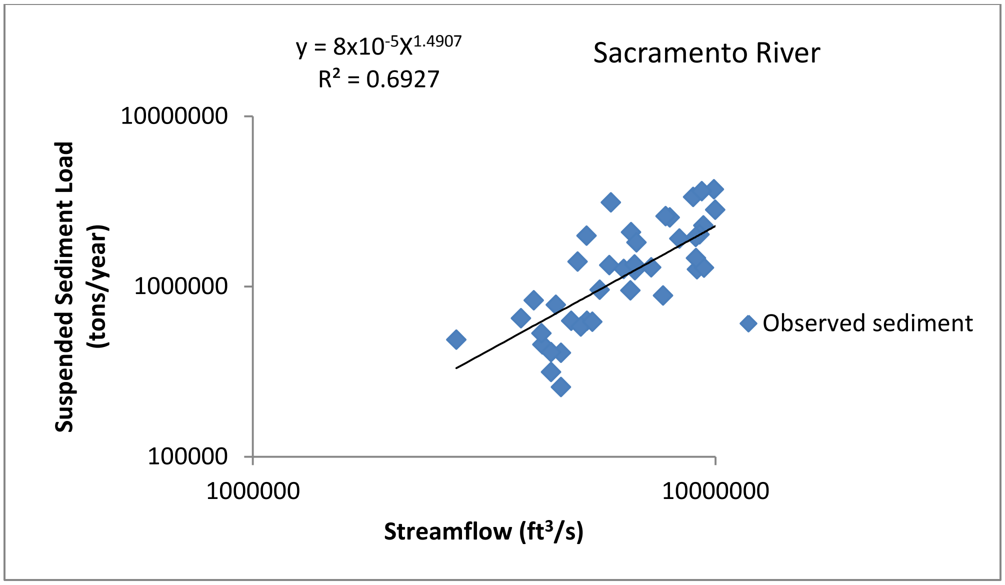

Figure 6. Regression relationship between the annual suspended sediment load (tons/year) and streamflow (ft3/s) of the Sacramento River.
The regression equation for this relationship is $Y = 8x10^{-5}X^{1.4907}$

4.1.2 Sediment Record length: 30 years and above

4.1.2.1 Cuyahoga River at Independence, Ohio

Figure 7 shows that the daily suspended sediment at Independence, Ohio of the Cuyahoga River correlates well with the streamflow, r = 0.90. The monthly relationship as illustrated in figure 8 also shows a good correlation between the suspended sediment and streamflow, r = 0.85. The annual correlation coefficient r = 0.78 also indicates good correlation between the annual suspended sediment load and streamflow as shown in figure 9. The monthly correlation coefficient is slightly lower than the daily coefficient. The daily variations are smoothed out in the monthly loads, the monthly relationship is considered more accurate for the prediction of suspended sediment load. This is not the case for this gaging station.

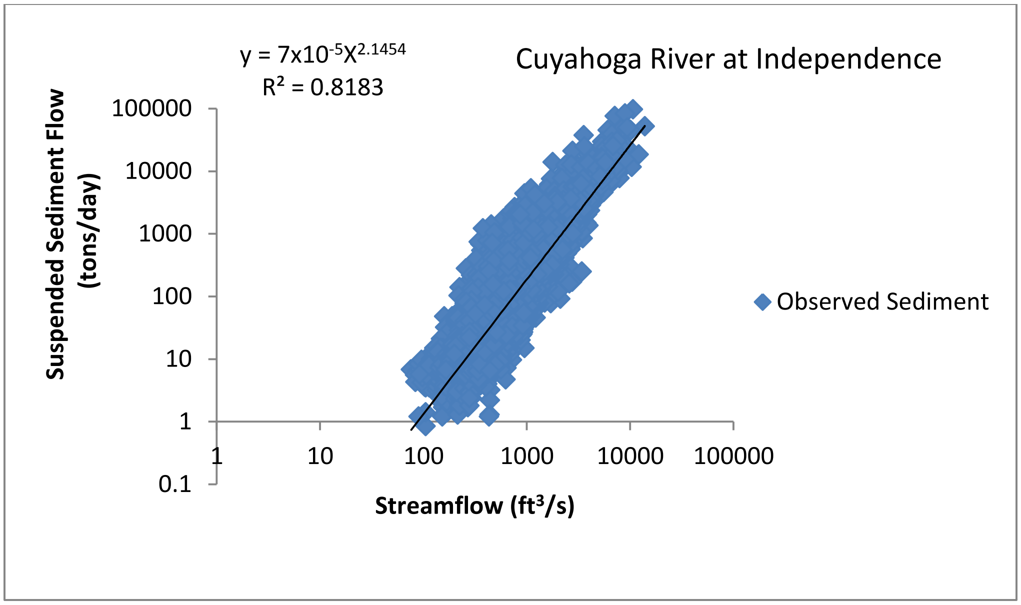

Figure 7. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Cuyahoga River at Independence.
The regression equation for this relationship is $Y = 7x10^{-5}X^{2.1454}$

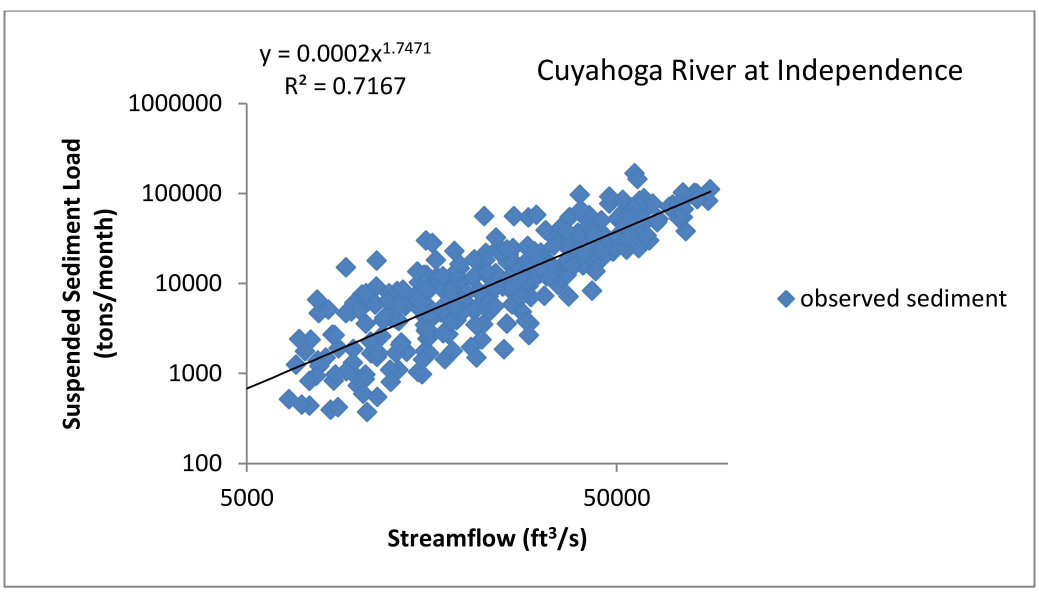

Figure 8. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft^3/s) of the Cuyahoga River at Independence.
The regression equation for this relationship is $Y = 0.0002X^{1.7471}$

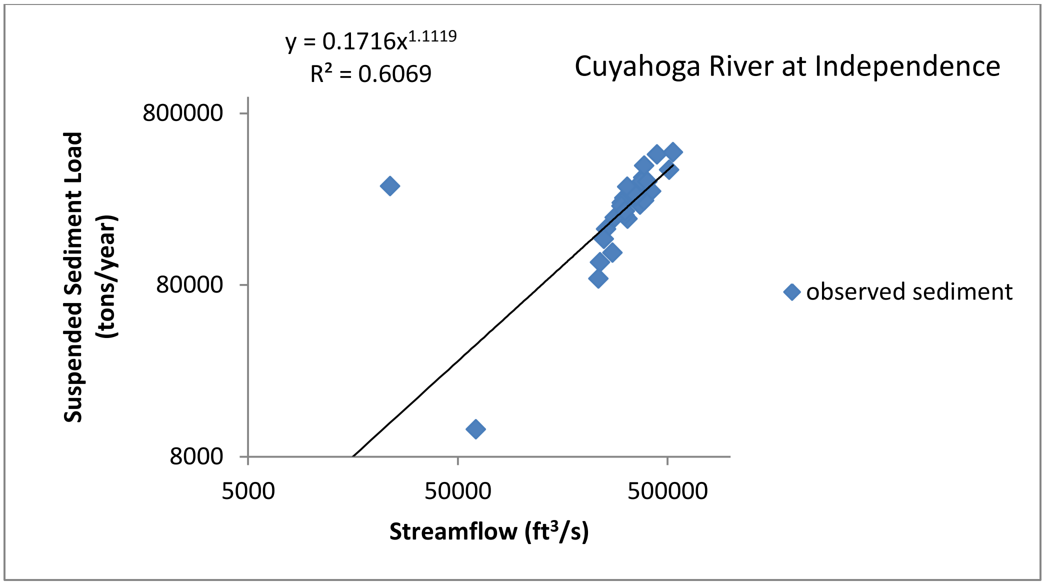

Figure 9. Regression relationship between annual suspended sediment load (tons/year) and streamflow (ft^3/s) of the Cuyahoga River at Independence.
The regression equation for this relationship is $Y = 0.1716X^{1.1119}$

4.1.2.2 Mississippi river at McGregor, Iowa

Figure 10 shows a somewhat lower correlation between the daily suspended sediment and streamflow with a correlation coefficient $r = 0.72$. The monthly and annual relationship of the suspended sediment and streamflow are illustrated in figures 11 and 12, respectively. The monthly suspended sediment is also not strongly correlated to the streamflow and has a correlation coefficient $r = 0.75$, but slightly higher than the daily correlation coefficient. As expected, the annual correlation coefficient $r = 0.59$ and is lower than both the daily and monthly correlation coefficients.

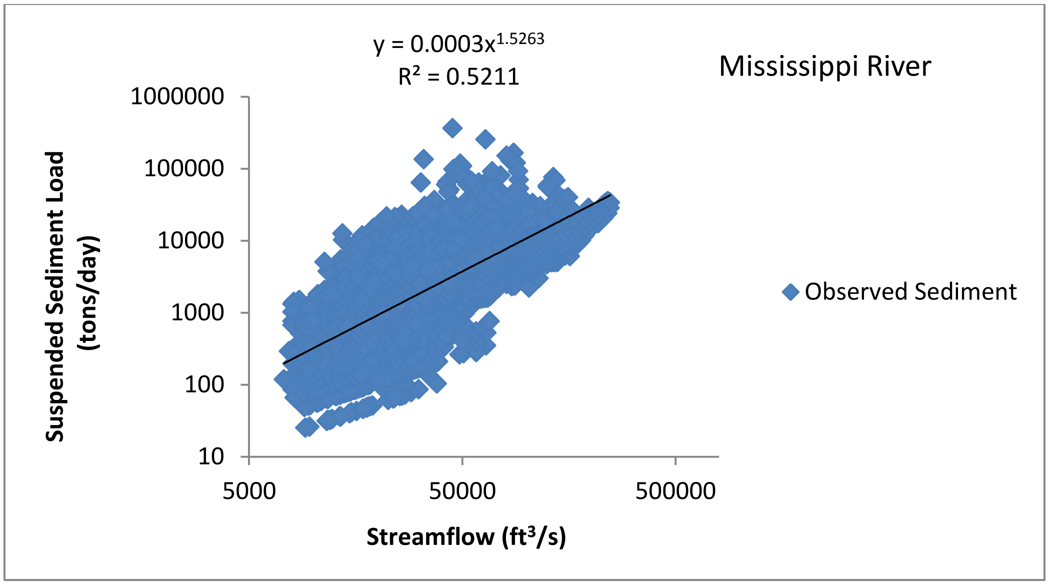

Figure 10. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Mississippi River.
The regression equation for this relationship is $Y = 0.0003X^{1.5263}$

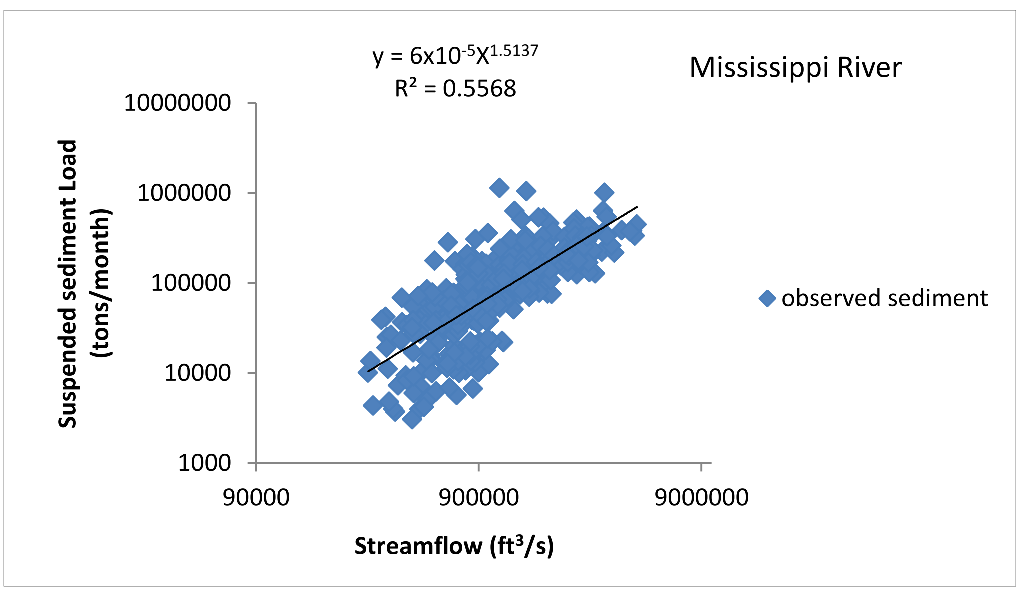

Figure 11. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft^3/s) of the Mississippi River.
The regression equation for this relationship is $Y = 6x10^{-5}X^{1.5137}$

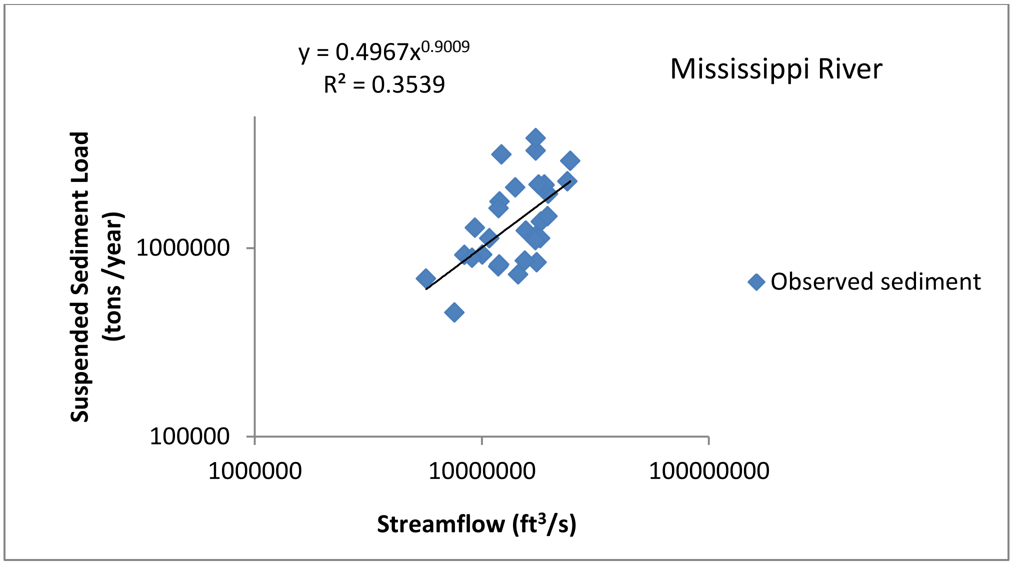

Figure 12. Regression relationship between annual suspended sediment load (tons/year) and streamflow (ft^3/s) of the Mississippi River.
The regression equation for this relationship is $Y = 0.4967X^{0.9009}$

4.1.3 Sediment Record Length: 15 Years and above

4.1.3.1 Kankakee River at Momence Illinois

Figure 13 shows that the daily suspended sediment at Momence, Illinois on the Kankakee River has a good correlation with the streamflow, r = 0.82. The monthly relationship as illustrated in figure 14 also shows a good correlation between the suspended sediment and streamflow, r = 0.88. The annual correlation coefficient r = 0.83 which also indicates good correlation between the annual suspended sediment load and annual streamflow is illustrated in figure 15. As explained earlier, the monthly correlation coefficient is higher than the daily and annual. The annual processes produced are also prone to various levels of hydrologic changes throughout the year.

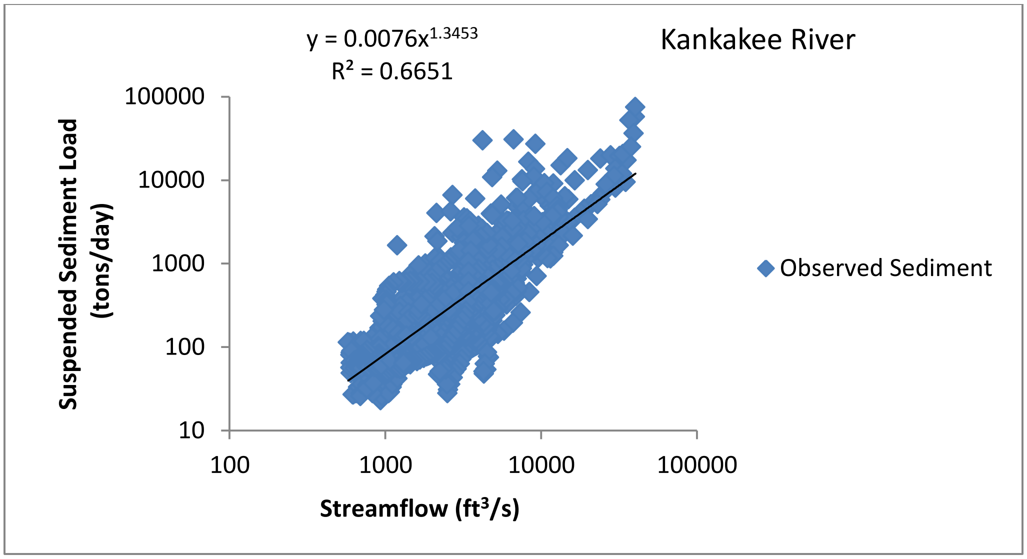

Figure 13. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Kankakee River.
The regression equation for this relationship is $Y = 0.0076X^{1.3453}$

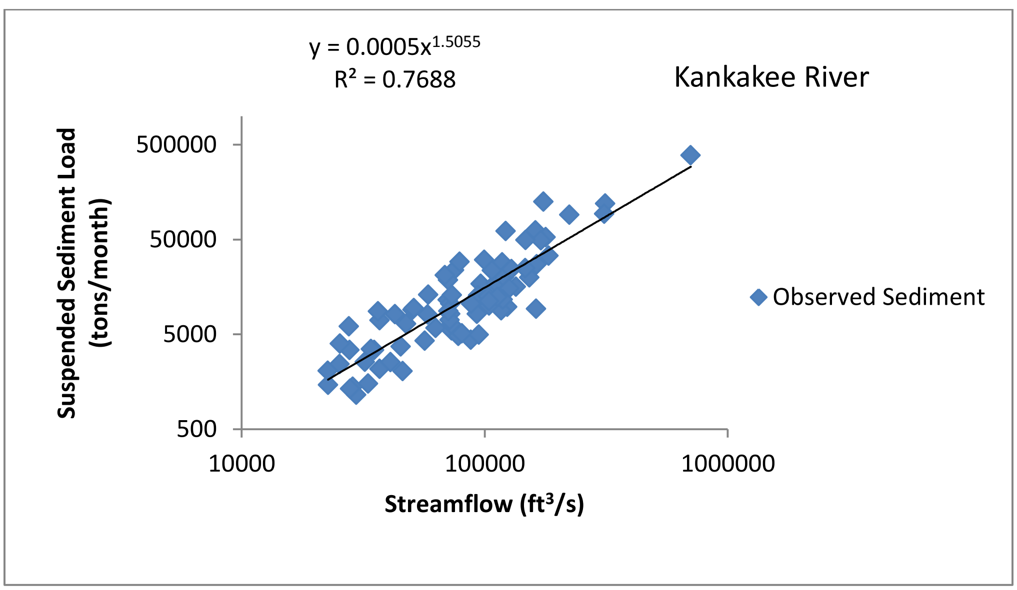

Figure 14. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft3/s) of the Kankakee River.
The regression equation for this relationship is $Y = 0.0005X^{1.5055}$

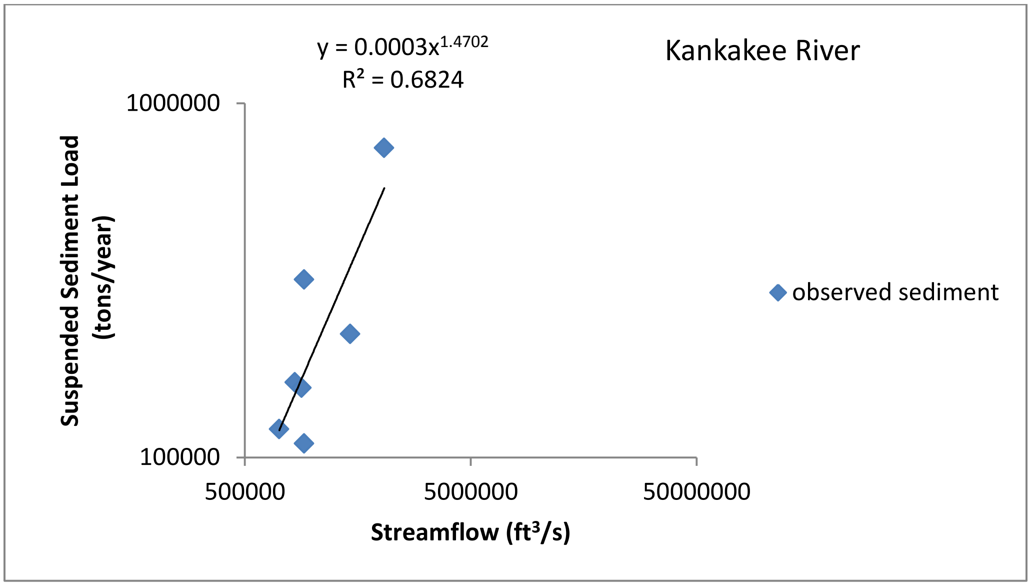

Figure 15. Regression relationship between annual suspended sediment load (tons/year) and streamflow (ft^3/s) of the Kankakee River.
The regression equation for this relationship is $Y = 0.0003X^{1.4702}$

4.1.3.2 Smoky Hill River at Lindsborg, Kansas

Figure 16 shows a correlation coefficient r = 0.91 for the daily suspended sediment and streamflow. The monthly relationship has a correlation coefficient r = 0.89 and the annual correlation coefficient r = 0.84 as illustrated in figures 17 and 18, respectively. The annual correlation coefficient is slightly lower than the monthly because even though both are produced by processes that depend on hydrologic variations, the annual relationship is more susceptible to these changes.

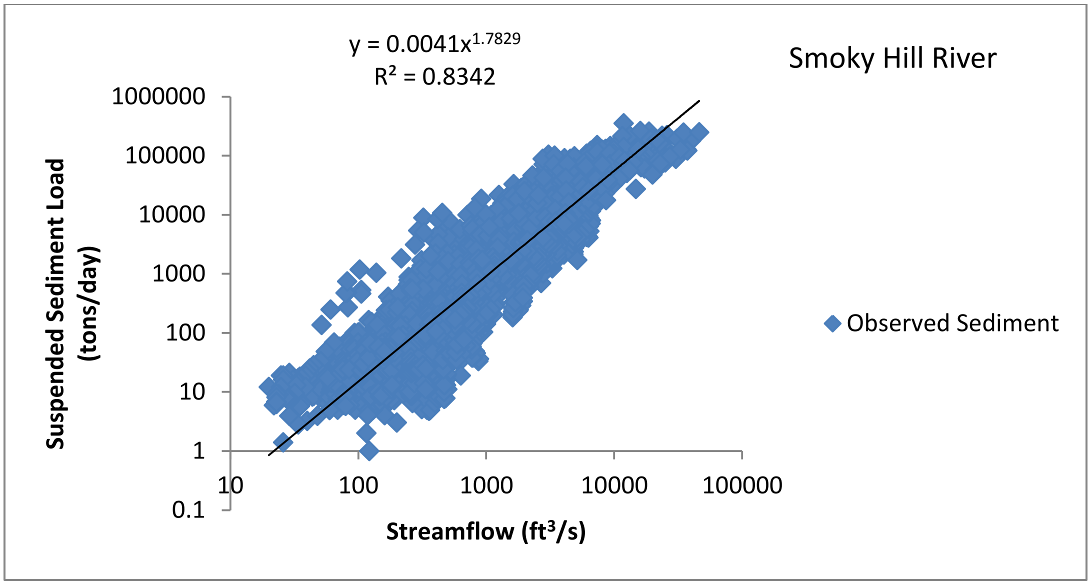

Figure 16. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Smoky Hill River.
The regression equation for this relationship is $Y = 0.0041x^{1.7829}$

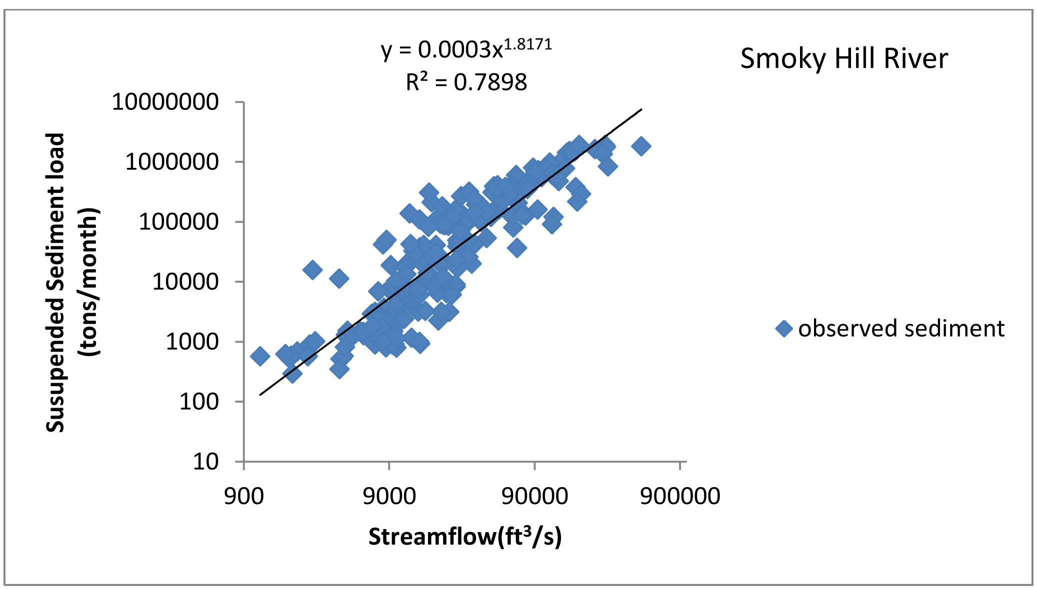

Figure 17. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft^3/s) of the Smoky Hill River.
The regression equation for this relationship is $Y = 0.0003X^{1.8171}$

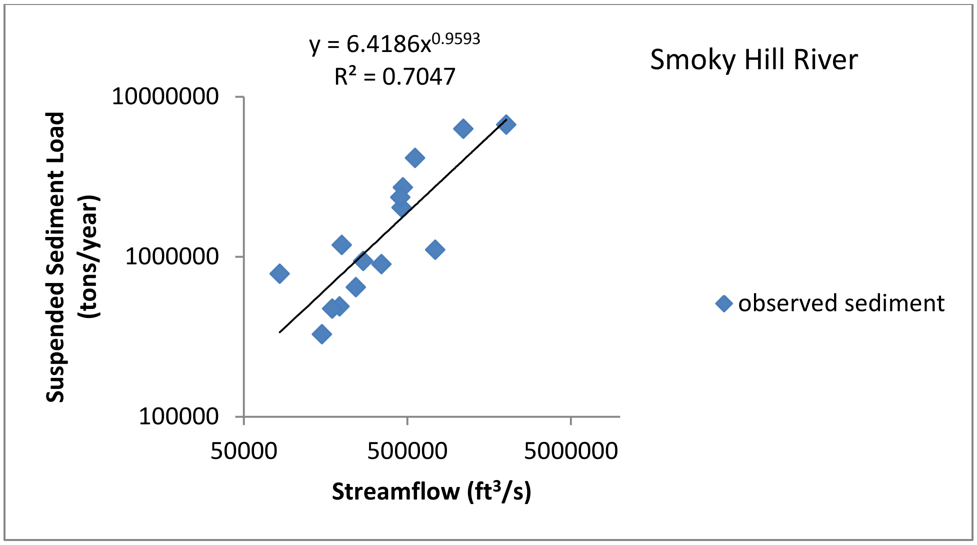

Figure 18. Regression relationship between annual suspended sediment load (tons/year) and streamflow (ft3/s) of the Smoky Hill River.
The regression equation for this relationship is $Y = 6.4186X^{0.9593}$

4.1.4 Sediment Record Length: 10 years

4.1.4.1 Cuyahoga River at Old Portage, Ohio

Figures 19, 20 and 21 present the regression relationships between the suspended sediment load and streamflow of the Old Portage gaging station on the Cuyahoga River. Figure 19 shows a good correlation between the daily suspended sediment load and streamflow, $r = 0.85$. The monthly and annual relationships also show a good correlation between the suspended sediment and streamflow, $r = 0.80$ and 0.93 respectively. The annual relationship shows better correlation between the suspended sediment and streamflow than the daily and monthly relationships probably due to its short 10 year sediment record length.

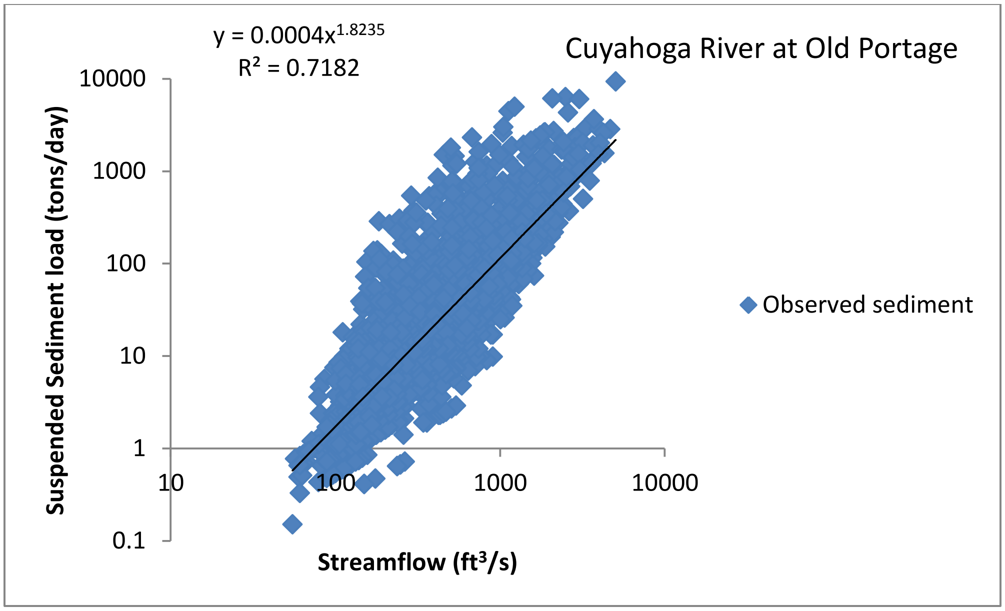

Figure 19. Regression relationship between daily suspended sediment load (tons/day) and streamflow (ft^3/s) of the Cuyahoga River at Old Portage.
The regression equation for this relationship is $Y = 0.0004X^{1.8235}$

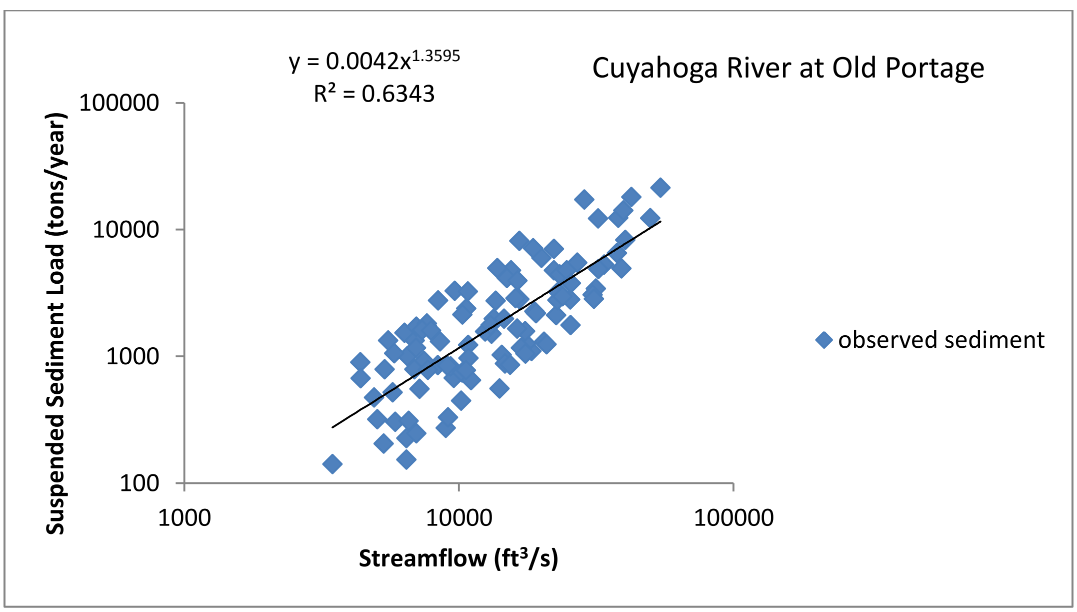

Figure 20. Regression relationship between monthly suspended sediment load (tons/month) and streamflow (ft^3/s) of the Cuyahoga River at old Portage.
The regression equation for this relationship is $Y = 0.0042X^{1.3595}$

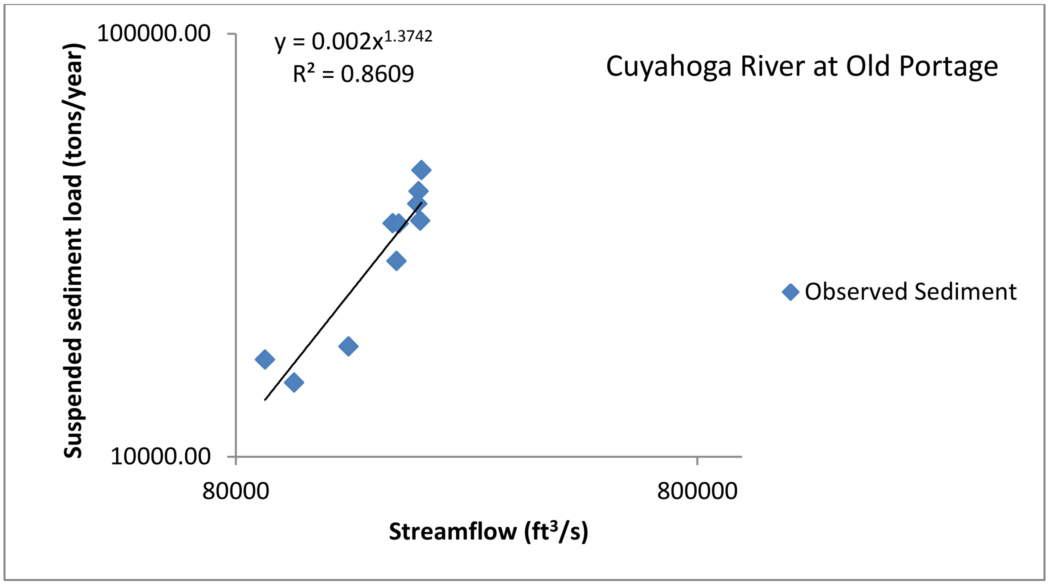

Figure 21. Regression relationship between annual suspended sediment load (tons/year) and streamflow (ft^3/s) of the Cuyahoga River at Old Portage.
The regression equation for this relationship is $Y = 0.002X^{1.3742}$

4.2 Discussion of Linear Regression

4.2.1 Deviation of the predicted sediment load from the observed values

Figures 22 to 28 show comparison between the annual observed sediment and the corrected predicted sediment for each of the investigated rivers. The percent deviation is the discrepancy between the predicted sediment and the observed sediment. The lower the percent deviation, the better the accuracy of the prediction. The annual percent deviation of the Maumee River is 38% and is illustrated in Figure 22. The annual percent deviation of the Sacramento River is also 38% and shown in Figure 23. For the Cuyahoga River at Independence, Ohio the percent deviation is 128% and is illustrated in Figure 24. The Mississippi River has 39% deviation as illustrated in Figure 25. The annual percent deviations for the Kankakee River, Smoky Hill River and the Cuyahoga River at old Portage are 31%, 53%, 13% and are shown in figures 26, 27 and 28 respectively.

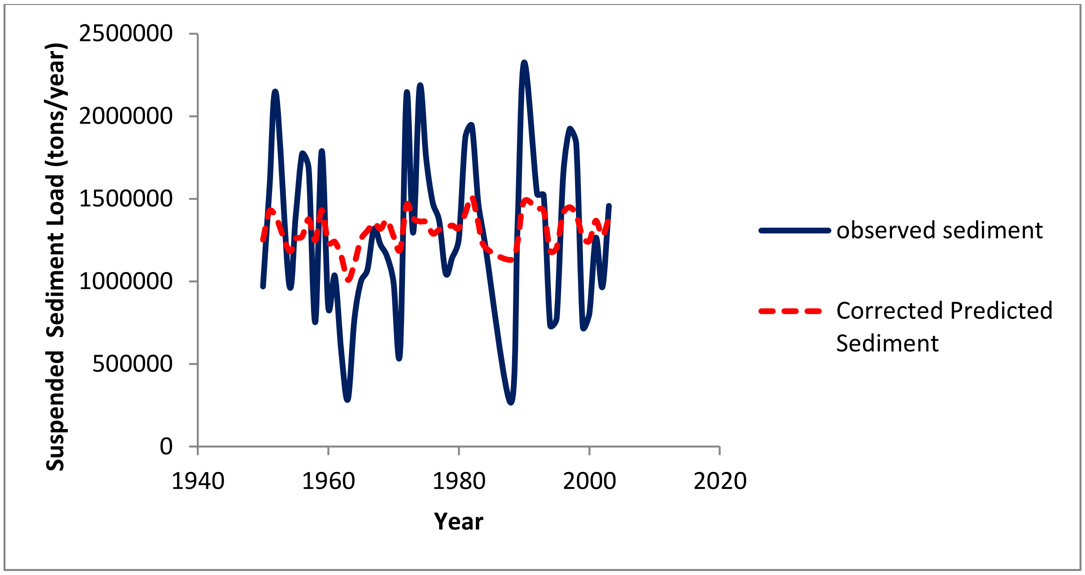

Figure 22. Comparison between the annual observed sediment and the corrected predicted sediment of the Maumee River at Waterville, Ohio

.

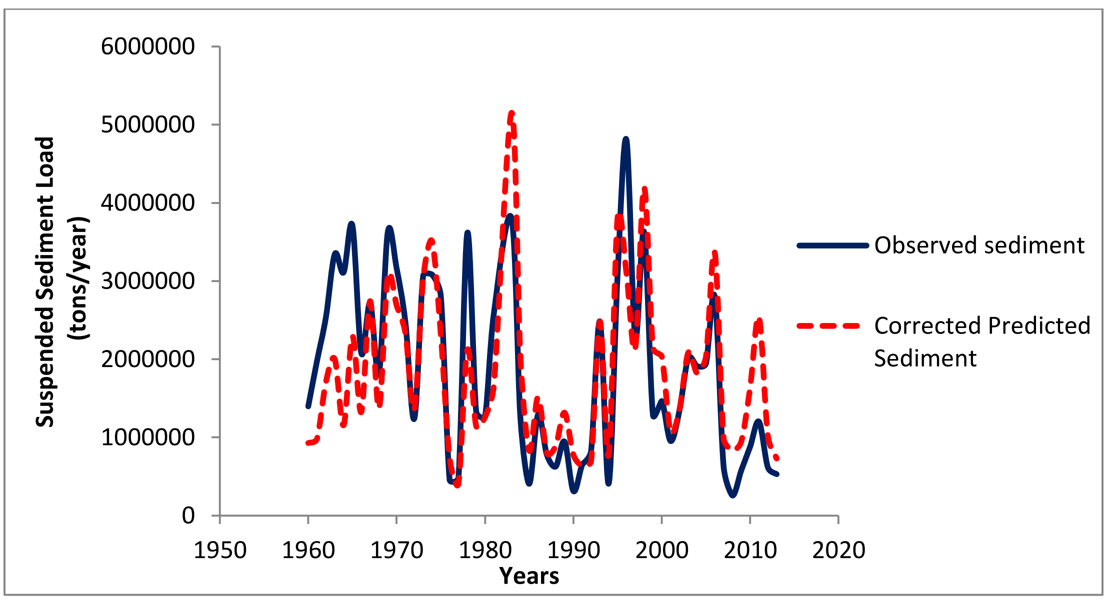

Figure 23. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Sacramento River near Red Bluff, California.

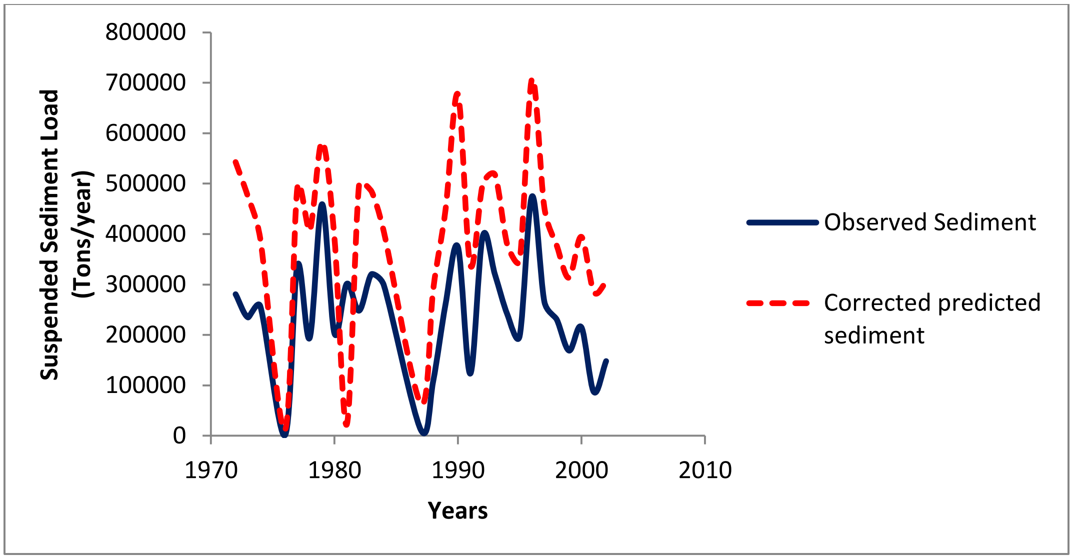

Figure 24. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Cuyahoga River at Independence, Ohio.

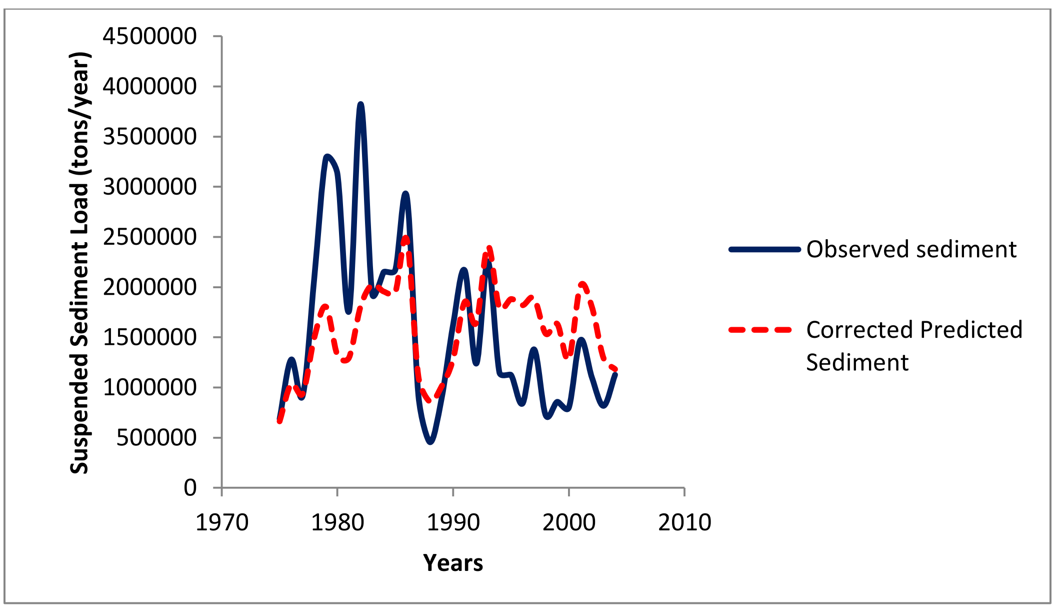

Figure 25. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Mississippi River at McGregor, Iowa

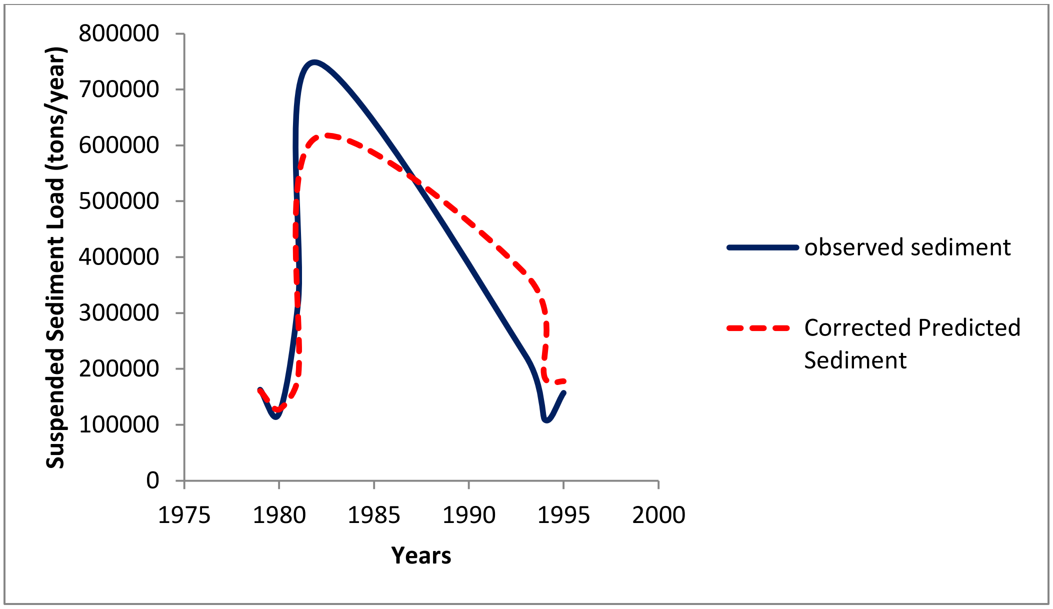

Figure 26. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Kankakee River at Momence, Illinois.

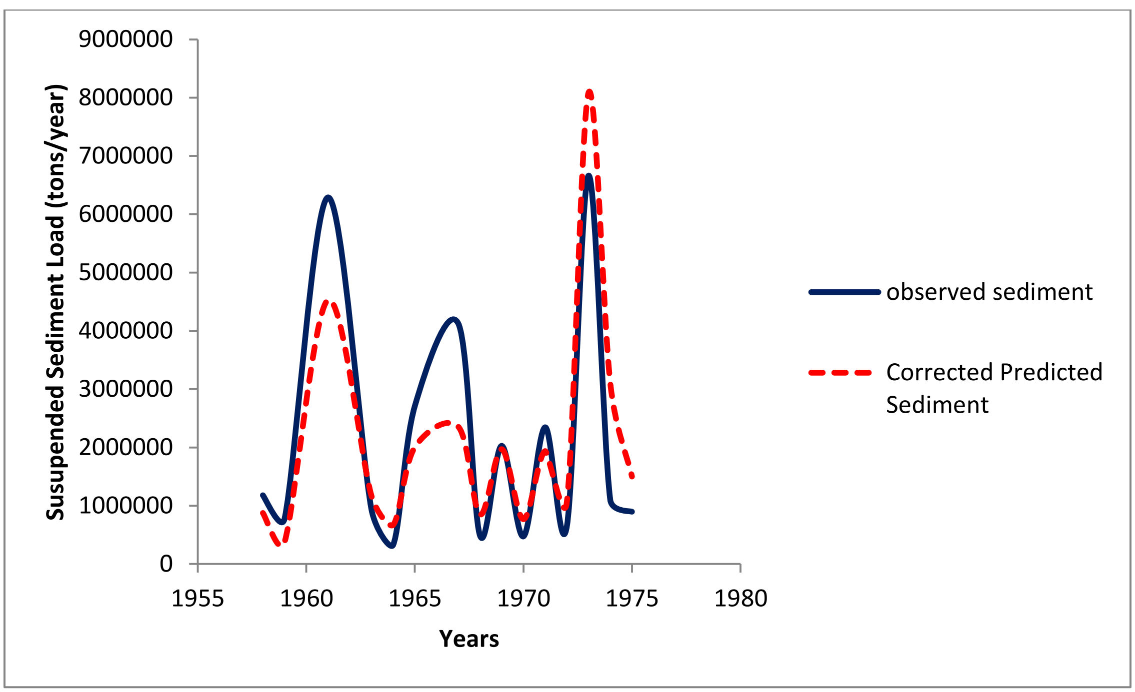

Figure 27. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Smoky Hill River at Lindsborg, Kansas.

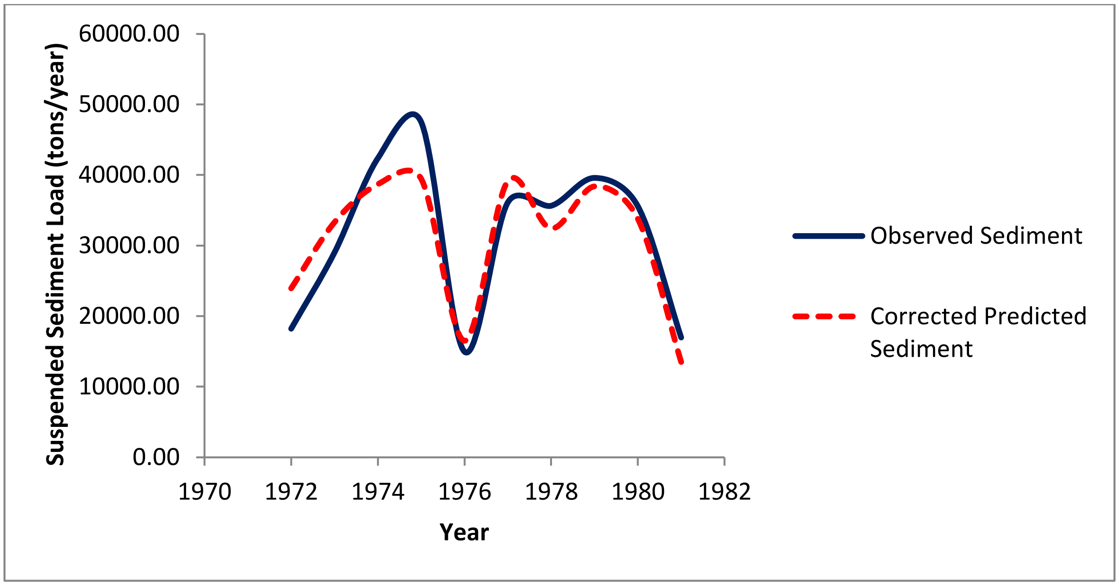

Figure 28. Comparison between the annual observed sediment and the annual corrected predicted sediment of the Cuyahoga River at Old Portage, Ohio

4.2.2 Prediction of the suspended sediment load

Prediction of the suspended sediment load is based on the annual data that is the sum of the monthly suspended sediment load of each year. Appendices A-G show the comparison of the observed and predicted sediment loads of the investigated rivers and is also displayed in figures 22 to 28 above. The corrected predicted sediment load and the observed sediment load of the seven rivers investigated follow almost the same trend, except that in some years the corrected predicted sediment is either a little higher or lower than the observed sediment load. The exception is the Cuyahoga River at Independence Ohio, with its corrected predicted sediment load appearing to be much different compared to the observed sediment. This can also be seen in its percent deviation value of 128%. The Kankakee River at Momence, Illinois did not produce a wavy trend graph like the other rivers because it has some data missing between the years.

Table 1. Assessment of river sediment record length, percent deviation and correlation coefficient

River		Sediment record length, Years	Percent Deviation, %	Correlation coefficient, r
1	Maumee River at Waterville, Ohio	54	38	0.86
2	Sacramento River near Red bluff, California	54	38	0.83
3	Cuyahoga River at Independence, Ohio	37	128	0.78
4	Mississippi River at McGregor, Iowa	30	39	0.59
5	Kankakee River at Momence, Illinois	17	31	0.83
6	Smoky Hill River at Lindsborg, Kansas	18	53	0.84
7	Cuyahoga River at Old Portage, Ohio	10	13	0.93

4.2.3 Relationship between the percent deviation and the sediment record length

Table 1 provides an assessment summary of the relationships between the sediment record length, percent deviation and correlation coefficient. As mentioned earlier, the accuracy of the prediction will be determined by analyzing the percent deviation and that the lower the percent deviation, the more accurate the prediction. Comparing the percent deviations of the seven investigated rivers, Cuyahoga River at Old Portage, Ohio has the lowest percent deviation, 13% and the Cuyahoga River at independence Ohio has the highest percent deviation, 128%. This very high percent deviation indicates an inaccurate prediction of the suspended sediment load at that gaging station on the Cuyahoga River relative to the Mississippi River at McGregor, Iowa (39%)

both of which were investigated for a record length of 30 years and above. The Cuyahoga River at old Portage, Ohio therefore can be said to have produced the best prediction for suspended sediment loads, maybe because of the shorter sediment record length which means less hydrologic variations. Maumee River at Waterville, Ohio and the Sacramento River near Red Bluff, California were both investigated for sediment record length of 54 years and interestingly produced the same percent deviation, 38%. Smoky Hill River at Lindsborg, Kansas produced a somewhat higher percent deviation than the Kankakee River at Momence, Illinois with the same sediment record length of 15 years and over.

As mentioned earlier, past studies have mentioned that in order to get good results using regression analysis, sediment data should preferably be collected on a daily or weekly basis over a period between 10 to 20 years. This is found to be the case for all the investigated rivers except the Cuyahoga River at Independence, Ohio and the Smoky Hill River at Lindsborg, Kansas. This can be explained by the fact that prediction of the suspended sediment load as a function of streamflow by linear regression assumes that all the suspended sediment loads come from erosion of the river or stream channel by the streamflow. In reality, however, some of the sediments are added to the channel by sediment sources while some of the sediments are removed from the channel by sediment sinks. Therefore, the quantity of the suspended sediment load eroded from the channel by the streamflow can be increased by sediment sources and/or decreased by sediment sinks. The effects of sediment sources and sinks are probably the cause of the high percent deviations of the Cuyahoga River at Independence, Ohio and the Smoky Hill River at Lindsborg, Kansas. It is possible that the other five rivers are also affected to some extent by sediment sources and sinks, however, the most profound effects appear to be at the Cuyahoga River and the Smoky Hill River.

4.2.4 Relationship between the correlation coefficient and the sediment record length

The higher the correlation coefficient the stronger the correlation between the suspended sediment and streamflow. It can be seen in table 1 that the Cuyahoga River at Old Portage, Ohio with the shortest sediment record length of 10 years has the highest correlation coefficient, 0.93 compared to the other rivers, and therefore is said to exhibit the strongest correlation, and as mentioned earlier this could be due to its short length of data which means less hydrologic variations. The Maumee River at Waterville, Ohio produced a correlation coefficient of 0.86 which is close to 0.83 produced by the Sacramento River near Red Bluff, California with the same sediment record length, 54 years. The Kankakee River at Momence, Illinois was investigated for a record length of 17 years and produced a correlation coefficient of 0.83 which is quite close to 0.84 for the Smoky Hill River at Lindsborg, Kansas with sediment record length of 18years. There is quite a discrepancy between the correlation coefficients 0.78 and 0.59 produced by the Cuyahoga River at Independence, Ohio and the Mississippi at McGregor, Iowa probably because they were subjected to intense hydrologic changes compared to the other rivers.

4.2.5 Relationship between the percent deviation and the correlation coefficient

According to table 1 the Cuyahoga River at old Portage, Ohio produced the most accurate prediction for the suspended sediment load with the lowest percent deviation, 13% and also the highest correlation coefficient, 0.93. The Maumee River at Waterville, Ohio and the Sacramento River near Red Bluff, California both produced the same percent deviation of 38% and have close correlation coefficients of 0.86 and 0.83, relatively. The Kankakee River at Momence, Illinois and the Smoky Hill River at Lindsborg, Kansas have correlation coefficients of 0.83 and 0.84, both of which are quite close to the coefficient correlations of the Maumee river at McGregor, Ohio and the Sacramento River near Red Bluff, California even though they produced different percent deviations of 31% and 53%, respectively. The Cuyahoga River at Independence has a percent deviation of 128% and a correlation coefficient of 0.78. The Mississippi River at McGregor, Iowa has the lowest correlation coefficient of 0.59 and produced a percent deviation of 39%. The low correlation coefficients of the Cuyahoga River at Independence, Ohio and the Mississippi River at McGregor, Iowa are probably due to their intense hydrologic changes, as explained above.

4.3 Results of Pearson Correlation

4.3.1 Maumee River at Waterville, Ohio

Table 2. Correlation matrix between selected physical and biological parameters of the Maumee River at Waterville, Ohio

		Site	Suspended sediment	Nitrate	Total Nitrogen	Flow	Total Phosphorus
Site	Pearson Correlation	1	.141	-.061	-.083	.351*	.186
Suspended sediment	Pearson Correlation	.141	1	-.212	-.200	.573**	-.043
Nitrate	Pearson Correlation	-.061	-.212	1	.844**	-.202	.395**
Total Nitrogen	Pearson Correlation	-.083	-.200	.844**	1	-.234	.295*
Flow	Pearson Correlation	.351*	.573**	-.202	-.234	1	-.063
Total Phosphorus	Pearson Correlation	.186	-.043	.395**	.295*	-.063	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 2 shows the correlation coefficients between the parameters studied including suspended sediment, nitrates, total nitrogen, flow and total phosphorus. Suspended sediment showed a strong significant correlation of 57.3% with flow. Nitrates also showed a strong correlation with total phosphorus and total nitrogen of 39.5% and 84.4% respectively, while total phosphorus and total nitrogen had a weak but significant correlation of 29.5%. Interestingly, there was no significant correlation between suspended sediment load and nitrates at this gaging station nor was there any significant relationship between total phosphorus and suspended sediment.

4.3.1.1 Principal Component Analysis (PCA)

Data reduction analysis using Principal Component Analysis (PCA) was applied to further explain the relationship between the variables. The total variance of the parameters studied is shown in Table 3. The data reduction analysis revealed two components which explain 65% of the data variance. The first principal component explained about 39% of the data variance and the second principal component explains 26% of the data variance.

Table 3. Total Variance Explained of parameters measured in the Maumee River at Waterville Ohio

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.339	38.976	38.976	2.339	38.976	38.976
2	1.543	25.722	64.699	1.543	25.722	64.699
3	.922	15.371	80.070			
4	.673	11.213	91.283			
5	.377	6.283	97.566			
6	.146	2.434	100.000			

Extraction Method: Principal Component Analysis.

4.3.2 Zoar Valley at Cattaraugus Creek, New York

Different parameters were used for the correlation analysis of the Zoar Valley at Cattaraugus creek. The results are shown in Table 4 and the parameters chosen include suspended sediment, streamflow, nitrates, turbidity and fecal coliform. In this case, suspended sediment and fecal coliform, suspended sediment and turbidity, turbidity and fecal coliform, nitrate and flow showed strong significant relationships with correlations of 81%, 92%, 85% and 86% respectively. There was not sufficient data to establish a relationship between nitrates and suspended sediment but to an extent, nitrates and turbidity appear to be correlated (41%).

Table 4. Correlation matrix between selected physical and biological parameters of the Zoar Valley at Cattaraugus Creek, New York

		Year	Suspended sediment	Turbidity	Fecal coliform	Flow	Nitrate
Year	Pearson Correlation	1	.264**	.216*	.224*	-.682**	.[c]
Suspended sediment	Pearson Correlation	.264**	1	.915**	.811**	.143	.[c]
Turbidity	Pearson Correlation	.216*	.915**	1	.850**	.187*	.407
Fecal coliform	Pearson Correlation	.224*	.811**	.850**	1	.110	.522*
Flow	Pearson Correlation	-.682**	.143	.187*	.110	1	.864**
Nitrate	Pearson Correlation	.[c]	.[c]	.407	.522*	.864**	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
c. Cannot be computed because at least one of the variables is constant.

4.3.2.1 Principal Component Analysis (PCA)

Since the majority of environmental variables were found to be related to a significant degree, a data reduction analysis by principal component analysis (PCA) was performed. The total variances of the parameters studied are shown in table 5. The data reduction analysis revealed the measured variables could be explained by two factors representing 89% of the data variance. The first principal component explained about 57% of the data variance and the second principal component explains 32% of the data variance.

Table 5. Total Variance Explained of parameters measured in the Zoar Valley at Cattaraugus, New York

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.854	57.075	57.075	2.854	57.075	57.075
2	1.599	31.987	89.062	1.599	31.987	89.062
3	.277	5.535	94.597			
4	.195	3.902	98.499			
5	.075	1.501	100.000			

Extraction Method: Principal Component Analysis.

4.4 Discussion of Pearson correlation

4.4.1 Maumee River at Waterville, Ohio

Table 6. Component Matrix of the studied variables of the Maumee River at Waterville, Ohio

Component Matrix-a		
	Component	
	1	2
Site	-0.244	0.625
SS	-0.555	0.513
it	0.848	0.365
TN	0.836	0.307
Flow	-0.596	0.615
TP	0.446	0.533
Extraction Method: Principal Component Analysis.		
a - 2 components extracted.		

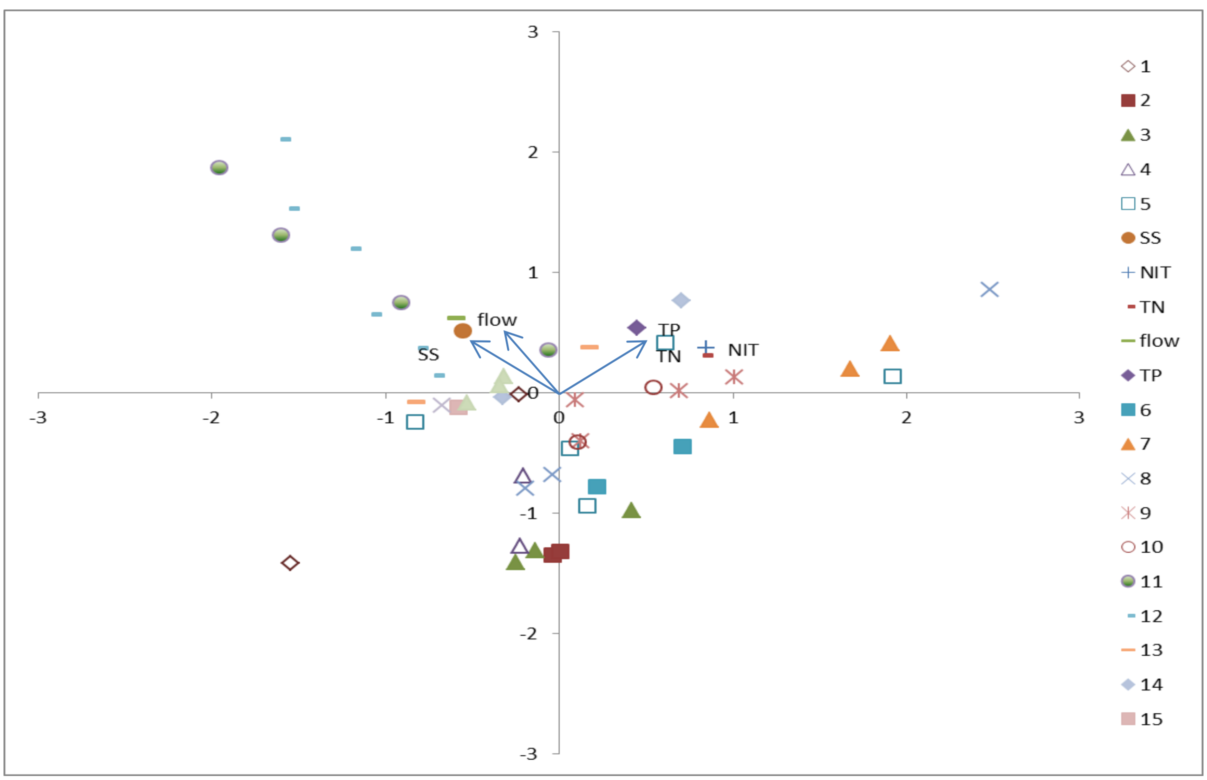

Figure 29. Maumee River at Waterville, Ohio PCA graph display.
The numbers represent the different sampling stations.

Table 6 shows the component matrix of the studied variables of the Maumee River at Waterville Ohio generated by PCA and displayed in Figure 29. The first component is strongly represented by nitrates and total nitrogen while the second principal component is equally represented by all the variables. Figure 29 shows a scattered plot of the data probably due to the heterogeneous nature and vast hydrologic variations of the river, the type of sampling method used and the time of the year and also due to the small size of the data used. However, at station 11 and 12, which had the highest, number of sampling events, it is clear that flow was the parameter that influenced the distribution of the data. Station 9 was influenced by nitrates, total nitrogen and phosphorus. No clustering is seen in the rest of the stations probably because of the few data points, for example, stations 1 and14 have only one data point and stations 4 and15 have two data points, stations 2, 3 and 7 each have three data points which do not say much about the stations.

4.4.2 Zoar Valley at Cattaraugus Creek, New York

Table 7.Component matrix of the studied variables of the Zoar Valley at Cattaraugus creek, New York

Component Matrix -a		
	Component	
	1	2
ss	0.942	0.143
turb	0.964	0.116
FC	0.922	0.078
flow	0.038	0.938
year	0.43	-0.824
Extraction Method: Principal Component Analysis.		
a - 2 components extracted.		

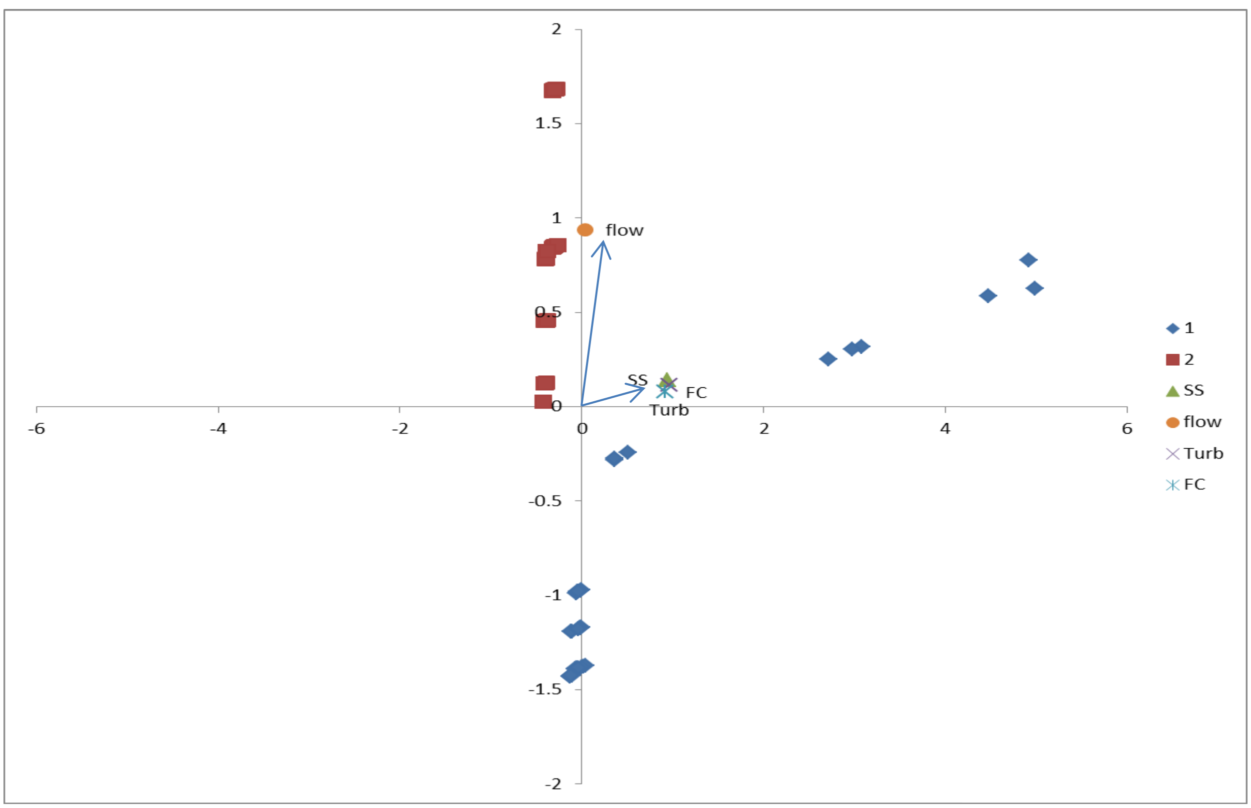

Figure 30. Zoar Valley of Cattaraugus creek, New York PCA graph display.
The numbers represent the different sampling years.

Table 7 shows the component matrix of the studied variables of the Zoar valley at Cattaraugus Creek, New York generated by PCA and shown in Figure 30. Flow is the main parameter greatly influencing the other parameters for principal component two because it has the highest coefficient of 0.93 compared to suspended sediment (0.14), turbidity (0.12) and fecal coliform (0.078), while for principal component one, the loadings of suspended sediment, turbidity and fecal coliform were tightly correlated as their coefficients are close to each other. The data points are clustered, unlike the Maumee River at Waterville Ohio, possibly because the physical conditions of the Zoar Valley at Cattaraugus creek, New York are more homogenous and it experiences less hydrologic variations, the data size per sampling is larger and sampling was done over a period of two years unlike the for the Maumee River at Waterville, Ohio which was all done in the same year and with less data collected per sampling station.

CHAPTER 5. Conclusions and Recommendations

5.1 Conclusions

1) Sediment sources/sinks and hydrologic variations that take place throughout the year play a vital role in the regression analysis of suspended sediment data. Each river or stream is subjected to unique sediment sources/sinks and hydrologic changes.

2) Five of the seven investigated rivers produced good predictions (with percent deviations ranging from 13% to 39%) of the suspended sediment loads using linear regression

3) The sediment record length can affect the value of the correlation coefficient between the streamflow and suspended sediment rate. Six of the seven investigated rivers yielded high correlation coefficients ranging from 0.78 to 0.93. This confirms that sediment and streamflow data should be collected for a period of at least 10 years (the shortest period investigated) in order to obtain accurate results from linear regression.

4) Four of the investigated rivers with sediment record length varying from 17 to 54 years resulted in approximately similar percent deviations ranging from 31% to 39%. Therefore, the percent deviation between predicted and suspended sediment is less likely affected by the sediment record length and most likely affected by the hydrologic variations and sediment sources/sinks.

5) Two of the investigated rivers show that significant hydrologic variations and sediment sources/sinks can increase the percent deviation, regardless of the sediment record length. Conversely, the lack of significant hydrologic variations and sediment sources/sinks, due to shorter sediment record lengths (10-20 years), can decrease the percent deviation.

6) The Pearson correlation analysis showed that there was no correlation between the phosphorus nor nitrates and the suspended sediment load of the Maumee River at Waterville, Ohio; meanwhile there was a strong correlation between fecal coliform and suspended sediment loads and turbidity of the Zoar Valley at Cattaraugus, New York, which therefore effects the transportation of fecal coliform in the stream.

7) The Maumee River at Waterville, Ohio and the Zoar Valley at Cattaraugus, New York rivers have different physical conditions but the principal component analysis revealed that they are both influenced by flow, which greatly determines the other parameters. It can therefore be said that water movements from land into the rivers have a greater potential of carrying along with it harmful pollutants.

5.2 Recommendations

1. Further studies should be done to include the effect of sources and sinks in the prediction of the sediment load in the streams.
2. The seven rivers could be investigated for correlations between the suspended sediment and fecal coliform, and suspended sediment and nutrients for broader comparison purposes.

References

American Rivers. 2014. America runs on the Mississippi River. Available at: http://www.americanrivers.org/rivers/fun/america-runs-on-the-mississippi-river/

American Rivers. 2002. The Ecology of Dam Removal: A Summary of Benefits and Impacts.

Amin IE, Jacobs AM. 2007. Accounting for sediment sources and sinks in the regression analysis of the suspended sediment load of streams: the Rio Puerco, New Mexico as an example. *Environmental Geosciences*. 14(1):1-14.

Araujo HA, Cooper AB, Hassan MA, Venditti J. 2012. Estimating suspended sediment concentrations in areas with limited hydrological data using a mixed-effects model. *Hydrological Processes*. 26: 3678–3688.

Basto Salgado GP. 2005. Assessment of water quality, benthic community structure and microbial indicators in Cattaraugus creek, Zoar Valley, NY. Masters Thesis. Youngstown State University.

Berkun M, Aras E, Akdemir UO. 2015. Water Runoff, Sediment Transport and Related Impacts in the Southeastern Black Sea Rivers. *Environmental engineering and Management.* 14(4):781-791.

Bhowmik NG, Bonini AP, Bogner WC, Bryne RP.1980. Hydraulics of flow and sediment transort in the Kankakee River in Illinois: Illinois State Water Survey, Report of Investigation 98p.70-94

Botkin DB, Keller EA. 2005. Environmental Science: Earth as a living planet, 5th Ed. New York. *John wiley and Sons Inc.* 664p.

City of Momence. 2015. Available at: http://www.cityofmomence.com/demographics.aspx

Cohn TA, Gilroy EJ. 1991. Estimating loads from periodic records: U. S. Geological survey branch of systems analysis technical memo. 91.01. 81p.

Cuyahoga Valley National Park Ohio. 2013. National Park Service. Available at: http://www.nps.gov/cuva/index.htm

Environmental Protection Agency. 2013. Cuyahoga River. Available at: http://www.epa.gov/greatlakes/aoc/cuyahoga/

Ferguson RI. 1987. Accuracy and precision of methods for estimating river loads. *Earth Surface Processes and Landforms.*12: 95-104.

Gray JR, Simoes FJM. 2008. Estimating Sediment Discharge: Appendix D in Sedimentation engineering: processes, measurements, modeling, and practice. *American Society of Civil Engineers.* 22p. (1065-1086).

Hansen D, Bray D. 1987. Generation of annual suspended sediment loads for the Kennebacasis using sediment rating curves. In: Proc. 8th Canadian Hydrotechnical Conference, Canadian Society of Civil Engineers (Quebec, May 1987).

Ivens JL, Bhowmik NG, Allison RB, Gross DL. 1981. The Kankakee River Yesterday and Today. Illinois Department of Energy and Natural Resources. Champaign Illinois.

Jain SK. 2001. Development of integrated sediment rating curves using ANNs. *Hydrologic Engineering*. 127(1):30-37.

Jie LC, Yu ST. 2011. Suspended sediment load estimate using support vector machines. In Kaoping River basin. 978-1-61284-459-6/11 2011 IEEE.

Jones BL, Hawley NL, Crippen JR. 1972. Sediment Transport in the Western Tributaries of the Sacramento River, California. *Geological survey water-supply paper* 1798-j.

Kansas Water Office. 2009. Smoky Hill Saline River Basin. Available at: http://www.kwo.org/kansas_water_plan/KWP_Docs/VolumeIII/SHS/Rpt_SHS_Entire_Basin_Section_KWP_2009.pdf

Kansas Water Office. 2011. Smoky Hill Saline River Basin. Available at: http://www.kwo.org/about_us/BACs/Smoky_Hill_Saline_River_Basin.html

Lamba J, Karthikeyan KG, Thompson AM . 2015. Apportionment of suspended sediment sources in an agricultural watershed using sediment fingerprinting. *Geoderma* 239–240: 25–33.

Laerd Statistics. 2013. Pearson product –moment correlation. Available at: https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php

Laubel A, Svendsen LM, Kronvang B, Larsen SE. 1999. Bank erosion in a Danish lowland stream system. *Hydrobiologia* 410:279–285.

Le C, Zha Y, Li Y, Sun D, Lu H, Yin B. 2010. Eutrophication of lake waters in China: cost, causes, and control, *Environmental Management.* (45):662–668.

Leopold LB, Maddock T. 1953. The hydraulic geometry of stream channels and some physiographic implications. *Geological Survey Professional Paper*. 252-A: 21.

Leopold LB, Miller JP. 1956. Hydraulic factors and their relation to the drainage net. *Geological Survey Professional Paper*. 252-A: 11-13.

Madsen B L. 1995. Danish watercourses – ten years with the new watercourse act. Ministry of Environment and Energy, Denmark. Danish Environmental Protection Agency, Copenhagen.

McGregor Weather. 2015. Available at: http://www.idcide.com/weather/ia/mcgregor.htm.

Merritt WS, Letcher RA, Jakeman AJ. 2003. A Review of Erosion and Sediment Transport Models. *Environmental Modelling & Software*. 18(8-9):761–799.

Mount J. n.d. Sacramento River Basin. Available at: http://www.sacriver.org/files/documents/roadmap/report/SacRiverBasin.pdf

Myint SW, Walker ND. 2002. Quantification of surface suspended sediments along a river dominated coast with NOAA AVHRR and SeaWiFS Measurements: Louisiana USA. *International Journal of Remote Sensing*. 23(16):3229-3249.

Ohio Environmental Protection Agency. n.d. The Maumee River Watershed. Available at: http://www.epa.state.oh.us/dsw/tmdl/MaumeeRiver.aspx

Ohio History Central. n.d. Maumee River. Available at: http://www.ohiohistorycentral.org/w/Maumee_River?rec=2164

Ongley ED. 1996a. Control of water pollution from agriculture. Food and agriculture Organization of the United Nations, irrigation and drainage paper 55.

Ongley ED. 1996b. Sediment Measurements. In Bartram J, Ballance R, Editors. Water Quality Monitoring - A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes. UNEP/WHO.

Palmer PB, O'Connell DG. 2009. Regression Analysis for Prediction: Understanding the Process. *Cardiopulmonary Physical Therapy*. 20(3): 23–26.

Russell MA, Walling DE, Hodgkinson RA. 2001. Suspended sediment sources in two small lowland agricultural catchments in the UK. *Journal of Hydrology.* 252(1-4):1-24.

Svendsen LM, Kronvang B, Kristensen P, Græsbøll P. 1995. Dynamics of phosphorus compounds in a lowland river system: importance of retention and non-point sources. *Hydrological Processes*. 9:119–142.

Tayfur G. 2003. Modelling sediment transport. In Watershed hydrology. V. P. Singh and R. N. Yadava, eds. *Allied*. 560p: 353–375.

The Kankakee River. 2010. Overview and some History. Available at:

http://www.thekankakeeriver.com/kankakee02.html

Trimble SW. 1999. Decreased rates of alluvial sediment storage in the Coon Creek Basin, Wisconsin, 1975 - 93. *Science* 285 (5431): 1244–1246.

Troeh FR, Hobbs JA, Donahue RL. 2004. Soil and Water Conservation for Productivity and Environmental Protection, 4th Ed. Upper Saddle River, New Jersey. *Prentice Hall.* 656p.

US Climate data. 2015. Climate Red Bluff, California. Available at:

http://www.usclimatedata.com/climate/red-bluff/california/united-states/usca0919.

United States Army Corps of Engineers. 2006. Mahoning River Ohio Environmental Dredging; Draft Feasibility Report and Environmental Impact Statement. Available at:

http://www.villageoflowellville.com/wp-content/uploads/2013/07/MAHONING-MAIN-REPORT-EIS.pdf

United States Environmental Protection Agency. 2006. Wadeable streams assessment: a collaborative survey of the nation's streams, Washington, D.C., 98 pp.

United States Geological Survey Sediment Portal. Available

at: http://cida.usgs.gov/sediment/.

Van Rijn LC. 1984. Sediment Transport Part 1: Bed load Transport. *Hydraulic Engineering*. 110(10):1431-1456.

Walling DE. 1977. Limitations of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers. In Erosion and solid matter transport in inland waters (Proceedings of the Paris symposium, July 1977). *IAHS* Wallingford, Oxfordsire. 122: 34–48.

Waterville, OH Weather. 2015. Available at:

http://www.idcide.com/weather/oh/waterville.htm

Yang CT. 1996. Sediment transport theory and practice. *McGraw-Hill*, New York.

Appendices

Appendix A. Comparison of the observed and predicted suspended sediment loads of the Maumee River at Waterville, Ohio

Year	observed sediment, Million tons	Corrected predicted sediment load, Million tons
1950	1.0	1.3
1951	1.6	1.4
1952	2.1	1.4
1954	1.0	1.2
1955	1.4	1.3
1956	1.8	1.3
1957	1.7	1.4
1958	0.8	1.2
1959	1.8	1.4
1960	0.8	1.2
1961	1.0	1.2
1962	0.6	1.2
1963	0.3	1.0
1964	0.8	1.1
1965	1.0	1.3
1966	1.1	1.3
1967	1.3	1.4
1968	1.2	1.3
1969	1.2	1.4
1970	1.0	1.3
1971	0.6	1.2
1972	2.1	1.5
1973	1.3	1.4
1974	2.2	1.4
1975	1.7	1.4
1976	1.5	1.3
1977	1.4	1.3
1978	1.0	1.3
1979	1.1	1.3
1980	1.3	1.3
1981	1.9	1.4
1982	1.9	1.5
1983	1.5	1.4
1984	1.2	1.2
1988	0.3	1.1
1989	1.3	1.2

1990	2.3	1.5
1992	1.5	1.4
1993	1.5	1.4
1994	0.7	1.2
1995	0.8	1.2
1996	1.6	1.4
1997	1.9	1.4
1998	1.8	1.4
1999	0.7	1.3
2000	0.8	1.2
2001	1.3	1.4
2002	1.0	1.3
2003	1.5	1.4

Appendix B. Comparison of the observed and predicted suspended sediment loads of the Sacramento River near Red Bluff, California

Year	Observed sediment, Million tons	Corrected Predicted sediment, Million tons
1960	1.4	0.9
1961	2.0	1.0
1962	2.5	1.7
1963	3.3	2.0
1964	3.1	1.2
1965	3.7	2.3
1966	2.1	1.3
1967	2.7	2.7
1968	1.8	1.4
1969	3.6	3.1
1970	3.2	2.7
1971	2.5	2.3
1972	1.2	1.4
1973	3.1	3.0
1974	3.1	3.5
1975	2.8	2.3
1976	0.5	0.7
1977	0.5	0.4
1978	3.6	2.1
1979	1.3	1.1
1980	1.3	1.3
1981	2.6	1.7
1983	3.8	5.2
1984	1.3	2.0
1985	0.4	0.8
1986	1.3	1.5
1987	0.8	0.8
1988	0.6	0.9
1989	0.9	1.3
1990	0.3	0.8
1991	0.6	0.6
1992	0.8	0.7
1993	2.5	2.5
1994	0.4	0.8
1995	3.1	3.8
1996	4.8	3.1
1997	2.3	2.1
1998	3.6	4.2
1999	1.3	2.1

2000	1.5	2.0
2001	1.0	1.1
2002	1.3	1.4
2003	2.0	2.1
2004	1.9	1.8
2005	1.9	2.0
2006	2.8	3.4
2007	0.6	1.0
2008	0.3	0.8
2009	0.6	0.9
2010	0.9	1.6
2011	1.2	2.5
2012	0.6	1.0
2013	0.5	0.7

Appendix C. Comparison of the observed and predicted suspended sediment loads of the Cuyahoga River at Independence, Ohio

Year	Observed Sediment, Million tons	Corrected predicted sediment, Million tons
1972	0.3	0.5
1973	0.2	0.5
1974	0.3	0.4
1976	0.0	0.0
1977	0.3	0.5
1978	0.2	0.4
1979	0.5	0.6
1980	0.2	0.4
1981	0.3	0.0
1982	0.2	0.5
1983	0.3	0.5
1984	0.3	0.4
1987	0.0	0.1
1988	0.1	0.3
1989	0.3	0.4
1990	0.4	0.7
1991	0.1	0.3
1992	0.4	0.5
1993	0.3	0.5
1994	0.2	0.4
1995	0.2	0.3
1996	0.5	0.7
1997	0.3	0.5
1998	0.2	0.4
1999	0.2	0.3
2000	0.2	0.4
2001	0.1	0.3
2002	0.1	0.3

Appendix D. Comparison of the observed and predicted suspended sediment loads of the Mississippi River at McGregor, Iowa

Year	Observed sediment, Million tons	Corrected predicted sediment load, Million tons
1975	0.7	0.7
1976	1.3	1.0
1977	0.9	0.9
1978	2.1	1.5
1979	3.3	1.8
1980	3.1	1.3
1981	1.8	1.3
1982	3.8	1.8
1983	1.9	2.0
1984	2.2	2.0
1985	2.2	1.9
1986	2.9	2.5
1987	0.9	1.1
1988	0.5	0.9
1989	0.9	1.0
1990	1.6	1.3
1991	2.2	1.9
1992	1.2	1.6
1993	2.3	2.4
1994	1.1	1.8
1995	1.1	1.9
1996	0.7	1.6
1997	1.4	1.9
1998	0.7	1.5
1999	0.9	1.6
2000	0.8	1.3
2001	1.5	2.0
2002	1.1	1.8
2003	0.8	1.3
2004	1.1	1.2

Appendix E. Comparison of the Observed and Corrected Predicted suspended sediment loads of the Kankakee River at Momence, Illinois

Year	observed sediment, Million tons	Corrected Predicted Sediment, Million tons
1979	0.2	0.2
1980	0.1	0.1
1981	0.3	0.2
1982	0.7	0.6
1993	0.2	0.4
1994	0.1	0.2
1995	0.2	0.2

Appendix F. Comparison of the observed and predicted suspended sediment loads of the Smoky Hill River at Lindsborg, Kansas

Year	observed sediment, Million tons	Corrected predicted sediment load, Million tons
1958	1.2	0.9
1959	0.8	0.4
1961	6.3	4.5
1963	0.9	1.2
1964	0.3	0.7
1965	2.7	2.0
1967	4.1	2.4
1968	0.5	0.8
1969	2.0	2.0
1970	0.5	0.8
1971	2.3	1.9
1972	0.6	1.1
1973	6.7	8.1
1974	1.1	3.1
1975	0.9	1.5

Appendix G. Comparison of the observed and predicted suspended sediment loads of the Cuyahoga River at Old Portage, Ohio

Year	Observed Sediment, Million tons	Corrected Predicted Sediment, Million tons
1972	0.02	0.02
1973	0.03	0.03
1974	0.04	0.04
1975	0.05	0.04
1976	0.01	0.02
1977	0.04	0.04
1978	0.04	0.03
1979	0.04	0.04
1980	0.04	0.03
1981	0.02	0.01

Appendix H. Calculations of the Prediction of the Annual Suspended Sediment Loads of the Maumee River at Waterville, Ohio

Year	Streamflow	observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	10^logy (Predicted Sediment)	Residual	10^Residual	Corrected predicted sediment	Absolute % deviation
1950	1575065	969302	6.197298	5.986459	6.052468	1128412.075	-0.06601	0.858996	1252537	29.22055
1951	2534472	1570277	6.403887	6.195976	6.109259	1286053.529	0.086717	1.221004	1427519	9.091236
1952	2298969	2136639	6.361533	6.329731	6.097616	1252033.144	0.232115	1.706535	1389757	34.95594
1954	1289659	984164	6.110475	5.993067	6.0286	1068070.501	-0.03553	0.921441	1185558	20.46349
1955	1603575	1390638	6.205089	6.143214	6.054609	1133990.496	0.088605	1.226322	1258729	9.485441
1956	1666909	1773509	6.221912	6.248833	6.059234	1146130.187	0.189599	1.547389	1272205	28.26625
1957	2221056	1683600	6.34656	6.226239	6.0935	1240222.429	0.132739	1.357498	1376647	18.23195
1958	1556760	753135	6.192222	5.876873	6.051072	1124791.713	-0.1742	0.669577	1248519	65.77623
1959	2561094	1789129	6.408426	6.252642	6.110507	1289753.002	0.142135	1.387187	1431626	19.98197
1960	1486599	841845	6.172194	5.925232	6.045566	1110622.476	-0.12033	0.757994	1232791	46.43918
1961	1518243	1034877	6.181341	6.014889	6.048081	1117071.81	-0.03319	0.926419	1239950	19.81614
1962	1159364	571291	6.06422	5.756857	6.015884	1037252.281	-0.25903	0.550773	1151350	101.5348
1963	718714	285911	5.856556	5.456231	5.958798	909489.4079	-0.50257	0.314364	1009533	253.0935
1964	1004138	770233	6.001793	5.886622	5.998723	997064.7871	-0.1121	0.7725	1106742	43.68924
1965	1581893	998322	6.199177	5.999271	6.052984	1129754.704	-0.05371	0.883663	1254028	25.61355
1966	1832804	1071412	6.263116	6.029957	6.070561	1176416.121	-0.0406	0.910742	1305822	21.8786
1967	2089804	1317108	6.320106	6.119621	6.086227	1219628.01	0.033394	1.079926	1353787	2.784828
1968	1895625	1221887	6.277752	6.087031	6.074585	1187365.743	0.012447	1.029074	1317976	7.863947
1969	2185705	1151335	6.339592	6.061202	6.091584	1234764.368	-0.03038	0.932433	1370588	19.04341
1970	1673744	994430.5	6.223689	5.997574	6.059722	1147420.19	-0.06215	0.866666	1273636	28.07697
1971	1311775	574202.5	6.117859	5.759065	6.03063	1073074.582	-0.27156	0.5351	1191113	107.4378
1972	2776394	2136330	6.443481	6.329668	6.120143	1318691.862	0.209525	1.620037	1463748	31.48306
1973	2255686	1295050	6.353279	6.112287	6.095347	1245508.431	0.01694	1.039776	1382514	6.753744
1974	2141544	2178670	6.330727	6.338191	6.089147	1227855.402	0.249044	1.77437	1362919	37.44259
1975	2116230	1744803	6.325563	6.241746	6.087728	1223848.344	0.154019	1.425669	1358472	22.14183
1976	1751361	1478184	6.243376	6.169729	6.065134	1161807.954	0.104594	1.272314	1289607	12.75738
1977	1868098	1365493	6.2714	6.13529	6.072838	1182600.716	0.062451	1.154653	1312687	3.867197

1978	1927767	1047867	6.285055	6.020306	6.076592	1192866.565	-0.05629	0.878445	1324082	26.35969
1979	2003900	1144718	6.301876	6.058699	6.081216	1205635.694	-0.02252	0.949473	1338256	16.90701
1980	1915509	1255120	6.282284	6.098685	6.07583	1190776.617	0.022855	1.054035	1321762	5.309624
1981	2439569	1870296	6.387313	6.27191	6.104703	1272631.724	0.167208	1.469629	1412621	24.47071
1982	3122689	1942455	6.494529	6.288351	6.134176	1361997.534	0.154175	1.426181	1511817	22.16975
1983	2074953	1471259	6.317008	6.167689	6.085376	1217239.24	0.082313	1.208685	1351136	8.164664
1984	1382335	1233180	6.140613	6.091027	6.036885	1088641.71	0.054142	1.13277	1208392	2.010095
1988	1084327	266189.2	6.03516	5.42519	6.007896	1018347.335	-0.58271	0.261393	1130366	324.6474
1989	1504407	1334210	6.177365	6.125224	6.046988	1114264.017	0.078236	1.197391	1236833	7.298451
1990	2917753	2324946	6.465049	6.366413	6.126072	1336817.776	0.240341	1.739164	1483868	36.17624
1992	2615730	1529516	6.417593	6.184554	6.113027	1297258.916	0.071527	1.179037	1439957	5.855356
1993	2547132	1521826	6.406051	6.182365	6.109854	1287816.297	0.072511	1.18171	1429476	6.068337
1994	1288276	734062.2	6.110009	5.865733	6.028472	1067755.515	-0.16274	0.687482	1185209	61.45888
1995	1368179	777371.3	6.136143	5.890629	6.035656	1085565.571	-0.14503	0.716098	1204978	55.00672
1996	2434873	1644470	6.386476	6.216026	6.104473	1271957.824	0.111553	1.292865	1411873	14.14417
1997	2677715	1924115	6.427764	6.284231	6.115823	1305638.033	0.168408	1.473697	1449258	24.67923
1998	2373641	1830749	6.375415	6.262629	6.101432	1263083.22	0.161197	1.449429	1402022	23.41811
1999	1653177	725488.8	6.218319	5.860631	6.058246	1143526.841	-0.19762	0.634431	1269315	74.95994
2000	1535887	803792.9	6.186359	5.905144	6.049461	1120625.586	-0.14432	0.717272	1243894	54.7531
2001	2169553	1264102	6.33637	6.101782	6.090699	1232249.237	0.011084	1.02585	1367797	8.202995
2002	1689686	965911	6.227806	5.984937	6.060854	1150414.225	-0.07592	0.83962	1276960	32.20263
2003	2249100	1457414	6.352009	6.163583	6.094998	1244507.683	0.068585	1.171076	1381404	5.215408

Appendix I. Calculations of the prediction of the annual suspended sediment loads of the Sacramento River near Red Bluff, California

Year	Daily Flow	Observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	10^logy (Predicted Sediment)	Residual	10^Residual	Corrected predicted sediment	Absolute % deviation
1960	5041930	1397006	6.702597	6.145198	5.896847	788582.6	0.248351	1.771541	928161.7	33.56065
1961	5271284	1982977	6.721916	6.297318	5.922623	836803.3	0.374694	2.369705	984917.5	50.33137
1962	7970570	2538193	6.901489	6.404525	6.16221	1452813	0.242315	1.747089	1709961	32.63079
1963	8965700	3343549	6.952584	6.524208	6.23038	1699732	0.293827	1.967104	2000584	40.16585
1964	5944930	3110896	6.774147	6.492885	5.992309	982446.9	0.500576	3.166477	1156340	62.82936
1965	9934100	3708260	6.997129	6.56917	6.289811	1948998	0.279359	1.902649	2293971	38.13889
1966	6579620	2081062	6.818201	6.318285	6.051086	1124828	0.267199	1.850116	1323922	36.38237
1967	11358500	2718378	7.055321	6.43431	6.367452	2330514	0.066858	1.166428	2743015	0.906323
1968	6758080	1813859	6.829823	6.258604	6.066593	1165716	0.192011	1.556004	1372048	24.35752
1969	12340700	3624573	7.09134	6.559257	6.415508	2603203	0.143749	1.392351	3063970	15.46673
1970	11192500	3162765	7.048927	6.500067	6.358921	2285183	0.141146	1.384031	2689661	14.95856
1971	10046800	2466838	7.002028	6.392141	6.296348	1978554	0.095793	1.246788	2328758	5.597439
1972	6708700	1239855	6.826638	6.093371	6.062343	1154366	0.031027	1.074057	1358688	9.584465
1973	12234000	3066025	7.087568	6.486576	6.410476	2573217	0.076099	1.191514	3028676	1.218153
1974	13664300	3083690	7.135587	6.489071	6.474543	2982244	0.014528	1.034017	3510101	13.82796
1975	9993200	2809760	6.999705	6.448669	6.293248	1964483	0.155421	1.430279	2312197	17.70838
1976	4228690	455196	6.626206	5.658198	5.794926	623629.1	-0.13673	0.729915	734011.4	61.25173
1977	2755160	484582	6.440147	5.685367	5.546686	352116.5	0.138681	1.376198	414441.1	14.47451
1978	9357900	3614496	6.971178	6.558048	6.255189	1799653	0.302859	2.008441	2118191	41.39732
1979	5903710	1329171	6.771125	6.123581	5.988278	973369	0.135303	1.365537	1145655	13.80678
1980	6336810	1262974	6.801871	6.101394	6.029298	1069790	0.072096	1.180582	1259142	0.303381
1981	7815720	2584020	6.892969	6.412296	6.150842	1415278	0.261454	1.825804	1665782	35.53525
1983	18216300	3803752	7.26046	6.580212	6.641148	4376717	-0.06094	0.869088	5151396	35.42933
1984	9130900	1257504	6.960514	6.099509	6.24096	1741645	-0.14145	0.722021	2049917	63.01472
1985	4641980	405853	6.666703	5.608369	5.848958	706249.3	-0.24059	0.57466	831255.4	104.8169
1986	7270000	1290335	6.861534	6.110702	6.108902	1284996	0.001801	1.004155	1512440	17.21298
1987	4521640	777070	6.655296	5.89046	5.833738	681927.8	0.056722	1.139519	802629	3.289156

1988	4882120	627000	6.688608	5.797268	5.878184	755412	-0.08092	0.830011	889120	41.80542
1989	6552660	944958	6.816418	5.975413	6.048707	1118683	-0.07329	0.844706	1316690	39.33842
1990	4419630	314077	6.645386	5.497036	5.820516	661479.5	-0.32348	0.47481	778561.4	147.8887
1991	3801650	649769	6.579972	5.812759	5.733241	541054.9	0.079518	1.20093	636821.6	1.992615
1992	4052750	826498	6.60775	5.917242	5.770302	589253.7	0.146939	1.402618	693551.6	16.08551
1993	10536300	2484319	7.022688	6.395207	6.323913	2108206	0.071294	1.178404	2481358	0.119177
1994	4417510	409941	6.645178	5.612721	5.820238	661056.2	-0.20752	0.62013	778063.2	89.79881
1995	14520600	3071834	7.161985	6.487398	6.509762	3234166	-0.02236	0.949807	3806613	23.91989
1996	12473100	4809150	7.095974	6.682068	6.421692	2640533	0.260377	1.82128	3107907	35.37513
1997	9423850	2279135	6.974228	6.35777	6.259258	1816594	0.098512	1.25462	2138132	6.186702
1998	15621700	3626288	7.193728	6.559462	6.552115	3565454	0.007347	1.017062	4196539	15.72547
1999	9461500	1287397	6.97596	6.109712	6.261568	1826284	-0.15186	0.704927	2149536	66.96764
2000	9086800	1463845	6.958411	6.165495	6.238154	1730432	-0.07266	0.845942	2036718	39.13481
2001	5625210	953376	6.750139	5.979264	5.960278	912594	0.018987	1.044688	1074123	12.66522
2002	6691790	1343465	6.825542	6.128226	6.060881	1150485	0.067345	1.167738	1354121	0.793182
2003	9242450	2015705	6.965787	6.304427	6.247996	1770091	0.056431	1.138758	2083397	3.358253
2004	8367900	1908752	6.922616	6.28075	6.190397	1550235	0.090352	1.231267	1824626	4.407377
2005	9071800	1945053	6.957693	6.288931	6.237197	1726621	0.051734	1.126508	2032233	4.482163
2006	13227070	2795725	7.121464	6.446494	6.455699	2855613	-0.0092	0.979028	3361056	20.22127
2007	5280390	627248	6.722666	5.797439	5.923623	838732.5	-0.12618	0.747852	987188.2	57.38403
2008	4643550	256097	6.66685	5.408404	5.849154	706568	-0.44075	0.362452	831630.5	224.7326
2009	5123170	578670	6.709539	5.762431	5.906109	805580.8	-0.14368	0.718326	948168.7	63.85309
2010	7719330	881382	6.88758	5.945164	6.143651	1392038	-0.19849	0.633159	1638429	85.8932
2011	10712100	1201025	7.029875	6.079552	6.333501	2155268	-0.25395	0.557251	2536750	111.2154
2012	5422870	617738	6.734229	5.790804	5.939051	869062.7	-0.14825	0.710809	1022887	65.58586
2013	4209440	529295	6.624224	5.723698	5.792283	619844.3	-0.06858	0.853916	729556.7	37.83556

Appendix J. Calculations of the prediction of the annual suspended sediment loads of the Cuyahoga River at Independence, Ohio

Year	Streamflow	observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	10^logy (Predicted Sediment)	Residual	10^Residual	Corrected predicted sediment load	Absolute % deviation
1972	418211	280785.5	5.621395	5.448375	5.484947	305454.8	-0.03657	0.919237	543038.7	93.39986
1973	370308	234787.1	5.568563	5.370674	5.426203	266810.3	-0.05553	0.879978	474336.4	102.0283
1974	310935	256936	5.49267	5.409825	5.341817	219693.2	0.068008	1.169522	390571.5	52.01118
1976	12873	1090.7	4.10968	3.037705	3.80407	6368.985	-0.76636	0.171252	11322.81	938.1229
1977	382637	338081.4	5.582787	5.529021	5.442018	276705.7	0.087003	1.221809	491928.5	45.50594
1978	322591	194045.2	5.508652	5.287903	5.359588	228869.4	-0.07168	0.847843	406884.9	109.6856
1979	445671	459323.2	5.649014	5.662118	5.515656	327835.8	0.146462	1.401077	582827.8	26.88838
1980	312912	203203.9	5.495422	5.307932	5.344877	221246.9	-0.03695	0.918448	393333.7	93.56601
1981	23772	301408.9	4.376066	5.479156	4.100265	12596.93	1.378891	23.92717	22394.87	92.56994
1982	387610	248128.2	5.588395	5.394676	5.448254	280707.3	-0.05358	0.883939	499042.5	101.1228
1983	377908	320466.3	5.577386	5.505782	5.436013	272905.9	0.069769	1.174274	485173.1	51.39599
1984	321315	298779.8	5.506931	5.475351	5.357674	227863	0.117677	1.311225	405095.8	35.5834
1987	61014	11508.5	4.785429	4.061019	4.555436	35928.27	-0.49442	0.320319	63873.43	455.0109
1988	238252	108474.9	5.377037	5.035329	5.213244	163397.1	-0.17791	0.663873	290487.9	167.7928
1989	352570	261209.1	5.547245	5.416988	5.402499	252638.4	0.014489	1.033925	449141.6	71.94715
1990	510225	374429.5	5.707762	5.57337	5.580978	381046.1	-0.00761	0.982636	677425.3	80.922
1991	273711	123168.2	5.437292	5.090499	5.280243	190652.5	-0.18974	0.646035	338942.8	175.1869
1992	386622	395045.2	5.587287	5.596647	5.447021	279911.8	0.149626	1.41132	497628.3	25.96744
1993	400724	319712.4	5.602845	5.504759	5.464321	291287	0.040438	1.097586	517851.1	61.97405
1994	302482	240755.1	5.4807	5.381575	5.328507	213062.5	0.053068	1.129974	378783.4	57.33143
1995	279921	197908.5	5.447035	5.296464	5.291076	195468.2	0.005388	1.012485	347504.1	75.58826
1996	531122	474778.2	5.725194	5.676491	5.600361	398438.1	0.07613	1.191599	708344.8	49.19488
1997	354986	265948.2	5.550211	5.424797	5.405797	254564.1	0.019	1.04472	452565.1	70.17038
1998	301597	230582.1	5.479427	5.362826	5.327092	212369.5	0.035733	1.085759	377551.4	63.73837
1999	255143	168952.3	5.406784	5.227764	5.24632	176327.5	-0.01856	0.958173	313475.7	85.54096
2000	313920	215194.6	5.496819	5.332831	5.34643	222039.5	-0.0136	0.969172	394742.8	83.43526

2001	234688	87067.9	5.370491	4.939858	5.205966	160681.6	-0.26611	0.541866	285660.4	228.0892
2002	248926	148104.2	5.39607	5.170567	5.234408	171556.7	-0.06384	0.863296	304994.3	105.9322

Appendix K. Calculations of the prediction of the annual suspended sediment loads of the Mississippi River at McGregor, Iowa

Year	Streamflow	Observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	10^logy (Predicted Sediment)	Residual	10^Residual	Corrected predicted sediment	Absolute % deviation
1975	5684300	686653	6.754677	5.836737	5.781383	604481	0.055355	1.135938	662129.9	3.571403
1976	9324740	1277910	6.969637	6.1065	5.97504	944147.6	0.13146	1.353507	1034190	19.07175
1977	8372850	919122	6.922873	5.963373	5.932911	856861.7	0.030462	1.072661	938579.9	2.117006
1978	14055800	2093641	7.147856	6.320902	6.135597	1366461	0.185305	1.532163	1496779	28.50831
1979	17300200	3285067	7.238051	6.516544	6.216854	1647610	0.29969	1.993838	1804741	45.06227
1980	12232000	3128330	7.087497	6.495313	6.081221	1205648	0.414092	2.594729	1320630	57.78483
1981	11981300	1759986	7.078504	6.245509	6.073118	1183364	0.172391	1.487274	1296220	26.35052
1982	17311900	3824722	7.238345	6.5826	6.217119	1648614	0.365481	2.319962	1805841	52.78505
1983	19651400	1936233	7.293393	6.286958	6.266712	1848044	0.020245	1.04772	2024291	4.547885
1984	18918000	2150852	7.276875	6.332611	6.251831	1785793	0.080779	1.204424	1956102	9.054539
1985	18849500	2172118	7.2753	6.336883	6.250412	1779966	0.086472	1.220314	1949720	10.23875
1986	24517300	2900362	7.389473	6.462452	6.35327	2255641	0.109182	1.285826	2470760	14.81201
1987	10052500	919942	7.002274	5.96376	6.004443	1010283	-0.04068	0.910579	1106632	20.29372
1988	7571670	454135	6.879192	5.657185	5.893558	782632.6	-0.23637	0.580266	857271.6	88.77022
1989	9078260	887623	6.958003	5.948229	5.964559	921634.5	-0.01633	0.963097	1009530	13.7341
1990	11847700	1626964	7.073634	6.211378	6.068731	1171470	0.142647	1.388823	1283192	21.12967
1991	17807800	2167031	7.25061	6.335865	6.228169	1691099	0.107696	1.281434	1852377	14.52004
1992	15618600	1237758	7.193642	6.092636	6.176846	1502610	-0.08421	0.823739	1645913	32.97534
1993	23834100	2257462	7.377199	6.35362	6.342213	2198936	0.011408	1.026616	2408646	6.697091
1994	17159700	1149783	7.23451	6.060616	6.213664	1635550	-0.15305	0.702995	1791531	55.81475
1995	18114900	1125155	7.258036	6.051212	6.234859	1717350	-0.18365	0.655169	1881132	67.18869
1996	14620500	709030	7.164962	5.850665	6.151009	1415822	-0.30034	0.50079	1550848	118.7281
1997	18189400	1379997	7.259818	6.139878	6.236465	1723711	-0.09659	0.800596	1888100	36.81915
1998	14465600	723111	7.160336	5.859205	6.146841	1402301	-0.28764	0.51566	1536037	112.4207
1999	15485700	856918	7.189931	5.932939	6.173503	1491087	-0.24056	0.574694	1633290	90.60053
2000	11830570	798609	7.073006	5.902334	6.068165	1169944	-0.16583	0.682605	1281520	60.46905

2001	19521500	1471534	7.290513	6.16777	6.264117	1837035	-0.09635	0.801037	2012232	36.74382
2002	17257600	1097898	7.23698	6.040562	6.21589	1643955	-0.17533	0.66784	1800737	64.01679
2003	11940440	816899	7.07702	5.912168	6.071782	1179728	-0.15961	0.692447	1292237	58.18814
2004	10814500	1126860	7.034006	6.05187	6.033031	1079023	0.018839	1.044334	1181928	4.886867

Appendix L. Calculations of the prediction of the annual suspended sediment loads of the Kankakee River at Momence, Illinois

Year	Streamflow	observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	Predicted Sediment (10^log Y)	Residual	10^Residual	Corrected Predicted Sediment	Absolute % Deviation
1979	831635	162766	5.92	5.21	5.18	151567.61	0.03	1.07	160661.66	1.29
1980	709777	120241	5.85	5.08	5.08	120071.71	0.00	1.00	127276.02	5.85
1981	914940	317682	5.96	5.50	5.24	174405.70	0.26	1.82	184870.04	41.81
1982	2071180	748218	6.32	5.87	5.76	579728.60	0.11	1.29	614512.32	17.87
1993	1464910	223055	6.17	5.35	5.54	348414.73	-0.19	0.64	369319.61	65.57
1994	914254	109332	5.96	5.04	5.24	174213.48	-0.20	0.63	184666.29	68.90
1995	891268	157080	5.95	5.20	5.22	167812.18	-0.03	0.94	177880.91	13.24

Appendix M. Calculations of the prediction of the annual suspended sediment loads of the Smoky Hill River at Lindsborg, Kansas

Year	Streamflow	observed sediment	Log Flow (X)	Log SSL (Y)	Log Y	10^logy (Predicted Sediment)	Residual	10^Residual	Corrected predicted sediment	Absolute % deviation
1958	198723	1179313	5.30	6.07	5.89	776336.04	0.18	1.519075813	875215.34	25.786
1959	82976	780114.1	4.92	5.89	5.53	335885.64	0.37	2.322558636	378666.26	51.4601
1961	1102450	6288760	6.04	6.80	6.60	4016753.74	0.19	1.565632452	4528354.11	27.9929
1963	269439	938236	5.43	5.97	6.02	1039635.33	-0.04	0.902466446	1172050.17	24.9206
1964	150427	326617	5.18	5.51	5.77	594359.14	-0.26	0.549528017	670060.66	105.152
1965	471481	2710375	5.67	6.43	6.25	1778256.95	0.18	1.524175119	2004747.54	26.0343
1967	557690	4126217	5.75	6.62	6.32	2089079.56	0.30	1.975136362	2355158.58	42.9221
1968	192676	487604.2	5.28	5.69	5.88	753659.97	-0.19	0.646981688	849651.10	74.2502
1969	462857	2026980	5.67	6.31	6.24	1747042.47	0.06	1.160234819	1969557.37	2.83289
1970	173621	473033	5.24	5.67	5.83	682010.02	-0.16	0.693586587	768875.33	62.5416
1971	454743	2345353	5.66	6.37	6.23	1717652.31	0.14	1.365441300	1936423.89	17.4357
1972	242317	641422.1	5.38	5.81	5.97	939030.66	-0.17	0.683068327	1058631.83	65.0445
1973	2017408	6662470	6.30	6.82	6.86	7171812.80	-0.03	0.928979908	8085262.41	21.3553
1974	743948	1100417	5.87	6.04	6.44	2754299.16	-0.40	0.399527044	3105104.95	182.175
1975	348701	897520.6	5.54	5.95	6.12	1331421.79	-0.17	0.674106889	1501000.49	67.2386

Appendix N. Calculations of the prediction of the annual suspended sediment loads of the Cuyahoga River at Old Portage, Ohio

Year	Streamflow	Observed Sediment	Log Flow (X)	Log SSL (Y)	Log Y	Predicted Sediment (10^Log Y)	Residual	10^Residual	Corrected Predicted Sediment	Absolute % Deviation
1972	140407	18204.37	5.15	4.26	4.37	23690.36	-0.11	0.77	23927.26	31.4
1973	178536	29004.59	5.25	4.46	4.52	32957.30	-0.06	0.88	33286.87	14.8
1974	199147	42355.37	5.30	4.63	4.58	38296.10	0.04	1.11	38679.06	8.7
1975	202045	47499.81	5.31	4.68	4.59	39064.01	0.08	1.22	39454.65	16.9
1976	107167	14953.12	5.03	4.17	4.21	16343.31	-0.04	0.91	16506.74	10.4
1977	200790	36106.10	5.30	4.56	4.59	38730.95	-0.03	0.93	39118.26	8.3
1978	175102	35630.34	5.24	4.55	4.51	32089.33	0.05	1.11	32410.22	9.0
1979	198025	39580.70	5.30	4.60	4.58	37999.92	0.02	1.04	38379.92	3.0
1980	180513	35564.12	5.26	4.55	4.52	33459.85	0.03	1.06	33794.45	5.0
1981	92544	16961.20	4.97	4.23	4.13	13359.36	0.10	1.27	13492.96	20.4