1975

A statistical study of the variations in Des Moines River water quality

Lewis Metzler Naylor
Iowa State University

76-1861

NAYLOR, Lewis Metzler, 1942-
A STATISTICAL STUDY OF THE VARIATIONS IN
DES MOINES RIVER WATER QUALITY.

Iowa State University, Ph.D., 1975
Engineering, civil

A statistical study of the variations in

Des Moines River water quality

by

Lewis Metzler Naylor

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Civil Engineering
Major: Water Resources

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1975

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Discharge of domestic and industrial wastes to surface waters has long been recognized as an environmental problem. Since the Royal Rivers Pollution Act was passed in England in 1876 (65), great effort and enormous sums of money have been expended to assess this problem and improve wastewater treatment operations. In 1972, the 92nd United States Congress passed Public Law 92-500. This law, which extended the Water Pollution Control Act of 1965 and its amendments, authorized $18 billion for publicly owned wastewater treatment plants (45). A price tag of this magnitude brings to mind a remark by George Bernard Shaw when invited to comment on a grandiose scheme, "This is probably the way God would have done it if He had the money" (18). Perhaps the most significant point of the Water Pollution Act of 1965 is the establishment of stream standards rather than effluent standards (64). One of the outcomes of this point is a goal of P.L. 92-500 which specifies that the discharge of pollutants into the navigable waters of the nation be eliminated by 1985.

In the midwestern states, large land areas are devoted to agricultural operations. In the State of Iowa, probably more than 90 per cent of the state's acreage is devoted to some type of agricultural use. Because of these extensive agricultural operations, the pollution potential from agricultural sources in the Des Moines River Basin, as an example, greatly exceeds that from domestic and industrial sources (7). Thus, it is unlikely that meeting the 1985 goal of P.L. 92-500 by industries and municipalities will have any measurable effect on surface water quality in Iowa where runoff is the primary problem.

Variations in surface water quality are a function of the agricultural activities and the natural events that take place within that basin, as illustrated by several examples. An intense spring thunderstorm closely following cultivation and fertilizer application will result in a host of changes in the limnological character of nearby rivers and streams. Surface runoff washes accumulated livestock waste and debris from other agricultural operations into rivers. Groundwater flowing through highly fertilized soils transports dissolved minerals and plant nutrients to adjacent streams. In a river basin containing many livestock and few people, many farms and few cities, surface runoff and not municipal wastewater discharges is expected to be the principal factor controlling the water quality of the basin's streams and rivers.

Seasonal climatological conditions and hydrologic factors are closely associated with the effects of dispersed or non-point source discharges (runoff) on surface water quality. Using this statement as a working hypothesis, the following objectives for the research were formulated.

Objectives of the Study

1. To evaluate the importance of hydrologic factors such as runoff and seasonal climatologic factors such as runoff on water quality in the Des Moines River above the Des Moines Metropolitan area.

2. To determine what portion of the variability in water quality could be accounted for considering solely seasonal relationships within the upper Des Moines River Basin and the hydrologic characteristics of the river.

3. To determine whether the impact of agricultural runoff on the upper Des Moines River could be evaluated in terms of seasonal and hydrologic relationships.

LITERATURE REVIEW

Domestic and industrial wastes are discharged on a nearly continuous basis. The discharge of these wastes into streams has the greatest impact on water quality during low-flow conditions. Because water use and waste production are relatively constant throughout the year, the effects of the wastewater discharge on receiving streams can be predicted based on average quality and quantity of the wastewater discharge and the temperature and discharge rate of the stream.

One of the earliest attempts to predict the effects of wastewater discharges on river water quality were the classical studies by Streeter and Phelps on the Ohio River in 1925 (55). From their studies were developed mathematical formulations which describe the interplay of the deoxygenation of polluted water as caused by its biochemical oxygen demand (BOD), and subsequent reoxygenation from the atmosphere. In a riverine environment, this relationship describes a spoon-shaped profile of the dissolved-oxygen deficit along the path of water movement, commonly called the dissolved-oxygen sag curve. Since the early work by Streeter and Phelps many dissolved oxygen models have been developed, and, in some cases, used to estimate the maximum permissible pollutional loading of streams (15, 21, 30, 48). Some models have been expanded to include prediction equations for a number of water quality substances (27, 30).

Surface Runoff Effects on Water Quality

Surface runoff, in contrast with wastewater inputs from point sources, is intermittent and coincides with rainfall and snowmelt events. Consequently, the impact of the pollutants in runoff will frequently be greatest during periods of high streamflow. (The qualification is given here because for some pollutants, although there may be tremendous quantities washed into the stream, the quantity of the runoff water dilutes the pollutant concentration to less than that observed during base flow conditions.)

Attempting to define relationships of water quality with a Streeter-Phelps type of formulation can be very complex, especially during periods of high runoff. Although inputs of pollutants may be expected to be quite large, the actual input is a function of many factors. Of these, runoff is but one, albeit an important one. The quality of land runoff depends on climatological, hydrological, geological, and land-use factors in the particular river basin. The relationship may be simplified by considering that for a given basin some of these factors, such as geological conditions and land-use patterns, do not change appreciably over a period of years.

Several studies have focused on the effects of agricultural land runoff on river water quality. Wallace and Dague (63) modeled the effects of land runoff on dissolved oxygen in the Iowa River during within-bank streamflow conditions. In the water quality model they developed, the Streeter-Phelps equation was used to predict dissolved-oxygen (DO) concentrations. Low DO in the river could not be attributed to domestic and industrial wastewater discharges. They felt that the only cause of low DO (less than the Iowa standard of 5 mg/l) was land runoff, and this occurred only during periods of high streamflow. At other times, the DO was consistently near saturation. Their model indicated that for low Do to occur during periods of high streamflow, a runoff ultimate BOD of 40 to 45 mg/l (5-day BOD of about 27 to 30 mg/l) was required. Although this value at first seemed too high, it was consistent with observed data from one part of the basin. They suggested that the nitrogenous oxygen demand may also be a significant cause of oxygen depletion in the river. The oxygen equivalent of 1 mg/l of ammonia nitrogen is about 4.5 mg/l.

Harms (31) made a thorough study of the physical and chemical quality of agricultural land runoff in South Dakota, but did not relate it to surface water quality. This study is of interest, however, because rainfall runoff and snowmelt runoff were considered individually. The two-year project, 1971 and 1972, included one of the wettest months on record, May 1972. Over 9 in of rain was recorded at one site. The period also included a very dry summer when very little runoff occurred. While rainfall accounted for only one-third of the total quantity of runoff, it was responsible for nearly all of the soil loss and two-thirds of the chemical oxygen demand (COD) lost in the runoff. However, snowmelt runoff caused nearly two-thirds of the nitrogen and one-half the phosphorus lost in runoff. The phosphorus loss was especially interesting. About one-half the phosphorus lost in snowmelt runoff was soluble, whereas in rainfall runoff only 10 percent was soluble. The results seemed to indicate that soil conservation practices directed at limiting soil loss may not appreciably reduce inputs of nutrients to streams in South Dakota. Most of the annual nutrient load for the area was contributed by snowmelt runoff, and a large percentage was soluble.

These results, however, appear to contradict findings of other researchers. Holt (33) in a summary paper, reported that soil lost during runoff was nutrient enriched by a factor of two or three compared to the surface soil from which the sediments were derived. In their survey, no attempt was made, however, to differentiate between the chemical nature of snowmelt runoff and rainfall runoff.

Results from a study of nitrogen losses in surface runoff from agricultural watersheds on Missouri Valley Loess indicated that water soluble nitrogen and sediment nitrogen losses in runoff were usually highest at the beginning of the cropping season (50). Amounts decreased progressively throughout the year reflecting a seasonal effect. It was felt that this was associated with nutrient removal by crops, leaching, and nitrogen being combined with organic matter.

Willrich (67) studied the properties of tile drainage water in Iowa. The concentrations of nitrogen and phosphorus found in the drainage water were sufficiently high to be conducive to the growth of algae and other aquatic plants. Practically all (99 per cent) of the nitrogen in the tile water samples was in the oxidized nitrate form. The median concentration of the total nitrogen ranged from 12 to 27 mg/l, and was found to be independent of flow rate. Median values for phosphorus concentrations were 0.2 to 0.3 mg/l. Beer (12) and Holt (32) have indicated that in a completely tile-drained area a large percentage of the water yield may be derived from subsurface drains.

A study of 14 streams in central Iowa considered factors statistically important in influencing suspended algae densities (38). Factors included were adjusted streamflow, water temperature, upstream watershed area, nutrients, and human population. Streamflow, water temperature, and upstream watershed area, all physical factors, accounted for 50 per cent of variance in a multiple linear regression equation. The only significant correlation coefficient of algal density, as represented by the concentration of chlorophyll a in mg/m^3, with several factors was upstream watershed area (r=0.73). The regression equation indicated an inverse relationship between chlorophyll a concentration and adjusted streamflow and a direct relationship with temperature. Addition of ammonia nitrogen to the regression increased the R^2 value by only two per cent. Human population effects were negligible. The conclusions were that algae are generated in the upstream watershed bottom or benthic areas and probably not within the flowing stream. Nutrient concentrations appeared to have little effect on algal populations.

Jones (37) made a study of the limnological characteristics and factors influencing the water quality of a reach of the Skunk River near Ames, Iowa. The Skunk River Basin, which lies entirely within the State of Iowa, is adjacent to and east of the Des Moines River Basin. Jones concluded that sewage effluents could cause ammonia nitrogen concentraions to exceed the Iowa water quality standard only during low flow periods. He also noted that the algae suspended in the river were benthic forms, and their density, as represented by chlorophyll a concentrations, was inversely correlated with streamflow. Highest chlorophyll a concentrations corresponded with a decrease in nutrient concentration. He suggested that this may indicate depletion of nutrients during low flow conditions.

Statistical Water Quality Models

Mathematical modeling of surface water quality may be approached in two different ways. A causal mathematical model may be developed which is based on known and suspected biological, chemical, and physical causes and effects. Alternatively, a statistical model may be developed which is based on statistical interdependence.

In statistical modeling a river basin may be considered a "black box" for which known inputs result in reliably predictable outputs. If nature is considered to be orderly and to respond in a similar manner to a particular set of environmental conditions, it should be possible to develop a statistical model which would predict water quality based on known statistical relationships and on existing conditions. The seasonal variation in water quality in response to typical climatological conditions is one example. Other examples are that rivers are laden with sediment following snowmelt and rainfall in the spring, and that during harsh winter conditions few plankton are found in the river. Year after year, the general pattern is repeated.

Where a considerable body of water quality data is available covering a number of years of river basin conditions influenced by a wide range of environmental conditions, water quality relationships may be developed statistically without resorting to the complexities of a causal mathematical model. Although statistical dependence among water quality and its determining factors does not necessarily imply causal dependence, these relationships may suggest important, but unapparent relationships that would be ignored, and could be of a causal nature.

Tirabassi (60) developed a statistical water quality model for the Passaic River Basin in New Jersey. The Passaic River is an old, slow moving river whose water quality is influenced, to a large degree, by effluents from domestic and industrial wastewater treatment plants. During critical low flow periods in the summer, these effluents make up about half the flow of the river. Seventeen water quality parameters were monitored biweekly over a five-year period, largely during base-flow conditions. Multiple linear regression equations of each of the water quality parameters were developed as functions of the other 16 parameters plus streamflow. The regression equations were developed in a stepwise fashion similar to that described by Draper and Smith (19). Tirabassi found that seasonal partioning of the data according to the natural warm-cold periods of June-November and December-May was statistically significant for about one-third of the parameters. Although his regression equations may have been good predictors of water quality (R^2 and F values were omitted from the results), the equations added little to the knowledge of the river's ecological relationships. For example, the best predictor of chloride was alkalinity, and of ammonia were odor (a highly subjective test) and alkalinity. In fact, alkalinity appeared in half the regression equations.

One of the values of this study is that it provides a good illustration of the potential abuse of statistics in regard to interpretation of the results of regression analysis. Causal inferences must be drawn only after careful consideration of the ecological relationships. In other words, it is highly unlikely that alkalinity caused changes in chloride concentration. At best, alkalinity was associated with another parameter which was the real cause of changes in chloride concentration.

Nour (46) developed a statistical water quality model for the Pearl River Basin in Mississippi and Louisiana. The principal factor influencing water quality was the discharge of domestic and industrial waste. Eight to twelve water samples were collected at eight sites during base-flow conditions over a period of about a year. The water samples were analyzed for 16 water quality substances.

In contrast with Tirabassi, Nour included as explanatory variables, temperature, streamflow, location, and month (January = 1, etc.), and did not include the actual water quality parameters. In his model, for example, nitrate was not used to predict nitrite, and vice-versa. Rather his unstated hypothesis was that water quality was a function of spatial, temporal, and physical factors. Water quality relationships as a function of the explanatory variables were developed using a stepwise multiple linear regression technique (19). His regression equations explained 47 to 95 percent of the variance contained in the original data, but the small sample size (8 to 12 observations) limits the applicability of the results. No attempt was made to infer causal relationships.

Another way of handling data which is to be analyzed by regression analysis is through the use of principal component analysis. This procedure searches for linear dependencies among the so-called independent variables and derives transforms for their elimination. In essence, new variables are statistically constructed so that the original data matrix is transformed into a set of uncorrelated vector components. The strength of this procedure is that, although the vector components contain the same information as the original data matrix, most of the information will be contained in fewer vectors than the number of the original variables. Thus, the dimensionality of the problem may be reduced. It may be found, for example, that 90 per cent of the information contained in 20 variables could be represented by two or three vector components.

Principal component analysis would have been very useful in a study such as that of Tirabassi (60) in which a large number of independent or explanatory variables were used to predict variations in a water quality parameter. Two excellent papers covering the statistical-mathematical aspects and the practical applications of principal component analysis have been written by Fiering (24) and Mahlock (42).

Ledbetter and Gloyna (41) presented a number of different methods for representing the concentration of chlorides and other materials as a function of streamflow in rivers of the Southwest. Of particular interest was their inclusion of an antecedent flow index which expressed the immediate past history of flow for the location under consideration. There was some indication from their results that separation of the data on the basis of seasonal wet-dry periods would be appropriate. Although they made extensive analyses of chloride relationships using several different transformations of streamflow, and reviewed and evaluated frequency and probability relationships of streamflow and chloride, statistical analyses were limited to correlations of flow with dissolved solids and chloride.

CHARACTERISTICS OF THE DES MOINES RIVER BASIN IN IOWA

Tracing the Flow of the Des Moines River

The Des Moines River Basin may be considered to consist of three sub basins: the lower Des Moines, the upper Des Moines, and the Raccoon River basins as shown in Figure 1. Although several rivers and many small streams drain into the lower Des Moines River basin, the limnological character of the lower Des Moines River is primarily a function of the water quality inputs from the basins above it. The upper Des Moines River basin is the larger of the two upper basins.

Row crops are the predominant agricultural activity in the upper Des Moines River basin, whereas livestock feeding predominates in the Raccoon River basin. These two agricultural activities and the Des Moines Metropolitan Area are the principal factors which influence water quality in the lower Des Moines River basin. The general plan of the entire Des Moines River Basin is shown in Figure 2.

The source of the West Fork of the Des Moines River is in the meadows of Murray and Lyon Counties in southwestern Minnesota at an altitude of 1800 to 1850 feet above sea level. The outlet of a large shallow lake, Lake Shetek, forms the initial stream in a flat plain area. Several small lakes drain to Lake Shetek. The northernmost of these is Long Lake in Lyon County, Minnesota, which lies less than 5 miles south of the Cottonwood River which flows into the Minnesota River at New Ulm, Minnesota, and ultimately into the Mississippi River at St. Paul,

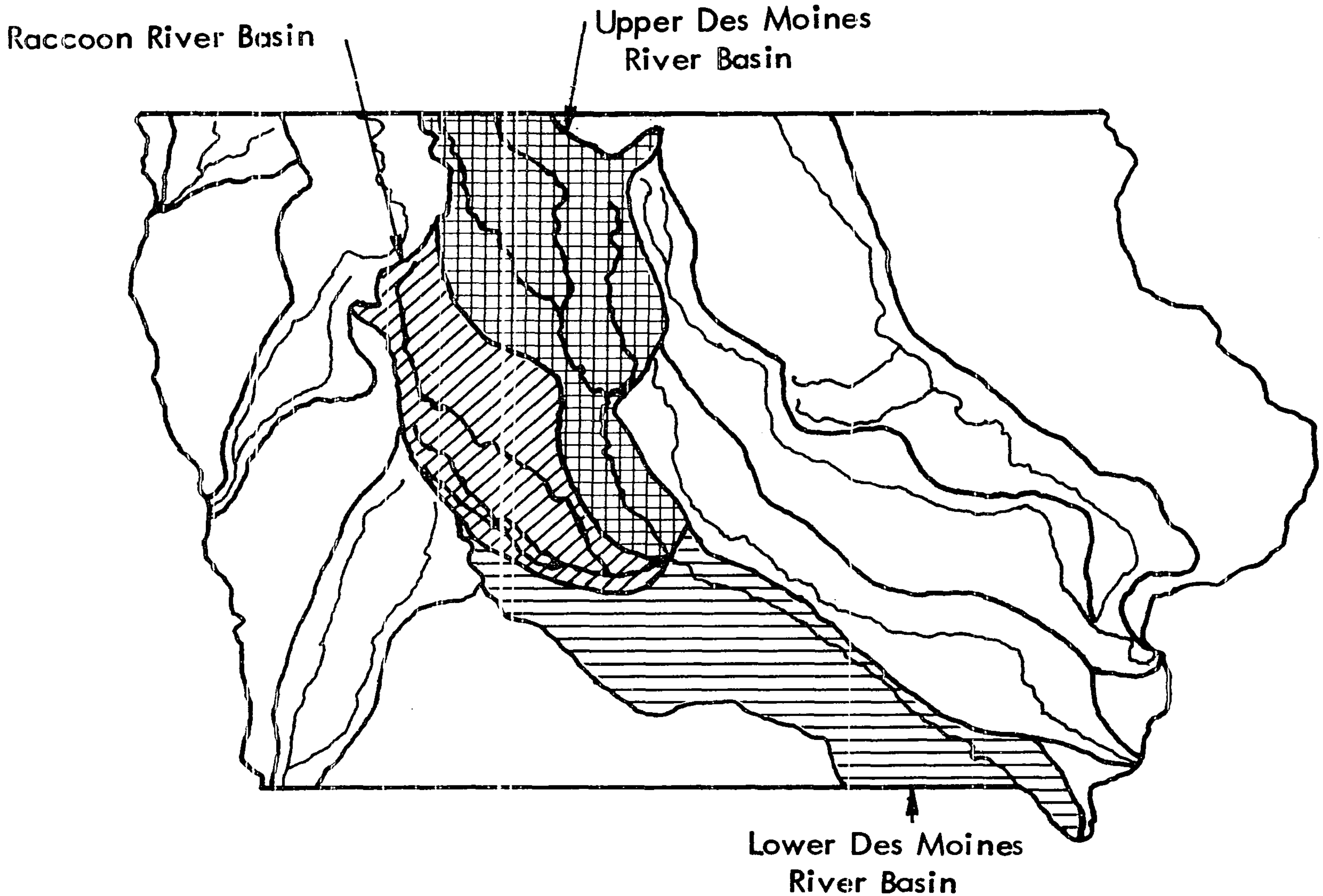


Figure 1. Locations of sub-basins in the Des Moines River Basin

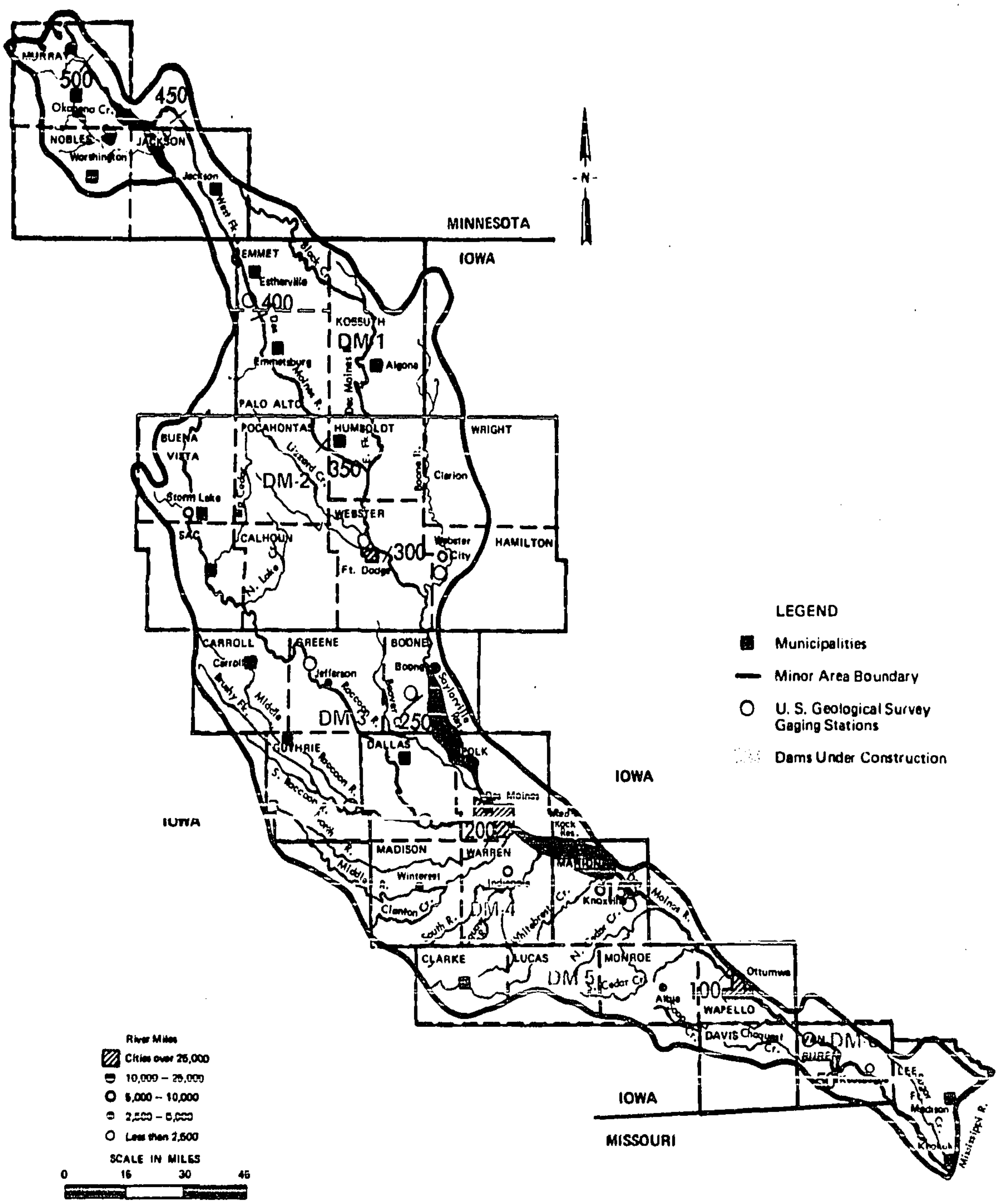


Figure 2. General plan of the Des Moines River Basin

Minnesota. The East Fork of the Des Moines River is formed by the outlet of Okamanpeden Lake near the Iowa-Minnesota border.

From Lake Shetek, approximately 40 miles north of the Iowa-Minnesota border, the West Fork of the Des Moines River flows in a southeasterly direction where it is joined by the East Fork a few miles below Humboldt, Iowa. The confluence of the Boone River and the Des Moines River is just above Stratford, Iowa, and the Raccoon River enters at Des Moines, Iowa. Below Des Moines, many smaller rivers flow into the major existing Des Moines River impoundment, the Red Rock Reservoir. The Des Moines River forms the boundary between Iowa and Missouri from Farmington, Iowa to Keokuk, Iowa, a distance of about 30 river miles. The total length of Iowa's largest river from its source in Minnesota to its mouth immediately below Keokuk, Iowa, is about 535 miles. There, the river empties into the Mississippi river, 486 miles below St. Paul, Minnesota.

Watershed Characteristics

More than 14,500 square miles of three states are drained by the Des Moines River, including 23 per cent of Iowa. The watershed has a long and relatively narrow crescent shape averaging about 40 miles in width from southwestern Minnesota to the Iowa-Missouri border. Figure 3 shows the location of continuous-record stream gaging and water quality stations located in the upper Des Moines River basin. Streamflow discharge rates used in this study were taken from the U.S.G.S. recording gage records at Saylorville, Iowa (65). The river at the Saylorville gaging station, about 8 miles north of Des Moines, drains 5,841 square miles.

From its source in Minnesota to its outlet at Keokuk on the Mississippi River, the Des Moines River falls nearly 1,370 feet. The stream slope averages 3.2 feet per mile from the source to river mile 300 near Fort Dodge, Iowa. The slope then becomes more gentle, about 1.6 feet per mile, from river mile 300 to the confluence with the Mississippi River. It is interesting to note that although the river is navigable only for small boats at the present time, a steamboat was able to bring supplies to Des Moines from Keokuk in 1851. During still another high-water period of yesteryear, the river was navigable as far north as Fort Dodge.

The Des Moines River watershed lies in a recently glaciated plain in which the valley cut into the glaciated area does not generally exceed 200 feet. Many lakes and ponds dot the headwater area and a rather poorly defined drainage pattern exists in northern Iowa (Figure 2). The stream has cut deeper near Humbolt, Iowa, exposing the limestone underlying the glacial till. In Boone County, the valley formed by the river is about 1/4 mile wide and 150 to 200 feet deep. Sandstone outcroppings in the Ledges State Park south of Boone, Iowa are a major scenic attraction in the area. The valley widens in Dallas and Polk Counties to about 1/2 mile where the river cuts through the drift in the vicinity of the Saylorville Reservoir.

Below Des Moines, where the valley is mature, the landscape changes dramatically, and the drainage pattern is well defined (Figure

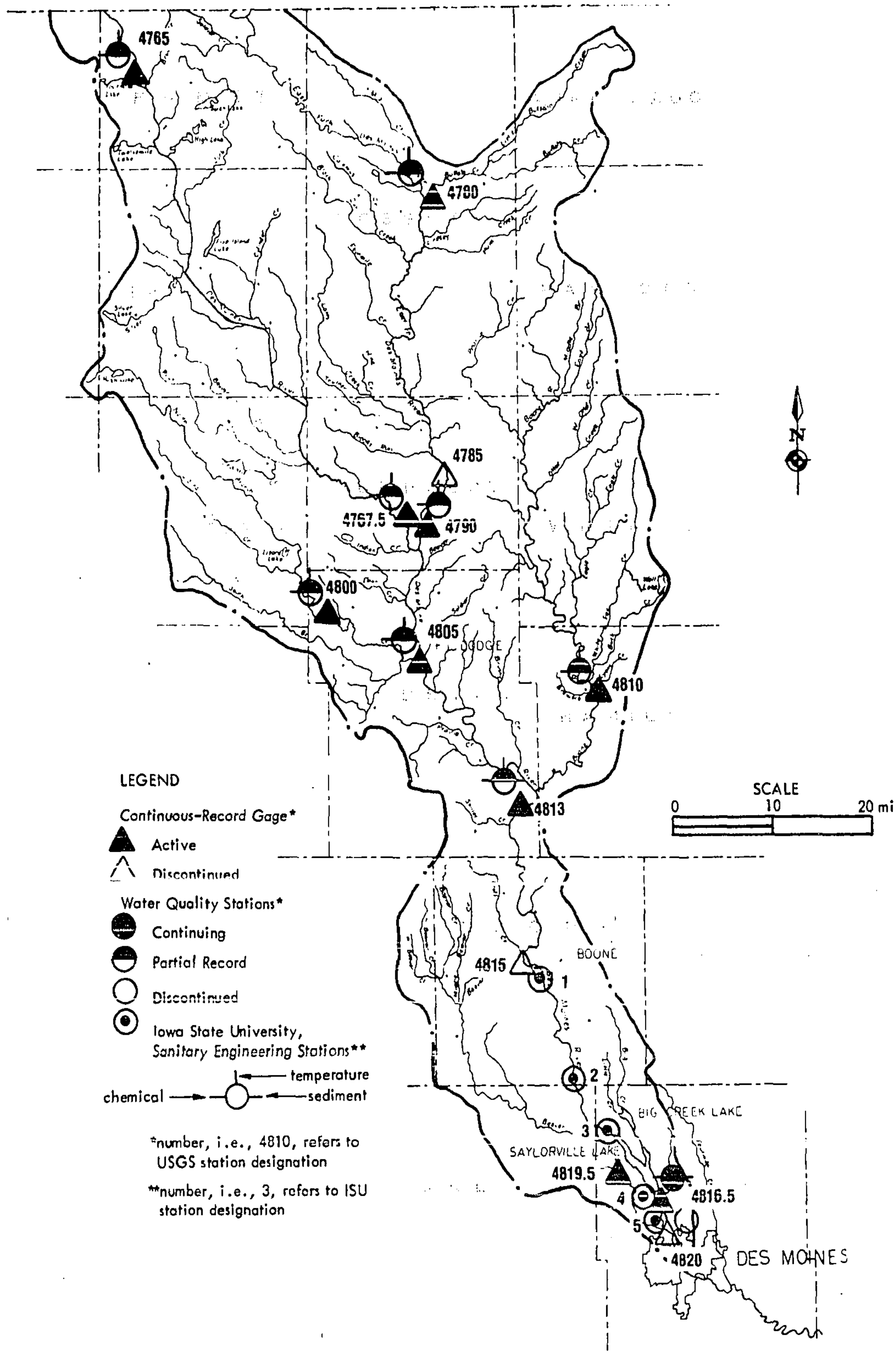


Figure 3. Location of continuous-record gaging stations and water quality sampling stations

2). In the area between Des Moines and Knoxville, Iowa, the river meanders through a flood plain 2 to 4 miles wide bordered by rounded cliffs. Red Rock Reservoir covers much of this area at high water levels, but at the conservation pool level, river meanders are still visible. Downstream near the site of Red Rock Dam in Marion County, the valley width is reduced to 1 to 3 miles, forming a deep flat-bottomed valley. Near Tracey, Iowa the river has cut into the limestone, forming a flood plain 1/2 to 2 miles in width. The stream valley in Mahaska County and Van Buren County becomes constricted to a width ranging from 1/3 to 1 mile wide, but below this reach the flood plain again becomes wider and is bordered by rounded bluffs in the vicinity of the confluence of the Des Moines River with the Mississippi River.

The soils found in the Upper Des Moines River watershed are moderately permeable. Pockets of sand and gravel are common, and these are highly permeable. Because of the moderate to high permeability, drainage problems during wet spring weather are localized. The water covered areas are generally confined to clay pan soils in pot hole areas and in places where the land is lower than the adjacent roads.

Large quantities of water percolate into the permeable soil and contribute to the groundwater supply rather than direct runoff into the streams. In poorly-drained areas, however, tile drains and open ditches divert much of this excess water from fields into the stream. It has been estimated that as much as half of the upper Des Moines River Basin is artificially drained with tiles and open ditches (61).

Population

Based on national census data, the state of Iowa is growing more slowly than the rest of the nation. In the 70 year interval between 1900 and 1970, the percentage of the U.S. population living in Iowa has steadily declined from 2.84 per cent to 1.39 per cent.

Many factors influence population growth and decline within the state. However, the fundamental factors are mortality, fertility, and migration. Mortality has not changed substantially for many years. Hence, fertility and net migration in effect control variations in the Iowa population. In Iowa, and throughout the rest of the nation, fertility is declining. The peak numbers of births (66,123) in Iowa occurred in 1951 following World War II. In 1974, the number of births was less than 39,000, the lowest since 1917. This represents a current trend toward smaller families begun in 1959. Thus, if the population of Iowa is to grow, the dominant factor appears to be the net migration. Between 1900 and 1970 one million more people moved out of Iowa than moved into the state.

Based on 1973 population estimates and considering the population to be distributed uniformly throughout each county, the population of the upper Des Moines River basin above Des Moines is approximately 150,000 including the rural population. The only cities over 10,000 population in this area are Boone (pop. 12,468) and Fort Dodge (pop. 31,263).

Population distribution

More than half a million people live in the cities and towns of the Des Moines River Basin, including the Raccoon River Basin according to the 1970 census records, and nearly 40 per cent of these people live in the Des Moines metropolitan area. Of the remaining 60 per cent, about 14 per cent live in the Raccoon River Basin, 23 per cent live in the upper Des Moines River Basin, and 23 per cent live in the lower Des Moines River Basin. Population distribution data are summarized in Table 1. Figure 4 shows the locations of municipalities in the upper

Table 1. Population distribution in the Des Moines River Basin - 1970

Area	Municipal Population	Per cent of Total
Des Moines River Basin	504,606	100.0
Des Moines River Basin above Red Rock Dam	436,241	85.6
Des Moines Metropolitan Area	201,404	40.0
Lower Des Moines River Basin	315,558	62.5
Upper Des Moines River Basin	116,647	23.1
Raccoon River Basin	72,401	14.4

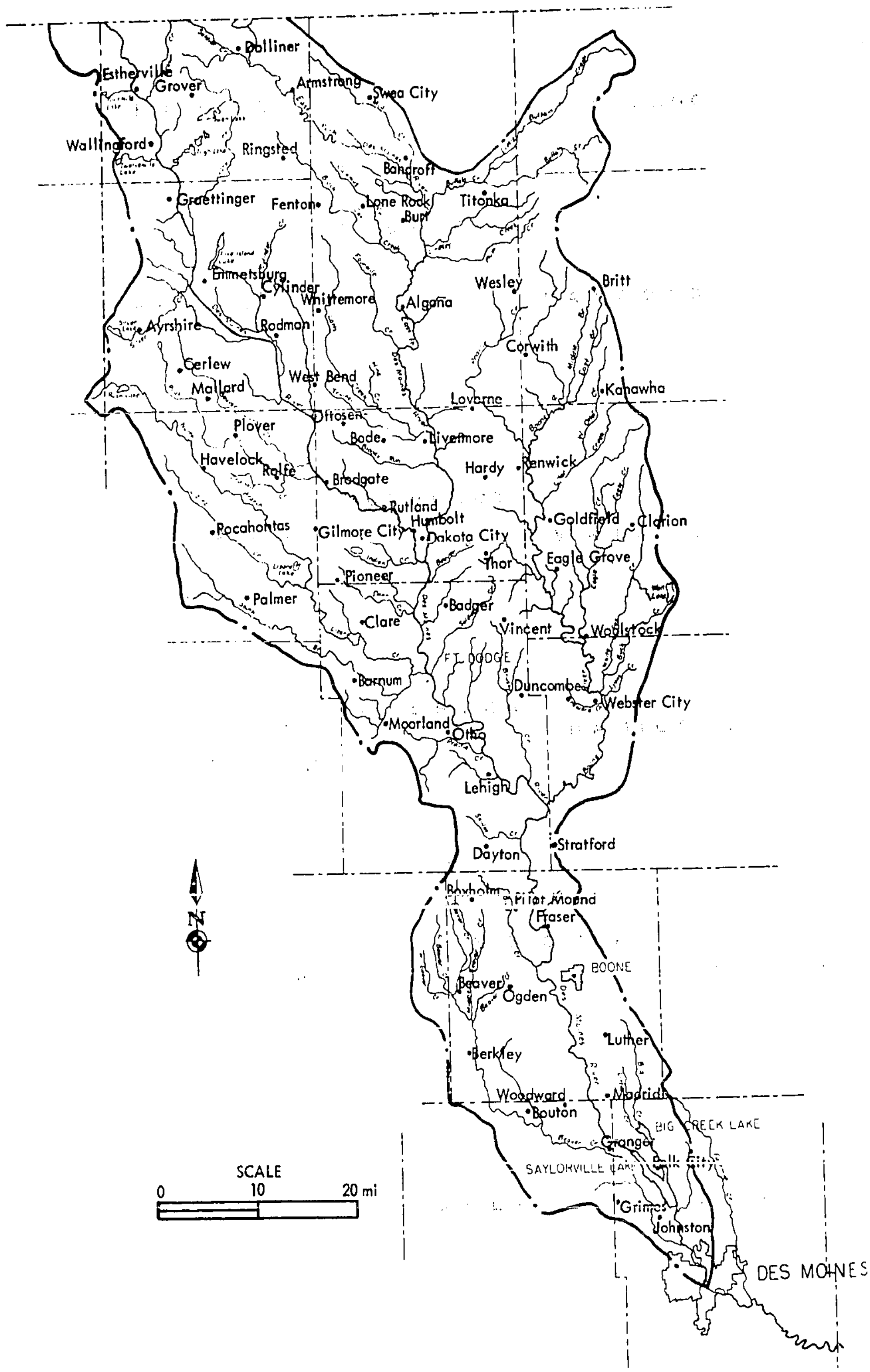


Figure 4. Municipalities in the upper Des Moines River Basin in Iowa

Des Moines River Basin. There are few large cities in the basin above the City of Des Moines.

Municipal wastewater treatment

The principal methods of treating municipal waste in the Des Moines River basin are trickling filters and waste stabilization ponds. Trickling filter plants serve the greatest population, including the Des Moines metropolitan area.

Communities in the upper Des Moines River basin contribute about 6,400 to 8,300 lb BOD/day, respectively, during the summer and winter to the rivers and streams of the basin (61). Few cities and towns in the basin area produce more than 100 lb BOD/day and nearly half the point sources add less than 10 lb BOD/day. Waste inputs from eight point sources contribute nearly 78 percent of the total BOD. One of these cities, Estherville, with a population just over 8,100, contributes nearly 2400 lb BOD/day. It has, however, submitted plans to the Iowa Department of Environmental Quality for polishing ponds and dual-media filters following secondary activated sludge treatment. Table 2 summarizes the BOD contributions to rivers and streams in the Des Moines River basin above Red Rock Dam.

Animal wastes The State of Iowa is consistently among the nation's leaders in the production of cattle, hogs, poultry, and other livestock. Livestock production in Iowa contributes greatly to the state's economic development, but it also has great potential for polluting the surface water of the state.

Table 2. Point source BOD contributions to rivers and streams in the Des Moines River Basin above Red Rock Dam (61)

Area	Average Daily BOD, lb/day	Per cent of Total
Lower Des Moines River Basin	11,394 (9,733)[a]	56.4 (48.2)[a]
Upper Des Moines River Basin	6,422	31.8
Raccoon River Basin	2,392	11.8
Total in Des Moines River Basin	20,208	100.0

[a]Des Moines metropolitan area.

Based on individual animal estimates for the period from January 1972 to January 1973, the equivalent of 2.4 million cows were on farms in the Des Moines River Basin (36). This figure was estimated by multiplying the number of each kind of animal by the factors given in Table 3 to give the number of equivalent 1000-lb cows.

These conversion factors were estimated from data derived from several sources (23, 36, 61) which compared individual animals and were based on the pollutional waste characteristics of livestock. The factors are equivalent roughly for BOD, phosphorous, and nitrogen. Since the factors were based on ranges in the data for individual nutrients, they should not be considered to be exact.

For comparison, the BOD of the livestock waste produced in the basin is equivalent to a human population of at least 20 million people, far exceeding the human population of 500,000 living in the Des Moines River Basin. The number of livestock and their density in number per

Table 3. Livestock waste production equivalents

Animal	Conversion Factor[a]
All cattle and calves	0.8[b]
Swine	0.4
Poultry	0.02

[a]Number of animals times conversion factor gives waste production equivalent to that of a 1000 lb cow (23, 36, 61).

[b]Confirmed by S. Melvin. Personal communication. Agricultural Engineering Extension, Iowa State University, Ames, Iowa, December 12, 1974.

Table 4. Livestock distribution in the Des Moines River Basin in Iowa - 1972 (34, 65)

River Basin	Numbers of Equivalent cows in thousands	Drainage area sq mi	Density, animals per sq mi
Raccoon River Basin	800	3590	223
Upper Des Moines River Basin	730	4695[a]	155
Lower Des Moines River Basin	744	4217	176
Des Moines River Basin	2274	12502[a]	182

[a]These drainage areas do not include the portion of the upper Des Moines River Basin in Minnesota, about 1550 sq mi

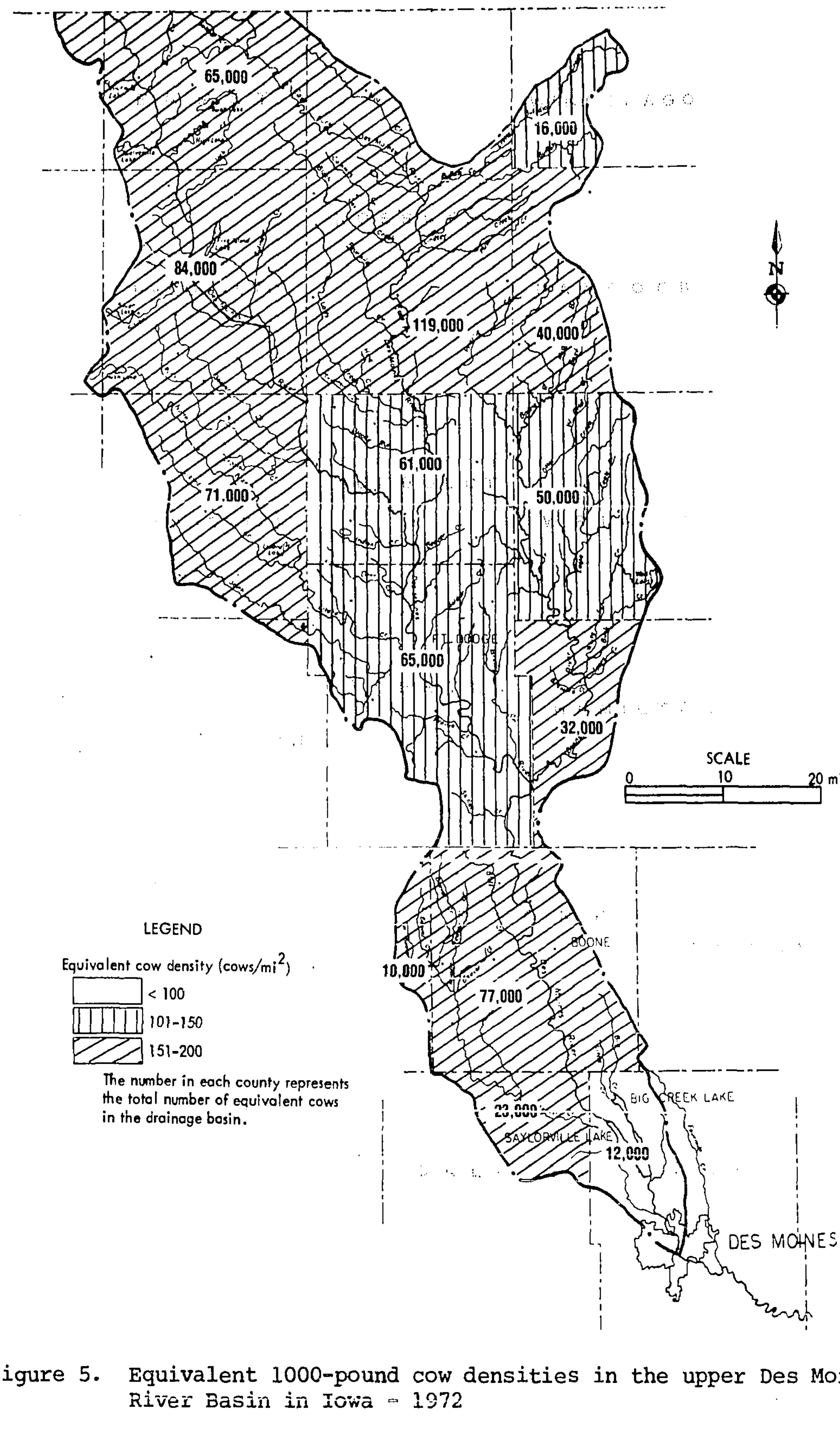


Figure 5. Equivalent 1000-pound cow densities in the upper Des Moines River Basin in Iowa - 1972

square mile varies from one part of the Des Moines River Basin to another, as shown in Table 4. Figure 5 shows pictorially the equivalent 1000-lb cow densities in the upper Des Moines River Basin.

Crop production The principal agricultural activity in the Des Moines River basin involves crop production - primarily that of corn and soybeans. The annual acreage devoted to production of each grain will vary somewhat from year to year, depending on expected market conditions and crop rotation practices. Weather conditions are the principal factor determining yield.

In the last 25 years, it is estimated that corn production has increased from about 60 to 100 bushels per acre (statewide average). As farmers shifted from the use of horses in farming to gasoline and diesel fuel powered farming equipment, the crop production has shifted from oats to soybeans. Average yields for soybeans (statewide average) now approximate 30 to 35 bushels per acre. Figure 6 shows pictorially the total acreage devoted to corn and soybean production in the upper Des Moines basin.

Corn and soybean production represent the most important cash grain crop in the State of Iowa. It does not, however, represent the only use of Iowa farm land. Table 5 summarizes the 1973 crop acreage devoted in Iowa to different farm uses.

Fertilizer application Data concerning fertilizer application by county or river basin are not readily available. The most important crops are corn and soybeans, and most of the fertilizer is used on these

Table 5. Statewide average crop acreage distribution in Iowa - 1971 (35)

Crop	Acreage	Per cent of farmland
Total Farm Land	33,705,189	100.00
Corn (field)	11,883,148	35.26
Soybeans	7,588,192	22.51
Oats	1,244,300	3.69
Sorghum	13,414	0.04
Wheat	26,724	0.08
Rye	2,752	0.008
Timothy seed	1,684	0.005
Red clover seed	2,487	0.007
White corn	9,304	0.03
Popcorn	31,496	0.09
Hay	2,465,313	7.31
Other crops	38,767	0.12
Pasture	6,465,709	19.18
All other land	3,933,948	11.67

crops. In 1968 an average of about 65 pounds of all plant nutrients per acre were applied to midwestern harvested acreage. Nearly 34 per cent of this amount was applied in the fall (43). By 1974, fertilizer application rates had increased to nearly 150 lb per acre for all plant nutrients. About 120 lb per acre of nitrogen and 26 lb per acre of

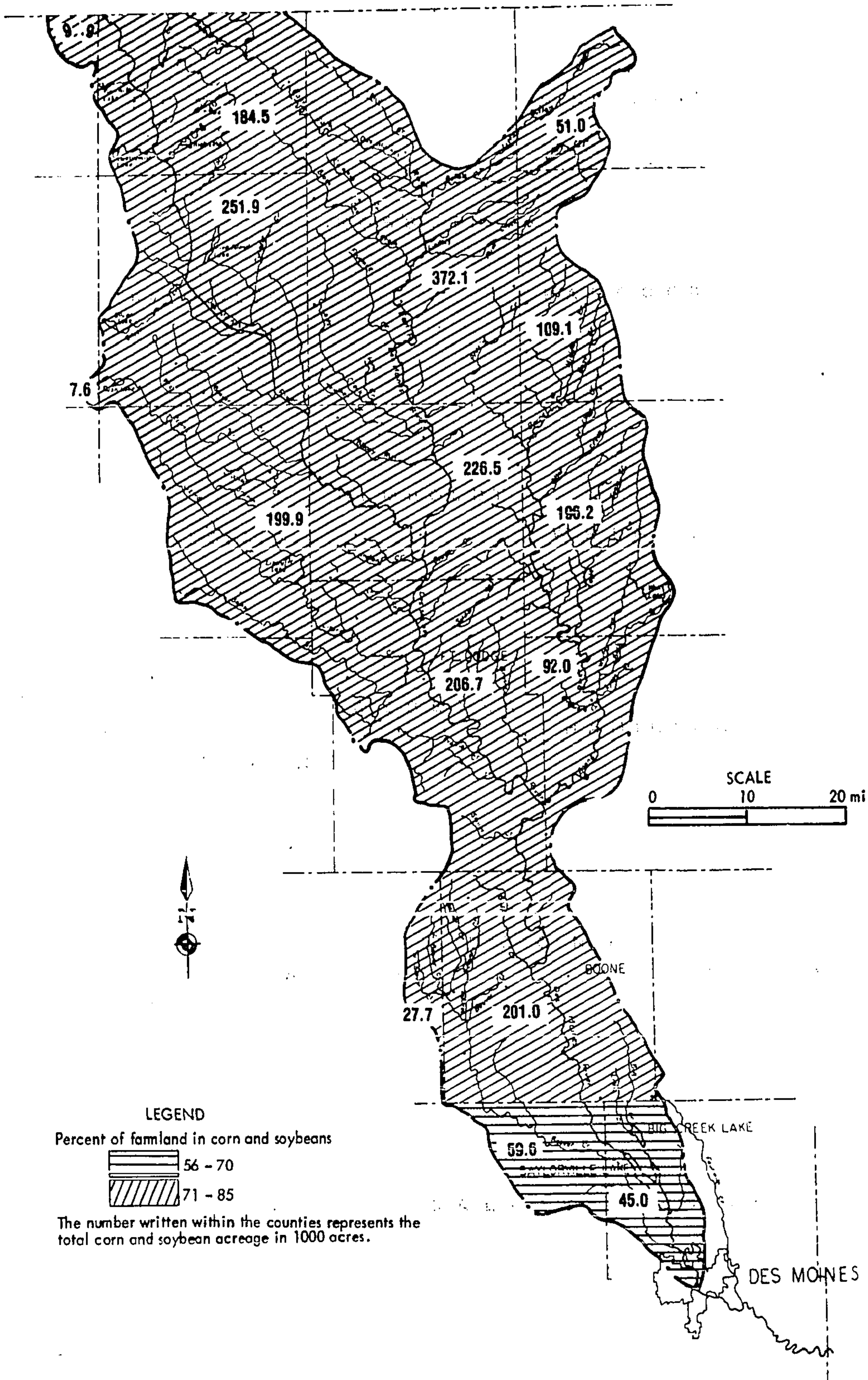


Figure 6. Crop production in the upper Des Moines River Basin - 1973

phosphorous were commonly applied.[1] Fertilizer application rates for corn acreage are much greater than that for soybeans. Because of the nitrogen-fixing quality of soybeans, nitrogen fertilizer is not normally applied. On farms where corn and soybeans are rotated, phosphorus, likewise, may not be applied to the soybean crop because of the carryover from the previous year's corn crop.

Climatology and hydrology

The State of Iowa receives an average of about 32 inches of rain each year. In the Des Moines River Basin, this annual amount varies typically from 28 to 36 inches, with the greater amounts received in the southern part of the basin. During the six-year period of this study, 1967 to 1973, the annual precipitation in the central part of Iowa has averaged 34.26 inches according to the records of the U.S. Weather Service. This area includes most of the upper Des Moines River Basin. The average annual precipitation for the central part of Iowa was 31.36 inches prior to 1965. The range for the 1967 to 1973 period was from 27.59 to 41.82 inches. One of Iowa's greatest assets is the timing of this rainfall. Nearly half of the annual precipitation occurs during the months of May, June, July, and August.

Streamflow in the upper Des Moines River has averaged 1,747 cubic feet per second (cfs) for the 53-years period of record at Stratford, Iowa. At Saylorville, Iowa, the average streamflow is 2,603 cfs, covering a shorter

[1]Voss, R. D. Personal communication. Department of Agronomy, Iowa State University, Ames, Iowa, March 10, 1975.

period of 12 years as shown in Table 6. The much higher streamflow at Saylorville reflects the higher than average precipitation during the past 12 years. During this same period, the streamflow has averaged 2,526 cfs and 2,749 cfs, respectively, at Stratford and Saylorville as given in Table 7.

Table 6. Discharge records in the upper Des Moines River Basin for period of record[a]

	Stratford 05-4813[b]	Saylorville 05-4816.5[b]
Average Annual Flow, cfs	1747	2603
Minimum Daily Flow	17	44
Maximum Flow	57,400 (June 22, 1954)	47,700 (April 10, 1965)

[a]Period of record: Stratford - 1920 to 1973
Saylorville - 1961 to 1973.

[b]U.S.G.S. stream gage number.

Table 7. Discharge records in the upper Des Moines River Basin for period from 1967 to 1973 (65)

	Stratford	Saylorville
Average Annual Flow, cfs	2,526	2,749
Minimum Average Annual Flow	409	466
Maximum Average Annual Flow	4,962	5,175
Minimum Average Monthly Flow	75	75
Maximum Average Monthly Flow	15,770	15,830
Minimum Daily Flow	46	44
Maximum Daily Flow	24,600	23,800

RESEARCH METHODS

Water Quality Sampling and Analysis

Water quality samples have been collected on a weekly basis at several locations along the Des Moines River since July, 1967 as a part of a preimpoundment study sponsored by the U.S. Army Corps of Engineers. The part of the Des Moines River Basin of interest in this study is the upper portion, north of the confluence of Beaver Creek with the Des Moines River (see Figure 2). Some of Iowa's richest farmland is located here, and a large percentage of the basin is used for the production of corn and soybeans.

Saylorville Reservoir, closure scheduled for the summer of 1975, will receive the runoff from the upper Des Moines River Basin. Thus, the limnological character of the water flowing into the reservoir is of considerable interest in regard to the use of the reservoir for recreation. Water quality of the river above the reservoir will also determine to a large extent, in association with the operational mode of the reservoir, water quality in the Des Moines River within the metropolitan area of Des Moines and, to a lesser extent, in Red Rock Reservoir below Des Moines.

The sampling location chosen for this study was located a few miles below Saylorville Dam near Saylorville, Iowa and is labelled as Station 5 in Figure 3. Water quality at this site and at another location near Boone, Iowa (Station 1 in Figure 3) is being monitored on a long term basis (starting in 1967) by personnel of the Sanitary

Engineering Section of the Iowa State University Engineering Research Institute (ERI) under contract with the Crops of Engineers, U.S. Army, Rock Island District.

From 112 to 309 weeks of data were available for analysis, depending on when the analysis of a particular limnological substance was begun and the number of weeks for which data were missing. Complete compilations of these data are contained in the preimpoundment study annual reports (3, 4, 6, 8, 9, 11).

During the period of the study the water samples were analyzed for 20 to 40 parameters. Analyses conformed to procedures described in Standard Methods (53) or in other standard references, as listed in the annual reports.

Statistical Analysis

Statistical relationships were developed for 17 limnological parameters as a function of parameters representative of climatological, hydrological, and seasonal conditions in the upper Des Moines River Basin. The description and definition of each of these parameters are detailed in another place in this thesis. A stepwise regression analysis was used to formulate statistical relationships (19).

In this particular procedure,

> ...the order of insertion (of the independent or explanatory variables) is determined by using the partial correlation coefficient as a measure of the importance of variables not yet in the equation.

During the parameter selection procedure there is

...re-examination at every stage of the regression of the variables incorporated into the model in previous stages. A variable which may have been the best single variable to enter at an early stage, may at a later stage, be superflous because of the relationship between it and other variables now in the regression. To check on this, the partial F criterion for each variable in the regression at any stage of calculation is evaluated and compared with a preselected percentage point of the appropriate F distribution. This provides a judgement on the contribution made by each variable as though it had been the most recent variable entered, irrespective of its actual point of entry into the model. Any variable which provides a nonsignificant contribution is removed from the model. This process is continued until no more variables will be admitted to the equation and no more are rejected (19).

Significant difference between means

In an attempt to explain a greater portion of the variability in water quality, the data were divided into two sets: one for which the water temperature was greater than 10°C and one for which the water temperature was less than or equal to 10°C. The t test was used to determine whether a statistically significant difference existed between the means of each biannual grouping and full year data set (13). The following statistical equation was used.

$$t(df = n_1+n_2-2) = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{n_1s_1^2+n_2s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

where: df = degrees of freedom

n_1 = number of data points for the annual grouping

n_2 = number of data points for biannual grouping

s_1 = standard deviation, annual

s_2 = standard deviation, biannual

$\bar{X}_1$ = mean of data, annual

$\bar{X}_2$ = mean of data, biannual

If the t test indicated no significant difference between the means ($P = 0.05$) when it was felt that a difference should exist, plots of the regression equations were examined for differences in the predicted values as compared to the observed values. A judgement was then made to determine whether the regression equation for annual grouping or the biannual grouping indicated the best relationship.

Missing Data

In some instances it was desirable to have complete data sets for the water quality parameters. To construct complete data sets, missing data values were calculated from regression equations developed from the original data. After the calculated data values had been added, regression equations were again developed. The two equations were compared in terms of the explanatory variables included in the regression and other statistics. It was felt that statistical similarity of the two equations indicated that the characteristics of the original data had not been altered by the inclusion of the calculated values. For most regressions, however, missing data were not of great concern since the missing values were scattered throughout the six years of data and not confined to a particular season or year.

Development of the Research

The broad objective of this research was to develop and evaluate the statistical relationships between water quality in the Des Moines River and climatological, hydrological, and seasonal factors in the upper Des Moines River Basin. The research was completed in two phases. In the first phase, the statistical relationships were developed as a water quality model using a stepwise linear regression procedure. After the results were evaluated it was clear that a second phase of the research was necessary. It was felt that the model could be refined by the addition of new variables. In addition, week to week variability in some water quality data appeared to be improbably great. It was felt that the data could be treated in such a way as to smooth improbably high peaks. Because of the great difference between summer and winter in terms of the aquatic chemistry and biology, weather conditions, and agricultural activities, it was considered that grouping water quality data into summer and winter seasons would allow treatment of the two groups as separate populations.

Phase One

A site on the Des Moines River about ten miles upstream from the of Des Moines and near the lower end of the basin was chosen for the collection of water samples. Water quality at this site was expected to be a function of the climatological, geological, hydrological, and land-use conditions in the upstream basin. Developing explanatory parameters which would be representative of these factors so that they could be used in the statistical analysis was the initial problem.

Some of the factors would be fairly uniform, whereas others would be variable. Basin-wide geological conditions and land-use patterns have not changed greatly in the past decade. Domestic and industrial waste contributions to the river are relatively uniform throughout the year. From personal observation it was felt that seasonal weather patterns, or climatology, and agricultural activities, which varied considerably through the year, would have the greatest impact on water quality. Precipitation and temperature appeared to be the weather conditions of importance because of their relationship to runoff, snowmelt, soil temperature, and water temperature. Although land-use patterns are uniform from year to year (about 95 percent of the basin is in farmland and 80 percent of this is devoted to the production of corn and soybeans), agricultural activities do change seasonally in relatively predictable ways. Fields are plowed, fertilizer is applied, and crops are planted in the spring. Crops sprout and mature during the summer, covering the fields with lush growth. In the fall, crops are harvested and the residue either remains on the fields or is plowed under. Fertilizer may also be applied in the fall (42).

Some solutions to the problem of representing these conditions began to form. While precipitation data for the basin were available, its use was cumbersome, especially since the distribution and intensity of the precipitation varied over the basin's 5841 sq. miles. Another way of representing precipitation intensity and duration is through runoff characteristics. Other than the actual precipitation data, runoff is probably the single best indicator of precipitation. The quantity of surface runoff is reflected by streamflow. Thus, streamflow was considered to be representative of the average quantity and intensity of precipitation falling on the basin. Accurate streamflow records from the recording gage at the sampling site were available on a continuous daily basis.

However, streamflow alone would not tell the whole story. Suspended sediment concentration in the river provides important information regarding the condition of the soil surface. During the spring time, surface runoff is heavily laden with suspended sediment and surface debris. But during the summer when extensive root systems of maturing crops hold the soil in place, surface runoff may contain lower concentrations of suspended solids. In addition, suspended sediment serves as a vehicle for transport of absorbed material to streams. Daily suspended sediment concentrations were available for the river at the sampling site.

Temperature was another climatological factor of interest. Both water temperature and air temperature data were available from the water quality study sampling records. However, it was felt that water temperature would be the more important because of its influence on aquatic life. It was also representative of air temperatures above the freezing point.

Representing agricultural activities posed a somewhat different problem. For example, how can spring field activities be quantified? What numbers should be placed on fertilizer use or planting of soybeans and corn?

Although fertilizer use has increased dramatically in the last 20 years, it was considered that during the period of the study from 1967 to 1973, basin-wide fertilizer use was essentially constant. It was considered that although cropping patterns could vary at any one place in the basin, these patterns would be fairly uniform when considered on a basin-wide scale. Similar reasoning was applied to livestock production. What remained were the seasonal variations in agricultural activities. It was felt that these activities in association with the normal seasonal weather patterns were the principal factors representing the effects of agricultural activities on water quality.

Developing variables representing agricultural activities appeared to be best approached by the use of wave functions having an annual cycle. Although another study had used a linear function in which a time variable was given a number from one to 12 representing the months of the year (60), this approach seemed too simplistic. Use of several sine wave functions of the day of the year seemed to be more appropriate.

Three sine wave functions were developed and were considered to represent seasonal variations. Two of these, one reaching a maximum in the spring and the other in the fall, represented respectively, the sum of spring and fall weather conditions and agricultural activities. The third peaked in late June and was related to the number of hours of sunlight received by the basin.

Since the initial hypothesis was that water quality is a function of the climate, hydrology, and season, other water quality parameters were not included in the statistical analysis. However, from the four basic parameters which were expected to influence water quality, namely: streamflow, suspended sediment concentration, water temperature, and season, were developed a total of 25 variables. The additional variables were functions of streamflow and enabled the representation of streamflow dynamics. Streamflow variables were divided into three components: a time series component, a hydrograph slope component, and an antecedent flow index component. Altogether there were six components: the three streamflow components, suspended sediment concentration, water temperature, and seasonal functions.

There were several advantages in being able to represent water quality by such basic parameters as flow, suspended sediment concentration, temperature and season. Of the great many physical, chemical and biological parameters which may be monitored in a stream, the most frequently measured parameters are flow, suspended sediment concentration, and temperature because of their usefulness for flood forecasting, soil loss computations, and estimation of sedimentation rates of reservoirs. Because of the widespread monitoring of these parameters, statistical relationships developed for one part of the basin could probably be applied to other parts of the basin. Basins could be compared using the statistical relationships developed, providing a type of index of the factors influencing water quality.

Statistical relationships between water quality parameters and the water quality determining factors were developed using the stepwise linear regression procedure described by Draper and Smith (19). The results of the statistical analysis from Phase One of the research provided useful information regarding the water quality determining factors.

> The variables developed were useful in explaining some of the variability in the water quality.
>
> In general, only one variable from each of the components was entered into the regression at the level of statistical significance selected.
>
> As indicated by the F test, the statistical relationships described by the regression equation were highly significant, generally at the 0.005 level or better. That is, the variables selected appeared to be statistically valid.
>
> While up to 60 percent of the variance could be accounted for in some parameters, the average was about 36 percent. It was felt that this proportion was too low to be of any real usefulness.

In addition, the water quality data were carefully examined for week to week variability in terms of both concentration and quantity. Possible causes of this variability were considered. It was discovered that for some parameters, week to week variability was improbably high. Some of this variation was felt to be artificially introduced as a result of the sampling procedure or as a result of normal variations of precision and accuracy of the analyses of water samples during the six-year water quality study period. There also was evidence that the statistical relationships of a number of the water quality substances were distinctly different during the winter as contrasted with the summer. It was felt that the reason for this was the different sources of the streamflow. During the summer, the source of the streamflow is, generally, runoff, shallow groundwater, and tile drainage. During the winter, runoff periods are infrequent and most of the streamflow comes from groundwater.

This portion of the study constituted the first phase of the research effort. During the second phase of the research, changes were made in the regression model and in the treatment of the data so as to incorporate knowledge from the first phase.

Phase Two

The first change made was in the treatment of the data. Data for the water quality parameters were divided into two groups based on water temperature. Earlier investigations of variation in plankton populations with streamflow indicated that there was a distinct difference for water temperatures above 10°C as contrasted with temperatures less than 10°C.

Although the percentage variance accounted for (R^2) was improved in most cases for either the cold or the warm weather data, the mean R^2 for the cold wetaher regression was only 43.3 percent and that for the warm weather was 38.4 percent. It was felt that the variance not accounted for was still too great.

Plots were drawn of the residual difference (the observed less the predicted value). Ideally this residual should equal zero, but in most cases the residuals averaged 20 to 40 per cent of the mean. Most interesting, however, were a few individual residuals which exceeded the mean by 3 to 4 times the standard error. Data associated with these outlying residuals were examined. In some cases, the data were found to be mistakes, but for others there was no evidence to indicate they were not the measured value. The mistakes were either corrected or the data were omitted, depending on the situation. Some of the data, primarily that for turbidity, were associated with heavy runoff events such as would occur with simultaneous snowmelt and rainfall in the spring. Regressions were rerun, omitting the outliers. The percentage variance accounted for improved, sometimes dramatically, for some of the limnological parameters, but for others the R^2 value decreased. This revision also revealed some interesting statistical relationships not evident from the regressions using the original data. However, for most parameters the low R^2 could not be attributed to outliers.

In an attempt to reduce random variation, the week-to-week values were smoothed using an averaging technique. This was an attempt to smooth the unexplainable high or low values which had been discovered during the first phase of the research of the data and for which there was no reason to reject. Of particular interest was the parameter chloride. Chloride is a conservative substance in that it is not changed to other forms in the aquatic environment as a result of chemical or biological reactions. No logical reason could be found for the large fluctuations of the quantity of chloride in the river, particularly during periods when there was little change in flow. For this reason, it was felt that the data smoothing procedure was valid and would probably be more representative of the actual variation of the data. It was noted that the water quality determining variables introduced into the regression did not change. This was interpreted to mean that the smoothing routine did not change the general character of the information contained in the original data.

At this point the variables representing climatological, hydrological and seasonal conditions in the basin were reviewed. One aspect which was missing were variables representing quantity inputs of the water quality materials. Two situations could effectively describe the quantity inputs: they could be variable or they could be essentially uniform on constant. Variable inputs would probably represent agricultural activity. Constant inputs, on the other hand, might be related to groundwater contributions to the river.

After consideration of the patterns of the variable inputs, it was felt that they would best be described by the seasonal parameters. Since the variable inputs were considered to be quantity or weight, and the relative input quantity to be described by the seasonal variables, concentration would be proportional to the value of the seasonal variable divided by the streamflow. The observed concentration resulting from a constant input was represented by a constant divided by the streamflow.

SELECTION AND INTERPRETATION OF PARAMETERS

Introduction

The principal objective of this research was to investigate the statistical relationships of various water quality parameters to a relatively small number of explanatory parameters which would be indicative of climatological and hydrological conditions, and the seasonal agricultural activities within upper Des Moines River Basin. The explanatory parameters were considered to have four basic components: streamflow, suspended sediment concentration in the river, river water temperature, and seasonal variations. Limnological parameters in the statistical analysis included a selection of physical, chemical and biological substances which would be important in the interpretation of variations in water quality.

Explanatory Parameters

Several constraints were placed on the desired explanatory parameters.

Data for the parameters must be readily accessible for most Iowa rivers.

The data should be available on a continuous daily basis.

The data should reflect principally climatological and hydrological conditions and agricultural activities in the basin as opposed to the inclusion of actual water quality parameters such as dissolved oxygen or nutrient concentration.

Consideration of parameters which would conform to these constraints led to the selection of five fundamental variables: streamflow, suspended sediment, water temperature, specific conductivity, and season. The inclusion of specific conductivity was eventually abandoned because of lack of continuity in the data. The other parameters conformed to the constraints. Their relationship to climatological conditions and agricultural activities is proposed as illustrated in Figure 7.

Many researchers refer to variables used in statistical analysis as dependent and independent. Frequently, however, so-called independent variables are highly correlated, not independent. This is especially true in this research since streamflows on several consecutive days are used individually as independent variables. Variables of this nature were considered to be explanatory in that they explained or accounted for a certain proportion of the variability in the data for a water quality parameter. Although these explanatory variables were not completely independent, attempts were made statistically to select variables which both explained variability of a dependent variable and were not highly correlated. Causal inferences were cautiously drawn because, in some instances, the relationships of the explanatory variables were secondary to a primary variable.

A hierarchy of the relationships of explanatory variables can be developed and is suggested in Figure 7. It is possible that additional variables exist which are highly related to water quality and, if properly interpreted, would explain all the variability in water quality. However, until those variables are discovered, we must be content to use lower order variables which are, more or less, functions of the variables above them. In Figure 7, season was considered

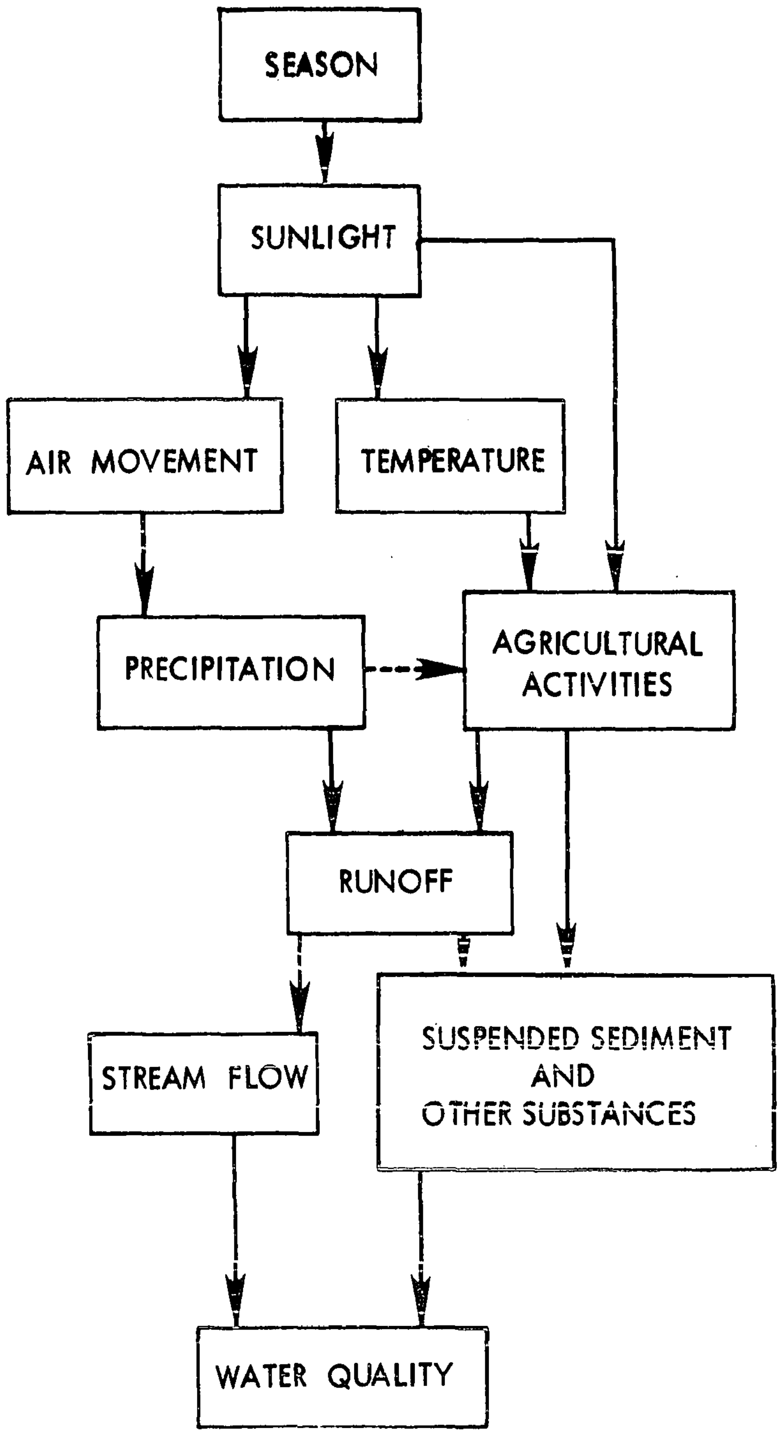


Figure 7. Interrelationships of climatological and hydrologic conditions, agricultural activities, and water quality

to be a primary variable, not so much in the sense of the four seasons, but in terms of the path of the earth around the sun in association with the tilt of the earth's axis relative to its plane of revolution about the sun. In this sense, the other variables are a function of season because of the effects of the sun on the warming and cooling of the earth. Many climatological changes are highly related to seasonal variations in temperature. These changes in association with the activities of man are believed to serve as the stimuli which result in a given water quality condition.

Discussion of explanatory parameters

Streamflow Streamflow is a function of a great many factors, some climatological and others related to the activities of man such as removal of vegetative cover. In general, the principal association is with precipitation. However, streamflow is also related to soil moisture. Following a dry period, a heavy rain may cause only a moderate increase in streamflow because most of the moisture is absorbed by the soil and little remains to run off. Typically this situation occurs during the autumn and contrasts markedly with spring and early summer conditions when much of the precipitation runs off wet soils.

Streamflow varies in relatively predictable ways during the year. During the winter, little runoff occurs and the flow is typically very low. In the spring snowmelt and rainfall runoff combine to produce very high flows which continue through the summer. By early autumn, the amount of rainfall received in the basin decreases, resulting in less runoff and lower streamflow.

In an attempt to quantify the annual pattern in streamflow variations, a plot of the six-year average of the mean daily streamflow was developed from records of the United States Geological Survey (65). For example, the mean daily streamflow on January 1 for the years 1968 to 1973 were averaged, and this average is one point on the graph in Figure 8. A smooth curve was then drawn through the points. The precise location of the curve was somewhat subjective, especially for the months of March through July, because of the great variation in streamflow during the six-year period. The overall trend of the curve, however, was determined to be representative of the typical variations in streamflow throughout the year.

Vegetative cover also influences streamflow. Extensive vegetative cover intercepts precipitation and allows time for evaporation of moisture and permits transpiration. It also extends the time for absorption of moisture by the soil. During periods of low soil moisture, runoff is delayed, smoothing the peaks of the stream hydrograph. To this extent, streamflow is a composite variable and representative of soil moisture, precipitation and vegetative cover.

In this context, streamflow was expected to be related to a number of limnological substances. Surface runoff conveys numerous dissolved and suspended materials to rivers. During the spring when soil moisture is high and little vegetative cover exists, sediment and surface debris account for much of these transported materials. At other times,

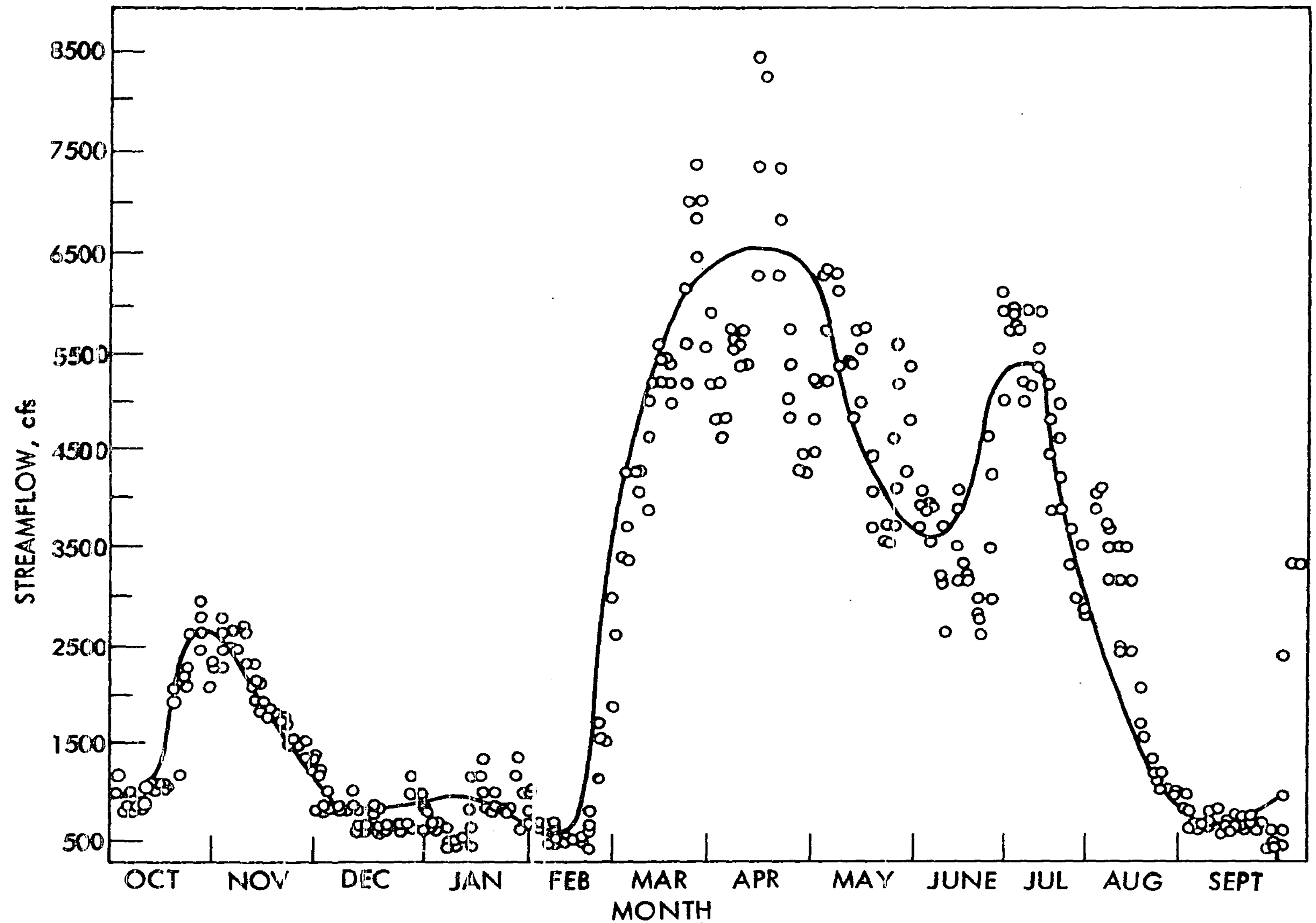


Figure 8. Smoothed average daily streamflow, Des Moines River at Saylorville, 1967-1973

when the soil is covered by extensive vegetation, a large proportion of the rainfall is absorbed and may eventually enter streams as groundwater and tile drainage. The time of the entrance of this water containing dissolved materials to streams will lag the occurrence of the rain storm. Thus, the limnological character of the stream is a function of not only streamflow, but of the timing of streamflow relative to the sampling date.

<u>Streamflow dynamics</u> Streamflow was considered to have three components: magnitude, change in flow with time, and time of occurrence relative to a given date. The magnitude was viewed in several ways. It was considered as the flow rate in cubic feet per second (cfs), the mean flow of the week prior to the sampling date, and the standardized flow. Standardized streamflows were calculated by dividing the streamflow on the sampling date by the mean flow for the six-year period of the water quality study, 2749 cfs. Averaging flows over seven days was an attempt to include information regarding streamflow over a longer period of time, in effect smoothing the hydrograph. Standardized flows were included in an attempt to incorporate information regarding the streamflow on the sampling date relative to the long-term average.

Determining the effects of runoff on the limnological character of the river was one of the objectives of the research. Therefore, it was of particular concern that parameters be developed which would accurately represent various types of runoff conditions. Runoff was related to streamflow dynamics.

Two methods were used to incorporate dynamic information. One method was to calculate the slope of the hydrograph on the sampling date. The difference in flow over a three-day and a five-day period, the sampling date being the mid-point of the interval, was divided by the flow on the sampling date as shown.

$$\text{Slope} = (Q_j - Q_k)/Q_i$$

or

$$DQn/Q$$

where

Q = streamflow, cfs

i = sampling day

n = length of interval, days

$$j = \left(\frac{n-1}{2}\right)\ (i+1)$$

$$k = \left(\frac{n-1}{2}\right)\ (i-1)$$

The five-day interval was eventually dropped as an explanatory variable because it provided little additional information regarding streamflow dynamics. The standardized slope variable indicated the relative magnitude of the increase or decrease in flow with respect to the magnitude of streamflow on the sampling date.

Another method used to incorporate dynamic information was to develop a ratio of streamflow on the sampling date with the average flow over various intervals from two days to 28 days prior to the sampling date. These variables were considered to be a relative antecedent flow index (RAFI).

$$\text{RAFI} = Q_i/\left(\sum_{k=1}^{n} Q_k\right)/n$$

where

Q_i = streamflow on the sampling ith day

Q_k = streamflow on the kth day prior to the sampling day

n = antecedent period

Over the shorter periods, the value was expected to be indicative of immediate changes in flow such as a runoff event. For longer periods, information of a somewhat different nature was sought. For example, if the flow on the sampling date was large relative to the mean flow of the previous 28 days, it could be indicative of rainfall and runoff following a dry period or a low-flow period such as would occur during the winter months when the river was ice covered.

Relationship of concentration to streamflow hydrograph

Streamflows on days other than the sampling date were included as explanatory variables. These values covered the period from five days prior to four days following the sampling date. The rationale for this procedure was that some water quality materials would correlate better with streamflow on days other than the sampling date because of a natural time variance of the substance reaching the stream from the soil surface or through the soil profile dissolved in the groundwater flow to the stream following precipitation. Several examples are shown in Figure 9.

If streamflow is related to changes in concentration of water quality parameter, three situations may occur. The maximum (or minimum) concentration of the substance may precede the hydrograph peak, it may

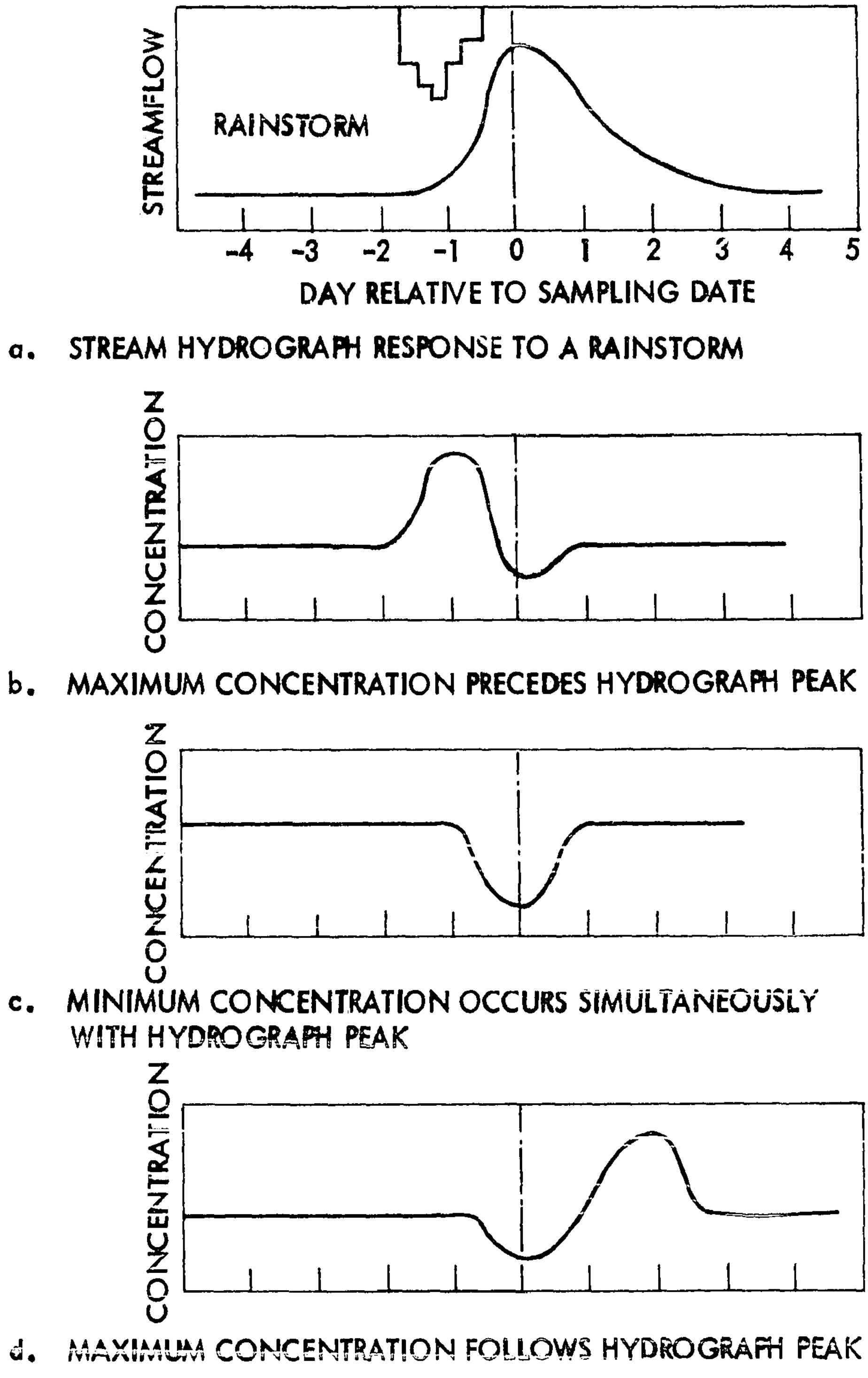


Figure 9. Time of occurrence of maximum or minimum concentration relative to hydrograph peak

occur simultaneously with the peak, or it may occur after the hydrograph peak. These three situations are illustrated in Figure 9.

One situation which could result in concentration increases preceding the hydrograph peak, as illustrated in Figure 9b would be the rapid translocation substances held loosely to the soil surface. These materials would be suspended or dissolved in the runoff water first entering the stream. The concentrations of the material in the stream would be more highly correlated with the time of the rainstorm than with the time of the hydrograph peak.

As an example, livestock waste materials are not tightly bound to the soil surface of a feedlot and would be expected to be washed rapidly into nearby streams by surface runoff.

Although it is difficult to make generalizations concerning feedlot runoff because of the variations for individual lots, important factors include stocking rates, lot slope, feed rations, depth of manure pack, evaporation rate, and antecedent moisture conditions. In general, one-half inch of rainfall is necessary to produce any runoff (28). When this one-half inch or more falls within a twenty four hour period, it almost invariably produces runoff (28). In the midwest, it has been found that when more than one half inch of rainfall occurs, 30 to 75 per cent will runoff.

A common Iowa practice has been to locate feedlots on slopes adjacent to streams because of the ease of disposal of wastes with the first rain. Materials held loosely in place are rapidly washed into the adjacent stream during a rainstorm. The concentration of the material in the stream would initially increase more rapidly than the rate of streamflow and would precede the greatest portion of the runoff flow. This would result in a stronger relationship between the concentration of the material in the stream and the flow prior to the hydrograph peak.

The second situation, one for which a concentration change would coincide with the hydrograph peak as illustrated in Figure 9c, would occur if the change in concentration were primarily due to dilution. In this case the concentration would decrease to a minimum. This might occur for a conservative substance such as chloride which enters the stream at a relatively constant rate. As the flow increases the concentration would decrease due to dilution by the runoff flow.

In the third situation, maximum concentration of the substance occurs after the hydrograph peak has passed as illustrated in Figure 9d. Intuitively, this might be the result for substances which flow to streams dissolved in groundwater. Precipitation would percolate into the soil, eventually reaching the interflow and tile drains. Interflow would move slowly towards the stream transporting the dissolved materials. Sometime later, after the hydrograph peak had passed, the interflow or tile drain effluents with the dissolved substances would enter the stream. Only at this later time would the increase in concentration be evident.

Nitrates, for example, are thought to reach the stream dissolved in groundwater flowing through the soil or through drainage tiles (32). Nitrification of ammonia to nitrates occurs in the soil. Rainfall percolating through this soil will dissolve the highly soluble nitrates and travel to nearby streams through tiles or as groundwater interflow. In one area it was shown that a groundwater contributed 52 per cent of the total nitrogen to a lake (33). A similar relationship probably exists also for rivers. Because of the relatively slow movement of water through the soil to the streams, it was expected that the greatest quantity of nitrates would reach streams at a time following a rainfall event, after the peak of the hydrograph.

Transformation of streamflow Preliminary analysis of the data indicated that the natural logarithm of flow correlated better with water quality variables than the untransformed flow value. As a result, the logarithm of the magnitude of streamflow was substituted. Where ratios were calculated, untransformed values of the magnitude of streamflow were used.

Summary Because of the anticipated importance of streamflow in regard to water quality, 20 different variables were initially derived from streamflow data and are listed in Table 8. These variables can be divided into two main groups and one sub-group.

The magnitude of the streamflow on various dates with respect to the sampling dates.

The magnitude of streamflow relative to the mean streamflow for varying periods prior to the sampling date.

> The change in the magnitude of the streamflow or slope relative to the magnitude of the streamflow on the sampling date.

Table 8. Interpretation of streamflow parameters

Parameter	Interpretation
Time series component:	
1. lnQ+4 2. lnQ+2 3. lnQ+1 4. lnQ 5. lnQ-1 6. lnQ-2[a] 7. lnQ-3 8. lnQ-5	$\ln Q \pm n$ where lnQ = natural logarithm of the mean daily streamflow n days from the sampling date
9. lnQA7	lnQA7 where QA7 = mean daily streamflow for the week prior to the sampling date. The natural logarithm is taken of the result
10. QSTD	$QSTD = Q_i/\overline{Q}$ where Q_i = the streamflow on the sampling date $\overline{Q}$ = the mean streamflow (2749 cfs) for the period of this study
Hydrograph slope component:	
11. DQ3/Q	$DQn/Q = (Q_j - Q_k)/Q_i$ where: Q = streamflow, cfs i = sampling day n = length of interval, days $j - (\frac{n-1}{2})(i+1)$ $k = (\frac{n-1}{2})(i-1)$
12. DQ5/Q[a]	

[a]These variables were eventually eliminated because they provided little additional information.

Table 8 (Continued)

Parameter	Interpretation
Relative antecendent flow index component:	
13. Q/QA2	$Q/QAn = Q_i/(\sum_{k=1}^{n} Q_k)/n$
14. Q/QA3[a]	where:
15. Q/QA4	
16. Q/QA5	Q_i = streamflow on the sampling day
17. Q/QA7	
18. Q/QA14	Q_k = streamflow on the kth day prior to the sampling day
19. Q/QA21	
20. Q/QA28	n = antecedent period

<u>Suspended sediment</u> Suspended sediment concentration in the river was expected to be important as an explanatory variable because of its relationship with the condition of the soil surface, the soil moisture, and the soil temperature.

The tenacity with which the soil is held in place and resists erosion is related to the vegetative cover slope and the temperature of this soil. In the early spring before the soil thaws, snowmelt runoff may not cause greatly elevated concentrations of suspended sediment because most of the soil is frozen firmly in place (31). On the other hand, during the summer when crops are maturing, the soil surface is again held in place by extensive root systems. At this time, however, other factors are also important in the prevention of translocation of soil, as were discussed in the section of streamflow. Greatest amounts of soil loss due to rainfall runoff would be expected to occur in the late spring following plowing and planting of crops because the soil is neither held in place by root systems nor because of frozen soil conditions.

Another soil condition important to the resistance of soil to translocation is soil moisture. Under similar conditions of rainfall, vegetative cover, and soil temperature a soil with a low moisture content will have a greater capacity to absorb precipitation. In this sense, the soil will be more resistant to translocation due to runoff forces than a similar soil under similar groundcover conditions, but with a high moisture content.

As a result, suspended sediment concentrations in the river were expected to be an important indicator of soil moisture and vegetative cover.

Suspended sediment was also expected to be of importance as an explanatory variable because of the substances which would be absorbed onto its surface. Of particular interest was phosphate which tend to bind chemically to clays. One study at an experimental farm in Indiana indicated that available phosphorus was present to the extent of about 405 ppm in sediment lost through erosion (62). Another study indicated that nutrients were eroded selectively (44). The concentrations of nutrients in the translocated soil was higher than that in the surface soil from which the sediment was derived. The conclusion here was that the most easily eroded soil is the richest in nutrients and has the greatest potential for pollution.

<u>Transformation of suspended sediment</u> Preliminary analysis of the data indicated that the natural logarithm of suspended sediment concentration correlated better with water quality variables than the untransformed suspended sediment value. Another study supported this transformation (2). As a result, the natural logarithm of suspended sediment was used in the statistical analyses.

<u>Summary</u> The concentration of suspended sediment in a stream is indicative of a variety of soil conditions related to soil moisture, soil temperature, and vegetative cover. Nutrient loss is also related to sediment loss because of adsorption of nutrients onto the sediment.

<u>Temperature</u> Water temperature is a key factor in the ecosystem of a river. The growth and activity of all forms of life, including bacteria, algae, benthic organisms, and fish are closely related to the temperature of their aquatic environment. Chemical parameters such as the solubility of dissolved gases and the kinetics of many reactions that occur in the river are also temperature dependent.

Water temperature is related to air temperatures above freezing. To this extent it is related to terrestrial seasonal changes such as snowmelt and the growth of vegetation. Agricultural activities in the basin are seasonally dependent, although not strictly a function of temperature.

Because of the diverse relationships of temperature to many of the materials related to water quality, temperature was believed to be a parameter which would be useful in the explanation of variation in water quality.

Seasonal effects   Agricultural activities are necessarily seasonal in character. Crops must be planted in the spring and harvested in the fall. Because 95 per cent of the basin is in farmland and because agricultural activities occur regularly from year to year, it is important to understand the effects of these seasonal agricultural activities on water quality.

During the spring, nearly all precipitation reaches recently plowed or cultivated soil surfaces because of the absence of vegetation. Consequently, transport of soil and other surface debris suspended in surface runoff is often greatest at this time. Also occurring in the late spring are general snowmelt and the spring rains. When these two situations occur simultaneously, soil loss is high and the result is very turbid streams. Temperature effects are important at this time because of the transition from cold weather biological forms to warm weather forms. All of these factors combine to cause great changes in water quality during the spring months as contrasted with the changes observed during the winter and fall months.

However, by late summer, when crops are mature, a portion of the precipitation will be intercepted by the vegetation. The arrival of the rain at the soil surface may be delayed, or prevented altogether because of evapotranspiration. Soil and other materials are more tightly bound by vegetation. Thus, the limnological character of the runoff water should be greatly different in the late summer than in the spring.

In the autumn, after the harvest, crop residues remain on the field. Vegetative growth is often slowed because of dry conditions, shorter days, and lower temperatures. Although soil is bound to a lesser extent by vegetation, crop residues and dry soil conditions tend to reduce runoff and soil loss because of greater moisture retention. At this time, much of the fertilizer applied in the spring has been removed through crop harvesting or washed off the soil during previous runoff periods. Consequently, runoff during the late fall is expected to have a lower nutrient concentration than would be found in either the spring or the summer runoff.

As a result, seasonal agricultural activities in association with seasonal climatological patterns were expected to be strongly related to the observed changes in water quality.

Consultation with personnel of the Agricultural Engineering Department at Iowa State University indicated that although the exact dates of agricultural activity consisting of plowing, applying fertilizer and pesticides, and planting varied from year to year depending on weather conditions, spring activities peaked in late April and fall activities peaked in late September.

In order to construct an explanatory variable which would simulate seasonal changes, the day of the year (January 1 set equal to one) was transformed through a sine wave function. This sine function was adjusted to produce either a spring maximum or a fall maximum. The exact dates chosen for computation of the sine function maxima were April 20 and September 20.

Additionally, a third season variable was constructed to be representative of the amount of solar radiation received by the basin. June 21, the summer solstice, was selected as the maximum to represent the longest day (greatest number of sunlight hours) of the year. It was considered that the sunlight variable would be related to plant growth or to other conditions peaking at this time.

It was anticipated that the minimum of these functions occurring six months later would be important because of the relationship to water quality minima or inversely related to water quality maxima. The methods of calculation of the seasonal variables are shown in Table 9, and are plotted as a function of the day of the year in Figure 10.

Interpretation of seasonal variables Seasonal inputs may be treated in two different ways. They may be considered to be a function of the concentration of the water quality substance in the river at any time or they may be considered to be a function of the total quantity or mass of the substance in the river at any time. No modification of the seasonal sine wave functions is necessary for the representation of the concentration of a material in the river. However, if the variable is considered to be representative of the total quantity of a material in the river, a function must be developed which would be representative of concentration. The total quantity of a material in

Table 9. Calculation of the values of seasonal parameters

Parameter	Date of Maximum/Minimum	Values of i and n[a]
1. SUN	June 20/ December 20	for $i \geq 82$, $n = -81$ for $i < 82$, $n = 284$
2. AGS	April 21/ October 21	$i \geq 21$, $n = -20$ $i < 21$, $n = 345$
3. AGF	September 21/ March 22	$i \geq 173$, $n = -172$ $i < 173$, $n = 193$

[a]General formula:

$$\sin\left[\frac{i+n}{365.25} \times 2\pi\right] + 1$$

where: i = the day of the year of the sampling date

n = a value which shifts the maximum of the sine function to the desired date.

the river M_T, is equal to the streamflow, Q, times the concentration, C, time a conversion constant, k.

$$M_T = QCk$$

The concentration would be represented

$$C = M_T/Qk$$

For some parameters one procedure may be preferable to the other. In the case of a substance such as chloride, the seasonal variable may be more useful as a quantity variable. It may be considered that

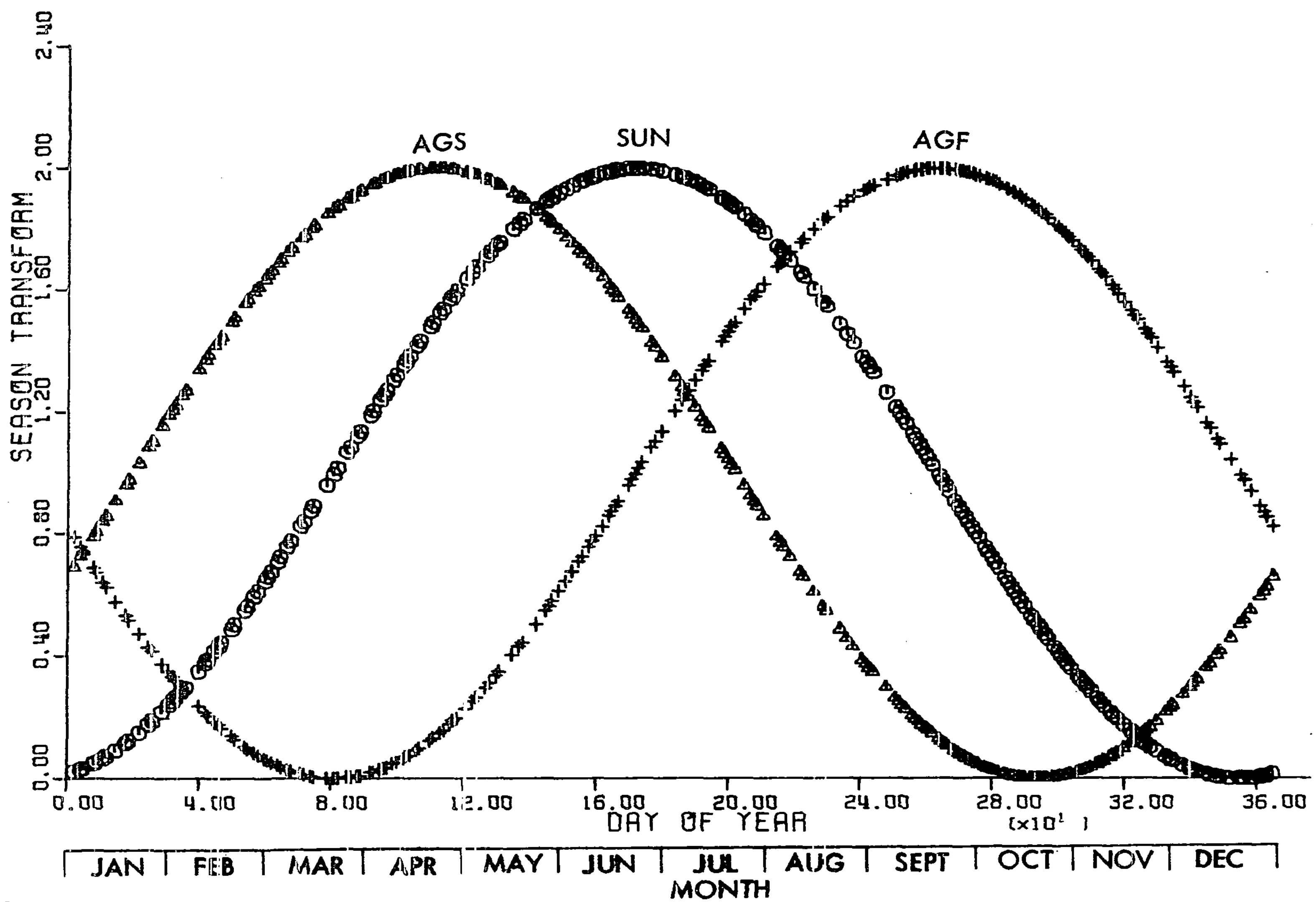


Figure 10. Season variables

winter applications of salt onto streets and highways in the basin results in the introduction of a certain amount of chloride into streams each year. The total quantity, M_T, of chloride in the stream would be a function of the background quantity from groundwater and other continuous inputs, M_C, plus the intermittent inputs from snow-melt, runoff, M_I.

$$M_T = M_C + M_I$$

The observed concentration in mg/l, C, would be a function of the total quantity in pounds per day, streamflow in cfs, and a conversion constant, k.

$$C = k\ M_T/Q \qquad k = 0.1862$$

To use seasonal parameters as variables indicative of the quantity of the water quality materials present in the river during various seasons, the value of the sine function was multiplied by 10,000 and divided by the streamflow value on the sampling date. It was also considered that for some water quality substances, a relatively constant amount entered the stream. The concentration would be determined by the extent of the dilution provided by the streamflow. In this case, an inverse relationship with streamflow would describe the concentration. The value of this parameter was taken as 100,000 divided by the streamflow on a given sampling date. The values were multiplied by large numbers to provide numbers greater than one.

The parameters representing seasonal concentration changes in the river, and seasonal and constant quantities in the stream are listed in Table 10.

Table 10. Seasonal parameters

Parameter	Description
SUN	Sine function of the day of the year to represent a maximum concentration on June 21.
AGS	Sine function of the day of the year to represent a maximum concentration on April 20.
AGF	Sine function of the day of the year to represent a maximum concentration on September 20.
SUNC	SUN x 10,000/Q[a]
AGSC	AGS x 10,000/Q
AGFC	AGF x 10,000/Q
M[b]/Q	100,000/Q

[a]Q denotes streamflow on sampling date.

[b]M represents a constant quantity of the material in the river.

<u>Summary of the interpretation of explanatory parameters</u> Water quality in the Des Moines River can be considered to be a function of four basic factors which are representative of climatological and hydrological changes and seasonal agricultural activities in the basin. Streamflow is related to precipitation over the basin, to surface runoff which washes the soil surface and transports dissolved and suspended materials to the river, and to groundwater flow which carries dissolved

materials to the river. While runoff adds surface materials to the river, it may also dilute the concentration of these materials dissolved and suspended in the river.

Suspended sediment concentration in the river relates in important ways to the condition of the ground surface. Some types of materials tend to be adsorbed onto the surface of soil particles. River water temperature determines the rate at which biological and chemical reactions occur within the aquatic ecosystem. Temperature is also related to certain seasonal terrestrial activities. Seasonal variations in water quality may be related to the seasonal agricultural activities in a highly agriculturally oriented region. The seasonal variables may be considered to be either concentration functions or quantity functions. A list of explanatory parameters used in the statistical analysis is given in Table 11. Some of the explanatory parameters included initially were omitted from this list because they contributed little additional information.

Water Quality Parameters

A variety of physical, chemical, and biological parameters were selected for the statistical analysis of their relationships with the explanatory variables representative of climatological and hydrological conditions, and seasonal activities within the basin. Briefly, these included turbidity, chloride, silica, two hardness parameters, three oxygen related parameters, three nitrogen parameters, two phosphorous parameters, three plankton parameters, and fecal coliform. A complete

Table 11. Explanatory parameters used in statistical analysis

Parameter Name	Description
1. lnQ + 4	Parameters 1-7 indicate natural logarithm of streamflow on dates relative to the sampling date.[a]
2. lnQ + 2	
3. lnQ + 1	
4. lnQ	
5. lnQ - 1	
6. lnQ - 3	
7. lnQ - 5	
8. lnQA7	Natural logarithm of the mean flow for the seven days prior to the sampling date of the study.
9. QSTD	Streamflow on the sampling date divided by the mean flow for the period.
10. DQ3/Q	Change in streamflow over a three-day period divided by the flow on the sampling date.
11. Q/QA2	
12. Q/QA4	
13. Q/QA7	Parameters 11-16 indicate streamflow on the sampling date divided by the mean flow of periods from two to 28 days prior to the sampling date.
14. Q/QA14	
15. Q/QA21	
16. Q/QA28	
17. TEMP	River water temperature, degrees Celsius

[a]Streamflow parameters are discussed more fully in Table 8.

Table 11 (Continued)

Parameter Name	Description
18. lnSED	Natural logarithm of suspended sediment concentration on the sampling date.
	Parameters 19-21 are sine functions of the day of the year producing a maximum on the date shown.[b]
19. SUN	June 21 Concentration function
20. AGS	April 20 Concentration function
21. AGF	Sept. 20 Concentration function
22. SUNC	SUN x 10,000/Q Variable quantity function
23. AGSC	AGS x 10,000/Q Variable quantity function
24. AGFC	AGF x 10,000/Q Variable quantity function
25. M/Q	100,000/Q Constant quantity function

[b]Seasonal parameters are discussed more fully in Tables 9 and 10.

list of these parameters and their means, standard deviations, and range is given in Table 12. These descriptive data apply to the period of the water quality study from July 6, 1967 through July 27, 1973 (3, 4, 6, 8, 9, 11). For the three divisions that are listed in Table 12, annual applies to the complete data set for the entire year, as opposed to the warm and cold season grouping. Warm season and cold season groupings are based on observed limnological differences in the river during, respectively, warm water periods when the water temperature is greater than 10°C and cold water periods when it is 10°C or

Table 12. Description of water quality parameters[a]

Parameter	Annual Mean	Annual Standard Deviation	Annual Range	Warm Season Mean	Warm Season Standard Deviation	Cold Season Mean	Cold Season Standard Deviation
1. Turbidity, JTU	39.8	40.9	0-480	51.5	46.3	25.2	26.7
2. Chloride, mg/l	27.3	12.3	7.1-67.7	26.5	11.5	27.8	12.5
3. Silica, mg/l	13.3	7.87	0-35.1	11.7	7.6	15.2	7.4
4. Total hardness, mg/l as $CaCO_3$	348	93	152-676	311	53	395	110
5. Calcium, mg/l as $CaCO_3$	228	78	82-440	197	59	269	81
6. Dissolved oxygen (DO), mg/l	11.1	3.55	4.5-28.9	9.53	2.71	13.0	3.54
7. Biochemical oxygen demand (BOD), mg/l	8.75	5.76	0.5-30.2	9.86	5.52	7.43	5.76
8. Chemical oxygen demand (BOD), mg/l	37.2	20.6	1.0-136.3	44.5	21.1	27.8	15.6
9. Ammonia, mg/l as N	0.34	0.32	0-2.49	0.27	0.24	0.44	0.38
10. Organic nitrogen, mg/l as N	0.90	0.90	0-8.46	1.12	0.98	0.61	0.68
11. Nitrate, mg/l as N	4.53	3.73	0-13.3	4.01	4.03	4.77	3.28

[a]Applies to Des Moines River water samples collected near Saylorville, Iowa, July 6, 1967 to July 27, 1973.

Table 12 (Continued)

Parameter	Annual			Warm Season		Cold Season	
	Mean	Standard Deviation	Range	Mean	Standard Deviation	Mean	Standard Deviation
12. Total phosphorus, mg/l as PO_4	1.29	0.80	0.7-5.8	1.17	0.65	1.43	0.97
13. Ortho phosphorus, mg/l as PO_4	0.38	0.36	0-2.06	0.21	0.19	0.60	0.40
14. Fecal coliform, no./100 ml	811	2233	0-20,000	1068	2900	471	573
15. Total plankton, no./0.01 ml	314	396	3-1986	468	440	122	205
16. Diatoms, no. / 0.01 ml	257	353	1-1938	378	401	104	191
17. Flagellates, no./ ml	2100	2304	32-18,688	2795	2685	1258	1328

lower. These differences will be discussed in detail later.

Although additional parameters have been routinely monitored in the water quality study, it was felt that the parameters selected would adequately serve as indicators of water quality.

Discussion of water quality parameters

Turbidity   Turbidity is a measure of the light scattering and light absorbing characteristics of water. It may be caused by a variety of suspended materials such as clays and silt from the translocation of topsoil or from the suspension of sediment from stream bottoms or banks. However, finely divided organic matter, colloids, plankton and other microorganisms may also account for a considerable portion of the turbidity. In general, surface waters of high turbidity are associated with poor water quality for esthetic and health reasons. Highest turbidities in the Des Moines River generally occur in the spring during periods of high surface runoff. Conversely, lowest turbidities occur during the winter when the river is ice covered.

Chloride   Chlorides occur in all natural waters in widely varying concentrations. In the Des Moines River groundwater may be the principal source of chloride. However, groundwater in the basin is considerably lower in chloride than the highest concentrations measured in the Des Moines River. Animal wastes and industrial wastes may account for the higher concentrations observed.

Silica Silica, next to oxygen in abundance in the earth's crust, is present in surface waters in both soluble and colloidal forms. A silica cycle is observed in surface waters containing diatoms which use silica in their skeletal structure. The silica removed from the water during the formation of the skeletons is slowly returned by dissolution of the dead diatoms. In the Des Moines River this inverse relation is observed in the spring and fall when diatom blooms frequently occur. A correlation coefficient of about -0.5 indicates the strength of this inverse relationship in the Des Moines River.

Hardness Hardness is caused by divalent metallic ions. Calcium and magnesium are the principal hardness producing ions in the Des Moines River. The hardness of a water reflects the nature of the geological formations with which it has been in contact. Because hardness is a component of groundwater, the highest values have been measured during cold weather periods of low flow and little precipitation. Although not reflected by the means listed in Table 12, there is a small but interesting difference in the proportion of the total hardness attributable to calcium during August and September (0.52) as contrasted with the winter months (0.62). Groundwater feeding the streams at this time is not greatly different in chemical composition from that during the winter when it has much the same hardness characteristics as the river. The decrease in the proportion of calcium hardness is most likely due to the photosynthetic activities of plankton, and will be discussed in greater detail later.

Dissolved oxygen   Dissolved oxygen is a necessary component of streams for a healthy ecosystem. Fish require a minimum concentration of about 4.5 mg/l. Five important factors interact to determine the dissolved oxygen concentration in a river.

Temperature determines the solubility of oxygen.

Algae contribute oxygen during photosynthesis.

Bacteria deplete oxygen during respiration.

The time of day the sample is collected is important because of the relationship of light and algae photosynthesis.

Turbulence of the stream causes a dissolved oxygen equilibrium to be reached between the water and the atmosphere.

There is considerable variation in the dissolved oxygen concentration from week to week because of variations in the time of day that samples were collected. Highest mean values of dissolved oxygen are recorded during the winter when temperature is the controlling factor. Maximum values, however, are recorded during low flow periods in the late summer when the photosynthetic activity of large numbers of plankton cause the stream water to become supersaturated with oxygen.

Biochemical oxygen demand   The biochemical oxygen demand (BOD) test is essentially a bioassay procedure involving the measurement of oxygen consumed by living organisms while using the dissolved organic matter present in the water. The BOD of a water is one yardstock that measures the existing level of pollution. In the Des Moines River, the BOD is generally less than the dissolved oxygen concentration.

Although bacterial respiration is probably the principal removal mechanism of the dissolved oxygen during the BOD test, it may not be the only one. The BOD analysis is run over a period of five days in the dark at a constant temperature of 20°C. Although the darkness prevents photosynthesis by algae, it does not prevent their respiration. During the BOD test, plankton will continue to respirate, metabolizing dissolved food supplies from within their cells. During periods of high plankton populations, this respiration could account for a large amount of dissolved oxygen uptake, an artifact since it would not be actually related to the metabolizable dissolved organics in the water. For the Des Moines River, the correlation coefficient between numbers of plankton and BOD is 0.64. While this does not confirm the causative agent, it does suggest a relationship which would invite further study. It is interesting to note that the highest BOD concentrations occur during the months of July, August, and September when plankton populations are greatest and not necessarily during runoff periods in the spring when great quantities of organic matter are washed into streams.

<u>Chemical oxygen demand</u> During the chemical oxygen demand (COD) analysis, organic matter is oxidized to carbon dioxide and water regardless of the biological assimilability of the substances and is one measure of the total organic carbon in the water. As a result, COD values are usually greater than BOD values, especially when significant amounts of organic materials are present which are resistant to biological degradation. This is one of the advantages of the test. In Iowa, large quantities of lignins are washed into rivers during heavy runoff

periods. These materials are not broken down in the BOD test, but are oxidized chemically in the COD test. Consequently, the greatest COD occurs during peak runoff periods. Some nitrogen compounds do cause interferences in the COD analysis. Amines are converted to ammonia and some forms of organic nitrogen are oxidized to nitrate. High concentrations of these interfering materials will result in artificially high estimates of the chemically oxidizable carbon in a water.

Nitrogen Nitrogen is an important nutritive element for aquatic plants and algae in a surface water ecosystem. An inorganic nitrogen concentration of less than 0.3 mg/l is generally considered growth limiting for algae (49). This situation occurs only rarely in the Des Moines River.

Nitrogen enters surface water in several different forms and from many different sources. Nitrogen may enter the river as the principal inorganic nitrogen forms - ammonia and nitrate or as the organic form - amines and proteinaceous materials. Effluents from wastewater treatment plants and runoff from agricultural lands are probably the two principal sources in Iowa. Baumann and Kelman estimated that about 6000 pounds of nitrogen per day enter the Des Moines River on a fairly uniform basis from domestic and industrial wastewater sources in the basin above Boone, Iowa (7). Most of this will initially be in the form of ammonia. Agricultural contributions are intermittent during the warm season, coinciding with rainfall and runoff. During the winter, nitrogen input into the river as nitrate is not related to runoff, but to a more complex set of variables. However, the source of this form of nitrogen is probably also agricultural in origin. The quantity of nitrogen from agricultural sources such as animal wastes and agricultural losses from the portion of the basin above Boone, Iowa has been estimated at 75,000 pounds per day (7). This quantity of nitrogen would be divided between ammonia, organic nitrogen, and nitrates. Because ammonia is oxidized in soil and water by autotrophic nitrifying bacteria to nitrate under aerobic conditions, it is difficult to assess the relative amounts of each form of nitrogen entering streams accurately. A third significant source of nitrogen is precipitation. Estimates of nitrogen inputs from precipitation range from three pounds per acre per year (26) to 18 pounds per acre per year (52). Since precipitation would wash agriculturally based nitrogen, primarily ammonia, from the atmosphere, the quantity added to streams is not necessarily in addition to that from agricultural sources. The amount is considerable. If even the lower rate is considered, the average daily contribution on the basin would be the considerable amount of nearly 29,000 pounds per day.

Ammonia concentrations in the river are generally greatest during the winter when water temperatures below 10°C inhibit nitrification of ammonia to nitrate. Since the river at this time is ice covered and receives little surface runoff, the source of this ammonia is considered to be primarily from wastewater treatment plants. During the warm season, ammonia concentrations decrease to about 75 per cent of the annual mean concentration (5). This decrease is probably due to the increased nitrification rate in the river caused by higher river water temperatures. Dilution by surface runoff and metabolic uptake are probably also factors.

The most important sources of organic nitrogen are animal wastes and decaying plant and animal tissue. These sources contain appreciable amounts of unassimilated proteinaceous material containing organic nitrogen as amino acids. Organic nitrogen concentrations are generally greatest during spring runoff periods, but high concentrations also occur during early fall when flows are less than the average annual flow. Part of this would consist of plankton which normally appear in very large numbers in the early autumn.

Nitrate, as does ammonia, serves as a nutrient for plants in the aquatic ecosystem. Nitrate enters the river dissolved in groundwater from interflow and tile drainage effluents. It is also formed within aquatic environment as a result of microbial oxidation of ammonia and nitrite. For the past several years a nitrate cycle has been exhibited in the Des Moines River. The concentration of nitrate generally drops in the early autumn to less than one mg/l, and then increases during the winter months to eight mg/l or more.

Phosphorus Phosphorus is a fertilizer or nutrient in the aquatic, as well as, the terrestrial environment, and is an essential element for aquatic plant growth. Phosphorus concentrations of less than 0.01 mg/l in surface waters may limit algal growth (49). During this study, water samples were analyzed for two forms of phosphorus; total phosphorus and filterable or soluble inorganic phosphorus (considered to be orthophosphate). Total phosphorus is a measure of the organic and inorganic forms of phosphorus. Organic phosphate is formed primarily in biological processes. Crop residues and animal wastes, likewise, contain phosphorus in the organic form. Inorganic phosphate, as P_2O_5, is applied to agricultural land and enters streams dissolved and suspended in runoff and adsorbed on sediment particles. Once in the stream, inorganic phosphate enters the food chain and is converted to the other forms of phosphorus.

In Iowa, the two major sources of phosphorus are agricultural runoff and wastewater treatment plant effluents. For the portion of the basin above Boone, Iowa it has been estimated that agricultural sources, which include cropland losses and animal wastes, account for about 3300 pounds of phosphorus per day. For the same area, treated wastewater contributes about 1600 pounds per day according to a study by Baumann and Kelman (7).

During the month of January in the years 1972 to 1974, the river was, for the most part, ice-covered. Little surface runoff occurred. It may be considered that most of the phosphorus in the river at this time was contributed by wastewater treatment plants since phosphorus is not a significant component of groundwater. The mean total phosphorus concentration during these months was 0.91 mg/l as PO_4 and the streamflow averaged 2100 cubic feet per second, higher than normal. These figures indicate average phosphorus inputs of about 9080 pounds of phosphate (3000 pounds of phosphorus) per day, somewhat greater than the quantities estimated by Baumann and Kelman (7). For the period of the study, orthophosphate concentrations indicate inputs of about 2730 pounds of phosphate (890 pounds of phosphorus per day). In this case, the inputs are less than the estimate of Baumann and Kelman (7). Phosphate binds to sediment and it may be this mechanism which introduces complexity into an attempt to sort out the sources of phosphorus inputs to streams.

Orthophosphate concentrations are greatest during the winter months of January through March. During January and February, the source of most of this phosphorus is wastewater treatment plant effluents. In March, the snowmelt runoff transports phosphorus from agricultural sources, but also dilutes the stream concentration. A second smaller rise in phosphate concentration occurs in the autumn when rainfall washes accumulated crop residues from the harvest and animal waste into streams. Some of this phosphorus may also come from fall fertilizer applications during years when favorable weather conditions exist. Variations in total phosphate concentrations in the stream parallel that of orthophosphate, but runoff conditions result in a greater change in concentration for total phosphate. This is probably because sediment containing adsorbed phosphates is included in the sample analyzed. In the orthophosphate analysis, the sample is filtered through a 0.45 micron filter, excluding all but colloidal sized sediment particles.

<u>Fecal coliform</u> The presence of fecal coliform bacteria in water indicates contamination by fecal materials excreted by warm-blooded animals such as livestock and humans. While coliform bacterial are relatively harmless as disease causing organisms, their presence is evidence of the possibility of pathogenic bacterial and viral contamination in the water.

There is not a great deal of variation in numbers of coliform per 100 ml over the year when the four-year monthly mean of the natural logarithm of the coliform count is considered. (This transformation tends to smooth very high peaks). For the four-year period, 1970 through 1973, greatest monthly mean counts were recorded for March and September. The high March mean is probably due to snowmelt runoff washing animal wastes accumulated during the winter into the river. The September maximum may be artificially high. One of the highest recorded counts, 20,000 organisms per 100 ml, was included in this mean. On this occasion, heavy runoff washed large amounts of sediment into the river and the date the sample was collected coincided with the peak of the hydrograph at that location. In general, very high coliform counts which are ten to 100 times the annual mean are recorded during peak runoff periods.

<u>Phytoplankton</u> Phytoplankton, or more commonly, algae, make possible important chemical changes and metabolic activities in the aquatic environment as a result of their photosynthetic activities. Oxygen released during algal photosynthesis and oxygen reaeration are

the two primary sources for renewal of oxygen in the riverine environment. During low flow periods in late summer, plankton counts may increase to 100,000 or more per ml. Their vigorous photosynthetic activities at this time frequently increase daytime dissolved oxygen to twice the saturation concentration. Another important chemical effect occurring simultaneously is removal of carbon dioxide from the water. With the removal of carbon dioxide, the buffering capacity of the water is reduced, and it is at this time that greatest phenolphthalein alkalinity occurs. The pH of the water may approach or exceed nine units. Also accompanying the carbon dioxide removal is a shift in the bicarbonate equilibrium towards the formation of carbonate, as evidenced by the increased phenolphthalein alkalinity. This process causes some of the carbonate to precipitate. Because of the difference in solubility products of calcium and magnesium carbonates, calcium will precipitate first and the magnesium will remain in solution. It is probably for this reason that during the late summer and early fall, the magnesium hardness (the total hardness less the calcium hardness) makes up a larger portion of the total hardness as contrasted with the rest of the year.

Plankton populations are influenced by a number of factors. Water temperature, the number of daylight hours and the presence of essential nutrients are all important for plankton growth and reproduction. Streamflow rate, however, appears to be the dominant factor regulating the number of plankton per ml from one week to the next. From one season to the next, however, water temperature is the principal factor controlling the growth of the floating or planktonic algal forms. During cold weather, algal populations appear to shift from planktonic forms to benthic or attached forms.

Many types of algae are included in the term phytoplankton. Phytoplankton are the floating algal forms, as contrasted with the attached forms or benthic algae. In the Des Moines River, diatoms are the most abundant form of algae and, on an annual basis, make up about 80 per cent of the algal species. During the winter months of January, February, and March, however, the motile, flagellated forms comprise 20 to 50 per cent of the phytoplankton. The green algae are normally not present during the winter. They begin to appear in larger numbers in April or May when water temperatures increase to 10 to 15°C and the number of hours of sunlight increases. They continue to make up 10 to 20 per cent of the plankton population until September when shorter days and lower water temperatures appear to limit their growth. Blue green algae growth patterns parallel that of the green algae. During the growth period of the blue green algae, they make up about three per cent of the total phytoplankton population.

STATISTICAL ANALYSIS - RESULTS AND DISCUSSION

Introduction

Limnological data were collected on 314 occasions at a site on the Des Moines River near Saylorville, Iowa from July 6, 1967 to July 27, 1973 as a part of a preimpoundment study associated with the Saylorville Reservoir. The preimpoundment study is being conducted on a long-term basis by personnel of the Sanitary Engineering Section of the Iowa State University Engineering Research Institute under contract with the Corps of Engineers, U.S. Army, Rock Island District. A complete compilation of the data, collected at approximately one week intervals throughout the year, is contained in annual reports submitted to the Corps of Engineers and on file with the Iowa State University Engineering Research Institute (3, 4, 6, 8, 9, 11).

During the study period, typical Iowa weather provided opportunities for observation of the river's response to a broad spectrum of hydrological conditions from nearly drought to floodstage flows. In the water year 1967-68, the mean annual flow was 466 cfs, the lowest recorded during the study period. The following year, 1968-69, the mean annual flow was 5175 cfs, the maximum recorded during the study period. For the period of this study, the streamflow has averaged about 2750 cfs.

Because of the great variation in hydrological and climatological conditions, the data are believed to be representative of the response of the river to the types of conditions which would occur over a longer period. Great variations in the limnological data resulted from this wide range of conditions, as shown in Table 12. A review of the data indicated that some values were greatly different than would normally be expected. In these cases, the original analytical results were checked for errors. If an error was clearly indicated, the data were omitted. Otherwise, all data values were included in the regression analyses. It was felt that for the purpose of representing accurately the river's response to basin conditions no data should be omitted simply on the basis that a given point appeared to be outside the expected range.

Because data sets for most of the limnological parameters were nearly complete, missing data values did not cause great problems. In most cases, the data set was compressed prior to the regression analysis. However, for the analysis of smoothed data, complete sets were desirable. For this type of analysis, missing values were calculated from regression equations, and the calculated value was entered in the place of the missing data value.

Grouping of Data

Preliminary analysis of the data indicated that the river responded differently to the environmental conditions of cold weather than to warm weather. Based on this observation, the data were analyzed on a bi-annual basis as well as on an annual or full year basis. This division is not without precedence (60).

The difference in the means of the biannual grouping compared to the annual grouping was tested statistically for significant differences at $P=0.05$, i.e. at a probability of 0.05. Parameters for which no significant difference was indicated were total phosphate, chloride, and fecal coliform. It was reasoned that for some of these parameters, the response should differ from the cold season to the warm season. Of particular interest were chloride and fecal coliform parameters. These were investigated in greater detail in order to determine whether the biannual grouping could be justified on some other basis.

Explanatory Parameters

Explanatory parameters were selected to represent hydrological and climatological conditions in the basin. These parameters were grouped into six components: streamflow, runoff, suspended sediment, temperature, seasonal variations, and seasonal-flow interaction parameters. An explanation of the individual parameters has been given previously.

Each of the limnological parameters was regressed on the entire set of 25 explanatory parameters using a computerized stepwise multiple regression technique (19). Explanatory parameters were selected at a significance level of $\alpha = 0.05$, unless otherwise noted. Because the stepwise regression technique tended to exclude explanatory parameters which were correlated highly, occasionally a somewhat undesirable selection occurred. For example, a fairly common selection was the streamflow on a day other than the sampling date. The reason for this selection was due to the higher correlation of the streamflow on the sampling date with other parameters included in the regression. Since it would be helpful to have streamflow on the sampling date included for ease of estimation of the limnological character of the water, for several parameters streamflow on the sampling date was forced into the regression in place of streamflow on a prior or following date. The resulting equation was evaluated to determine whether the initial selection was fully justified or whether the substitution could be made with little loss of statistical significance in terms of the standard error, the R^2 value, and the F test. In any case, the initial selection of the preset α level always included those parameters which minimized the error sums of squares. In this sense, the initial regression selected the explanatory variables which were best representative of the relationship of the limnological parameter with hydrological and climatological conditions.

Intercorrelation of Explanatory Parameters

Most of the explanatory parameters used in this study were correlated to some extent. For some, the correlation was nearly unity, as in the case for streamflows on successive days. For others, the correlation was zero. Knowledge of the correlation between explanatory parameters aids the interpretation of relationships which are indicated in the regression equations. Appendix A lists these correlations.

Correlation within the component groups varied from one group to another. Highest intragroup correlation existed for the streamflow component because of the strong relationship between streamflow on successive days. The lowest correlation within the streamflow group was between the parameters Q+4 and Q-5 for which r = 0.88, 0.91, respectively, for the warm season and the cold season. It should be noted that although the parameters of the form Q±n which are natural logarithm transforms of flow correlated highly (r = 0.8) with the parameter QSTD, essentially untransformed streamflow, the distribution of the two parameters would be different. The slight seasonal difference of the intercorrelation of the Q±n parameters resulted from a less variable flow during the cold season when runoff events were less frequent. Parameters indicative of a change in flow (runoff events) correlated to a lesser extent as given in Appendix A. Again, it is noted that correlation is higher during the cold season for parameters of the form Q/QAn. In contrast, the parameter DQ3/Q, which was function of the slope of the hydrograph on the sampling date, correlated best with those of the form Q/QAn during the warm season. It may be that this parameter provides information of a somewhat different nature than the other parameters in the group. Correlation within the seasonal parameters indicated that the parameters AGS and AGF are nearly inversely related. One of these parameters could probably have been omitted with little loss of statistical significance. For the seasonal-flow interaction parameters which were reciprocal functions of streamflow, the highest correlation was between the parameters M/Q and AGFC, r = 0.91, during the warm season.

The extent of correlation between parameters of different component groups was also of interest. Correlation of the streamflow and the runoff component groups with other parameters was of particular interest as an aid in the interpretation of the regression equations. Results of regression analysis come close to implying case and effect (54). However, in a multivariate study which includes highly correlated explanatory variables, proper interpretation requires a good understanding of how these variables are related.

Streamflow correlated most highly with season and season-streamflow parameters. Highest correlations were indicated for the warm season. It was apparent that during this season, the parameters SUNC, AGFC, and M/Q were essentially reciprocal streamflow functions. Of the season parameters, AGS was most highly related to streamflow. This was the result of highest streamflows generally occurring in the spring. The concentration of suspended sediment also correlated well with streamflow (r = 0.80).

Correlation of streamflow with other explanatory parameters was somewhat lower during the cold season. Temperature was an exception and correlated to a much greater extent because of the increase in streamflow accompanying warmer spring temperatures and runoff conditions.

Of particular interest was the lack of high correlation between streamflow and runoff parameters. This was considered to be an indication of a certain amount of independence between streamflow and runoff. It also indicated that the inclusion of runoff parameters in the regression equation represented a different type of relationship in contrast with reciprocal flow functions such as SUNC and M/Q which were highly related to streamflow.

Runoff parameters did not correlate particularly well with any of the other parameters. The highest correlation was with sediment (R = 0.46 warm, 0.40 cold). It is of interest that this correlation was much lower than for streamflow and sediment. This lends statistical support to the observation that although the concentration of suspended sediment does increase dramatically with runoff, following the runoff period sediment concentration fairly well parallels streamflow rather than dropping immediately to pre-runoff levels. This may also indicate that suspension of benthic materials and bank erosion contribute to a considerable extent to suspended sediment loads in the river.

In summary, the interrelationships of explanatory parameters and runoff parameters were not highly related and appeared to indicate different relationships in the regression equations. However, streamflow parameters and season-flow interaction parameters provided similar information, particularly during the warm season, and should be interpreted as such in regression equations.

Relationships of Limnological Parameters

In this section a brief overview of the different methods of treating the data prior to the regression analysis will be presented, and will be followed by a detailed analysis of the results and the conclusions based on the regression analysis for the individual parameters. For several of the limnological parameters, 10 to 15 regression equations were developed during the progress of the research.

Those regression equations which were statistically most significant or which were important to the development of conclusions regarding the limnological relationships between water quality and hydrologic and environmental conditions are included in the individual sections and discussed in detail. A complete listing of all the regression equations developed during the progression of the research is contained in Appendix B.

Normally the regression equation of best fit as indicated by the statistical parameters R^2, F and the standard error of the estimate was selected as best representing the statistical relationships between the dependent parameter and the independent or explanatory parameters. However, selecting the regression equation of best fit was sometimes difficult and required a good deal of intuitive judgement regarding the general character of the data. For example, nitrate and turbidity both had a standard deviation about equal to the mean. The distributions of the data, however, were considerably different. The turbidity distribution was skewed to the right by extreme values recorded during runoff events.

The maximum turbidity recorded was 480 JTU. This value exceeded the annual mean, 40 JTU, by more than ten standard deviations. Nitrate data, in contrast, were approximately normally distributed over the range of values from one to 15 mg/l. However, nearly 25 per cent of the nitrate values were less than one mg/l.

Outliers greater than three or four standard deviations from the mean, as mentioned for turbidity, have a considerable effect on the regression equation. Because the best regression equation is that for which the error sum of squares (the sum of squares of the distance from the regression line) is minimized, the regression line is biased strongly by data which exceeds the mean by several standard deviations. It was considered that for the purpose of this study, data with this character were out of statistical control.

Two methods were used to bring these outliers into statistical control. Data were transformed using the natural logarithm. These transformed values were then entered into the regression equation for each of the limnological parameters. It has been mentioned previously that flow data parameters of the form $Q \underline{+} n$ and sediment data were also transformed, effectively normalizing their distribution (54). The mathematical relationship between these parameters and an independent variable would be depicted as following:

$Y = a\, X^{b}$ or

$\log Y = \log a + b \log X$

A second method used to bring the limnological data into statistical control was simply that of omitting outliers exceeding the mean by several standard deviations (29). This provided a data set with, sometimes, a considerably different character. As a result, the resulting regression equations were different since the extreme values could not bias the regression.

Another problem of interest was the week to week variability in the data. Some parameters appeared to have an unreasonable amount of variability from week to week. This was first noticed for the chloride data. Chloride is a conservative substance in that it is not biodegraded or removed from the water by any known natural processes. Yet it was observed that the quantity apparently present in the river, as judged from the analytical results of the water samples, varied on some occasions by many tons from week to week. This occurred during periods of little change in streamflow, as well as during periods of more dynamic streamflow. That the indicated change represented what was actually happening in the river seemed highly unlikely.

An attempt was made to remove some of the apparently unjustified variation by smoothing the data for several limnological parameters. (The smoothing routine has been discussed previously.) It is assumed that the raw data possessed excessive variation because of a number of sources of error such as the collection of unrepresentative water samples, chemical and biological changes in the sample prior to analysis and lack of precision and accuracy in the analysis. Although the percentage variance accounted for (R^2), the F ratio, and the standard error of the estimate (SE) were improved as a result of smoothing the data, it should be noted that these statistical parameters apply to the smoothed data set and not to the raw data set. However, it was felt that, in some cases, the regression equations based on the smoothed data might provide a better representation of the actual relationships of the limnological parameters with the environmental and hydrologic parameters.

The data were analyzed on both an annual and a biannual basis. For parameters having a significant difference between the biannual seasonal distribution and the annual distribution, an equation is listed for each season. Equations are also listed for both the annual and the biannual groupings where no significant difference was indicated by the t test, but where it was considered that separation of the data might be justified on some other basis.

Discussion of statistical relationships

Turbidity   Because of the wide range in the turbidity data, see Table 12, a number of regression equations were developed in an attempt to normalize the data and bring it into statistical control. Application of the t test to the data indicated that the biannual distributions of the data were significantly different from the annual distribution ($P = 0.05$).

The regression equations selected as those of best fit for the warm and the cold seasons are listed below.

Turbidity, JTU

Warm season, natural log

$$\ln Turb = -0.171 \ln Q{+}4 + 0.212 \ln Q{-}5 + 0.347\ Q/QA14 + 0.313 \ln SED + 1.421$$

Cold season

$$Turb = 18.9 \ln Q{+}1 - 11.2 \ln Q{-}5 + 5.09\ QSTD + 26.0\ Q/QA14 + 21.5\ Q/QA28 + 3.20 \ln SED + 0.454\ SUNC + 0.204\ AGFC - 51.74$$

Table 27 in Appendix B provides a complete listing of the regression equations developed from the turbidity data.

During the warm season the turbidity of the river varies greatly, covering a range of 0 to 480 JTU. Thus, it was not unexpected that the natural log transform provided the best fit. Turbidity appears to be a function of streamflow, runoff, and suspended sediment concentration. Although the two streamflow parameters, Q+4 and Q-5, are included in the regression equation, it is important to note that the correlation of turbidity and streamflow on the sampling date is the greatest of the streamflow component group.

One interpretation of the inclusion of the two streamflow parameters is that a change in flow over the 10 day interval indicated by the terms Q+4 and Q-5 is important in turbidity considerations, as well as the actual measured streamflow. Since the regression coefficients have opposite signs, the net value will be a function of the flow on the two days ($-0.171 \ln Q{+}4 + 0.212 \ln Q{-}5$). However, for the net value to be negative, the streamflow on the day indicated by Q+4 would need to be several times that on the day indicated by Q-5. Using the unsigned coefficients

$$0.171 \ln Q{+}4 = 0.212 \ln Q{-}5$$

$$\ln Q{+}4 = \frac{0.212}{0.171} \ln Q{-}5 = 1.24 \ln Q{-}5$$

Since log functions of flow are used

$$Q{+}4 = (Q{-}5)^{1.24}$$

That a streamflow increase of this magnitude would not increase turbidity seems unlikely in the light of the understanding of causes of turbidity. However, also to be taken into account are the runoff parameter and the sediment parameter both having positive coefficients. Under the circumstances of this large change in flow, both of these parameters would very likely increase.

Few such situations were recorded during the six-year period of this study. On one occasion, June 14, 1968, the suspended sediment concentration was 900 mg/l. Although the turbidity data for that sampling date was missing (no analysis), the predicted value for turbidity, 82 JTU, appeared to be a reasonable estimate when compared to values at nearby sampling stations, 83 and 68 JTU, included in the study at that time. These stations located above the sampling site at Saylorville (Station 5) are labelled as stations 3 and 4 in Figure 3.

During the cold season, turbidities are lower and have a smaller variation. Extreme values are less frequent because the river is ice covered and the soil surface is frozen. Fewer runoff events occur. In this case the untransformed data were considered to be normally distributed. The regression equation has been given previously.

As for the warm season, streamflow, runoff, and sediment concentration are included in the regression equation. Additionally included are the two season-flow interaction parameters, SUNC and AGFC. These two parameters are of greatest importance during very low flow periods, but at high flows will add little to the regression equation in terms of forecasting turbidity. This can be shown as follows:

$$\text{SUNC} = \frac{10{,}000 \text{ X SUN}}{Q}$$

$$\text{AGFC} = \frac{10{,}000 \text{ X AGF}}{Q}$$

On March 1, the values for SUN and AGF are, respectively, about 0.6 and 0.1. If the flow on the sampling date, Q, is considered to be 200 cfs the sum of the values are only 1.46 JTU:

$$0.454 \text{ SUNC} + 0.204 \text{ AGFC}$$

and

$$\frac{(454)(0.6)}{200} + \frac{(204)(0.1)}{200} = 1.46$$

At higher flows the sum would be, for practical purposes, negligible, and the parameters could probably be omitted. Thus the principal parameters related to turbidity are streamflow and suspended sediment concentration.

A number of other regression equations were developed in an attempt to study relationships within a smaller range of turbidity. These are listed in Appendix B, Table 27. As a first attempt, only turbidity levels less than 176 JTU were included in the regression and indicated as W < 176 in Appendix B. This range effectively eliminated turbidities greater than about three standard deviations from the mean. No values were omitted for the cold season, but for the warm season, four values were dropped. Of importance here is that the R^2 and F values did not change greatly from that for the untransformed data, indicated as W in Appendix B. The important difference appears in the standard error of the estimate. The standard error of the estimate (SE) for the treatments W, and W < 176 are, respectively, 37.4 and 19.9. With the elimination of the four outliers, a large source of variation is removed from the data, and the accuracy of the estimate is improved.

In a second step, the turbidity range was narrowed even more. Only turbidities less than 89 JTU were included in the regression. Four values for the cold season and 14 values for the warm season were omitted. Little change in the statistical significance was noted for the cold season regression. For the warm season, the R^2 and F values were increased, while the SE was decreased. Although the parameters included in the regression equations varied for different treatments, the parameter common to all, and which emerged as the most important in turbidity relationships was suspended sediment.

As another consideration, why was the regression equation based on the log transformation of the data selected for the warm season, but not for the cold season? Two factors were important in this decision. One was the size of the SE in relationship to the standard deviation, S, of the actual turbidity data. Table 13 summarizes these data.

Table 13. Relationship of the standard error to the standard deviation for turbidity

Treatment[a]	SE	S	SE/S
W	37.4	46.3	0.81
W,L	0.386	0.567	0.68
C	10.5	26.7	0.39
C,L	1.62	2.07	0.80

[a]W = warm season,
C = cold season,
L = natural log transformation of data used in the regression.

[b]S = standard deviation.

It is clear that the better estimate is provided by the regression using the untransformed data. (Similar reasoning was used for selection of the "best" regression equations for the other parameters.) A second consideration which led to the rejection of the regression equation for the log transformed data was the omission of the sediment parameter. It was felt that this parameter was of considerable importance, and particularly so during the cold season when plankton and colloidal material would contribute little to turbidity.

<u>Sources of error</u> Although the method of analysis for turbidity has advanced greatly since the early use of the standard technique (53, p. 253), turbidity measurements for streams are still subject to error and imprecision. If the sample contains a high concentration of inorganically based suspended sediment, as would be the case during high runoff periods, a certain proportion of

these solids will settle out in the sample container. At the time of analysis, the sample must be shaken vigorously to redistribute the sediment, and a representative sample then must be selected. Assuming that a representative sample is selected, settling may occur in the sample tube during analysis. Errors of this nature tend to distort the turbidity - sediment relationship.

River sampling errors also will occur. The sampling bucket is lowered from a bridge to the river's surface. Although surface samples are to be collected, during low streamflow the bucket may drop to the stream bottom, stirring up and suspending the sediment. This sediment is drawn into the sampling bucket and strongly biases the sample in regard to the limnological characteristics of the river for several other parameters, in addition to turbidity, which appear to be influenced by sediment.

Some factors causing sampling errors cannot be controlled. The sand on the stream bottom is constantly shifting. What may be a good sampling location one week may be a poor location the next because of the formation of relatively unmixed pools downstream of sandbars. Even within the well mixed portion of the stream, the concentration of suspended sediment and other limnological components will vary because of the eddy currents. Thus many factors may bias the evaluation of water quality, and the analytical results, even if perfectly accurate, probably never reflect perfectly the actual water quality at the time the water sample was collected.

<u>Chloride</u> Chloride is a conservative parameter in that it is not changed to other forms by physical, chemical or biological processes in the aquatic environment. Because of this fact it was felt that a considerable portion of the variance could be accounted for by the regression equation. This ultimately proved to be a correct assessment. However, two questions arose as the analysis of the data progressed. The t test indicated that no significant difference existed in the data distribution of the biannual groupings as compared to the annual grouping. Secondly, the week to week variation in concentration appeared to be excessive.

Inspection of Table 12 shows only slight seasonal differences in the annual mean as compared to the biannual means. Results of the t test indicated that only a 30 percent probability existed of a difference in the distributions. On the other hand, the biannual R^2 values of 80.4 and 74.0 percent given in Appendix B, Table 28 were considerably greater than the annual R^2, 59.2 percent. The standard error for the annual grouping, 6.91 mg/l, is somewhat larger than for the warm and cold season grouping, respectively 4.60 and 4.88. These two statistical factors do indicate a better fit of the data for the biannual grouping. The regression equations are fairly similar, but this is to be expected since chloride is a function of flow only.

That the concentration of chloride appears to have an unreasonable amount of variation is shown in Table 14 for some representative data collected during the fall of 1972. To account for this variation is difficult. It has been estimated that the domestic contribution of

Table 14. Variation in weekly chloride concentration

Date	Flow, CFS	Concentration, mg/l	Mass, Tons/day
9/28/72	3890	22.3	233
10/5/72	2400	13.8	89
10/12/72	2550	19.6	135
10/19/72	1750	29.6	140
10/26/72	5520	22.2	330

chloride would be about two tons per day (5). It was thought that some industries might contribute significant quantities of chloride to the river. For example, packing houses use chloride during processing of meat, and this chloride appears in the wastewater at concentrations of about 300 mg/l. However, the volume of this wastewater is not very great. According to data from a recent report (61) industrial contributions of chloride in the upper Des Moines River Basin amount to about two tons per day. Contributions from domestic and industrial sources would be expected to be fairly uniform during the year, and the sum of the contributions from these sources, about four tons per day, is only a small percentage of the quantity observed in the river. It seems unlikely that industrial or municipal sources could cause the variations of many tons per day as indicated in Table 14. Animal waste is another possible source, but the potential contributions from this source have not been evaluated. However, it seems unlikely that chloride from animal waste could result in the observed variations, particularly during periods of relatively constant flow as shown in Table 14. Thus the principal source of variation is unknown, but may be attributable in part to non-homogeneity of the water sample with respect to actual river water quality and to random variation and error in the laboratory analysis.

Chloride concentration data were smoothed in an attempt to remove some of this variation. It was expected that this procedure would provide a better evaluation of the river's response in terms of chloride concentration to changes in streamflow. Because variation in the total quantity was the main source of interest, quantity of chloride present in the stream was calculated, smoothed, and then converted back to concentration, as discussed elsewhere. This modified data set was then entered into the regression.

The t test was used to check for a significant difference in the means of the smoothed data as compared to the raw data. No significant difference was indicated for the smoothing of the concentration data. However, a significant difference ($P = 0.2$) was indicated for the smoothing of the quantity data. Thus the latter technique, discussed on p. 66, was rejected as a statistically valid method of smoothing out the variation since the distribution of the raw data had been altered significantly.

Both smoothing routines were evaluated in regard to their effectiveness in reducing the variation of the indicated quantity of chloride in the river. The smoothed concentration data were used to calculate the indicated quantity for the dates listed in Table 14. The same pro cedure was followed for the smoothed quantity data. Although smoothing the quantity data did result in less variation of the indicated quantity of chloride in the river, smoothing the concentration data did not. However, smoothing the quantity data tended to significantly distort the distribution of the data. Thus neither routine was entirely satisfactory. The regression equations based on the smoothed concentration data are listed below and will be used as a basis for further discussion since this routine did not distort the distribution of the data. Table 28 in Appendix B contains a complete listing of the regression equations describing chloride relationships.

<u>Chloride, mg/l, smoothed data</u>

Annual

$$Cl = -4.21 \ln Q\text{-}2 - 4.00 \ln QA7 + 1.38\ QSTD + 86.53$$

Warm season

$$Cl = -3.36 \ln Q\text{-}5 + 4.12\ SUN - 0.290\ SUNC + 0.0714\ M/Q + 41.64$$

Cold season

$$Cl = -6.33 \ln Q{+}1 + 73.26$$

To test the relative accuracy of the relationships at various streamflows, the equations were applied to known data taken from both the data used in the regression analysis and from other data collected at the same sampling location. As anticipated the predicted chloride concentrations were nearly identical for streamflows near the mean. At very high and very low streamflows greater differences were discernible. The equation for the smoothed data appeared to be representative of the actual relationships at all but very low streamflows. For streamflows less than 150 to 200 cfs the regression equations developed from the non-smoothed data and the log transformed data provided better estimates. These two sets of equations are shown below.

<u>Chloride, mg/l</u>

Warm season

$$Cl = -1.79\ QSTD + 3.40\ SUN - 0.342\ SUNC + 0.0924\ M/Q + 18.07$$

Cold season

$$Cl = -6.80 \ln Q{+}1 + 76.51$$

<u>Chloride, mg/l, natural log</u>

Warm season

$$\ln Cl = -0.101\ QSTD + 0.00150\ M/Q + 3.163$$

Cold season

$$\ln Cl = -0.280 \ln Q{+}1 + 5.289$$

The range of the chloride concentration estimates based on the three sets of regression equations for the biannual grouping of the data was investigated for different streamflows and for different times during the year. Within the expected range of streamflow for the cold season, all regression equations developed from the cold season data provided essentially the same estimate. Greater differences in the estimates were noted for the regression equations based on the warm season data. These differences were provided by the seasonal parameters SUN and SUNC in the regression equations based on the smoothed and non-smoothed data, and were important at streamflows less than 400 cfs. For a given flow condition, highest concentrations were predicted during late October and lowest during late June. For example, the indicated range at 200 cfs was 37 to 59 mg/l as estimated from the regression equation developed from the non-smoothed data. While there may be some causal basis for these seasonal variations in concentration for a given streamflow, no good explanation could be developed other than that the trend was seasonally related to the average streamflow.

For streamflows of 400 cfs to about 15,000 cfs, all warm season regression equations gave very similar estimates, and were within 2 to 4 mg/l of each other. Beyond 15,000 cfs the regression equation developed from the smoothed data provided the best estimate, while the other equations tended to underestimate the observed concentration.

A preliminary hypothesis was that the quantity of chloride in the river was relatively constant and that concentration changes were regulated by streamflow. Since the regression equations accounted for a large proportion of the variance in the data (70%), it was considered that chloride concentration estimates based on these equations would be reasonable and representative of observed concentrations. To test the hypothesis, estimated chloride concentrations corresponding with streamflows in the range of 400 to 15,000 cfs were calculated from regression equations based on the biannual groupings and the annual grouping of the data. From these data, quantities were calculated. Not a great deal of difference in the estimated quantities was evident for a specific streamflow. What was evident was that there was a great deal of difference in the estimated quantity at different streamflows. In the streamflow range of 400 to 15,000 cfs the estimated quantity of chloride flowing past a given point along the river was about 40 to 400 tons per day. Thus the hypothesis was rejected.

On the basis of the preceding analysis it was concluded that although runoff was not related significantly to chloride concentration, higher streamflows tended to result in the contribution of additional chloride to the river. The overall effect, however, was that concentration was decreased due to dilution.

Although the source of the additional chloride is not known positively, animal waste is one potential source. The amount from animal waste may be estimated roughly from the amount of potassium in the excreta. Based on 1000 pound of live animal weight, the amount of potassium excreted by cattle is 0.31 lb/day (57). The number of equivalent cattle in the basin was estimated at 740,000. If the chloride content of the waste is considered to be equal to that for potassium, the total quantity of chloride from excreta would be about 47.5 tons per day, far less than that observed in the river at the higher streamflows. In addition it would be expected that chloride concentration would be related to runoff. No runoff parameter was included in the regression. Thus no firm conclusion could be drawn regarding the source of chloride ions which would cause the wide variations observed in the river in the quantity of this substance.

Because of the relative simplicity of the chloride relationships, it was felt that it would be desirable to check the adequacy of the regression equation in regard to describing these relationships. This can be accomplished using the analysis of variance information. In essence, this procedure permits the statistical comparison of the variance of the data within repeat observations (pure error) with the variance associated with the lack-of-fit for a regression equation (19). Repeat observations were selected from the six years of data included in this study. A repeat observation was considered to be that for which the repeat flows differed by no more than 5 per cent from their mean value. This prerequisite was necessary because of the great variability of streamflows recorded on the sampling dates. The procedure was applied to the full date set because it permitted greater variation in the repeat flows. The variance within the repeat measurements was compared to the lack-of-fit of the regression equation for the full year data set which is shown below.

Chloride, mg/l

Annual

$$Cl = -0.153\ \underset{\sim}{Q}STD + 0.0557\ M/\underset{\sim}{Q} + 22.66$$

The results of this statistical analysis are depicted in Table 15. A complete discussion of this statistical technique is beyond the scope of this thesis and is discussed thoroughly in Applied Regression Analysis by N. R. Draper and H. Smith (19).

Table 15. Analysis of variance showing pure e ror and lack-of-fit for chloride data (annual basis)

Source	df	SS	MS	F ratio
Total	246	37284		
Regression	2	20190	10095	144, significant at $\alpha = 0.05$
Residual	244	17094	70	
Lack of fit	228	16057	70	1.08, not significant
Pure error	16	1037	65	

(F (16,228, 0.95) = 2.01

For 16 and 228 degrees of freedom for respectively, pure error and lack-of-fit, the lack-of-fit F-test value, 1.08 does not exceed F (16,228, 0.95). Thus the relationship indicated by the regression equation must be regarded as adequate. In other words, addition of more variables or use of exotic transformations of the raw data would probably not result in great improvements in the model.

Although this procedure is useful for checking the validity of regression equations, it is limited to those limnological parameters for which it is possible to obtain repeat observations. Unfortunately, results of the regression analyses for the other parameters do not lend themselves to this type of analysis because of the number of different explanatory parameters included in the regression equation.

Silica   Silica is not a conservative substance in the aquatic environment. Diatoms use silica as the structural material in their cell walls. Thus the growth and death of diatoms is inversely related to silica concentration. This greater complexity was reflected in the regression equations as shown below and Table 29 in Appendix B.

Silica, mg/l, natural log

Warm season

$$\ln SiO_2 = 1.41 \ln Q\text{-}1 - 0.221\ QSTD + 0.512\ AGF + 0.0533\ SUNC - 0.0622\ AGSC - 7.092$$

Cold season

$$\ln SiO_2 = 0.0664\ SUNC + 0.0148\ AGSC - 0.0221\ AGFC + 2.878$$

Results of the *t* test indicated that the means of the bi-annual grouping were significantly different from that of the annual grouping. This division also seemed intuitively reasonable. The relationships of dissolved silica that would be observed during the cold season would be expected to be different from those during the warm season because of the different character of the water in regard to plankton population.

Both the raw data and the natural log transforms of the data were regressed on the explanatory parameters. The regression equations for the transformed data were selected as best representative of the silica relationships. The warm weather equation indicated that streamflow and seasonal variations were the principal factors related to the silica concentration. The season variable AGF had a positive coefficient and was probably related to the diatom die-off with the onset of cold weather in the autumn.

The regression equation based on the cold season data was of particular interest in that it included season-streamflow interaction parameters only. Because of this rather unusual grouping of parameters the equation will be discussed in some detail. This equation can be simplified, by reference to the definition of the parameters SUNC, AGSC, and AGFC.

$$\ln SiO_2 = \frac{148}{Q} (-4.49\ SUN + AGS - 1.49\ AGF) + 2.878$$

Consideration of the sum of the three season variables as adjusted by their coefficients and streamflow appears to indicate that for a given streamflow highest concentrations would occur in about February. For a given date, however, higher streamflows are associated with higher silica concentrations, as shown in Table 16.

The pattern which emerges for the variation during the cold season is that silica concentrations peak about February, and then decline through through April. The cause of these variations is probably related to the diatom growth pattern in the river. Because most of the diatoms are attached rather than planktonic during the winter, the relatively low correlation coefficient found between silica and the planktonic diatom concentration (-0.19) may not be entirely indicative of the true relationship. The negative coefficient does support the suggestion that the higher diatom populations observed during the fall and the spring when environmental conditions are more favorable in comparison to winter,

Table 16. Relationship of silica concentration with season and streamflow

Month	Value of Season Variables			Total
	SUN[a] x k_1[b]	AGS x k_2	AGF x k_3	
November	-1.80	0.00	-2.68	-4.48
December	-0.45	0.20	-1.94	-2.19
January	0.00	0.60	-1.19	-0.59
February	-0.90	1.00	-0.60	-0.50
March	-2.69	1.60	-0.15	-1.24
April	-4.49	1.90	0.00	-2.59

Flow, cfs	Sum of Season Variables	Month	SiO_2 mg/l[c]
1000	-4.48	November	9.16
1000	-0.50	February	16.51
1000	-2.59	April	12.11
100	-4.48	November	0.02
1000	-4.48	November	9.16
10000	-4.48	November	16.64

[a] Values for SUN, AGS, and AGF taken from Figure 10.

[b] $k_1 = -4.49$, $k_2 = 1.00$, $k_3 = -1.49$.

[c] Predicted concentration using the regression equation for the cold season.

are associated with lower silica concentrations.

<u>Hardness</u> Total and calcium hardness are considered together since calcium is one component of total hardness. Most of the hardness minerals in the river are contributed by groundwater, although there may be some contribution from the dissolution of limestone outcroppings within or near the river. Groundwater is the principal source of water for streams during the winter. At other times during the year, except during late summer dry spells, snowmelt and rainfall runoff are the principal sources. For these reasons it was expected that different relationships would exist between hardness and the explanatory variables for the warm and the cold seasons. The results of the *t* test did indicate that for both hardness parameters the biannual distributions were significantly different from the annual distribution.

Regression equations were developed using the raw data and the transformed data. Transforming the data did not make a great deal of difference in R^2 values for total hardness. The R^2 value for calcium during the cold season was increased by about 9 per cent, and was selected as the equation of "best fit". For the other three cases the regression equations for the untransformed data were selected. These equations are shown below and in Tables 30 and 31 in Appendix B.

Total hardness, mg/l as $CaCO_3$

Warm season

$$\text{T-Hard} = 26.6 \ln Q{+}4 + 85.6 \ln Q{-}1 - 45.6\ \text{QSTD} - 19.5\ \text{AGS} + 0.456\ M/Q - 497.8$$

Cold season

$$\text{T-Hard} = 68.2\ Q/QA7 - 53.4\ Q/QA28 - 13.8 \ln \text{SED} - 125.\ \text{SUN} - 2.36\ \text{SUNC} - 2.24\ \text{AGFC} + 0.174\ M/Q + 517.2$$

Calcium hardness, mg/l as $CaCO_3$

Warm season

$$\text{Ca} = 96.1 \ln Q{+}2 - 228. \ln Q{+}1 + 226. \ln Q - 29.9\ \text{QSTD} - 42.9\ \text{SUN} + 1.76\ \text{SUNC} + 0.771\ \text{AGFC} - 436.4$$

Cold season, natural log Ca

$$\ln \text{Ca} = 0.187\ Q/QA7 - 0.119\ Q/QA28 - 0.304\ \text{SUN} - 0.0131\ \text{SUNC} + 0.00349\ \text{AGSC} - 0.00452\ \text{AGFC} - 0.0839 \ln \text{SED} + 6.084$$

The regression equations for total and calcium hardness are complex, and a precise interpretation of the significance of each parameter is difficult. Streamflow, runoff, and season are the principal factors which are related to these two limnological substances.

It us interesting to note the different relationships between total hardness and the flow parameters during the warm season. The parameters Q+4 and Q-1 indicate a direct relationship, whereas QSTD and M/Q indicate an inverse relationship. At high flows, the first two terms and QSTD are of greatest importance. During low flow periods, as might occur during late summer when groundwater supplies much of the streamflow, the term M/Q becomes of importance.

Two conclusions may be drawn regarding these flow parameters. The direct relationship appears to indicate that runoff water which is normally associated with high streamflow does contribute to the hardness of the stream. This may be the result of the rain percolating into the soil profile, dissolving hardness minerals, and then flowing as interflow to nearby streams where the hardness minerals are introduced to the stream. The term M/Q appears to be related to groundwater contributions of hardness minerals since it only becomes significant during low flow periods.

The seasonal term AGS appears to be related to dilution effects of the spring runoff on river hardness, as indicated by the term's greatest negative value in the spring when melted snow and rain runoff frozen or wet soils. A small portion of the runoff may move through the soil profile at this time of year as compared to the summer, and would probably be dilute relative to the concentration of hardness minerals in the river.

Similar types of relationships exist for calcium hardness in terms of flow and seasonal relationships. One difference is the replacement of the term M/Q with the two terms SUNC and AGFC which are a function of season and reciprocal flow. It may be shown by reference to the definitions of the terms SUNC and AGFC that the sum of these terms times their regression coefficients can be combined into the following form.

$$1.76 \text{ SUNC} + 0.771 \text{ AGFC} = (7710/Q)(2.28 \text{ SUN} + \text{AGF})$$

The sum of the terms 2.28 SUN + AGF reaches a maximum of 5.78, about July and then decreases rapidly through the months of August through October to 3.04. As for the term M/Q in the total hardness regression equation, the terms SUNC and AGFC are most important in the equation during low flow periods. Thus they appear to be associated with hardness contributions by groundwater flowing into the river. One difference, however, is the dependence on seasonal factors.

It has been suggested previously that plankton metabolic activities may cause a reduction in the calcium hardness during late summer when populations are greatest. This typically will occur during the months of August through October. The sum of the terms SUNC and AGFC, for a fixed flow, during these months is declining and would indicate lower calcium hardness for a given flow than for earlier in the year. Thus it would appear that these terms permit the regression equation to accommodate the effects of plankton activities on calcium hardness.

Runoff and sediment parameters are included only in the regression equations representing the cold season hardness relationships. It is expected that these are related to the spring runoff period when streamflow is highly variable. The sediment parameter in the total hardness equation reflects streamflow and runoff, but is probably most important in its relationship to the spring runoff contribution to streamflow. The negative regression coefficient of the sediment term tends to support the earlier suggestion that the effect of spring runoff is dilution. These relationships could be summarized as follows. If a runoff event occurs during a period when the soil is frozen or very wet, most of the runoff is direct and does not flow through the soil profile. Sediment production by runoff from wet soil might be expected to be high, especially in the spring when there is little vegetation to hold the soil in place. The effect of the runoff would be to dilute the hardness of the river. On the other hand, when a change in flow occurs, but is not accompanied by an increase in the suspended sediment concentration, rain has probably moved through the soil profile to the streams and carried with it dissolved hardness minerals. Dilution effects would be less in this case.

The season parameter, SUN, is included in the cold season regression equations for both hardness parameters, and is representative of the normal variations in hardness. Greatest hardness is observed during mid-winter when groundwater supplies most of the streamflow. Lower hardness is typically observed in the fall and the spring.

Dissolved oxygen   Dissolved oxygen (DO) relationships were expected to be different during the warm season than during the cold season because the effects of the photosynthetic activities of plankton. Because of the seasonal differences, it was also expected that the biannual distribution of the data would be different from that for the annual distribution. Application of the t test to the data indicated that a significant difference did exist ($P = 0.1$)

Water samples were not collected at the same time each day. During the warm season plankton photosynthetic activities influence greatly the observed DO. For example, during late summer when plankton populations were greatest, the DO concentration occasionally exceeded 25 mg/l, nearly three times its saturation concentration. These very high concentrations were measured in samples collected in the afternoon when plankton photosynthetic activities were at their maximum. In contrast, samples collected during the morning contained a much lower DO, usually at about the saturation concentration, 8 to 10 mg/l. Because of the variations, the DO data were smoothed in an attempt to remove this source of variation. Smoothing the data resulted in an improvement of the R^2 value of 37.4 per cent to 49.3 per cent. Transforming the DO data using natural logs prior to regression yielded an R^2 value for the regression of 45.2 per cent. Because the R^2 value was not greatly different, the latter equation was felt to adequately describe DO relationships. Similar terms were included in both equations as shown below and in Table 32 in Appendix B.

<u>Dissolved oxygen, mg/l - Warm season</u>

Smoothed

$$DO = -0.333\ QSTD + 2.98\ DQ3/Q - 5.28\ Q/QA2 + 3.34\ Q/QA4 - 0.947\ Q/QA28 - 3.91\ SUN - 0.867\ AGF - 0.0345\ AGFC + 20.71$$

Transformed using natural logs

$$\ln DO = -0.0509\ QSTD + 0.193\ DQ3/Q - 0.275\ Q/QA2 - 0.0969\ \ln SED - 0.421\ SUN + 0.174\ AGS - 0.00088\ M/Q + 3.665$$

The same treatment was applied to the cold season data. Although smoothing the data gave an improved R^2 value, 47.5 per cent, as compared to 38.9 per cent for the regression on the raw data and 28.2 per cent for the regression on the transformed data, there did not appear to be adequate justification for the smoothing of the data. These three equations do not differ greatly as is shown below and in Table 32 in Appendix B.

<u>Dissolved oxygen, mg/l - Cold season</u>

Smoothed

$$DO = 0.478 \ln SED + 1.65\ AGF - 0.0486\ AGFC + 0.00952\ M/Q + 8.092$$

Untransformed

$$DO = 0.566 \ln Q\text{-}5 + 1.67\ AGF - 0.0486\ AGFC + 0.0104\ M/Q + 6.184$$

Transformed using natural logs

$$DO = 0.0463 \ln SED + 0.135\ AGF - 0.00290\ AGFC + 0.00059\ M/Q + 2.129$$

The regression equations for the two seasons reveal some interesting relationships. Runoff appears to play an important, but complex part in the dissolved oxygen content of the river during the warm season. Three of the runoff parameters DQ2/Q, Q/QA2, and Q/QA4 all indicate short term, but immediate effects. This in accordance with the understanding of the causes of changes in dissolved oxygen. The value of the term DQ3/Q is seldom as great as one and probably contributes little to the predicted value of the dissolved oxygen during most periods. The other two terms probably are of greater importance. It is apparent that higher streamflows are associated with decreased DO, as indicated by the coefficients of the term QSTD. Together with the runoff terms, the relationship could be interpreted simply as indicating that high runoff results in a reduced dissolved oxygen. From the coefficients of the seasonal variables AGF and SUN, in the regression using the smoothed data it is apparent that higher DO occurs during the fall, other factors being the same. For example, on June 20, the DO reduction due to these two terms (isolated from the rest of the equation)

$$-3.91\ SUN - 0.867\ AGF$$

would be equivalent to

$$-(3.91)(2.0) - (0.867)(1.0) = -8.69,$$

but for September 20

$$-(3.91)(1.0) - (0.867)(2.1) = -5.64$$

This is in accordance with the observed seasonal variation of DO concentrations.

However, another factor which would be important in regulating the DO is plankton population, which is typically greatest during the months of August through October. Streamflow during these months is lower than during the rest of the warm season. Runoff events are less frequent. Thus it is felt that a part of the inverse relationship with streamflow and the complex relationship with runoff must be due to the effects of plankton photosynthesis.

During the cold season, the DO is relatively constant. The equations indicated are probably unnecessarily complex, although simpler in form than the equation for the warm season. For prediction purposes, DO during the cold season could probably be estimated reasonably well from the more simple equation based on the untransformed or raw data incorporating only the parameter M/Q. This equation, shown below, while accounting for a smaller percentage of the variance relative to the more complex equations, did give an R^2 of 31.5 per cent and may adequately describe the cold season DO relationships.

$$DO, mg/l = 0.00554 M/Q + 11.57$$

In essence, the regression equation indicates that a lower DO accompanies higher streamflow. This relationship appears to be straightforward. Highest DO is recorded during cold winter weather when streamflows are typically lower than during the spring or fall. During the fall and spring, higher flow and warmer water temperatures combine to produce low DO.

For the purpose of comparing the predicted DO of the four cold season equations, five sampling dates were selected and the predicted and observed DO on each date is shown in Table 17.

There is not a great deal of difference in the DO predicted by the four equations. All are within 1 to 2 mg/l of the observed DO and provide reasonable estimates for the small sample shown.

In summary, the warm weather DO does appear to be influenced by runoff. Runoff and high streamflow are associated with a lower dissolved oxygen. Greater dissolved oxygen concentrations were predicted for the fall, a period of typically lower streamflow and less

Table 17. Predicted and observed dissolved oxygen concentrations during the cold season

Date	Predicted DO, mg/1				Observed DO, mg/1
	S[a]	T[b]	U[c]	UB[d]	
11-15-72	13.7	13.9	13.8	11.7	12.4
12-14-72	12.4	12.3	12.6	11.9	13.0
1-11-73	11.7	11.6	12.0	11.9	11.9
2-15-73	10.9	10.8	11.4	11.9	12.6
3-15-73	11.0	11.1	11.7	11.6	10.4

[a]S = regression equation using smoothed data.

[b]T = regression equation using log transformed data.

[c]U = regression equation using untransformed data.

[d]UB = regression equation using untransformed data containing only the parameter M/Q.

frequent runoff. It is felt that plankton must also be related to the DO at this time, and that their contribution to DO is included in the seasonal, and possibly the streamflow and runoff relationships. During the cold season, the DO is relatively constant. The most fundamental relationship is with streamflow. Higher streamflows are associated with lower DO. This does not appear so much as a causal, but rather an associational relationship in that higher flows in the spring and fall are associated with lower DO in contrast with high DO and lower flows in the winter. A more complex, and perhaps causal relationship was also developed. The direct relationship of DO with streamflow or sediment indicates that a higher DO would be expected in the fall and perhaps the spring. These parameters may be related to the photosynthetic activities of plankton.

Biochemical oxygen demand The principal source of BOD in the river varies during the year. During the warm season surface runoff was expected to be the principal source. In contrast, point sources such as wastewater treatment plant effluents were expected to provide the principal contributions during the cold season. Because of these seasonal differences it was not surprising to find that the biannual seasonal means were significantly different from the annual means as determined by the *t* test.

The data were treated in various ways prior to regression analysis. Regression equations developed from the different treatments of the data are given in Table 33 in, Appendix B. One method used was to smooth the data. The reason for this will become clear.

Standard Methods (53) indicates that the BOD test has relatively poor precision. It was reported that for a glucose-glutamic acid mixture analyzed by 34 laboratories, the geometric mean of all results was 184 mg/l and the standard deviation of that mean was ± 31 mg/l (17%). The precision obtained by a single analyst in his own laboratory was ± 11 mg/l (5%) at a BOD of 218 mg/l (34 mg/l greater than the mean of all tests). Within the BOD range recorded for river water samples collected in this study (0.5 to 30.2 mg/l), the precision was found to be no better. The standard deviation for four replicate samples was about 13 per cent of the mean value (14). The accuracy could not be determined. Causes of poor precision and accuracy are the collection of unrepresentative samples, time delay before analysis, and lack of precision and error during analysis.

Because of the inadequacies of the BOD test, the data were smoothed in an attempt to remove some of the variance which may have been the result of errors. The warm season means for the raw data and the smoothed data were not greatly different, 9.86 and 9.97 mg/l, respectively, and the standard deviations were 5.52 and 4.42 mg/l, respectively. For the cold season the means of the raw data and the smoothed data were, respectively, 7.43 and 7.61 mg/l. The respective standard deviations were 5.76 and 5.05. Thus, the smoothing routine removed a portion of the variation from the BOD data and it was felt that the regression equations based on the smoothed data might provide a better manifestation of the actual relationships existing in the river. These regression equations are shown below and in Appendix B, Table 33.

<u>Biochemical oxygen demand, mg/l, smoothed data</u>

Warm season

$$BOD = -3.06 \ln Q^{-1} - 0.117\,TEMP - 0.135\,AGFC + 0.0141\,M/Q + 35.42$$

Cold season

$$BOD = -0.785\,QSTD - 2.30\,Q/QA14 + 3.02\,Q/QA28 + 6.28\,SUN - 2.46\,AGS + 0.190\,SUNC + 0.0850\,AGFC + 4.118$$

The relationships indicated by the regression equation for the warm season were somewhat puzzling. No runoff parameters were included regression equation during a period when runoff water was expected to be the principal source of BOD in the river. In contrast, runoff terms were included in the regression equation during a period when point source discharges were expected to provide a relatively constant contribution of materials which would exert a BOD.

Closer examination of the regression equation for the warm season indicates that BOD is related inversely with streamflow. In other words, the effect of higher streamflow is dilution. One explanation is that although runoff may wash large quantities of material into the river which could exert a BOD, this runoff water has a lower BOD than the river. Example calculations are shown in Table 18.

Table 18. Relationship of BOD and streamflow during the warm season

Parameter x coefficient	Streamflow, cfs 100	1000	10000
-3.06 ln Q-1	-14.1	-21.1	-28.2
+0.0141 M/Q	14.1	1.41	0.141
Total[a]	0.00	-19.7	-28.1

[a]Estimated value as influenced by streamflow only. The parameter AGFC is not included as it has a seasonal variation.

Another approach in explaining the BOD-streamflow relationship is that the BOD of the water samples was influenced in some way so as to produce artificially high results. For example, it was found that plankton population was related strongly to BOD ($r = 0.64$). Highest plankton populations occurred typically in late summer and early autumn when the streamflow was much lower than the annual average. This was also the period when the river water samples exhibited a BOD which was much greater than average.

One way of viewing the BOD-plankton relationship is to assume that the BOD was externally introduced and that the organic matter which caused the higher BOD in some way stimulated plankton growth. Thus the direct association. There is some support for this view in the literature (39). Studies of the relationships of nutrients and plankton population are generally concerned with lakes rather than rivers, however.

Because the correlation between BOD and plankton provides no implication in regard to cause and effect, consideration must be given also to the possible effect of large numbers of plankton within the BOD bottle on the final results of the BOD analysis. The conditions under which the BOD test are run are favorable to the normal dark phase metabolic activities of algae. During algal respiration stored organic compouned are oxidized to carbon dioxide as the algae remove oxygen from their environment. In the dark environment of the BOD bottle algae would continue to respirate until some essential metabolic substance was limiting. The oxygen uptake of algae during the BOD analysis is not able normally to be differentiated from that of bacteria. However, use of oxygen by algae is not considered to be a component of BOD according to the definition of the biochemical oxygen demand (16).

> ...the amount of molecular oxygen required to stabilize the decomposable matter present in a water by aerobic biochemical action.

Whereas bacteria are able to use dissolved organics originating outside their cell wall during respiration, algae use soluble organics stored within the cell wall and do not have the capacity to stabilize decomposable materials outside their cell wall.

From a study of nighttime respiration rates of algae in streams of Central Iowa, Swanson (56) developed the following relationship between chlorophyll a concentration and oxygen uptake due to their respiration.

Oxygen uptake, mg/l/hr = 0.00209 x chl a, mg/l.

In another study by Kilkus et al. (38) plankton counts were related to chlorophyll a using a least-squares regression analysis for several rivers in Central Iowa. The resulting relationship is expressed below.

Chl a, mg/M^3 = 398 x number of cells/ml

They found that this relationship was constant over the range of chlorophyll a values observed in their study, 14 to 152 Mg/m^3. Using the results of these two studies it was possible to estimate the potential effect of plankton on the results of the BOD analysis.

In the present study plankton were counted by genus and reported as the number of cells per ml. Typical counts recorded during the months of August through October ranged from 10,000 to 200,000 plankton/ml. The oxygen demand was calculated based on the assumption that the plankton respiration rate was linear for 24 hours. At a plankton count of 10,000 per ml, the oxygen demand would be only about 1.3 mg/l. But at 100,000 per ml, the oxygen demand would be 12.6 mg/l and could account for a significant portion of the BOD.

A number of very high plankton counts were recorded during the months of August through October of 1971. On September 24, the count was 127,300 per ml and corresponded with an observed BOD of 18.0 mg/l. Using the procedure outlined above, the potential oxygen demand due to plankton respiration in this situation would be 16 mg/l. While the environmental conditions within the river are not identical to those in a BOD bottle, it is clear that interferences in the BOD analysis due to algal respiration are not negligible. Another study has shown lesser effects of plankton on BOD (25).

In summary, an inverse relationship between BOD and streamflow is indicated by the warm season regression equation. A similar inverse relationship exists for plankton and streamflow. High plankton counts and high BOD values are obtained during low flow periods in late summer and early autumn. On the basis of the previous discussion it is suggested that the relationship between BOD and streamflow is not necessarily that which would exist in the river. Rather, the BOD-streamflow relationship may be artificial because of the oxygen demand due to algal respiration within the BOD bottle during analysis. Interference by algal respiration would be expected to be greatest during low flow periods when highest plankton populations are typically observed.

The variables included in the cold season regression equation for BOD include runoff parameters, as well as streamflow and season parameters. The negative coefficient of the flow parameter, -0.785 QSTD, indicates an inverse relationship between BOD and streamflow. The term QSTD has been defined previously as

$$QSTD = Q/\bar{Q}$$

where Q is the streamflow on the sampling date and $\bar{Q}$ is the six-year mean flow at the sampling location, about 2750 cfs. As the streamflow increases the value of QSTD becomes more negative.

Two season parameters are included in the equation, SUN and AGS. Their sum during January and February goes to the larger negative values, as shown in Table 19. Later in the winter, the sum of season variables becomes less negative, apparently an indication of an increase in the BOD with the onset of spring. Runoff events in early spring would add oxygen demanding carbonaceous materials to the river and the BOD of these materials is exerted more readily as the water temperature increases. The value of the streamflow parameter becomes more negative at this time because the flow exceeds the mean flow. See Figure 8.

The season-streamflow interaction parameters have the coefficients

Table 19. Relationship of BOD with streamflow and season during the cold season

	1 + 6.28 x SUN[a]	2 - 2.46 x AGS	Subtotal Columns 1 & 2	3 - 0.785 x QSTD[b]	4 850/Q x (2.24 SUN + AGF)	Subtotal Columns 3 & 4	TOTAL[c] columns 1-4 + b_0
November	2.02	-.06	1.96	-0.77	0.77	0.00	6.08
December	0.37	-0.59	-0.22	-0.29	1.22	0.93	4.83
January	0.09	-1.62	-1.53	-0.26	0.80	0.54	3.13
February	1.57	-2.96	-1.39	-0.21	1.03	0.82	3.55
March	4.03	-4.02	-0.01	-1.00	0.37	-0.63	3.48
April	7.42	-4.80	2.62	-1.78	0.37	-1.41	5.33

[a]Values for season variables estimated from Figure 10.

[b]Streamflow values estimated from Figure 8.

[c]Not including runoff parameters, b_0 = 4.118.

0.190 SUNC + 0.0850 AGFC

It can be shown by reference to the definition of SUNC and AGFC that this is equivalent to

850/Q (2.24 SUN + AGF)

When values on the first of each month for streamflow, season, and season-streamflow parameters are considered over the period of the cold season, the relationship of the components to BOD is clearer, as shown in Figure 11 and Table 19. Effects of the runoff parameters are not included in the analysis as they would be related to individual runoff events, as opposed to the consideration of the more general relationships over the entire cold season. In other words, the runoff parameters add dynamics to the general trend indicated by the other explanatory variables, and their effects would be superimposed.

There appears to be a trend towards lower BOD during the winter months as indicated by the season parameters. The observed concentration, however, is a function of the dilution by streamflow. During December through February, groundwater is the source for most of the streamflow. The source of BOD in the river at this time is mainly the effluents from wastewater treatment plants. In late autumn and early spring, an additional source of BOD is runoff water. During the early spring, the effect of the runoff is principally dilution, whereas during late autumn it has the opposite effect. One interpretation of this observation is that the components of the autumn runoff

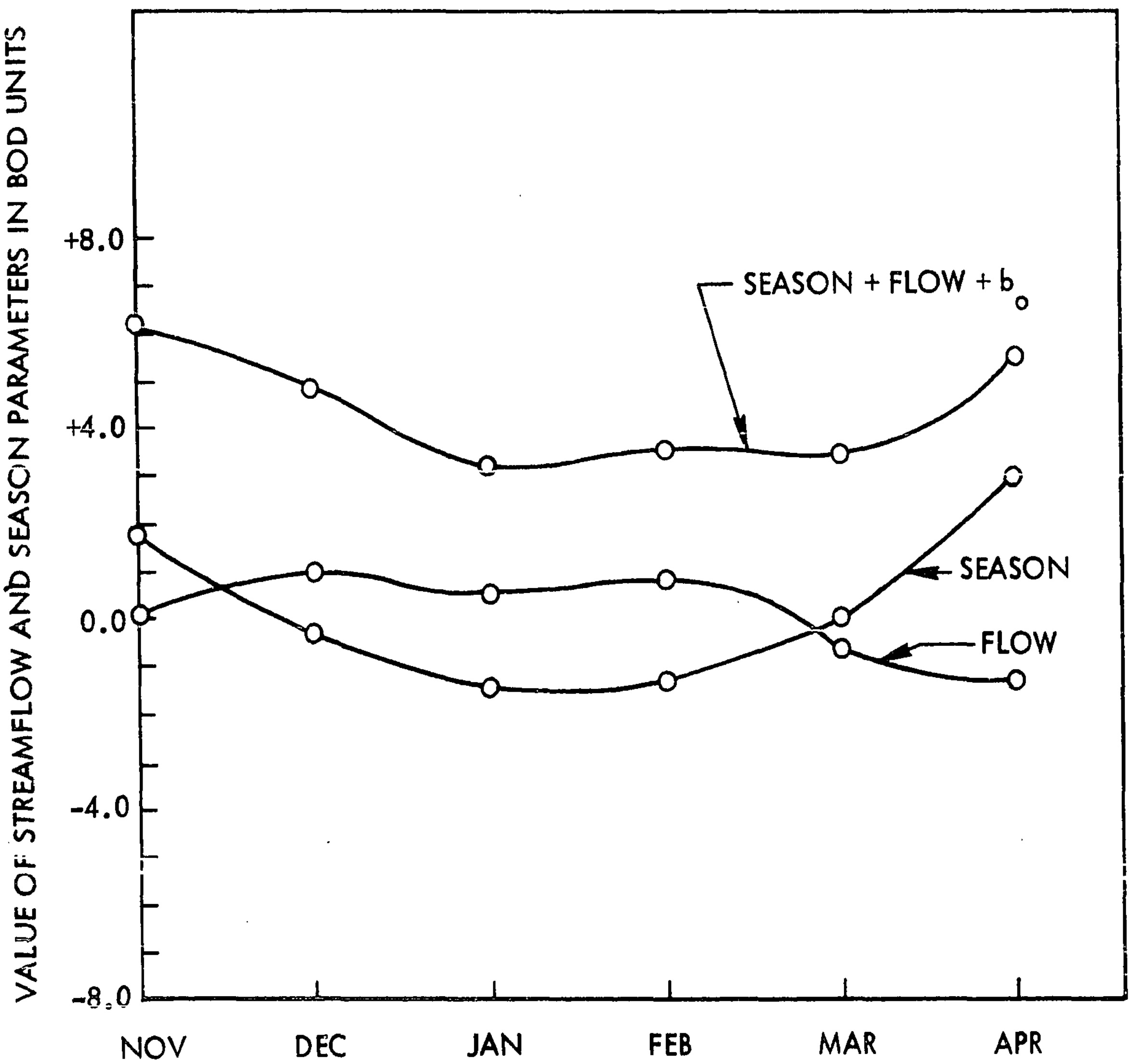


Figure 11. Relationship of BOD with streamflow and season during the cold season

such as crop residues from the harvest are biodegradable. By the spring these materials have been washed previously from the soil surface, and the major component of the spring runoff water is sediment and soluble inorganic materials which have little effect on the average BOD.

Chemical oxygen demand   In contrast with BOD, materials exerting a COD do include some of the substances associated with sediment. These materials would be expected to differ seasonally. Application of the *t* test to the COD data indicated that the data should be grouped on a biannual basis. As shown below and in Table 34 in Appendix B, the regression equations for the transformed and the untransformed data are nearly identical.

Chemical oxygen demand, mg/l

Warm season

Untransformed

$$COD = 27.3\ Q/QA4 + 0.608\ TEMP + 0.283\ AGSC + 0.07971$$

Transformed using natural logs

$$\ln COD = 0.473\ Q/QA4 + 0.0136\ TEMP + 0.00614\ AGSC + 2.858$$

Cold season

Untransformed

$$COD = 6.81\ Q/QA28 + 4.22 \ln SED + 1.16\ TEMP + 0.246\ SUNC + 0.0233\ M/Q - 15.58$$

Transformed using natural logs

$$\ln COD = 0.188\ Q/QA28 + 0.279 \ln SED + 0.0775\ TEMP + 0.00122\ M/Q + 1.086$$

The equations developed from the untransformed data were selected for further study on the basis of a lower standard error of the estimate relative to the respective standard deviation.

Only three parameters were included in the regression equation representing the warm weather relationships of the COD parameter. These included runoff (Q/QA4), temperature, and a season-streamflow interaction parameter (AGSC). The regression equation for the cold season included runoff (Q/QA28), sediment, temperature, season-streamflow interaction (SUNC), and reciprocal streamflow. It is of interest to compare the relationships indicated by these two equations.

All terms have positive coefficients in both equations. Thus each term could be considered to be directly related to the COD concentration of the river water. Both equations contained a temperature term. This is the only parameter for which this is the case. During the cold season it is possible that the COD is indirectly related through temperature to snowmelt and the spring rains. Associated with snowmelt and spring rains, are runoff laden with sediment materials, higher streamflow, and an increase in river water temperature. Part of the increase in river temperature is probably caused by the addition of a large volume of runoff water which is relatively warmer than the groundwater which had served as the source of the base flow during the winter. Thus the relationship between COD and temperature would seem to indicate an indirect relationship of the type:

temperature $\rightarrow$ runoff $\rightarrow$ COD

The reason for the inclusion of temperature in the regression equation for the warm season is not clear. The t value for temperature is barely significant at $\alpha = 0.05$. The relationship may be as simple as that on several very warm days runoff events occurred which resulted in a high COD of the river water. At any rate, the relationship, as for the cold season, would appear to be indirect rather than causal.

The appearance of sediment in the cold season regression, but not in the warm season regression may be related to several factors. It has been suggested previously that the character of the runoff differs from season to season. In the spring, runoff is laden with soil and soluble materials (31). The soil is not bound to the surface by vegetation to the extent that it would be during the warm season. During the warm season, runoff may contain debris from the cutting of hay and grasses, animal waste, and crop residues from the autumn harvest, but relatively smaller quantities of soil. Thus the warm season river COD would be more strongly related to runoff than to sediment concentration in the river.

The warm season runoff term (Q/QA4) contrasts with cold season term (Q/QA28). It is suggested that although both of these terms indicate runoff, they may be related to the general nature of streamflow. During the winter, flow, on the average, is uniformly low, as shown in

Figure 8. When a runoff event does occur, it is much greater than the previous month since snowmelt and rainfall runoff often occur together near the end of the cold season (about March). This may be the principal runoff event as noted by changes in streamflow. On the other hand, during the warm season, the flow is, on the average, fairly high. Although the streamflow may be highly variable, the 28-day average may not change greatly. Not many events occur which would cause the value of the term Q/QA28 to be very large, as would be the case for a spring runoff event. However, because of the variable nature of the warm season flow as caused by summer rains, a runoff event may cause the flow to increase markedly above the mean flow for the prior four days (Q/QA4). Many of the short-term runoff events would occur during the summer, but few runoff events would cause the flow to be markedly greater than the average flow for the previous 28 days (Q/QA28).

Season-streamflow parameters included in the regression equations (SUNC, AGSC, M/Q) indicate dilution effects. At higher streamflows the value for these terms becomes small and the terms become of somewhat lesser importance in contrast with low flow periods which would typically occur during mid-winter and late summer.

During the low-flow periods of late summer and early fall, plankton populations are generally much higher than average. It seems likely that the inverse relationship between COD and flow as indicated by the term AGSC may be indirect and what the relationship actually reflects is the large number of plankton in the samples during low-flow periods.

In summary, runoff is the major factor related to the COD of the river water throughout the year. The materials dissolved and suspended in the runoff water are different for the warm and the cold season. During the cold season COD is related to sediment, but during the warm season other materials associated with runoff, probably of biological origin, are of importance.

Ammonia Ammonia concentrations averaged nearly twice as great during the cold season as during the warm season. This was probably related to better nitrification with warmer soil and water temperatures. Application of the t test to the ammonia data indicated that a bi-annual grouping of the data was justified statistically and probably reflected the different extent of nitrification.

The statistical accuracy of the regression was the poorest established for any parameter. Less than ten per cent of the variance was accounted for during the warm season. This lack of any strong relationship may be the result of several factors. Standard Methods (53) indicates that the standard deviation of the laboratory test results for the ammonia analysis was ±16 to 39 per cent of the mean. These results were for synthetic samples which were fixed so that a change in the ammonia concentration was unlikely, rather than for natural waters in which ammonia concentrations may change due to biological activity. During the warm season ammonia has a mean concentration of 0.27 mg/l and a standard deviation of 0.24 mg/l. Compared to the warm season mean of 0.27 mg/l, laboratory variation of 39 per cent would be equivalent to

0.11 mg/l or about half the standard deviation recorded for all the warm season data. Thus a considerable portion of the variance could be accounted for by lack of precision in the analysis.

Compounding the problem are the interactions of ammonia with aquatic life. Ammonia plays an important part in the biochemistry of aquatic plants, algae, and certain bacteria. During the sampling procedure, the water sample to be analyzed for ammonia is fixed with acid, reducing the possibility of the change of ammonia to other nitrogen forms. If this fixing process is inadequate, ammonia may be metabolized and converted to nitrate or organic nitrogen. In the past, water samples have been held for some time before their analysis because of laboratory limitations and this would provide time for a change in concentration to occur. These problems would be especially severe during the warm season when the water temperature averages about 20°C.

Another problem which may be related to the lack of variance accounted for are the different analytical procedures used and the number of different analysts participating in the analyses. Several different analytical techniques have been used for ammonia analysis during the six-year period of this study as new and improved methods became available. Variations in the results of the ammonia analysis associated with these two situations are difficult to assess, but could be significant.

Ammonia data, considered on an annual basis, were grouped in different ways in an attempt to determine if statistical precision would appear to be related to analytical techniques. From 1967 to 1971, water samples were analyzed for ammonia by direct nesslerization (53, p. 226). After 1971 an automated phenate method was used (53).

The results of the regression analysis of the data considered on an annual basis for the two periods revealed better statistical accuracy for the latter two years, as noted in Table 20.

Table 20. Results of statistical analyses of ammonia data for several periods

Period of Analysis	Technique	Annual		Biannual			
				Warm season		Cold season	
		R^2 [a]	SE [b]	R^2	SE	R^2	SE
1967 - 1971	- [c]	12.4	0.277	-	-	-	-
1971 - 1973	- [d]	50.1	0.263	-	-	-	-
1968 - 1973	- [c,d]	28.2	0.288	4.6	0.224	51.0	0.286
1967 - 1973	- [c,d]	16.4	0.293	3.7	0.232	30.4	0.322

[a] R^2 values as per cent and SE values taken from Table 35 in Appendix B.

[b] SE = standard error of the estimate.

[c] Direct nesslerization (53, p. 226).

[d] Automated phenate method (53, p. 232).

For the period 1967 to 1971, the R^2 for the regression equation was 12.4 per cent, and for the period 1971 to 1973 it was 50.1 per cent. However, it was expected that statistical relationships could change from year to year because of variations in rainfall, fertilizer application rates and timing of these applications, and unresolved factors. Perhaps these differences were less during the latter period, as it seemed unlikely that the great increase in the R^2 value could be attributed entirely to a change in the analytical procedure. Thus a different view of the data was taken.

Although there was no known reason to believe that the data from the first year of the sampling program was in error, it was felt that errors might be more likely since the laboratory and sampling procedures were just being established. Results of the regression analysis for the period, 1968-1973, gave an R^2 value of 28.2 per cent for the data considered on an annual basis. The data were then examined on a biannual basis. The results of these analyses are summarized in Table 20.

Considered on the annual basis, a considerable portion of the variation does appear to be contained in data from the first year of the study (1967-1968). On a biannual basis, the R^2 value is much better for the cold season data, but nearly the same for the warm season data. The standard error is improved in both cases. Two conclusions were drawn from these results. The first was that considerable variance was contained in the data collected the first year of the study. However, further analysis of the first year's data would be required to substantiate this conclusion. The second observation was that whatever the source of variation in the data collected during the warm season, it appears to be fairly uniform and not specific to any given year. This may indicate that more care will be required in collecting, preserving, and analyzing the water samples during the warm season. It may be that all the data collected during the warm season should be viewed as approximate and that no great significance can be justified for individual values.

In an attempt to remove some of the variance, the data covering the entire period of the study were smoothed. The R^2 values improved, but this was expected since some of the variance had been removed artificially. The regression equations resulting from the statistical analysis of the raw data and the smoothed data are given below. A list of all regression equations developed from ammonia data is contained in Appendix B, Table 35,

Ammonia, mg/l as N

Warm season

Raw data

$NH_3\text{-}N = 0.0990\ SUN + 0.1139$

Smoothed data

$NH_3\text{-}N = 0.267\ SUN + 0.1195$

Cold season

Raw data

$$NH_3\text{-}N = -0.153 \ln QA7 - 0.319\ Q/QA7 + 0.180\ Q/QA28 - 0.244\ AGF - 0.00034\ M/Q + 1.849$$

Smoothed data

$$NH_3\text{-}N = -0.0887\ QSTD - 0.246\ AGF + 0.0108\ Q/QA28 + 0.5539$$

Although not a great deal of significance could be attached to the data for the warm season, it did appear that a seasonal trend does exist, indicated by the term SUN. Reference to Figure 10, a plot of the seasonal variables against the day of the year, would indicate that peak ammonia concentrations occur about the end of June.

The regression equation developed for the cold season data indicated ammonia relationships with streamflow, runoff, and season. The relationship with flow was inverse, as indicated by the coefficient of QSTD, -0.0887.

This implied that dilution of the ammonia concentration in the streams was associated with increased streamflow. During the winter the principal source of ammonia in streams was the effluents from wastewater treatment plants. Because of the low water temperature, little nitrification occurred in the stream. Aquatic plant growth and bacterial activity which could convert ammonia to other forms of nitrogen when the river was warmer, was minimal. Since the wastewater effluent contributions were relatively uniform, their effects would be greatest during low flow periods. Any increase in streamflow would tend to dilute the ammonia in the streams.

The implications of the runoff term (Q/QA28) and the seasonal term (AGF) were considered jointly. Their respective regression coefficients,

+ 0.108 Q/QA28 - 0.245 AGF

imply an increased concentration which is associated with runoff and an average seasonal maximum occurring in the spring about April (see Figure 10). This is in accordance with the understanding of ammonia contributions to the stream. Animal wastes accumulate on the watershed during the winter on the frozen soil surface. Because of the cold temperature, little nitrification occurs in the soil. Snowmelt and spring rains wash the wastes which contain ammonia into streams. Since the soil and water are still cold, little nitrification occurs. The concentration of the ammonia observed in the river is increased, although diluted by the higher streamflow.

<u>Organic nitrogen</u>   Very high organic nitrogen concentrations frequently accompanied runoff. These high concentrations, sometimes several standard deviations greater than the mean, tended to skew the distribution and bias the regression. It was felt that omission of these outliers would provide a better evaluation of the more general relationship of organic nitrogen to the explanatory parameters. Thus the data were treated in several ways in order to determine the different relationships. Prior to regression analysis the data were divided according to warm and cold seasons on the basis of the *t* test. A

complete listing of the resulting regression equations is given in Appendix B, Table 36.

During the six-year study period, five organic-nitrogen values were recorded which exceeded 3.86 mg/l, or about 3.3 standard deviations greater than the annual mean. Three of these values were recorded during the warm season. The regression equation developed from the data set with the three values omitted accounted for 12.3 per cent of the variance. This R^2 value was much lower than when the three values were included, 37.4 per cent. This would indicate a rather severe biasing of the regression equation by the three outliers. Consideration of the regression equations, shown below, developed from the two treatments reveals few basic differences.

<u>Organic nitrogen, mg/l as N - Warm season</u>

Complete data set

$$\text{Org-N} = 3.04 \ln \text{Q+2} - 4.27 \ln \text{Q+1} + 0.354\ \text{QSTD} + 0.657 \ln \text{SED} - 0.00416\ \text{M/Q} + 6.916$$

Outliers greater than 3.87 mg/l omitted

$$\text{Org-N} = -0.253 \ln \text{Q-5} + 0.270 \ln \text{SED} + 1.464$$

Both equations contain streamflow and sediment parameters only. Although the regression equation developed from the complete data set is more complex due to the fitting of the outliers, the qualitative aspects of both equations indicated that these two parameters were the major ones associated with organic nitrogen.

The relationship with sediment was probably an indirect indication that runoff water which washed sediment into streams also washed materials containing organic nitrogen into streams. These kinds of materials would include animal waste, crop residues, and other types of plant debris. This would appear to indicate that a portion of the sediment is of biological origin, a reasonable conclusion for watersheds used principally for agricultural operations. It is of interest that the regression coefficient of sediment is less than one in both equations. Since suspended sediment concentrations were transformed using logarithms, a strictly linear relationship did not exist between organic nitrogen and sediment. This is reasonable. At lower sediment concentrations a large proportion of the sediment could be expected to be of biological origin. During periods of very high sediment load in the river, as would occur during intense runoff, most of the sediment would consist of soil and sand particles of an inorganic nature and would contribute to a lesser extent to the organic nitrogen concentration in the river. Streamflow is inversely related to organic nitrogen. This may be interpreted in two ways, although both may be true. While high streamflows may wash additional organic nitrogen into streams, the resulting concentration in the river may decrease due to dilution. The relationship may also be indicative of the effects of high plankton populations during low flow periods because of the organic nitrogen content of the plankton.

For example, during the late spring or early summer organic nitrogen concentrations were typically higher than average. Plant residues and animal wastes which had accumulated over the winter would contribute to these higher concentrations. Although the streamflow at this time of year was high, sediment concentrations in the river were also high. The regression equation reflects the interaction of these two parameters. Somewhat higher than average organic nitrogen was again observed about September. Streamflow and sediment concentration were lower at this time of year, but plankton populations were frequently very high. Thus the inverse relationship between organic nitrogen and streamflow may reflect the higher populations of plankton which would be a source of organic nitrogen.

Only two organic nitrogen values were greater than 3.87 mg/l during the cold season. Omitting these two values resulted in a drop in the R^2 value from 32.7 per cent to 12.2 per cent, an indication that outliers did tend to bias the regression as was the case for the regression equations developed from the warm season data. Comparison of the two equations shown below and in Appendix B, Table 36, does indicate some basic similarities, which are not readily apparent.

Organic nitrogen, mg/l as N - Cold season

Complete data set

$$\text{Org-N} = 0.0509\ \text{TEMP} + 0.0108\ \text{AGSC} + 0.2403$$

Outliers greater than 3.87 mg/l omitted

$$\text{Org-N} = -0.154\ \ln Q\text{-}5 + 0.390\ \text{SUN} + 1.488$$

The temperature parameter included in the regression equation for the complete data set would be similar to the season parameter, SUN, in the regression equation for which the outliers are omitted. On an annual basis, the correlation between temperature and the parameter, SUN, is 0.83. The maximum temperature during the cold season is, by definition, 10°C, resulting in a maximum value of about 0.51 for the term in the regression equation.

$$+0.0509 \text{ TEMP} = 0.0509 \times 10 = 0.509$$

In the other equation, the maximum value of SUN during the cold season, 1.5 occurs at the end of the cold season. When this is substituted into the equation, a maximum value of about 0.55 is indicated.

$$0.390 \text{ SUN} = 0.390 \times 1.5 = 0.55$$

Temperature, of course, is more dynamic than the sine function, SUN, and would allow more variance to be accounted for in the regression including the outliers. The inclusion of the temperature parameter is probably closely associated with the high organic nitrogen concentrations which would accompany snowmelt runoff on warm spring days, and the early spring rains.

For the two treatments, with and without outliers, the streamflow terms of similarity are, respectively, AGSC and Q-5. Both indicate an inverse relationship with flow, or dilution effects. When the outliers are omitted, the relationship is simply one of dilution - higher flow is associated with lower organic nitrogen concentrations. The relationship is not strictly linear since the coefficient of Q^{-5} is less than one and the logarithm of streamflow is used.

$$-0.157 \ln Q\text{-}5$$

This could also be written

$$-\ln(Q\text{-}5)^{0.157}$$

It is apparent that at higher streamflows the dilution effect is not as great as for lower streamflows. For example, a 100-fold increase in flow would result in only a doubling of the value of the streamflow term.

When the outliers are included, the term, AGSC, in the regression equation indicates a similar effect, although there is a seasonal variation which is associated with higher organic nitrogen values in the early spring. The term, AGSC, is additive to the equation

$$+ 0.0108 \text{ AGSC}$$

By reference to the definition of the term AGSC, this would be equivalent to

$$+ \frac{108}{Q} \text{ AGS}$$

The maximum value of AGS during the cold season is 2 (April 20). Thus in the spring only for flows less than 216 cfs will the term be greater than one. During April, a typical average flow would be 6000 cfs. At this flow, the contribution of the term AGSC to the equation would be bearly negligible (0.036). In December, an average flow of 1000 cfs and a value of 0.2 for the term AGS would result in an even smaller contribution (0.0216) to the equation. It is clear that this term will have little significance except during low flow periods.

That no sediment term is included in the regression equations is of interest. This may indicate that the sediment in the river during the cold season consists primarily of inorganic solids.

In summary, the major factors related to the concentration of organic nitrogen during the warm season are flow and suspended sediment. The relationship with streamflow is inverse, and may be related to dilution at higher flows and to the larger number of plankton in the stream during periods of low flow. During the cold season streamflow and temperature, associated with a seasonal trend, are the major factors related to organic nitrogen in the river. The inverse relationship of streamflow indicates dilution. Temperature and season are related directly and indicate higher organic nitrogen concentrations in the spring and fall in contrast with the winter. The regression was biased strongly by outliers. When these outliers were removed, only about 12 per cent of the variance was accounted for. This would indicate that the parameters included in regression analysis are significantly related to the organic nitrogen concentration but may not be the major factors related to the variations in the measured concentration of organic nitrogen in the river.

Nitrate   Variations of the nitrate concentration in the river were relatively uniform from year to year. Lowest concentrations were nearly always observed in the late summer during the months of August or September. During years when the flow at this time was low (less than 500 to 600 cfs), nitrate was virtually absent. Typical summer levels were four to five mg/l. This depletion may have been the result of plankton uptake since plankton populations were frequently very high at this time. By late September or October, nitrate levels generally increased to at least average concentrations, and continued to increase through the month of December or January when high concentrations (8 to 12 mg/l) were recorded. Year to year comparisons indicated that concentrations during these winter months were greatest in those years when the average winter streamflow was greatest.

Depending on the length and timing of snowmelt and the spring rains, the concentration declined from about 10 mg/l to 4 mg/l or less during the peak runoff months of February through April. As the quantity of runoff diminished, concentrations again increased, although showing considerable variation from week to week, to levels equal to or greater than those recorded during the winter.

Thus nitrate concentrations appeared to be influenced by at least three factors. Lowest nitrate levels observed during late summer when streamflow was also low appeared to be related to plankton activity. Winter nitrate levels were a function of streamflow. Groundwater is the principal source of streamflow during the winter when the soil surface is frozen and the river is ice covered. Nitrogen compounds dissolved in soil water infiltrate into the soil profile during the summer. There, by the action of nitrifying bacteria, these compounds are converted to nitrate. This storehouse of nitrate appears to be gradually washed from the soil profile during the winter by groundwater as it moves to nearby streams. With greater groundwater flow, as manifested in higher streamflows, increasingly large concentrations of nitrate occur in the river. Whether the higher nitrate levels are simply a function of higher winter streamflows, or whether the levels are more highly related to greater rainfall and the subsequent percolation of this rainfall into the soil profile during the summer is not entirely clear.

Nitrate concentrations observed during the runoff period in late winter and spring appear to be a function of the quantity of runoff water which is far in excess of the groundwater contribution. The dilution effect of the runoff water is the principal factor controlling the nitrate concentration at this time. Once the runoff events diminish, dilution effects are of less importance. Groundwater inputs of nitrate then appear to become the factor controlling nitrate concentration.

Because of the differences in nitrate levels during the cold and the warm seasons, it was anticipated that these seasons should be considered separately. Application of the $\underline{t}$ test indicated that this would be justified statistically.

Statistical analysis of the nitrate data indicated that a large percentage of the variance could be accounted for by the regression equations. The R^2 values for both the warm and the cold seasons were greater than 75 per cent. Based on the R^2 values it was anticipated that a rather accurate evaluation of the relationship between nitrate concentration and the explanatory parameters could be made.

The nitrate data were treated in a number of ways prior to regression analysis. Table 37 in Appendix B lists the resulting regression equations. Because there was considerable variation in the nitrate concentration the data were smoothed. The R^2 value improved somewhat compared with the non-smoothed data. However, it was felt that there was not adequate justification for this procedure. The observed variations did appear to be real, and not the result of sampling or analytical errors. Smoothing the concentration data which had been converted to quantity indicated that the procedure had resulted in a significantly different distribution of the data. Thus both smoothing routines were rejected as meaningful ways of handling the data.

Transforming the concentration data using natural logarithms improved the R^2 value as compared with the untransformed data for the cold season regression (82.6% and 73.8%), but not for the warm season (73.1% and 76.4%). However, the greater simplicity of the warm season regression which was based on the transformed data made it the more preferable of the two since the R^2 values differed by only a small percentage.

Reference to Table 37 in Appendix B or to the regression equations shown below indicate that the principal association of nitrate was with streamflow.

Nitrate mg/l as N, natural log

Warm season

$$\ln NO_3\text{-}N = 2.42 \ln Q - 0.766\ QSTD - 17.07$$

Cold season

$$\ln NO_3\text{-}N = 0.264 \ln Q{+}4 - 0.224\ Q/QA28 - 1.24\ SUN$$
$$- 0.0378\ SUNC - 0.0595\ AGFC + 1.0522$$

Both equations indicate a direct relationship between nitrate concentration and streamflow. For ease of illustration, the regression equation for the warm season data will be used to demonstrate this relationship. the predicted nitrate concentration for various streamflows is given in Table 21.

Table 21. Predicted nitrate concentration at various streamflows - warm season

Flow, cfs	Nitrate, mg/l (predicted)
100	0.00
1000	0.49
10000	11.0

No seasonal term is included in the regression equations for the warm season data. This would seem to indicate that any of the apparent seasonal trends which have been observed are more directly related to flow.

The regression equations for the untransformed data, shown below, indicate a more complex relationship for the warm season.

Nitrate, mg/l as N

Warm season

$$NO_3\text{-}N = 3.16 \ln Q + 3.74 \ln Q\text{-}1 - 2.16\ QSTD + 1.05\ Q/QA21 + 0.0229\ M/Q - 48.23$$

Cold season

$$NO_3\text{-}N = -3.38 \ln Q\text{-}5 + 5.21 \ln QA7 - 0.790\ Q/QA28 - 3.83\ SUN - 0.0281\ AGFC - 4.705$$

Two of the flow terms in the regression equation for the warm season, Q and Q-1, indicate a direct relationship between nitrate and flow. two other flow terms, QSTD and M/Q, indicate an inverse relationship, the former becoming more negative and the latter less positive with increased flow. Both indicate dilution with increases in streamflow. These two terms appear to provide a correction to permit the equation to conform to a variety of flow conditions.

The runoff term, Q/QA21 contained in the regression equation for the warm season, has a relatively small coefficient and the product of this term times its coefficient would be small in comparison to that contributed by the flow terms. For example, the maximum value of Q/QA21 times its coefficient, about 3.2, was much less than the sum of the flow terms times their coefficients, 51.0, for a sampling date selected from the warm season data. It would appear that runoff does not contribute greatly to the nitrate concentration in the river.

By way of further illustration, the value of the term Q/QA21 was as large as three in the warm season on only five occasions during the six-year period of this study. Most values were equal to one, or less. Thus, the contribution of this runoff term to the regression equation would usually be about one, or less.

Perhaps of more significance is the sign of the regression coefficient. The positive sign indicates that higher nitrate concentrations are associated with runoff events. This relationship may be indirect. As an example, part of the rainfall which would cause a change in flow would percolate through the soil, dissolve nitrates within the soil profile, and then enter the groundwater flow. Some time later the groundwater, containing a higher nitrate concentration than the stream (67), would feed into the stream and, in this way, increase the nitrate concentration.

The significance of the specific runoff term selected by the regression procedure, that is, the ratio of the flow on the sampling date to the mean flow for the prior 21 days, may indicate that a several week period of little rainfall is required in order for that soil permeability to be sufficiently great to permit the process just described to be of importance. In other words, if two rainfall events occur within a period of less than several weeks, the soil moisture may be too great to permit large amounts of percolation through the soil profile. In this case the streamflow parameters would be the controlling factors in the regression equation.

The regression equations describing nitrate relationships during the cold season are more complex, particularly in comparison to the warm season regression equation incorporating the log transformed data.

The equations for the untransformed and the transformed data listed previously, contain streamflow, runoff, and season parameters. As for the warm season, high streamflows are associated with high nitrate levels. One important difference does emerge, however.

The sign of the regression coefficient of the runoff term is negative. This appears to indicate dilution by runoff water and is consistent with the hypothesis offered earlier concerning dilution by spring runoff. During the cold season runoff period, typically February through April, the soil is either frozen or has a high moisture content. The amount of water infiltrating through the soil profile and eventually entering the stream would usually be small relative to the quantity of runoff entering the stream. The principal effect of snowmelt on rainfall runoff would be that of diluting the concentration of nitrate in the river.

Vastly different environmental conditions exist during the cold season-relatively constant flow in winter, and heavy runoff with great changes in flow during early spring. The season variables may permit the regression to accommodate these variations and the parameters probably indicate general trends regarding the nitrate concentration in the river. Superimposed on these general trends, then, are the effects described by the other parameters included in the regression equation.

Because of the differences in the regression equations for the warm season and the cold season data, and the high proportion of the variance accounted for by each equation, it was of interest to test the hypothesis that the variables for one season would provide an equally good estimate of nitrate concentration for the other season. To do this, parameters which were selected by the regression procedure for one season were forced into the regression equation for the other season. These results are shown in Table 22. In both cases, interchanging the parameters resulted in a poorer fit of the regression line, as indicated by R^2, F, and the standard error.

Considering Equations 1 and 3, and 2 and 4 it will be noted that runoff parameters retained the same sign for a given season. This would appear to indicate that runoff relationships are specific to each season. For the equations containing the same parameters, that is equations 1 and 4, and 2 and 3, differences in both the sign and the value of the coefficient are evident for the season and the runoff parameters. Thus the hypothesis that parameters for one season would apply equally well to the other season was rejected.

A second hypothesis was tested - that the regression equations could be simplified by substitution of streamflow on the sampling date for streamflow on days other than the sampling date. For example in Table 22, Equation 2, the parameter ln Q would be substituted for the parameters ln Q-5 and ln QA7. The results of this simplification of the regression equation are given in Table 23 and can be compared with Equations 1 and 2 in Table 22.

Table 22. Results of nitrate regression when parameters for warm season and cold season are reversed[a]

Season	Regression equation	R^2, %	F	SE
1. warm	3.16 ln Q + 3.74 ln Q-1 - 2.16 QSTD + 1.05 Q/QA21 + 0.0229 M/Q - 48.23	76.4	104.8	1.99
2. cold	-3.38 ln Q-5 + 5.21 ln QA7 - 0.790 Q/QA28 - 3.38 SUN - 0.0281 AGFC - 4.705	73.8	73.9	1.71
3. warm[b]	-0.686 ln Q-5 + 3.39 ln QA7 + 1.41 Q/QA28 + 0.657 SUN + 0.0236 AGFC - 18.76	68.0	69.0	2.32
4. cold[c]	1.06 ln Q + 1.43 ln Q-1 - 0.917 QSTD - 1.08 Q/QA21 + 0.00010 M/Q - 10.07	59.3	38.2	2.13

[a] Untransformed nitrate data entered into the regression.

[b] Cold season parameters forced into regression for warm season data.

[c] Warm season parameters forced into regression for cold season data.

Table 23. Substitution of streamflow on the sampling date for streamflow on days other than the sampling date

Season	Regression equation	R^2,%	F	SE
1. warm	6.80 ln Q - 2.07 QSTD + 0.508 Q/QA21 + 0.0225 M/Q - 47.10	74.1	116	2.08
2. cold	1.87 ln Q - 0.923 Q/QA28 - 3.77 SUN -0 0.256 AGFC - 4.820	72.3	87.8	1.74

The statistical results did not change greatly, although the F ratios in both cases were greater because fewer parameters had been included in the regression equation. Although the regression coefficients differed slightly, the same signs were retained. It is of interest to note that the sum of the coefficients of the terms Q and Q-1, respectively 3.16 and 3.74, for the warm season regression (Table 22, Equation 1) was approximately equal to the coefficient of Q in the simplified equation, 6.80 (Table 23, Equation 1). The same was found to be true for the regression equations for the cold season data. Based on the general similarity of the regression equations and the statistical results, the hypothesis was retained that the simplified regression equations would adequately describe the relationships between nitrate concentration and the explanatory variables.

In summary, the nitrate concentration of the river water is related directly to streamflow. Higher nitrate concentrations are associated with higher streamflows. Surface runoff does not appear to be a major factor regulating the nitrate concentration. Rather it is thought that groundwater contributions of nitrate are the principal source. Although there are some seasonal differences in nitrate relationships, the basic relationships with flow are the same.

<u>Phosphorus</u> Phosphorus enters the aquatic environment in two ways. Point source discharges of treated domestic and industrial wastewater contribute phosphorus to streams at a relatively steady rate throughout the year. In contrast, non-point-source discharges such as runoff from fields and feedlots contribute phosphorus to streams only during periods of snowmelt and rainfall runoff. Estimating the phosphorus contributions from point source discharges is largely a matter of collecting periodically samples from individual sources in a watershed. Analysis of these samples will provide a reasonably good estimate of the total phosphorus contributions. However, estimating phosphorus contributions from non-point-source discharges during runoff is a more difficult task because of the problems of obtaining representative samples in an extensive river basin system.

One approach to the estimation of phosphorus contribution from non-point-source discharges is the collection and analysis of water samples, and evaluation of the causes of variations in the phosphorus concentration of the river at a given site. Relating statistically the phosphorus concentration to runoff patterns would provide a means of estimating contributions under different types of climatological and seasonal conditions. This approach is essentially that taken in this research.

Phosphorus may be present in several forms in water. Two forms were considered in this research: total phosphorus and ortho or filtratable phosphorus. Phosphorus compounds adsorb strongly onto sediment particles, particularly the clays. Adsorption may occur on the land or within the aquatic environment. For example, Taylor et al. (58) have shown that dissolved phosphorus in a stream is depleted as a result of adsorption onto sediments which were derived from subsoils and streambanks.

Different analytical techniques were used for these two forms of phosphorus. Separation of the ortho-phosphorus was accomplished by filtering the water samples through membrane filters of 0.45 μ pore size (53). The filtrate was then analyzed for ortho-phosphorus using the ascrobic acid method (53). To analyze for total phosphorus, an unfiltered sample was subjected to perchloric acid digestion and then analyzed for phosphorus using the vanadomolybdophosphoric acid colorimetric method (53).

Variations in the total phosphorus concentration do not appear to exhibit any particular seasonal patterns, other than that which is related to runoff which normally occurs in late winter. Because of the tendency of phosphorus to adsorb onto sediment, it was not surprising to find that peak total phosphorus concentrations coincided with peak runoff periods when suspended sediment concentrations were greatest.

In contrast, orthophosphorus concentrations were greatest during "clear water" periods in the winter. It is felt that these high concentrations, two to three times as great as the annual mean, were the result of the interaction of several factors. Because of little surface runoff, the impact of point-source discharges was more evident since dilution effects would be expected to be smaller. Suspended sediment concentration would also be low and a lesser opportunity would exist for depletion of the phosphorus due to adsorption.

Lowest concentrations of ortho-phosphorus occurred in September when streamflow was low. Since plankton populations were greatest at this time, it is suggested that the low phosphorus concentration was the result of depletion by the large numbers of plankton. This apparent depletion pattern is identical to that observed for nitrate and similar to that for ammonia (5).

Differences in the seasonal distribution of the two forms of phosphorus were investigated. Application of the t test to the phosphorus data indicated that for ortho-phosphorus the biannual season distributions were significantly different from the annual grouping, and could be considered separately. No significant difference existed for the total phosphorus data. However, it should be noted that the number of data for total phosphorous was about a third of that for ortho-phosphorus since analysis for total phosphorus did not begin until 1971.

The regression equation developed for total phosphorus indicated relationships with the runoff parameters Q/QA2 and Q/QA21, and with the seasonal variable AGF, as shown below and in Appendix B, Table 38.

Total phosphate, mg/l as PO_4

$$Total\text{-}PO_4 = 1.03\ Q/QA2 + 0.344\ Q/QA21 - 0.216\ AGF + 0.01046$$

Reference to Table 38 in Appendix B indicates similar relationships for the warm and the cold seasons. The negative coefficient for the seasonal term AGF indicated an inverse relationship with this parameter. As shown in Figure 10, the seasonal variable AGF goes to a minimum about the end of March. Taking into consideration this inverse relationship, the indication is that higher concentrations of total phosphorus would be expected about the end of March. Deviations from this general trend are contributed by the runoff terms which provide the dynamics of the regression equation.

Because runoff appears to be an important factor controlling the concentration of total phosphorus an example will be provided to aid in the interpretation of the runoff parameters. Between February 26 and March 3, 1972, one to two inches of rain fell on the Des Moines River Basin north of Boone, Iowa. Lesser amounts fell between Boone and Saylorville, Iowa where the sampling site was located. Ten inches of snow was on the ground at Fort Dodge, Iowa on February 25. Because of the heavy rain and warmer air temperatures, the snow had completely melted by March 7.

Prior to the rainfall and snowmelt, the flow of the Des Moines River at Saylorville had varied between 210 and 230 cfs. When a river water sample was collected on February 24, the flow at Saylorville was

230 cfs. By the time the water sample was collected the following week, March 2, runoff had swelled the streamflow to 1600 cfs. The suspended sediment concentration during that interval increased dramatically from 6 mg/l to 151 mg/l.

The value of the two runoff parameters included in the regression equation, Q/QA2 and Q/QA21, were respectively, 1.00 and 1.07 on February 24, and indicated a relatively steady flow condition in the river. By the following week, because of the heavy runoff and much higher flow, the runoff parameters Q/QA2 and Q/QA21, had increased, respectively, to 1.68 and 5.83 and indicated that heavy runoff had occurred.

The calculated and the measured values for total phosphate on February 24, prior to runoff, were respectively 1.4 mg/l and 1.6 mg/l. On March 2, following the runoff event, the calculated and measured values were 3.6 mg/l and 2.8 mg/l. The contribution of the term AGF, about 0.01 mg/l, to the regression estimate was very small in comparison to that provided by the sum of the runoff terms, 1.4 mg/l and 3.6 mg/l.

Although the suspended sediment concentration was expected to be related strongly to the phosphorus concentration, the sediment parameter was not included in the regression equation. It is felt that the reason for this was the correlation between the sediment concentration and runoff terms, r = 0.45 warm season and r = 0.35 cold season as listed in Appendix A. Inclusion of the more highly related runoff terms would tend to cause the exclusion of the sediment term for statistical reasons.

It is concluded that runoff is the principal factor controlling the total phosphate concentration in the river. Although sediment may contribute to the total phosphate concentration, this contribution is effectively accounted for by the runoff terms.

The relationships indicated by the regression equations for ortho-phosphorus, as shown below and in Table 39 in Appendix B, were somewhat different than for total phosphorus.

Ortho-phosphate, mg/l as PO_4

Warm season

$$O\text{-}PO_4 = 0.137 \ln Q\text{-}3 + 0.107\ Q/QA28 - 0.0920\ AGS + 0.00407\ SUNC - 0.8736$$

Cold season

$$O\text{-}PO_4 = -36.32\ SUN + 71.47\ AGS + 61.65\ AGF - 0.000226\ M/Q - 96.536$$

During the warm season streamflow, runoff, and seasonal parameters are included in the regression equation. Increases in flow and runoff both appear to be associated with increases in ortho-phosphorus concentration during the warm season. Reference to Figure 10 shows that the season variable AGS goes to a minimum about October. The negative coefficient of AGS, -0.0920, indicates an inverse relationship between ortho-phosphorus and AGS. The range of values for AGS is 0.0 to -0.184. Largest negative values occur about mid-April, and smallest negative values occur about mid-October. If the term AGS is considered apart from the other terms in the equation, smallest concentrations would be indicated in the spring when runoff washes large amounts of sediment into the streams. It has been shown that sediment particles will adsorb phosphates (58). Highest phosphorus concentrations would be indicated for October when AGS goes to its smallest negative value. During this month, concentrations of ortho-phosphate are typically greater than those measured during the summer. Plankton die-off and runoff containing decomposing harvest residues may contribute to these higher concentrations. But, during the previous month, September, lowest concentrations are measured frequently. Thus the trend indicated by the term AGS does not appear to be entirely accurate. It is apparent that isolating individual parameters, and assigning to them specific relationships is not the best approach. All terms in the regression equation must be considered.

For example, although the term AGS approaches its smallest negative value in September, the flow at this time of year is low as shown in Figure 8. Runoff events are infrequent normally and the value of the runoff term, Q/QA28, would also be small. Consider the following hypothetical illustration.

About mid-September a streamflow of 1000 cfs was recorded. It had been declining gradually during the past month, and the value of the runoff term, Q/QA28, was 0.8. The value for the terms AGS and SUN were calculated to be, respectively, 0.15 and 1.2. As calculated from the regression equation representing warm season relationships, the ortho-phosphate concentration was estimated to be 0.19 mg/l. By the following month, October, autumn rainstorms had swelled the streamflow to 3500 cfs, and the runoff term was calculated to be 2.5. The values calculated for the seasonal terms AGS and SUN were, respectively, 0.1 and 0.42. Based on these conditions the concentration of ortho-phosphate was estimated to be 0.51 mg/l.

Thus, when the terms are considered as a group, the relationships indicated by the regression equation do appear to portray what is actually observed.

Variations in ortho-phosphorus concentrations during the cold season do not appear to be related to flow or runoff, but are a function of the three seasonal variables SUN, AGS, and AGF. The term M/Q is of minor importance except at streamflow less than 1000 cfs because of the very small regression coefficient. The indication is that fluctuations in concentration during the winter are fairly uniform from year to year. Streamflow and runoff do not appear to be of great importance.

To summarize, variations in the total phosphorus concentration are strongly influenced by runoff, possibly because of the phosphorus containing materials which are adsorbed into sediment that is washed into the river. Ortho-phosphorus concentration is related to runoff during the warm season, but the relationship includes streamflow and seasonal parameters as well. During the cold season, ortho-phosphorus concentration is related to runoff and sediment only as they are also related to the three seasonal parameters included in the regression equation.

One question which came up during evaluation of the phosphorus regression equations was whether the total phosphorus concentration could ever be less than the ortho-phosphorus concentration. The answer to this question is not easily obtained because of the dissimilarity of the equations. One approach was to select several dates of which samples were collected and which represented a variety of environmental and hydrologic conditions.

Conditions considered included runoff and non-runoff periods during spring, summer, autumn, and winter. In order to determine the nature of the runoff conditions, a stream hydrograph for the six-year sampling period was developed. Runoff and non-runoff periods were then selected for each season. A runoff condition was considered to exist when a sample was collected during the peak or the rising limb of the hydrograph. Non-runoff conditions were considered to exist when a sample was collected during the falling limb of the hydrograph.

The results of this brief, but hopefully, representative analysis are given in Table 24. Both the predicted and the observed values for total-phosphate and ortho-phosphate are listed as an indication of the accuracy of the relationships expressed in the regression equations.

Two conclusions can be drawn from the information presented in Table 24. For the dates selected, no ortho-phosphate concentration was greater than the total phosphate concentration. Because of the limited number of cases selected, it cannot be positively concluded that the estimated ortho-phosphate concentration will never be greater than that for total phosphate. However, this occurrence does seem unlikely. It should be mentioned that, because of the difference in analytical techniques for the analysis of the two forms of phosphorus,

Table 24. Predicted and observed concentrations of total and ortho-phosphate during runoff and non-runoff conditions

Date	Hydrologic condition[a]	$O\text{-}PO_4$ Observed	$O\text{-}PO_4$ Predicted	$Total\text{-}PO_4$ Observed	$Total\text{-}PO_4$ Predicted
<u>Winter</u>					
1-26-72	NRO	1.0	0.98	1.2	1.22
12-28-72	RO	0.6	0.68	0.9	1.16
<u>Spring</u>					
4-4-72	NRO	0.4	0.41	1.2	1.19
5-3-72	RO	0.2	0.35	2.8	2.53
<u>Summer</u>					
6-28-72	NRO	0.0	0.15	1.1	0.90
7-5-73	RO	0.3	0.26	1.6	1.61
<u>Fall</u>					
8-9-72	NRO	0.4	0.51	1.9	1.56
9-17-71	RO	0.2	0.13	0.7	1.10

[a]NRO denotes non-runoff, RO denotes runoff.

laboratory results do occasionally indicate a higher ortho-phosphate concentration than that for total phosphate.

A second conclusion drawn from the information in Table 24 is that good agreement exists between the predicted and observed concentration of the two forms of phosphorus. Again the limited number of data should be taken into consideration.

Fecal coliform Livestock waste is the principal source of fecal coliform organisms in the Des Moines River, as measured at the sampling site near Saylorville, Iowa. Since the waste production of livestock is by far in excess of that for the human population in the basin, it is suggested that contributions of coliform organisms from municipal wastewater treatment plants is of significance in regard to water quality only during the winter when runoff events are infrequent. At other times, livestock waste is almost certainly the principal source.

In order for livestock waste to have a significant impact on a stream a runoff event must occur. Runoff events may result from snowmelt, rainfall, or a combination of the two. In the interval between runoff events, animal waste accumulates on the soil surface. When the next runoff event occurs, much of the accumulated waste is washed into nearby streams, increasing greatly the number of coliform organisms in the stream. This increase is frequently as large as several orders of magnitude.

In order to visualize the variation in coliform counts during the year, the geometric mean of all coliform counts recorded for a given month during the period 1970 to 1973 was plotted against the month of the year (Figure 12). Because of the undue influence of one or two very high counts, the variations shown must be considered as approximations only. Although the use of the geometric mean tended to reduce the influence of counts as great as 20,000 organisms/100 ml, very large

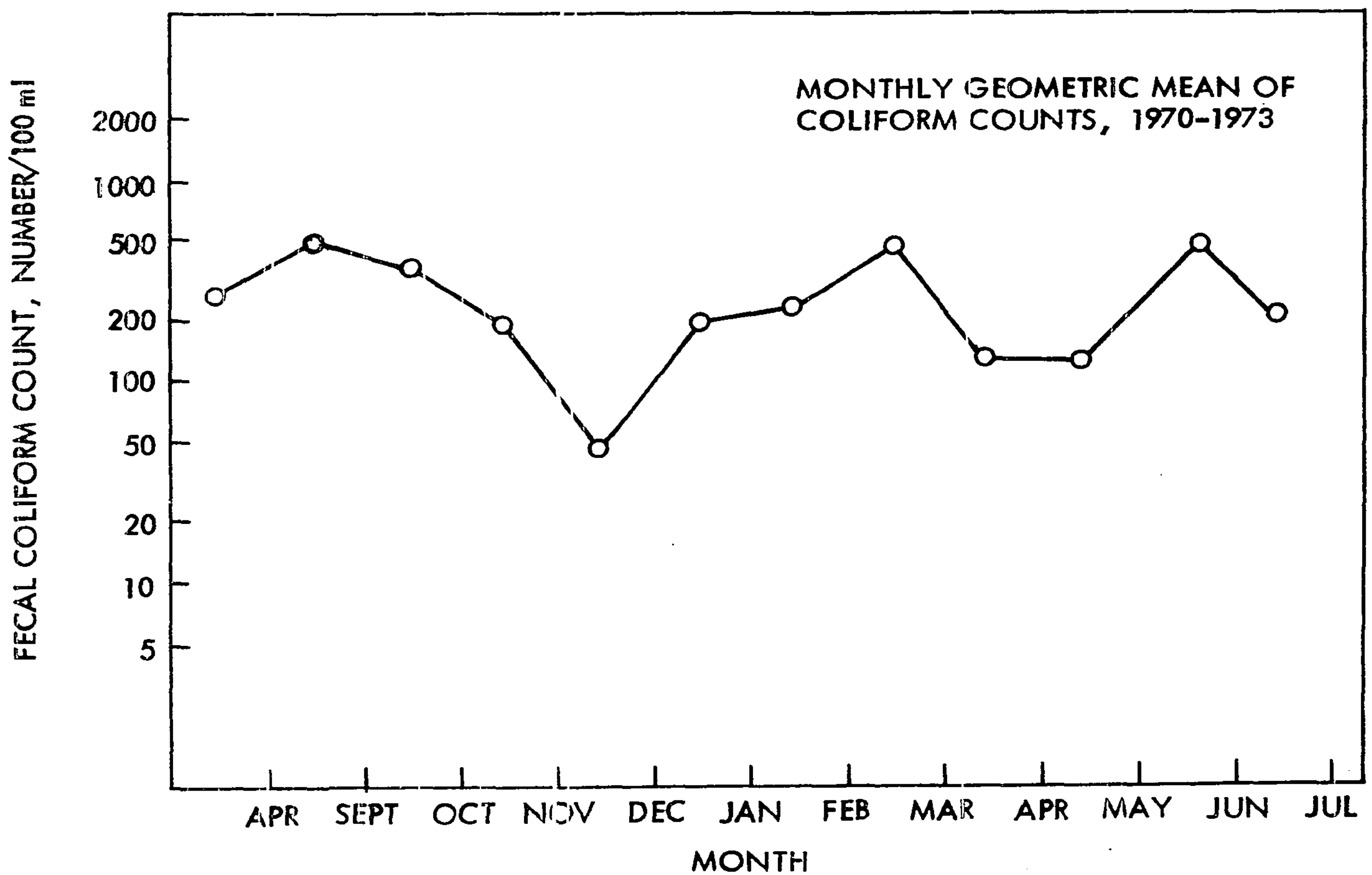


Figure 12. Variation in fecal coliform counts by month

values still tended to bias the average.

The peaks shown in Figure 12 correspond fairly well with expected runoff periods. The exception may be the peak in early autumn. An explanation for this peak may be that streamflow is generally low at this time, and that runoff events would be expected to have a greater impact on water quality in regard to the coliform count. The two large dips in the graph correspond to two different sets of circumstances. In December, normally a very cold month, little runoff occurs. Contributions of coliform organisms are mainly from domestic sources. In April and May, a second dip occurs. This is probably the result of depletion and dilution. Most of the animal waste which accumulated during the winter had been previously washed into the river, as indicated by the peak in March. Additional amounts washed into the river would tend to be diluted by the normally high streamflow.

The central point to be drawn from the discussion in this section is the association between runoff events and higher coliform counts. This association appeared to apply to all times during the year except for late spring events, as explained above.

Animal waste production is continuous throughout the year. Likewise runoff events which wash this waste to nearby streams occur during both seasons. As a consequence, although a difference existed between the biannual means, application of the t test to the fecal coliform data indicated that this difference was not significant. Regression equations developed from the log transformed and the untransformed data considered on an annual basis are given below.

Fecal coliform count, number/100 ml - annual

Transformed data, natural logs

$$\ln \text{F. coli} = 0.989 \ln Q{+}4 + 1.22\ Q/QA28 + 0.666 \ln SED + 0.0491\ AGFC - 8.352$$

Untransformed data

$$\text{F. coli} = -2982\ DQ3/Q + 7283\ Q/QA3 - 1103\ Q/QA14 - 5697$$

Results of the regression analyses for several other ways of handling the data are given in Appendix B, Table 40.

The two regression equations given above differed greatly in the amount of variance accounted for. Using the transformed data, the R^2 value for the regression was 26.2 per cent, whereas the regression based on the raw or untransformed data had an R^2 value of 68.5 per cent. It was felt that the equation representing the untransformed data was biased by several very large values in excess of several standard deviations greater than the mean. Regression analysis of a data set for which coliform counts greater than about 8000 organisms/100 ml were eliminated gave an R^2 of 31.4 per cent, as shown in Appendix B, Table 40. It was felt that this indicated that although the regression equation using the untransformed data gave a higher R^2, the equation developed for the transformed data would be representative of a greater variety of conditions. In this context, it is interesting to examine the two equations in greater detail.

The regression equation for the untransformed data contains only runoff parameters. (No attempt will be made to suggest specific associations for each term. Rather the equation will be considered in its entirety.) It would appear that this equation, precisely because it is biased by several large coliform counts, should provide a better assessment of the relationships during periods when existing environmental conditions would be expected to cause very high coliform counts. On the other hand, the equation developed from the transformed data would be of more general application.

Six examples are provided as an illustration in Figure 13 and listed in Table 25. Examples a, b, and c shown are considered to represent periods of very high runoff, as would occur during a heavy rainstorm or widespread snowmelt. Examples d and e are considered to represent non-runoff conditions. Example f represents a hydrologic condition for which the streamflow is very high because of a runoff event. However, this event immediately follows a previous event. This succession appears to be of some importance in regard to the numbers of coliform organisms observed in the river.

Hydrograph segments used to illustrate the types of hydrologic conditions are shown in Figure 13.

Table 25 lists the predicted on the observed coliform counts, and the streamflow for each of the hydrologic conditions. While these six examples cannot be representative of the great variety of hydrologic conditions observed in the river, they do illustrate the strengths and the limitations of the regression equations.

Runoff conditions b and c were chosen because the coliform counts on these two occasions were in excess of three standard deviations from

Figure 13. Hydrologic conditions on several sampling dates

a. Runoff, July 19, 1972

b. Runoff, August 28, 1972

c. Runoff, September 28, 1972

d. Non-runoff, May 25, 1973

e. Non-runoff, November 19, 1972

f. Two successive runoff events, March 8, 1973

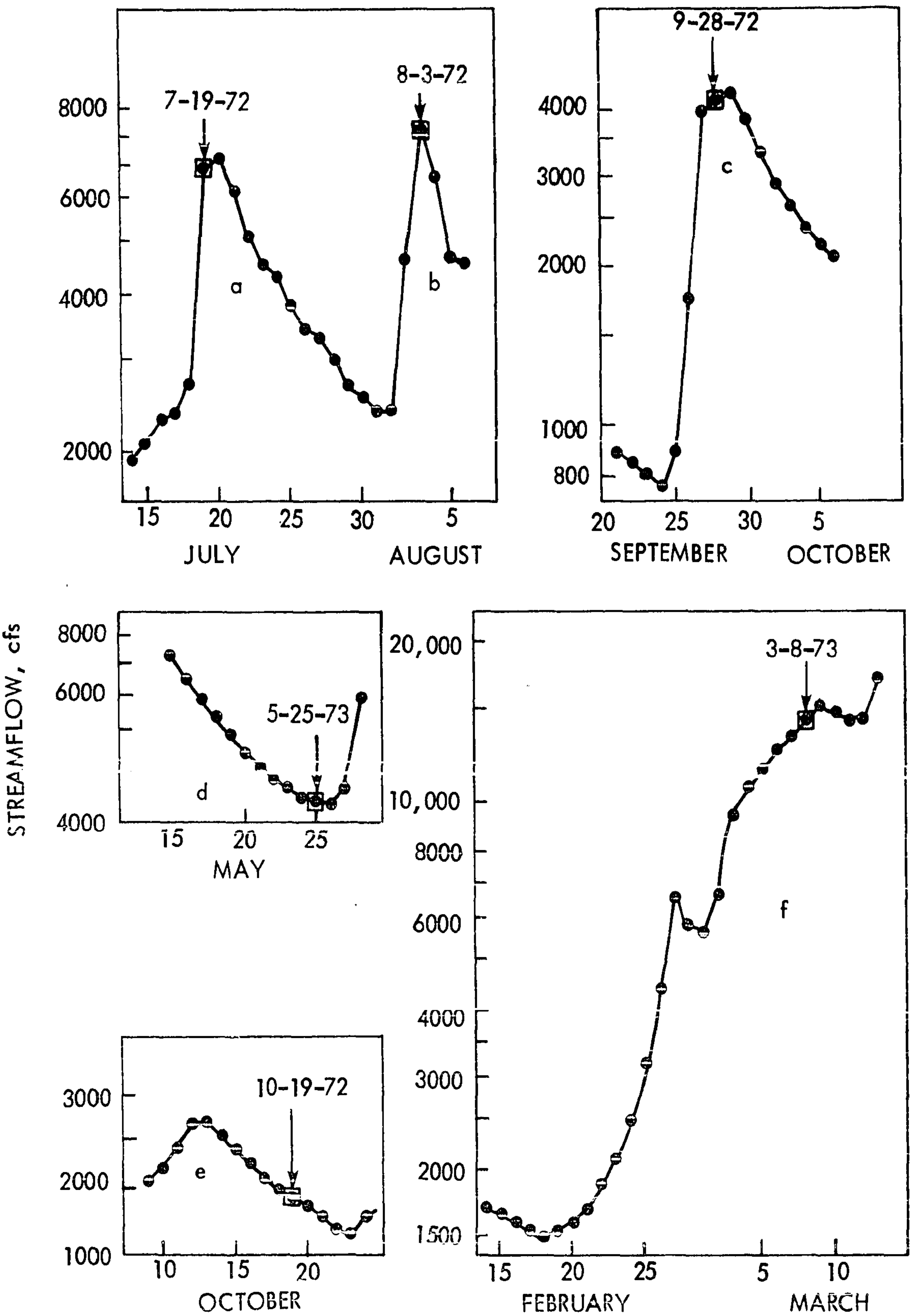

7-19-72
8-3-72
a
b
8000
6000
4000
2000
15
20
25
30
5
JULY
AUGUST
9-28-72
c
4000
3000
2000
1000
800
20
25
30
5
SEPTEMBER
OCTOBER
STREAMFLOW, cfs
5-25-73
d
8000
6000
4000
15
20
25
MAY
10-19-72
e
3000
2000
1000
10
15
20
OCTOBER
3-8-73
f
20,000
10,000
8000
6000
4000
3000
2000
1500
15
20
25
5
10
FEBRUARY
MARCH

Table 25. Predicted and observed fecal coliform counts for several hydrologic conditions

Hydrologic condition	Streamflow cfs	Coliforms/100 ml		
		Observed	Predicted, U[a]	Predicted, T[b]
a. Runoff	6,880	5,000	7,733	3,670
b. Runoff	8,180	12,600	9,962	3,816
c. Runoff	1,050	20,000	14,686	449
d. Non-runoff	4,390	100	510	135
e. Non-runoff	1,750	50	421	65
f. Runoff	14,600	330	4	16,149

[a]U = regression equation based on untransformed data.

[b]T = regression equation based on transformed data.

the mean. These high counts were felt to have biased the regression equation based on the untransformed data. It will be noted that the regression equation for the untransformed data does provide the better estimate. Runoff condition a, although manifesting a large increase in streamflow, for some reason did not appear to result in a coliform count as large as that recorded during runoff condition b. Both equations give reasonable estimates.

Water samples collected during periods of receding flow, non-runoff conditions d and e, contained relatively few coliform organisms. In both cases the regression equation based on the transformed data gave the better estimate, demonstrating its application to the non-runoff hydrologic conditions.

The final example, hydrologic condition f, indicates a large runoff event. The dramatic increase in streamflow was the result of snowmelt and rainfall runoff during February and March. Although a large coliform count might have been expected, the observed count was only 330 coliform/100 ml, as given in Table 25. It can be seen from Figure 13 that two runoff events actually occurred. By the time the sample was collected most of the animal waste had apparently been washed into the stream during the previous event. Additional amounts washed into the stream were probably diluted by streamflow which had swelled to nearly 15,000 cfs. The regression equation for the transformed data indicated an expected count of more than 16,000 coliform/100 ml, far in excess of the number actually observed, while the regression equation for the untransformed data indicated an expected count of just 4 coliforms/100 ml. Neither equation describes this somewhat unusual situation very well.

Because the flow was not changing rapidly, the runoff terms in the regression equation for the untransformed data did not predict a high count. In contrast, the equation for the transformed data was influenced greatly by the very high streamflow. The hydrologic condition also resulted in the streamflow on the sampling date to be more than three times the mean flow for the previous 28 days, represented by the parameter Q/QA28. This increased the expected count even more.

Because of the large amount of error in the estimate for hydrologic condition f, the regression equations developed from the biannual groupings of the data shown below, were also used to calculate the expected coliform count.

Fecal coliform count, number/100 ml

Warm season

F. coli = -2731 DQ3/Q + 7802 Q/QA4 - 4239 Q/QA14 + 2455 Q/QA21 - 5428

Cold season

F. coli = 774 DQ3/Q + 250 Q/QA14 + 189

The coliform count estimated from the cold season equation, 790 per 100 ml, was nearer the observed count than either of the other estimates. Although hydrologic condition f occurred during the cold season, the regression equation based on the warm season data was also used to estimate the coliform count as an indication of different seasonal relationships. The estimated count, 2200 per 100 ml, was interpreted to mean that for a given runoff condition higher coliform counts might be expected during the warm season, as estimated from the regression equation developed from the untransformed data.

These inaccuracies of the estimate provided by the several regression equations point out the difficulty of describing exactly the complex relationships between coliform counts and hydrologic and environmental conditions. A great many factors are involved. Some of these factors are related to hydrologic conditions such as runoff, streamflow, and suspended sediment concentration. Others are related to the length of time since the previous runoff event and the size of the previous event.

Further complicating this inability to describe these relationships exactly is the lack of precision and accuracy of the coliform test itself. River water samples are frequently highly turbid because of the suspended sediment load. Since bacteria tend to adsorb onto surfaces of the sediment, it is important to select a sample for analysis which is representative of the size distribution of the particles in the original sample. If the test sample has a large and unrepresentative number of fine particles, the total surface area of these sediment particles, and the number of bacteria indicated by the analysis may be greater than that of the original river water sample.

A recent study by Burnett (14) of the precision of the fecal coliform test results (membrane filter technique) for data collected in this study has indicated that for a given sample, the standard deviation was about 21.5 per cent of the mean. Standard Methods (53) lists no comparable test data for the membrane filter technique.

In summary, surface runoff is related directly to high coliform populations in the river. During periods of little runoff both streamflow and runoff are related directly to the coliform population.

Plankton Diatoms are the dominant plankton group in the Des Moines River (20). The next most abundant is the flagellate group. Together these two groups account for 90 per cent or more of the total plankton population observed in the Des Moines River. Although the relative proportion of these two forms differs from year to year, diatoms account for about 70 to 90 per cent, and flagellates account for about 5 to 30 per cent of the total number of plankton found in the river at any one time.

Plankton populations[1] are greatest generally in September or October, and lowest during the winter months of January through March. A large increase in the total plankton population occurs about mid-April when aquatic environmental conditions become more favorable.

Plankton growth is almost never nutrient limited in the Des Moines River. However, wide fluctuations in the total plankton population were observed. One explanation which has been provided by Kilkus (38) was based on a study of plankton populations of some central Iowa streams during months when the river was not ice covered. They concluded that physical factors assumed the dominant role in regulating the population of suspended algae, or plankton. Watershed area, streamflow, and temperature were identified as the important physical factors. Kilkus et al. suggested that algal material was being generated on the bottom areas of upland streams, and that little additional algal production occurred within the principal rivers. In this way, plankton populations measured at a point along the course of a stream were considered to be related to the amount of algal material lost from the stream system above.

Kilkus et al. identified an inverse relationship between plankton population and streamflow. They suggested that this was the result of

[1]Unless defined otherwise, plankton population should be interpreted as the total number of plankton per 0.01 ml.

dilution. If upstream production of algal material was considered to be constant on an areal basis, then increases in flow would have the effect of diluting the suspension of algae in the river. Increases in temperature, up to a physiological optimum would tend to increase the upstream growth rate, and consequently the total number of plankton moving downstream.

In a study of diatoms in the Des Moines River, Drum (20) suggested that light intensity, as regulated the suspended sediment concentration and ice and snow cover, may be a factor limiting algal growth. He noted that in effect, dry years on land are "light years" in the river and that wet years on land are "dark years" in the river. He also observed the inverse relationship between numbers of diatoms and streamflow, but suggested that following heavy rains, during which the higher streamflow had removed diatoms by scouring, renewed growth was prevented because of the maintenance of turbidity at levels which would restrict light penetration. In his study of winter diatom species, Drum found that the cold temperatures favored certain species. In general these species were attached forms which could develop long slender spines. The spines were absent during the warmer months. He attributed this cold season development to increased floatation potential and, because of the increased surface area, to improved nutrient absorption during the lower energy situation of the winter.

On the basis of these two studies it seemed certain that physical factors were of considerable importance in regulating plankton populations in the Des Moines River. What was not certain was that the relationship between plankton population and streamflow was the same throughout the year.

Since the study of Kilkus et al. (38) had excluded the months when the river was ice-covered, December through February, it was questioned whether the inverse relationship of plankton population with flow held true during the winter months. What made this question of even greater interest was an observation made by Drum. He noted that the diatom Gomphonema olivaceum, the most important attached diatom in the Des Moines River, developed massive colonies only when the water temperature was 10°C or less. It was abundant from late fall through early spring, but was nearly non-existent in warm weather collections. A second diatom, Stephanodiscus hantzschii, occurred abundantly throughout the year, but grew best when the water temperature was 5°C or less with the apparent concurrent development of slender siliceous spines.

From Drum's study it was evident that distinct differences existed between warm weather and cold weather algal forms. During the warm weather the dominant forms were planktonic, but during the cold weather the dominant forms were attached. It was felt that the effect of higher streamflow during the winter would not necessarily be dilution since the scouring action would tend to suspend the attached filamentous diatoms, possibly increasing the number of plankton in the river.

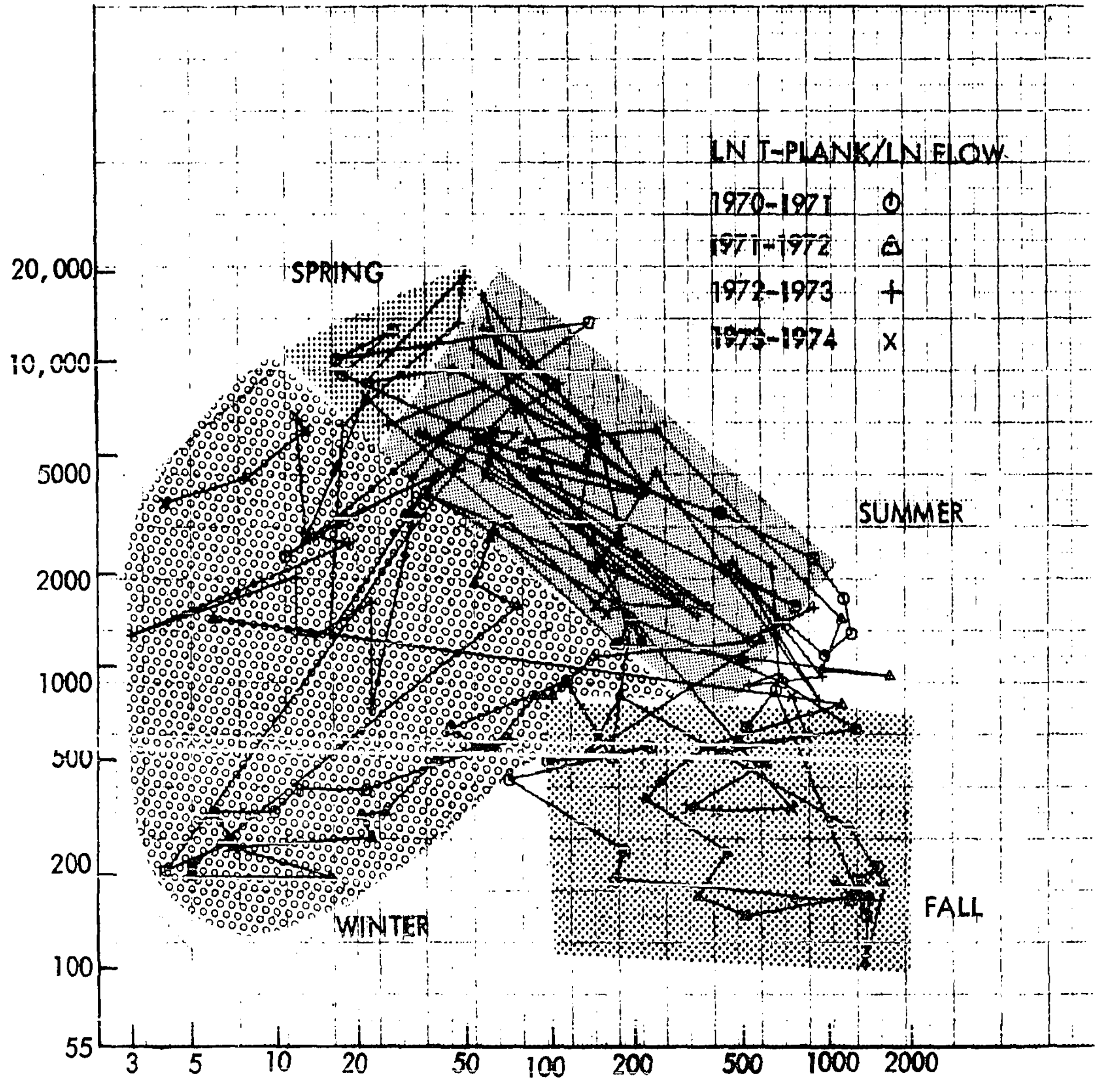
In order to investigate the hypothesis that the relationship between plankton population and streamflow was the same throughout the year, streamflow in cfs was plotted against the plankton population using a log-log plot. The graph, based on data collected from 1970 to 1974, is shown in Figure 14. Data collected on consecutive sampling dates are connected by a straight line an aid to the interpretation of week to week changes. The importance of the plot is not so much in the details, but in the slopes of these lines.

Several observations will be pointed out since the plot appears, initially, to consist only of a maze of lines.

In the upper right portion of the graph labelled summer, it will be noted that the connecting lines have a negative slope of about 40°. The points in this region represent data collected, for the most part, during the months of May or June through July. The relationship indicated is consistent with that suggested by Kilkus et al. (38), that is, plankton population is related inversely to streamflow. In the lower left portion of the graph labelled winter, the general trend of the slope of the connecting lines is positive. These data were collected, for the most part, during the months of December through February or March. This appears to indicate a direct relationship between the plankton population and flow. In the two other portions of the graph labelled spring and fall, plankton population appears to be independent of streamflow.

A final observation is that during the winter months there is considerable variation from year to year in the observed plankton population at a given streamflow. This may be associated with physical factors other than streamflow such as increased stream bottom area or less ice cover with the higher flows. Both of these factors would tend to increase the total plankton production.

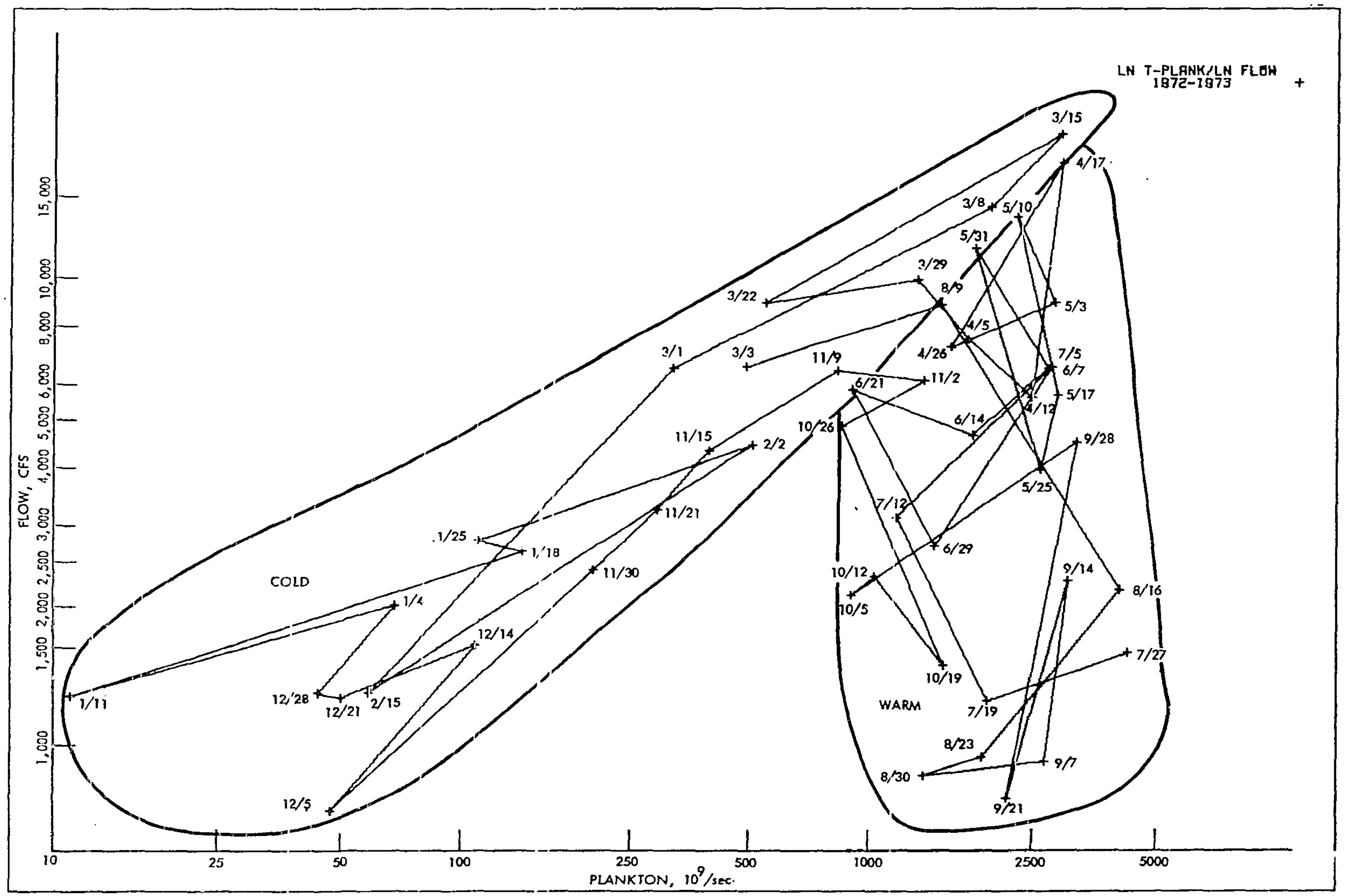
Figure 14. Relationship between plankton population and streamflow, 1970-1974


LN T-PLANK/LN FLOW
1970-1971 O
1971-1972 Δ
1972-1973 +
1973-1974 x
SPRING
SUMMER
FALL
WINTER
20,000
10,000
5000
2000
1000
500
200
100
55
3
5
10
20
50
100
200
500
1000
2000

In summary, during the winter plankton populations are related directly to streamflow for any given year. Higher flows scour the predominantly attached algae from the river bottom, resulting in a greater concentration of plankton in the river. In contrast during the summer, plankton populations are related inversely to streamflow. Higher flows dilute the predominantly suspended algae, reducing the observed concentration. During the spring and the fall, algal forms are changing from attached to suspended (or vice versa), and the length of daily light periods and the water temperature are changing. As water temperature and light take on increased importance, in regard to plankton population, the plankton-streamflow relationship is obfuscated. Based on these interpretations, the hypothesis that the relationship between plankton population and streamflow was the same throughout the year was rejected.

A conclusion by Kilkus et al. was that during their sampling period, algal production remained constant on an areal basis. This conclusion was tested by plotting the natural logarithm of the total number of plankton moving past the sampling location against the natural logarithm of the mean streamflow on the sampling date using data collected during the year 1972-1973. This graph is shown in Figure 15. It was felt that although variations might be observed from year to year, the fundamental relationship would be the same. In Figure 15 individual data points are identified by the month and day on which the sample was collected. As in Figure 14, consecutive sampling dates are connected by straight lines.

Figure 15. Relationship between the quantity of plankton in the river and streamflow, 1972-1973


LN T-PLANK/LN FLOW
1872-1873
FLOW, CFS
PLANKTON, 10^9/sec
COLD
WARM
1,000
1,500
2,000
2,500
3,000
4,000
5,000
6,000
8,000
10,000
15,000
10
25
50
100
250
500
1000
2500
5000

Two distinct relationships between the quantity of plankton in the river and streamflow are evident in Figure 15. These relationships were highly specific to the warm and the cold season, labelled as such on the plot. During the warm season the total number of plankton was relatively independent of streamflow, but during the cold season the total number of plankton was related strongly to streamflow. The division of the seasons was based on a water temperature of 10°C.

Within each season the variation in streamflow from minimum to maximum was virtually identical, about 750 cfs to 17,000 cfs. However, the variation in the total number of plankton was greatly different. The cold season minimum to maximum ratio was about 250, whereas the warm season ratio was only about five. Thus it would appear from the analysis of the plankton data as a function of streamflow that the total number of plankton in the river during the warm season is relatively constant. This supports the hypothesis of Kilkus et al. (38). As the streamflow increased, the concentration of the plankton decreased due to dilution, but the total number of plankton in the river remained relatively constant.

It is concluded on the basis of this analysis that what is commonly referred to as an "algal bloom" may be simply the result of a low streamflow condition. During low flow periods, algal production in the upland bottom areas is at the normal warm season rate, but less water is available for dilution.

In contrast, during the cold season the variation in the total number of plankton in the river is many times that for the warm season, and is almost certainly too large to be the result of chance. Increases in streamflow break loose the attached filamentous algal forms, primarily diatoms (20), from their base increasing both the total number in the river and the concentration.

In summary, variations in the plankton population, in terms of both concentration and total numbers, may be explained by changes in temperature and streamflow. Effects of temperature appear to be described best on the basis of two seasons: a cold season when the water temperature is less than 10°C and a warm season when the water temperature is greater than 10°C. Within each season streamflow is the dominant factor controlling the plankton population. During the cold season attached algal forms are scoured from their base and an increase in total numbers of plankton and concentration is observed as the streamflow increases. During the warm season, an increase in flow is associated with a decrease in concentration because of dilution, but the total number of plankton in the river remains relatively constant.

Because of the distinct seasonal difference in the response of the plankton population to environmental and hydrologic conditions, it was of particular interest how these differences would be manifested in the regression analysis. Application of the $\underline{t}$ test to the plankton data had indicated that the biannual distributions were significantly different from the annual distribution. The range of the total plankton data for the full year data set was 3 to about 2000 organisms/0.01 ml. As might be expected, the range of the data during the more rugged conditions of the cold season (3 to 300) was less than that recorded for the warm season (50 to 2000).

Regression equations were developed using both the natural log transformed and the untransformed data. The R^2 value for the warm season regression was somewhat higher for the untransformed data (54.9%) than for the transformed data (47.2%). However, it was felt that several very large plankton counts recorded during the warm season could have biased the regression based on the untransformed data. Thus the regression equation developed from the transformed data was selected as being representative of a greater variety of hydrologic conditions.

For the regression based on plankton data collected during the cold season the R^2 value obtained using the transformed data in the regression analysis (77.4%) was much greater than for that based on the untransformed data (49.5%). The regression equations based on the natural log transforms of the total plankton data are given below and in Appendix B, Table 41.

<u>Total plankton, number/0.01 ml, natural log</u>

Warm season

$$\ln \text{Plank} = -0.555\ \text{QSTD} - 0.482 \ln \text{SED} - 0.0509\ \text{TEMP} - 0.0327\ \text{SUNC} + 10.32$$

Cold season

$$\ln \text{Plank} = -1.03 \ln \text{QA7} + 2.60\ \text{SUN} + 1.18\ \text{AGF} - 0.0701\ \text{AGSC} + 10.35$$

Regression equations developed from other treatments of the total plankton data, as well as for diatoms and flagellates are listed in Tables 41, 42 and 43 in Appendix B.

Explanatory parameters contained in the regression equation for the warm season were flow, suspended sediment, temperature, and a season-flow interaction parameter. The relationship with flow, represented by the parameter QSTD, was inverse, as indicated by the negative regression coefficient. This supported the hypothesis that increasing streamflows diluted the plankton concentration. Suspended sediment was also related inversely with plankton population and appears to support the contention of Drum (20) that greater suspended sediment concentrations would decrease the light available to the plankton for photosynthetic activities, and hence reduce their growth rate. However, inadequate data were available to positively confirm this relationship.

Surprisingly, temperature was related inversely to plankton population. Although no good explanation could be found for this relationship, it was noted on examination of the plankton distribution data that a number of very high plankton counts were recorded during the month of October. The associated water temperature at this time of year, 15 to 18°C, was 10 to 15°C lower than the river temperatures recorded during the summer months when plankton populations were much lower. This relationship appeared to be indirect, rather than causal and would account for the inverse relationship indicated by the regression equation.

Interpretation of the season-flow interaction parameter SUNC is difficult. By reference to the definition of this term in Table 11, it will be noted that SUNC is a function of the season parameter SUN and reciprocal streamflow, 1/Q. Thus for a given date, the term SUNC times the regression coefficient, -0.0327, will have its largest negative value during periods of lowest streamflow. If values for the parameter SUN and mean monthly streamflow are used to calculate SUNC, an approximate range for the value of SUNC may be obtained. During the warm season this range is -0.08 to -0.62, but could be greater depending on the value of specific streamflows. It is interesting to note that the largest negative value is obtained for the month of September, the month for which greatest plankton populations and lowest streamflows are frequently observed. Thus, the term SUNC may provide a seasonal correction to the estimate contributed to the equation by the other flow parameter, QSTD.

While the regression equation developed from the warm season data supported the hypothesis developed for the relationship between plankton population and flow, that for the cold season appeared initially to contradict the hypothesis developed for the cold season relationship.

Streamflow, two seasonal parameters, and a season-flow interaction parameter were included in the regression equation developed from the cold season data. The negative regression coefficient for the streamflow parameter, ln QA7, was not anticipated since it indicated an inverse relationship between plankton population and streamflow, similar to that found for the warm season. River water temperature, considered to be of importance during the cold season, was not included. An explanation was sought for these two apparent inconsistencies.

The parameter AGSC, a function of reciprocal streamflow was included in the equations. The negative coefficient of this term, -0.0701, indicated the expected direct relationship between plankton population and flow. For a given date the term would have the largest negative value for the lowest streamflow. Expected values were calculated for the term using the mean monthly streamflow and the mid-month value of the term AGS. The range of expected values calculated for AGSC was -0.04 to -0.79. Largest negative values were obtained for January and February, generally the two coldest months of the year. The mean monthly streamflow for January was the lowest of the cold season months. From this analysis, the term AGSC appeared to be related inversely with temperature and streamflow. This relationship was supported by the correlation of SUNC and TEMP ($r = -0.341$ and of SUNC and ln Q ($r = -0.51$) listed in Appendix A.

The importance of the seasonal terms, SUN and AGF, was next investigated. During the cold season the range of values for the term SUN is 0.0 to about 1.4. That for AGF is 0.0 to about 1.7. Since the values of these two terms are fixed for a given date their sum can be considered as one term. Again using the sum of the mid-month values times the respective regression coefficients of the two seasonal terms, a range of 0.97 to 5.26 was obtained. The lowest value calculated was for January (0.97), typically a very cold month characterized by low streamflow and very low plankton populations. The highest value (5.26) was calculated for April, a month during which the transition between the cold and the warm seasons occurs. Streamflows, in April are generally very high and the plankton population, in terms of both total numbers and concentration, increases, sometimes dramatically. From this analysis it would appear that the sum of the two seasonal variables were related directly to streamflow and temperature.

Thus, analysis of the regression equation for the cold season plankton data would indicate that the relationship between plankton population and streamflow is more complex than a simple direct relationship. Although water temperature as such was not included in the regression equation it did appear to be an important factor which was related directly to plankton population. Streamflow is an apparently related to plankton populations although in different modes during the cold season. The scouring action of the higher streamflow does tend to increase both the total number and the concentration of plankton. However, dilution with increased flow is also indicated. The most probable explanation in this case is that once large numbers of the attached algae are broken loose and suspended, a period of time is required for the benthic forms to again establish a large population. Thus further increases in streamflow would tend to dilute the concentration of the suspended algal forms.

CONCLUSIONS AND RECOMMENDATIONS

The hypothesis on which this research was based was that variations in the limnological characteristics of the Des Moines River at Saylorville, Iowa were a function of several hydrologic factors and the normal climate conditions observed in the upper Des Moines River Basin. It was hoped that the results of this research would lead to a new method for relating the effects on the river of non-point source discharges such as runoff from agricultural lands. Probably the most important conclusion reached is that it was possible to develop parameters which represent different types of runoff events, and that these events could be related statistically to changes in water quality in the Des Moines River.

The hypothesis was developed from observations made while collecting river water samples over a period of nearly three years. Agricultural activities appeared to be the major factors associated with changes in water quality. Relating these activities mathematically to water quality was considered to be important. For example, evaluating the effects of point source discharges on the river is conveniently approached by collecting wastewater treatment plant effluent samples periodically from sources in the watershed. Analysis of these samples coupled with the associated flow volume provides a good estimate of contributions of nutrients, materials exerting a BOD, and other biological materials to the river. Based on this information the effects of the point source discharge on the river can be evaluated.

In contrast, evaluating the effects on the river of widespread agricultural activities in a highly agriculturally oriented area is more complex. Contributions of materials associated with agricultural activity occur principally during runoff periods. Evaluating the effects of these dispersed or non-point source discharges is more difficult because of the problems of obtaining representative samples in an extensive river basin system.

The approach used in this research was to evaluate the impact of non-point source discharges on the river based on statistical analysis of six years of water quality and streamflow data. Because surface runoff served as the major channel through which materials associated with seasonal agricultural activities entered the rivers and streams of basin, it was believed that elements of the streamflow dynamics and season would be associated strongly with water quality. The wealth of information available from the six-year study dictated statistical analysis as the tool of choice for the evaluation of the relationships between water quality and hydrologic and climatological conditions.

Because of the contrasting hydrologic and climatological conditions occurring throughout the year, the limnological data were divided into warm season and cold season groupings. A water temperature of 10°C was considered as the basis for the division of the data. Mean values of the biannual groupings of the data were compared, using the t test, to that for the full year data set to check for significant differences (P = 0.05) in their distributions. For those for which a significant difference was found, the data were analyzed on a biannual basis. Otherwise the full year or annual data set was considered for analysis.

A number of explanatory parameters were developed to represent streamflow dynamics, seasonal variations, and the interactions of these two factors. In essence the explanatory parameters were considered to be grouped into five components: streamflow, runoff, water temperature, season, and season-streamflow interaction.

Seventeen limnological substances were considered to represent the basic physical, chemical and biological elements of water quality.

Data for each of the limnological substances was regressed on the explanatory variables using a stepwise regression routine. Only those explanatory parameters which were statistically significant ($\alpha = 0.05$) were included in the final regression equation because fundamental relationships were sought.

The regression equations which were developed accounted for 0 to 85 per cent of the variance of the limnological parameters, as listed in Table 26. Although only a small percentage of the variance could be accounted for for some of the limnological parameters, the explanatory variables selected by the regression analysis were statistically significant in all cases and could be considered to be related to variations in water quality.

No distinct pattern emerged regarding the strength of the statistical relationship for conservative and non-conservative parameters, nor was any pattern evident indicating that the statistical relationships were better for the cold season than for the warm season. It had been suspected that the cold weather relationships might be stronger.

Table 26. Percentage variance accounted for by regression equations for 17 limnological substances

Parameter	R^2, %		
	Annual	Warm Season	Cold Season
1. Turbidity	51.8	54.7,L[a]	85.5
2. Chloride	54.2	70.3	69.7
3. Silica	19.7	52.0,L	54.5,L
4. Total hardness	45.2	43.4	67.5
5. Calcium	40.5	39.4	59.3,L
6. DO	44.2	45.2,L	31.5
7. BOD	32.5	46.7	56.4
8. COD	48.2	32.9	60.2
9. Ammonia	16.0	4.5,L	30.4
10. Organic nitrogen	29.2	37.4	32.7
11. Nitrate	61.7	76.4	82.6,L
12. Total phosphate	43.5	75.7	38.8,L
13. Orthophosphate	38.6	43.0	38.4
14. Fecal coliform	68.5	82.8	24.4
15. Total plankton	54.5,L	54.9	77.4,L
16. Diatoms	49.4,L	60.3	76.8,L
17. Flagellates	24.0,L	N.S.[b]	53.4

[a]L, regression equation based on the natural log transformation of the data.

[b]N.S., no parameters significant at $\alpha = 0.05$.

Biological activity is greater during the warm season and the possibility of degradation of some of the limnological substances prior to analysis would be greater at this time. For example, only four per cent of the variance in the ammonia concentration could be accounted for during the warm season. It was felt that at least part of this variance was due to change in concentration caused by biological activity during the interval between sample collection and sample analysis. Water samples to be analyzed for ammonia nitrogen are acidified at the time of sample collection to stop biological activity. If this fixing process is inadequate, ammonia may be metabolized and converted to other forms of nitrogen. Delay in analysis of the samples would provide time for the ammonia concentration to change. As the water temperature increased, the change in concentration would be more likely.

It was not surprising that all of the variance in the concentrations of the limnological parameters could not be accounted for. The objectives of the research were limited in this respect. The original intent of the research was to determine the extent of the relationship of the limnological substances with hydrologic and climatological conditions. It was hoped that the results of this evaluation would provide a new method of evaluating the effects of non-point source discharges as manifested in surface runoff and other factors and the statistical relationship with the limnological parameters. The results of this research are believed to show that this has been accomplished.

Nonetheless, it would be well to consider some of the factors which prevented the complete lack of fit of the regression equations. These factors fall, more or less, into four general categories: sampling and analysis, cultural details, hydrologic factors, and seasonal variations.

Factors associated with sampling and analysis of samples is probably the major source of error related to lack of fit. One of the problems associated with the use of data collected over a period of six years is the inconsistency of the numerous individuals involved in collecting and analyzing the samples. Slightly different sampling techniques have been used because of varying river conditions, weather, and the nature of the individual collecting the sample. Many individuals have been involved in the analysis of these samples. Several different analytical techniques for a limnological substance have been used as one method was replaced by another because of some desirable characteristic. Incorporated into the data because of these variations are random and indeterminate errors, systematic errors, errors resulting from personal bias, mistakes, and improper omission of data. These errors are difficult to identify and rectify after the data has once been entered into the data bank on computer cards. In a study such as this, outliers are fairly easy to identify, to check their validity, and to confirm their accuracy or reject them. However, the inclusion of an unrepresentative sample or minor errors in analysis are difficult, if not impossible to identify and correct or eliminate.

Cultural details such as variation in fertilizer application rates, timing of application and changing farm practices are other factors related to lack of fit. It is difficult to identify these variations in a large river basin. For the purpose of this research it was considered that on a basin-wide scale these variations averaged out, and were essentially constant for the six year period of the study.

Hydrologic factors such as the distribution of rainfall are important. For example, water quality would be expected to be considerably different depending on whether rain fell near the sampling site causing a given hydrologic condition as contrasted with the same hydrologic condition being caused by rainfall 100 miles upstream of the sampling site. All similar hydrologic conditions were effectively lumped together by the regression procedure and were considered to have a similar cause.

Seasonal variations include agricultural activities, meterological conditions, and other types of seasonally related factors. An attempt was made to account for these variations through the use of season parameters. However, these parameters were based on fixed dates which represented general trends. For example, maximum effects of spring agricultural activities and runoff conditions were associated with April 20, although these effects may have been manifested at greatly different times during the six-year period. The regression procedure occasionally included two or more of the season parameters, and effectively shifted the maximum to some other date. Thus it is believed that errors resulting from the seasonal variations were minimal.

An early concern was whether the explanatory parameters developed to represent runoff and season would delineate actual runoff situations and seasonal variations. The explanatory parameters streamflow, water temperature, and suspended sediment were straightforward. At the completion of the research it was concluded that the runoff and season parameters had indeed simulated actual situations. This conclusion was based primarily on the observation that parameters included in the regression equation, such as runoff, successfully described known runoff relationships with the limnological parameters. For example, it had been observed on a great many occasions that runoff was strongly related to changes in the fecal coliform count. The regression equation developed from the fecal coliform data strongly supported this relationship. Similar conclusions were reached for the other parameters.

Although the conclusions related to the individual parameters were discussed in their particular section, it should be mentioned that some unsuspected relationships were also indicated. For example, it was found that orthophosphate was not always related to runoff. Indirect relationships were also indicated. Evaluation of the regression results indicated that plankton populations may influence strongly some of the changes in the concentrations of limnological parameters such as nitrate, calcium, BOD, silica, orthophosphate and possibly others.

All of the explanatory parameters were included in one or more regression equations. Streamflow variables were included more frequently during the warm season when both very high and very low flows occurred. The runoff parameters were included in about equal frequency for the warm and the cold season regressions. However, the longer term runoff parameters such as Q/QA28 were included more frequently in the cold season equations. It was felt that this was due to the timing of the runoff events as related to the uniformly low streamflow during most of the winter.

Season and season-streamflow interaction parameters were included in most equations. It was felt that their usefulness was the indication of general trends in the concentration data. Superimposed on these general trends were the effects of the other parameters. In several equations only season parameters were used to describe the variations in the concentration of the limnological parameters.

Temperature was included in relatively few equations. In most cases, this relationship appeared to be associated to some extent with seasonal factors. Two factors resulted in the exclusion of temperature from the regression equation. Temperature is correlated with the parameter SUN ($r = 0.4$, biannual basis, and $r = 0.8$, annual basis). Thus the parameter SUN probably served as a kind of smoothed water temperature parameter. More importantly, the data were divided into seasons on the basis of water temperature, effectively removing most of the variation resulting from temperature differences.

At some point in any research effort it becomes desirable to tie up loose ends, and write a report of the research progress, its results and conclusions reached. No creative research project is ever completed. As the research progresses improvements and additions become apparent. In a short term project many questions are left unanswered. Occasionally, the researcher, at the end of his study, discovers that he has finally learned what important questions to ask. This research is regarded as the first major step in developing an improved method of evaluating the effects of surface runoff on stream water quality.

Spring runoff conditions introduce a large amount complexity into limnological relationships. Little runoff occurs during the winter months and surface materials are not greatly disturbed. With the onset of spring many types of events and activities occur which influence water quality such as snowmelt, heavy rainfall, and agriculturally related field work. A possible improvement in handling the complexities of this period would be to isolate the spring runoff period and consider it as a separate unit or season. An adverse effect of this treatment is that the relatively few data would reduce the statistical confidence. A second approach in dealing with runoff events and one which would partially remedy the problems associated with few data points is to consider all heavy runoff events as a group.

Parameters which provide a different approach to runoff should be tested. One approach would be based on the familiar antecedent precipitation index, but the parameter would be a function of streamflow rather than precipitation. The term would be based on the following relationship:

$$Q_s = b_1Q_1 + b_2Q_2 + \dots b_nQ_n$$

where

Q_s = antecedent flow index (AFI) for the sampling date

Q_n = streamflow on the nth day prior to the sampling date

b_n = coefficients to be determined which are based on the limnological effects of the streamflow up to n days prior to the sampling date and whose sum would equal one

That the limnological characteristics of a stream vary with flow has been fairly well established. In this research the relationship of the concentration of a limnological parameter, c, with flow, Q, has been described as

$$c = aQ^b$$

in which a and b are regression parameters. The parameter b is considered to be constant at all streamflows. It is expected that the parameter b should be a variable which is a function of streamflow such that

$$b = cQ^d$$

and c and d are regression coefficients. This approach has been described in greater detail by Ledbetter and Gloyna (41).

A final recommendation, and indeed an important one is that a thorough error analysis should be made of the data. At the outset of this research it was considered that the data, in essence, were correct. As the research progressed there were a number of clues which indicated this was not true. Outliers were evaluated for accuracy. Some were eliminated while others were corrected, but there was no attempt to attempt to determine the precision of the replicate results for individual data points. For example, there is doubt as to the validity of the ammonia data during the warm season because of the very low percentage variance accounted for by the regression equation.

The research is believed to have merit regarding the evaluation of the effects of non-point source runoff on river water quality during a variety of weather conditions. However, further work needs to be done to establish this validity. A similar analysis should be applied to other rivers and sampling locations in the basin in order to develop the method's potential for evaluating the effects of non-point source discharge on water quality in an agricultural state.

LITERATURE CITED

1. Armstrong, D. E. and G. A. Rolich. "Effects of agricultural pollution on entroplication." In Agricultural Practices and Water Quality, T. L. Willrich and G. E. Smith, Eds. Ames, Iowa: Iowa State University Press, 1970, pp. 314-330.

2. Barrows, H. L. and V. J. Kilmer. "Plant nutrient losses from soils by water erosion." Advances in Agronomy, 15 (1963), 303.

3. Baumann, E. R. Preimpoundment Water Quality Study, Saylorville Reservoir, Des Moines River, Iowa-Annual Report. ERI-558. Engineering Research Institute, Iowa State University, Ames, Iowa, 1969.

4. Baumann, E. R. and J. DeBoer. Preimpoundment Water Quality Study, Saylorville Reservoir, Des Moines River, Iowa-Annual Report. ISU-ERI-AMES-72262. Engineering Research Institute, Iowa State University, Ames, Iowa, 1972.

5. Baumann, E. R., J. DeBoer, and C. S. Oulman. Five-Year Study of Des Moines River Water Quality. Unpublished Report, Department of Civil Engineering, Iowa State University, Ames, Iowa, 1972.

6. Baumann, E. R. and M. D. Dougal. Preimpoundment Water Quality Study, Saylorville Reservoir, Des Moines River, Iowa-Annual Report. ERI-346. Engineering Research Institute, Iowa State University, Ames, Iowa, 1968.

7. Baumann, E. R. and S. Kelman. "Effects of surface runoff on the feasibility of municipal advanced waste treatment." In Agricultural Practices and Water Quality, T. L. Willrich and G. E. Smith, Eds. Ames, Iowa: Iowa State University Press, 1970, Chapter 25.

8. Baumann, E. R. and S. Kelman. Preimpoundment Water Quality Study, Saylorville Reservoir, Des Moines River, Iowa-Annual Report. ERI-80700. Engineering Research Institute, Iowa State University, Ames, Iowa, 1970.

9. Baumann, E. R. and S. Kelman. Preimpoundment Water Quality Study, Saylorville Reservoir, Des Moines River, Iowa-Annual Report. ISU-ERI-AMES-71076. Engineering Research Institute, Iowa State University, Ames, Iowa, 1971.

10. Baumann, E. R., L. M. Naylor, and D. A. Willis. Des Moines-Raccoon River Basin Characteristics Related to Red Rock and Saylorville Reservoirs, Des Moines River, Iowa-Special Report. ISU-ERI-AMES-74261. Engineering Research Institute, Iowa State University, Ames, Iowa, 1974.

11. Baumann, E. R., C. S. Oulman, and L. M. Naylor. Water Quality Studies-Red Rock and Saylorville Reservoirs, Des Moines River, Iowa-Annual Report. ISU-ERI-AMES-73280. Engineering Research Institute, Iowa State University, Ames, Iowa, 1973.

12. Beer, C. E. Preliminary Report on Nutrient Load in Three Small Iowa Streams. Department of Agricultural Engineering, Iowa State University, Ames, Iowa, 1972.

13. Blommers, P. and E. F. Lindquist. Elementary Statistical Methods in Psychology and Education. Boston: Houghton Mifflin Co., 1960, pp. 346-348.

14. Burnett, John. Unpublished research data. Department of Civil Engineering, Iowa State University, Ames, Iowa, 1974.

15. Busch, A. W. "Use and abuse of natural water systems." Journal Water Pollution Control Federation, 43 (1971), 1480-1483.

16. Clark, J. W., W. Viessman, Jr., and M. J. Hammer. Water Supply and Pollution Control, 2nd ed. Scranton: International Textbook Co., 1971, p. 265.

17. Clark, R. N. and B. A. Stewart. "Amounts, composition, and management of feedlot runoff." Symposium on Animal Waste Management, U.S. Department of Agriculture Southwestern Great Plains Research Center, Bushland, Texas, 1973.

18. Cleary, E. S. "Evolution of social attitudes and action on water pollution control." Journal Water Pollution Control Federation, 44 (1972), 1301-1306.

19. Draper, N. R. and H. Smith. Applied Regression Analysis. New York: John Wiley & Sons, 1966, pp. 169-171.

20. Drum, R. W. "Ecology of diatoms in the Des Moines River Basin." Ph.D. dissertation, Iowa State University, Ames, Iowa, 1964.

21. Ecologic Successions and Their Application to Water Resource Planning and Management - First Progress Report. Walnut Creek, California: Water Resources Engineers, Inc., 1970.

22. Efroymson, M. A. Mathematical Methods for Digital Computer. New York: John Wiley & Sons, 1960, Chapter 17.

23. Fact Sheet - National Pollutant Discharge Elimination System - Agricultural Permits. Iowa Department of Environmental Quality, Des Moines, Iowa, 1973.

24. Fiering, M. B. "Multivariate techniques for synthetic hydrology." Journal of the Hydraulics Division, ASCE, 90 (1964), 43-60.

25. Fitzgerald, G. P. "The effect of algae on BOD measurements." Journal Water Pollution Control Federation, 36 (1964), 1524-1542.

26. Gold berg, M. C. "Sources of nitrogen in water supplies." In Agricultural Practices and Water Quality. T. L. Willrich and G. E. Smith, Eds., Ames, Iowa: Iowa State University Press, 1970, Chapter 7.

27. Goodman, A. S. and R. J. Tucker. "Time varying mathematical model for water quality." Water Research, 5 (1971), 227.

28. Graber, R. Agricultural Animals and the Environment. Feedlot Waste Management Regional Extension Project, Oklahoma State University, Stillwater, Oklahoma, 1974, pp. 31-40.

29. Grant, E. I. and R. S. Leavenworth. Statistical Quality Control. New York: McGraw-Hill, Inc., 1972, pp. 103-106.

30. Grantham, G. R., J. C. Schooke, Jr., and E. E. Pyatt. "Water quality simulation model." Journal of the Sanitary Engineering Division, ASCE, 97 (1971), 569-586.

31. Harms, L. L., J. N. Dornbush, and J. R. Anderson. "Physical and chemical quality of agricultural land runoff." Journal Water Pollution Control Federation, 46 (1974), 2460-2470.

32. Holt, R. F. "Surface water quality is influenced by agricultural practices." Transactions American Society of Agricultural Engineers, 16 (1973), 565.

33. Holt, R. F., H. P. Johnson, and L. L. McDowell. "Surface water quality." Proceedings of the National Conservation Tillage Conference, Des Moines, Iowa, March 28-30, 1973, pp. 141-156.

34. "Iowa Annual Farm Census - 1972." Iowa Department of Agriculture Bulletin No. 92-AH, 1972.

35. Iowa Annual Farm Census - 1973 (Preliminary). Iowa Crop and Livestock Reporting Service, Iowa Department of Agriculture, Des Moines, Iowa, 1973.

36. Iowa Livestock and Poultry - County Estimates 1971-1973. Iowa Crop and Livestock Reporting Service, Iowa Department of Agriculture, Des Moines, Iowa, 1974.

37. Jones, J. R. "A limnological survey of the upper Skunk River, Iowa." M.S. thesis, Iowa State University, Ames, Iowa, 1972.

38. Kilkus, S. P., J. D. LaPerriere, and R. W. Bachmann. "Nutrients and algae in some central Iowa streams." Journal Iowa Agricultural and Home Economics Experiment Station, Paper No. J-7795, 1974.

39. Kuentzel, L. E. "Bacteria, carbon dioxide and algae blooms." Journal Water Pollution Control Federation, 41 (1969), 1737-1749.

40. Kunkle, S. H. "Effects of road salt on a Vermont stream." Journal American Water Works Association, 64 (1972), 290-299.

41. Ledbetter, J. O. and E. F. Gloyna. "Predictive techniques for water quality inorganics." Journal of the Sanitary Engineering Division, ASCE, 90 (1964), 127-151.

42. Mahlock, W. L. "Multivariate techniques for water quality analysis." Journal of the Environmental Engineering Division, ASCE, 100 (1974), 1119-1132.

43. Martin, W. P., W. E. Fenster, and D. L. Hanson. "Fertilizer management for pollution control." In Agricultural Practices and Water Quality, T. L. Willrich and G. E. Smith, Eds. Ames, Iowa: Iowa State University Press, 1970, pp. 142-158.

44. Massey, H. F. and M. L. Jacobsen. "Selective erosion of soil fertility constituents." Proceedings of the Soil Society of America, 16 (1952), 343.

45. Miller, S. S. "Water pollution law - 1972 style." Environmental Science and Technology, 6 (1972).

46. Nour, A. A. "A statistical methodology for predicting the pollutants in a river." Water Resources Bulletin, 8 (1972), 15-23.

47. Olson, R. A. "Influences of agricultural practices on water quality in Nebraska: a survey of stream, groundwater, and precipitation." Water Resources Bulletin, 9 (1973), 301.

48. Pingry, D. E. and A. B. Whinston. "Multigoal water quality planning model." Journal of the Environmental Engineering Division, ASCE, 99 (1973), 909-922.

49. Sawyer, C. N. "Fertilization of lakes by agricultural and urban drainage." Journal New England Water Works Association, 61 (1947), 109.

50. Schuman, J. E., et al. "Nitrogen losses in surface runoff from agricultural watersheds on Missouri Valley loess." Journal of Environmental Quality, 2 (1972), 299.

51. Schwab, G. O. et al. "Quality of drainage water from heavy textured soils." American Society of Agricultural Engineers, Paper No. 72-733, 1972.

52. Shannon, E. E. and P. L. Brezonik. "Entrophication analysis: A multivariate approach." Journal Sanitary Engineering Division, ASCE, 98 (1972).

53. Standard Methods for Examination of Water and Wastewater. 13th ed. New York: American Public Health Association, 1971.

54. Steel, R. G. D. and J. H. Torrie. Principles and Procedures of Statistics. New York: McGraw-Hill, Inc., 1960, p. 187.

55. Streeter, H. W. and E. B. Phelps. "A study of the pollution and natural purification of the Ohio River." U.S. Public Health Bulletin No. 146, 1925.

56. Swanson, C. D. "Algae and oxygen of some central Iowa streams." M.S. thesis, Iowa State University, Ames, Iowa, 1973.

57. Taiganides, E. P. and T. E. Hazen. Properties of Farm Animal Excreta. American Society of Agricultural Engineers, Paper No. 64-315, June 1964.

58. Taylor, A. W., et al. "Phosphate movement from an agricultural waterbed during two rainfall periods." Journal of Agricultural and Food Chemistry, 20 (1972), 900-905.

59. Thomann, R. V. "Time series analysis of water quality data." Journal of the Sanitary Engineering Division, ASCE, 93 (1967), 1-23.

60. Tirabassi, M. A. "A statistical methodlology for predicting the pollutants in a river." Water Resources Bulletin, 8 (1972), 15-23.

61. Upper Des Moines River Basin - Waste Load Allocation Study. Iowa Department of Environmental Quality, 1972.

62. Viets, F. G. "Fertilizers." Journal of Soil and Water Conservation, 26 (1971), 51.

63. Wallace, D. A. and R. R. Dague. "Modeling of land runoff effects on dissolved oxygen." Journal Water Pollution Control Federation, 45 (1973), 1795-1809.

64. Ward, R. C., S. R. Nicholas, and G. U. Skogerboe. "Surveillance in water quality management." Journal Water Pollution Control Federation, 45 (1973), 2081-2087.

65. Water Resources Data for Iowa - Part 1. Surface Water Records. U.S. Department of the Interior, Geological Survey.

66. Watson, J. D. "Amendment of the Rivers Pollution Prevention Act, 1876." Surveyor, 75 (1929), 509-510.

67. Willrich, T. L. Properties of Tile Drainainge Water. American Society of Agricultural Engineers, Paper No. 70-752, 1970.

ACKNOWLEDGMENTS

A number of individuals have been instrumental to the completion of this dissertation and my graduate program. Recognition is especially due Dr. T. Al Austin - good friend, counselor, and critic. Dr. Chuck Oulman who has served as my advisor provided many creative ideas during the development of the research. Dr. E. Robert Baumann, Acting Director of the Engineering Research Institute and Head of the Sanitary Engineering section has aided my graduate progress in many ways. Thanks is also due George Johnson of the Corps of Engineers.

The cooperation of the personnel of the Analytical Services Laboratory at Iowa State University is gratefully acknowledged.

This research was supported by the Engineering Research Institute at Iowa State University, Ames, Iowa in part through funds made available under contract numbers D.A.C.W. 25-75-C-0007 and D.A.C.W. 25-75-C-0008 from the United States Army Corps of Engineers, Rock Island, District.

APPENDIX A

Intercorrelations of Explanatory Parameters

Intercorrelations of explanatory parameters:

Warm season - lower triangular matrix, N = 168
Cold season - upper triangular matrix, N = 136

	lnSED	TEMP	SUN	AGS	AGF	SUNC	AGSC	AGFC	M/Q
lnSED		.35	.46	.24	-.14	-.35	-.47	-.36	-.79
TEMP	.02		.40	-.12	.31	-.05	-.34	-.07	-.29
SUN	.48	.45		.80	-.64	.25	-.02	-.35	-.24
AGS	.45	-.16	.73		-.97	.27	.25	-.43	-.08
AGF	-.37	.38	-.51	-.96		-.26	-.33	.41	.01
SUNC	-.71	.06	-.24	-.29	.27		.74	.12	.51
AGSC	-.28	-.08	.13	.29	-.31	.66		.29	.84
AGFC	-.74	-.01	-.59	-.67	.61	.71	-.01		.76
M/Q	-.79	-.09	-.53	-.49	.41	.89	.40	.91	
lnQ+4	.76	-.08	.47	.55	-.51	-.81	-.38	-.77	-.85
lnQ+2	.79	-.06	.49	.55	-.49	-.84	-.41	-.77	-.87
lnQ+1	.80	-.04	.51	.55	-.49	-.84	-.42	-.78	-.88
lnQ	.80	-.03	.52	.55	-.49	-.85	-.41	-.79	-.89
lnQ-1	.75	-.03	.51	.56	-.50	-.84	-.41	-.79	-.89
lnQ-3	.68	.03	.54	.57	-.50	-.81	-.39	-.78	-.87
lnQ-5	.65	.08	.56	.57	-.49	-.78	-.38	-.75	-.84
lnQA7	.69	.04	.55	.57	-.50	-.81	-.39	-.78	-.87
QSTD	.51	.14	.33	.41	-.38	-.54	-.28	-.47	-.53
DQ3/Q	.25	-.07	.02	.01	.00	-.07	-0.8	-.01	-.04
Q/QA2	.36	-.02	.03	-.04	.06	-.13	-.08	-.09	-.11
Q/QA4	.43	-.11	.00	-.03	.04	-.17	-.09	-.12	-.14
Q/QA7	.45	-.19	-.05	-.03	.02	-.20	-.11	-.14	-.16
Q/QA14	.47	-.26	-.09	-.01	-.03	-.24	-.13	-.17	-.19
Q/QA21	.46	-.28	-.11	.01	-.05	-.26	-.13	-.18	-.20
Q/QA28	.47	-.29	-.11	.02	-.07	-.28	-.13	-.20	-.22

	$\log_e$								
	Q+4	Q+2	Q+1	Q	Q-1	Q-3	Q-5	QA7	QSTD
lnSED	.79	.80	.80	.80	.79	.75	.72	.75	.64
TEMP	.37	.39	.39	.39	.40	.40	.40	.40	.36
SUN	.30	.48	.48	.48	.47	.43	.40	.43	.57
AGS	.49	.27	.27	.27	.25	.21	.18	.21	.41
AGF	-.20	-.17	-.16	-.16	-.15	-.11	-.09	-.11	-.30
SUNC	-.48	-.50	-.51	-.51	-.51	-.52	-.52	-.53	-.27
AGSC	-.61	-.64	-.65	-.65	-.65	-.65	-.65	-.65	-.31
AGFC	-.60	-.60	-.60	-.61	-.60	-.53	-.57	-.58	-.31
M/Q	-.74	-.76	-.77	-.77	-.77	-.75	-.75	-.76	-.37
lnQ+4		.98	.97	.97	.96	.94	.91	.94	.77
lnQ+2	.98		.996	.99	.98	.96	.94	.96	.79
lnQ+1	.97	.997		.997	.99	.97	.95	.97	.80
lnQ	.96	.99	.99		.996	.98	.96	.98	.80
lnQ-1	.95	.97	.98	.99		.98	.97	.99	.80
lnQ-3	.91	.94	.94	.96	.98		.99	.998	.76
lnQ-5	.88	.90	.91	.93	.95	.98		.995	.73
lnQA7	.91	.94	.94	.96	.98	.995	.99		.76
QSTD	.79	.81	.82	.82	.81	.76	.73	.76	
DQ3/Q	.18	.19	.18	.10	-.02	-.10	-.13	-.10	.10
Q/QA2	.16	.18	.18	.15	.01	-.09	-.10	-.07	.13
Q/QA4	.22	.24	.24	.21	.09	-.06	-.09	-.05	.21
Q/QA7	.26	.27	.27	.24	.13	-.04	-.11	-.04	.25
Q/QA14	.30	.32	.32	.29	.21	.05	-.04	.04	.35
Q/QA21	.32	.33	.32	.30	.23	.10	.00	.08	.34
Q/QA28	.34	.35	.34	.32	.26	.14	.04	.12	.36

	DQ3/Q	Q/QA2	Q/QA4	Q/QA7	Q/QA14	Q/QA21	Q/QA28
lnSED	.09	.27	.30	.32	.36	.41	.44
TEMP	-.01	-.04	-.04	-.05	-.06	-.04	-.01
SUN	.11	.23	.25	.28	.31	.34	.36
AGS	.13	.24	.26	.28	.31	.33	.34
AGF	-.12	-.22	-.23	-.26	-.28	-.29	-.29
SUNC	-.03	-.05	-.03	-.03	-.04	-.06	-.07
AGSC	-.05	-.10	-.07	-.07	-.10	-.12	-.14
AGFC	-.07	-.20	-.18	-.17	-.18	-.20	-.21
M/Q	-.08	-.18	-.15	-.14	-.17	-.20	-.22
lnQ+4	.18	.26	.26	.27	.32	.36	.40
lnQ+2	.19	.26	.27	.28	.32	.37	.41
lnQ+1	.16	.25	.24	.26	.30	.35	.39
lnQ	.10	.13	.23	.24	.29	.34	.38
lnQ-1	.02	.15	.15	.18	.24	.29	.34
lnQ-3	-.03	.03	.01	.04	.11	.18	.23
lnQ-5	-.04	.01	-.03	-.03	.04	.11	.17
lnQA7	-.02	.06	.04	.05	.11	.18	.23
QSTD	.08	.20	.23	.27	.33	.38	.43
DQ3/Q		.68	.59	.49	.42	.38	.35
Q/QA2	.76		.96	.87	.77	.73	.69
Q/QA4	.72	.95		.96	.88	.82	.77
Q/QA7	.67	.86	.96		.95	.89	.84
Q/QA14	.53	.64	.78	.90		.98	.95
Q/QA21	.43	.54	.68	.80	.96		.99
Q/QA28	.38	.48	.62	.74	.91	.98	

APPENDIX B

List of Regression Equations for Water Quality Parameters

Table 27. Regression equations -- turbidity

Treatment[a]	Regression equation[b]	N	R^2	F	SE
A	53.4 lnQ-2 - 93.8 lnQ-3 + 36.1 lnQ-5 + 7.49 QSTD + 13.2 Q/QA14 + 12.8 ln SED + 0.867 TEMP - 27.27	304	51.8	45.5	28.7
A,L	0.212 QSTD + 0.489 lnSED + 0.0388 TEMP + 0.346 AGF - 0.3258	304	41.4	53.8	1.20
A, <176	6.27 QSTD - 21.7 Q/QA21 + 26.5 Q/QA28 + 6.52 lnSED + 0.947 TEMP - 19.33	300	67.0	119.3	15.9
A,B	0.736 WQB + 2.22 QSTD - 18.4 DQ3/Q + 3.81 Q/QA28 + 5.91 lnSED - 21.17	304	76.0	157.0	20.3
W	23.9 Q/QA14 + 19.5 lnSED - 76.54	169	35.3	45.4	37.4
W,L	-0.171 lnQ+4 + 0.212 lnQ-5 + 0.347 Q/QA14 + 0.313 lnSED + 1.421	169	54.7	49.4	0.386
W, <176	4.97 QSTD + 11.3 lnSED - 19.12	165	35.9	45.2	19.9
W, <89	-12.2 DQ3/Q + 7.34 Q/QA28 + 13.1 lnSED - 33.11	156	58.0	69.9	12.0
C	18.9 lnQ+1 - 11.2 lnQ-5 + 5.09 QSTD - 26.0 Q/QA14 + 21.5 Q/QA28 + 3.20 lnSED + 0.454 SUNC + 0.204 AGFC - 51.74	135	85.5	93.2	10.5

[a]An explanation of the treatment description and other abbreviations is provided in Table 44 at the end of Appendix B.

[b]Only the right side of the regression equation containing the explanatory parameters will be shown in tables listed in Appendix B. The form of the dependent parameter should be interpreted from the treatment description.

Table 27 (Continued)

Treatment[a]	Regression equation[b]	N	R^2	F	SE
C,L	1.11 lnQ + 0.0973 TEMP + 0.0454 SUNC + 0.0303 AGFC - 6.277	135	38.2	20.1	1.62
C, <176	equation identical with treatment C	135	85.5	93.2	10.5
C, <89	9.74 lnQ+2 - 11.5 lnQ-5 + 7.72 QSTD - 28.3 Q/QA7 + 14.9 Q/QA28 + 4.15 lnSED + 21.6 SUN - 9.17 AGS + 22.03	131	82.4	71.3	9.45

Table 28. Regression equations - chloride

Treatment	Regression equation	N	R^2	F	SE
A	-0.153 QSTD + 0.0557 M/Q + 22.66	247	54.2	144.0	8.37
A,L	-0.252 lnQ-1 + 5.076	247	44.4	195.6	0.347
A, S_c	-4.21 lnQ-2 - 4.00 lnQA7 + 1.38 QSTD + 86.53	247	59.2	117.3	6.91
A,B	0.787 WQB - 0.844 lnQ-2 + 11.62	247	85.2	702.9	4.75
W	-1.79 QSTD + 3.40 SUN - 0.342 SUNC + 0.0924 M/Q + 18.07	134	70.3	76.2	6.37
W,L	-0.101 QSTD + 0.00150 M/Q + 3.163	134	62.1	107.3	0.255
W,E,S_c	-3.36 lnQ-5 + 4.12 SUN + 0.290 SUNC + 0.0714 M/Q + 41.64	140	80.4	138.2	4.60
W,E,S_m	15.8 lnQ+4 - 16.1 lnQ+2 + 4.31 lnQ-5 - 3.24 QSTD + 0.0762 M/Q - 4.913	140	74.6	78.6	6.86
C	-6.80 lnQ+1 + 76.51	108	63.4	183.8	6.29
C,L	-0.280 lnQ+1 + 5.289	108	69.7	244.0	0.225
C,E,S_c	-6.33 lnQ+1 + 73.26	118	70.0	270.8	5.14
C,E,S_c	-5.10 lnQ+1 - 0.400 SUNC + 0.0246 M/Q + 62.48	118	74.0	108.4	4.88
C,E,S_m	36.9 lnQ+4 - 42.0 lnQ+2 + 0.0286 M/Q + 62.61	118	75.9	119.5	8.96

Table 29. Regression equations -- silica

Treatment	Regression equation	N	R^2	F	SE
A	1.93 lnQ-5 - 1.27 lnSED - 0.147 SUNC + 7.637	285	19.7	23.0	7.36
A,B	0.944 WQB - 0.350 lnSED + 2.015	284	90.6	1359	2.51
W	11.9 lnQA7 - 2.43 QSTD - 8.20 Q/QA4 + 13.0 Q/QA7 + 5.06 AGF 0.153 SUNC + 0.0308 M/Q - 91.22	161	51.2	22.9	5.45
W,L	1.14 lnQ-1 - 0.221 QSTD + 0.512 AGF + 0.0533 SUNC - 0.0622 AGSC -7.092	161	52.0	33.6	1.00
C	-11.5 lnQ + 9.68 lnQ-1 - 3.10 AGF - 0.373 SUNC - 0.115 AGFC + 33.84	122	36.0	13.1	6.08
C,L	-0.0664 SUNC + 0.0148 AGSC - 0.0221 AGFC + 2.878	122	54.5	48.0	0.728

Table 30. Regression equations – total hardness

Treatment	Regression equation	N	R^2	F	SE
A	-11.4 QSTD - 20.7 Q/QA28 - 16.3 lnSED - 2.12 TEMP - 2.97 SUNC				
	+ 0.731 AGSC + 518.2	278	45.2	37.2	69.5
A,B	0.945 WQB - 3.23 lnSED + 28.72	276	93.6	2008	23.5
W	26.6 lnQ+4 + 85.6 lnQ-1 - 45.6 QSTD - 19.5 AGS + 0.456 M/Q - 497.8	155	43.4	22.9	40.7
W,L	0.0886 lnQ+4 + 0.281 lnQ-1 - 0.149 QSTD - 0.0616 AGS				
	+ 0.00147 M/Q + 3.054	155	42.5	22.0	0.135
C	68.2 Q/QA7 - 53.4 Q/QA28 - 13.8 lnSED - 125. SUN - 2.36 SUNC				
	-2.24 AGFC + 0.174 M/Q + 517.2	123	67.5	34.1	64.6
C,L	0.229 Q/QA7 - 0.182 Q/QA28 - 0.454 SUN - 0.00563 AGFC				
	+ 0.00034 M/Q + 6.121	123	65.5	44.5	0.191

Table 31. Regression equations -- calcium

Treatment	Regression equation	N	R^2	F	SE
A	-7.00 QSTD - 19.2 lnSED - 2.07 TEMP - 1.89 SUNC - 0.448 AGFC + 388.7	282	40.5	37.6	60.8
A,B	0.915 WQB - 0.310 TEMP + 19.92	279	89.5	1171	25.5
W	96.1 lnQ+2 - 228 lnQ+1 + 226 lnQ - 29.9 QSTD - 42.9 SUN + 1.76 SUNC + 0.771 AGFC - 436.4	160	39.4	14.1	47.1
W,L	0.141 lnQ-1 - 0.151 SUN + 4.432	160	26.3	28.0	0.264
C	-24.6 lnSED - 96.7 SUN - 1.28 AGFC + 439.1	122	50.2	39.6	57.9
C,L	-0.0839 lnSED + 0.187 Q/QA7 - 0.119 Q/QA28 - 0.304 SUN -0.0131 SUNC + 0.00349 AGSC - 0.00452 AGFC + 6.084	122	59.3	23.7	0.229

Table 32. Regression equations - dissolved oxygen

Treatment	Regression equation	N	R^2	F	SE
A	-2.04 Q/QA2 - 1.91 SUN - 0.0452 SUNC + 0.0547 AGSC + 14.91	308	44.2	60.0	2.67
A, S_c	-0.443 lnQ+4 + 111. SUN - 226. AGS - 197. AGF + 326.6	308	41.3	53.2	2.45
A,B	0.624 WQB + 0.0979 lnSED - 2.45 SUN - 1.22 AGF + 6.321	307	65.0	140.3	2.11
W	-0.352 QSTD + 2.69 DQ3/Q - 3.26 Q/QA2 -0.879 lnSED - 3.00 SUN -0.0561 AGFC + 23.83	171	37.4	16.3	2.18
W,L	-0.0509 QSTD + 0.193 DQ3/Q - 0.275 Q/QA2 - 0.0969 lnSED -0.421 SUN + 0.174 AGS - 0.00088 M/Q + 3.665	171	45.2	19.2	0.200
W,E,S_c	-0.333 QSTD + 2.98 DQ3/Q - 5.28 Q/QA2 + 3.34 Q/QA4 - 0.947 Q/QA28 - 3.91 SUN - 0.867 AGF - 0.0345 M/Q + 20.71	172	49.3	19.8	1.63
C	0.566 lnQ-5 + 1.67 AGF - 0.0486 AGFC + 0.0104 M/Q + 6.184	137	38.9	21.0	2.81
C,L	0.0463 lnSED + 0.135 AGF - 0.00290 AGFC + 0.00059 M/Q + 2.129	137	28.2	12.9	0.219
C,E, S_c	0.478 lnSED + 1.65 AGF - 0.0486 AGFC + 0.00952 M/Q + 8.092	142	47.5	31.0	2.27

Table 33. Regression equations -- biochemical oxygen demand

Treatment	Regression equation	N	R^2	F	SE
A	-1.51 lnQ+4 + 1.81 Q/QA28 + 0.158 TEMP + 0.0622 SUNC				
	+ 0.0259 AGSC + 14.20	304	32.5	28.7	4.77
A,S_c	-1.20 lnQ+4 - 2.04 lnQA7 - 3.91 Q/QA14 + 3.43 Q/QA28 + 0.799 lnSED				
	+ 3.19 SUN + 24.83	304	46.1	42.3	3.61
A,B	0.888 WQB - 0.211 lnQ+4 + 1.33 Q/QA2 - 0.607 AGF + 1.969	304	79.6	291.5	2.62
W	-3.62 lnQ-1 - 0.236 AGSC - 0.407 AGFC + 0.0670 M/Q + 37.88	168	46.7	35.8	4.08
W,L	-0.433 lnQ-1 - 0.00949 AGFC + 5.500	168	44.0	64.9	0.427
W,E,S_c	-3.06 lnQ-1 - 0.117 TEMP - 0.135 AGFC - 0.0141 M/Q + 35.42	172	56.8	54.9	2.94
C	2.08 Q/QA28 + 0.258 TEMP + 0.209 SUNC + 0.0977 AGFC + 0.9822	135	56.4	42.0	3.86
C,L	0.305 Q/QA28 + 0.0663 TEMP + 0.0218 SUNC + 0.0132 AGFC + 0.7323	135	28.9	13.2	0.910
C,E,S_c	-0.785 QSTD - 2.30 Q/QA14 + 3.02 Q/QA28 + 6.28 SUN - 2.46 AGS				
	+ 0.190 SUNC + 0.0850 AGFC + 4.118	142	72.5	50.5	2.72

Table 34. Regression equations - chemical oxygen demand

Treatment	Regression equation	N	R^2	F	SE
A	17.9 Q/QA4 + 2.91 Q/QA28 + 4.52 lnSED + 0.848 TEMP + 0.217 AGSC				
	+ 0.139 AGFC - 24.18	305	48.2	46.2	15.0
A, <106	-5.45 lnQ+4 + 17.4 Q/QA5 - 11.1 Q/QA14 + 9.15 Q/QA28 + 2.70 lnSED	300	42.9	31.3	13.3
	+ 0.306 TEMP + 8.76 SUN + 31.70				
W	27.3 Q/QA4 + 0.608 TEMP + 0.283 AGSC + 0.07971	171	32.9	27.3	17.4
W,L	0.473 Q/QA4 + 0.0136 TEMP + 0.00614 AGSC + 2.858	171	24.9	18.3	0.382
C	6.81 Q/QA28 + 4.22 lnSED + 1.16 TEMP + 0.246 SUNC + 0.0233 M/Q				
	+ 15.58	134	60.2	38.7	10.0
C,L	0.188 Q/QA28 + 0.279 lnSED + 0.0775 TEMP + 0.00122 M/Q + 1.086	134	25.3	10.9	0.932

Table 35. Regression equations -- ammonia nitrogen

Treatment	Regression equation	N	R^2	F	SE
A	0.0785 Q/QA28 - 0.0598lnSED - 0.00448 TEMP + 0.00391 AGSC				
	- 0.00039 M/Q + 16.42	309	16.4	11.9	0.293
A,L	-0.719 lnQ-1 + 0.585 lnQA7 + 0.356 Q/QA28 - 0.399 AGF - 0.5125	309	13.7	12.0	0.976
A, <1.06	-0.0332 lnQ-5 - 0.101 Q/QA7 + 0.0790 Q/QA28 - 0.0925 AGF + 0.6538	299	16.0	14.0	0.199
A,S_c	-0.114 lnQ + 0.0838 lnQ-5 - 0.0318 QSTD + 0.110 Q/QA28				
	- 0.158 AGF + 0.6241	309	29.5	25.4	0.219
A,B	0.607 WQB + 0.0461 lnQ-3 + 0.0274 Q/QA28 - 0.0428lnSED - 0.06499	308	58.3	105.8	0.207
A[a]	-0.145 AGF + 0.5042	208	12.4	29.0	0.274
A[b]	-0.0635 QSTD + 0.192 Q/QA28 - 0.153 lnSED - 0.208 AGF + 1.164	101	50.1	24.1	0.263
A[c]	-0.327 lnQ-1 + 0.223 lnQA7 + 0.121 Q/QA28 - 0.216 AGF + 0.8460	258	28.2	24.8	0.288

[a]Includes data for 1967 to 1971.

[b]Includes data for 1971 to 1973.

[c]Includes data for 1968-1973.

Table 35 (Continued)

Treatment	Regression equation	N	R^2	F	SE
W	0.0990 SUN + 0.1139	172	3.7	6.6	0.232
W,L	0.408 SUN - 2.244	172	4.5	8.1	0.860
W,E,S_c	0.267 SUN + 0.1195	172	7.1	13.0	0.164
W^c	0.109 SUN + 0.08059	144	4.6	6.8	0.224
C	-0.153 lnQA7 - 0.319 Q/QA7 + 0.180 Q/QA28 - 0.244 AGF -0.00034 M/Q + 1.849	137	30.4	11.4	0.322
C,L	-0.242 lnQ-5 - 0.965 Q/QA7 + 0.550 Q/QA28 - 0.528 AGF + 1.062	137	18.5	7.5	1.11
C,E,S_c	-0.0887 QSTD + 0.0108 Q/QA28 - 0.246 AGF + 0.5539	142	38.5	28.8	0.251
C^c	-0.259 Q/QA7 + 0.219 Q/QA28 + 0.0168 AGSC + 0.2490	114	51.0	38.2	0.286

Table 36. Regression equations - organic nitrogen

Treatment	Regression equation	N	R^2	F	SE
A	-0.453 lnQ-1 + 0.935 lnQ-3 - 0.763 lnQ-5 + 0.330 lnSED + 0.0281 TEMP + 0.00745 AGSC + 0.7932	294	29.2	19.7	0.765
A,L	-0.346 lnQ-5 + 0.536 lnSED + 0.0720 TEMP - 2.171	294	18.2	21.5	1.98
A, <3.87	-0.239 lnQ-5 + 0.197 lnSED + 0.0226 TEMP + 1.271	290	21.1	25.4	0.628
A, S_C	-0.339 lnQ+1 + 0.222 lnSED + 0.00781 TEMP + 0.307 SUN + 1.800	294	31.0	32.5	0.584
A,B	0.681 WQB - 0.262 lnQ-5 - 0.364 DQ3/Q + 0.261 lnSED + 0.9137	290	48.8	68.0	0.648
W	3.04 lnQ+2 - 4.27 lnQ+1 + 0.354 QSTD + 0.657 lnSED - 0.00416 M/Q + 6.916	166	37.4	19.2	0.790
W,L	-0.486 lnQ-5 - 0.00255 M/Q + 3.597	166	7.7	6.8	1.24
W, <3.87	-0.253 lnQ-5 + 0.270 lnSED + 1.464	163	12.3	11.2	0.668
C	0.0509 TEMP + 0.0108 AGSC + 0.2403	128	32.7	30.3	0.563
C,L	1.45 lnQ+2 + 1.40 AGF + 0.0676 SUNC + 0.00425 M/Q - 14.41	128	26.3	11.0	2.35
C, <3.87	-0.154 lnQ-5 + 0.390 SUN + 1.438	126	12.2	8.6	0.580

Table 37. Regression equations -- nitrate nitrogen

Treatment	Regression equation	N	R^2	F	SE
A	3.62 lnQ-1 - 0.935 QSTD - 0.411 lnSED - 1.62 TEMP + 0.0232 AGSC - 16.95	305	61.7	96.2	2.32
A,B	0.925 WQB + 0.267 lnQ-5 - 0.737 Q/QA2 - 0.141 lnSED - 0.204 SUN + 3.728	303	94.0	931.2	0.922
W	3.16 lnQ + 3.74 lnQ-1 - 2.16 QSTD + 1.05 Q/QA21 + 0.0229 M/Q - 48.23	168	76.4	104.8	1.99
W,L	2.42 lnQ - 0.766 QSTD - 17.07	168	73.1	224.7	1.41
W, E, S_c	0.937 lnQ+4 + 3.71 lnQ + 1.57 lnQ-1 - 1.86 QSTD + 0.0188 M/Q - 41.87	172	79.2	126.1	1.69
W,E, S_m	3.98 lnQ+4 + 3.15 lnQA7 - 2.58 QSTD + 0.0180 M/Q - 46.73	172	68.1	89.2	2.65
C	-3.38 lnQ-5 + 5.21 lnQA7 - 0.790 Q/QA28 - 3.83 SUN - 0.0281 AGFC - 4.705	137	73.8	73.9	1.71

Table 37 (Continued)

Treatment	Regression equation	N	R^2	F	SE
C,L	0.264 lnQ+4 - 0.224 Q/QA28 - 1.24 SUN - 0.0378 SUNC - 0.0595 AGFC + 1.052	137	82.6	124.4	0.986
C,E,S_c	2.21 lnQA7 + 2.57 Q/QA14 - 2.08 Q/QA28 - 4.04 SUN + 0.0113 AGSC -0.0226 AGFC - 8.784	142	80.6	93.2	1.40
C,E, S_m	9.59 lnQ+4 - 7.09 lnQ+2 - 0.542 QSTD - 1.22 Q/QA28 - 3.52 SUN - 0.260 AGFC - 7.732	142	72.6	59.5	2.45

Table 38. Regression equations - total phosphate

Treatment	Regression equation	N	R^2	F	SE
A	1.03 Q/QA2 + 0.344 Q/QA21 - 0.216 AGF + 0.01046	111	43.5	27.5	0.613
A, <3.96	2.60 Q/QA2 - 1.49 Q/QA3 + 0.349 Q/QA28 - 0.3188	~~109~~ 118	51.4	37.4	0.449
A,B	0.676 WQB + 0.122 QSTD + 0.359 Q/QA7 - 0.00795 TEMP + 0.06710	111	51.2	27.8	0.579
W	-0.124 QSTD + 0.684 Q/QA2 + 0.663 Q/QA21 + 0.0344 TEMP - 0.296 AGF - 0.4394	64	75.7	36.2	0.377
W,L	-0.115 QSTD + 0.616 Q/QA21 - 0.0264 TEMP - 0.240 AGF - 0.698	64	57.6	20.0	0.314
W, <3.96	equation identical with treatment W	64	75.7	36.2	0.377
C	0.964 Q/QA7 + 0.675 SUN + 0.04656	47	35.8	12.3	0.797
C,L	0.499 Q/QA7 + 0.252 AGS - 0.607	47	38.8	14.0	0.408
C, <3.96	-0.294 Q/QA28 + 0.8805	45	25.1	14.4	0.532

Table 39. Regression equations - orthophosphate

Treatment	Regression equation	N	R^2	F	SE
A	0.0768 Q/QA21 - 0.0473 lnSED - 0.0167 TEMP - 0.00334 AGFC + 0.8107	309	38.6	47.8	0.283
A, <1.19	0.0362 lnQ+4 + 0.0387 Q/QA28 - 0.0155 TEMP + 0.2429	297	34.5	51.3	0.228
A,B	0.614 WQB + 0.128 lnQ - 0.0577 QSTD + 0.0284 Q/QA21 - 0.0580 lnSED - 0.00529 TEMP - 0.4227	308	70.5	120.0	0.197
W	0.137 lnQ-3 + 0.107 Q/QA28 - 0.0920 AGS + 0.00407 SUNC - 0.8736	172	43.0	31.5	0.142
W,L	1.05 lnQA7 + 1.11 Q/QA28 - 0.0364 AGFC - 12.32	172	26.2	19.9	1.91
W, <1.19	Equation identical with treatment W	172	43.0	31.5	0.142
C	-36.3 SUN + 71.5 AGS + 61.6 AGF - 0.000226 M/Q - 96.54	137	38.4	20.6	0.322
C,L	-0.876 SUN 1.01 AGF - 0.00692 AGFC + 0.3089	137	22.9	13.2	1.09
C, <1.19	-15.5 SUN + 30.0 AGS + 25.7 AGF + 0.0677 lnSED - 40.39	125	35.6	16.6	0.240

Table 40. Regression equations - fecal coliform

Treatment	Regression equation	N	R^2	F	SE
A	-2982 DQ3/Q + 7283 Q/QA3 - 1103 Q/QA14 - 5697	130	68.5	91.2	1269
A,L	0.989 lnQ+4 + 1.22 Q/QA28 + 0.666 lnSED + 0.0491 AGFC - 8.352	130	26.2	11.1	3.07
A, <8181	1422 Q/QA4 + 242 lnSED + 238 AGF - 2392	128	31.4	19.0	844
A,B	0.208 WQB - 3120 DQ3/Q + 6668 Q/QA4 + 1775 Q/QA14 - 4566	129	70.7	74.8	1232
W	-2731 DQ3/Q + 7802 Q/QA4 - 4239 Q/QA14 + 2455 Q/QA21 - 5428	74	82.8	82.9	1238
W,L	1.07 Q/QA28 + 2.65 lnSED + 0.196 SUNC - 0.476 AGSC - 10.69	74	53.6	19.9	2.48
W, <8181	1240 Q/QA14 + 272 lnSED - 2016	72	41.8	24.8	962
W,L,E,S_c	1.07 lnQ-3 + 1.43 Q/QA28 + 0.0752 TEMP + 0.107 SUNC - 0.268 AGSC - 6.261	85	50.7	16.3	1.54
C	774 DQ3/Q + 250 Q/QA14 + 190	56	24.4	8.5	507
C,L	1.38 lnQ+4 - 5.881	56	22.4	15.6	3.13
C,L,E,S_c	1.31 lnQ+4 - 5.421	70	42.5	50.4	1.96

Table 41. Regression equations - total plankton

Treatment	Regression equation	N	R^2	F	SE
A,L	-0.851 lnQ-5 - 0.453 Q/QA7 + 0.077 TEMP - 0.0751 AGSC				
	-0.0160 AGFC + 11.48	147	54.5	33.7	1.02
W	-550 lnQ - 231 AGF - 29.3 SUNC + 5290	81	54.9	31.3	301
W,L	-0.555 QSTD - 0.482 lnSED - 0.0509 TEMP - 0.0327 SUNC + 10.32	81	47.2	17.0	0.842
W,L,E,S_C	-0.402 lnQ-5 - 0.279 QSTD - 0.272 lnSED - 0.469 SUN				
	- 0.458 AGF - 0.0339 SUNC + 12.13	85	58.5	18.3	0.594
C	-55.4 QSTD + 46.8 TEMP + 33.78	66	49.5	30.9	148
C,L	-1.03 lnQA7 + 2.60 SUN + 1.18 AGF - 0.0701 AGSC + 10.35	66	77.4	52.4	0.648
C,L,E,S_C	-0.460 lnQ+2 - 0.547 lnQA7 + 2.79 SUN + 1.24 AGF - 0.0635 AGSC				
	+ 10.03	70	82.6	61.0	0.531

Table 42. Regression equations -- diatoms

Treatment	Regression equation	N	R^2	F	SE
A,L	-0.567 lnQ-5 - 0.245 QSTD + 0.0836 TEMP - 0.0312 SUNC - 0.0615 AGSC + 8.645	147	49.4	27.5	1.19
W	-450 lnQ - 17.3 TEMP - 32.7 SUNC + 27.2 AGSC + 4426	81	60.3	28.8	259
W,L	-1.71 lnQ - 0.674 AGF - 0.0947 SUNC + 20.17	81	56.9	33.8	0.898
C	-53.2 QSTD + 43.2 TEMP + 26.21	66	49.4	30.8	138
C,L	-0.951 lnQA7 + 2.63 SUN + 1.17 AGF + 0.109 AGFC - 0.0110 M/Q + 9.097	66	76.8	39.7	0.756

Table 43. Regression equations - flagellates

Treatment	Regression equation	N	R^2	F	SE
A	79.7 TEMP + 1069	146	11.4	18.5	2176
A,L	0.833 DQ3/Q + 0.0334 TEMP - 0.0308 AGSC + 6.906	146	24.0	15.0	1.04
A,S_c	932 Q/QA4 + 76.5 TEMP + 116.3	146	19.4	17.2	1692
W	N.S.	80	-	-	-
W,L	N.S.	80	-	-	-
C	-340 QSTD + 1873 Q/QA4 + 130 TEMP + 1421 SUN - 2.35 M/Q - 1028	66	53.4	13.8	943
C,L	0.902 Q/QA7 + 0.134 TEMP - 0.0198 AGSC + 5.424	66	49.0	19.8	0.785

Table 44. Explanation of abbreviations used in Appendix B, Tables 27 to 43

A = regression equation developed from the annual or the complete data set, 314 weeks, 1967 to 1973

B = data for water quality parameter collected at an upstream sampling location at Boone, Iowa (labelled as Station 1 in Figure 2) included in the regression analysis as an explanatory variable

C = regression equation developed from the cold season data set, for which the river water temperature less than or equal to 10°C

E = missing data for dependent parameter estimated from regression equation developed from the raw data and substituted into the data set

F = ratio of two independent estimates of the same variance (54)

L = natural logarithm of dependent parameter data used in the regression analysis

<n = only dependent parameter data less than n included in the regression analysis

N = number of weeks of data used in the regression analysis

N.S. = no explanatory parameters were statistically significant at $\alpha = 0.05$

R^2 = per cent of the total variation about the mean of the dependent parameter explained by the regression equation (19)

S_c = smoothing routine applied to the concentration data for the dependent parameter, by averaging the concentration of the dependent parameter, C for the ith week with those for the prior and following week:

$$\overline{C}_i = (C_{i-1}+C_i+C_{i+1})/3,\ \overline{C}_i = \text{smoothed value}$$

Table 44 (Continued)

S_m = smoothing routine applied to values of the dependent parameter representative of mass or quantity, i.e., streamflow, Q, times concentration, C, by averaging the mass for the dependent parameter for the ith week with those for the prior and following week:

$$\bar{M}_i = (Q_{i-1}C_{i-1}+Q_iC_i+Q_{i+1}C_{i+1})/3, \quad \bar{M}_i = \text{smoothed value}$$

the smoothed concentration value is then calculated:

$$\bar{C}_i = \bar{M}_i/Q_i$$

SE = standard error of the estimate, S (54):

$$S = \sqrt{\frac{(Y-\hat{Y})^2}{N-2}}, \quad \begin{array}{l} Y = \text{value for dependent parameter} \\ \hat{Y} = \text{estimated value for dependent parameter} \end{array}$$

W = regression equation developed from the warm season data set, for which the river water temperature is greater than 10°C

WQB = data for water quality parameter at Boone, Iowa, included as an explanatory parameter