ANALYSIS OF SURFACE-WATER QUALITY DATA FOR PUERTO RICO

by Abraham Rodríguez

and

Rafael Muñoz
Department of Chemical Engineering

Project No. 04
Grant Agreement No. 14-08-0001-G-1611

FINAL TECHNICAL REPORT
TO
U.S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

The activities on which this report is based were financed in part by the Department of the Interior, U.S. Geological Survey, through the Puerto Rico Water Resources Research Institute.

September 1990

ABSTRACT

The purpose of this study was to analyze water quality data for several rivers in Puerto Rico to determine, through statistical methods, the presence of significant trends in the most significant quality parameters. The data analyzed correspond to samples which have been gathered over the years by the U. S. Geological Survey in its sampling stations network in Puerto Rico.

The rivers studied were the Grande de Arecibo, Grande de Añasco, Bayamón, Cagüitas, Fajardo, Gurabo, Grande de Loiza, Grande de Manatí, La Plata, and the Rosario. The water quality parameters selected for consideration were pH, dissolved oxygen, total nitrogen, total phosphorous, dissolved potassium, and fecal coliforms. Volumetric flow was also considered, to determine the potential use for public water supply that the rivers could show.

A preliminary analysis of the water quality data produced three important findings: (1) graphical methods are not adequate for determining the presence and magnitude of trends; (2) Lineal regression analysis can be used to determine the trend directions when the calculations do not involve extreme values but cannot be used to determine the magnitudes of these trends; and (3) Non-parametric analysis is the best alternative for detecting the presence, direction, and magnitude of trends in surface water quality data.

The Kruskal-Wallis test was used to determine the

GENERAL INDEX

presence of seasonality in the data and it was found that in most cases the data showed a dependency on the seasons of the year which characterize Puerto Rico's climate. To take into account this seasonality phenomenon, different time bases were used to analyze the data.

Non-parametric techniques permitted determining the presence of trends and their magnitudes. The results obtained in these tests clearly demonstrate that the water quality of rivers in Puerto Rico undergoes much greater change-for better or worse-at testing sites closer to the sea than at upstream sites, where water quality usually remains extremely stable. The trends depend largely on the geographical locations of the sampling stations, as well as on the location of the rivers. The results also indicate that the water quality in the Rio Grande de Añasco, Rio Grande de Arecibo, Rio Bayamón, and Rio Rosario could be satisfactory if their fecal coliform levels are reduced. These streams do not show any significant trends to negative changes in their water quality. These findings, added to the fact that volumetric flow of these streams is considerable, suggest that they have a great potential for additional water supply use.

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

One of the most important problems that people all over the world have to face today is the enormous pollution caused by wastes to the air, water, and soil environments. Even a small island like Puerto Rico is not free from this problem, which is being caused mainly by the desire for comfort and luxuries in the daily lives of its inhabitants. Because of this, it is necessary to do something about the threat to our environment. Water pollution is one of the critical areas of major concern, because we are causing the direct contamination of surface waters and the indirect contamination of rain and ground waters. The main objective of this study is to analyze the problem of surface water quality in Puerto Rico.

Surface water sources in Puerto Rico are sufficient to meet the water use demands of the island only when rainfall is normal. Although the total usable surface water resources in the island (1.1 billion gallons/day) are more than ample to supply existing and anticipated future needs, many of the available resources are not located near demand areas. As a result, the sources under utilization have to be drawn upon continuously and intensely for water supply for domestic, industrial, and agricultural applications.

The judicious use of these surface water sources requires that the regional water quality patterns be followed over time and projected into the future to assure that applicable water quality criteria are always met and to indicate if there is the need to tap other available surface water resources which remain unused. Toward these objectives, this project involved the statistical analysis of a very large number of water quality measurements which have been made over more than three decades using a sampling network which covers the principal rivers of Puerto Rico. This network of surface water quality stations includes over 70 long-term stations in streams and in eleven lake sites. The analysis of its data could also provide a reliable technical base for improving the water resources management process in the island.

II. WATER QUALITY INFORMATION FOR PUERTO RICO

A. Description of the Puerto Rico Water Quality Network

The Geological Survey has published since 1941 records of data for the chemical and physical characteristics of Puerto Rico's surface waters. The data on the quality of surface water include chemical quality, biological and microbiological quality, and water temperatures and sediment, and are collected from designated sampling sites at predetermined intervals, such as monthly or less frequently. The locations of the surface water quality stations are shown in Figure 1.

Stations are listed in downstream direction along each main river. Stations on tributaries are listed between stations on the main stream in the order in which those tributaries enter the main stream. Each water quality station has an assigned station number. The numbering begins at the northwest cape of Puerto Rico and proceeds, basin by basin, in a clockwise direction around the island.

Water samples for analyses are collected at or near gaging stations. The discharge records at these stations are used in conjunction with the computations of the chemical constituents and sediment loads.

The chemical quality information includes concentrations of individual dissolved constituents and certain properties or characteristics such as hardness, sodium adsorption ratio, specific conductance, and pH. The biological quality includes

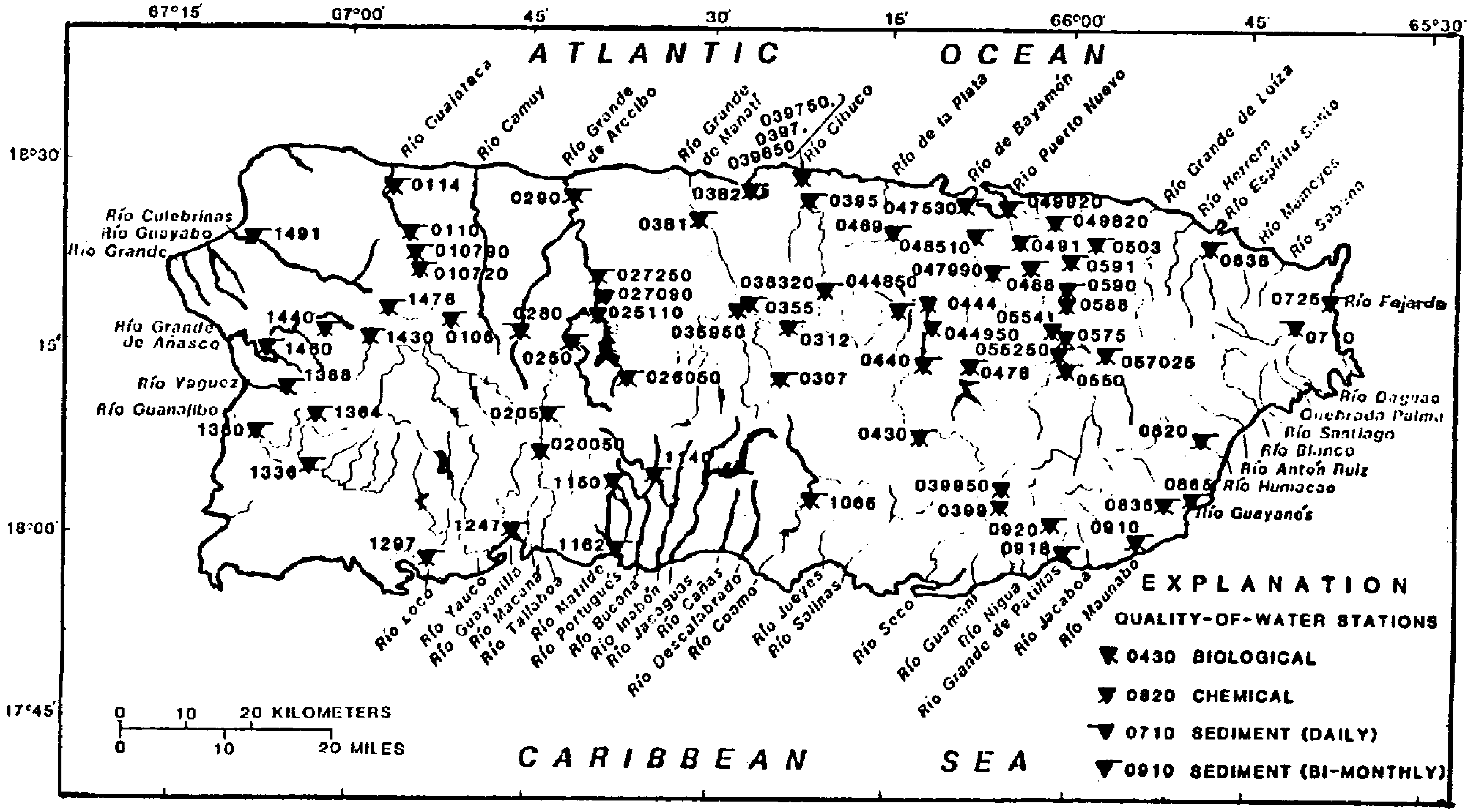

FIGURE 1: Location of Water Quality Stations in Puerto Rico

qualitative and quantitative analyses of plankton, bottom organisms, and particulate inorganic and amorphous matter present. Microbiological quality includes quantitative identification of certain bacteriological indicator organisms.

B. Selection of Rivers and Stations for Study

The selection of the surface water bodies and the stations to be considered in this study was based on the following factors: the availability of data, the number of water quality parameters reported for various stations, the length and continuity of the water quality records (1) (2) (3), and the importance of the stream as an actual or a potential water supply source. Depending on these factors, the specific stations and parameters were chosen, the period of time for data analysis was defined, and the method of statistical data analysis to be employed was identified.

Based on all these considerations, a selection of 22 sampling stations distributed along 10 different rivers was made for this study. The rivers and stations selected are shown in Figure 2. Those chosen for their present importance as water supply sources were the Rio Grande de Loiza, Rio Cagüitas, Rio Gurabo, Rio de Bayamón, Rio La Plata, Rio Grande de Manatí, and Rio Grande de Arecibo. It should be noted that Rio Cagüitas and Rio Gurabo are tributaries to the Rio Grande de Loiza. The Rio Grande de Añasco, Rio Fajardo, and Rio Rosario were chosen for study because they remain largely

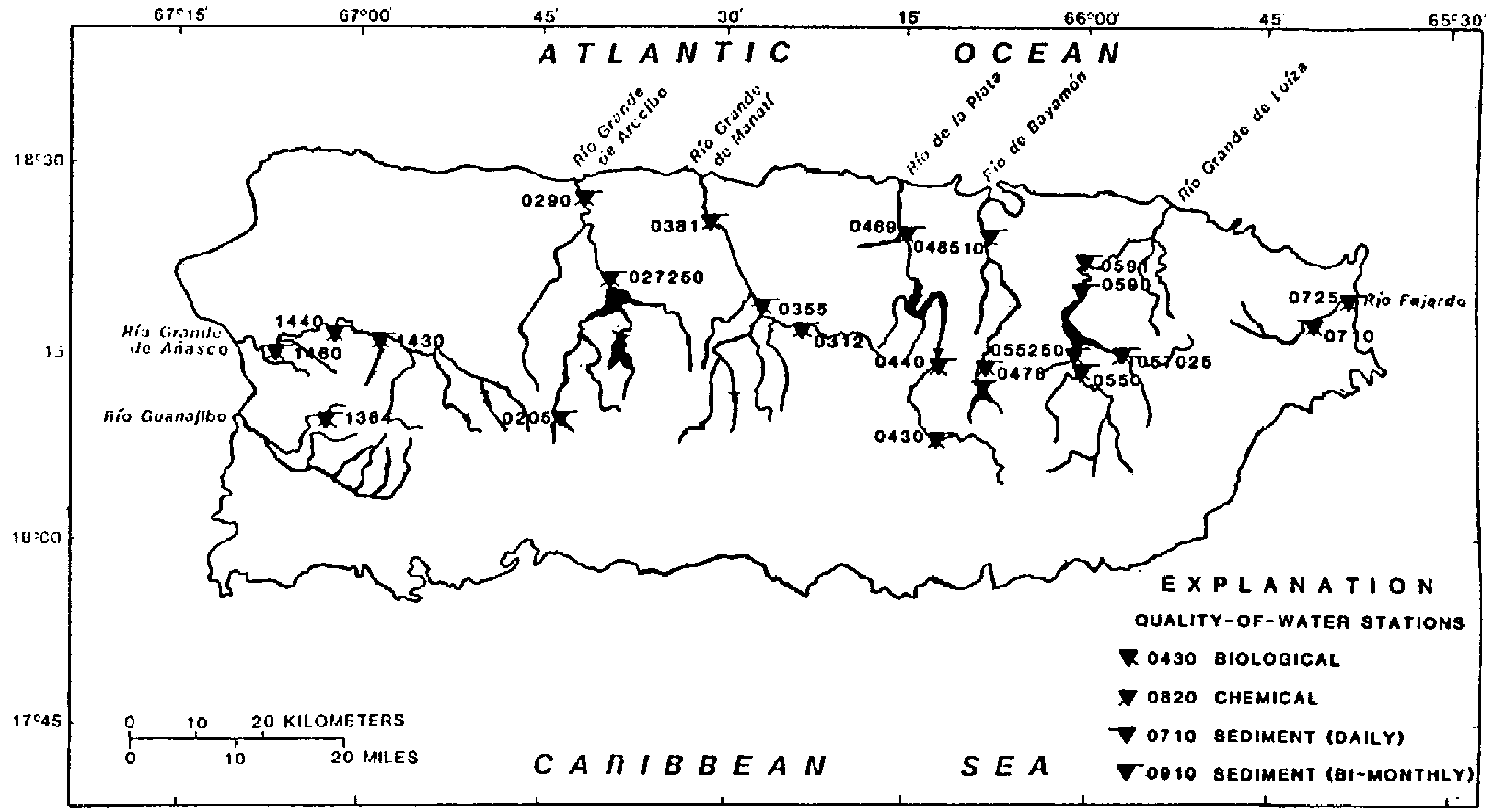

FIGURE 2: Location of Selected Rivers and Water Quality Stations

untapped for water supply uses. For each of these 10 rivers the Geological Survey has maintained continuous water quality records for more than 40 parameters over long periods of time. In each of these rivers the sampling stations were chosen to allow the comparison of statistical trends between upstream and downstream zones of the streams.

To better understand the maps, tables, and graphs in this report it is necessary to indicate the form in which the stations are represented. Numbers from 1 to 22 are used in the tables, 8-digit numbers are used in graphs, and 4-to 6-digit numbers are used in maps. Maps and graphs follow the same numbering nomenclature employed by the U.S. Geological Survey in its publications. The relationship between the three station identification numbers is as follows:

Table Number	River basin	Graph Number	Map Number
Station 1	Rio Grande de Añasco	501 43 000	1430
Station 2	Rio Grande de Añasco	501 44 000	1440
Station 3	Rio Grande de Añasco	501 46 000	1460
Station 4	Rio Grande de Arecibo	500 20 500	0205
Station 5	Rio Grande de Arecibo	500 27 250	027250
Station 6	Rio Grande de Arecibo	500 29 000	0290
Station 7	Rio de Bayamón	500 47 600	0476
Station 8	Rio de Bayamón	500 48 510	048510
Station 9	Rio Cagüitas	500 55 250	055250

Station 10	Rio Fajardo	500 71 000	0710
Station 11	Rio Fajardo	500 72 500	0725
Station 12	Rio Gurabo	500 57 025	057025
Station 13	Lago Loiza at Damsite	500 59 000	0590
Station 14	Rio Grande de Loiza	500 55 000	0550
Station 15	Rio Grande de Loiza	500 59 100	0591
Station 16	Rio Grande de Manatí	500 31 200	0312
Station 17	Rio Grande de Manatí	500 35 500	0355
Station 18	Rio Grande de Manatí	500 38 100	0381
Station 19	Rio de La Plata	500 43 000	0430
Station 20	Rio de La Plata	500 44 000	0440
Station 21	Rio de La Plata	500 46 900	0469
Station 22	Rio Rosario	501 36 400	1364

C. Data Access and Selection of Water Quality Parameters

The surface water quality data for the selected stations was accessed directly from the USGS data bank in San Juan, Puerto Rico. Typically, the USGS assays sample for over 40 different parameters. Out of these, the 7 parameters shown in Table 1 following streamflow are considered to be the most important indicators of water quality in Puerto Rico and were therefore selected for this study. Dissolved oxygen is an indicator of organic pollution, coliforms are a measure of fecal contamination, and nitrogen, phosphorous, and potassium measure environmental impacts from agricultural activities.

TABLE 1: Selected Water Quality Parameters

Streamflow (cfs)
pH (standard units)
Dissolved Oxygen (mg/l)
Dissolved Oxygen (% saturation)
Fecal Coliforms (cols./100ml)
Dissolved Potassium (mg/l as K)
Total Nitrogen (mg/l as NO_3)
Total Phosphorous (mg/l as P)

III. RELATED WORK

The analysis of surface water quality data in this study consists of determining if the data analyzed present trends to decrease or to increase through a determined period of time.

No systematic analysis of the surface water quality data gathered in Puerto Rico had been undertaken prior to this study. However, many such statistical studies of regional water quality patterns have been done for other parts of the United States. For the purpose of simplification, previous studies will be presented in a format which illustrates the alternative statistical techniques which are applicable for determining if the data show trends to change as a function of time.

A. Selection of Applicable Statistical Techniques

Most hypothesis-testing procedures are based on the assumption that random samples are selected from normal populations. Fortunately, most of these tests are still reliable when we experience slight departures from normality, particularly when the sample size is large. Traditionally, these procedures have been referred to as parametric methods. In addition there exists a great number of alternative test procedures, called non-parametric or distribution-free methods, that often assume no knowledge whatsoever about the distributions of the underlying populations, except perhaps that they are continuous (4)(5).

Non-parametric tests have gained a certain appeal in recent years for several reasons. First, the computations involved are usually very quick and easy to carry out. Second, the data need not be quantitative measurements but can be in the form of qualitative responses such as "defective" versus "nondefective," "yes" versus "no," and so forth. Frequently they are values of an ordinal scale to which we assign ranks. On an ordinal scale the subjects are ranked according to a specified order, and a non-parametric test analyzes the various ranks. A third and perhaps the most important advantage in using non-parametric tests is that they are encumbered with less restrictive assumptions than their parametric counterparts (5).

However, some parametric methods have been utilized in the analysis of surface water quality data. These include the following: (i) analysis of variance, (ii) simple and multiple linear regression (6), (iii) t-tests, and (iv) multivariate-normal procedures such as discriminant function analysis (3) (7).

Most investigators concur on the advantages in using non-parametric rather than parametric methods when analyzing water quality data. For example, Berryman et al. (8) conclude that parametric tests are limited by the assumption of normality and that the non-parametric tests are considered better adapted to water quality time series. Van Belle and Hughes (9) conclude that non-parametric tests have been developed because the assumption of classical parametric methods (i.e., normality, linearity, independence) are usually not met by typical water quality data. They go on to say that additional idiosyncrasies of the data, such as missing values, censored data, and seasonality, compound the analysis problems and that non-parametric methods, being more flexible, can handle these problems more easily.

Lettenmaier (10) conducted a study on the detection of trends in water quality data from records with dependent observations. He developed a method for summarizing the power of the parametric tests against step and linear trends using dimensionless numbers.

Hirsch et al. (11) evaluated non-parametric methods that can manage the problems due to missing values, seasonality, and values below the limit of detection. They conclude that these methods are intended to be exploratory methods for identifying and quantifying changes in water quality time series. As the number and length of water quality time series grow, it is desirable to have a set of objective automatic procedures that are reasonably powerful over a wide range of situations for identifying trends. They believe that the non-parametric methods are useful and appropriate for these purposes.

Hipel (12), applying enhanced approaches for non-parametric methods to water quality time series, as well as employing well-designed simulation experiments, clearly demonstrated the efficacy of utilizing non-parametric tests in environmental assessments.

Alexander and Smith (13) used the Seasonal Kendall test for determining trends in lead concentration in some rivers in the United States. Their studies serve to illustrate the applicability of non-parametric techniques and to advise us on the presence of important factors in surface water quality data that deserve consideration, such as missing values, ties, censoring values, and seasonality.

Hughes and Millard (14) found that a common problem arises in testing for trends in water quality when observations are reported as "less than detection limit." They proposed a non-parametric statistic similar to Kendall's law for correcting this problem. Monte Carlo simulations show that the normal approximation to the distribution of this statistic is quite good, even for small samples and a large proportion of censored observations. The statistics also demonstrate a greater power in treating as ties all observations less than the target censored observations.

The previous review shows that very often the normality assumptions are not justified in water quality data and that this type of data is usually subject to missing values, censored data, and seasonality. Upon these considerations, it was decided to use non-parametric methods in this study instead of parametric ones. A great number of non-parametric methods have been applied to surface water quality data, among them, Spearman's rho test, Mann Witney's test (8) (10), Kendall test (8), Seasonal Kendall test (11) (13) (15) (16), Seasonal Kendall Slope Estimator (11), Kruskal Wallis test (11) (17), Mann-Kendall test (11) (18), Sen's test (11), Wilcoxon signed rank test (17), and the Seasonal Hodges-Lehmann estimator (19).

In order to select the specific non-parametric techniques for use in this study, an examination of the data to be analyzed had to be undertaken. This was needed because surface water quality data may exhibit several properties that can affect the results in the determination of long-range trends. The techniques to be applied must be capable of allowing for the presence of these properties in the data. The most important of these properties are the following: (i) missing values, that is, values that are not reported at a time interval in the selected period; (ii) censoring values, which correspond to values reported as less than the detection limit of the instrument used; (iii) ties, which are repeated values in some intervals of the period studied; and (iv) seasonality, which is the existence of different distributions for different times of the year. After a careful analysis of the water quality data, which is represented in Appendix D, the following techniques were finally selected:

(i) Kruskal-Wallis test (for determining the presence of seasonality in the data)

(ii) Mann-Kendall test (for investigating trends in those data which do not show seasonal behavior)

(iii) Seasonal Kendall test (for investigating trends in the data which present seasonal behavior)

(iv) Sen's test and Seasonal Kendall Slope Estimator (for determining the magnitudes of the trends)

B. Description of the Non-parametric Techniques Selected

Before describing in detail the non-parametric techniques which were selected previously, it is pertinent to know some important unique characteristics of each one of them.

The Kruskal-Wallis test is essentially an analysis of variance for non-parametric data. The sets of data used need not be of the same size and the number of sets is not restricted. This technique is useful in determining whether or not the sets of data that come from the same population present an equal distribution.

The Mann-Kendall test and the Seasonal Kendall test are utilized for determining trends over the course of a specific time period. The Mann-Kendall test is capable of handling missing values, censored data, and tie values, but cannot be used in situations in which the seasonality phenomenon is present. This technique is easy to utilize and has good reliability. The Seasonal Kendall test is slightly more complicated than the Mann-Kendall test, but it has the advantage that it can be used without losing reliability in situations in which the seasonality phenomenon is present.

The Sen's test and the Seasonal Kendall Slope Estimator test are utilized for determining the magnitude of the trends detected by the Mann-Kendall and the Seasonal Kendall tests, respectively. The Sen's test is utilized in cases where seasonality does not apply. The Seasonal Kendall Slope Estimator was developed from the Sen's test with the purpose of taking into consideration the effect of seasonality. Both tests are capable of managing missing values, censored data, and tie values. The advantage of the Seasonal Kendall Slope Estimator over the Sen's test is based on the applicability of the first in cases where the seasonality effect is present. Both tests possess the advantage that they can be incorporated along with the Mann-Kendall or Seasonal Kendall tests in the same computational program.

To summarize, if the data under analysis do not present the seasonality phenomenon, the Mann-Kendall and the Sen's tests should be used for detecting the trends and magnitudes. If, on the other hand, the data present the seasonality phenomenon, the Seasonal Kendall and the Seasonal Kendall Slope Estimator should be used for the determination of trends and magnitudes.

1. Kruskal-Wallis Test (20) (21)

The Kruskal-Wallis test is an extension of the Wilcoxon rank sum test from two to k independent data sets. These data sets need not be drawn from underlying distributions that are normal or even symmetric, but the k distributions are assumed to be identical in shape. The null hypothesis is

H_0: The populations from which the k data sets have been drawn have the same mean.

The alternative hypothesis is

H_A: At least one population has a mean larger or smaller than at least one other population.

The data take the form,

Population

1	2	3	...	k
X_{11}	X_{21}	X_{31}	...	X_{k1}
X_{12}	X_{22}	X_{32}	...	X_{k2}
⋮	⋮	⋮	⋮	⋮
X_{1n1}	X_{2n2}	X_{3n3}	...	X_{knk}

The total number of data is $m = n_1 + n_2 + \ldots + n_k$, where the n_i need not be equal. The steps in the testing procedure are as follows:

a. Rank the m data from smallest to largest; that is, assign the rank 1 to the smallest datum, the rank 2 to the next largest, and so on.

b. Compute the sum of the ranks for each data set. Denote this sum for the jth data set by R_j.

c. Compute the Kruskal-Wallis statistic as follows:

$$K_w = \left[\frac{12}{m(m+1)} \sum_{i=1}^{k} \frac{R_j^2}{n_j}\right] - 3(m+1) \qquad \textbf{(1)}$$

d. For an α level test, reject H_0 and accept H_A if $K_w \geq X^2_{1-\alpha,k-1}$, where $X^2_{1-\alpha,k-1}$, is the $1-\alpha$ quantile of the chi-square distribution with k-1 degrees of freedom, as obtained from Appendix A, where k is the number of data sets.

2. Mann-Kendall Test (11)(21)

The Mann-Kendall test uses only the relative magnitudes of the data rather than their measured values. We consider the case in which only one datum per time period is taken, where the time period may be a day, week, month, and so on.

The first step is to list the data in the order in which they were collected over time: X_1, X_2,...,X_n, where X_i is the datum at time i. Then the sign of all n(n-1)/2 possible differences $X_j - X_k$, where j>k is detected. These differences are X_2-X_1, X_3-X_1,...,X_n-X_1, X_3-X_2, X_4-X_2,..., X_n-X_{n-2}, X_n-X_{n-1}. A convenient way of arranging the calculations is shown in Table 2.

Let sgn $(X_j - X_k)$ be an indicator function that takes on the values 1, 0, or −1 according to the sign of $X_j - X_k$:

$$\begin{aligned} \text{sgn}\ (X_j - X_k) &= 1 \quad \text{if} \quad X_j - X_k > 0 \\ &= 0 \quad \text{if} \quad X_j - X_k = 0 \\ &= -1 \quad \text{if} \quad X_j - X_k < 0 \end{aligned}$$

Then compute the Mann-Kendall Statistic

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn\ (x_j - x_k) \qquad (2)$$

which is the number of positive differences minus the number of negative differences.

TABLE 2: Differences in Data Values Needed for Computing the Mann-Kendall Statistic S to Test for Trend

Data Values Listed in the Order Collected Over Time							
X_1	X_2	X_3	X_4	...	X_n	No. of + Signs	No. of - Signs
	$X_2 - X_1$	$X_3 - X_1$	$X_4 - X_1$	...	$X_n - X_1$		
		$X_3 - X_2$	$X_4 - X_2$	...	$X_n - X_2$		
			$X_4 - X_3$	...	$X_n - X_3$		
					$X_n - X_4$		
					$X_n - X_{n-1}$		
					S =	(Sum of + Signs)	− (Sum of - Signs)

Suppose we want to test the null hypothesis H_0, of no trend against the alternative hypothesis, H_A, of an upward trend. For this, the variance of S is computed by using the following equation:

$$VAR(S) = \frac{1}{18}[\, n(n-1)(2n+5)] \tag{3}$$

Then S and VAR(S) are used to compute the test statistic Z as follows:

$$Z = \frac{S-1}{[VAR(S)]} \qquad \text{if} \quad S > 0 \tag{4}$$

$$Z = 0 \qquad \text{if} \quad S = 0 \tag{5}$$

$$Z = \frac{S+1}{[VAR(S)]} \qquad \text{if} \quad S < 0 \tag{6}$$

A positive (negative) value of Z indicates an upward (downward) trend. If the null hypothesis, H_0, of no trend is true, the statistic Z has a standard normal distribution and hence we use Appendix B to decide whether to reject H_0. To test for upward trend (a one-tailed test (22)), H_0 is rejected if Z is greater than $Z_{1-\alpha}$. We reject H_0 in favor of the alternative hypothesis of a downward trend if Z is negative and the absolute value of Z is greater than $Z_{1-\alpha/2}$.

3. Seasonal Kendall test (11) (17)

The Seasonal Kendall test for trend is insensitive to the existence of seasonality. This test is a generalization of the Mann-Kendall test. The Seasonal Kendall test may be used for months, but conceptually the test may also used for other "seasons" (eg., four quarters of the year or the three 8-hour periods of the day).

Let X_{il} be the datum for the ith season of the lth year, K the number of seasons, and L the number of years. The data for a given site (sampling station) are shown in Table 3. The null hypothesis, H_0, is that the X_{il} are independent of the time (season and year) they were collected. The alternative hypothesis, H_A, is that for one or more seasons the data are not independent of time.

Let S_i be the statistic computed for season i, that is

$$S_i = \sum_{k=1}^{n_i-1} \sum_{l=k+1}^{n_i} sgn(x_{il} - x_{ik}) \qquad (7)$$

where $l > k$, n_i is the number of data (over years) for season i, and

$$\begin{aligned} \text{sgn } (X_{il} - X_{ik}) &= 1 \quad \text{if} \quad X_{il} - X_{ik} > 0 \\ &= 0 \quad \text{if} \quad X_{il} - X_{ik} = 0 \\ &= -1 \quad \text{if} \quad X_{il} - X_{ik} < 0 \end{aligned}$$

TABLE 3: Data for the Seasonal Kendall Test at One Sampling Station

		SEASON			
		1	2	...	K
YEAR	1	X_{11}	X_{21}	...	X_{K1}
	2	X_{12}	X_{22}	...	X_{K2}
	⋮	⋮	⋮	⋮	⋮
	L	X_{1L}	X_{2L}	...	X_{KL}
		S_1	S_2	...	S_K

$VAR(S_i)$ is computed as follows:

$$VAR(S_i) = \frac{1}{18}\left[n_i(n_i-1)(2n_i + 5)\right] \qquad \textbf{(8)}$$

After the S_i and $VAR(S_i)$ are computed, we sum across the k seasons:

$$S' = \sum_{i=1}^{k} S_i \qquad \textbf{(9)}$$

$$VAR(S') = \sum_{i=1}^{k} VAR(S_i) \qquad \textbf{(10)}$$

Next, compute

$$Z = \frac{(S' - 1)}{[VAR(S')]^{1/2}} \qquad \text{if } S' > 0 \qquad \textbf{(11)}$$

$$Z = 0 \qquad \text{if} \quad S' = 0 \qquad \textbf{(12)}$$

$$Z = \frac{(S' + 1)}{[VAR(S')]^{1/2}} \qquad \text{if} \quad S' < 0 \qquad \textbf{(13)}$$

If the alternative hypothesis is for an upward trend at the α level (a one-tailed test), we reject H_0 if Z is greater than $Z_{1-\alpha}$. Reject H_0 in favor of downward trend if Z is negative and the absolute value of Z is greater than $Z_{1-\alpha}$.

4. Sen's Test and Seasonal Kendall Slope Estimator (11) (21)

Sen's test is not greatly affected by gross data errors or outliers.

First, compute the N' slope estimates, Q, for each station:

$$Q = \frac{x_{i'} - x_i}{i' - i} \qquad \textbf{(14)}$$

where, $X_{i'}$ and X_i are data values at times (or during time periods) i' and i, respectively, and where $i' > i$; N' is the number of data pairs for which $i' > i$. The median of these N' values of Q is Sen's estimator of slope.

The median of the N' slope estimates is obtained in the usual way. That is, the N' values of Q are ranked from smallest to largest.

Sen's estimator = median slope

$$= Q_{[(N'+1)/2]} \qquad \text{if} \quad N' \quad \text{is odd}$$

$$= \tfrac{1}{2}(Q_{[N'/2]} + Q_{[(N'+2)/2]}) \quad \text{if} \quad N' \quad \text{is even}$$

The Seasonal Kendall Slope Estimator is a generalization of Sen's estimator of slope. First, the individual N_i slope estimates for the ith season are computed from;

$$Q_i = \frac{X_{il} - X_{ik}}{l - k} \qquad (15)$$

where, as before, X_{il} is the datum for the ith season of the lth year, and X_{ik} is the datum for the ith season of the kth year, where l > k. Do this for each of the K seasons. Then the $N_1' + N_2' + \ldots N_k' = N'$ individual slope estimates are ranked and their median is found. This median is the seasonal Kendall slope estimator.

5. Modification to Accommodate Missing Values (11) (16)

To accommodate missing values, it is necessary to extend the definition of the sgn function given in the Kendall and Seasonal Kendall tests to handle missing values. Define sgn $(X_l - X_i)$ in Kendall test or sgn $(X_{il} - X_{ik})$ in seasonal Kendall test to be zero if either X_l or X_i is missing or X_{il} or X_{ik} is missing, respectively. In essence, we say that since we cannot tell whether a missing value is greater or less than any actual value, it is neither. The Mann-Kendall test statistic S is unchanged, and its variance remains the same but n does not include the missing values.

6. Modification to Accommodate Censoring and Ties Values (11) (16)

When data are reported as "less than" a limit of detection, they may be arbitrarily set to some constant value which is less than the limit of detection for purposes of non-parametric trend testing. This is because the non-parametric tests are based on ranks rather than magnitudes, all censored values are viewed as showing the same rank, and this rank is less than the rank of any non-censored value. Thus the problem of censoring reduces to a problem of dealing with ties.

When we have to correct for the presence of ties, the statistic S does not vary, but the variance (S) is calculated in the following way,

$$VAR(S) = \left[\frac{n(n-1)(2n+5) - \sum_{t} (t-1)(2t+5)}{18} \right] \qquad \textbf{(16)}$$

where t is the extent of any given tie (number of X's involved in a given tie) and Σ_t denotes the summation over all ties. (For example, if there were four ties of two and one tie of three, then Σ_t t (t-1) (2t + 5) = 4 * 18 + 1 * 66 = 138).

IV. PROCEDURE

A. Preliminary Analysis of the Surface Water Quality Data

The obvious indicator of trends as a function of time is the graphical representation of the data under study. To determine if trends were visually detectable in the water quality data being analyzed, sampling stations were chosen at random and each of the values reported for each parameter were plotted versus time. Figure 3 shows a typical example of the results obtained. As can be seen, graphical representation is useless for the detection of trends, since too many points are crowded together in a small space, and their distribution appears to be totally random. To improve this, the annual average value for each parameter at each sampling station was calculated and plotted as a function of time expressed in years. Figure 4 illustrates that the visual analysis and inspection improve somewhat, but that opposite trends can arise from one year to the other, thus weakening and confusing the interpretation of the results.

An examination of the data to determine trends was also done using the linear regression method, although this was considered to be less reliable than the non-parametric methods because it assumes that there exists a straight line correlation of the data, which is not valid for water quality data. Correlation coefficients ranging from 0.751 to 0.006 were obtained (Figure 5).

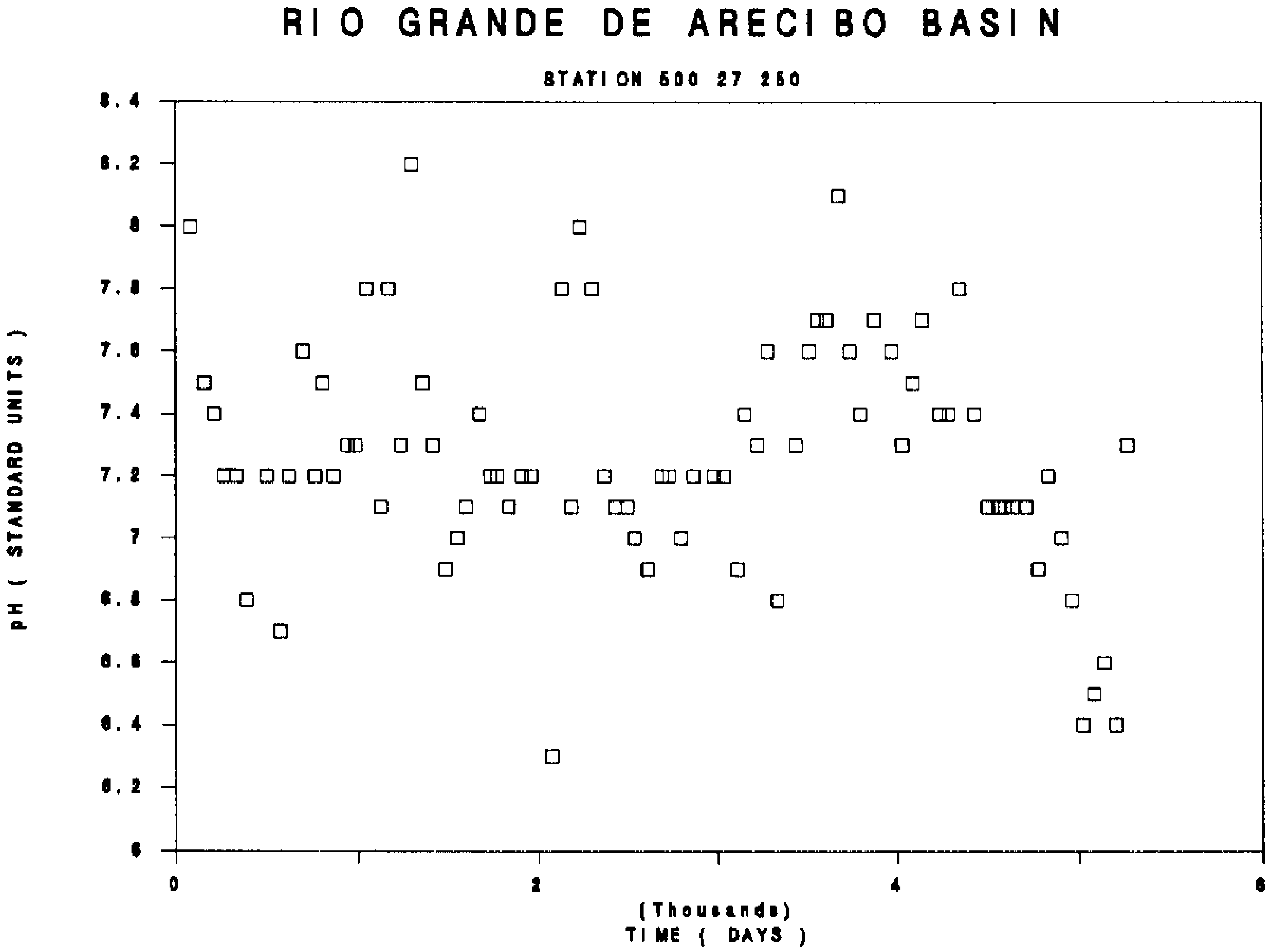

FIGURE 3: Graph of Monthly pH Values vs. Time for Station 500 27 250 at Rio Grande de Arecibo Basin

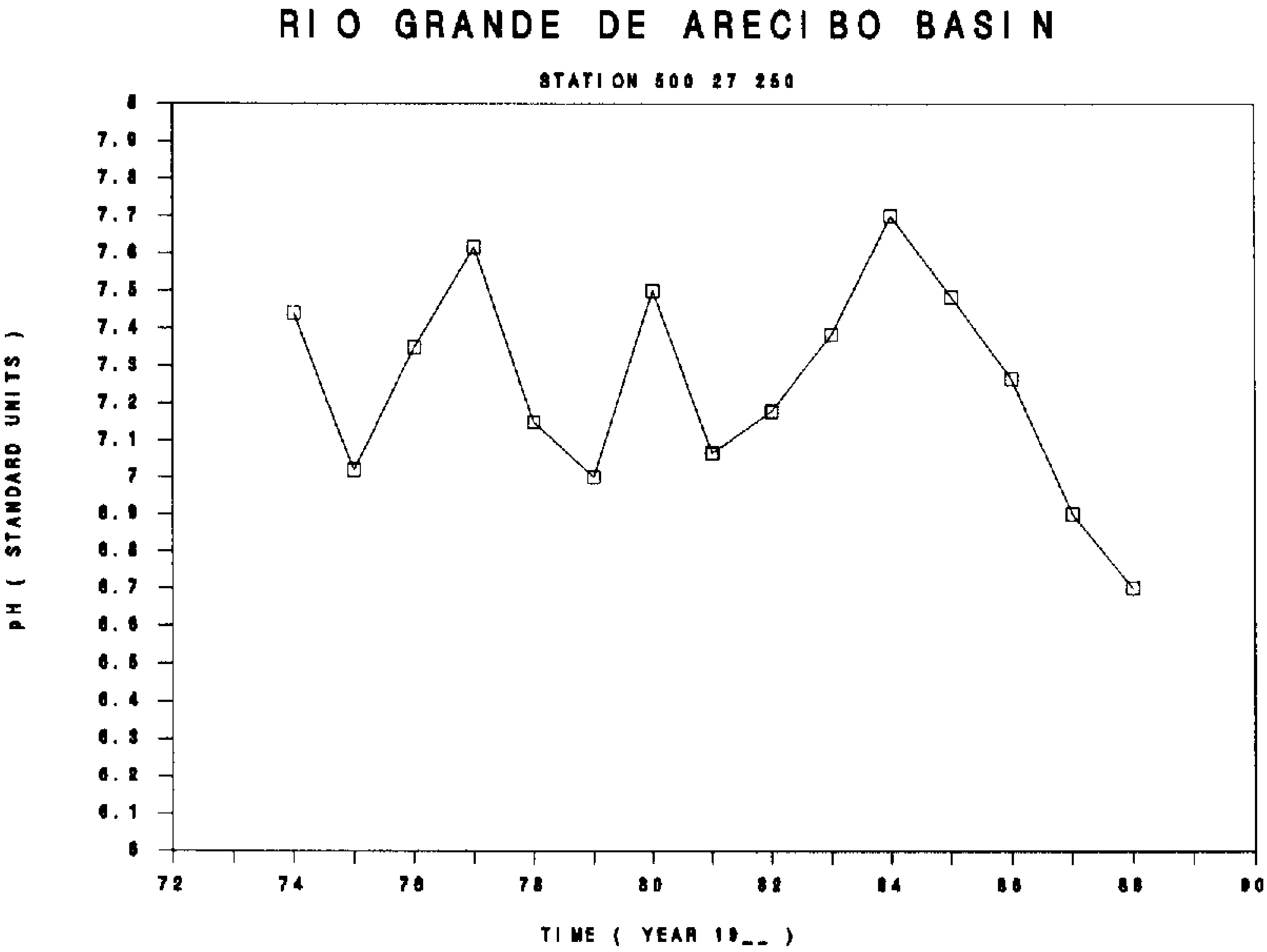

FIGURE 4: Graph of Annual Average pH Values vs. Time for Station 500 27 250 at Rio Grande de Arecibo Basin

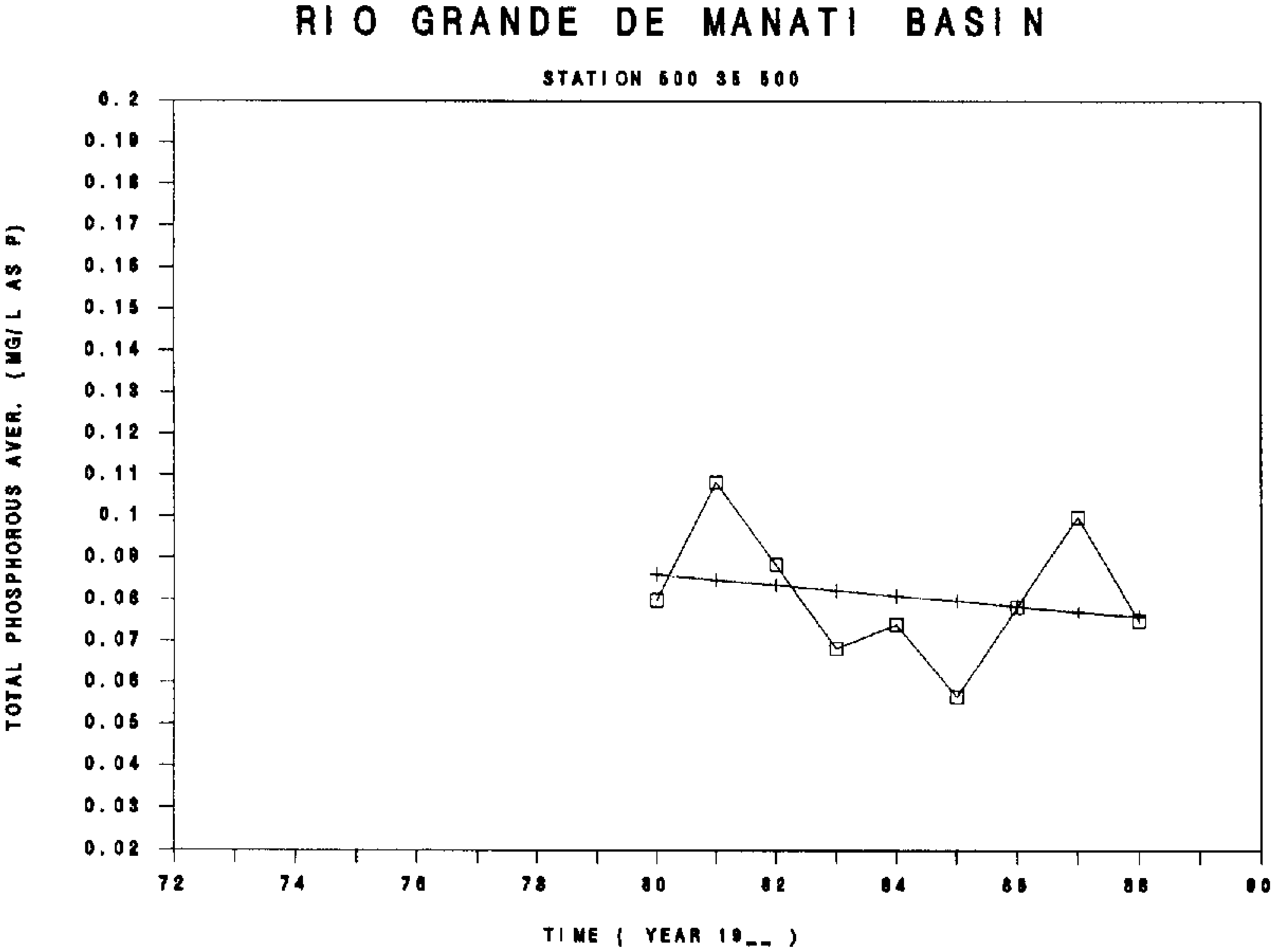

FIGURE 5: Graph of Lineal Regression of Annual Average of Total Phosphorous Concentration Values vs. Time for Station 500 35 500 at Rio Grande de Manatí Basin

These results are compared later on in this report with the values obtained using the non-parametric techniques.

B. Selection of Time Frequency Base

The surface water quality data reported by the USGS are taken about six times during the year. From one year to the next sampling may be done in different months. This means that when the monthly data are analyzed, as much as 40% may be missing values. Because the monthly data include more values, it also includes a large number of ties and censoring values. If the average annual values of the data are used, the number of ties and of censoring values is diminished and seasonal variations do not apply. For these reasons the use of the annual data was considered to be more adequate than the use of the monthly data in this study.

For data in which seasonality is not involved, the Mann-Kendall method is best suited to determine upward or downward trends, and the Sen's test can be used to determine the magnitude of the trends. These tests were done in this work with the help of the computer program which is shown in Appendix C. The results obtained show that there were upward or downward trends in 13% of the total number of data sets studied, where a data set is all the data for one specific parameter at a given station.

C. Statistical Analysis to Determine Seasonality

As stated on page 16, the Kruskal-Wallis statistical method may be used to determine seasonal behavior of a given set of data. The data corresponding to each parameter of each station was analyzed for seasonality with the help of the Statgraphics statistical package. It was found that 16% of the data sets were positive for seasonal behavior. Based on this, it was concluded that the trends present in the annual average data were being masked by the seasonal behavior influence. Those parameters that had been excluded because of seasonal behavior on the data were then analyzed for trends.

D. Statistical Analysis of the Three-Month Average Data

At this point, it was considered necessary to examine the data for seasonal behavior. The utilization of the monthly data with all its missing values was not considered appropriate. To go about this, the monthly precipitation data for Puerto Rico which is shown in Table 4 were divided into four seasons of three months each. Table 5 shows the seasons thus established for the study. The Seasonal Kendall test for trends and the Seasonal Kendall Slope Estimator for magnitudes were applied to the average values obtained for each season selected. These tests were done with the help of the computer program shown in Appendix C. A description of the tests is presented on page 21. By using these tests in which the seasonal behavior is taken into account, trends were found in

23% of the data sets analyzed. Although the validity of this study is not based upon the number of parameters that show a trend, we should note that to this point a 33% of the total data sets analyzed showed a trend.

TABLE 4:+ Islandwide Monthly Precipitation and Annual Averages for 1987 Water Year and the 30-Year Reference Period, 1951-80.

Months	Water Year 1987 (inches)	30-Year Normal (inches)
October	7.12	7.74
November	8.29	5.95
December	2.95	4.32
January	2.99	3.08
February	3.86	2.35
March	3.99	2.63
April	5.83	4.63
May	9.61	6.48
June	8.78	5.58
July	3.59	5.48
August	3.88	7.28
September	5.18	7.78
TOTAL	66.07	63.30

+ Reference (23)

TABLE 5: Seasons Selected

SEASONS	MONTHS INCLUDED
1	August, September, October
2	November, December, January
3	February, March, April
4	May, June, July

E. Statistical Analysis of the Monthly Data

Although the monthly data had been previously ignored because of their many missing values, it was decided at this point to analyze them within the context of the seasonal behavior effects. It should be noted that the larger the number of periods into which the year is subdivided, the easier it is for the seasonal behavior to mask the trends.

When a test is applied to the monthly data it should be kept in mind that the real value of the missing value could be greater or smaller than the previous or the following one and because of this a false trend result might be obtained. This procedure is acceptable in cases where there are no missing values.

The statistical analysis of the monthly data was done utilizing the same computer program and technique as with the analysis of the three-month data.

V. RESULTS AND DISCUSSION

A. Results of Analysis of Seasonal Test

As indicated in the procedure, it was necessary to determine if the water quality data obtained by the USGS showed seasonal behavior of any kind. For this purpose the Kruskal-Wallis test was applied to each set of data. It was found that 16% of the data sets analyzed showed a seasonal behavior. These results suggest that seasonality is a common phenomenon in surface water quality data for Puerto Rico. The data sets that present seasonal behavior are identified with the letter S in Table 6. The stations and parameters that did not show seasonal behavior were excluded from this table.

It was observed that streamflow is a parameter that shows the highest seasonal behavior among all the parameters considered. This happens because this parameter is highly dependent on the rainy seasons prevalent in Puerto Rico, during which increases in streamflow occur directly. Nevertheless, the fact that a seasonal behavior is observed does not mean that there is a trend to increase or to decrease as a long range function of time.

The fact that 16% of the data sets present a seasonal behavior is enough to take seasonality into consideration, even including the data sets and stations that did not show this pattern. It should be considered that there were many missing values in the monthly reports and this could mask the

TABLE 6: Results of Analysis of Seasonal Test

Parameter\Station	1	2	3	4	5	6	7	8	10	11	12	14	17	18	21	22
Streamflow (cfs)	S#	S	S	S				S		S	S	S				S
pH (standard units)			S							S						
Dissolved Oxygen (mg/l)	S					S								S		S
Dissolved Oxygen (% saturation)																
Fecal Coliform (cols./100ml)	S	S	S			S							S			
Dissol. Potassium (mg/l as K)	S	S													S	
Total Nitrogen (mg/l as NO_3)		S			S	S	S		S							
Total Phosphorous (mg/l as P)		S				S									S	

\# S means the presence of seasonality in this data set seasonal trend. Also, it may be that the trend cannot be observed in a month-to-month lapse but it may be obvious from season-to-season over the course of a year. This is the reason why the Seasonal Kendall tests were applied to the monthly and the tri-monthly values without considering if these parameters had shown seasonal behavior when the Kruskal-Wallis test was applied to them.

B. Results of Analysis of Long-term Averages

To have a clear idea of the changes in the water quality parameters that occur as the rivers approach the ocean, the average value of each parameter was determined using all the data available for each station. Table 7 presents the results obtained.

C. Results of Analysis of Tests for Trends

As described in the procedure, four different tests were applied to try to detect the presence and magnitude of trends in the data. The results of these tests are shown on Table 8. The values presented in this table indicate the magnitude of the annual trends even though some of them were determined using tri-monthly and monthly values. The negative sign in some of the values indicate a downward trend while a positive one shows an upward trend. The letters indicate which set of values was analyzed; AA stands for annual data values, TA for the three-month period, and MA for the one-

TABLE 7: Results of Long-term Averages

Parameter\Station	1	2	3	4	5	6	7	8	9	10	11
Streamflow (cfs)	53.8	224	258	40.1	303	375	27.4	65.7	24.5	56.3	47.7
pH (standard units)	7.87	7.76	7.66	7.13	7.25	7.70	7.93	7.34	7.32	7.42	7.22
Dissolved Oxygen (mg/l)	8.93	8.68	8.32	8.21	5.82	8.40	8.69	5.66	3.86	8.61	7.54
Dissolved Oxygen (% saturation)	112	107	103	100	64	104	105	67	58	109	102
Fecal Coliform (cols./100ml)	1565	2844	3614	7964	1844	2388	3245	109	1524449	2054	103271
Dissol. Potassium (mg/l as K)	1.65	1.45	1.72	2.15	2.03	1.65	2.61	3.31	4.95	1.17	4.98
Total Nitrogen (mg/l as NO_3)	6.10	5.23	6.40	7.38	4.10	4.36	4.63	10.2	30.9	2.99	5.60
Total Phosphorous (mg/l as P)	.058	.075	.100	.159	.044	.057	.076	.455	1.52	.070	.228

TABLE 7: Continued

Parameter/Station	12	13	14	15	16	17	18	19	20	21	22
Streamflow (cfs)	71.4	123	268	41	79	193	349	133	162	218	32
pH (standard units)	7.28	6.90	7.64	7.67	8.08	7.91	7.28	7.87	7.90	7.54	8.22
Dissolved Oxygen (mg/l)	4.89	2.30	7.50	7.49	8.45	8.36	7.59	8.98	8.59	6.29	8.87
Dissolved Oxygen (% saturation)	62.2	24	97	96	105	104	96	119	107	86	109
Fecal Coliform (cols./100ml)	24926	584	28252	6717	4546	3372	13042	4752	9908	5497	3060
Dissol. Potassium (mg/l as K)	4.41	2.90	2.03	2.88	2.18	2.74	2.07	2.23	2.67	2.46	1.25
Total Nitrogen (mg/l as NO_3)	10.8	5.30	6.06	5.92	6.05	4.64	5.80	8.79	8.23	4.43	4.09
Total Phosphorous (mg/l as P)	.618	.200	.240	.260	.096	.081	.177	.480	.340	.170	.055

TABLE 8: Magnitudes and Directions of Trends[a] Detected with Non-parametric Tests

Parameter\Station	1	2	3	5	6	7	8	9	10	11
Streamflow (cfs)				48[*] MA		4.0 TA[+]	9.8 TA	3.92 TA		
pH (standard units)	.023 AA[#]	.052 TA	-.212 TA		.120 MA			.013 AA	.04 TA	.100 TA
Dissolved Oxygen (mg/l)				-.15 AA	.132 TA	-.20 TA		.554 TA		
Dissolved Oxygen (% saturation)							8.4 TA			-41.4 MA
Fecal Coliform (cols./100ml)		-400 MA	-654 MA	156 MA	-690 MA	-373 MA		-206833 AA	-383 AA	-16720 AA
Dissol. Potassium (mg/l as K)	.039 AA	.032 AA			.100 TA	-.06 AA			.048 TA	-.20 TA
Total Nitrogen (mg/l as NO_3)	.214 AA		.800 TA	.488 TA	.484 TA	-.008 TA		1.708 AA	.332 TA	
Total Phosphorous (mg/l as P)		.004 TA	-.024 MA						-.003 AA	-.003 AA

TABLE 8: Continued

Parameter/Station	12	13	14	16	17	18	19	20	21
Streamflow (cfs)				53.33 MA			-2.43 TA		31.0 MA
pH (standard units)			0.0 MA	.252 MA			.016 AA		-.04 TA
Dissolved Oxygen (mg/l)								-.454 TA	.996 MA
Dissolved Oxygen (% saturation)	-8.64 TA	14.0 TA	-7.20 TA			-24.0 MA			
Fecal Coliform (cols./100ml)		-84.0 AA	-23500 MA	708.0 AA		-4517 TA			-558.0 AA
Dissol. Potassium (mg/l as K)		-4.88 TA	.088 TA	.091 AA	.069 AA	.120 TA	.158 TA	.050 AA	.080 TA
Total Nitrogen (mg/l as NO_3)	-3.95 MA					.484 TA	.064 TA		.306 AA
Total Phosphorous (mg/l as P)			-.032 TA		-.024 MA	-.012 TA			-.006 AA

a The numbers show the magnitudes of the annual trends even though some of them were determined using tri-monthly and monthly values. The signs indicate the direction of the trends.

* MA = The trends were detected using the annual averages of data sets.

\+ TA = The trends were detected using the tri-monthly averages of data sets.

\# AA = The trends were detected using the monthly values of data sets.

month period. The table does not include any of the stations in which no trends were observed. The type of data grouping used to measure the direction of the trends does not influence the determination of the direction, since for any parameter that shows a positive or negative trend using the annual data values there will be a large probability that it will show the same trend in direction when the three-month and the one-month data values are used. However, the grouping of the data affect the magnitude of the trends. For example, most of the trends observed using the monthly data prove to be absurd. On the other hand, the magnitudes determined using yearly values agree well with the data set used. The reason for this, is the large number of missing values in the monthly data reported, which is not the case in the yearly values data.

D. Discussion of the Results

For the purpose of the discussion of the results each river studied will be considered at a time. The long-term averages will be analyzed first, and the trends detected will then be analyzed.

1. Rio Grande de Añasco Basin

Three sampling stations on this river were analyzed. Figure 6 shows the location of these stations. The results obtained in the analysis of long-term averages are presented in Table 7. An increase in streamflow is seen to occur as the

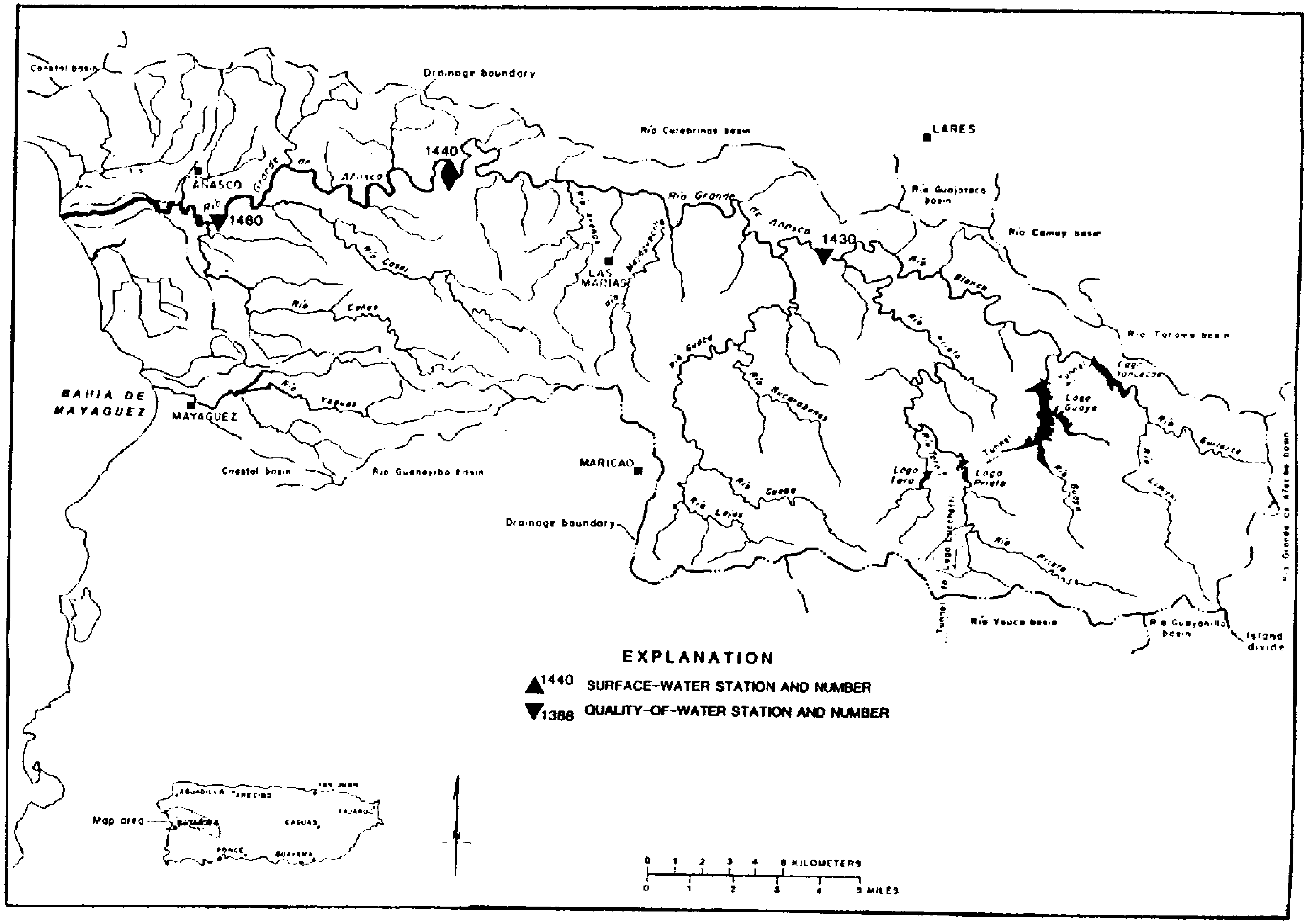

FIGURE 6: Rio Yagüez and Rio Grande de Añasco Basins

stream approaches the sea, while the pH decreases about a tenth of a unit from one station to the next. The dissolved oxygen and the fecal coliforms increase in the downstream direction, while the concentration of potassium, nitrogen, and phosphorous undergo different changes among each other except from the second to the third station in which the three increase in value. The increase, and that in the fecal coliforms, can be attributed to increase in agricultural and human activities respectively.

The trends found for this river are presented in Table 8, which also shows that as the river gets closer to the ocean the water quality changes. The results demonstrate that trends are dependent upon the sampling station location, as in the case of pH and total phosphorous concentration.

The results obtained show that the water quality in this stream could be satisfactory if the fecal coliform concentration is reduced and that there are no significant trends to negative changes. This factor, plus the fact that the volumetric flow of the stream is large, suggest that the river has high potential for additional water supply use.

2. Rio Grande de Arecibo Basin

Three stations of this river were included in the study, with locations as presented in Figure 7. The streamflow increases as the river nears the ocean, the large increase from station 4 to 5 occurring because from one to the other

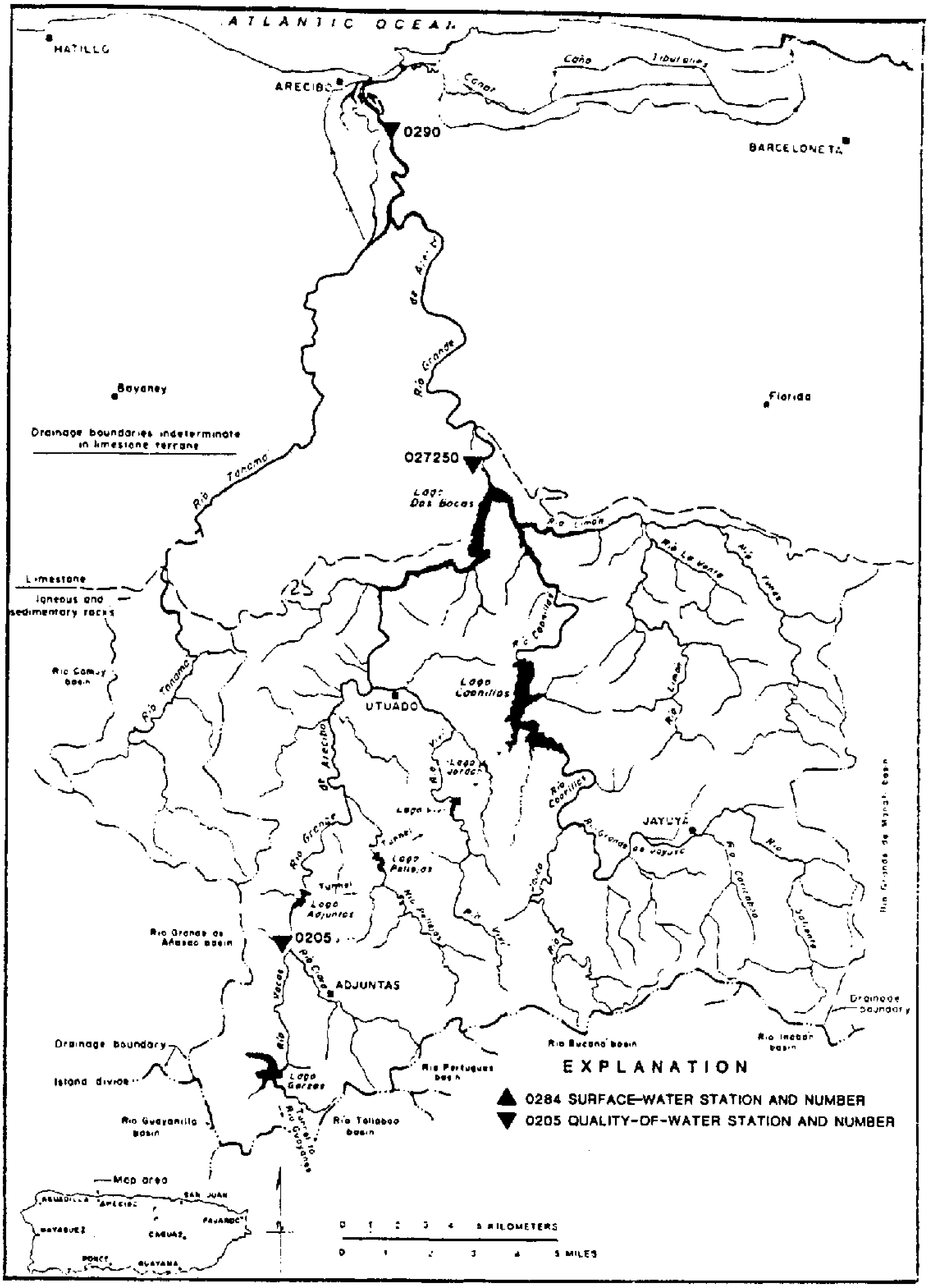

FIGURE 7: Rio Grande de Arecibo Basin

this river joins two other rivers to form Dos Bocas Reservoir. The pH increases significantly from station 4 to 6. Dissolved oxygen and fecal coliforms decrease from the fourth to the fifth station, possibly due to the retention period in the reservoir or simply to a dilution process.

The results shown in Table 8 show that station 4 does not present trends in any of the parameters included in the study, while station 6 had the largest number of parameters in which trends could be observed. This fact demonstrates that the water quality parameters of this river are not being affected upstream, but as it gets to the ocean some trends are observed although they are not necessarily negative. This was to be expected, since in its upstream zone these are minimal impacts from man, industry, or agriculture. The trends observed show that the changes in water quality are dependent upon the location of the sampling station. This was the case for the trends in dissolved oxygen and in fecal coliforms.

3. Rio Bayamón Basin

Two sampling stations were considered in this study, as shown on Figure 8. From one to the other there is a significant increase in streamflow, from the large number of tributaries reaching the stream. From station 7 to 8 there is a considerable decrease in pH, in dissolved oxygen, and fecal coliforms. On the other hand, the concentrations of dissolved potassium, total nitrogen, and total phosphorous increase in

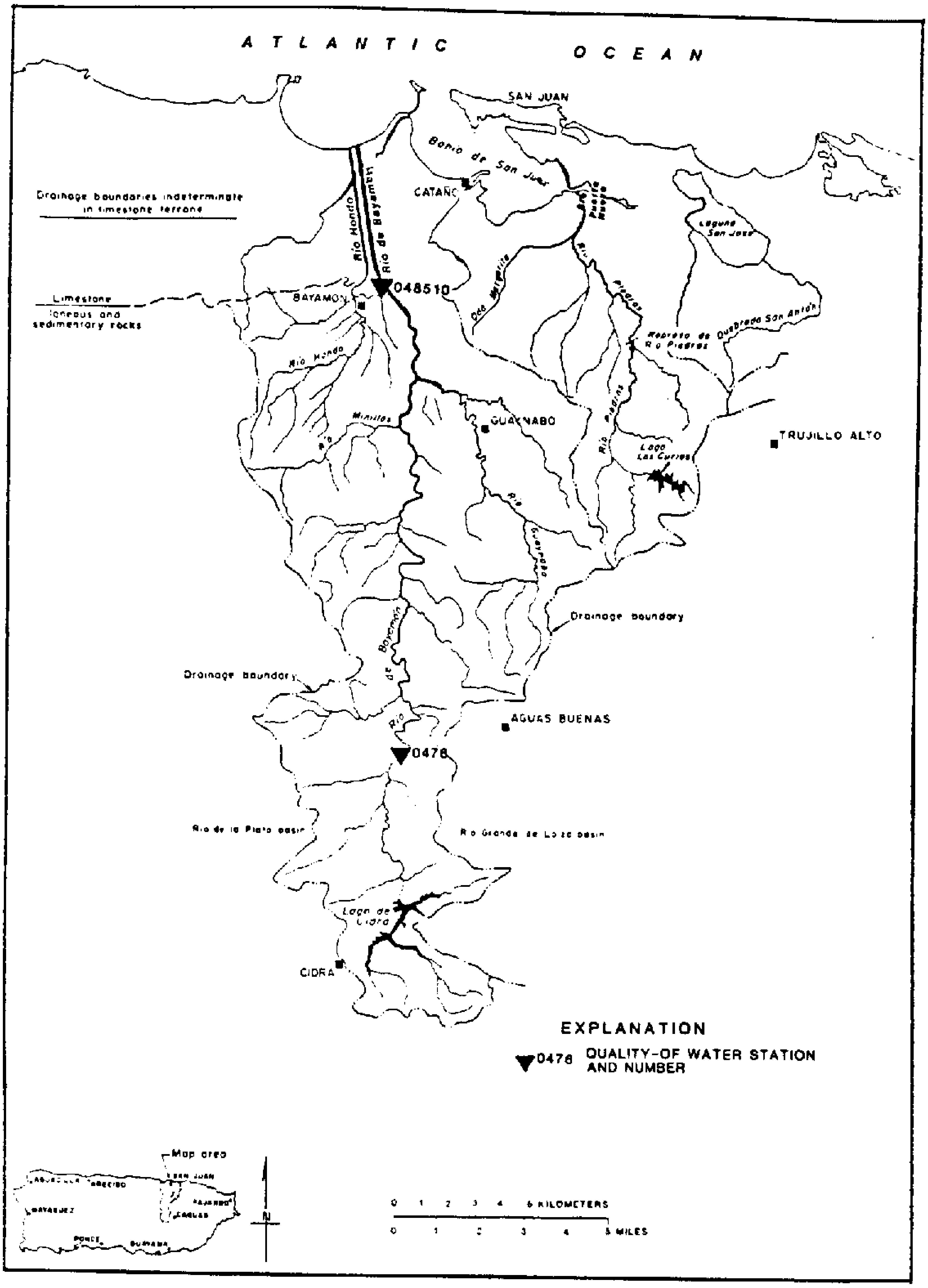

FIGURE 8: Rio Hondo to Rio Puerto Nuevo Basins

the downstream direction. In general, this river has an overall increase in quality as it flows downstream, which differs from the case of most of the other rivers.

The trends detected show that this river is also unique in that the upstream station shows a larger trend to changes in the parameters studied than the downstream one.

4. Rio Fajardo Basin

The location of the sampling stations considered for this river are shown in Figure 9, their numbers in the tables being stations 10 and 11. For this river the streamflow in the downstream station is smaller than in the upstream one. The dissolved oxygen contents and the pH decrease as the river approaches the ocean. From station 10 to station 11 the fecal coliform concentration increase by a factor of about 50, and, the concentrations of dissolved potassium, total nitrogen, and total phosphorous also increase significantly. This can be explained on the basis of the increased population and human activities as the river nears the ocean.

Six out of the seven water quality parameters studied show a change trends in at least one of the stations considered. The concentration of fecal coliforms in both stations shows a trend to decrease over time, perhaps due to improved sanitary practices in the region.

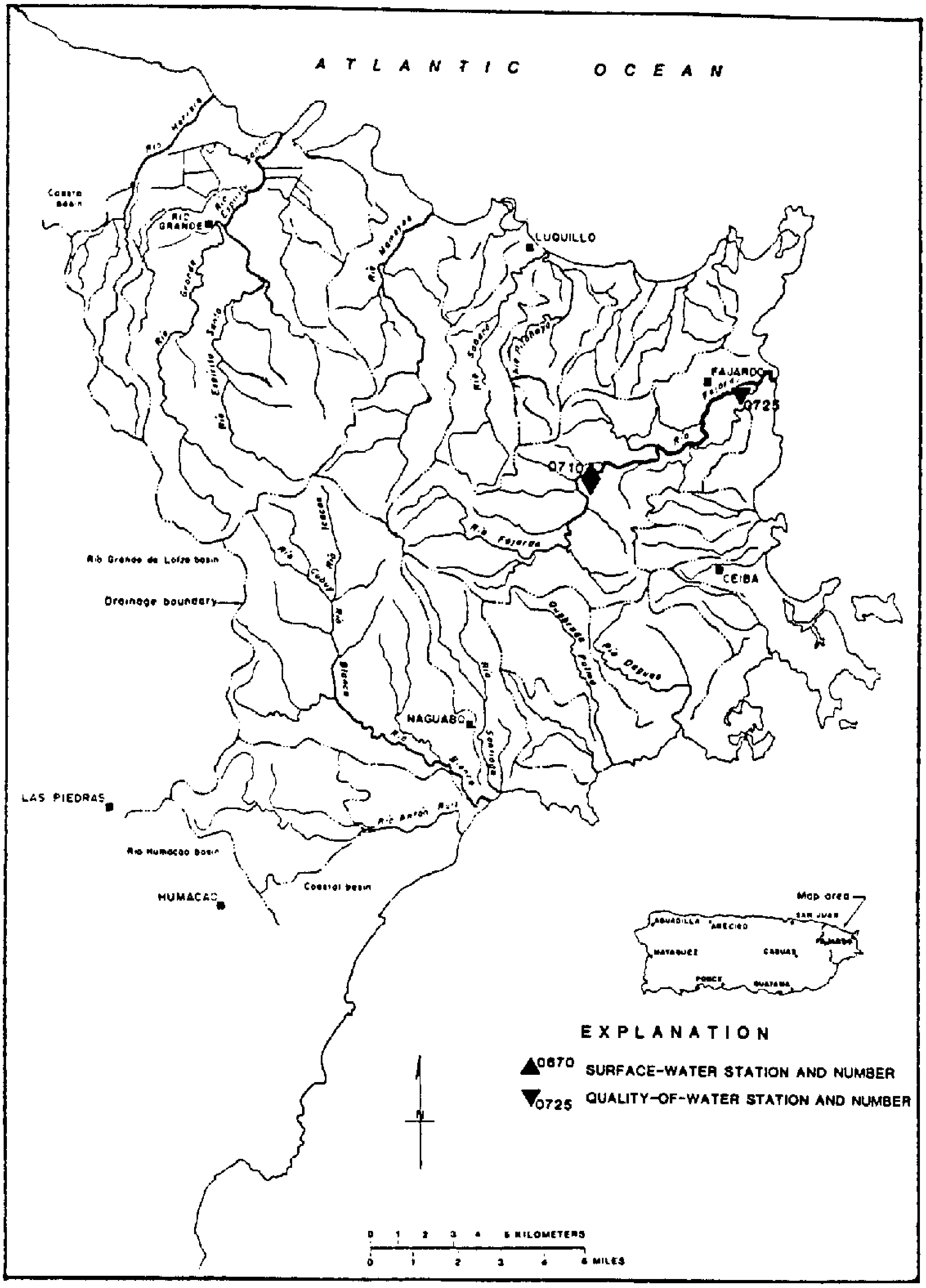

FIGURE 9: Northeastern River Basins--Rio Herrera to Rio Antón Ruiz Basins

5. Rio Grande de Loiza, Rio Cagüitas, and Rio Gurabo

The locations of the sampling stations studied in these rivers are shown on Figure 10, all forming part of the Rio Grande de Loiza Basin. Stations 9, 12, and 14 are located at points just before the Cagüitas, Gurabo, and Grande de Loiza rivers flow into the Loiza Lake, respectively. Station 9 evidences an adequate pH, a low concentration of dissolved oxygen, extremely high concentration of fecal coliforms, and normal values of dissolved potassium and total phosphorous. This pattern is also observed at station 12. At station 14 the concentration of fecal coliforms is on the high side, while at station 13 and 15 the values of all the parameters reflect an acceptable water quality. There is a large decrease in flow from station 13 to station 15 because water is drawn off for public supply use immediately after station 13. Between these two stations there are increases in the concentration of dissolved oxygen and fecal coliforms, while those of dissolved potassium, total nitrogen, and total phosphorous remain practically unchanged.

The trends observed in the Cagüitas River show that the concentration of fecal coliforms decreases with time, even though their absolute values are high because the river receives the affluent from Caguas sewage treatment plant. At stations 12 and 14 there are also treated sewage discharges from the towns of Las Piedras-Juncos and San Lorenzo, respectively. The trends detected at station 13 do not impair

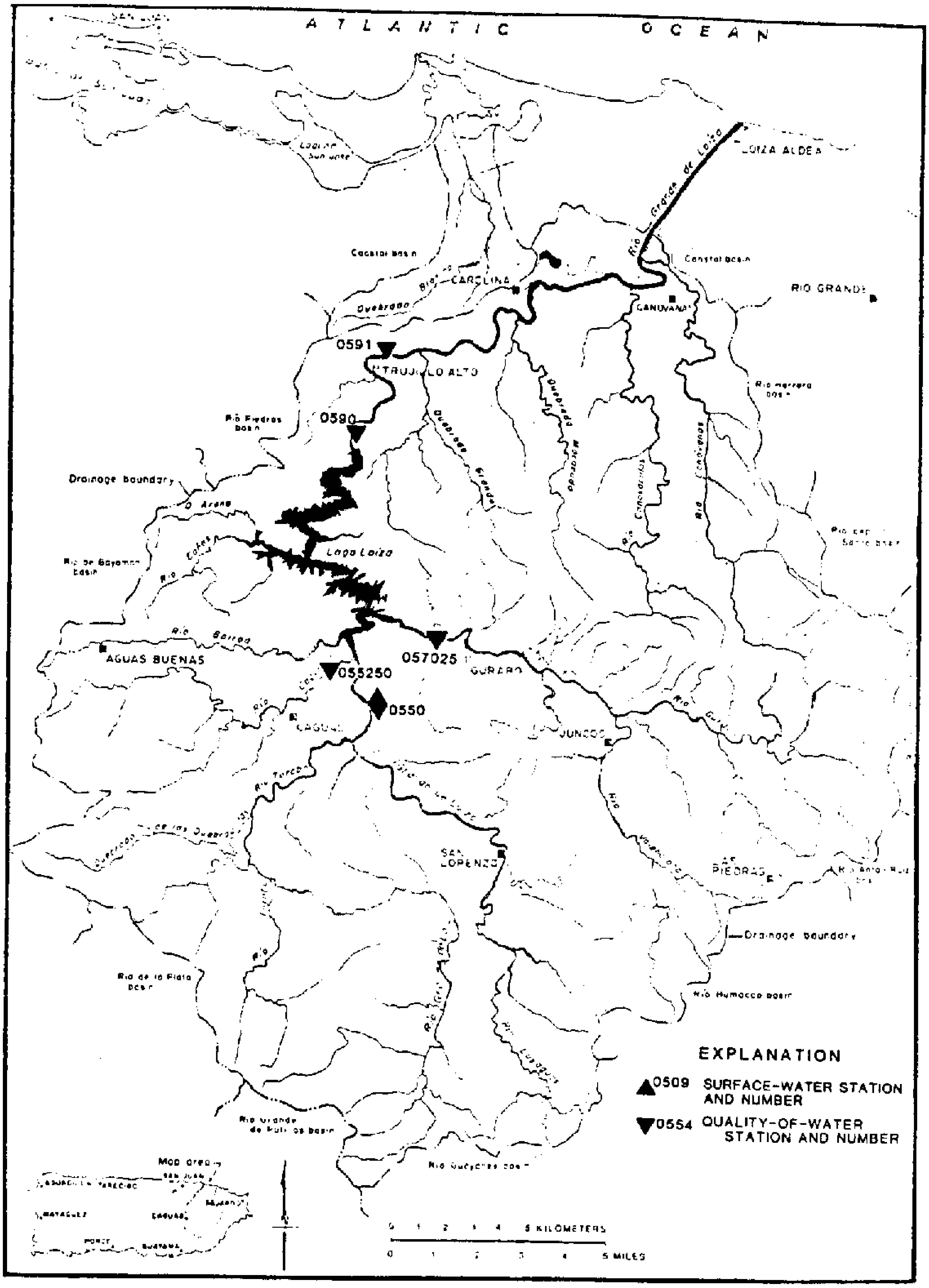

FIGURE 10: Rio Grande de Loiza Basin

the quality of the water, while at station 15 no trend of any kind was found for any of the parameters considered. It is interesting to note that although the three rivers present trends in some of the parameters tested before they flow into the main stream these trends disappear once they reach it.

6. Rio Grande de Manatí Basin

Three sampling stations were analyzed in the basin, with locations as shown in Figure 11. Table 7 shows once more the considerable increase in streamflow that occurs in the downstream direction all the way over into the sea. The dissolved oxygen and the pH show significant decrease from station 16 to station 18. Fecal coliform concentration decrease from station 16 to station 17, but they increase considerably from number 17 to number 18 due to the discharge of treated sewage from the towns of Ciales and Manatí. The changes in dissolved potassium, total nitrogen, and total phosphorous between these stations are not significant in magnitude.

The results show that the station closest to the ocean, (18), is the one that shows trends in a larger number of parameters. The effect of station location on the trend was observed for the fecal coliforms parameter. It was interesting to find that the concentration of dissolved potassium tends to increase at the sampling stations down river.

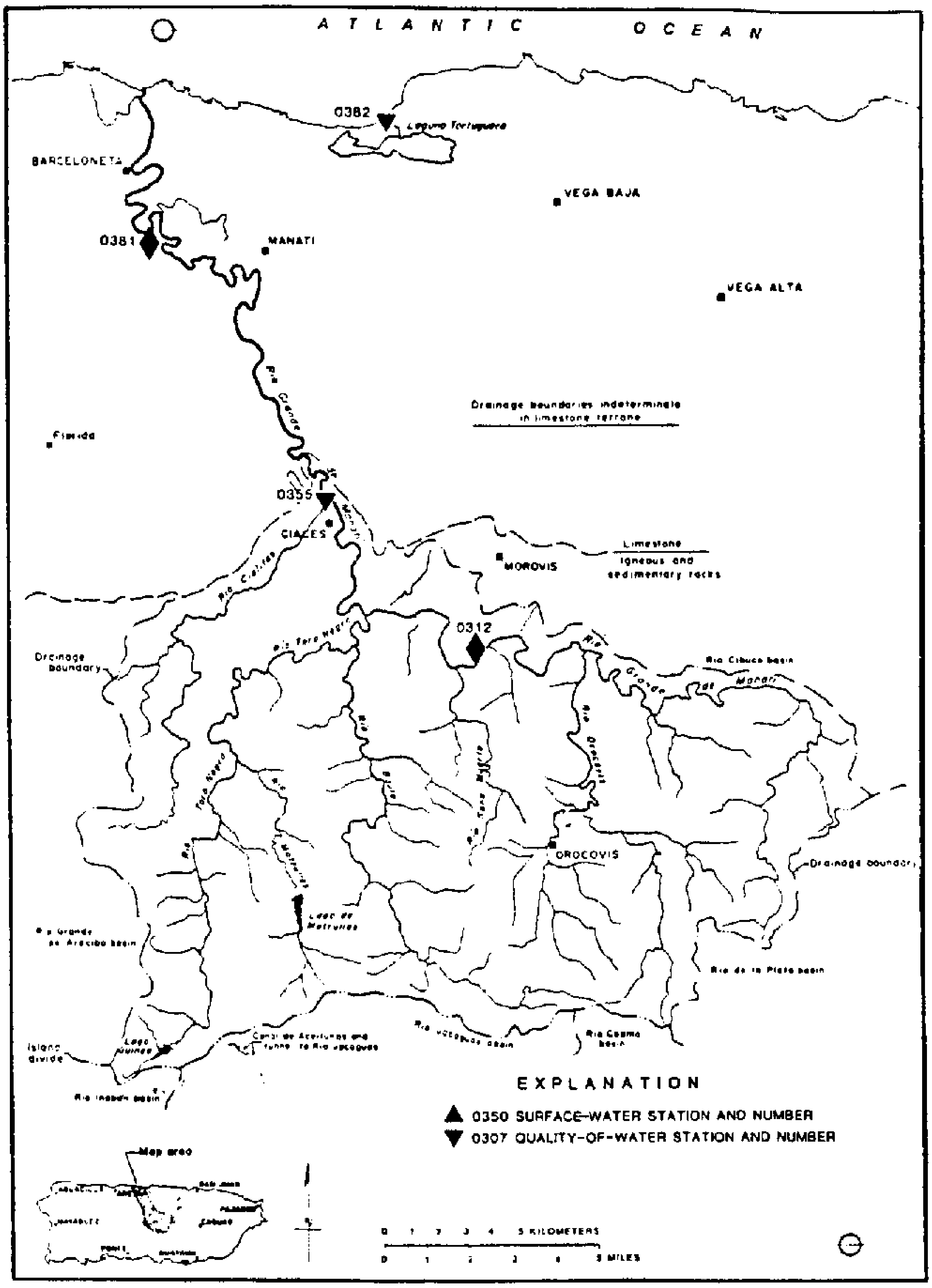

FIGURE 11: Rio Grande de Manatí Basin

7. Rio La Plata Basin

The three sampling stations analyzed in this basin were located as shown in Figure 12. The long-term averages show an increase in streamflow in the downstream direction, as would be expected from the large number of tributaries entering it. They also show minimal changes in pH from station to station, considerable decreases in dissolved oxygen from station 19 to station 21, and increases in fecal coliforms concentration from station 19 to station 20, followed by decreases from station 20 to station 21. It is to be noted that the concentration of fecal coliforms decreases as streams flow through reservoirs, as observed in this basin.

Station 21 presents trends for most of the parameters considered. The trends follow the pattern that the closer to the ocean the more the water quality is affected. The trends observed in the total nitrogen, dissolved potassium, and pH for station 19 may be due to the presence of fertilizers used in surrounding areas. Upstream of this station there are numerous effluent streams which bring their loads of fertilizer contamination into the river. The trend found in the streamflow is attributed mostly to precipitation. However, the streamflow trend found for station 21 is dependent on the flow control of the dam. It was found that the streamflow, pH, and dissolved oxygen trends are highly dependent on the location of the sampling station.

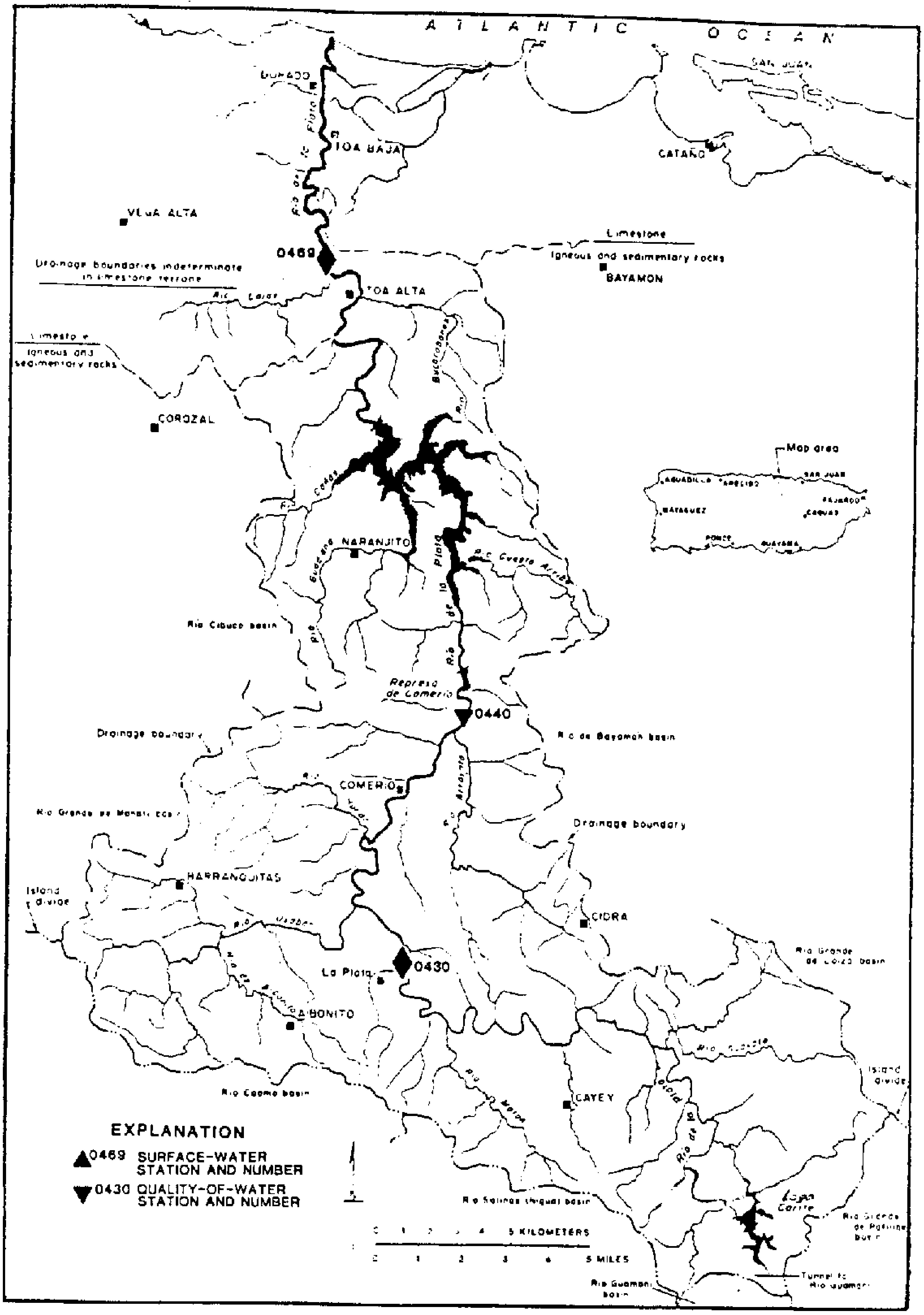

FIGURE 12: Rio de La Plata Basin

8. Rio Rosario Basin

Only one sampling station was included in this study as shown in Figure 13. All the water quality parameters considered in this work show acceptable long-term average values, as observed in Table 7. No trends were detected in the parameters analyzed. This may be due to the fact that the river flows through an area where there are no industrial sites and with very little agricultural activity. Because of this, the river shows a large potential for use as a public water supply to serve rural communities in the region. It also appears that station 22 could be eliminated or relocated, within the scope of the water quality network operated by the U. S. Geological Survey.

E. Comparison Between the Non-parametric and the Lineal Regression Analyses.

A comparison of the trends found using the non-parametric and the lineal regression analyses was undertaken. Table 9 shows the results obtained using the lineal regression analysis, while the results obtained using the non-parametric methods are presented in Table 8. Comparing these tables we can see that with the exception of five results, both tables show trends toward the same direction. The five exceptions are the concentration of total nitrogen at station 3 and 10, the concentration of total phosphorous in station 2 and 21, and the value of pH at station 14. These five exceptions can

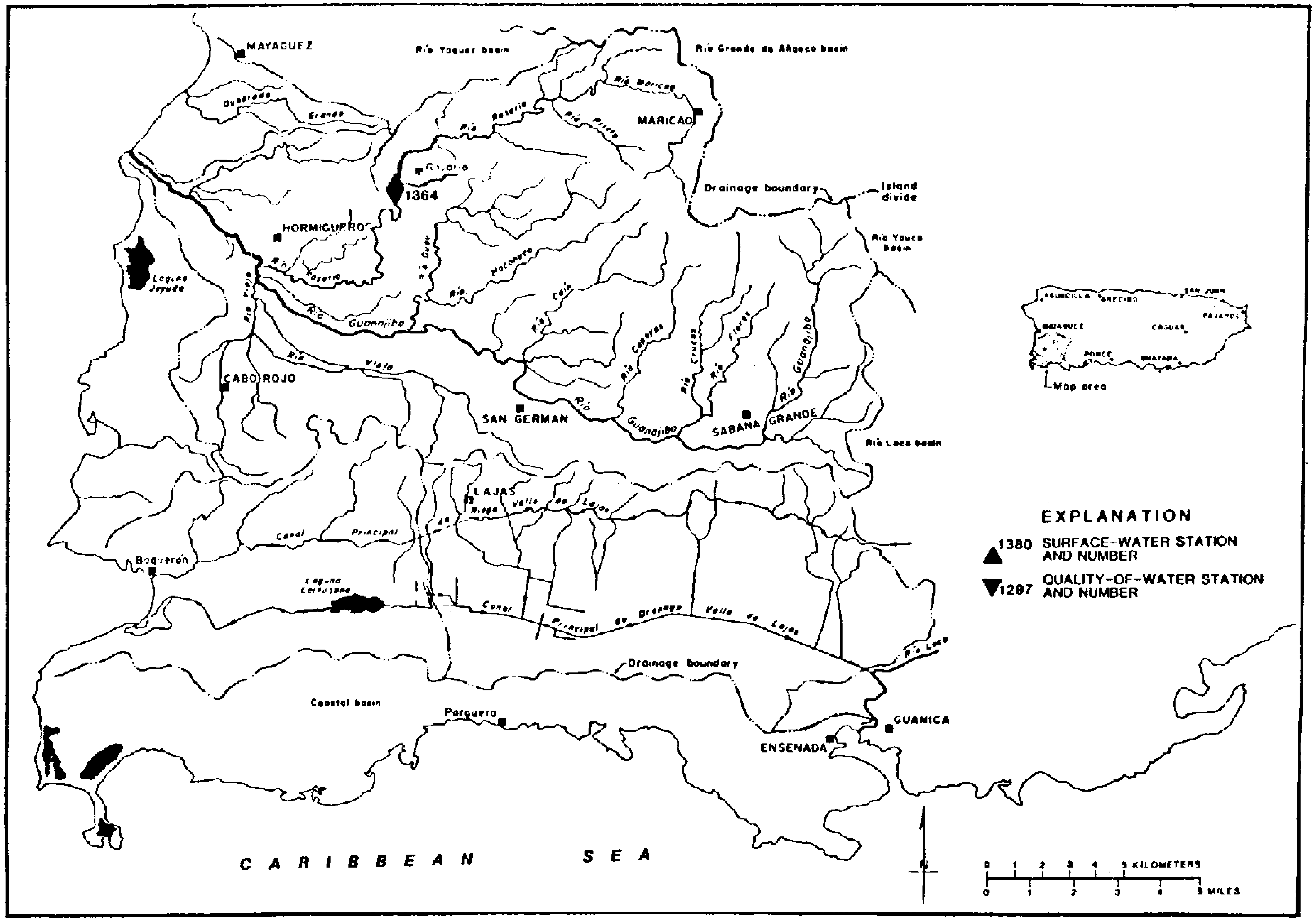

FIGURE 13: Rio Guanajibo Basin

be attributed to the great influence of the extreme values in the lineal regression calculations. This does not happen when the non-parametric method is used. Based on this we can conclude that when the calculations do not involve extreme values, the method of lineal regression may be an expedient alternative to determine increasing or decreasing trends and this could save considerable time and effort. For instance, if a particular parameter is such that a decrease in its concentration indicates an increase in water quality, then finding a negative slope for that parameter through lineal regression analysis tells us that water quality with respect to that parameter is increasing over time and we need not analyze that parameter at that site any further.

Also, the magnitudes of the trends found using the lineal regression analysis compare favorably with those found by the non-parametric methods for those parameters studied using annual data. However, this was not the case for those parameters in which tri-monthly or monthly data were analyzed. The fact that similar results are obtained with lineal regression and non-parametric methods for the data expressed as annual values but not for tri-monthly or monthly ones is due to the high effect which extreme values have on the linear regressions, and to the influence of missing values, which appear in larger numbers as the time period used decreases.

TABLE 9: Results of Linear Regression Analysis for Trends

Parameter\Station	1	2	3	5	6	7	8	9	10	11
Streamflow (cfs)				40.27		1.118	3.33#	.280		
pH (standard units)		.013	-.051		.007			.012	.010	.023
Dissolved Oxygen (mg/l)				.438	.042	-.039		.159		
Dissolved Oxygen (% saturation)							1.838			-2.54
Fecal Coliform (cols./100ml)		-209	-496	223	-95.0	-76.0		-407488	-497	-34141
Dissol. Potassium (mg/l as K)		.032			.014	-.061			.011	-1.178
Total Nitrogen (mg/l as NO_3)	.267		-.062	.122	.122	-.024		2.247	-.064	
Total Phosphorous (mg/l as P)		-.002	-.009						-.012	-.044

TABLE 9: Continued

Parameter\Station	12	13	14	16	17	18	19	20	21
Streamflow (cfs)				6.94			-4.57		7.284
pH (standard units)			-.002	.012			.014		-.007
Dissolved Oxygen (mg/l)								-.116	.061
Dissolved Oxygen (% saturation)	-2.53	2.87	-1.75			-.827			
Fecal Coliform (cols./100ml)		-121	-1811	2533		-2349			-565
Dissol. Potassium (mg/l as K)		-.129	.010	.046	.462	.024	.008	.048	.019
Total Nitrogen (mg/l as NO_3)	-.245					.201	.205		.364
Total Phosphorous (mg/l as P)			-.003		-.001	-.002			.006

The numbers indicate the magnitude of the annual trends and the signs the direction of these trends.

F. Compliance with Water Quality Criteria

It is pertinent to use the results of this study to determine if the water quality standards established by Puerto Rico's Environmental Quality Board are met. Similarly, the form in which the trends detected in the parameters which were considered could point to the risk of violating the standards should also analyzed. The two issues are analyzed best through Tables 7, 8, and 10.

TABLE 10: Puerto Rico Water Quality Standard Regulation (24)

Parameters	Standard
pH	6 - 9 Standard Units
Fecal Coliform	Not Exceed 2,000 Cols./100ml
Dissolved Oxygen	Not Less Than 5.0 mg/L
Total Nitrogen	Not Exceed 10.0 mg/L
Total Phosphorous	Not Exceed 1.0 mg/L

Table 10 shows the water quality standards established by the Environmental Quality Board, while Table 7 summarizes the long-term averages for the data considered in the study. Comparing the two tables it is observed that the standards for pH and total phosphorous concentrations are met in full in all the stations except for total phosphorous in station 9. However, to prevent accelerated eutrophication, total phosphorous as P should not exceed 0.05 mg/L in any stream at the point where it enters a lake or reservoir (U.S. Environmental Protection Agency, 1976). Stations 4, 9, 12, 14, 19, and 20 all of which precede reservoirs, violate this criterion. Of these, station 9, which is located at Cagüitas River, exceeds the 0.05 mg/L by a factor of 30. For these same parameters, Table 8 shows that there are no trends which could result in predictable violations of those standards in the future.

The tables show that the standard for dissolved oxygen is not met at stations 9, 12, and 13, in each of which the concentration is less than 5.0 mg/L. At the other stations the trends in dissolved oxygen concentration do not point to future risks of failure to meet the established standards.

The standard for fecal coliforms is set by law at an MPN value of 2,000 colonies/100 ml. Except for stations 1, 5, 8, and 13, this standard is not met at any of the other sampling points. Most of the trends observed in the stations point to decreases in fecal coliform concentrations, with the exception of numbers 5 and 16 which show trends to increase.

The total nitrogen standard of 10.0 mg/L is not met by the water at stations 8, 9, and 12. At station 9 specifically not only is the long-term average three times the value of the standard, but in addition there is a trend of 1.708 mg/L to increase with time. At the other stations neither the present values nor the trends point to situations in which the total nitrogen standard could not be met.

In summary, the most common surface water quality problem in the rivers studied was the high concentration of fecal coliforms. Low concentrations of dissolved oxygen and high total nitrogen contents are also common, although to a lesser extent. Overall, there is a trend to decreases in the concentration of fecal coliforms and to increases in total nitrogen. Cagüitas River is the most polluted stream among all those considered, showing below standard values for dissolved oxygen, fecal coliforms, total nitrogen, and total phosphorous concentrations.

VI. CONCLUSIONS

The results obtained in a preliminary analysis of the water quality data indicate that the graphical representation of surface water quality data is not an adequate method to detect the presence and magnitude of trends. The actual distribution of this data cannot be followed visually, largely because seasonality is present in most cases.

In the course of this work we concluded that for any statistical analysis of surface water quality data it is necessary to take into account the presence of missing values, ties, censoring points, and seasonality. Leaving out any of these factors could lead to one of two possibilities: detection of trends in data that are not really present, or failure to detect trends in data that actually are present. When the calculations do not involve extreme values, the method of lineal regression may be an expedient alternative to determine the trend directions of the data sets, but it cannot be used to determine the magnitudes of these trends. The best method for determining the presence, direction, and magnitude of trends for surface water quality data in all moments is non-parametric analysis.

Results of non-parametric analysis for trends demonstrate that the water quality of rivers in Puerto Rico tends to change dramatically as the rivers get closer to the sea. Results also indicate that the trends toward change in the parameters studied are very dependent on the location of the stations, as well as on the location of the rivers. This is due to factors such as industrial and agricultural activities, and the increased population density in the coastal plains that the rivers traverse before entering the sea.

Results of analysis of long-term averages show that each parameter behaves in different ways for the various rivers. The only obvious case in which this did not apply is in streamflow, since all main stream increase their volumetric flow as they approach the sea. No generalization of this type is possible for any of the other parameters analyzed.

It was concluded that the most common surface water quality problem in the rivers studied was the high concentration of fecal coliforms. Low concentrations of dissolved oxygen and high total nitrogen contents are also common, although to a lesser extent. Cagüitas River is the most polluted stream among all those considered, showing below standard values for dissolved oxygen, fecal coliforms, total nitrogen, and total phosphorous.

It was found that some stations are not contributing useful information to the management and conservation of surface waters in Puerto Rico. Examples of this are Station 0590 on the Loiza River and 1364 on the Rosario River which present no trend or change in any of the parameters studied. The reasons for this in each of these two cases are quite distinct. Station 0590 is located immediately past the dam at Lake Loiza, which doesn't allow any external factor to affect the quality of the water between the lake and this station. As for the station on the Rosario River, the failure to detect any trends in the data is due to the fact that the river is located in a sparsely populated region with very few factories and very little agricultural activity, therefore being free of impacts that could change the quality of the water.

VII. RECOMMENDATIONS

The results of this study support the following recommendations for management of the network of sampling stations and of the data they collect:

1. Samples of water must be taken at each station at least once a month each and every month of the year. All due effort should be made to assure that the time lapse between samples is between 30 and 31 days consistently throughout the year.

2. Data files should be stored not by dates but by time versus concentration. A reference date can be taken and the samples labeled according to the number of months that have passed since the reference date; that is, with 1, 2, etc., instead of with January, February, etc. Organizing the data in this way will make statistical analyses faster and more efficient.

3. The same statistical analysis presented in this study should be performed using data from a number of recent previous years. This is advisable because a trend toward change in a particular parameter might have begun only in recent years and in such case it might not be detected in this study due to the use of all available data.

4. A similar analysis should also be performed for all available parameters so as to discover which are truly

essential and which can be eliminated without jeopardizing the reliability of results.

5. The water quality in the Rio Grande de Añasco, Rio Grande de Arecibo, Rio Bayamón, and Rio Rosario could be satisfactory if their fecal coliform levels are reduced and shows no significant trends to negative changes. Since the volumetric flow of these streams is considerable, they should be regarded as prime sources for future water supply needs in Puerto Rico.

6. The stations judged to be unnecessary should be relocated, or they could be eliminated to allow a greater concentration of effort on those stations that demonstrate changes in water quality.

REFERENCES

1. Smith, Richard A., Alexander, Richard B., and M. Gordon Wolman, "Analysis and Interpretation of Water Quality Trends in Major U.S. Rivers, 1974-81," U.S. Geological Survey Water Supply Paper 2307, 1985, pp. 1-25.

2. Steele, Timothy D., "Water Quality Monitoring Strategies," Journal des Sciences Hydrologiques, Vol. 32, No. 2, June, 1987, pp. 207-213.

3. Ravichandran, S., "Water Quality Studies on Buckingham Canal (Madras, India) - A Discriminant Analysis," Hydrobiologia, Vol. 154, May, 1987, pp. 121-126.

4. Gibbons, Jean Dickinson, Non-parametric Statistical Inference, McGraw-Hill Inc., New York, N.Y., 1971, pp. 1-21.

5. Walpole, Ronald E., Myers, Raymond H., Probability and Statistic for Engineers and Scientists, Macmillan Publishing Company, New York, N.Y., 1989, pp. 615-645.

6. Gross, Alan J., and Jeffrey R. Folts, "Analysis of Selected, Existing, Water Quality Data on the Connecticut River," Water Resources Research, Completion Report FY-73-2, 1980, pp. 1-29.

7. Helsel, Dennis R., "Advantages of Nonparametric Procedures for Analysis of Water Quality Data,"Journal des Sciences Hydrologiques, Vol. 32, No.2, June, 1987, pp. 179-190.

8. Berryman, David, Bernard Bobée, D. Cluis, and J.Haemmerli, "Nonparametric Tests for Trend Detection in Water Quality Time Series," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 545-556.

9. Van Belle, Gerald, and J.P. Hughes, "Nonparametric Tests for Trend in Water Quality," Water Resources Research, Vol. 20, No. 1, January, 1984, pp. 127-136.

10. Lettenmaier, Dennis P., "Detection of Trends in Water Quality Data from Records with Dependent Observations," Water Resources Research, Vol. 12, No. 5, October, 1976, pp. 1037-1046.

11. Hirsch, Robert M., J.R. Slack, and R.A. Smith, "Techniques of Trend Analysis for Monthly Water Quality Data," Water Resources Research, Vol. 18, No. 1, February, 1982, pp. 107-121.

12. Hipel, Keith W., "Nonparametric Approaches to Environmental Impact Assessment," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 487-492.

13. Alexander, Richard B., and R.A. Smith, "Trends in Lead Concentrations in Major U.S. Rivers and Their Relation to Historical Changes in Gasoline-Lead Consumption," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 557-569.

14. Hughes, James P., and Steven P. Millard, "A Tau-Like Test for Trend in the Presence of Multiple Censoring Points, " Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 521-531.

15. Smith, R.A., R.M. Hirsch,and J.R. Slack, "A Study of Trends in Total Phosphorous Measurements at Stations in the NASQAN Network," Water Supply Paper 2190, U.S. Geological Survey, Reston, Va., 1982, pp. 1-10.

16. Hirsch, Robert M., and J.R. Slack, "A Nonparametric Trend Test for Seasonal Data with Serial Dependence," Water Resources Research, Vol. 20, No. 6, June, 1984, pp. 727-732.

17. Hipel, Keith W., A. Ian McLeod, and R.R. Weiler, "Data Analysis of Water Quality Time Series in Lake Erie," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 533-544.

18. Lettenmaier, Dennis P., "Multivariate Nonparametric Test for Trend in Water Quality," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 505-512.

19. Hirsch, Robert M., "Statistical Methods and Sampling Design for Estimating Step Trends in Surface Water Quality," Water Resources Research, Vol. 24, No. 3, June, 1988, pp. 493-503.

20. Hays, William L., and Robert L. Winkler, Statistics: Probability, Inference, and Decision, Holt, Rinehart and Winston, Inc., New York, N.Y., 1971, pp. 835-837.

21. Gilbert, Richard O., Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, N.Y., 1987, pp. 204-252.

22. Neville, Adamm., and J.B. Kennedy, Basic Statistical Methods for Engineers and Scientists, International Textbooks Company, Scranton, PA., 1964, pp. 146-147.

23. Curtis, Russell E., Jr., Aquino, Zaida, Diaz, Pedro L., and Garcia, René, "Water Resources Data Puerto Rico and the U. S. Virgin Islands, Water Year 1986," U. S. Geological Survey Water- Data PR-86-1, April, 1988, p. 3.

24. Junta de Calidad Ambiental, "Ammendments to the Puerto Rico Water Quality Standard Regulation," Commonwealth of Puerto Rico / Office of the Governor, Environmental Quality Board, February, 1983, pp. 26-28.

APPENDICES

APPENDIX A

Quantiles of the Chi-Square Distribution with v Degrees of Freedom

Degrees of Freedom v	Probability of obtaining a value of χ^2 smaller than the tabled value													
	0.005	0.001	0.025	0.050	0.100	0.250	0.50	0.750	0.900	0.950	0.975	0.990	0.995	0.999
1	....	....	....	....	0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75	20.52
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96	26.12
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76	31.26
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	4.60	5.23	6.27	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00	45.32
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.64	55.48
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.80	55.76	59.34	63.69	66.77	73.40
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.42	104.22	112.32
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.64	107.56	113.14	118.14	124.12	128.30	137.21
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45
X	-2.576	-2.326	-1.96	-1.645	-1.282	-0.674	0.0	0.674	1.282	1.645	1.96	2.326	2.576	3.090

Source: After Pearson and Hartley, 1966.

For $v > 100$, take $\chi^2 = v[1 - 2/9v + X\sqrt{2/9v}]^3$ or $\chi^2 = 1/2\ [X + \sqrt{2v - 1}]^2$ if less accuracy is needed, where X is given in the last row of the table.

APPENDIX B

Cumulative Normal Distribution (Values of p Corresponding to Z_p for Normal Curve)

Z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5674	.4714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5967	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8595	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9671	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9963	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9967	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9956	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Source: After Pearson and Hartley, 1966.

APPENDIX C

PROGRAM FOR DETECTING MONOTONIC TRENDS USING SEASONAL KENDALL TEST AND SEASONAL KENDALL SLOPE ESTIMATOR

```
C
C               THIS IS THE PROGRAM TO CALCULATE THE
C              STATISTICS TESTING FOR MONOTONIC TREND.
C
C
C                 VARIABLES UTILIZED
C
C     VARSN  =  Used to do the summation of the variances
C         S  =  Statistic calculated for the Kendall and
C               Seasonal Kendall tests
C        LO  =  Number of seasons
C         N  =  Total quantity of data
C        MV  =  Number of missing values
C      TIES  =  Correction for variance
C         L  =  Used for the second summation of sgn
C      T(J)  =  Data Values
C      T(I)  =  Data Values
C         X  =  Difference between two data values
C       DIF  =  Seasonal change between two values
C        NV  =  Number of non-missing values
C      VARS  =  Variance of data
C         Z  =  Statistic for the normal distribution test
C      ZABS  =  Absolute value of Z
C
      DIMENSION T(30)
      VARSN = 0.0
      S = 0.0
C
      READ (5,*) LO
      DO 12 NUM = 1,LO
C
      READ (5,*) N, MV, TIES
           DO 3 J = 1,N
           READ (5,*) T(J)
3          CONTINUE
C
      L = N-1
C
                DO 10 I = 1,L
                NO = I+1
                     DO 8 J = NO,N
                     IF (T(J).EQ.0.OR.T(I).EQ.0) THEN
                          X = 0.0
                     ELSE
                     X = T(J) - T(I)
```

```
                        DIF = X/(J-I)
                        WRITE (6,*) DIF
                        ENDIF
                           IF (X.GT.0.0) THEN
                             X = 1.0
                           ELSE
                             IF (X.EQ.0.0) THEN
                                  X = 0.0
                             ELSE
                                  X = -1.0
                             ENDIF
                           ENDIF
     S = S + X
8                       CONTINUE
10                  CONTINUE
C
     NV = N-MV
     VARS = NV*(NV-1)*(2*NV+5)/18 - TIES
     VARSN = VARS + VARSN
12   CONTINUE
C
C
C

                   IF (S.GT.0.0) THEN
                        Z = (S-1)/VARSN**.5
                   ELSE
                        IF (S.EQ.0.0) THEN
                             Z = 0.0
                        ELSE
                             Z = (S+1)/VARSN**.5
                        ENDIF
                   ENDIF
C
     ZABS = ABS (Z)
          IF (ZABS.LT.1.65) THEN
               WRITE (6,20)
          ELSE
               WRITE (6,30)
          ENDIF
20   FORMAT (//,3X,'ACCEPT THE HYPOTHESIS WITH A S VALUE
        * OF',2X,F9.2,/)
30   FORMAT (//,3X,'REJECT THE HYPOTHESIS WITH A S VALUE
   * OF,,2X,F10.9,/)
STOP
END
```

APPENDIX D

ORIGINAL MONTHLY DATA VALUES

RIO DE LA PLATA BASIN
RIO DE LA PLATA AT PROYECTO LA PLATA, P.R.
WATER-QUALITY RECORDS, STATION 500 43 000

LOCATION.--Lat 18 09'37", long 66 13'44", Hydrologic Unit 21010005, at upstream side of bridge on Highway 173, 0.4 mi (0.6 km) northeast of Proyecto La Plata, and 2.5 mi (4.0 km) upstream from Rio Usabón.

DRAINAGE AREA.--54.8 mi (141.9 km), excludes 8.2 mi (21.1 km) upstream from Carite Reservoir, the flow of which is diverted to Rio Guamaní.

DATE (MM/DD/YY)	STREAMFLOW, INSTAN-TANEOUS (CFS)	pH (STANDARD UNITS)	DISSOL. OXYGEN (MG/L)	DISSOL. OXYGEN (PERCENT SAT.)	FECAL COLIFORM 0.7 UM-UF (COLS./100ML)	DISSOL. POTASS. (MG/L AS K)	TOTAL NITROGEN (MG/L AS NO3)	TOTAL PHOSPHOROUS (MG/L AS P)
10/06/59		7.7						
11/10/59		7.6						
12/09/59		7.4						
01/18/60		7.6						
02/02/60		7.7						
03/14/60		7.9						
04/11/60		7.9						
05/10/60		7.5						
06/02/60		7.4						
06/09/60		7.7						
07/12/60		7.7						
08/12/60		7.3						
09/17/60		7.0						
10/07/60		7.2						
11/14/60		7.3						
12/13/60		7.3						
01/08/61		7.4						
04/10/61		7.6						
05/08/61		7.6						
06/09/61		7.4						
07/15/61		7.4						
07/30/61		7.7						
09/13/61		8.0						
11/15/61		7.6						
12/13/61		7.4						
02/13/62		7.3						
02/20/62		8.3						
02/27/62		7.3						
03/09/62		7.5						
04/03/62		7.5						
05/08/62		7.5						
06/11/62		7.7						
07/10/62		7.2						
08/01/62		7.4						
09/14/62		7.1						
10/17/62		7.6						
11/19/62		7.7						
12/11/62		7.7						
01/09/63		8.1						
02/07/63		7.9						
03/06/63		7.8						
04/05/63		8.0						
05/10/63		8.3						
06/04/63		7.5						
07/03/63		8.4						
08/13/63		8.0						
09/19/63		7.4						
10/09/63		8.6						

ORIGINAL MONTHLY DATA VALUES - CONTINUED

DATE	STREAMFLOW	pH	DISSOL. OXY. MG/L	% SAT.	FECAL COLIFORM	DISSOL. POTASSIUM	TOTAL NITROGEN	TOTAL PHOSPHOROUS
11/20/63		7.6						
12/27/63		7.4						
01/20/64		8.5						
03/03/64		8.2						
04/16/64		7.9						
05/13/64		7.6						
05/27/64		7.7						
07/01/64		7.5						
08/12/64		7.8						
09/15/64		7.5						
10/06/64		7.8						
11/10/64		7.4						
12/07/64		7.7						
01/18/65		7.9						
02/04/65		7.4						
03/08/65		8.0						
04/08/65		8.0						
05/10/65		7.3						
06/09/65		7.3						
07/13/65		7.6						
08/18/65		7.9						
09/08/65		7.6						
10/13/65		7.6						
11/17/65		7.8						
12/29/65		7.7						
02/25/66		7.8						
03/16/66		7.9						
04/18/66		7.6						
05/18/66		7.6						
06/22/66		8.4						
07/19/66		8.4						
09/01/66		7.7						
10/19/66		7.9						
11/16/66		8.4						
12/19/66		7.9						
01/20/67		7.5						
02/21/67		8.3						
03/20/67		8.6						
04/18/67		7.7						
05/05/67		8.1						
07/09/67		8.0						
07/10/67		7.9						
08/02/67		8.0						
09/05/67		8.2						
10/02/67		7.9						
11/16/67		8.1				4.2		
12/11/67		8.2				2.6		
01/08/68		7.7				2.4		
02/01/68		7.8				6.5		
03/12/68		8.1				2.6		
04/02/68		8.1				2.8		
05/10/68		8.1				0.6		
06/10/68		7.7				2.4		
07/15/68		8.6				1.9		
08/19/68		7.8				3.5		
09/18/68		8.0				2.7		
10/18/68		9.0				1.5		
12/19/68		8.0				2.0		
01/10/69		7.6				1.5		
06/19/69		8.1				1.8		
01/22/70		9.0	8.5			1.3		
03/04/70		8.6				1.3		
04/09/70		7.8				1.4		
07/14/70		8.0	7.5			1.6		
08/19/70		7.2				1.0		

ORIGINAL MONTHLY DATA VALUES - CONTINUED

DATE	STREAMFLOW	pH	DISSOL. OXY. MG/L	% SAT.	FECAL COLIFORM	DISSOL. POTASSIUM	TOTAL NITROGEN	TOTAL PHOSPHOROUS
09/04/70		7.5	9.0			1.4		
02/02/71		7.8				1.6		
03/19/71		9.2				1.4		
04/14/71		8.4				1.4		
06/11/71		8.8				1.3		
07/06/71		7.3				1.7		
08/09/71		8.0				2.1		
09/07/71		7.6				1.4		
10/18/71		7.4				3.0		
01/17/72		7.4				1.7		
06/02/72		8.2				1.4		
06/28/72		7.6				1.2		
08/29/72		7.8				2.0		
05/01/73		7.4				2.3		
10/05/73		7.3				1.7		
02/12/74		7.2						
04/10/74		7.4	5.9			1.9		
04/16/74	10	8.2	5.2			1.9		
06/10/74		7.3	8.0			1.1	1.8	0.46
08/23/74		7.6	5.4			0.8	1.4	0.50
10/11/74	82	7.3	8.6			3.0	8.1	0.57
12/18/74		7.5	8.2			1.6	4.1	0.18
02/18/75	23	7.9	8.7			1.9	5.7	0.29
04/11/75	15	7.8	4.7			2.4	8.9	0.62
06/23/75	9	8.2	10.9			1.8	6.5	0.63
08/11/75	77	7.4	6.7			2.5	8.2	0.25
10/28/75	27	7.5	9.1			2.0	6.2	0.26
12/11/75	479	7.5	7.2			2.3	7.7	0.23
02/20/76	82	7.8	9.4			1.8	7.7	0.36
04/30/76	51	8.2	9.2			2.0	9.2	0.55
06/25/76	18	8.1	9.4			1.4	4.9	0.29
08/17/76	15	8.2	11.6			2.2	8.0	0.64
10/15/76	71	7.6	6.4			2.5	8.4	0.33
12/13/76	36	8.1	8.4			2.6	6.9	0.38
02/16/77	13	8.4	11.2			1.7	8.5	0.65
04/18/77	3	8.5	10.0			1.8	2.5	0.72
06/17/77	4	8.6	12.2			3.9	1.9	1.10
08/11/77	4	8.1	9.6			2.5	2.7	0.75
10/07/77	576	7.4	7.8			3.2	8.6	0.21
12/16/77	123	7.8	8.4			1.9	6.2	0.20
02/21/78	30	8.2	9.4			4.0	8.3	0.32
04/21/78	32	7.5	8.4		600	1.9	7.7	0.37
06/09/78	28	8.0	8.8		510	2.8	9.2	0.52
08/14/78	9	8.2	10.0		190	2.8	4.8	0.03
10/16/78	10	8.1	10.8		90	2.9	9.1	0.74
12/18/78	25	8.2	10.2		800	2.3	11.0	0.39
02/22/79	10	7.7	8.4		53	2.6	10.0	0.80
04/12/79	10	7.6	8.2		90	4.8	20.0	1.60
06/18/79	54	8.0	8.9		800	2.8		0.38
08/22/79	88	7.8	7.9		32000	1.6	8.2	0.32
11/21/79	92	8.1	8.6		48000			
01/22/80	26	8.1	8.7		910			
03/10/80	22	7.8	8.4		360		28.0	0.53
05/20/80	18	7.8	8.8		510			
07/02/80	10	8.3	8.6		160			
09/04/80	14	8.3	8.6		110	3.0	8.8	0.02
11/12/80	11	8.3	9.6		210	2.2	9.5	0.77
01/14/81	20	7.2	7.9		1400		11.0	0.38
03/19/81	9	8.3	10.6		170	3.6	17.0	1.10
05/21/81	32	8.0	8.2	108	800		8.5	0.54
07/23/81	88	7.9	8.0	99	1400	2.0	5.5	0.30
09/18/81	18	8.1	8.6	115	120	2.5	8.3	0.42
11/10/81	21	7.9	10.9	140	5700		11.0	0.81
01/26/82	22	8.2	9.8	120	72	2.3	8.7	0.52
03/05/82	33	7.9	9.3	113	82		10.0	0.39

APPENDIX E

ANNUAL DATA VALUES

RIO DE LA PLATA BASIN
RIO DE LA PLATA AT PROYECTO LA PLATA, P.R.
WATER-QUALITY RECORDS, STATION 500 43 000

LOCATION.--Lat 18 09'37", long 66 13'44", Hydrologic Unit 21010005, at upstream side of bridge on Highway 173, 0.4 mi (0.6 km) northeast of Proyecto La Plata, and 2.5 mi (4.0 km) upstream from Rio Usabón.

DRAINAGE AREA.--54.8 mi (141.9 km), excludes 8.2 mi (21.1 km) upstream from Carite Reservoir, the flow of which is diverted to Rio Guamaní.

TIME YEARS	STREAMFLOW AVERAGE	pH AVE.	OXYGEN AVERAGE	% SAT. AVERAGE	COLIFORM FEC. AVERAGE	POTASSIUM AVERAGE	NITROGEN AVERAGE	PHOSPH. AVERAGE
59		7.5667						
60		7.5000						
61		7.5667						
62		7.5231						
63		7.9167						
64		7.7818						
65		7.6750						
66		7.9600						
67		8.0417				3.4000		
68		8.0818				2.6273		
69		7.8500				1.6500		
70		8.0167	8.0000			1.3333		
71		8.0625				1.7375		
72		7.7500				1.5750		
73		7.3500				2.0000		
74	46.0000	7.5000	6.8833			1.7167	3.8500	0.4275
75	104.9500	7.7167	7.8833			2.1500	7.2000	0.3800
76	45.5000	8.0000	9.0667			2.0833	7.5167	0.4250
77	120.4000	8.1333	9.8667			2.5000	5.0667	0.6050
78	22.3000	8.0333	9.6000		438.0000	2.7833	8.3500	0.3950
79	1045.6167	7.8400	8.4000		16188.6000	2.9500	12.7333	0.7750
80	16.8333	8.1000	8.7833		376.6667	2.6000	15.4333	0.4400
81	31.3667	7.9000	9.0333	115.5000	1598.3333	2.7000	10.2167	0.5917
82	34.1667	7.7000	8.7333	109.3333	597.3333	2.1333	8.0833	0.4100
83	28.3333	8.2000	9.5333	122.8333	368.0000	2.0500	9.9167	0.4750
84	183.6800	8.2800	10.7600	141.6000	4964.0000	2.3667	8.4500	0.6560
85	57.9500	8.3000	10.1750	129.5000	1058.6667	1.6000	6.6167	0.3217
86	33.1667	8.2833	9.4333	124.0000	330.1667	2.6667	9.2000	0.3333
87	59.6667	7.9167	8.3167	107.1667	1215.0000	2.3333	9.1400	0.4133
88	165.2500	7.6500	8.2250	104.2500	25132.5000	2.1000	10.1250	0.5525