# LAND USE, WATER QUALITY AND RIVER-MOUTH LOADINGS: A SELECTIVE OVERVIEW FOR SOUTHERN ONTARIO

by

Edwin D. Ongley
Department of Geography
Queen's University
Kingston, Ontario

with assistance from
Prof. L.H. Broekhoven
Department of Mathematics
Queen's University

A Technical Report to the International Reference Group on GREAT LAKES POLLUTION FROM LAND USE ACTIVITIES of the International Joint Commission. Prepared under contract DSS OSU77-00129, Department of Fisheries and Environment, Canada.

March, 1978

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ACKNOWLEDGEMENTS

The many agencies which have provided data for this study are acknowledged within the body of the text. To them I extend my thanks for their co-operation since 1973 when I first became involved in the PLUARG study. Over the years, Mr. Robert Perkins has provided the systems and data processing expertise necessary to creatively manage the Task D data bank held at Queen's University. Mr. P. Hooper, graduate student in the Department of Mathematics, performed the factor analysis reported herein.

In revising an earlier draft of this report (Ongley, 1977b) I have profited from comment and constructive criticism from numerous individuals. Inevitably, however, not all readers of a summary report of a lengthy investigative program will agree on analytic methods, data presentation, subject matter included/omitted, or even, for that matter, on general research philosophy for overview purposes. Justification of research strategy and defense of views expressed herein remain, therefore, my sole responsibility.

For those who may use the material presented on the following pages I offer a piece of conventional though often sadly overlooked wisdom; statistical inference is no substitute for common sense and scientific insight.

## DISCLAIMER

The study discussed in this document was carried out as part of the efforts of the Pollution From Land Use Activities Reference Group, an organization of the International Joint Commission, established under the Canada - U.S. Great Lakes Water Quality Agreement of 1972. Funding was provided under several contracts from the Department of Fisheries and Environment, Canada.

Findings and conclusions are those of the author and do not necessarily reflect the views of the Reference Group or its recommendations to the Commission.

# INTRODUCTION

The Great Lakes Water Quality Agreement, with Annexes, Texts and Terms of Reference between the United States of America and Canada, signed at Ottawa on April 15, 1972, included a reference to study pollution in the Great Lakes System from agriculture, forestry, and other land use activities. The reference asked that the study assess whether the boundary waters of the Great Lakes System were being polluted by land drainage and, if so, where and to what extent and what remedial measures would provide improvements in controlling pollutants from land usage. Accordingly, the International Reference Group on Great Lakes Pollution from Land Use Activities was established in late 1972, and produced a detailed study plan (February, 1974) outlining an extensive study scheduled for completion by mid 1977 with a final report in July 1978.

The Reference Group established four task groups to examine various aspects of the problem. These Task groups were directed to:

Task A. To assess problems, management programs and research and to attempt to set priorities in relation to the best information now available on the effects of land use activities on water quality in boundary waters of the Great Lakes.

Task B. Inventory of land use and land use practices, with emphasis on certain trends and projections to 1980 and, if possible, to 2020. Present land use report to be completed in 1974, report on trends to be completed in 1975.

Task C. Intensive studies of a small number of representative watersheds, selected and conducted to permit some extrapolation of data to the entire Great Lakes basin and to relate contamination of water quality, which may be found at river mouths on the Great Lakes, to specific land uses and practices. Preparation of activities in 1974, intensive surveys in 1975 and 1976.

Task D. Diagnosis of degree of impairment of water quality in the Great Lakes, including assessment of concentrations of contaminants of concern in sediments, fish and other aquatic resources. Activities during 1974-1976.

The attempt to identify causality and to establish spatial and temporal trends within the PLUARG program has taken two distinctly different approaches. One, Task C, has been the detailed and site specific study of land use, land management practices, soil, geomorphic and hydrologic variables and their deterministic relationships with water quality attributes in a variety of representative drainage basins. This approach has involved intensive data collection for a period of two to three years.

The second approach (within Task D) has taken a wholistic view of the Great Lakes watershed. For the Canadian component of this study historical data sets of water quality (nearshore and fluvial) and quantity, together with land use, demographic, geologic, geomorphic and atmospheric data generated for PLUARG have been combined in a unique data bank. Manipulation of this data bank has permitted the identification of certain spatial and temporal trends of land use and water quality across (in particular) Southern Ontario.

Whereas causality identified in the site-specific approach loses specificity as one extrapolates into larger areas, the regional perspective cannot identify specific cause and effect except in a general sense (and largely by inference) because of the inevitable homogenization both of cause and effect which occurs as one moves downstream within ever larger drainage areas. The site-specific and regional perspectives are, therefore, complementary and must be viewed together.

## BACKGROUND

This report[1] is based upon a long-term PLUARG program in the Department of Geography at Queen's University under the auspices of Canadian Task D, for the assembly, management and analysis of data holdings pertinent to the assessment of regional trends for the Canadian side of the Great Lakes. These assessments have in the past involved computations of tributary loadings of a variety of nutrients, solids, contaminants and other water quality attributes, the evaluation of routinely available data for the purpose of developing regional trends, the statistical analysis of water quality-landuse relationships, and the development both of file management and of analytical methods for Task D purposes. The work performed under this program has been reported in a series of reports to (Canadian) PLUARG Task D of the International Joint Commission. Some of the work is summarized in a number of journal papers.

One of the objectives of the Task D program was to assess the spatial variability both of diffuse pollutant loads from Canadian tributary river systems to the Great Lakes and of the relationships between land use and water quality. The perspective adopted was that of the drainage basin unit. There were identified some 178 Canadian basins tributary to the Great Lakes and Connecting Waterways for which water quality data are available. Of these, 115 are in Southern Ontario, an area severely impacted by human occupancy patterns within the past two hundred years. Maps showing the location of all basins appear in Ongley (1974). Whereas river-mouth loads of a selected group of water quality variables have been produced for all years of record for all tributary basins, most statistical analyses reported herein reflect only those 115 basins[1] in the

---

[1] The present report is a revision of an earlier draft submitted to Canadian Task D (Ongley, 1977b).

southern agricultural part of the Province of Ontario.

## LIMITATIONS OF THE REGIONAL APPROACH

Any attempt to define regional land use-water quality linkages and trends inevitably employs a static perspective (i.e., the use of data averages) in what is in reality a dynamic environment which includes cyclic, random and sequential changes within data sets. Noise in the analysis is therefore an expected commodity generated, in part, by the static model employed. Dynamic models over the short term (i.e., several years) are precluded in this type of study for much the same reasons that additional noise is generated in the static model by the nature of the variables, by the kinds of data which describe the variables and by certain assumptions used (for example, demographic stability; lack of technological innovation in agriculture). Available water quality data are typically those collected under government surveillance programs in which, in Ontario, over a thousand sites are sampled once or twice a month. Moreover, not only do sampling programs change but so too do analytical programs evolve, incorporating innovations in laboratory technology with attendant changes in accuracy and precision of water quality data. As is illustrated below, phosphorus levels have changed dramatically in river systems, apparently as a result of phosphorus abatement programs introduced in the Province. The question of adequacy of surveillance-type data is noted below.

This study was further restricted by the use of whole drainage basins tributary to the Great Lakes. The enormous variations in basin size mitigates against isolation of precise land use-water quality trends which might result from use of small, relatively homogeneous basins. Therefore, further elimination of data points was often necessary to reduce excessive noise caused by large, heterogeneous systems. Although most data are held in spatially distributed form within drainage systems, the adoption of a whole basin perspective was made on the grounds that the cost of analysing spatially distributed data for the whole of Southern Ontario was unjustified until an initial appraisal of the value of such data using aggregate data by basin was made. Moreover, the purpose of the study was not to develop statistical models, but rather to gain insight into land use-water quality linkages which, with site-specific information, might guide the development of regional remedial strategies in terms of river contributions to Great Lakes' water quality. Furthermore, lack of resolution of land use-water quality relationships at mouths of larger rivers underscores the minimal expectation of beneficial results at the stream mouth if applying limited land use-specific or land management-specific remedial programs in lieu of comprehensive measures in large river basins.

---

[1] In practice, only 101 of the 115 basins have satisfactory records for statistical analysis.

## REMARKS ON REPORT CONTENT

This report, although limited to selected results of a statistical evaluation of phosphorus trends and of land use-water quality interactions, contains enough primary information to permit full understanding without recourse to earlier reports. The study is intended primarily as a source document to be used with other PLUARG information in the preparation of the final PLUARG report dealing with evaluation of regional trends of loadings and land use linkages and the development of remedial strategies. Therefore, much of the information is presented with explanatory notes but with little detailed comment for, prior to receipt of technical reports from other PLUARG activities, detailed synthesis at this stage in terms of remedial programs would be premature. It should also be noted that many relationships found amongst variables are noted herein but left unexplained because they are beyond this writer's experience.

Although PLUARG is primarily interested in substantive results, the multivariate analysis of the Task D data bank held at Queen's University has provided insight into analytical procedures involving regional data sets. Therefore, in addition to substantive conclusions drawn from the analysis, methodological experience gained through care and handling of this data bank is germane to potential post-PLUARG activities. It seems likely that future land use-water quality models for regional public policy development will involve a statistical component in which data of the type used here are the best which are likely to be available for regional analysis.

In most of the work reported here, aggregate basin characteristics are compared with mean (usually mean annual) river-mouth load and/or concentration data. Although solids data, which are particularly susceptible to seasonal (and event) hydrologic variations, have been assessed for seasonal resolution, land use-water quality linkages are here assessed using mean annual data. Sufficient data do exist to explore seasonal linkages, however there are good scientific and statistical reasons for refraining from doing so until the distributed data within individual basins can be fully exploited.

# SUMMARY CONCLUSIONS

## A. WATER QUALITY DATA EVALUATION

The use of water quality surveillance data such as those collected routinely by the Ontario Ministry of the Environment, for the calculation of river loads of nutrients and contaminants and the identification of spatial and temporal patterns of water quality has been the subject of considerable debate both within PLUARG and elsewhere. In particular, suspended solids data are prone to considerable error in view of the discharge-dependent nature of suspended mineral sediment. Because suspended solids are particularly important with respect to phosphorus and trace metal flux, the evaluation of the adequacy of suspended solids surveillance data has been an important part of this program (see, for example, Ongley, Ralston and Thomas, 1977).

A1. Surveillance data collected some 12 to 20 times per year per site over a period of years has proved to be adequate for identification of regional and temporal trends.

A2. Suspended solids data (mineral plus organic), although providing correct relative values amongst basins, cannot be used to estimate absolute values for suspended (mineral) sediment loads. However, such data do:
  (i) mirror observed (actual) seasonal sediment distributions,
  (ii) allow correct rank-ordering of basins in terms of total load and unit load (yield).

A3. Estimates for mean annual load either for solids or solutes can be obtained either from annual or monthly aggregations of data (bearing in mind the restrictions of A2), although monthly aggregations are preferred in order to account for large seasonal variations in concentration and discharge data.

A4. The use of statistical averages for computation of river loads for the Great Lakes area (as per A3) recognises that:

  (i) surveillance data do not reflect hydrodynamic factors involved in suspended solids transport except in a very general sense,
  (ii) mineral sediment in particular does not in general closely follow hydrodynamic theory due to:
    (a) preponderance of sub-sieve particles,
    (b) sediment exhaustion effect,

(c) sediment flux generally is controlled by supply factors.

A5. There being no independent data sets for evaluation of solute load calculations, it is an opinion[1] that solute estimates from surveillance data are more accurate than for sediment loads. Whether such estimates for solute loads are sufficiently accurate for lake response modelling is not known. Total phosphorus loads which involve a substantial particulate component may be considerably underestimated.

B. PHOSPHORUS (see also C11 - C15, D5 - D8)

Phosphorus flux to the Great Lakes is produced by point sources and a variety of diffuse sources including atmosphere, natural (background), agriculture and urban runoff. The results below reflect only that part of the Great Lakes' phosphorus load which is introduced by tributaries. The results do not include point sources with outfalls direct to the Lakes. Tributary data reflect virtually all diffuse phosphorus flux with the exception of atmospheric input to Lake surfaces and groundwater discharge directly into each Lake.

The effect of different land use categories upon phosphorus levels can be determined by using only those basins which are known to be free from point sources (i.e., nonpoint source-NPS-basins). Phosphorus-land use relationships are established for data averaged over the period 1968-1972, a period during which phosphorus levels commenced to fall as a result of abatement programs. However, land use data reflect the late 1960's with some update in 1970-71; population is drawn from the 1971 census. Temporal trends of phosphorus concentration are established through use of annual data for the period 1969-1974. All relationships need re-evaluation with post-1974 data.

By incorporating known point sources of phosphorus into river data sets, the relative importance of diffuse sources can be estimated by simple mass balance calculations. Partitioning of tributary-mouth phosphorus loads into point and diffuse components is carried out for 1974, a year for which atmospheric inputs were calculated in another PLUARG study and in which phosphorus levels in the Lower Lakes tributary drainage represent compliance with detergent and sewage treatment effluent standards.

---

[1] Based, for example, on comparison of published and observed soluble and total phosphorus load ratios with those observed in these data. Also, solutes do not generally behave like solids in their response to hydrodynamic factors produce large but often unsampled solid loads during large discharge events.

B1. Using a year by year comparison for the period 1969-1974, the phased reduction in detergent phosphorus commencing in 1970 has been accompanied by a reduction of some 60 to 70% in soluble (filterable orthophosphate) and total phosphorus in rivers tributary to the Lower Lakes. Reduced stream levels of phosphorus are not restricted to those rivers having point sources but are also observed in the group of streams with no identifiable point sources. Observed reductions are not related to rainfall-runoff variations (concentration-dilution effects). Although precise links between detergent-related phosphorus and stream levels of phosphorus in diffuse (nonpoint) systems are speculative, likely sources are:

(i) urban diffuse runoff,
(ii) septic tank leakage,
(iii) small private point sources from domestic and commercial establishments in basins classified as diffuse.

There has been no reduction in the use of phosphorus fertilizers during the same period.

B2. Assuming that the general decline in phosphorus levels over the period 1969-1974 is causally related to detergent phosphorus reduction, the failure of point source basins tributary to Lake Erie[1] to respond to the detergent abatement program suggests that either a significant amount of the phosphorus input to these rivers does not conform to abatement requirements, or there is a substantial increase in phosphorus use. The imposition of phosphorus removal in certain sewage treatment plants in 1974 has resulted in a reduction to 1969 levels, indicating a temporary arrest of an otherwise rising phosphorus trend.

B3. The long-term relative importance of various sources upon river- mouth levels of phosphorus is difficult to accurately determine. With the exception of municipal effluent data, quantitative estimates of other point source inputs to tributary phosphorus loads are not well documented but are thought to be minimal. An additional and probably large source of error is the transmission loss of phosphorus due to chemical, biological and sedimentary processes during transport downstream.

B4. The potential contribution (1974) of terrestrial deposition of atmospheric phosphorus to river-mouth loads of total phosphorus for the Lower Lakes is approximately 10% and rises to a maximum of 34% for drainage to the St. Mary's River in the Upper Lakes. It is not known, however, to what extent, if any, phosphorus of atmospheric origin is mobilized across the land surface into fluvial systems.

---

[1] Lake Erie in this context includes Canadian drainage tributary to Lake St. Clair and to the St. Clair and Detroit Rivers.

B5. Assuming zero transmission loss, point source (sewage treatment plant) contribution to 1974 river-mouth loads of total phosphorus is 33% in Lake Ontario and 12% for Lake Erie.[1]

B6. Background (essentially non-controllable) diffuse phosphorus (includes atmospheric inputs) accounts for 43% of 1974 river- mouth loads of total phosphorus to Lake Ontario and 22% to Lake Erie[1].

B7. Diffuse above-background (potentially controllable) phosphorus accounts for a minimum of 25% of 1974 river-mouth loads of total phosphorus to Lake Ontario and 65% to Lake Erie[1].

B8. Development of optimal remedial programs for diffuse phosphorus control should consider the importance of geographic variation of, in particular, urban and cropland land use categories. The known importance of diffuse urban runoff on levels of stream phosphorus concentration and the proximity of urban conurbations to Lake Ontario suggest that any attempt to reduce the 25% of total phosphorus load which is considered to be of above-background diffuse origin, should initially concentrate upon diffuse urban controls. In contrast, the lack of urban centres adjacent to Lake Erie and the positive relationship between cropland and total phosphorus levels in streams, suggests that an agricultural policy would initially have the greatest effect in reducing the 65% of the total phosphorus load to Lake Erie[1] which is thought to be controllable and of diffuse origin. The urban centres adjacent to Lake St. Clair and the Detroit River would require an approach similar to that for Lake Ontario, bearing in mind that there is considerable non-urban impact to Lake St. Clair by, in particular, the Thames River.

## C. LAND USE-WATER QUALITY RELATIONSHIPS

The objective of this part of the study is not to develop statistical models; rather, it is to assess the degree to which regional data sets may be used to gain insight into land use-water quality linkages. The study is, therefore, explorative. Substantive results (i.e., regression models) are restricted by the relatively small number of observations (101 basins in Southern Ontario) from which further selections by basin characteristics must be made, by the large variations in basin size with increasingly complex land use patterns in larger basins, and by noise inherent both in water quality data and in Canada Land Inventory land use categories.

---

[1] Lake Erie in this context includes Canadian drainage tributary to Lake St. Clair and to the St. Clair and Detroit Rivers.

The search for land use-water quality relationships was carried out primarily with simple regression procedures. It was found that where strong trends exist between two variables, the addition of subsequent independent variables using multiple regression seldom produced much improvement in the explained variance. This suggests that:

C1. Canada Land Inventory land use categories, although broadly defined, are sufficiently different that particular land uses which affect specific water quality characteristics tend to be lumped within single Canada Land Inventory categories.

The study has utilized concentration data rather than unit loads (yield = mass per unit area) both to avoid use of discharge estimators for small ungauged basins and because it can be demonstrated that,

C2. Concentration data tend to react to dominant land use rather than to hydrologic factors, and,

C3. Concentration and yield are highly correlated.

The nature of the study is explorative and, although the majority of the relationships noted in the main body of the report are statistically significant, the changing relationships between water quality and land use can be more usefully explored by examining trends in correlation coefficients rather than a rigorous application of hypothesis testing and statements of standard errors. For example, the possibility that certain trends may prove to be important is far more important in this study than avoidance of Type II errors, particularly given the restrictions inherent in these data (as noted above) and the inevitable transgression of certain statistical requirements.

A subset of the 228 variables available to this study was chosen to represent characteristic dependent (water quality) and independent (basin characteristics) variables. These were further examined for covariance and a final selection made as follows:

Dependent Variables[1]: total dissolved solids, suspended solids, total nitrogen, nitrate nitrogen, total phosphorus, chloride.

---

[1] Other contaminant data such as biocides and trace metals are generally too limited to include in this study. These form the basis of a forthcoming loads summary to Canadian Task D.

Independent Variables: population, basin area, percent of basin area in: large urban, agriculture, woodland, cropland, improved pasture, sand, loam, clay.

The land use and population variables reflect, in general, conditions in 1971. Of the 115 basins in agricultural Southern Ontario which are tributary to the Lower Lakes, southern Lake Huron and Georgian Bay, and to Lake St. Clair, Detroit and St. Clair Rivers, only 101 basins have sufficient data for the study period 1968-1972. Forty-nine of the 101 basins are classified as nonpoint (diffuse) source systems (NPS basins) and 52 as point source (PS) basins.

C4. Overburden categories of sand, loam and clay are largely unrelated to water quality or agricultural variables. This probably reflects the fact that the NPS category, which was used to assess trends in the absence of point discharges, contains few basins small enough to display primarily a single overburden type.

C5. Total Dissolved Solids, although moderately correlated with Large Urban (+0.55) and with Woodland (-0.60), are not uniquely related to any of the rural, urban or overburden variables.

C6. Suspended Solids are particularly related to Cropland (+0.74 for NPS basins) and strongly negatively correlated with Woodland and Improved Pasture.

C7. Disturbance associated with tillage of Improved Pasture is not evident in suspended solids concentrations at river mouths.

C8. Sediment production commonly associated with urban construction does not produce any correlation between river-mouth concentrations of suspended solids and the Large Urban land use category. Evidently, the effects of Cropland dominate general trends of river-mouth concentrations of suspended solids. Although the data encompass considerable urban construction, the water quality sampling strategy, the long record period and the relatively short-lived effect of construction mitigate against significant inflation of long-term urban-related sediment concentrations.

C9. Point sources are sufficiently associated with total nitrogen to obscure relationships between total nitrogen and other diffuse variables. However, in the absence of point sources, evidence suggests that total nitrogen is strongly related both to diffuse urban runoff and to Cropland.

C10. Unlike total nitrogen, nitrate nitrogen is not related to urban runoff but is somewhat dependent upon Cropland. The lack of strong relationships suggests that mineralization of nitrogen forms during downstream transport tends to obscure specific source-river-mouth relationships.

C11. At a time when reductions in phosphorus levels are observed due to implementation of phosphorus abatement programs, phosphorus is only moderately related to Large Urban when considering all (PS + NPS) basins.

C12. For nonpoint source basins, phosphorus is strongly associated with Large Urban (+0.69) implying that diffuse urban runoff (a spatial variable) is more directly related to phosphorus concentrations than population *per se.* The trend requires verification using additional data for large population centres.

C13. Although phosphorus is moderately correlated with Cropland (+0.63) in the absence of urban conditions, the presence of urban effects totally obscures phosphorus-Cropland relationships. This suggests that for diffuse systems, urban runoff is the dominant variable. This relationship should be re-evaluated using post-1974 data.

C14. Phosphorus concentrations tend to be negatively associated with Improved Pasture.

C15. Following from C5, C6 and C13, the negative relationship between Improved Pasture and both Phosphorus and Suspended Solids suggests that economies of minimum tillage practices should be evaluated for potential reduction of Cropland-related concentrations of suspended solids and phosphorus.

C16. Chloride concentrations are positively related to urban point sources. These data do not indicate a close relationship between river-mouth concentrations of chloride and diffuse urban runoff, probably because surveillance data are not sufficiently sensitive to pulses of excessive chloride runoff (from road salt application) during melt events.

## D. DATA FILE FILTERING

Once data files have been appropriately summarized for statistical purposes, the single major drawback in their use for the purpose of quickly establishing water quality-land use relationships, is the number of steps necessary and, accordingly, the amount of time expended to assess large numbers of combinations of variables. Notwithstanding the interactive statistical package developed for ease of regression analysis for these data files, several routine multivariate techniques were evaluated for their ability to identify general data relationships.

D1. A clustering algorithm in which variables are grouped according to distance functions was not found to be particularly useful.

D2 Principal component analysis verifies that variables (R mode) and observations (Q mode) tend to cluster as one would expect on *a priori* grounds. However, the data exhibit large variance which is unexplained by the first few principal components. In part, this result reflects inherent noise in the data (noted above). Observed R mode clusters reflect the fact that the variables commonly available in water quality data sets have already been selected on *a priori* grounds as having expected interrelationships.

D3. Principal component analysis in Q mode offers a particularly useful means for *a posteriori* inference by its ability to group basins (observations) which display similar/dissimilar characteristics. It is therefore a potentially valuable tool for identifying groups of basins which have similar/ dissimilar responses (water quality) or characteristics (independent variables) from which 'representative' basins could be chosen for further study.

D4. Discriminant analysis has proved to be a particularly efficient and effective method for the identification and ranking of those independent variables which contribute to the variance of each selected water quality (dependent) attribute. The results obtained from discriminant analysis were always substantiated by subsequent regression analysis (reported in Section C).

D5. Those basins in Southern Ontario with exceptionally high concentrations of total nitrogen, total phosphorus and soluble phosphorus invariably relate to urban conditions.

D6. Those basins within the upper third of concentration values for total nitrogen, total phosphorus and soluble phosphorus relate both to urban and to Cropland variables.

D7. Following from D4 and D5, remedial strategies may be selectively applied to urban situations for those basins with the highest concentrations of total nitrogen, total and soluble phosphorus. Where a reduction is required for these basins within the top one-third of concentration values, remedial measures will have to consider agricultural in addition to urban factors.

D8. When considering the large variations in river-mouth concentrations of nutrient forms, solids (dissolved and suspended) and chloride, basins with high concentrations of one substance do not necessarily display high concentrations of the others. Within the context of political and administrative expediency, a least-cost application of remedial programs for particular substances might consider basin-specific application in addition to regional application (see B6).

D9. The urban sub-categories (such as Medium Density Residential) of the Canada Land Inventory scheme and which in this analysis are part of discriminating functions, should be interpreted as surrogates for pollutant producers which are either located within, for example, Medium Density Residential, or are causally associated (although not necessarily linked in a locational sense) within the Medium Density Residential category.

# DATA DESCRIPTION

## INTRODUCTION

The analyses reported here are of two types. One is the examination of temporal trends of water quality at tributary mouths and the second is the identification of water quality-land use linkages. For the former, average annual mean values of phosphorus and chloride are derived from a large number of basins, whereas for the latter, data are averaged for individual basins over the period 1968-1972. In both cases, averaged data are required to overcome aberrations inherent in estimates of average water quality conditions derived from conventional stream monitoring programs (10-20 samples per site-year).

## WATER QUALITY-LAND USE LINKAGES

An attempt to statistically relate[1] stream-mouth loadings and concentration to watershed physiography, demography and land use has required two preliminary steps. One was the creation of an interactive statistical package (described in Ongley, 1977a) and the second way a judicious selection of variables from Task D data holdings. The variables selected are specified by watershed and are formatted into a Master Data Matrix (Table 1) which contains all available information for statistical analysis by watershed. For the purpose of this analysis none of the data, such as land use, etc., are spatially resolved within watersheds. The analysis is restricted to Southern Ontario (basin #1-115 of Table 2) where land use data are tabulated by Task B. There are 101 usable watersheds (having loadings data for all or part of the period 1968-72). Basins without data prior to 1972 have been omitted as they are not thought to represent the 1971 target year for land use and demography, nor do they have a sufficiently lengthy record for calculation of mean seasonal and annual loads.

## WATER QUALITY

Variables #1-195 (Table 1) are mean annual and mean seasonal values for the period 1968-72 calculated from Ontario Ministry of the Environment raw quality data and Water Survey of Canada discharge values. Descriptive statistics for these variables appear in Appendix 2. Phosphorus and chloride trends discussed in Sections 6 and 7 represent annual mean data obtained from the same sources. For the water quality parameters reported here, laboratory accuracy is

---

[1] This applies to Section 8 (covariance in Variable Subsets), Section 9 (Factor Analysis), Section 10 (Discriminant Analysis) and Section 11 (Correlation- Regression Analysis).

**TABLE 1:** MASTER DATA MATRIX

MASTER DATA MATRIX - Dependent Variables

| Variable | | V# | V# | V# | V# | V# | V# | V# |
|---|---|---|---|---|---|---|---|---|
| Par. | | TDS | Susp. Sol. | T. Kjeld. | $NO_3$-N | S.R.P. | Alk. | T. Iron |
| YR. | N | 1 | 31 | 61 | 91 | 121 | 151 | 181 |
| | C | 2 | 32 | 62 | 92 | 122 | 152 | 182 |
| | T | 3 | 33 | 63 | 93 | 123 | 153 | 183 |
| SPR. | N | d | 34 | 64 | 94 | 124 | 154 | 184 |
| | C | 5 | 35 | 65 | 95 | 125 | 155 | 185 |
| | T | 6 | 36 | 66 | 96 | 126 | 156 | 186 |
| SUM. | N | 7 | 37 | 67 | 97 | 127 | 157 | 187 |
| | C | 8 | 38 | 68 | 98 | 128 | 158 | 188 |
| | T | 9 | 39 | 69 | 99 | 129 | 159 | 189 |
| FALL | N | 10 | 40 | 70 | 100 | 130 | 160 | 190 |
| | C | 11 | 41 | 71 | 101 | 131 | 161 | 191 |
| | T | 12 | 42 | 72 | 102 | 132 | 162 | 192 |
| WINT. | N | 13 | 43 | 73 | 103 | 133 | 163 | 193 |
| | C | 14 | 44 | 74 | 104 | 134 | 164 | 194 |
| | T | 15 | 45 | 75 | 105 | 135 | 165 | 195 |

| Par. | | T. Solids | $NH_3$-N | $NO_2$-N | T.P. | Hard. | Chloride |
|---|---|---|---|---|---|---|---|
| YR. | N | 16 | 46 | 76 | 106 | 136 | 166 |
| | C | 17 | 47 | 77 | 107 | 137 | 167 |
| | T | 18 | 48 | 78 | 108 | 138 | 168 |
| SPR. | N | 19 | 49 | 79 | 109 | 139 | 169 |
| | C | 20 | 50 | 80 | 110 | 140 | 170 |
| | T | 21 | 51 | 81 | 111 | 141 | 171 |
| SUM. | N | 22 | 52 | 82 | 112 | 142 | 172 |
| | C | 23 | 53 | 83 | 313 | 143 | 173 |
| | T | 24 | 54 | 84 | 114 | 144 | 174 |
| FALL | N | 25 | 55 | 85 | 115 | 145 | 175 |
| | C | 26 | 56 | 86 | 116 | 146 | 176 |
| | T | 27 | 57 | 87 | 117 | 147 | 177 |
| WINT. | N | 28 | 58 | 88 | 118 | 148 | 178 |
| | C | 29 | 59 | 89 | 119 | 149 | 179 |
| | T | 30 | 50 | 90 | 120 | 150 | 180 |

KEY

| | | | |
|---|---|---|---|
| V# | Variable Number | N | Number of Samples |
| Vn | Variable Name | C | Mean Concentration |
| Par. | Water Quality Parameter | T | Mean Tons (short) |

MASTER DATA MATRIX - Independent Variables

| V# | NAME | |
|---|---|---|
| 196 | Basin Population | |
| 197 | Basin Area | |
| | LANDFORM CLASSES | (% of Basin Area) |
| 198 | Very Weakly Broken Plains | -Sand |
| 199 | | -Loam-Silt |
| 200 | | -Clay |
| 201 | Weakly Broken Plains | -Bare Bedrock |
| 202 | | -Sand |
| 203. | | -Loam-Silt |
| 204 | | -Clay |
| 205 | Moderately Broken Uplands | -Sand |
| 206 | | -Loam-Silt |
| 207 | | -Clay |
| 208 | | -Bare Bedrock |
| 209 | Strongly Broken Uplands | -all materials |
| 210 | All peat categories | |
| 211 | All escarpment categories | |
| | CANADA LAND INVENTORY (%of Basin Area) | |
| 212 | Urban Area >25,000 pop. | |
| 213 | Urban Area <25,000 pop. | |
| 214 | Extractive | |
| 215 | Agriculture | |
| 216 | Woodland | |
| 217 | Marsh, Swamo, Barren, Outdoor Recreation | |
| 218 | Low Density Commercial | |
| 219 | Medium Density Commercial | |
| 220 | High Density Commercial | |
| 221 | Low Density Residential | |
| 222 | Medium Density Residential | |
| 223 | High Density Residential | |
| 224 | Transportation | |
| 225 | Orchards, Horticulture, etc. | |
| 226 | Cropland | |
| 227 | Improved Pasture | |
| 228 | Unimproved Pasture | |

**TABLE 1** (cont'd)

| | |
|---|---|
| TDS | Total Dissolved Solids |
| T. Solids | Total Solids (mineral and organic) |
| Susp. Sol. | Suspended Solids |
| $NH_3$ - N | Ammonia Nitrogen |
| T. Kjeld. | Total Kjeldahl Nitrogen |
| $NO_2$ - N | Nitrite Nitrogen |
| $NO_3$ - N | Nitrate Nitrogen |
| T.P. | Total Phosphorus (as P) |
| S.R.P. | Soluble Reactive Phosphorus (= filterable orthophosphate) as P |
| Hard. | Hardness |
| Alk. | Alkalinity |
| T.Iron | Total Iron |

**TABLE 2:** TASK D WATERSHEDS USED IN STATISTICAL ANALYSES

| Basin No. | | Basin Name |
|---|---|---|
| 1 | NPS | Collins Creek |
| 2 | NPS | Millhaven Creek |
| 3 | NPS | Wilton Creek |
| 4 | | Napanee River |
| 5 | NPS | Salmon River |
| 6 | | Moira River |
| 7 | | Trent River |
| 8 | NPS | Smithfield Creek |
| 9 | | Butler Creek |
| 10 | NPS | Salem Creek |
| 11 | | Colborne Creek |
| 12 | NPS | Shelter Valley Brook |
| 13 | NPS | Brookside Creek |
| 14 | | Cobourg Brook |
| 15 | | Gage Creek |
| 16 | NPS | Ganaraska River |
| 17 | NPS | Graham Creek |
| 18 | NPS | Wilmot Creek |
| 19 | | Bowmanville Creek |
| ( | | Bowmanville C.(Soper)) |
| 20 | | Harmony Creek |
| 21 | | Oshawa Creek |
| 22 | | Pringle Creek |
| 23 | NPS | Lynde Creek |
| 24 | NPS | Carruthers Creek |
| 25 | | Duffin Creek |
| 26 | | Rouge River |
| 27 | | Highland Creek |
| 28 | | Don River |
| 29 | | Humber River |
| 30 | | Mimico Creek |
| 31 | NPS | Etobicoke Creek |
| 32 | | Credit River |
| 33 | | Oakville Creek |
| 34 | NPS | Fourteen Mile Creek) |
| 35 | NPS | Bronte Creek |
| 36 | | Rambo Creek |
| 37 | | Grindstone Creek |
| 38 | | Spencer Creek |
| 39 | | Redhill Creek |
| 40 | NPS | Stoney Creek |
| 41 | | 40 Mile Creek |
| 42 | NPS | 30 Mile Creek |
| 43 | | 20 Mile Creek |
| 44 | NPS | 16 Mile Creek |
| 45 | NPS | 15 Mile Creek |
| 46 | | 12 Mile Creek |
| 47 | NPS | 8 Mile Creek |
| 48 | NPS | 6 Mile Creek |
| 49 | | 4 Mile Creek |
| 50 | NPS | 2 Mile Creek |
| 51 | NPS | 1 Mile Creek |
| 52 | | Welland River) |
| 53 | NPS | Usshers Creek |
| 54 | NPS | Black Creek |
| 55 | NPS | Bakers Creek |
| 56 | NPS | Millers Creek |
| 57 | | Frenchman's Creek |
| 58 | | Grand River |
| 59 | NPS | Stoney Creek |
| 60 | | Sandusk Creek |
| 61 | | Nanticoke Creek |
| 62 | | Lynn River |
| 63 | NPS | Young Creek ) |
| 64 | NPS | Fishers Creek ) |
| 65 | | Dedrich Creek |
| 66 | | Big Ck(Norfolk) |
| 67 | NPS | Clear Creek |
| 68 | | South Otter Creek |
| 69 | | Big Otter Creek |
| 70 | NPS | Silver Creek ) |

| | | |
|---|---|---|
| 71 | | Catfish Creek |
| 72 | | Kettle Creek |
| 73 | NPS | Talbot Creek |
| 74 | NPS | Brock Creek |
| 75 | | 16 Mile Creek |
| 76 | NPS | OAC nr. Merlin ) |
| 77 | NPS | Muddy Creek |
| 78 | NPS | Sturgeon Creek |
| 79 | NPS | Cedar Creek |
| 80 | NPS | Big Creek(Essex) |
| 81 | | Canard River |
| 82 | NPS | Turkey Creek |
| 83 | | Little River |
| 84 | NPS | Pike Creek |
| 85 | | Puce River |
| 86 | NPS | Belle River |
| 87 | NPS | Ruscom River |
| 88 | | Thames River |
| 89 | | Sydenham River |
| 90 | NPS | Talford Creek |
| 91 | | Hickory Creek |
| 92 | | Ausable River |
| 93 | | Bayfield River |
| 94 | | Maitland River |
| 95 | NPS | Lucknow River |
| 96 | NPS | Pine River |
| 97 | | Penetangore River |
| 98 | | Saugeen River |
| 99 | NPS | Sauble River |
| 100 | NPS | Pottawatomi River |
| 101 | NPS | Sydenham River |
| 102 | NPS | Telfer Creek |
| 103 | NPS | Waterton Creek) |
| 104 | NPS | Orchard Creek) |
| 105 | NPS | Bighead River |
| 106 | NPS | Beaver River |
| 107 | NPS | Silver Creek |
| 108 | NPS | Black Ash Creek ) |
| 109 | NPS | Pretty River ) |
| 110 | NPS | Batteaux River) |
| 111 | | Nottawasaga River |
| 112 | NPS | Hog Creek) |
| 113 | NPS | Sturgeon River) |
| 114 | NPS | Coldwater River ) |
| (114a | | North River) |
| 115 | | Severn River |

NPS Nonpoint Source Basins
(All others are Point Source Basins)

( ) omitted from 1968-72 analytical
period because of lack of data

generally consistent over the 1968-72 period. With introduction of updated equipment some improvement in precision has occurred since 1970 (King, 1975).[1]

POPULATION (Variable #196 - Table 1)

Population by watershed is obtained by accessing the 1971 Census of Population by enumeration area (EA). One of the Task D data sets is a listing of EA's by watershed together with the population of the EA falling within each watershed. Total population of a watershed is a summation of population by EA's. The proportion of an EA falling within a watershed is used to pro-rate the EA population where less than 100% of the EA falls within the watershed. A number of EA's are empty, either because of suppression of data or by reason of land use activity (e.g., institutions, factories, etc.).

## BASIN AREA

These values are taken from the tabulations provided for land use data by watershed on tape by Task B.

## LANDFORM CLASSES

The Land Classification system of Ontario Ministry of Natural Resources (OMNR) contains a possible 100 categories. These data were put on magnetic tape and subjected to a frequency analysis in which the data (expressed as % of basin area) per land category were expressed as a frequency distribution for the 101 basins used in the analysis (e.g., how many basins have 25% of their area in deep clay [category #49]). Table 3 identifies the OMNR Land Classification categories in terms of Task D variables. Square brackets contain the number of basins in which each category is found. The frequency analysis indicates that the data can be grouped to reduce the number of variables from 100 to 14 (Table 1). All categories combined with P (peat) and cumulated by basin into variable #210 (Table 1) and, similarly, all Escarpment (E) combinations become variable #211. All land classification data on the Master Data Matrix are expressed as percent of basin area.

## LAND USE DATA

This information consists of eight major and nineteen minor land use categories (Table 4) mapped by the Canada Land Inventory and provided by Task B on magnetic tape, grouped by Task D watersheds. Land use data, although originally mapped in the late 1960's was updated (principally in urban areas) in 1970-71 for PLUARG. These data have been interfaced with existing Task D files and, after frequency analysis, the number of categories reduced (Table 4) from the original 27 (8 + 19) to 17 by eliminating redundant categories (appearing under both major and

---

[1] Soluble phosphorus (soluble reactive phosphorus - SRP) as used in this report implies filterable orthophosphate.

## **TABLE 3:** GROUPINGS OF OMNR LAND CLASSES INTO VARIABLES

### VERY WEAKLY BROKEN PLAINS(2)

| Unconsolidated Materials | Unit Number | Variable # |
|---|---|---|
| Peat and muck | P | |
| Sand | 1 [5] | 198 |
| Loam | 2 [11] | 199 |
| Silt | 3 [1] | |
| Clay | 4 [5] | 200 |
| All inorganic materials | 5 [0] | |

### WEAKLY BROKEN PLAINS

| Consolidated Inorganic Material (Bare Bedrock) | Unit Number | |
|---|---|---|
| Weakly resistant limestone bedrock | 6 [1] | |
| Weakly resistant argillaceous bedrock | 7 [0] | |
| Weakly resistant sandstone bedrock | 8 [0] | 201 |
| Moderately resistant bedrock | 9 [0] | |
| resistant bedrock | 10 [0] | |

| Unconsolidated Inorganic Material | Unit Number | | | | |
|---|---|---|---|---|---|
| | Bare and Shallow | Shallow (with bare and some deep) | | Deep & Shallow with some bare) | Deep |
| Sand (coarse and medium) | 11 [0] | 21 [0] | | 31 [0] | 41 [0] |
| Sand (fine and silty) | 12 [0] | 22 [0] | 202 | 32 [0] | 42 [8] |
| Sand (coarse to fine & silty) | 13 [5] | 23 [4] | | 33 [3] | 43 [8] |
| Sand, with other materials | 14 [2] | 24 [1] | | 34 [2] | 44 [51] |
| Loam | 15 [0] | 25 [0] | | 35 [4] | 45 [16] |
| Loam, with other materials | 16 [4] | 26 [10] | 203 | 36 [15] | 46 [45] |
| Silt | 17 [0] | 27 [0] | | 37 [0] | 47 [0] |
| Silt, with other materials | 18 [0] | 28 [0] | | 38 [0] | 48 [4] |
| Clay | 19 [0] | 29 [0] | 240 | 39 [6] | 49 [47] |
| Clay, with other materials | 20 [0] | 30 [0] | | 40 [7] | 50 [52] |

KEY

(1) Landforms are differentiated on the basis of their surface relief and nature of the unconsolidated materials or bedrock which give substance to their relief.

(2) Brokenness of relief patterns refer to the degree and frequency with which elevations vary from the mean level of the land surface. The five relief classes used here were established to provide groupings of slope patterns which are significant in land use because of their influence upon surface drainage and local climate.

[ ] Number of basins in which landform appears.

### MODERATELY BROKEN UPLANDS (2)

| Unconsolidated Inorganic Material | Unit Number | | | | |
|---|---|---|---|---|---|
| | Bare and Shallow | Shallow (with bare and some deep) | Variable # | Deep and Shallow some bare) | Deep |
| Sand (coarse and medium) | 51 [0] | 61 [0] | | 71 [0] | 81 [0] |
| Sand(fine and silty) | 52 [0] | 62 [0] | 205 | 72 [0] | 82 [0] |
| Sand (coarse to fine & silty) | 53 [7] | 63 [5] | | 73 [3] | 83 [12] |
| Sand, with other materials | 54 [0] | 64 [1] | | 74 [1] | 84 [17] |
| Loam | 55 [0] | 65 [0] | | 75 [0] | 85 [12] |
| Loam, with other materials | 56 [1] | 66 [1] | 206 | 76 [0] | 86 [19] |
| silt | 57 [0] | 67 [0] | | 77 [0] | 87 [0] |
| Silt, with other materials | 58 [0] | 68 [0] | | 78 [0] | 88 [0] |
| Clay | 59 [0] | 69 [0] | 207 | 79 [1] | 89 [5] |
| Clay, with other materials | 60 [0] | 70 [0] | | 80 (0) | 90 [12] |

| Consolidated Inorganic Material (Bare bedrock) | Unit Number | |
|---|---|---|
| Weakly resistant limestone bedrock | 91 [0] | |
| Weakly resistant argillaceous bedrock | | |
| Weakly resistant sandstone bedrock | 92 [0]<br>93 [0] | 208 |
| Moderately resistant bedrock | 94 [0] | |
| Strongly resistant bedrock | 95 [1] | |

### STRONGLY BROKEN UPLANDS

| Inorganic Material | Unit Number | |
|---|---|---|
| All unconsolidated materials, deep | 96 [2] | |
| All unconsolidated materials, deep to bare bedrock | 97 [0] | 209 |
| Bare bedrock and shallow unconsolidated materials | 98 [0] | |

### VERY STRONGLY BROKEN LAND

| Inorganic Material | Unit Number |
|---|---|
| All inorganic materials | 99 [0] |

### Classes added for Task D purposes

| | Unit Number |
|---|---|
| All P categories | 100 |
| All Escarpment categories | 101 |

**TABLE 4:** CANADA LAND INVENTORY: LAND USE CATEGORIES

| | Major Land Uses | Fate of Category for Task D Use |
|---|---|---|
| 1 | - Urban area > 25,000 population | |
| 2 | - Urban area < 25,000 population | |
| 3 | - Extractive | |
| 4 | - Agriculture | |
| 5 | - Woodland | |
| 6 | - Marsh and Swamp | combined |
| 7 | - Barren | |
| 8 | - Outdoor recreation | |
| | Minor Land Uses | |
| 1.1 | - Low density commercial | |
| 1.2 | - Medium density commercial | |
| 1.3 | - High density commercial | |
| 1.4 | - Low density residential | |
| 1.5 | - Medium density residential | |
| 1.6 | - High density residential | |
| 1.7 | - Transportation | |
| 2.1 | - Urban area < 25,000 population | eliminated (see #2) |
| 3.1 | - Extractive | eliminated(see #S) |
| 3.2 | - Slag heaps | |
| 4.1 | - Orchards, horticulture etc. | |
| 4.2 | - Cropland | |
| 4.3 | - Improved pasture | |
| 4.4 | - Unimproved pasture | |
| 5.1 | - Productive woodland | eliminated(see #5) |
| 5.2 | - Non-productive woodland | |
| 6.1 | - Marsh and swamp | eliminated (see #6,7,8) |
| 7.1 | - Barren | |
| 8.1 | - Outdoor recreation | |

minor headings) and by combining relatively unimportant categories having small total areas (e.g., #214 of Table 1). No information has been eliminated by this data reduction. All land use data are recorded as percent of total basin area.

## POINT AND NONPOINT BASINS

In the identification both of trends in pollutant-land use relationships and of priorities in dealing with pollutant fluxes from tributary drainage, it is useful to separate the 101 basin data set into those basins having point sources (PS) and those without (Table 2). The nonpoint source (NPS) subset can be used to establish relationships which are unique to diffuse sources of pollution whereas the point source subset combines point and diffuse sources which, together, may obscure meaningful trends.

Separation of the data set into the point and nonpoint categories was carried out by the Ontario Ministry of the Environment (Mr. J.G. Ralston, personal communication) on the basis of presence or absence of municipal wastewater (sewage) treatment plants and industrial outfalls. No attempt was made to ascertain the type of pollutant associated with industrial outfalls. Hence, some basins lumped into the point source category by virtue of an industrial outfall may or may not be affected by that outfall depending upon the water quality attribute under consideration. The only outfalls considered are those upstream of the water quality site from which data have been used to calculate basin loadings.

# RANKED LOADINGS, CONCENTRATIONS AND YIELDS OF RIVER-MOUTH DATA

In Ongley (1977a) the thirteen quality parameters (plus total nitrogen) for which mean annual loadings and concentrations were calculated for the period 1968-72, were ranked in terms of concentration and total tonnage (short tons), and unit load (yield) for the 101 Task D basins in Southern (agricultural) Ontario. These data have been resorted and ranked (Appendix 1)[1] for the 52 point source basins and the 49 nonpoint source basins identified in Table 2. In addition, all basins plus point and nonpoint subsets have been ranked for concentration and tonnage[2] in which each is weighted (multiplied) by basin area. This has been done to facilitate two approaches for identifying 'worst case' basins. Ranked yield (unit load) information is also provided to complete the picture.

One approach is to locate those basins with the highest concentrations and/or the largest tonnages and apply appropriate remedial measures. From the point of view of identifying and rectifying sources of high concentrations, the smaller basins will generally be more easily managed than larger basins. However, one may adopt the position that small basins with high concentrations produce a relatively insignificant portion of total tributary load to Lake. Therefore, the second approach for identifying 'problem' basins is to rank the basins using concentrations weighted by basin area. If one is able to identify the sources of higher concentrations (particularly in the case of diffuse sources) in the larger basins (top ranked basins), one may appreciably reduce total load to the Lake. Using the weighted approach the top ranked basins are not always the largest. Therefore, one may use the rankings to decide which basins are of an appropriate size for which remedial measures might be expected to produce beneficial results, and which are too large and therefore too complex to attempt a remedial program. This may be a particularly useful strategy to apply to NPS basins if it is decided to selectively apply remedial measures for the purpose of evaluating the impact of policy recommendations.

The weighting of tonnage by basin area is <u>not</u> immediately useful in identifying problem watersheds. 'Worst' case situations are most easily identified by unweighted ranking of tonnages

---

[1] Because of bulk, this Appendix is bound separately and has been made available to PLUARG-Task D in limited quantity.

[2] For reasons of reporting continuity, computer listings use short tons. Analyses in this report have been converted to SIS units.

or ranked unit loads (yield). However, top ranked unit loads tend to be the smallest basins, whereas the largest basins are generally top ranked in total tonnage. Therefore, the weighted tonnages are provided solely for the purpose of evaluating the relative importance of a basin's size, its tonnage and unit load. This may be important when implementing remedial programs in deciding trade-offs amongst basin size (complexity), expected reduction due to remedial programs and relative importance to total Lake loadings. However, 'worst' case priorities cannot be initially assigned on the basis of area-weighted tonnages.

# PHOSPHORUS TRENDS AND RESPONSE TO ABATEMENT PROGRAMS, 1969-1974

## INTRODUCTION

The role of phosphorus in the well-documented eutrophic trend in the Lower Great Lakes has been the subject of intensive study over the past decade. Legislative response to the need for phosphorus control has been (1) detergent phosphorus limitations, introduced by the Province of Ontario in 1970 and, at the time of writing, enforced by three States and several municipalities bordering the Great Lakes and, (2) the implementation of phosphorus effluent standards for municipal sewage treatment plants to be achieved through phosphorus removal in primary and/or secondary treatment. The control of phosphorus within the Great Lakes area has, in part, been a response to studies of the International Joint Commission culminating in the 1972 (International) Great Lakes Water Quality Agreement.

The work reported here examines response to phosphorus controls of point and diffuse phosphorus loads at river-mouth locations for Canadian drainage tributary to the Great Lakes. The implementation of further remedial measures to control diffuse phosphorus sources must be predicated by knowledge both of the relative importance and appropriate control costs of diffuse versus point sources of phosphorus and of trends in river phosphorus loads. Because the state-of-art of diffuse phosphorus sources, transport through fluvial systems, and bioavailability of sediment-related phosphorus is inexact, this study adopts a spatial perspective and, for Southern Ontario, examines the effectiveness of phosphorus control in terms of changes of concentration levels at river mouths. The choice of river-mouth locations to evaluate the effect of phosphorus legislation reflects the PLUARG mandate which was to evaluate the impact of diffuse (nonpoint) pollution upon water quality of the Great Lakes.

Although both total and soluble phosphorus forms are examined here, it is recognised that inputs of large amounts of soluble phosphorus into fluvial systems are accompanied by a rapid reduction in the soluble form downstream due to chemical, biological and sedimentary processes. Although the ensuing total phosphorus load may be large, the available fraction is significantly reduced. Although phosphorus has a high affinity for particulates and indeed up to 95% of total phosphorus may be transported in this form (Logan, 1977) the amount of particulate-related phosphorus which is bioavailable is not well understood and can vary considerably depending on the fluvial/lacustrine depositional environment. This study, therefore, makes no attempt to evaluate the ecological impact of changes in river-mouth concentrations of phosphorus.

Using EPA data, phosphorus from U.S. domestic sources is estimated by Hopson (1975) to be about 3.5 lb/capita /year, comprised of 2.3 lb. from synthetic detergents with phosphorus builders and 1.2 lbs. from human excretions. Assuming a water use of 100 gallons per capita per year, Hopson claims that the calculated mean phosphorus concentration of 11.5 mg/L agrees with measured values in some municipal wastewaters. A full ban on detergent-related phosphorus should, therefore, produce a 65% reduction in phosphorus concentration in municipal wastewater to a level of some 4.0 mg/L. The latter value compares with 5.0 mg/L predicted by Hetling and Carcich (reported in Sweeney, 1973). The actual standard achieved in individual sewage treatment plants will, in part, depend upon the phosphorus concentration in influent wastewaters reflecting the forms and sources of phosphorus. Data from Hopson for Erie County, New York, indicate that plants with conventional biological treatment can achieve a total phosphorus effluent concentration of from 2.2 to 3.8 mg/L following a full ban on detergent-related phosphorus. These values represent a 60% decrease from 'no ban' concentrations.

The effect of the detergent-related phosphorus ban on stream water quality in Erie County was studied by Sweeney (1973). Using June-August data collected at 152 stream sites reflecting municipal point source pollution, Sweeney found a 60 and 67% reduction for total and orthophosphorus following imposition of a full ban. Sweeney concluded that the reductions could not be attributed to hydrologic variation (concentration-dilution effects), variation in industrial effluent, nor to improvements in sewage treatment facilities.

Phosphorus control legislation (limited to laundry detergent phosphorus) in the Province of Ontario is summarized in Tables 5 and 6 and is in accordance with the Canada Water Act (1972). The 5 mg/L standard for municipal wastewater effluent is a *de facto* standard reflecting reduction (to 2.2% by weight) of phosphorus in detergent. The 1.0 mg/L standard is achieved by phosphorus removal in the treatment plant.

## DATA

Phosphorus data provided by the Ontario Ministry of the Environment (MOE) are here reported as Soluble (Reactive) Phosphorus (filterable orthophosphate using the conventional 0.4511 diameter as the particulate/solute boundary) and Total Phosphorus in mg/L of elemental phosphorus. Data produced since 1969 have been analysed using auto-analysers, therefore, data precision is good. For the purpose of this paper, 1969 has been chosen as a benchmark year and the data over the period 1969-1974 have been utilized.

Quality data are available for 115 river basins tributary to the Lower Lakes, Connecting Waterways) and the Southern parts of Lake Huron and Georgian Bay. This geographical area

---

[1] Connecting Waterways in this chapter implies drainage tributary to Lake St. Clair, and to the St. Clair and Detroit Rivers.

**TABLE 5:** LEGISLATED LEVELS OF PHOSPHORUS IN LAUNDRY DETERGENT, ONTARIO

| Date | Detergent by weight as P (%) | Detergent Contribution to [a] Municipal Sewage Loads of Total Phosphorus (%) |
|---|---|---|
| pre 1970 | 16.0 | 45 |
| 1970 | 8.7 | 30 |
| 1973 | 2.2 | 10 |

[a] Mr. J.G. Ralston, Ontario Ministry of the Environment, personal communication

**TABLE 6:** MUNICIPAL WASTEWATER EFFLUENT STANDARDS, ONTARIO[a]

|  | Drainage to | 1+ mgd [b] Plants | < 1 mgd Plants |
|---|---|---|---|
| 1974 | Lake Erie and Connecting Waterways [c] | 1.0 mg/L | 1.0 mg/L |
|  | Trent System of Lake Ontario | 10 mg/L | 1.0 mg/L |
|  | Elsewhere | 5.0 mg/L [d] | 5.0 mg/L [d] |
| 1976 | As above plus Direct Discharge to Lake Ontario | 1.0 mg/L | 5.0 mg/L [d] |

[a] As of January 1 of respective year. Standards are expressed as elemental P. Source is Ontario Ministry of the Environment

[b] mgd is million gallons per day

[c] drainage tributary to Lake St.Clair, and to the St. Clair and Detroit Rivers

[d] some municipal wastewater plants may comply with a 1.0 mg/L standard where local problems exist. 5.0 mg/L is a *de facto* standard reflecting reduced detergent loadings.

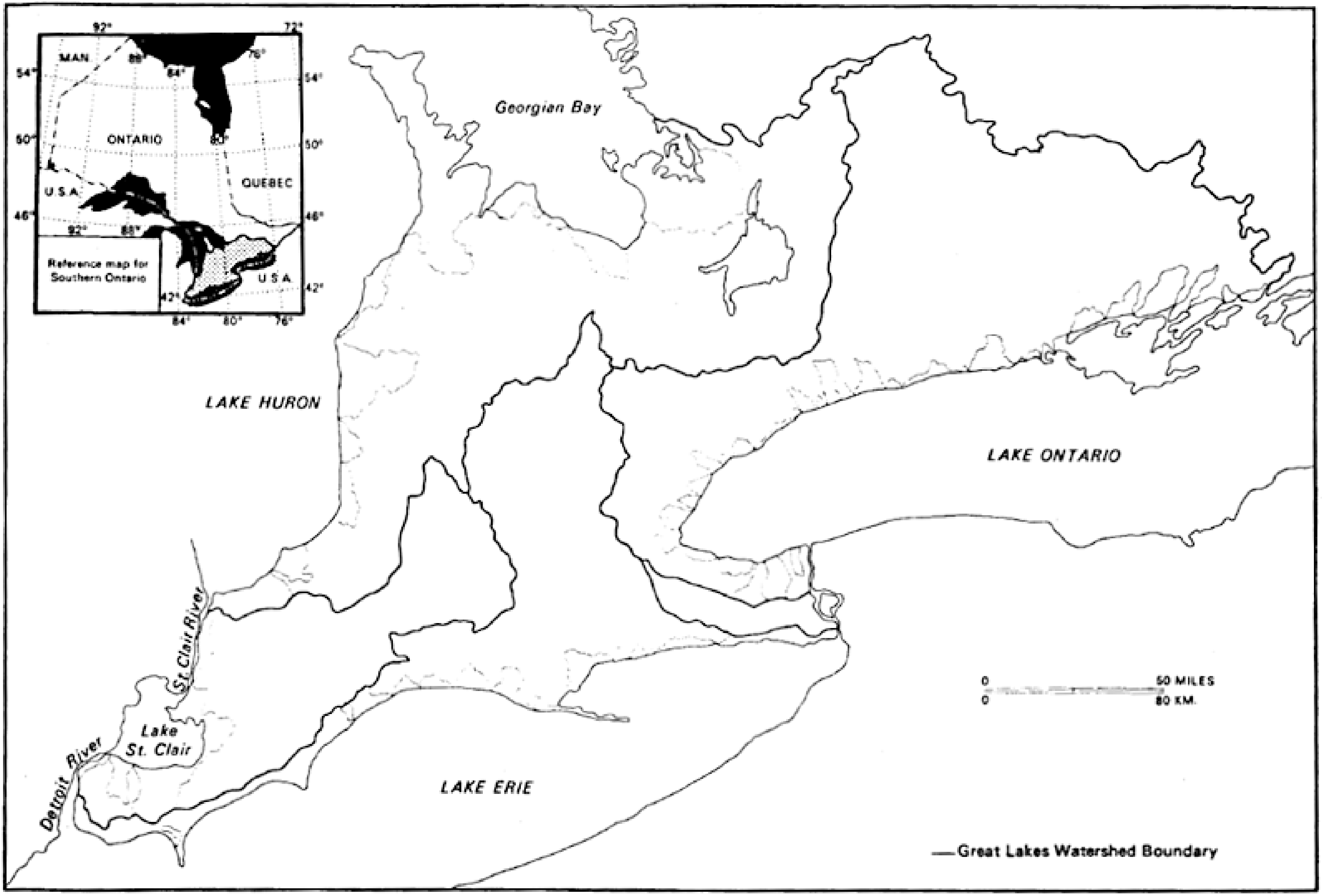

**Figure 1:** Tributary drainage area in Southern Ontario for which phosphorus concentration data are available for the entire period 1969-1974.

comprises that part of Ontario which is primarily south of the Canadian Shield and which has some agricultural potential. Of the 115 basins, 78 have data for the entire 1969-74 period (Figure 1). Because temporal trends are identified, only the latter group of basins is used in order to eliminate undue bias introduced through use of partial records. Table 7 indicates the sample size for each water body and the range in basin areas for point and nonpoint source basins.

The basis for evaluating the effectiveness of phosphorus legislation is to compare basins with point sources and those without. As noted above, using MOE records of municipal treatment plant and industrial outfall locations, 43 of the 78 basins were defined as PS basins[1] and the remaining 35 as NPS basins. Inevitably, small unidentified point sources associated with private and commercial establishments are included in the NPS category. Urban diffuse runoff whether or not collected in separate storm sewers is regarded as a nonpoint source. No attempt was made to evaluate the importance of municipal or industrial sewage effluent at the stream-mouth in cases where the outfall is some distance upstream. Phosphorus in industrial effluent (in tributaries) is considered minimal[2] though a few industries are responsible for intermittent discharges. However, in the absence of industrial effluent quality data, all such basins were lumped into the point source category. All basins contain a variety of land uses, ranging up to 90% cropland and, in one extreme NPS case, 63% urban.

Although phosphorus loads are the important variable in lake response models, the calculation of stream-mouth loads require discharge data which are not available in over half the basins used. It follows that yield (unit load), which is a derivative of load and which is generally the unit used to study diffuse source production of pollutants, cannot be accurately defined for those basins lacking discharge data. Although neither dimensionless nor characteristic of an average basin condition, yield does usefully describe inter-basin variability, especially where land use and physical variables such as soil and relief are reasonably uniform. It can be shown, however, that when using annual mean values as in this study, concentration is closely related to yield and can be used to compare land use-water quality relationships amongst basins.

Load is defined as

$$L = C \bullet Q \bullet k \qquad (1)$$

where C is concentration, Q is discharge and k is a constant. Generally, annual or monthly mean discharge is directly proportional to basin area (A) for Great Lakes tributary drainage (Ongley, 1974)

---

[1] Hereafter, PS and NPS refer to point and nonpoint (diffuse) source respectively. Assignment of basins to the point source category was based on 1973-74 records.

[2] Mr. J.G. Ralston, Ontario Ministry of the Environment, personal communication.

**TABLE 7:** NUMBERS AND SIZES OF BASIN SETS[a], SOUTHERN ONTARIO

| Drainage to | Numbers of Basins<br>Point Source | <br>Nonpoint Source | Basin Areas ($km^2$)<br>Point Source<br>Min. | <br><br>Max. | <br>Nonpoint Source<br>Min. | <br><br>Max. |
|---|---|---|---|---|---|---|
| Lake Ontario | 26 | 17 | 11 | 12715 | 4 | 932 |
| Lake Erie | 9 | 5 | 82 | 6653 | 24 | 152 |
| Lake Erie, Lake St.Clair, St.Clair & Detroit Rivers | 14 | 7 | 48 | 6653 | 24 | 152 |
| Southern: Lake Huron and Georgian Bay | 8 | 6 | 67 | 5064 | 57 | 612 |

[a] Basins having six year's phosphorus data for the period 1969 - 1974.

TABLE 8: WEIGHTED[a] UNIT AREA DISCHARGE FOR TRIBUTARY DRAINAGE

(meters$^3$/second/km$^2$ x $10^2$)

| Drainage to | 1969 | 1970 | 1971 | 1972 | 1973 | 1974 |
|---|---|---|---|---|---|---|
| Lake Ontario | 0.98 | 0.76 | 0.84 | 1.21 | 1.16 | 1.15 |
| Lake Erie | 1.08 | 0.81 | 0.78 | 1.02 | 1.07 | 1.11 |
| Lake Huron | 1.19 | 0.97 | 0.95 | 1.10 | 1.17 | 1.19 |

[a] Values are derived from discharge records which are weighted by tributary basin areas.

in which case Equation 1 can be rewritten as

$$L = C \bullet A \bullet k \tag{2}$$

for monthly and annual aggregations of data. From the formulation for yield (Y)

$$Y = \frac{C \bullet A \bullet k}{A} \tag{3}$$

it is seen that yield is a function of concentration. Moreover, as noted in Sections 9 and 10, phosphorus concentrations vary over three orders of magnitude depending largely upon land use, particularly in the case where basins contain even a small fraction of their area in urban land use. For the purpose of this report, therefore, concentration is the variable of interest and allows inter-basin comparisons without resorting to estimates of discharge in ungauged basins.

All data are reported as annual mean values. Although it is known that where sediment is involved, land use-water quality relationships are more likely to be seen in spring data when contributing area approaches 100%, it has been shown (Ongley *et al.*, 1977) that annual data provide a summary of basin load and yield characteristics which is adequate for comparative purposes. Data are presented on a lake-by-lake basis wherein all data are aggregated according to recipient Great Lake. In addition, Lake St. Clair and the Detroit and St. Clair Rivers are grouped with Lake Erie insofar as they are strongly linked in terms of generation, transfer and fate of phosphorus within the aquatic system. For the same reasons, municipal effluent standards established in 1974 apply universally to this group, the effect of which should be seen in 1974 data.

## HYDROLOGIC FACTORS

Temporal trends in phosphorus concentration in lacustrine environments must consider dilution or concentration effects due to natural changes in the hydrologic regime and the subsequent differences in relative mass balances of phosphorus and water. It is to be expected that under certain conditions, increasing levels of phosphorus concentration are to be expected regardless of effluent standards. This effect is also seen in fluvial systems where rural environments have been seen to produce greater concentrations of phosphorus under increased runoff conditions (the runoff effect), whereas urban point sources of phosphorus are diluted. One should, therefore, consider phosphorus flux through time in light of natural hydrologic variation. Table 8 illustrates the annual mean unit area weighted discharge[1] over the period 1969-1974. Although there is some difference amongst the various tributary groups reflecting climatic zonation in Southern Ontario, all illustrate a wet 1969, two successive dry years in 1970-71 and a return to wet conditions in 1972 through 1974. Not coincidentally, 1974 water levels on the Lower Lakes reached record highs.

---

[1] Weighted by basin area to reflect the influence of large basins.

## WORKING HYPOTHESIS

As a working hypothesis it is reasonable to expect that a reduction in phosphorus as a result of legislation should be seen in those basins having point sources. Principally, this should occur due to reduction in detergent-related phosphorus and made apparent in stream water quality by the resulting reduction in phosphorus loadings from municipal treatment plants. Additionally, the 1974 effluent standard (Table 6) of 1.0 mg/L of phosphorus in municipal wastewater plants in the Lake Erie and Connecting Waterways group ought to be noticeable in PS basins. Although some improvement in sewage handling facilities may have occurred over the record period, any beneficial effects may be obscured in individual plants by increasing industrial and commercial use of phosphorus.

It is proposed that NPS basins, which one might regard as being least affected by detergent legislation and unaffected by effluent standards, should provide a water quality performance standard against which PS basins can be compared. The concentration of phosphorus in NPS basins should reflect hydro-chemical pathways of natural and anthropogenic diffuse sources of phosphorus into fluvial systems. Also, year to year fluctuations of phosphorus concentration in NPS basins should reflect terrestrial and aquatic mobilization of phosphorus under changing climatic (runoff) conditions.

It should be noted that wide variations in phosphorus concentrations do exist in NPS basins for, as noted below, concentration varies considerably with land use. No attempt has been made to sort NPS basins on the basis of dominant land use, size, physiography or pedology although phosphorus transfer mechanisms are known to reflect these characteristics.

## DISCUSSION

The temporal trends of Figure 2 (Table 9) indicate a marked decline in phosphorus concentration for all of Lake Ontario data, for NPS basins both in the Lake Erie (+ Connecting Waterways) and the Lake Huron-Georgian Bay groups, and for PS soluble phosphorus in both the latter groups. The total phosphorus trends for Lake Erie (+ Connecting Waterways) drainage are less clear, but they do suggest a declining trend in the period 1972-74 which could arguably be influenced by the imposition of the 1974 effluent phosphorus removal program. Nevertheless, the reduction of some 53 to 73% in concentrations of both total and soluble phosphorus forms in NPS basins is comparable to values found by Sweeney (1973) for point source drainage. One might look for changes in reporting procedure (e.g., $PO_4$ to P) which would cause an apparent reduction in load. These data, however, all employ phosphorus reported as P. No significant changes in monitoring strategy occurred over the record period. There is no evidence for a decline[1] in agricultural use of phosphorus (Statistics Canada).

---

[1] I am indebted to Dr. D.R. Coote for procuring this information.

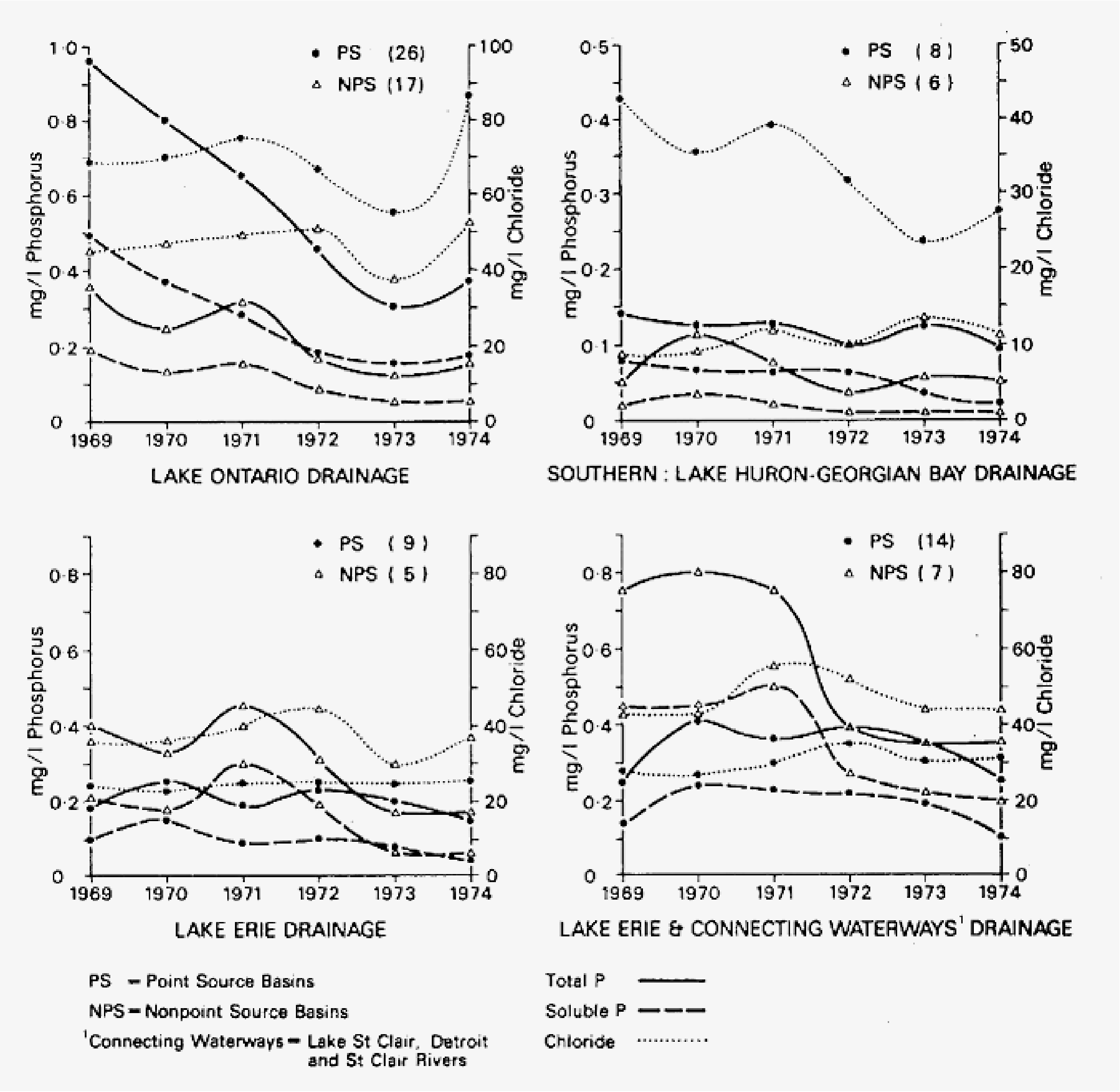

**Figure 2:** Phosphorus and chloride trends for the period 1969-1974.

**TABLE 9:** ANNUAL MEAN CONCENTRATIONS OF PHOSPHORUS AND CHLORIDE (mg/L) FOR TRIBUTARY DRAINAGE[a]

| | Point Source Basins | | | | | | | Nonpoint Source Basins | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N[b] | '69 | '70 | '71 | '72 | '73 | '74 | N[b] | '69 | '70 | '71 | '72 | '73 | '74 |
| Lake Ontario | | | | | | | | | | | | | | |
| Chloride | 26 | 68.8 | 70.6 | 75.5 | 66.4 | 55.5 | 86.5 | 17 | 45.5 | 47.4 | 48.9 | 50.0 | 37.4 | 52.4 |
| SRP | 26 | .492 | .367 | .279 | .177 | .146 | .170 | 17 | .194 | .134 | .152 | .082 | .053 | .053 |
| TP | 26 | .960 | .807 | .646 | .446 | .308 | .367 | 17 | .352 | .247 | .308 | .168 | .122 | .149 |
| Lake Erie | | | | | | | | | | | | | | |
| Chloride | 10 | 24.8 | 23.3 | 25.0 | 25.2 | 24.4 | 25.7 | 5 | 36.3 | 36.1 | 40.0 | 44.2 | 30.6 | 36.8 |
| SRP | 9 | .105 | .149 | .092 | .100 | .077 | .044 | 5 | .213 | .184 | .299 | .194 | .071 | .062 |
| TP | 9 | .176 | .251 | .186 | .229 | .198 | .152 | 5 | .398 | .329 | .452 | .312 | .172 | .174 |
| Lake Erie and Connecting Waterways | | | | | | | | | | | | | | |
| Chloride | 15 | 28.1 | 27.5 | 29.8 | 35.1 | 29.5 | 31.5 | 7 | 43.3 | 43.2 | 55.8 | 52.5 | 44.1 | 43.3 |
| SRP | 14 | .138 | .242 | .232 | .222 | .194 | .104 | 7 | .446 | .456 | .507 | .272 | .218 | .200 |
| TP | 14 | .245 | .415 | .359 | .393 | .352 | .252 | 7 | .752 | .795 | .745 | .397 | .355 | .356 |
| Southern Lake Huron and Georgian Bay | | | | | | | | | | | | | | |
| Chloride | 8[d] | 42.6 | 35.6 | 39.1 | 31.6 | 23.4 | 27.9 | 8 | 8.3 | 9.1 | 12.5 | 10.6 | 13.2 | 11.4 |
| SRP | 8 | .080 | .068 | .062 | .062 | .038 | .022 | 6 | .020 | .036 | .020 | .010 | .011 | .010 |
| TP | 8 | .140 | .127 | .126 | .106 | .127 | .095 | 6 | .050 | .116 | .073 | .038 | .058 | .050 |

[a] Data taken from basins having six years data over the period 1969 - 1974.

[b] N = Number of tributary basins.

[c] Tributary basins having some or all of their area in agricultural Southern Ontario.

[d] Profoundly influenced by salt mining in Maitland River

The decline in phosphorus concentration values in NPS basins can be evaluated relative to rainfall/runoff variability by comparison with chloride (Figure 2). This chemically conservative substance, although considerably influenced by anthropogenic sources and having different transfer mechanisms than phosphorus, reflects climatically induced chemical flux from NPS watersheds. The profound changes in ratios of soluble and total phosphorus to chloride (Table 10) reveals that both phosphorus forms have undergone a marked reduction which cannot be attributed to variations in naturally occurring transfer mechanisms.

Phosphorus removal in municipal effluent in 1974 for Lake Erie (+ Connecting Waterways) is clearly illustrated (Figure 2) by the reduction in concentration to values approaching those of 1969. The Lake Erie (+ Connecting Waterways) point source data are curious both because the concentration values are substantially less than those for NPS basins and because these data do not indicate the reduction seen in Lake Ontario drainage nor that observed in Lake Erie NPS drainage. The failure of detergent regulations to cause a visible downward trend suggests an increasing use of detergent and/or unregulated phosphorus cleaning compounds in the Lake Erie area. The lower concentrations in PS systems relative to NPS basins is thought to reflect the method of grouping basins into the PS and NPS categories. For example, as reported below, it has been found that urban diffuse drainage severely affects stream levels of phosphorus.

Where urban centres are adjacent to tributary mouths such as for much of Lake Ontario drainage, phosphorus concentration values are twice those found in Lake Erie drainage where urban centres are a considerable distance upstream. In addition, PS basins for Lake Erie (+ Connecting Waterways) include many small rural watersheds which are classified as PS due to the presence of an industrial outfall but which is phosphorus-free. In the absence of urban diffuse runoff and with low-intensity agricultural land use, phosphorus concentrations are depressed. Because data are intentionally not weighted, the Grand and Thames River systems which dominate PS tributary flow to Lake Erie and Lake St. Clair respectively, are averaged together with very low concentration values characteristic of essentially rural PS basins. The affect of considerably reduced population and agricultural potential upon phosphorus concentrations can be seen for both PS and NPS groups within Huron and Georgian Bay drainage.

The effect of phosphorus legislation in PS basins relative to the NPS group is most easily seen in a PS/NPS ratio using normalized annual mean concentration values by lake group--that is, each annual mean concentration (Table 9) is divided by the 1969 value. A ratio of unity (Table 11) indicates that PS basins are behaving collectively in a manner similar to NPS basins. Ratios exceeding unity indicate a relatively poorer performance by PS basins whereas values less than unity reflect better performance by the PS basins. Table 11 indicates that with the exception of the Lake Huron and Georgian Bay group, the hypothesis that PS basins should reflect better performance in terms of phosphorus reduction consequent to detergent legislation is not substantiated.

**TABLE 10:** PHOSPHORUS/CHLORIDE x $10^2$ RATIOS FOR NONPOINT SOURCE BASINS

| | Year | | | | | |
|---|---|---|---|---|---|---|
| | '69 | '70 | '71 | '72 | '73 | '74 |
| Lake Ontario | | | | | | |
| SRP/Chloride | .43 | .28 | .31 | .16 | .14 | .10 |
| TP/Chloride | .77 | .52 | .63 | .34 | .33 | .28 |
| Lake Erie | | | | | | |
| SRP/Chloride | .59 | .51 | .75 | .44 | .23 | .17 |
| TP/Chloride | 1.10 | .91 | 1.13 | .71 | .56 | .47 |
| Lake Erie and Connecting Waterways | | | | | | |
| SRP/Chloride | 1.03 | 1.06 | .91 | .52 | .49 | .46 |
| TP/Chloride | 1.74 | 1.84 | 1.33 | .76 | .80 | .82 |
| Lake Huron | | | | | | |
| SRP/Chloride | .24 | .40 | .16 | .09 | .08 | .08 |
| TP/Chloride | .60 | 1.28 | .59 | .36 | .44 | .44 |

[a] Data taken from basins having six years data over the period 1969 - 1974.

## **TABLE 11:** POINT SOURCE/NONPOINT SOURCE RATIOS OF NORMALIZED ANNUAL MEANS [a]

| | Year | | | | | |
|---|---|---|---|---|---|---|
| | '69 | '70 | '71 | '72 | '73 | '74 |
| Lake Ontario | | | | | | |
| SRP | 1 | 1.08 | .72 | .85 | 1.09 | 1.27 |
| TP | 1 | 1.20 | .77 | .98 | .93 | .90 |
| Chloride | 1 | .99 | 1.02 | .88 | .98 | 1.09 |
| Lake Erie | | | | | | |
| SRP | 1 | 1.64 | .62 | 1.05 | 2.20 | 1.44 |
| TP | 1 | 1.72 | .93 | 1.66 | 2.60 | 1.98 |
| Chloride | 1 | .95 | .92 | .83 | 1.17 | 1.02 |
| Lake Erie and Connecting Waterways | | | | | | |
| SRP | 1 | 1.72 | 1.48 | 2.64 | 2.88 | 1.68 |
| TP | 1 | 1.60 | 1.48 | 3.04 | 3.04 | 2.18 |
| Chloride | 1 | .98 | 1.54 | 1.54 | 1.54 | 1.54 |
| Lake Huron | | | | | | |
| SRP | 1 | .47 | .78 | 1.55 | .86 | .55 |
| TP | 1 | .39 | .62 | l.00 | .78 | .68 |
| Chloride | 1 | 1.04 | .71 | .91 | .69 | .67 |

[a] Data are from Table 9.

Although the decline in phosphorus concentrations observed in some PS basin groups can be explained in terms of phosphorus abatement programs, the widespread and marked decline in NPS basins is perplexing. Although data aggregated both in time and space and in terms of human and physical influences within drainage systems must be approached with considerable caution, the reduction of phosphorus concentrations in some PS groups and virtually all NPS groups concurrent with a phased reduction in detergent-related phosphorus suggests that both PS and NPS groups are responding to the same measures. If this proves to be true, it carries profound implications for assessing the spatial impact of detergent controls. Examining individual tributary basins, it is clear that NPS watersheds have varying responses over the study period. Although certainly not causally conclusive, discriminant and correlation analyses indicate that stream-mouth levels of phosphorus respond to cropland and urban diffuse variables. Although a positive relationship exists between concentration and basin area in cropland, the presence of even a small amount of diffuse urban runoff profoundly affects phosphorus levels and obscures any relationship between phosphorus and the cropland variable. It is suggested, therefore, that explanation for reductions in stream-mouth levels of phosphorus due to a detergent abatement program should first be sought in urban diffuse source systems, followed by linkages with agriculture. It is also possible that septic tank leakage and the presence of small private (commercial) point sources not included in lists of conventional point sources may affect stream- mouth levels of phosphorus.

Assuming that it is detergent legislation which has not been selectively effective in PS relative to NPS basins, it may be argued that:

(1) additional sources (either unregulated phosphorus or increased detergent use) of phosphorus have been made available in the PS group to offset legislated reductions in detergent-related phosphorus, or

(2) the relative importance of diffuse sources within PS watersheds overshadows any reduction in municipal treatment plant loads due to runoff effects in the wetter 1972-74 period.

It is this writer's opinion that the former argument may indeed apply to limited numbers of watersheds but does not obscure the general downward trend when large numbers of basins are considered (e.g., Lake Ontario group). The second argument is not supported by data from predominantly agricultural watersheds such as those of southwestern Lake Erie drainage where, despite large phosphorus loads, there has been a reduction in the 1972-74 period commensurate with that seen in other NPS and PS basins.

## CONCLUSIONS

The division of southern Ontario drainage to the Great Lakes into those basins having no point sources (NPS) and those with point sources (PS) has permitted an evaluation of the effectiveness of phased introduction of phosphorus control through legislated detergent and municipal effluent standards. In spite of the range of land uses contained within the NPS group, it was hypothesized that NPS basins should reflect annual concentrations of soluble and total phosphorus which, although commensurate with land use patterns and hydrologic activity, are unrelated to detergent controls and effluent abatement programs. The hypothesis that river-mouth phosphorus concentration levels in NPS basins could be used as a standard against which to evaluate the performance of PS basins in response to phosphorus abatement programs, is not substantiated. Although there has been a reduction of up to 72% in phosphorus levels in PS basins, there has been a comparable decline in the NPS group. Using chloride for comparison, the reduction in soluble and total phosphorus levels cannot be attributed to natural concentration/dilution effects.

Surprisingly, not only has there been no marked improvement in PS basins relative to NPS basins in the Lower Lakes but, in fact, the reverse tends to be true. While increasing use of phosphorus by industrial and commercial establishments may explain the failure of some PS basin groups to respond to abatement programs, the data suggest that detergent control has had a generally beneficial effect not only in PS basins but also in NPS basin groups. If this latter supposition is substantiated, legislation which affects domestic use of phosphorus at source has a wide-spread though probably a land use-specific beneficial affect. The alternative of legislation which controls effluent standards through phosphorus removal affects only those basins which receive point source loadings.

# SOURCE CONTRIBUTIONS TO 1974 RIVER-MOUTH PHOSPHORUS LOADS, ONTARIO

## INTRODUCTION

Although detergent legislation has been directed towards one of the most significant contributors to fluvial phosphorus budgets, the cost-effectiveness of further phosphorus control strategies, whether directed to effluent standards or to diffuse source control within tributary watersheds depends to a large measure on the relative importance of various phosphorus sources. Such relativity is, however, exceedingly difficult to accurately determine. Point source data are generally limited to municipal waste loadings. In Ontario, industrial sources of phosphorus to river systems are considered to be minimal.[1] A simple mass balance calculation of diffuse phosphorus wherein point source phosphorus is deducted from river-mouth loadings is complicated by the now well-known (but not readily quantifiable) transmission losses of point and diffuse sources of phosphorus during downstream transport through chemical, biological and sedimentary processes (Baker and Kramer, 1975; Logan, 1977).

Of the phosphorus originating from diffuse sources, some is background and non-controllable, some is due to a variety of anthropogenic influences, and a further fraction is attributable to loadings from atmospheric sources. In the latter two cases the phosphorus pathways from source to stream are complex and not fully understood. Therefore, in the discussion below, atmospheric contributions should be considered only as potential loads and their effect can be considered as part of background load calculations. The evaluation of phosphorus loads in this report is restricted to those loads transmitted from stream mouths of Canadian tributary drainage. Although data are tabulated by Lake and Connecting Waterways, analysis and discussion is restricted to Lakes Ontario and Erie plus Connecting Waterways[2] where phosphorus generation, transmission and fate has historically had a significant impact on trophic levels. All significant tributary drainage to the Lower Lakes is included (Figure 3). Measured river-mouth loads include all upstream municipal effluent to the extent these are transmitted to the mouth, and virtually all diffuse phosphorus sources with the exception of direct-to-Lake atmospheric deposition and groundwater discharge into the Lakes.

---

[1] Mr. J.G. Ralston, Ontario Ministry of the Environment, personal communication.

[2] Tributary drainage to Lake St. Clair, and to the St. Clair and Detroit Rivers.

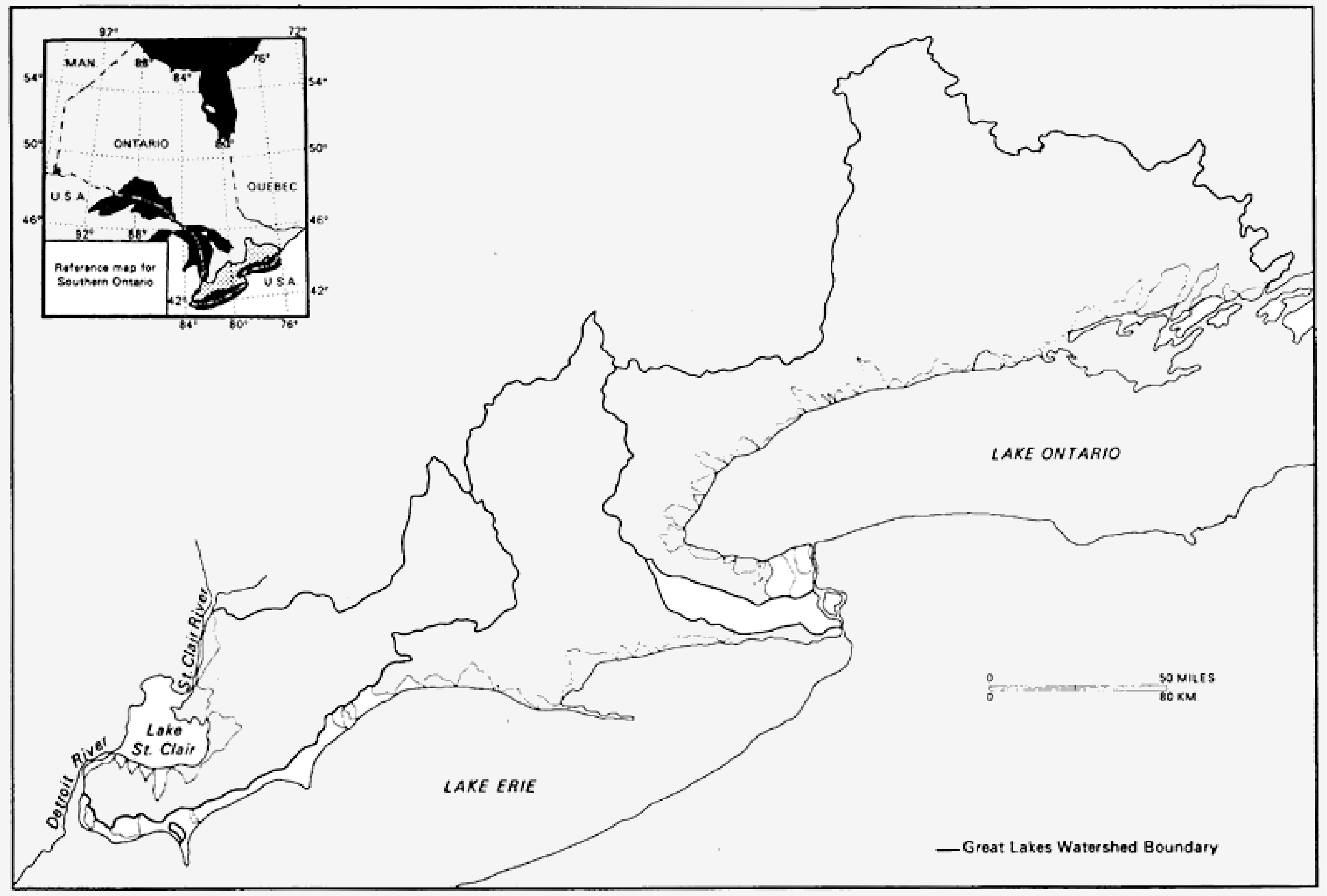

**Figure 3:** Tributary drainage to the Lower Great Lakes and Connecting Waterways (Lake St. Clair, Detroit and St. Clair Rivers) for which there are phosphorus data, 1974.

River-mouth loads should not be confused with total load of phosphorus to the Lakes which includes tributary, direct-to-Lake diffuse and direct-to-Lake point loads from municipal and industrial outfalls. Direct-to-Lake loads are, for Lake Ontario, particularly large, therefore any cost-effective arguments for diffuse versus point control must eventually include these.

The choice of 1974 for this study was predicated by available data. A realistic appraisal of relative load contribution by tributary and direct-to-Lake discharges must await availability of 1976 data which will reflect the imposition of a 1.0 mg/L phosphorus effluent standard for all direct-to-Lake Ontario municipal effluent discharges which exceed one million gallons per day.

As illustrated in Section 6 (above) annual mean phosphorus concentration values at river-mouth locations respond to land use types within drainage basins. The declining trend in concentration values of Figure 2 represents a net reduction in basin loads to the Lakes which appear to be a response to detergent legislation phased in over the period 1970-1973 and the imposition in 1974 of a 1.0 mg/L municipal effluent standard for all treatment facilities in Lake Erie and Connecting Waterways' drainage and for the Trent River system within the Lake Ontario watershed (Tables 5 and 6). Phosphorus levels in river-mouth locations appear to level off in 1974 and, although no later data are available, appear to be representative of phosphorus concentrations in Canadian tributary drainage of the Lower Lakes following legislated reductions in phosphorus loadings. It is recognised, of course, that a single year is not necessarily representative of total pollutant levels as diffuse contribution varies considerably from year to year depending upon human interference and hydrologic behaviour; therefore, the values recorded below are merely diagnostic and not absolute. Although 1974 was the third reasonably wet year in succession (Table 8), phosphorus concentration trends do not appear to be significantly affected by above-average runoff.

## ATMOSPHERIC PHOSPHORUS

Atmospheric loadings for total phosphorus and selected other contaminants in 1974 were modelled for the Lower Lakes watershed by Acres Consulting Services Ltd. (1977) under contract to PLUARG. Actual loadings from precipitation and bulk samples were compared with estimated loadings calculated from source emission data and air trajectories. Although some uncertainties remain concerning the interpretation of the precipitation chemistry (Dr. F.C. Elder, personal communication), the results are the best estimates currently available. For the purpose of atmospheric loadings estimates, the Acres study divided the Great Lakes drainage basin (land area) into 37 subareas and the loadings calculated using the Thiessen polygon weighting technique. For our purposes here, tributary basin boundaries were overlain on the Acres model boundaries and the phosphorus loading per model subarea prorated into each drainage basin. The potential atmospheric loadings per watershed appear in Table 12.

RIVER-MOUTH PHOSPHORUS LOADS

River-mouth loads of phosphorus are generally calculated [1] in this study for individual rivers by the equation

$$L = \sum_{i=1}^{12} C_i \bullet \bar{Q}i \bullet k$$

where L is annual load, C is monthly mean concentration, Q is monthly mean discharge and k is a constant. Data are drawn from surveillance records of the Ontario Ministry of the Environment which represent some ten to twenty samples per site year for more than 1000 sites across the province. Samples have been analysed for phosphorus with auto-analyses facilities since 1969, therefore data precision is good.

Annual loads calculated from surveillance data of this type may underestimate, perhaps considerably so, actual phosphorus loads. Firstly, such data may fail to accurately portray concentration/dilution effects for the range of discharge conditions in any one year, in which case C and Q are not appropriately related. Secondly, surveillance data although providing good interbasin comparisons, seriously underestimate absolute annual loads of total solids due to the non-linear and discharge-dependent nature of sediment-discharge relationships[2]. For similar reasons, larger nutrient loads which have been found under high runoff conditions (Baker and Kramer, 1973; U.S. Army Corps of Engineers, 1975) may be under-represented in surveillance data. Although all but the very highest discharges are represented over a number of years of surveillance data (Ongley *et al.*, 1977), individual storm events are usually not represented. Moreover, because 30 to 95% of total phosphorus has been reported to be associated with particulates (Ongley, 1976; Sweeney, 1973; Logan, 1977), any error in characterizing the particulate load may accrue to particulate-related phosphorus. However, errors in particulate load estimation can not be simply transferred to phosphorus because the relationships amongst discharge, phosphorus, organic solids, mineral sediment and their various and often pronounced seasonal variations in Great Lakes tributary fluvial systems are complex and poorly understood. Nevertheless, soluble/total ratios are similar to those reported for phosphorus in the Great Lakes literature.

With present data, therefore, it is not possible to state an error term for calculated phosphorus loads. Although such loads may differ somewhat from those published by the International Joint Commission, both are drawn from the same data set and reflect different computational procedures

---

[1] A full explanation of the load estimation procedure is found in Ongley, 1974.

[2] Ongley *et al.*, 1977, discuss several loading estimates and errors due to surveillance-type data for suspended solids.

and, implicitly, different assumptions within the estimating equations. For example, 1972 total and soluble phosphorus loads for Canadian tributaries to Lake Ontario as reported by Casey and Salbach (1974) are 26 and 12.6 percent higher than loads calculated for this study. Both are drawn from the same data set but reflect the difference between use of mean annual and cumulative mean monthly procedures. Using an untuned loading model, Chapra (1977) reports remarkable close agreement with published total phosphorus loads. Such comparison could, however, lead to circular argument and potentially erroneous conclusions about model or measured loads. It is beyond the scope of this report to examine methods of phosphorus load estimates and the errors inherent in surveillance data of the type collected routinely by government agencies. Nevertheless, potential for error must be borne in mind, particularly when estimates of diffuse contributions to total phosphorus load are derived from a mass balance approach in which total load is only a best-available estimate.

## MUNICIPAL WASTE LOADS

Point source total phosphorus loads of Tables 12 and 13 are those provided by the Ontario Ministry of the Environment to the International Joint Commission from municipal waste treatment plants which exceed one million gallons per day to tributary waterways. Loadings are summations of municipal discharges where, generally, individual plant loads are calculated from average (measured) discharge and phosphorus concentrations. Inclusion of loadings from smaller treatment plants may raise the values of Tables 12 and 13 a small amount. Actual municipal loads to the Lakes are undoubtedly less than reported due to transmission losses of phosphorus where treatment facilities are some considerable distance upstream.

## DISCUSSION

The maximum potential contribution of atmospheric sources to river-mouth loads of total phosphorus in the Lower Lakes is some 10 to 13 percent (Table 12) and rises as one moves into the Upper Lakes, reflecting the reduction in local anthropogenic loads. The proportion of atmospheric to total river-mouth load may rise slightly in Lake Ontario in 1976 when municipal loads to all tributaries comply with the 1.0 mg/L effluent standard. Assuming complete transmission of municipal waste loadings of phosphorus to river mouths, municipal point sources (Table 12) account for one-third of tributary phosphorus loads for Lake Ontario, about 12 percent for Lake Erie (plus Connecting Waterways) and much less than 10 percent for the Upper Lakes.[1] To what extent industrial phosphorus is included in the non-municipal component of river-mouth loads of, in particular, the Upper Lakes, is not known to this writer. Therefore, any mass balance computations from Table 12 of diffuse source contribution to Upper Lakes tributary phosphorus loads are probably very much in error.

---

[1] Proportionally less if river-mouth phosphorus loads are underestimated by surveillance data.

Until 1976 data are available which include full implementation of the 1.0 mg/L effluent standard for Lake Ontario data, it is difficult to accurately compare the relative importance of municipal effluent to river-mouth loads of Erie and Ontario drainage. Nevertheless, the bulk of municipal effluent for the urban conurbations of the western end of Lake Ontario is direct-to-Lake, and official projections of municipal loadings to tributaries in 1976 suggest a reduction of point source contribution in Table 13 to approximately 27 percent. It is, therefore, considered that the 1974 value of 285 tonnes (Tables 12 and 13) is a reasonable first approximation for comparative purposes, particularly when considering uncertainty attributed to downward bias due to transmission losses within the fluvial system.

Assuming complete transmission both of point and atmospheric sources of phosphorus, a mass balance calculation (Table 12) indicates that between 67 and 88 percent of tributary phosphorus loads to the Lower Great Lakes is attributable to diffuse sources in Southern Ontario. The larger amount for Lake Erie (plus Connecting Waterways) relative to Lake Ontario, although due in small measure to compliance of all Erie municipal treatment plants to the 1.0 mg/L standard, undoubtedly reflects the greater influence of agriculture relative to urban variables.

The difference between diffuse contribution to phosphorus loads to Lakes Erie and Ontario becomes even more apparent when one considers background, essentially non-controllable [1] levels of phosphorus. The calculation of background levels is somewhat speculative. However, using a criterion of constant annual mean concentration for the period 1969-1974, five small and essentially nonagricultural rural watersheds without point sources in Southern Ontario provide representative background concentrations (which would include any atmospheric imputs) of total phosphorus averaging 0.044 mg/L. This is virtually identical to values found independently by Sweeney (1973) in Erie County, New York. Obviously there is no true nor single background value for tributary basins in Southern Ontario which can be obtained from averaged downstream monitoring locations in an area as geomorphologically and geologically diverse as this (see, for example, Dillon and Kirchner, 1975). The only justification for using the criterion of constancy of concentration values over the period 1969-74 is the observation that basins with significant agricultural and/or urban land uses illustrate phosphorus concentrations at river-mouth locations which are generally at least an order of magnitude or more higher than background and which display a downward trend (Figure 2) due, presumably, to phosphorus abatement programs.

Applying the background concentration to total tributary discharge in 1974 to Lakes Ontario and Erie (plus Connecting Waterways), the estimated background contribution to total tributary phosphorus flux is 374 and 287 tonnes (43 and 22 percent of total) respectively to each Lake (Table 13). These values will, of course, change from year to year in proportion to runoff, however such

---

[1] Except insofar as atmospheric emissions, which are reflected in background flux, are reduced in future.

## **TABLE 12:** SUMMARY BY LAKE OF POTENTIAL MUNICIPAL AND ATMOSPHERIC CONTRIBUTION TO RIVER-MOUTH LOADS OF TOTAL PHOSPHORUS, 1974

(metric tonnes)

| Drainage tributary to: | 1 Atmosphere | 2 Municipal Waste Loads | 3 Tributary Load to Lake | Ratio ⅓ | Ratio ⅔ |
|---|---|---|---|---|---|
| Lake Ontario | 116.3 | 284.8 | 873.7 | .133 | .326 |
| Niagara River | 0.9 | -- | 6.3 | .143 | -- |
| Lake Erie | 59.3 | 156.0 | 604.7 | .098 | .122 |
| Lake St.Clair, Detroit and St.Clair Rivers | 59.2 | | 672.4 | .088 | |
| South: Lake Huron and Georgian Bay[a] | 102.4 | -- | 568.5 | .180 | -- |
| Total: Lake Huron and Georgian Bay | 327.8 | 92.5 | 1469.2 | .223 | .063 |
| St. Mary's River | 8.2 | -- | 23.9 | .343 | -- |
| Lake Superior | 139.2 | 39.0 | 1091.8 | .127 | .036 |

[a] Basins Numbers 90 - 115 (Table 2)

Data are from Acres Consulting Services Ltd.,1977; Ontario Ministry of the Environment; Ongley, 1977a

**TABLE 13:** SOURCE CONTRIBUTIONS TO TRIBUTARY TOTAL PHOSPHORUS LOADS, 1974

(metric tonnes)

| | Lake Ontario | | Lake Erie and Connecting Waterways[a] | |
|---|---|---|---|---|
| Load to Lake | 873 | 100.0 % | 1277 | 100.0% |
| Point Sources [b] | 285 | 32.6 % | 156 | 12.2% |
| Diffuse Sources: | | | | |
| Atmosphere | (116) | (13.3 %) | (119) | (9.3 %) |
| Background (estimated)[c] | 374 | 42.8 % | 287 | 22.4% |
| Above-background [d] (non-atmospheric) | 215 | 24.6 % | 834 | 65.3% |

[a] Drainage to Lake St.Clair, and to the St.Clair and Detroit Rivers.

[b] Municipal treatment plants; assumes 100% transmission to river mouth.

[c] Includes atmospheric component. Background phosphorus is estimated using phosphorus values from representative rural but essentially nonagricultural watersheds in Southern Ontario

[d] Residual after summation of other entries less atmospheric (contained in background values). Values increase if point source transmission is less than 100%.

amounts must figure in any formulation of control strategies for diffuse sources of phosphorus. The calculated proportion of 1974 river-mouth loads attributable to above-background (and, presumably, potentially controllable) diffuse phosphorus sources is some 65 percent to Lake Erie and 25 percent to Lake Ontario--higher[1] if transmission loss of point source phosphorus is assumed. Using projected municipal loadings to Lake Ontario tributary drainage for 1976 to reflect general compliance with the 1.0 mg/L effluent standard, and reducing the load to Lake accordingly (assuming 100 percent transmission), the above-background non-atmospheric diffuse contribution to river-mouth phosphorus loads is estimated only at 30 percent.

A different approach is to use the 'residual from background' flux which identifies not only the maximum controllable phosphorus both from point and diffuse sources but also the maximum diffuse contribution if one assumes zero transmission of municipal effluent. Using this approach, potentially controllable phosphorus from all tributary sources as a proportion of river-mouth loads is 78 and 57 percent respectively for Lake Erie (plus Connecting Waterways) and Lake Ontario.

## CONCLUSIONS

Although estimates for 1974 diffuse phosphorus loads to the Lower Lakes are admittedly crude and not necessarily representative of long-term averages, such tables are useful when assessing cost-effective controls within a broader economic context, and particularly in respect to diffuse relative to point source control. Load reduction by diffuse control measures, when considering unit area cost, the area involved and the probability (taking into consideration inexact science) of a beneficial result, must be compared against potentially controllable loads. Their difference defines realistic thresholds of phosphorus flux which must be accepted if one is to maintain present land tenure patterns.

With additional information, one can obtain from Table 13 some geographic resolution of kinds of strategies which might be effective for diffuse source control. As noted below for nonpoint source basins, river-mouth concentrations of phosphorus are positively related to area of basin in cropland. However, when even a small proportion of basin area is subject to diffuse urban runoff, phosphorus levels increase dramatically and obscure agricultural effects. This suggests that for Lake Ontario, where large urban areas are situated adjacent to the Lake, diffuse control might initially focus upon urban drainage. But for Lake Erie where urban centres are generally some considerable distance upstream, diffuse source control will have to consider agricultural effects in addition to urban factors. Within a spatial context, such general considerations can only be translated into least-cost remedial programs when the effect upon the Lake of point and diffuse sources at varying distances upstream can be expressed as a distance-decay function for representative drainage in Southern Ontario.

---

[1] Up to 77.5% for Lake Erie and 57.5% for Lake Ontario if point source contribution is considered to be zero (0.0% transmission).

In evaluating phosphorus trends, the following general comments may assist future computations.

(1) Poorly documented transmission effects of point source effluents in rivers of Southern Ontario.

(2) Unknown (fluvial) transmission effects of phosphorus resulting from diffuse sources.

(3) It follows that future mass balance arguments which focus upon Great Lakes' water quality must take into account a distance-decay function in the calculation of effect upon the Lake of upstream point and diffuse sources. Although one may assume transmission losses from point sources, it is also equally probable that above-background, diffuse loads, do not equally impact the Lake but are biased towards downstream sources. Remedial programs should consider the distance-decay factor in order to reduce implementation costs.

(4) The mobility of atmospheric sources is not known and therefore the atmospheric contribution to river- mouth load is not easily assessed except by calculating background values.

(5) Remedial measures will have to consider geographic trends not only of point vs. diffuse contribution but also in terms of types of diffuse sources.

(6) Although it remains for others to assess the relative merits of point vs. diffuse control strategies in terms of appropriate phosphorus reduction, such consideration ought to include an assessment of these data on an individual basis basis for those watersheds, identified in Appendix 1 as having the worst phosphorus conditions.

# COVARIANCE IN VARIABLE SUBSETS

Before proceeding with further statistical analyses of the variables of Table 1, it is useful to select those variables which are likely to be meaningful in the development of land use-loadings relationships. After a judicious selection of variable subsets, the variables within each subset have been compared by correlation procedures in order to assess the degree to which they exhibit covariance. Not only is this information useful in assessing statistical results, but initially gives some insight into combinations of variables which, being statistically linked, may suggest linkages with specific land use activities.

In the analysis which follows, the 228 variables of Table 1 have been reduced to nine dependent and twenty-one independent variables (Table 14). The nine dependent variables include those variables most commonly associated with anthropogenic influence upon water quality (suspended solids; nutrients--TN, $NO_3$, TP, SRP; chloride and iron) and two general indicators (TDS and Hardness). The covariance amongst these variables is illustrated in Tables 15 and 16 in which mean annual values for concentration (C) and yield (Y) for the period 1968-72 reflect:

1) All 101 basins[1] of the data set,
2) 49 NPS[2] basins,
3) 52 PS[2] basins.

The correlation matrices are provided as indicators of covariance; statistical confidence in the correlation coefficients was not generally determined. These correlation matrices were used to identify potential relationships for which correlation and regression analysis (below) was pursued. Although not reproduced here, the correlation between C and Y for individual variables is virtually always very high (>0.85).

Comparing all 101 basins, the PS, and NPS basin data sets (Tables 15 and 16), all dependent variables correlate with TFe in the NPS set to a greater extent than for 101 or PS groups.

---

[1] As defined in Section 3.

[2] NPS and PS are defined in Section 4.

Generally, TFe is most closely associated with Suspended Solids for all data groups and with TN ($r = +0.65$) for the NPS group. TDS is associated with Chloride and Hardness in each of the three basic groups and to a moderate extent with TN, TP and SRP in the PS group. TN is closely correlated with TP and SRP in all basin groups and with NO3 and Suspended Solids for NPS basins; elsewhere it is not particularly related to other variables. TP is highly correlated with SRP in all groups ($r > 0.98$) and with TN especially in PS basins. Hardness and Chloride are related only to TDS in all groups, although Chloride and TN and TP are more closely related in PS than NPS basins. As noted below, this suggests that urban and industrial-related sources (point and/or diffuse) are responsible for increased amounts of nitrogen and phosphorus together with chloride.

The high degree of correlation of concentration with yield causes the same variable relationships noted above to apply to yield data (Table 16). Although not always so, the correlation coefficients for yields tend to be slightly lower than for concentration data.

The correlation matrices for the overburden-rural variables are provided for all 101 basins and for the NPS group (Table 17). Correlation coefficients are generally low, with expected negative relationships amongst cropland and woodland, and cropland and improved pasture. The overburden types (sand, loam, clay) are not meaningfully related to any of the other agricultural variables, probably because of the small numbers of basins with the majority of their area in a single overburden type.

The correlations amongst urban categories (Table 18) of the PS and NPS group offer a number of interesting comparisons. Although Large Urban is moderately to highly correlated with commercial, residential and transportation variables in NPS basins, the pattern is broken in the PS group by the apparent non-association of Low Density Commercial with Large Urban in PS basins, and, in the same group, the much more clearly defined trend between Transportation and Large Urban. Nevertheless, the large positive relationships amongst Large Urban and individual urban variables makes Large Urban a useful generalized variable both for PS and NPS basin groups in the identification of urban-related pollutant generators. Small Urban is unrelated to any of the individual variables for both PS and NPS groups rendering it unsuitable as a generalized variable.

As noted in the discussion (below) of discriminant analysis, the individual urban variables suggest a degree of specificity between land use types and pollutant generators which is probably unwarranted and misleading. Although there are basic similarities between PS and NPS groups amongst the correlations between various commercial, residential and transportation variables, the differences which do exist may reflect differences between internal urban structures of PS and NPS basins. An example is the observation that positive associations with NPS basins between Low, Medium and High Density Residential on the one hand, and Low and Medium Density Commercial on the other, are not reflected in the PS group.

**TABLE 14:** COVARIANCE IN VARIABLE SUBSETS: VARIABLES AND ABBREVIATIONS

| Dependent | (Mean annual concentration and yield for period 1968 - 1972) | Reference in Table 1. |
|---|---|---|
| TDS | Total dissolved solids | 2,3 : 197 |
| SS | Suspended solids | 32,33 : 197 |
| TN | Total nitrogen | 62 + 77+ 92,(63 + 78 +93): 197 |
| $NO_3$ | Nitrate nitrogen | 92,93: 197 |
| TP | Total phosphorus | 107, 108 : 197 |
| SRP | Soluble (reactive) phosphorus | 122, 123 : 197 |
| Ha | Hardness | 137,138: 197 |
| Chl | Chloride | 167, 168 : 197 |
| T.Fe | Total Iron | 182, 183 : 197 |
| Independent (Overburden-rural subset; % of basin area) | | |
| Sand | All sand categories | 198,202,205 |
| Loam | All loam categories | 199,203,206 |
| Clay | All Clay categories | 200, 204, 207 |
| Agr | Agriculture | 215 |
| Wood | Woodland | 216 |
| Orch | Orchard and horticulture | 225 |
| Crop | Cropland | 226 |
| I. Pas | Improved pasture | 227 |
| Un. Pas | Unimproved pasture | 228 |
| Independent (Urban subset; % of basin area) | | |
| Lg. Urb | Large Urban | 212 |
| S. Urb | Small Urban | 213 |
| L. Com | Low density commercial | 218 |
| M. Com | Medium density commercial | 219 |
| H. Com | High density commercial | 220 |
| L. Res | Low density residential | 221 |
| M. Res | Medium density residential | 222 |
| H. Res | High density residential | 223 |
| Tran | Transportation | 224 |

## **TABLE 15:** CORRELATION MATRICES FOR CONCENTRATION DATA

All (101)Basins Mean Annual Data for Period 1968-72

| | TDS | SS | TN | $NO_3$ | TP | SRP | Ha | Chl | TFe |
|---|---|---|---|---|---|---|---|---|---|
| TDS | 1.0000 | 0.3100 | 0.5495 | 0.4361 | 0.4473 | 0.4372 | 0.7803 | 0.8869 | 0.4009 |
| SS | 0.3100 | 1.0000 | 0.5290 | 0.5976 | 0.2881 | 0.2531 | 0.2316 | 0.2178 | 0.6972 |
| TN | 0.5495 | 0.5290 | 1.0000 | 0.5267 | 0.8902 | 0.8397 | 0.2570 | 0.5553 | 0.3917 |
| $NO_3$ | 0.4361 | 0.5976 | 0.5267 | 1.0000 | 0.2992 | 0.3017 | 0.4400 | 0.2589 | 0.4046 |
| TP | 0.4473 | 0.2881 | 0.8902 | 0.2992 | 1.0000 | 0.9838 | 0.1358 | 0.5090 | 0.1895 |
| SRP | 0.4372 | 0.2531 | 0.8397 | 0.3017 | 0.9838 | 1.0000 | 0.1365 | 0.4970 | 0.1508 |
| Ha | 0.7803 | 0.2316 | 0.2570 | 0.4400 | 0.1358 | 0.1365 | 1.0000 | 0.4462 | 0.2189 |
| Chl | 0.8869 | 0.2178 | 0.5553 | 0.2589 | 0.5090 | 0.4970 | 0.4462 | 1.0000 | 0.3814 |
| TFe | 0.4009 | 0.6972 | 0.3917 | 0.4046 | 0.1895 | 0.1508 | 0.2189 | 0.3814 | 1.0000 |
| NPS Basins | | | | | | | | | |
| TDS | 1.0000 | 0.3380 | 0.4839 | 0.5113 | 0.3329 | 0.3229 | 0.8542 | 0.8624 | 0.4356 |
| SS | 0.3380 | 1.0000 | 0.6971 | 0.6987 | 0.3101 | 0.2457 | 0.2245 | 0.2570 | 0.7969 |
| TN | 0.4839 | 0.6971 | 1.0000 | 0.6603 | 0.8134 | 0.7651 | 0.2609 | 0.4868 | 0.6491 |
| $NO_3$ | 0.5113 | 0.6987 | 0.6603 | 1.0000 | 0.2523 | 0.2454 | 0.4817 | 0.3312 | 0.5906 |
| TP | 0.3329 | 0.3101 | 0.8134 | 0.2523 | 1.0000 | 0.9915 | 0.0960 | 0.4079 | 0.3439 |
| SRP | 0.3229 | 0.2457 | 0.7651 | 0.2454 | 0.9915 | 1.0000 | 0.0947 | 0.3944 | 0.2884 |
| Ha | 0.8542 | 0.2245 | 0.2609 | 0.4817 | 0.0960 | 0.0947 | 1.0000 | 0.5186 | 0.2213 |
| Chl | 0.8624 | 0.2570 | 0.4868 | 0.3312 | 0.4079 | 0.3944 | 0.5186 | 1.0000 | 0.4316 |
| Tfe | 0.4356 | 0.7969 | 0.6491 | 0.5906 | 0.3439 | 0.2884 | 0.2213 | 0.4316 | 1.0000 |
| PS Basins | | | | | | | | | |
| TDS | 1.0000 | 0.2798 | 0.5852 | 0.3518 | 0.5081 | 0.5014 | 0.7404 | 0.9163 | 0.3539 |
| SS | 0.2798 | 1.0000 | 0.4362 | 0.4483 | 0.3004 | 0.2793 | 0.2521 | 0.1964 | 0.6344 |
| TN | 0.5852 | 0.4362 | 1,0000 | 0.4447 | 0.9198 | 0.8709 | 0.2883 | 0,5693 | 0.2334 |
| $NO_3$ | 0.3518 | 0.4483 | 0.4447 | 1.0000 | 0.3349 | 0.3449 | 0.4021 | 0.1982 | 0.2272 |
| TP | 0.5081 | 0.3004 | 0.9198 | 0.3349 | 1.0000 | 0.9819 | 0.1884 | 0.5349 | 0.0893 |
| SRP | 0.5014 | 0.2793 | 0.8709 | 0.3449 | 0.9819 | 1.0000 | 0.1890 | 0.5307 | 0.0534 |
| Ha | 0.7404 | 0.2521 | 0.2883 | 0.4021 | 0,1884 | 0.1890 | 1.0000 | 0.4319 | 02419 |
| Chl | 0.9163 | 0.1964 | 0.5693 | 0.1982 | 0.5349 | 0.5307 | 0.4519 | 1.0000 | 0.3317 |
| Tfe | 0.3569 | 0.6344 | 0.2334 | 0.2272 | 0.0893 | 0.0534 | 0.2419 | 0.3317 | 1.0000 |

## **TABLE 16:** CORRELATION MATRICES FOR YIELD DATA

All (101) Basins — Mean Annual Data for Period 1968-72

| | TDS | SS | TN | $NO_3$ | TP | SRP | Ha | Chl | TFe |
|---|---|---|---|---|---|---|---|---|---|
| TDS | 1.0000 | 0.3090 | 0.5188 | 0.3787 | 0.4109 | 0.4054 | 0.7312 | 0.7957 | 0.3020 |
| SS | 0.3090 | 1.0000 | 0.4730 | 0.4739 | 0.2568 | 0.2295 | 0.1938 | 0.2202 | 0.6260 |
| TN | 0.5188 | 0.4730 | 1.0000 | 0.5128 | 0.8665 | 0.8096 | 0.1716 | 0.5463 | 0.3815 |
| $NO_3$ | 0.3787 | 0.4739 | 0.5128 | 1.0000 | 0.2448 | 0.2679 | 0.3075 | 0.1795 | 0.3273 |
| TP | 0.4109 | 0.2568 | 0.8665 | 0.2448 | 1.0000 | 0.9799 | 0.0598 | 0.5165 | 0.2012 |
| SRP | 0.4054 | 0.2295 | 0.8096 | 0.2679 | 0.9799 | 1.0000 | 0.0716 | 0.4965 | 0.1669 |
| Ha | 0.7312 | 0.1938 | 0.1716 | 0.3075 | 0.0598 | 0.0716 | 1.0000 | 0.2517 | 0.0765 |
| Chl | 0.7957 | 0.2202 | 0.5463 | 0.1795 | 0.5165 | 0.4965 | 0.2517 | 1.0000 | 0.3601 |
| TFe | 0.3020 | 0.6260 | 0.3815 | 0.3273 | 0.2012 | 0.1669 | 0.0765 | 0.3601 | 1.0000 |
| NPS Basins | | | | | | | | | |
| TDS | 1.0000 | 0.2646 | 0.3269 | 0.4047 | 0.2178 | 0.2339 | 0.7648 | 0.7181 | 0.3034 |
| SS | 0.2646 | 1.0000 | 0.6778 | 0.7066 | 0.3005 | 0.2546 | 0.0831 | 0.1718 | 0.7764 |
| TN | 0.3269 | 0.6778 | 1.0000 | 0.5470 | 0.8184 | 0.7677 | 0.0648 | 0.3951 | 0.6131 |
| $NO_3$ | 0.4047 | 0.7066 | 0.5470 | 1.0000 | 0.1593 | 0.1852 | 0.3502 | 0.1733 | 0.5445 |
| TP | 0.2178 | 0.3005 | 0.8184 | 0.1593 | 1.0000 | 0.9866 | -0.0182 | 0.3545 | 0.2960 |
| SRP | 0.2339 | 0.2546 | 0.7677 | 0.1852 | 0.9866 | 1.0000 | 0.0139 | 0.3371 | 0.2488 |
| Ha | 0.7648 | 0.0831 | 0.0648 | 0.3502 | -0.0182 | 0.0139 | 1.0000 | 0.1731 | -0.0160 |
| Chl | 0.7181 | 0.1718 | 0.3951 | 0.1733 | 0.3545 | 0.3371 | 0.1731 | 1.0000 | 0.3510 |
| TFe | 0.3034 | 0.7764 | 0.6131 | 0.5445 | 0.2960 | 0.2488 | -0.0160 | 0.3510 | 1.0000 |
| PS Basins | | | | | | | | | |
| TDS | 1.0000 | 0.3555 | 0.6291 | 0.3835 | 0.5176 | 0.5085 | 0.7617 | 0.0511 | 0.3.122 |
| SS | 0.3555 | 1.0000 | 0.3489 | 0.3107 | 0.2266 | 0.2060 | 0.3734 | 0.2515 | 0 .5111 |
| TN | 0.6291 | 0.3389 | 1.0000 | 0.4779 | 0.8839 | 0.8294 | 0.3441 | 0.5943 | 0.2463 |
| $NO_3$ | 0.3835 | 0.3107 | 0.4779 | 1.0000 | 0.2596 | 0.2921 | 0.4039 | 0.1556 | 0.1817 |
| TP | 0.5176 | 0.2266 | 0.8839 | 0.2596 | 1.0000 | 0.9709 | 0.1550 | 0.5712 | 0.1319 |
| SRP | 0.5005 | 0.2060 | 0.8294 | 0.2921 | 0.9789 | 1.0000 | 0.1766 | 0.5625 | 0.1012 |
| Ha | 0.7617 | 0.3734 | 0.3441 | 0.4039 | 0.1850 | 0.1766 | 1.0000 | 0.3992 | 0.2311 |
| Chl | 0.8511 | 0.2515 | 0.5943 | 01556 | 0.5712 | 0.5625 | 0.3992 | 1.0000 | 0.3519 |
| TFe | 0.3122 | 0.5111 | 0.2463 | 0.1817 | 0.1319 | 0.1012 | 0.231.1 | 0.3519 | 1.0000 |

**TABLE 17:** CORRELATION MATRICES FOR OVERBURDEN-RURAL VARIABLES

All (101) Basins

| | Σ Sand | Σ Loam | Σ Clay | Agr. | Wood. | Orch. | Crop. | I.Pas. | Un.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| Σ Sand | 1.0000 | -0.1512 | -0.5796 | -0.2471 | 0.4805 | -0.2142 | 0.0245 | -0.1788 | 0.0454 |
| Σ Loam | -0.1512 | 1.0000 | -0.6375 | -0.1884 | 0.1589 | 0.1155 | -0.4190 | 0.2478 | 0.1792 |
| Σ Clay | -0.5796 | -0.6375 | 1.0000 | 0.4057 | -0.5897 | 0.0488 | 0.4096 | -0.1190 | -0.1596 |
| Agr. | -0.2471 | -0.1884 | 0.4057 | 1.0000 | -0.4883 | 0.1774 | 0.6330 | 0.0683 | -0.0808 |
| Wood. | 0.4805 | 0.1589 | -0.5897 | -0.4883 | 1.0000 | -0.2914 | -0.4514 | 0.2715 | 0.0989 |
| Orch. | -0.2142 | 0.1155 | 0.0488 | 0.1774 | -0.2914 | 1.0000 | 0.0228 | -0.3556 | -0.2688 |
| Crop. | 0.0245 | -0.4190 | 0.4096 | 0.6330 | -0.4514 | 0.0228 | 1.0000 | -0.5802 | -0.4431 |
| I.Pas. | -0.1788 | 0.2478 | -0.1190 | 0.0603 | 0.2715 | -0.3556 | -0.5802 | 1.0000 | 0.2721 |
| Un.Pas. | 0.0454 | 0.1792 | -0.1596 | -0.0808 | 0.0989 | -0.2688 | -0.4431 | 0.2721 | 1.0000 |
| | | | | | | | | | |
| NPS Basins | | | | | | | | | |
| Σ Sand | 1.0000 | -0.1363 | -0.5209 | -0.2765 | 0.4471 | -0.2373 | 0.0226 | -0.0860 | 0.1204 |
| Σ Loam | -0.1363 | 1.0000 | -0.7054 | -0.2550 | 0.1534 | 0.1410 | -0.3832 | 0.1721 | 0.0960 |
| Σ Clay | -0.5209 | - 0.7054 | 1.0000 | 0.4806 | -0.5666 | 0.0286 | 0.4223 | -0.1549 | -0.1062 |
| Agr. | -0.2765 | 0.2550 | 0.4306 | 1.0000 | -0.6361 | 0.1922 | 0.6334 | -0.1390 | -0.0962 |
| Wood. | 0.4471 | 0.1534 | -0.5666 | 0.6361 | 1 .0000 | -0.3818 | 0.5672 | 0.4455 | 0.1602 |
| Orch. | -0.2373 | 0.1410 | 0.0286 | 0.1922 | 0.3818 | 1.0000 | 0.0712 | -0.4550 | -0.3799 |
| Crop. | -0.0226 | -0.3832 | 0.4223 | 0.6334 | 0.5672 | 0.0712 | 1.0000 | -0.6921 | -0.4361 |
| I. Pas . | -0.0860 | 0.1721 | -0.1549 | -0.1390 | 0.4455 | -0.4550 | -0.6921 | 1.0000 | 0.3082 |
| Un.Pas. | 0.1204 | 0.0960 | -0.1062 | -0.0962 | 0.1602 | -0.3799 | 0.4361 | 0.3082 | 1.0000 |

## **TABLE 18:** CORRELATION MATRICES FOR URBAN VARIABLES

All (101) Basins

| | Lg. Urb | S. Urb. | L. Com. | M. Com. | H. Com. | L. Res. | M. Res. | H. Res. | Tran. |
|---|---|---|---|---|---|---|---|---|---|
| Lg. Urb | 1.0000 | -0.0940 | 0.2034 | 0.6742 | 0.9070 | 0.5790 | 0.8707 | 0.7680 | 0.6701 |
| S. Urb. | -0.0940 | 1.0000 | 0.0206 | -0.1412 | -0.1296 | -0.0125 | -0.0308 | -0.0696 | -0.1785 |
| L.Com. | 0.2034 | 0.0206 | 1.0000 | 0.0633 | 0.1529 | 0.2813 | 0.1919 | 0.0200 | 0.2408 |
| M.Com. | 0.6742 | -0.1412 | 0.0633 | 1.0000 | 0.7647 | 0.1829 | 0.4711 | 0.4778 | 0.6014 |
| H. Com. | 0.9070 | -0.1296 | 0.1529 | 0.7647 | 1.0000 | 0.2824 | 0.6564 | 0.7502 | 0.7519 |
| L.Res. | 0.5790 | -0.0125 | 0.2819 | 0.1829 | 0.2824 | 1.0000 | 0.6822 | 0.2335 | 0.1844 |
| M. Res. | 0.8707 | -0.0308 | 0.1919 | 0.4711 | 0.6564 | 0.6822 | 1.0000 | 0.4303 | 0.4572 |
| H. Res. | 0.7680 | -0.0696 | 0.0200 | 0.4778 | 0.7502 | 0.2335 | 0.4383 | 1.0000 | 0.4256 |
| Tran. | 0.6701 | -0.1785 | 0.2408 | 0.6014 | 0.7519 | 0.1844 | 0.4572 | 0.4256 | 1.0000 |
| NPS Basins | | | | | | | | | |
| Lg.Urb | 1.0000 | 0.0291 | 0.7491 | 0.8554 | 0.8208 | 0.7645 | 0.9248 | 0.8006 | 0.5344 |
| S. Urb. | 0.0291 | 1.0000 | 0.0756 | 0.0256 | -0.0119 | -0.0117 | 0.0736 | 0.0614 | -0.0768 |
| L.Com. | 0.7491 | 0.0756 | 1.0000 | 0.9068 | 0.6930 | 0.4571 | 0.6597 | 0.7472 | 0.3101 |
| M.Com. | 0.8554 | 0.0256 | 0.9068 | 1.0000 | 0.9252 | 0.4439 | 0.6736 | 0.8849 | 0.5142 |
| H. Com. | 0.8208 | -0.0119 | 0.6930 | 0.9252 | 1.0000 | 0.3545 | 0.5788 | 0.8877 | 0.6274 |
| L.Res. | 0.7645 | -0.0117 | 0.4571 | 0.4439 | 0.3545 | 1.0000 | 0.8389 | 0.3535 | 0.1230 |
| M. Res. | 0.9248 | 0.0736 | 0.6597 | 0.6736 | 0.5788 | 0.8389 | 1.0000 | 0.5927 | 0.2837 |
| H. Res. | 0.8006 | 0.0614 | 0.7472 | 0.8049 | 0.8877 | 0.3535 | 0.5927 | 1.0000 | 0.5160 |
| Tran. | 0.5344 | -0.0768 | 0.3101 | 0.5142 | 0.6274 | 0.1230 | 0.2837 | 0.5160 | 1.0000 |
| PS Basins | | | | | | | | | |
| Lg. Urb. | 1.0000 | -0.2514 | 0.0791 | 0.6439 | 0.9334 | 0.5677 | 0.8467 | 0.7949 | 0.8058 |
| S. Urb. | -0.2514 | 1.0000 | -0.0164 | -0.2888 | -0.2794 | 0.0082 | -0.1742 | -0.1886 | -0.3305 |
| L.Com. | 0.0791 | -0.0164 | 1.0000 | -0.0512 | 0.0454 | 0.3088 | 0.0685 | -0.0601 | 0.2524 |
| M.Com. | 0.6439 | -0.2888 | -0.0512 | 1.0000 | 0.7368 | 0.1397 | 0.4315 | 0.4255 | 0.7434 |
| H. Com. | 0.9334 | -0.2794 | 0.0454 | 0.7368 | 1.0000 | 0.3368 | 0.6866 | 0.7421 | 0.9038 |
| L.Res. | 0.5677 | 0.0082 | 0.3088 | 0.1397 | 0.3368 | 1.0000 | 0.6341 | 0.3353 | 0.3029 |
| M. Res. | 0.8467 | -0.1742 | 0.0685 | 0.4315 | 0.6866 | 0.6341 | 1.0000 | 0.4308 | 0.6061 |
| H. Res. | 0.7949 | -0.1806 | -0.0601 | 0.4255 | 0.7421 | 0.3353 | 0.4308 | 1.0000 | 0.5024 |
| Tran. | 0.8058 | -0.3305 | 0.2524 | 0.7434 | 0.9038 | 0.3029 | 0.6061 | 0.5024 | 1.0000 |

# FACTOR ANALYSIS

## INTRODUCTION

Factor analytic methods are used in this study as a data description technique. The method used is that of principal components.

The difference between "true" factor analysis and principal components lies in the nature of the assumptions about the error terms. Factor analysis makes specific assumptions about the normality of errors and the covariances and develops a criterion for estimating the parameters of the model based on these assumptions. Principal components uses an arbitrary mathematical measure to achieve the same result. Since the purpose of using these methods is to obtain insight into the clustering of the observations and the variables, the differences between the two methods is not considered important. Efficient computational programs are available for principal components but are not readily available for factor analysis.

The factor analysis mode1 is described mathematically as follows. Let $\underline{X}$ be an nxp matrix. The rows represent different observations (basins) and the columns represent different variables. The factor model is

$$\underline{X} = \underline{F} \bullet \underline{L} + \underline{E}$$

where $\underline{F}$ is an nxm matrix of factor scores
$\underline{L}$ is an mxp matrix of factor loadings and
$\underline{E}$ is an nxp matrix of errors.

In the above model the $\underline{F}$ and $\underline{L}$ are estimated, then the $\underline{E}$ matrix can be estimated as the difference between observed and predicted values of the observations.

The benefit of this model only arises if the value of m is small, say 2 or 3. In this case the data have been reduced from a p-dimensional space to an m-dimensional space and if m is 2 then the observations can be plotted.

The model stated above is unfortunately not unique and consequently further conditions must be imposed. The usual conditions are that the columns of $\underline{F}$ are orthogonal and of unit length. With this condition a solution can be obtained (though it is still not unique).

The matrix of loadings can now be analysed. This matrix tells how the variables are related to the factors and if the loadings for each variable are plotted, then points which are close together indicate variables which are behaving similarly. To eliminate the influence of scale in this analysis, the variables have been transformed to have zero mean and unit standard deviation (Z values); the factors have the same properties.

Factor analysis can be considered successful if the E matrix is small and the number of factors (m) is also small. In this case an important reduction in the data occurs.

The method described above is generally called R mode. Its purpose is to analyse relations between variables. A second method is Q mode analysis. In this case the observations and variables are interchanged but otherwise the analysis proceeds in the same way. The two methods are mathematically equivalent but the interpretations are different; in one case the variables are clustered and in the other the observations are clustered. The Q mode analysis is computationally more difficult because of the very large matrices which arise.

## DISCUSSION

The factor analysis employs all 13 dependent and 33 independent variables listed in Table 1. The data represent average conditions over the period 1968 - 1972 for the 101 tributary basins of Southern Ontario described in Section 3. No attempt was made to remove the potential for grouping by presence/absence of point source by partitioning the basins into point and nonpoint subsets for this analysis. Of the dependent variables only mean annual concentration data are used. Table 19 identifies the abbreviations used in the illustrations below.

### R Mode Analysis

R mode analysis does not produce two or three dominant principal components. For example, with all 33 independent variables, the first two principal components account for only 30.1 percent of the variation in the data. The loadings (after varimax rotation) in the first two principal components are plotted in Figure 4. In this and other similar plots, one is able to identify clustering of variables. Each figure gives a two-dimensional approximation of the relative positions of the variables. The approximation is fairly good for those points (such as UAG) which lie far from the origins and very poor for those (e.g., MBUC) lying close to the origin. Therefore, one looks for clusters lying distant from the origin. In the ideal case where the first two principal components account for 100 percent of data variance, all points would lie on the axis of the plot and all would be equal (maximum) distance from the origin. In Figure 4 there is a clustering of urban variables (UAG, HDC, MDC, HDR, MDR and TRANSPOR). This is consistent with correlations (Table 18) of urban variables in which those clustered variables are generally highly correlated. Each of the urban variables has a high positive loading on the first (horizontal) principal component and small loading in the second (vertical). The variable AGR has a large negative correlation with the first component. A number of the landform classes cluster together on the second component. Generally, in Figure 4 the factor analysis has produced about what might have been expected on the grounds of *a priori* knowledge or information from variable correlations.

## TABLE 19: VARIABLES[1] AND ABBREVIATIONS USED IN FACTOR ANALYSIS

| | | |
|---|---|---|
| <u>Dependent Variables</u> | | |
| TDS | Total Dissolved Solids | |
| TSOL | Total Solids | |
| SUSPSOL | Suspended Solids | |
| NH3N | Ammonia Nitrogen | |
| TKJELD | Total Kjeldahl Nitrogen | |
| NO2N | Nitrate Nitrogen | |
| NO3N | Nitrate Nitrogen | |
| TP | Total Phosphorus | |
| SRP | Soluble Reactive Phosphorus | |
| HARD | Hardness | |
| ALK | Alkalinity | |
| CHLOR | Chloride | |
| TIRON | Total Iron | |
| | | |
| <u>Independent Variables</u> | | |
| BPOP | Basin Population | |
| BAREA | Basin Area | |
| <u>Landform Classes (% of Basin Area)</u> | | |
| VWBPS | Very Weakly Broken Plains | - Sand |
| VWBPLS | | - Loam-Silt |
| VWBPC | | - Clay |
| WBPBB | Weakly Broken Plains | - Bare Bedrock |
| WBPS | | - Sand |
| WBPLS | | - Loam-Silt |
| WBPC | | - Clay |
| MBUS | Moderately Broken Uplands | - Sand |
| MBULS | | - Loam-Silt |
| MBUC | | - Clay |
| MBUBB | | - Bare Bedrock |
| SBUAM | Strongly Broken Uplands | - All Materials |
| APC | All Peat Categories | |
| AEC | All Escarpment Categories | |
| | | |
| <u>Canada Land Inventory (% of Basin Area)</u> | | |
| UAG | Urban Area >25,000 pop. | |
| UAL | Urban Area <25,000 pop. | |
| EXT | Extractive | |
| AGRI | Agriculture | |
| WOOD | Woodland | |
| MARSHETC | Marsh, Swamp, Barren, Outdoor Recreation | |
| LDC | Low Density Commercial | |
| MDC | Medium Density Commercial | |
| HDC | High Density Commercial | |
| LDR | Low Density Residential | |
| MDR | Medium Density Residential | |
| HDR | High Density Residential | |
| TRANSPOR | Transportation | |
| ORCHARD | Orchards, Horticulture, etc. | |
| CROPLAND | Cropland | |
| IMPAS | Improved Pasture | |
| UNIMPAS | Unimproved Pasture | |

---

[1] Variables are identified in Table 1.

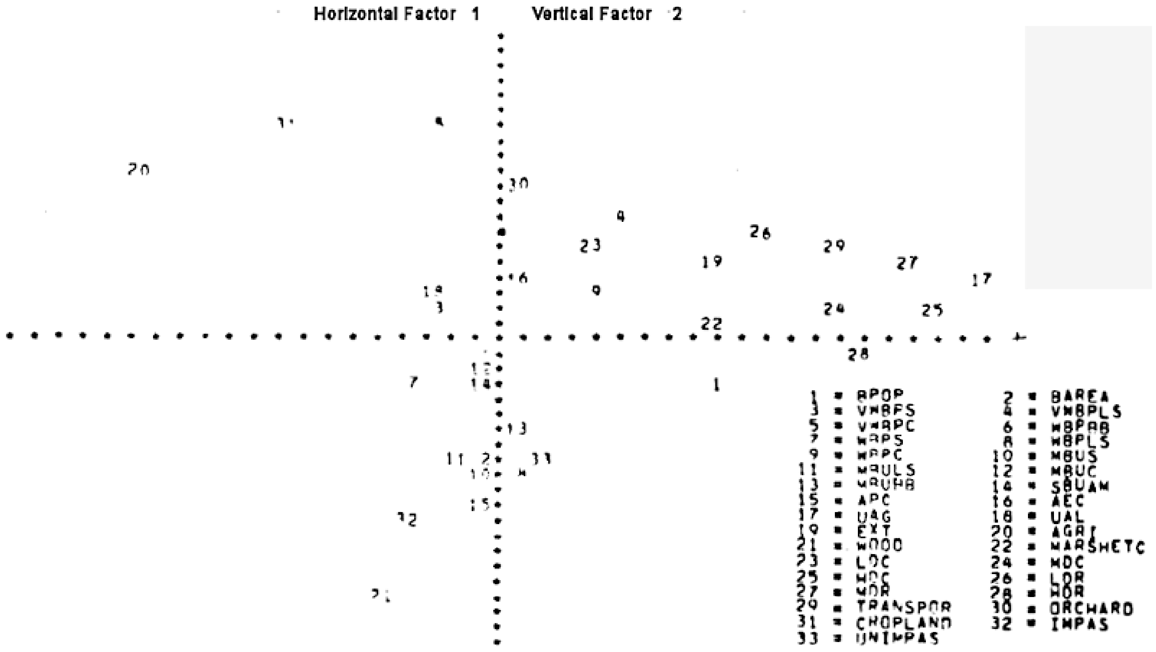

**Figure 4:** Plot of factor loadings on first two principal components (after rotation) of all independent variables. 30.1% of variation is accounted for.

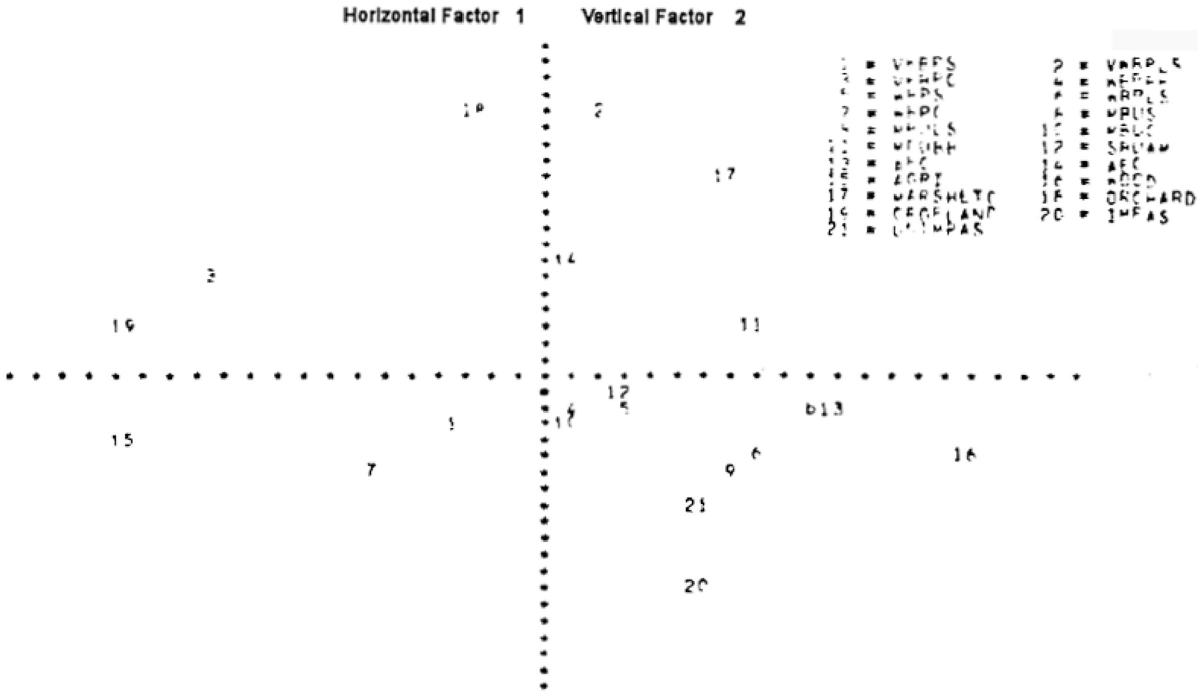

**Figure 5:** Plot of factor loadings on first two principal components (after rotation) on non-urban independent variables. 28.9 % of variation is accounted for.

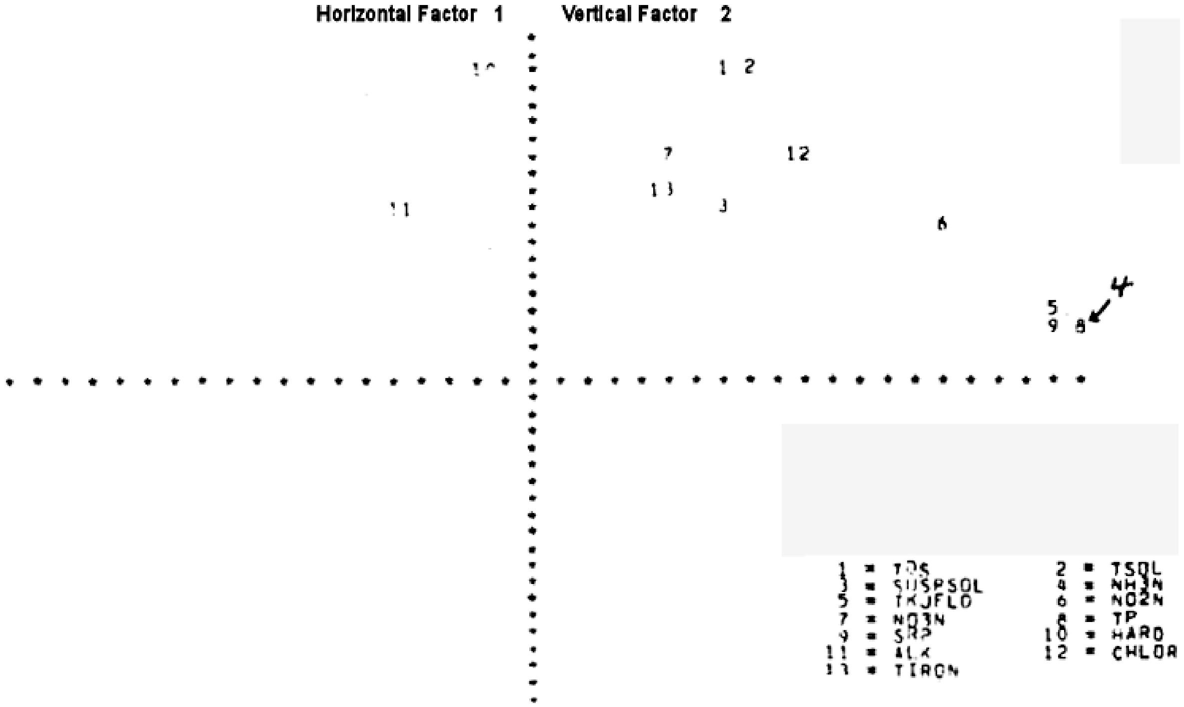

**Figure 6:** Plot of factor loadings on first two principal components (after rotation) of dependent variables. 68.7% of variation is accounted for.

Removal of landform variables produces no groups other than those identified in Figure 4. Using only landform and agricultural variables (Figure 5), the first (horizontal) component suggests degree of agricultural activity. Cropland and Agriculture lie to the far left of the origin, whereas Peat and Woodland lie to the far right. The strong positive relationship between Orchard and the second principal component suggests that Orchard is a separate case of agricultural land use continuum. Significantly, neither pasture category is strongly related to either component. Use of landform categories alone produces no groups.

R mode analysis of dependent variables (Figure 6) produces a strong positive cluster of $NH_3$-N, TKJELD, TP and SRP. Ammonia and Total Kjeldahl nitrogen are not included in parameter correlations (Table 15), however TP and SRP are very highly corrected. It is demonstrated below that TP and nitrogen forms are related both to urban and to rural diffuse drainage. Possibly, separation of the data to reflect point and nonpoint influences may allow interpretation of the first principal component.

Q Mode Analysis

Because of limitations imposed by matrix dimensions acceptable to the SPSS factor analysis program, analysis of basins was conducted on subsets for 'Lake Ontario' (# 1-50) and 'other than Lake Ontario' (# 51-115) basins.[1]

Using dependent variables for Lake Ontario basins a group of basins representing largely urban and intensive agriculture (# 26, 35, 36, 41, 42, 47, 48, 49 of Table 2) from the western end of the Lake cluster on the positive side of first component, whereas largely rural and low intensity agricultural basins (# 5, 6, 7, 22, 27, 40 of Table 2) from eastern Lake Ontario (Highland Creek is an exception) group together on the negative side. Another group of rural basins from eastern Lake Ontario group on the positive side of the second component (# 1, 8, 10, 11, 12, 15, 17, 18) but why these are different from the rural group negatively related to the first component is not clear. The first two principal components account for 64 percent of the variation. The first component appears to be related to intensity of land use whereas the second is not readily explained. None of the groups appear to reflect presence or absence of point sources. As is noted below, the interpretation of Q mode analysis can be made more explicit by judicious selection of variables which define the criteria upon which grouping will be performed.

Using independent variables for Lake Ontario data, an 'urban-intensive agriculture' group of basins is identified together with several loose clusters of other basins. Only 34 percent of the variance is explained by the first two components, suggesting that a better selection of variables might have improved matters. A comparable 'urban-intensive agriculture' group appears when all variables are considered. Summarizing Lake Ontario data, urban and intensive agriculture tend

---

[1] Refer to Table 2 for basin listings.

to group together. This observation is supported by correlation-regression studies which link large concentrations of, for example, total phosphorus and nitrogen forms both with Cropland and urban situations.

Using Lake Erie and Lake Huron data, Q mode analysis groups together agricultural basins within Lake Erie drainage with a second pronounced group comprising Lake Huron and Georgian Bay tributaries. Although the composition of each group varies somewhat, these general groups emerge when using dependent, independent and all variables. The variance explained by the first two principal components is 65 percent for dependent variables and falls to 37 percent both for independent and all variables. Apart from the clusters on a geographical basis which appear to reflect agricultural intensity, there is a discernible tendency for urban basins of Lake St. Clair to group together.

In assessing the results of the factor analysis it is apparent that principal components and their associated basin clusters behave about as one would expect on *a priori* grounds. The data exhibit large variance which is unexplained by the first few principal components. In part, this situation may reflect poor selection of variables to be entered into the analysis. However, it is more likely that the kinds of data employed are very 'noisy'. In comparison with discriminant analysis, the latter is far more efficient in identifying linkages amongst groups of independent variables in terms of their relative effects upon dependent variables. Noisy data reflect the use of data generated only at tributary mouths. These, necessarily, combine basins of all sizes which, obviously, increases variability both in basin characteristics (independent variables) and in their effect upon water quality. One might expect a more efficient factor analytic approach when basin data are selected to reduce variability imparted by excessive size. An entirely similar argument (discussed below) applies to problems of land use-loadings linkages.

## CONCLUSIONS

Although the application of factor analytic methods to these data has not produced much insight into associations of variables or observations (basins), a number of comments concerning the potential for this technique can be drawn from these results.

The dependent variables such as those used here in R mode factor analysis are usually employed in a field research program because they are known *a priori* to have some meaning, either singlely (e.g., chloride) or as groups (e.g., SRP, TP) which can be expected to relate to the subject under investigation. The use of R mode with such variables is likely to group those variables which are already known to have some association. Similar comments may apply to independent variables, except that a much larger number of variables are involved. Nevertheless, in any well-designed field investigation, the independent variables are chosen in order to shed light on known or suspected relationships involving the dependent variables.

In summary, R mode analysis using variables which have been selected according to some *a* *priori* knowledge, is unlikely to produce new insight into variable associations except in the circumstance where a "shotgun" approach has been used in the initial selection of variables , or selection of variables has been made on criteria other than that for which the data are being screened by R mode analysis. An example of the latter is where data from diverse sources are being combined into data sets in order to examine a question where *a priori* expectations are lacking.

Q mode analysis offers a particularly useful means for *a posteriori* inference by its ability to group basins (observations) which display similar/dissimilar characteristics. Although the examples used above illustrate groups of basins, the question remains, "what is there which is common to each group member?" In the example, the first two factors were chosen using all dependent, all independent, and the total of both dependent and independent variables. It is to be expected that, under such circumstances, basin groups are relatively amorphous because groups which might be defined by certain combinations of variables may be differently clustered by other combinations of variables. This suggests that the application of Q mode analysis should be preceded by a rigorous selection of variables according to *a priori* criteria upon which one wishes to cluster basins into groups. This can also be done by using the groups of variables which are identified in R mode.

Q mode appears to be a potentially valuable strategy for identifying groups of basins which have similar responses (dependent variables) or characteristics (independent variables) from which "representative" basins could be chosen. Of course, this application depends on the existence of a spatially diverse data set such as that used here which includes relevant variables. Alternatively, Q mode can be used to test *a priori* assumptions of "representativeness" established independently on other criteria.

In conclusion, the factor model used here does not provide much in the nature of substantive information about variables or basin groups within the Task D data set. It does, however, illustrate a methodology which may be very effective in optimizing future field sampling strategies by its power to categorize basin groups on scientifically meaningful variables. Of course, basins which do not fit into groups are also clearly identified. One drawback of available factor analytic computer programs (such as the SPSS routine used here) is their inefficiency at producing the best possible factor groups. Although the mathematical algorithms exist, routinely available computer programs have sacrificed mathematical rigor for computational efficiency. For most purposes, however, this is probably not a major drawback.

# DISCRIMINANT ANALYSIS

## INTRODUCTION

Discriminant analysis is one of the techniques used to investigate which of the independent variables have an effect on the dependent variables. It is used here as a screening procedure only to identify important variables. The procedure of discriminant analysis was developed for quite different purposes. Its use here is an adaptation of the original method.

The original problem of discriminant analysis is as follows. Two (or more) multivariate populations exist and have known probability density functions. If measurements are made on an unknown individual, then what is the best procedure for classifying this individual as belonging to one of the populations? Typical examples occur in the field of medicine. Suppose it is known that people with a certain type of cancer respond in a known way to various diagnostic tests (temperature, concentrations of various substances in the blood, x-rays, etc.) and that people without the cancer respond differently. A new individual is given the tests and the problem is to classify him as having the cancer or not based on the tests alone.

The problem has been stated in quite general terms. It is usual, to make other assumptions, namely that the two populations have the same multivariate normal distributions except that the mean levels of some or all of the variables might be different. In this case the discriminant function becomes a simple linear function of the observations and the classification becomes that of classifying into one population if the linear function is greater than a certain value and into the other population if it is less.

In practice a number of problems exist. It is not known *a priori* what the distributions are and hence the parameters must be estimated. It is also not known which, if any, of the measured variables are important for discrimination.

The procedure that is followed is, first, to collect individuals that are known to belong to the two different population. A discriminant function is developed for each of the variables separately and the variable which does the best job of classifying the known populations correctly is chosen. The basis of the selection is that the total variance of each variable can be broken down into a variance between populations (explained variation) and a variance within populations (unexplained variations). The statistical significance of the explained variation can be tested by an F-test (ratio of explained variance to unexplained variance). The F statistic is scale free and hence the variable with the largest F value is chosen as being the most important, provided it is significant (exceeds a critical value obtained from tables--in the present case 4.0 was used).

The procedure is then repeated using the variable chosen as most important in combination with all the other variables. In this way the second most important variable is chosen. A third variable is chosen as the best in combination with the previous two and then a fourth and so on until no new variable adds significantly to the explained variation.

As a final step the procedure uses the discriminating function (probability model) and classifies the individuals into the two populations and counts the number of errors. One would expect a 50 percent correspondence between observed and predicted on the basis of chance alone. The percent of basins correctly assigned by the function to each population is an overall measure of the ability of the variables to be used for classification purposes.

The technique described above was used on selected water quality data. In this case there did not exist two *a priori* known populations and, hence, those had to be created artificially. In this study the mean annual concentrations for the period 1968-72 for each of the dependent variables for all 101 basins of Southern Ontario (Table 2) were ranked (Ongley, 1977a) and it was found that a large range exists between minimum and maximum values (Table 20). This suggests that there is reasonable cause to expect that the independent variables are in some way responsible for the range in each of the dependent variables. In order to identify the cause for large concentrations of a particular substance it is useful to attempt to discriminate between basins having high concentrations and those with low concentrations in terms of the controlling (independent) variables. Therefore, the two basin populations per variable were defined as those basins having concentrations falling into the lower third of ranked data and either (1) upper third or (2) highest values were a group of extreme concentration values stand apart from the ranked data. Using these selection criteria the SPSS discriminant analysis routine performs on operator-selected independent variables a step-wise analysis wherein the discriminating variables are identified (using an F test') in the order in which they contribute to the discriminating model. The minimum level of acceptance of F was set at 4.0 (95 percent confidence limits for degrees of freedom appropriate to these data).

In most cases the top and bottom thirds were used to define the two populations. In discriminant analysis there are two conflicting requirements. It is desirable that the two populations be as different as possible and also that there be as many observations as possible in each group. If the observations were divided in half then there would be a maximum number in each group but the least possible separation. If they were divided into fourths or fifths there would be greater separation but fewer observations. It is known that if the data are fairly evenly scattered a division into thirds achieves a desirable balance.

---

[1] Ratio of explained to unexplained variance.

**TABLE 20**: RANGE OF MEAN ANNUAL VALUES FOR WATER QUALITY ATTRIBUTES FOR 101 BASIN MOUTHS, 1968-72

| Parameter | Concentration (mg/L) | | Tonnes (1000) | | Tonnes/Sq. Km. | |
|---|---|---|---|---|---|---|
| | Min. | Max. | Min. | Max. | Min. | Max. |
| Dissolved Solids | 117. | 835. | .268 | 196. | 3.74 | 307. |
| Total Solids | 125. | 895. | .285 | 206. | .49 | 335. |
| Suspended Solids | 4.50 | 148. | .027 | 156. | 1.35 | 70.7 |
| $NH_3$ - N | .037 | 6.88 | .0001 | .681 | .013 | 2.03 |
| Total Kjeldahl - N | .368 | 11.33 | .001 | 2.638 | .124 | 3.57 |
| $NO_2$ - N | .004 | .368 | .00002 | .092 | .001 | .104 |
| $NO_3$ - N | .031 | 4.55 | .0002 | 5.143 | .012 | 1.34 |
| Total P | .027 | 5.8 | .0001 | 1.147 | .009 | 1.84 |
| Soluble Reactive P | .006 | 3.42 | .00002 | .838 | .002 | .998 |
| Hardness - $CaCO_3$ | 91.2 | 458. | .169 | 619. | 28.7 | 162. |
| Alkalinity | 70.0 | 260. | .112 | 396. | 22.6 | 117. |
| Chloride | 4.25 | 219. | .026 | 96. | 2.01 | 93.5 |
| Iron | .202 | 4.81 | .001 | 3.449 | .067 | 1.59 |

Basin Area ($km^2$) (3.1 - 12714.)

Discriminant analysis as used in this study is intended chiefly as a diagnostic technique to identify important variables. The discriminant functions themselves should not be used for predictive purposes because of the invalidity of many of the assumptions. Predictive relations, in so far as it is possible to obtain these from data of this kind, should be developed using regression analysis. The technique, however, is robust enough for the purpose of identifying important variables.

## VARIABLES[1]

The dependent variables used in this analysis are:

Total Dissolved Solids (TDS)
Suspended Solids (SS)
Total Kjeldahl (Tkj)
Nitrate Nitrogen ($NO_3$)
Total Nitrogen (TN)
Total Phosphorus (TP)
Soluble Reactive Phosphorus (SRP)
Chloride (Chl)

Total Nitrogen, Total Phosphorus and a Soluble Reactive Phosphorus each displayed a small group of basins with extremely high concentration values which, in addition to being grouped with the top one-third concentration values (B groups of Tables 21 and 22), were grouped separately (A groups of Tables 21 and 22) on the grounds that causation might be more readily apparent in comparison with basins with low concentrations.

Discriminant analysis was performed on several sets of independent variables beginning with an overview approach employing:

Population
Area
Large Urban[2] (>25,000 pop.)
Agriculture
Woodland
Orchard
Cropland

These variables encompass the major categories of land use plus spatial and demographic information from the 101 drainage basins. Log transformations of population and area[3] did not particularly change any results especially for the first three discriminating variables; therefore, raw means were used throughout.

---

[1] See Table 1.

[2] Combination of Small (<25,000 pop.) and Large Urban categories made no difference to the results; therefore, for ease of computation, Large Urban alone was used. See also Correlation discussion for additional information.

[3] All other variables are expressed as percent of basin area.

Following the overview approach the overburden-rural variables were used (see also Table 19):

Σ Sand
Σ Loam
Σ Clay
Woodland
Marsh
Orchard
Cropland
Improved Pasture
Unimproved Pasture

Finally, the urban subcategories were employed (see also Table 19):

Low Density Commercial
Medium Density Commercial
High Density Commercial
Low Density Residential
Medium Density Residential
High Density Residential
Transportation

## DISCUSSION

The results of the discriminant analysis are tabulated to show sample size in each of the two groups (N), the number of basins correctly discriminated in each group by the model (n)[1] and the overall success of the model (%). The mean and standard deviation of each concentration group is shown together with the F value associated with each discriminating variable. Because the F ratio for second and subsequent discriminating variables is calculated on the remaining unexplained variance, it is possible that such F values may exceed that for the first discriminating variable. In such rare cases, the order of selection of the discriminating variables is denoted by a subscript. Elsewhere, the F values indicate the order of selection, commencing with the highest. Scale of F as a measure of importance of the selected discriminating variables is nonlinear, however, the relative values of F are an indication of the relative importance of each of the discriminating variables. The tables are constructed so that the F value is entered opposite either the lower or upper group of concentration values. This position reflects that group for which the discriminating variable has the largest mean value. For example, for basins associated with the bottom third of the total dissolved solids concentrations, the first discriminating variable (woodland) has a mean value of 31.0%, compared with 9.0% for those basins having the upper third of the concentration values. The F value of 44.7 is entered opposite the lower group indicating that the mean percent of basins occupied by woodland is larger for basins with low concentrations than for those with high concentrations. The mean concentrations associated with each group for each of the dependent variables is provided in order that any inferences drawn from the discriminant analysis concerning controlling variables can be immediately placed in the context of whether the difference between mean concentration of lower and upper groups is sufficiently large to contemplate remedial measures.

---

[1] In some cases the SPSS program failed to compute the number of basins correctly classified by the model. It is thought that this is due to extremely small probabilities generated by the algorithm causing a 'zero divide' situation (a denominator which approaches zero) which terminates the computation.

The results of the discriminant analysis are readily apparent in the tables; therefore, only a brief summary is presented here. The discriminant models are almost always over 80% efficient and usually exceed 90%, indicating that these data provide useful information concerning variables associated with either low or high concentrations of each of the water quality (dependent) variables.

## SUMMARY OF RESULTS

### Overview

Woodland is the first discriminating variable (Table 13A) for TDS, T.Kj, TN (A and B), TP(B), and Chloride. In all cases (and not unexpectedly) the largest amounts of Woodland are associated with the low concentration groups. Because this exercise is concerned with causes for high concentrations (assuming that we do not wish to recommend complete reversion to woodland across the province), the overview was repeated with the Woodland category removed (Table 13B) with the result that the discriminating variables virtually always were associated with the large concentration groups. Urban is the first discriminating variable for T.Kj, TN(A), TP(A), SRP(A) and Chloride. It is significant that the small number of exceptionally high concentrations (A groups of Table 21) for TN, TP and SRP invariably related to the urban categories, an observation which, for phosphorus, is entirely consistent with data tabulated in Uttormark et al. (1974). When examining the top one-third of the concentration data (B groups) for TN, TP and SRP, Cropland becomes the first discriminating variable (also for $NO_3$) followed very closely by the Urban variable. This suggests that remedial strategies aimed at the highest concentrations should be directed toward urban areas within the basins ranked in Appendix 1 (and Ongley, 1977a).[1] Alternatively, remedial programs aimed at a general reduction in the upper third concentration group will have to consider agricultural as well as urban problems.

In summary, for most dependent variables the Cropland and Urban variables are the two most important discriminating variables. Cropland does not enter the discriminating model for TDS or T.Kj, nor does the urban variable enter into Suspended Solids. This latter observation may seem peculiar in view of an abundant literature identifying urban sedimentation as a significant environmental problem. It is known, however, that erosion and sediment transport is primarily due to construction activity in urban areas and that the temporal scale of such activities is rather short. These data, covering the period 1968-72, undoubtedly encompass considerable urban construction; however, the Ontario Ministry of Environment sampling strategies, the long record period and the relatively short-lived effect of construction mitigate against significant inflation of long term

---

[1] A note of caution; these data, ending in 1972, do not reflect subsequent changes in phosphorus levels due to phosphorus control legislation. The interpretation is provided here by way of illustrating the application of the statistical tool to rational formulation of remedial strategies.

**TABLE 21:** DISCRIMINANT ANALYSIS: OVERVIEW

| | | | | | (A) Initial Selection - F Values | | | | | | | | | (B) Second Selection - F Values | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{C}$ | $\delta$ | N | n | % | Pop. | Area | Lg.Urb | Agr. | Wood. | Orch. | Crop. | n | % | Pop. | Area | Lg.Urb | Agr. | Orch. | Crop. |
| TDS | 1 | 250.7 | 45.8 | 33 | 29 | 84.9 | | | | | 44.7 | | | 28 | 81.8 | | | | | | |
| | 2 | 527.9 | 118.2 | 33 | 27 | | | | | 5.4 | | 5.03 | | 26 | | | | 14.11 | 22.72 | 5.9 | |
| SS | 1 | 16.1 | 5.9 | 33 | 31 | 93.9 | | | | | | | | 31 | 93.9 | | | | | | |
| | 2 | 100.9 | 26.8 | 16 | 15 | | 5.4 | | | | | | 71.3 | 15 | | 5.4 | | | | | 71.3 |
| Tkj | 1 | .573 | .080 | 32 | 30 | 93.2 | | | | 9.1 | 32.9 | | | 29 | 88.6 | | | | 9.2 | | |
| | 2 | 6.105 | 3.020 | 12 | 11 | | | | | | | | | 10 | | | | 26.9 | | | |
| $NO_3$ | 1 | .356 | .159 | 33 | 30 | 90.9 | | | | | 15.3 | | | 30 | 87.9 | | | | | | |
| | 2 | 1.961 | .725 | 33 | 30 | | | | | 4.8 | | 4.9 | 47.7 | 28 | | | | 12.2 | | 10.7 | 47.7 |
| TN(A) | 1 | 1.036 | .212 | 33 | 31 | 90.9 | | | | 11.4 | 28.7 | | | 32 | 94.5 | | | | | | |
| | 2 | 8.349 | 2.518 | 11 | 9 | | | | | | | | 7.5 | 10 | | | | 27.3 | | 6.5 | 16.7 |
| TN(B) | 1 | 1.036 | .212 | 33 | 30 | 90.9 | | | | | 57.5 | | | 30 | 89.4 | | | | | | |
| | 2 | 4.989 | 2.820 | 33 | 30 | | | | | 11.43 | | 7.94 | 8.02 | 29 | | | | 35.02 | | 12.4 | 30.11 |
| TP(A) | 1 | .062 | .019 | 38 | 37 | 97.9 | | | | | | | | 37 | 97.9 | | | | | | |
| | 2 | 3.378 | 1.614 | 9 | 9 | | | | 50.4 | | | 5.0 | 12.9 | 9 | | | | 50.4 | | 5.0 | 12.9 |
| TP(B) | 1 | .058 | .017 | 33 | 27 | 83.3 | | | | | 41.5 | | | 29 | 84.9 | | | | | | |
| | 2 | 1.243 | l.563 | 33 | 28 | | | | | 6.2 | | | 5.4 | 27 | | | | 27.32 | | 5.0 | 14.41 |
| SRP(A) | 1 | .013 | .003 | 25 | 24 | 96.9 | | | | | | | | 24 | 96.9 | | | | | | |
| | 2 | 2.091 | .874 | 7 | 7 | | | | 51.0 | | | 5.8 | 4.1 | 7 | | | | 51.0 | | 5.8 | 4.1 |
| SRP(B) | 1 | .015 | .005 | 33 | 27 | 86.4 | | | | | 58.0 | | | 27 | 86.4 | | | | | | |
| | 2 | .634 | .866 | 33 | 30 | | 4.03 | | | | | | 4.57 | 30 | | | | 25.62 | 7.7 | | 14.71 |
| Chl | 1 | 11.0 | 4.1 | 33 | - | | | | | | 61.7 | | | 29 | 81.8 | | | | | | |
| | 2 | 90.0 | 48.3 | 33 | - | | | | | | | | | 25 | | | | 16.81 | 32.82 | 7.2 | |

F value entered in group position associated with larger mean discriminating variable

F accept = 4.0 F reject = 3.9

$\overline{C}$ = Mean Annual Concentration (mg/L)

$\delta$ Sample Standard Deviation

N = Number in group n = Number correctly classified by discriminating function

% Percent of group predicted by discriminant function

1 Lower Concentration Group 2 Upper Concentration Group

Data are means for 1968-72

**TABLE 22:** DISCRIMINANT ANALYSIS

| | | | | | (A) Overburden-Rural Selection - F Values | | | | | | | | | | | (B) Urban. Subcategories -F Values | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\overline{C}$ | $\delta$ | N | n | % | Sand | Loam | Clay | Wood. | Marsh. | Orch. | Crop. | I.Pas. | Un.Pas. | n | % | L.Com. | M.Com. | H.Com. | L.Res. | M.Res. | H.Res. | Tran. |
| TDS | 1 | 250.7 | 45.8 | 33 | 23 | 78.8 | | | | | | | 44.7 | | | | | | | | | | | |
| | 2 | 527.9 | 118.2 | 33 | 29 | | | | | | | 5.0 | | | | - | | | | | | 12.7 | | |
| SS | 1 | 16.1 | 5.9 | 33 | - | | | | | | | | | | | - | | | | | | | | |
| | 2 | 100.9 | 26.8 | 16 | - | | | | | | | | 71.3 | | | - | | | | | None Significant | | | |
| TKj | 1 | .573 | .080 | 32 | 30 | 93.2 | | | | 32.5 | | | | 9.7 | | - | | | | | | | | |
| | 2 | 6.105 | 3.020 | 12 | 11 | | | | | | | | 4.9 | | | - | | | | | | 28.8 | | |
| $NO_3$ | 1 | .356 | .159 | 33 | 28 | 89.4 | | | | 15.3 | | | | | 8.0 | - | | | | | | | | |
| | 2 | 1.961 | .725 | 33 | 31 | | | | | | | | 47.7 | | | - | | | | | None Significant | | | |
| TN(A) | 1 | 1.036 | .212 | 33 | 30 | 90.9 | | | | 28.7 | | | | 18.2 | | - | | | | | | | | |
| | 2 | 8.349 | 2.518 | 11 | 10 | | | | | | | | | | | - | | | | | | 32.3 | | |
| TN(B) | 1 | 1.036 | .212 | 33 | 30 | 93.9 | | | | 57.5 | | | | 31.6 | | - | | | | | | | | |
| | 2 | 4.989 | 2.820 | 33 | 32 | | | | | | | | | | | - | | | | | | 8.2 | | |
| TP(A) | 1 | .062 | .019 | 38 | 37 | 97.9 | | | | 21.9 | | | | 18.3 | | - | | | | | | | | |
| | 2 | 3.378 | 1.614 | 9 | 9 | | | | | | | | 4.6 | | | - | | | | | | 62.7 | | |
| TP(B) | 1 | .058 | .017 | 33 | 27 | 84.9 | | | | 41.5 | | | | 21.3 | | - | | | | | | | | |
| | 2 | .243 | 1.563 | 33 | 29 | | | | | | | | | | | _ | | | | | | 12.0 | | |
| SRP(A) | 1 | .013 | .003 | 25 | - | | | | | | | | | | | 24 | 93.8 | | | | | | | |
| | 2 | 2.091 | .874 | 7 | - | | | | 28.8 | | | | | | | 6 | | | | | | 51.8 | | 7.5 |
| SRP(B) | 1 | .015 | .005 | 33 | 28 | 88.4 | | | | 58.0 | | | | 9.6 | | - | | | | | | | | |
| | 2 | .634 | .866 | 33 | 29 | | | | 5.1 | | | | | | | - | | | | | | 11.7 | | |
| Chl | 1 | 11.0 | 4.1 | 33 | 29 | 87.9 | 7.0 | | | 61.8 | | | | 7.2 | | - | | | | | | | | |
| | 2 | 90.0 | 48.3 | 33 | 29 | | | | | | | | | | | | | | | | | 12.9 | | |

F value entered in group position associated with larger mean discriminating variable

F accept = 4.0 F reject = 3.9

$\overline{C}$ = Mean Annual Concentration (mg/L)

$\delta$ Sample Standard Deviation

N = Number in group n = Number correctly classified by discriminating function

% Percent of group predicted by discriminant function

1 Lower Concentration Group 2 Upper Concentration Group

Data are means for 1968-72

urban-related sediment concentrations. On the other hand, cropland is an annual producer of sediment--a fact readily discernible in these data.

### Overburden-Rural

As a second stage in examining the role of cropland, only the landform overburden type and rural categories were used (Table 22). Because of the limited number of basins (observations), each of sand, loam and clay classes were grouped, thereby eliminating any differentiation on the basis of relief. Woodland, Marsh, Orchard, Cropland, Improved Pasture and Unimproved Pasture were the remaining independent variables. Although Woodland is the first discriminating variable for TDS, T.Kj, TN(A and B), TP(A and B), SRP(B) and Chloride, it is usually combined with Improved Pasture to form the discriminating function. The largest values both of Woodland and Improved Pasture are always associated with the low concentration groups. As noted below in the correlation procedures, Cropland and Improved Pasture generally have opposite relationships to concentrations and yields of the dependent variables. That Improved Pasture should also be associated here with low concentration values is peculiar in view of the fact that the definition of 'Improved Pasture' is that it must exhibit evidence of ploughing, seeding or other improvement. Nevertheless, these results are substantiated in correlation procedures. This suggests that the economies of minimum tillage practices should be evaluated within the context of potential reduction of large (upper third) concentrations of variables such as TN(B), TP(B) and SRP(B). Cropland remains uniquely associated with large concentrations of suspended solids. Of the overburden types, clay is the first discriminating variable for SRP(A). Whereas high concentrations of SRP(A) are also discriminated on the basis of 'Urban' (Table 21B), it can be demonstrated that clay is generally common to basins which contain large urban conditions--an association which can be explained by the occurrence of historical settlement patterns on clay plains bordering the Lower Lakes. The addition of further rural variables in Table 22A does not change the association of large concentrations of $NO_3$ with croplands observed in Table 21B.

### Urban Subcategories

The above discussion has revealed that the Urban category is associated with high concentrations of many of the quality variables. A pass at the data (Table 22B) using the urban categories of Low, Medium and High Density Commercial; Low, Medium and High Density Residential, and Transportation identifies only Medium Density Residential as the discriminating function for most of the upper concentration groups with the exception of 'Transportation' which is a distant second discriminating function for SRP(A). No discriminating function can be identified for Suspended Solids or NO3 using urban variables. Significantly, these are the two variables for which Cropland is the principal discriminating function. These results illustrate the importance of correct scientific inference as it would seem unlikely that Medium Density Residential *per se* is a prime phosphorus producer. Although, 'Urban' is a loose yet comprehensive category, it is very likely that, in comparison, Medium Density Residential is a surrogate for phosphorus producers which are either located within the Medium Density Residential class (e.g., small plazas, car washes, etc.) or statistically (i.e., functionally) associated (though not linked in a locational sense) with the Medium

Density Residential category. It is likely that present definitions of the urban categories, while sensible within economic and/or demographic classifications, do not produce the best possible classification system for pollutant generation.

In these data there is also the problem of point and nonpoint sources within the data set. It is likely, therefore, that further filtering with discriminant analysis using a larger data set composed of smaller basins may not only offer useful insight into causal relationships between pollutant generation and independent variables but will also lead to better definitions of (independent) variables.

In conclusion, discriminant analysis appears to be most useful for rapid association of independent variables with groups of basins having high or low concentrations of water quality (dependent) variables. The results are entirely consistent with those obtained from the more rigorous but much more time-consuming correlation-regression analysis discussed below. The F values for the discriminating variables allows an assessment of the relative importance of the variables forming the discriminant function. Used in conjunction with tables of ranked concentration data (Appendix 1 and Ongley, 1977a) the specific basins involved may be identified, which in a sense, offers an alternative (though not necessarily a replacement) to the Q mode factor analytic approach. Causal inferences must be carried out cautiously for the variables may well be spatially linked to (in a locational sense), or statistical surrogates for pollutant generators.

## BASIN COMMUNALITY

A significant question is the extent to which basins located in the upper or lower group of concentration values for a particular water quality variable are common to upper or lower groups for other variables. The matrix of Table 23 identifies the percent of basins common to each of the Lower and Upper concentration group for all combinations of variables. Generally, there is moderate to high communality within Lower concentration groups. In comparison, however, the Upper concentration groups have considerably fewer basins in common. For example, 64.8% of the basins included in the Lower concentration groups for each of TDS and TP(A) are common to each group, whereas only 16.7% of the basins are common to the two Upper concentrations groups. As a rule, basins with low concentrations of one substance also have low concentrations of other substances. In contrast, high concentrations of one substance do not necessarily reflect high concentrations of other substances. As seen in discriminant, factor and correlation analyses, high concentrations of the dependent variables are related to different sets of independent variables. In terms of causality, it is not surprising that basin communalties are high for low concentrations but very variable for high concentrations.

Obviously, where two variables are highly correlated, the degree of communality will be large. For example, TP and SRP which are highly correlated, have a large proportion of basins in common. Other variables which are poorly correlated, particularly when a small number of the highest concentrations are considered (such as TP(A) and $NO_3$) have few (23.8%) basins in common. In

**TABLE 23:** PERCENT OF BASINS IN COMMON IN DISCRIMINANT ANALYSIS

| | TDS | SS | TKj | NO3 | TP(A) | TP(B) | SRP(A) | SRP(B) | Chl | TN(A) | TN(B) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TDS | | 70.0 | 55.4 | 72.7 | 64.8 | 63.6 | 62.1 | 60.6 | 72.7 | 69.7 | 69.7 |
| SS | 20.4 | | 61.5 | 69.7 | 88.5 | 72.7 | 65.5 | 66.7 | 60.6 | 75.8 | 78.8 |
| TKj | 40.0 | 21.4 | | 57.6 | 71.4 | 73.8 | 56.1 | 64.6 | 73.8 | 70.8 | 70.8 |
| NO$_3$ | 48.5 | 49.0 | 22.2 | | 62.0 | 54.5 | 62.1 | 57.6 | 60.6 | 84.8 | 84.8 |
| TP(A) | 16.7 | 24.0 | 85.7 | 23.8 | | - | 73.0 | - | 72.3 | 70.4 | - |
| TP(B) | 66.7 | 36.7 | 53.3 | 48.5 | - | | - | 81.8 | 75.8 | - | 69.7 |
| SRP(A) | 30.0 | 17.4 | 73.7 | 25.0 | 87.5 | - | | - | 65.5 | 65.5 | - |
| SRP(B) | 63.6 | 32.6 | 44.4 | 45.5 | - | 87.9 | - | | 72.7 | - | 69.7 |
| Chl | 84.8 | 24.5 | 48.8 | 42.4 | 38.1 | 60.6 | 35.0 | 63.6 | | 69.7 | 69.7 |
| TN(A) | 36.4 | 29.6 | 87.0 | 27.3 | 90.0 | - | 66.7 | - | 45.5 | | - |
| TN(B) | 60.6 | 44.9 | 53.3 | 72.7 | - | 75.8 | - | 66.7 | 57.6 | - | |

Lower Concentration Group

Upper Concentration Group

a sense, Table 23 allows one to express the variance inherent in the Lower (smaller values) and Upper (larger values) portions of a regression of two variables in terms of a spatial attribute (here, basins in common).

## RECOMMENDATIONS

1. The data set should be segregated to reflect presence and absence of point sources in order to more clearly assess the role of non- point sources.

2. The phosphorus analysis reflects the period prior to phosphorus controls. The results are meaningful but only in a general sense. They should be evaluated in the context of the discussion above of phosphorus control using 1974 data which reflects implementation of phosphorus regulations for detergent and municipal wastewater effluents.

3. A similar analysis should be carried out for post-1974 data, particularly for phosphorus.

4. An expanded data set containing smaller basins (i.e., reduced variability) should further clarify significant relationships amongst variables.

5. The urban categories used are probably not appropriate for identification of phosphorus producers. For example, Medium Density Residential may be a surrogate for parking lot, local plaza and car wash (fictitious example).

6. With a rational selection of independent variables, discriminant analysis offers the most rapid overview of relationships amongst dependent and independent variables with respect to groups of basins having high and low concentrations of water quality (dependent) attributes.

# CORRELATION-REGRESSION ANALYSIS

## INTRODUCTION

Of the four statistical techniques employed in this report, regression is the most explicit in the sense of hypothesis testing. Regression approaches to land use-water quality relationships are numerous. In a recent comprehensive survey of such literature Omernik (1976) reports that:

> "In attempting to develop systems for estimating nutrient runoff from land use based on coefficients developed entirely, or in part, from the literature, most reviewers have summarized their findings by presenting a range of values and, in some cases, midpoints or averages. Generally, these ranges are quite wide and the midpoints, or other indicators of central tendency, do not vary from one land use type to another as appreciably as one might expect."

Omernik's views are probably not entirely shared by Dillon and Kirchner (1975) who, for total phosphorus, found in their own studies significant differences in mean annual yields, for geologic, forest and pasture categories, and substantial differences amongst, variously, their categories and published yields from agricultural and urban systems.

Omernik's (1976) study is perhaps the most comprehensive of its type, using field data drawn from 473 basins in eastern United States (east of the Mississippi River) representing nonpoint source watersheds. His data include water quality (collected monthly in general), geology, land use, slope, and domestic animal and fowl density (various species). Yields and concentrations of nitrogen and phosphorus (total and inorganic) were regressed against a variety of land use types and, although certain trends emerged, the correlation coefficients were disappointingly low (although statistically significant due to the large number of data points). Although frequency polygons indicate that mean nitrogen and phosphorus concentrations can be attributed to different land use types, regression of mean concentration was performed with '% agriculture plus % urban' land use on the grounds that, 'Generally, relationships between these ratios and nutrient levels in streams were found to be more significant than those considering only one land use'. Although correlation coefficients (r) of +0.73 and +0.83 were obtained for phosphorus and nitrogen concentrations respectively, r values were generally much lower for yields (dependent variable) and for other combinations of land use variables.

In view of the amount of data involved, Omernik's study is disappointing in so far as the regression approach is intuitively an attractive way of synthesizing data for the purposes of predicting concentration or yields from land use data. The coefficients of determination are always less (usually much less) than 65% indicating considerable unexplained variance. Moreover, the use of '% agriculture plus % urban' is not useful for estimating the influence of different categories of land use. It follows that his regression models are unsuitable either for gaming with alternative remedial strategies or for generating loading functions.

Although this study also uses the linear regression model, the variables herein are more numerous and more specific (e.g., agricultural and urban subtypes). Moreover, these data are not as severely limited as is the case for Omernik's study, by paucity of urban information. The variables used here reflect the outcome of initial filtering by correlation (Section 8). Only one of two or more variables which are closely correlated (such as total phosphorus and soluble reactive phosphorus) is reported below. The dependent variables are concentrations and yields of:

- Total Dissolved Solids
- Suspended Solids
- Total Nitrogen
- Nitrate Nitrogen
- Total Phosphorus
- Chloride

Although yield (unit load) and concentration are very closely correlated in this study, the discussion of each analysis is generally limited to concentration. In part, this decision reflects the fact that some 50% of the basins used in this study lack discharge gauges. Flow information has, therefore, been estimated on the basis of an estimating equation reflecting basin area (Ongley, 1974).[1] It follows, therefore, that some linearity is to be expected between concentration and yield for those basins for which discharge is estimated from a linear model. Also, as illustrated in Section 6, concentration values reflect dominant land uses for annual (or mean annual) aggregation of data.

A more pressing reason for avoiding the presentation of explicit regression models linking yield with independent variables is to preclude the indiscriminate use of the model without a full understanding of the circumstances under which it may or may not apply. Because of the potential for such models and the limitations imposed by the data employed herein, they ought to be developed on a 'demand' basis and used with full cognizance of the circumstances surrounding the kinds of data employed in the models. Because concentration and yield are highly correlated, the

---

[1] The estimating function for discharge in ungauged basins employs mean annual data and has extremely small variance. The significance of the correlation between concentration and yield herein is probably not particularly affected by the use of a linear estimating model for discharge. Further discussion is found in Ongley (1974).

discussion of concentration relationships applies also to yields. The mathematical constants are not, however, directly transferrable.

The nature of the study is explorative and, although the majority of the relationships noted in the main body of the report are statistically significant, the changing relationships between water quality and land use can be more usefully explored by examining trends in correlation coefficients rather than a rigorous application of hypothesis testing and statements of standard errors. For example, the possibility that certain trends may prove to be important is far more important in this study than avoidance of Type II errors, particularly given the restrictions inherent in these data (as noted above) and the inevitable transgression of certain statistical requirements.

The data set was divided into the following subsets for correlation- regression analysis.

1. All 101 basins regardless of presence or absence of point sources.

2. 49 nonpoint source basins (as defined in Section 4).

The point source basins have not been considered separately because of the many differences in types of point sources. The nonpoint data set allows explicit examination of diffuse source contributions. For each data subset, each dependent variable (concentration and yield) was correlated with respect to:

(1) Overview: 7 independent variables representative of major land use and demographic categories:

Population
Area
Large Urban
Agriculture
Woodland
Cropland
Improved Pasture

It should be recalled that 'Large Urban' is a useful surrogate for a variety of urban subcategories as indicated in the discriminant analysis. 'Cropland' and 'Improved Pasture', although subject to tillage, behave statistically very differently with respect to generation of certain pollutants. This fact will be commented upon below.

(2) Overburden-Rural (see Table 14): 7 independent variables reflecting:

Area
$\Sigma$ Sand
$\Sigma$ Loam
$\Sigma$ Clay
Woodland
Cropland
Improved Pasture

The correlation matrices permit an examination of overburden relationships with respect both to concentration and yields, and to agricultural land uses.

Using the interactive regression program described in (Ongley, 1977a) relationships which initially appeared of interest in the correlation matrix were examined by selecting further basin subsets on the basis of other criteria and examining regressions, histograms and plots of relevant data associations. Two examples of criterion frequently used are: (a) basins less than a stated area (in order to reduce spatial variance in land use) and (b) basins for which 'Large Urban' >0.0% or, alternatively = 0.0% (the presence or absence of 'Large Urban', particularly in nonpoint source basins, has a profound influence on phosphorus generation). Meaningful[1] regressions were examined on a two-dimensional plot to establish whether or not the trends were the result of a single or a few data points distant from a trendless cluster of points. It was found that where strong trends exist between two variables, the addition of subsequent independent variables using multiple regression seldom produced much improvement in the explained variance.

### TOTAL DISSOLVED SOLIDS (TDS) All Basins (Table 24)

Total dissolved solids concentrations are moderately correlated with Large Urban (+0.55) and negatively with Woodland (-0.60). The correlation coefficient does not improve when only basins with non-zero values for Large Urban are used. This implies that variance is not restricted to other variables when Large Urban 0 0.0% and that Large Urban does not have a unique relationship to TDS. None of the overburden variables relate to concentration or yield although sand is positively related to Woodland, whereas clay is negatively related to Woodland and positively to Cropland. These latter two relationships are, of course, common to all overburden-rural correlation matrices and will not be further discussed.

### NPS Basins (Table 25)

The reduction from +0.55 to +0.32 of the correlation coefficient of TDS concentration versus Large Urban implies that TDS, while attributable in part to point sources, is weakly related to urban diffuse runoff. Because of the control of natural as well as anthropogenic factors upon TDS, it is not surprising that none of the rural, urban or overburden characteristics are uniquely related to TDS concentrations. As noted above, Improved Pasture and Cropland display opposite signs in their respective relationships to TDS.

---

[1] Generally statistically significant but, where data are very limited, trends which are scientifically reasonable are examined.

## SUSPENDED SOLIDS (SS)

### All Basins (Table 26)

Discriminant analysis indicates the Cropland is the sole discriminating variable. Correlation results confirm this outcome with r = +0.66 for SS concentration and Cropland. No other variable except for the more general variable Agriculture is significantly (positively) related to SS, either for overview or for overburden variables.

### NPS Basins (Table 27)

Elimination of point source basins raises the correlation coefficient for SS concentration with Cropland from +0.66 to +0.74, suggesting that point sources and/or urban diffuse drainage associated with point source basins is responsible for limited SS which obscures the relationship with Cropland in those basins having point sources. None of the overburden variables are related to SS concentrations when all NPS basins are considered. Figure 7 illustrates the trend between SS concentration and Cropland.

### Discussion

Because Population and urban variables are generally unrelated to Suspended Solids concentrations, the effects of overburden type and basins size on variance associated with concentration-Cropland trends are conveniently illustrated using all (101) basins. Because Cropland incorporates a host of management practices and crop types which are frequently related to soil and slope characteristics, one would expect considerable unexplained variance in the concentration-Cropland trend. By restricting the overburden (soil) variable to a large proportion of sand or clay, variability introduced by mixed soil types is reduced. In Table 28 only those basins with in excess of 50% of their area in sand have been used. The correlation coefficient between SS concentration and Cropland remains unchanged at +0.66. However, when only basins under 50 sq. miles are used, the correlation coefficient improves to +0.77 (Figure 8). There are too few basins to test for the effect of increasing the % sand category to an amount greater than 50%.

In comparison with sand, basins with greater than 50% of their area in clay have a lower correlation coefficient between SS concentration and Cropland (r = +0.59: Table 28). Increasing the % clay to 60%, 70%, 80% and 90% respectively produces correlation coefficients of +0.58, 0.60, 0.64 and 0.68 (n = 19). When basins having 90% of their area in clay are restricted to those of less than 50 sq. miles, r = +0.83 (Figure 9).

These results illustrate the dilemma facing those who seek unit load (yield) values for the purpose of assessing the role of agriculture on water quality at some downstream location. Whereas small areas with relatively similar and homogeneous conditions may be expected to produce yield values with small variance, such values change to an unknown but probably a considerable degree in a downstream direction due to transmission effects (enrichment, sedimentation, bio-uptake, phase and chemical changes, etc.). An alternative (such as that use by EPA, 1976), to small area

**TABLE 24:** CORRELATION COEFFICIENT MATRIX - TOTAL DISSOLVED SOLIDS

ALL BASINS

(A) OVERVIEW

| | C | Y | Pop. | Area | Lg. Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.8423 | 0.2420 | -0.2016 | 0.5520 | -0.0160 | -0.5995 | 0.0691 | -0.2731 |
| Y | 0.8423 | 1.0000 | 0.1688 | -0.2499 | 0.4391 | -0.0350 | -0.4490 | 0.0068 | -0.1722 |
| Pop. | 0.2420 | 0.1688 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.2016 | -0.2499 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.5520 | 0.4391 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | -0.0160 | -0.0350 | -0.2742 | -0.1939 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.5995 | -0.4490 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.0691 | 0.0068 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I. Pas. | -0.2731 | -0.1722 | -0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

(B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I. Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.8423 | -0.2016 | -0.2332 | -0.0238 | 0.2266 | -0.5995 | 0.0691 | -0.2734 |
| Y | 0.8423 | 1.0000 | -0.2499 | -0.2792 | 0.0284 | 0.1968 | -0.4490 | 0.0068 | -0.1722 |
| Area | -0.2016 | -0.2499 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.2332 | -0.2792 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | -0.0238 | 0.0284 | 0.0390 | -0.1512 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.2266 | 0.1968 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.5995 | -0.4490 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.0691 | 0.0068 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I. Pas. | -0.2731 | -0.1722 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

## **TABLE 25:** CORRELATION COEFFICIENT MATRIX - TOTAL DISSOLVED SOLIDS

### NONPOINT SOURCE BASINS

### (A) OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.8736 | 0.2945 | -0.3289 | 0.3182 | 0.2248 | -0.5603 | 0.2151 | -0.3934 |
| Y | 0.8736 | 1.0000 | 0.1367 | -0.2666 | 0.1924 | 0.1629 | -0.3918 | 0.1034 | -0.2073 |
| Pop. | 0.2945 | 0.1367 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | 0.3289 | -0.2666 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.3182 | 0.1924 | 0.6400 | -0.1120 | 1.0000 | -0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | 0.2248 | 0.1629 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.5603 | -0.3918 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.2151 | 0.1034 | -0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.3934 | -0.2073 | -0.0315 | 0.2872 | -0.2767 | -0.1390 | 0.4455 | -0.6921 | 1.0000 |

### (B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.8736 | -0.3289 | -0.2244 | -0.0819 | 0.2436 | -0.5603 | 0.2151 | -0.3934 |
| Y | 0.8736 | 1.0000 | -0.2666 | -0.3043 | -0.0021 | 0.2002 | -0.3918 | 0.1034 | -0.2073 |
| Area | -0.3289 | -0.2666 | 1.0000 | 0.0085 | 0.1136 | -0.1960 | 0.5283 | -0.3259 | 0.2872 |
| Σ Sand | -0.2244 | -0.3043 | 0.0085 | 1.0000 | -0.1363 | -0.5209 | 0.4471 | -0.0226 | -0.0860 |
| Σ Loam | -0.0819 | -0.0021 | 0.1136 | -0.1363 | 1.0000 | -0.7054 | 0.1534 | -0.3832 | 0.1721 |
| Σ Clay | 0.2436 | 0.2002 | -0.1960 | -0.5209 | -0.7054 | 1.0000 | -0.5666 | 0.4223 | -0.1549 |
| Wood. | -0.5603 | -0.3918 | 0.5283 | 0.4471 | 0.1534 | -0.5666 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.2151 | 0.1034 | -0.3259 | -0.0226 | -0.3832 | 0.4223 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.3934 | -0.2073 | 0.2872 | -0.0860 | 0.1721 | -0.1549 | 0.4455 | -0.6921 | 1.0000 |

## **TABLE 26:** CORRELATION COEFFICIENT MATRIX - SUSPENDED SOLIDS

ALL BASINS

### (A) OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9028 | 0.0574 | -0.1649 | 0.1242 | 0.3241 | -0.5313 | 0.6555 | -0.4987 |
| Y | 0.9028 | 1.0000 | 0.0935 | -0.1449 | 0.0948 | 0.2786 | -0.4479 | 0.5218 | -0.3607 |
| Pop. | 0.0574 | 0.0935 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.1649 | -0.1449 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.1242 | 0.0948 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | 0.3241 | 0.2786 | -0.2742 | -0.1989 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.5313 | -0.4479 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.6555 | 0.5218 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.4987 | -0.3607 | -0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

### (B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9028 | -0.1649 | -0.2203 | -0.2188 | 0.4121 | -0.5313 | 0.6555 | -0.4987 |
| Y | 0.9028 | 1.0000 | -0.1449 | -0.2380 | -0.1135 | 0.3382 | -0.4479 | 0.5218 | -0.3607 |
| Area | -0.1649 | -0.1449 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.2203 | -0.2380 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | -0.2188 | -0.1135 | 0.0390 | -0.1512 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.4121 | 0.3382 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.5313 | -0.4479 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.6555 | 0.5218 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.4987 | -0.3607 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

## **TABLE 27:** CORRELATION COEFFICIENT MATRIX - SUSPENDED SOLIDS

### NONPOINT SOURCE BASINS

### (A) OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9849 | 0.1143 | -0.1949 | 0.1266 | 0.3940 | -0.5554 | 0.7433 | -0.5701 |
| Y | 0.9849 | 1.0000 | 0.1372 | -0.1684 | 0.1104 | 0.3756 | -0.5225 | 0.7112 | -0.5082 |
| Pop. | 0.1143 | 0.1372 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | -0.1949 | -0.1684 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.1266 | 0.1104 | 0.6400 | -0.1120 | 1.0000 | -0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | 0.3940 | 0.3756 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.5554 | -0.5225 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.7433 | 0.7112 | -0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.5701 | -0.5082 | -0.0315 | 0.2872 | -0.2767 | -0.1390 | 0.4455 | -0.6921 | 1.0000 |

### (B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9849 | -0.1949 | -0.2135 | -0.2042 | 0.3893 | -0.5554 | 0.7433 | -0.5701 |
| Y | 0.9849 | 1.0000 | -0.1684 | -0.2225 | -0.1838 | 0.3702 | -0.5225 | 0.7112 | -0.5082 |
| Area | -0.1949 | -0.1684 | 1.0000 | 0.0085 | 0.1136 | -0.1960 | 0.5283 | -0.3259 | 0.2872 |
| Σ Sand | -0.2135 | -0.2225 | 0.0085 | 1.0000 | -0.1363 | -0.5209 | 0.4471 | -0.0226 | -0.0860 |
| Σ Loam | -0.2042 | -0.1838 | 0.1136 | -0.1363 | 1.0000 | -0.7054 | 0.1534 | -0.3832 | 0.1721 |
| Σ Clay | 0.3893 | 0.3702 | -0.1960 | -0.5209 | -0.7054 | 1.0000 | -0.5666 | 0.4223 | -0.1549 |
| Wood. | -0.5554 | -0.5225 | 0.5283 | 0.4471 | 0.1534 | -0.5666 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.7433 | 0.7112 | -0.3259 | -0.0226 | -0.3832 | 0.4223 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.5701 | -0.5082 | 0.2872 | -0.0860 | 0.1721 | -0.1549 | 0.4455 | -0.6921 | 1.0000 |

## **TABLE 28:** CORRELATION COEFFICIENT MATRIX - SUSPENDED SOLIDS

### ALL BASINS

### (A) SAND >50%

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.8821 | 0.3392 | -0.0026 | -0.5313 | 0.5345 | -0.4491 | 0.6673 | -0.5601 |
| Y | 0.8821 | 1.0000 | 0.2352 | 0.2432 | -0.3999 | 0.1560 | -0.2033 | 0.5824 | -0.6179 |
| Area | 0.3392 | 0.2352 | 1.0000 | -0.3115 | -0.2446 | 0.5222 | -0.2839 | 0.4331 | -0.2490 |
| Σ Sand | -0.0026 | 0.2432 | -0.3115 | 1.0000 | -0.4964 | -0.5217 | -0.0081 | 0.4318 | -0.4971 |
| Σ Loam | -0.5313 | -0.3999 | -0.2446 | -0.4964 | 1.0000 | -0.4803 | 0.3460 | -0.8898 | 0.7784 |
| Σ Clay | 0.5345 | 0.1560 | 0.5222 | -0.5217 | -0.4803 | 1.0000 | -0.3341 | 0.4289 | -0.2534 |
| Wood. | -0.4491 | -0.2033 | -0.2839 | -0.0081 | 0.3460 | -0.3341 | 1.0000 | -0.3904 | 0.0241 |
| Crop. | 0.6673 | 0.5824 | 0.4331 | 0.4318 | -0.8898 | 0.4289 | -0.3904 | 1.0000 | -0.8638 |
| I.Pas. | -0.5601 | -0.6179 | -0.2490 | -0.4971 | 0.7784 | -0.2534 | 0.0241 | -0.8638 | 1.0000 |

### (B) SAND >50% and AREA <50 Mi.$^2$

| | C | Y | Area | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9815 | 0.1607 | -0.3441 | 0.7720 | -0.8011 |
| Y | 0.9845 | 1.0000 | 0.1019 | -0.4479 | 0.7506 | -0.7604 |
| Area | 0.1607 | 0.1018 | 1.0000 | 0.2482 | 0.4994 | -0.5181 |
| Wood. | -0.3441 | -0.4418 | 0.2482 | 1.0000 | -0.1852 | -0.0113 |
| Crop. | 0.7720 | 0.7500 | 0.4994 | -0.1852 | 1.0000 | -0.9116 |
| I.Pas. | -0.8011 | -0.7604 | -0.5181 | -0.0113 | 0.9116 | 1.0000 |

### (C) CLAY >50%

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9542 | 0.0434 | 0.1724 | -0.1996 | 0.1564 | -0.3109 | 0.5939 | -0.4439 |
| Y | 0.9542 | 1.0000 | 0.1192 | 0.1684 | -0.1875 | 0.1788 | -0.2578 | 0.5233 | -0.3390 |
| Area | 0.0434 | 0.1192 | 1.0000 | 0.3835 | 0.1403 | -0.1955 | 0.1289 | 0.0829 | 0.0591 |
| Σ Sand | 0.1724 | 0.1684 | 0.3835 | 1.0000 | 0.0534 | -0.3856 | 0.0403 | -0.1976 | -0.0301 |
| Σ Loam | -0.1996 | -0.1875 | 0.1403 | 0.0534 | 1.0000 | -0.7319 | 0.0374 | -0.1295 | -0.0616 |
| Σ Clay | 0.1564 | 0.1788 | -0.1955 | -0.3856 | -0.7319 | 1.0000 | -0.1272 | 0.3040 | 0.1215 |
| Wood. | -0.3109 | -0.2578 | 0.1289 | 0.0403 | 0.0374 | -0.1272 | 1.0000 | -0.2769 | 0.3873 |
| Crop. | 0.5939 | 0.5233 | 0.0829 | -0.1976 | -0.1295 | 0.3040 | -0.2769 | 1.0000 | -0.4670 |
| I.Pas. | -0.4439 | -0.3390 | 0.0591 | -0.0301 | -0.0616 | 0.1215 | 0.3873 | -0.4670 | 1.0000 |

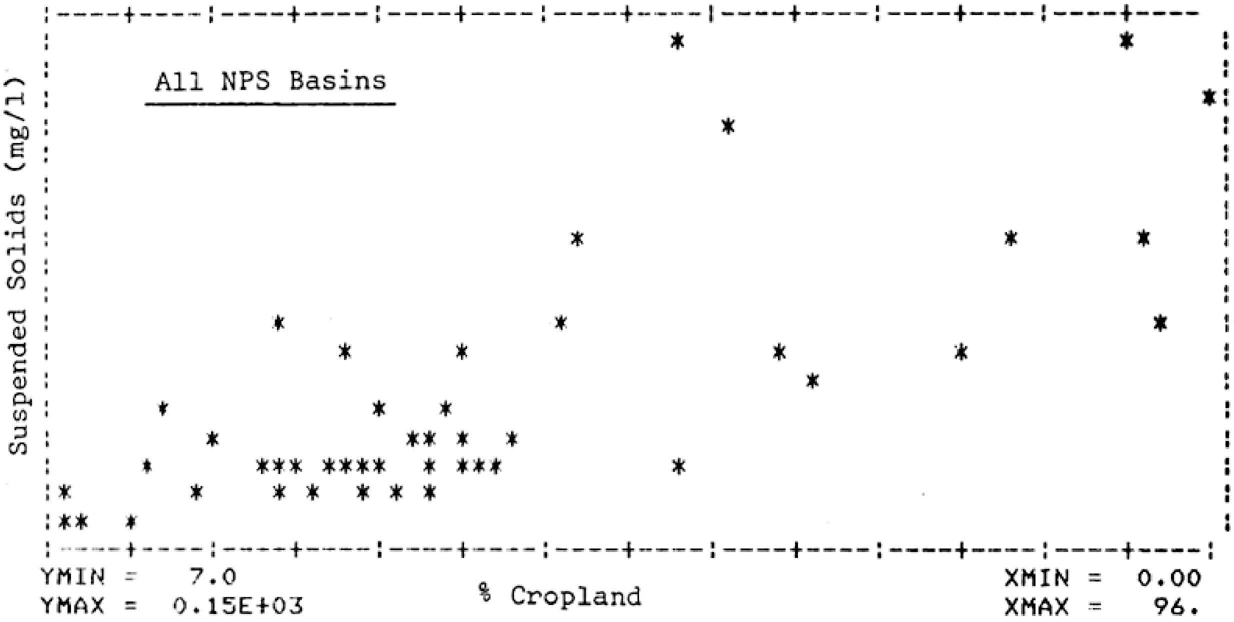

**Figure 7:** Suspended Solids concentration vs. % Cropland.

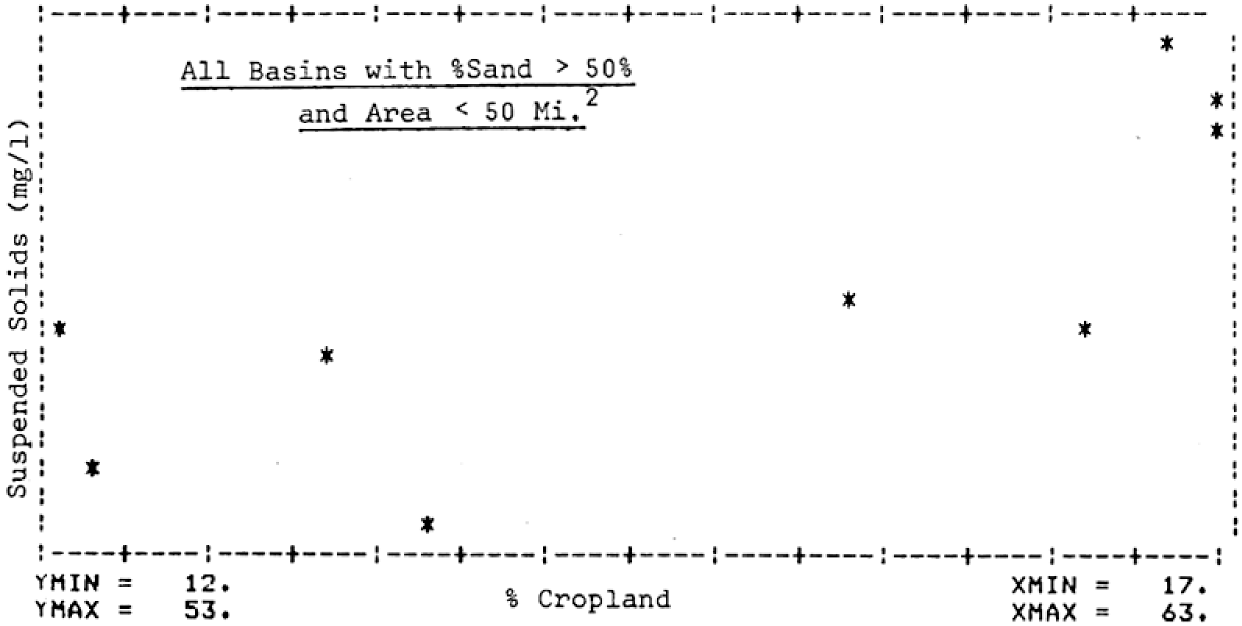

**Figure 8:** Suspended Solids concentration vs. % Cropland

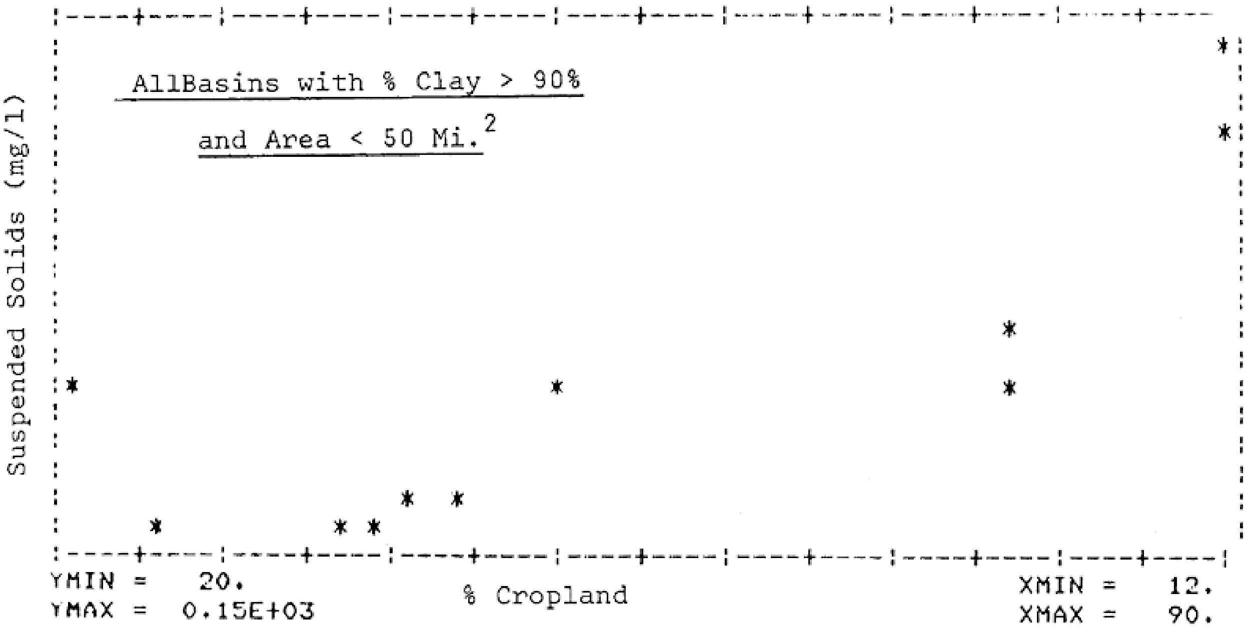

**Figure 9:** Suspended solids concentration vs. % Cropland.

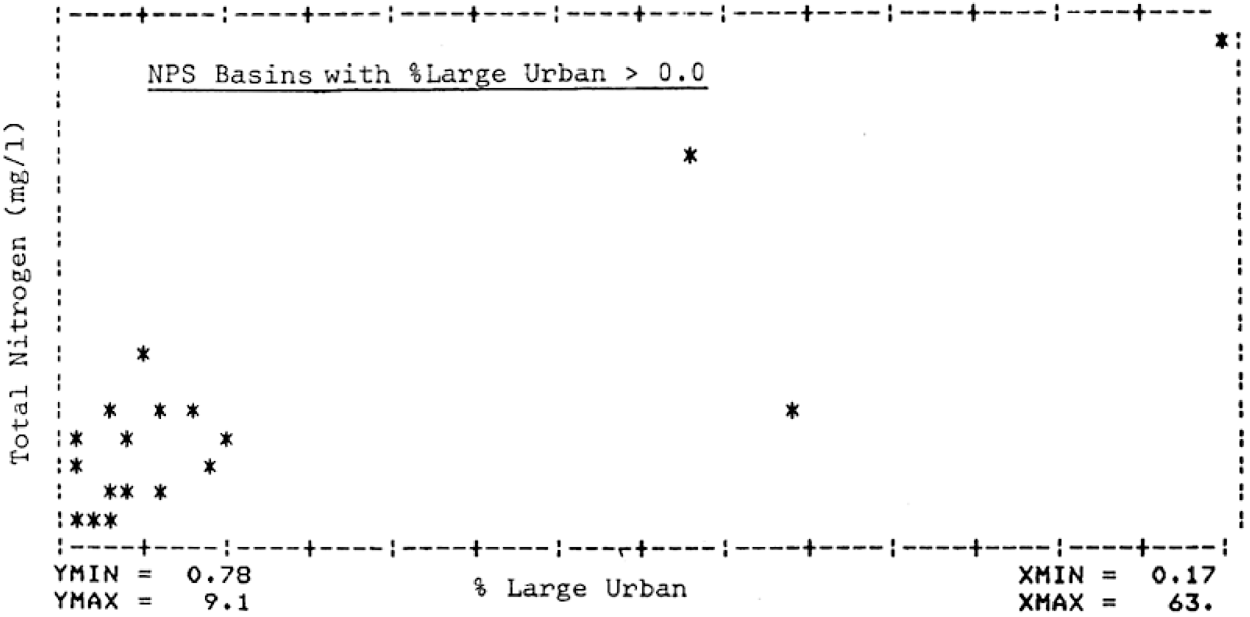

**Figure 10:** Total nitrogen concentration vs. %Large Urban.

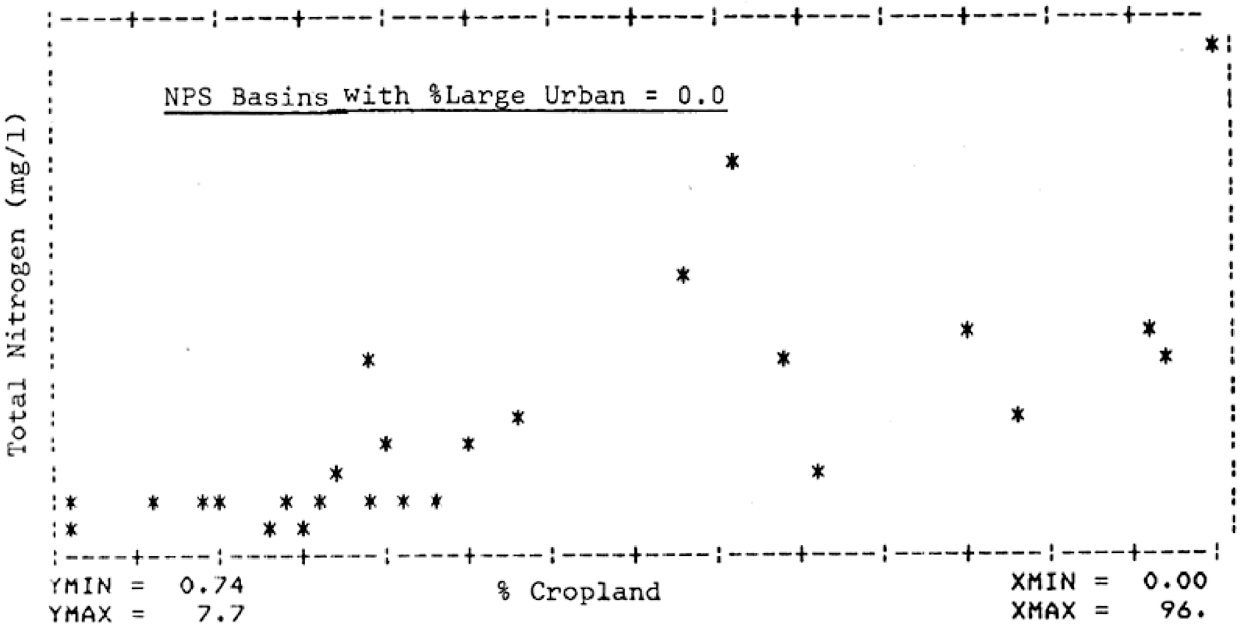

**Figure 11:** Total nitrogen concentration vs. %Cropland.

calculation of yield, is the use of large areas which implicitly takes into account transmission effects upstream of the monitoring site. Nevertheless, one sacrifices the moderate precision in yield values from small area and must accept large variance associated with large area calculations and which results from the inability to specify the variations due to physical, management and crop factors within the larger areas.

Suspended solids, like most variables, illustrate the opposing effects upon concentration of Cropland and Improved Pasture. Improved Pasture is not necessarily tilled each year and therefore is a low sediment producer. It stands to reason that, for basins containing primarily Cropland and Improved Pasture, an increase in area of Improved Pasture will cause a reduction in source of solids (Cropland). Although minimum tillage might be conceived as behaving similar to Improved Pasture (relative to Cropland), one must take cognizance of the fact that agricultural surfaces are primarily spring producers of sediment due to the effect of partial contributing area considerations upon overland flow. Regression of seasonal data should indicate the links between Cropland and suspended solids concentrations in terms of other runoff mechanisms.[1] High correlations in non-spring periods would probably indicate alternative sources of sediment such as underdrainage, bank erosion or site specific problems.

## TOTAL NITROGEN (TN)

### All Basins (Table 29)

Total nitrogen is slightly correlated both with Large Urban (+0.48) and Woodland (-0.48) but not with overburden types.

### NPS Basins (Table 30)

Eliminating point sources raises the correlation coefficient for Large Urban to +0.59 and from +0.23 to +0.48 for Cropland and from -0.38 to -0.57 for Improved Pasture. Hence, it would appear that not only is considerable variability imparted by point sources, but that urban and cropland factors merit investigation. Selecting only NPS basins with Large Urban >0.0% so as to restrict the sample to those basins having some proportion of their area in diffuse urban runoff, results in improvement of the correlation coefficient for Large Urban to +0.85. However, it must be noted that the trend is identified primarily by 3 basins having up to 63% of their area in Large Urban (Figure 10). Considerably more data are required to verify this trend. Whereas Cropland is unimportant (r = +0.09) when basins are selected to display urban effects, selection of basins for which Large Urban = 0.0% produces a correlation of +0.74. The distribution of data (Figure 11) is adequate.

---

[1] As explained in Section 1, seasonal data have not been employed until such time as spatially distributed data within watersheds can be utilized in order to overcome variance introduced by use of large basins within the present data set.

Notably, if one makes a similar selection using all (101) basins, the correlation coefficient for TN concentration and Cropland reduces to +0.59 indicating the role of point and diffuse urban sources. The moderate to high negative correlation of TN concentration with Improved Pasture in many of the selections is not easily explained.

In summary, point sources are sufficiently associated with total nitrogen to obscure relationships between TN and other variables. Although data are limited, TN appears to be directly related to urban diffuse drainage, and in the absence of urban conditions, TN becomes relatively strongly associated with Cropland and inversely with Improved Pastures.

## NITRATE NITROGEN ($NO_3$)

### All Basins (Table 31)

Unlike total nitrogen, $NO_3$, concentration is moderately correlated with Cropland (+0.51) but not with Large Urban. This may imply that point sources are less important than for TN in so far as trends with Cropland are not obscured as is the case for total nitrogen (all basins).

### NPS Basins (Table 32)

Although the relationship between Large Urban and TN concentration for basins with Large Urban >0.0% was limited by lack of data, a similar selection for $NO_3$ produces no relationship with Large Urban (r = +0.10). If the trend for TN proves to be correct and urban diffuse runoff is not responsible for $NO_3$, the nitrogen form in urban TN runoff must be largely of an organic nature. Again, for basins having Large Urban >0.0% the relationship between $NO_3$ concentration and Cropland is moderate (+0.52) yet higher than for TN. Because of the lack of relationship with Large Urban, elimination of basins with Large Urban (Large Urban = 0.0%) does not improve the correlation of $NO_3$ with Cropland. One concludes that, although $NO_3$ does not appear to be particularly related to urban runoff, it is only moderately correlated with cropland. In comparison TN is better correlated with Cropland for rural basins (Large Urban = 0.0%). These results may indicate that mineralization of other nitrogen forms to $NO_3$ during fluvial transport obscures unique associations of $NO_3$ with land use types. If so, the rate of mineralization must form part of the transport function which, in turn, will dictate the optimal distance (basins size) from a diffuse or point source for which $NO_3$ can be expected to bear some unique relationship.

## TOTAL PHOSPHORUS (TP)

Because total and soluble reactive phosphorus are highly correlated (Table 15), total phosphorus only has been selected for discussion purposes. It must also be recalled that these data encompass the years 1968-72. As discussed above (Section 6), phosphorus values fell very considerably in 1972 and subsequent years due, presumably, to phosphorus control measures applied to detergent and municipal wastewater effluent standards. Although the trends indicated here reflect phosphorus flux in the pre-control period, it may be expected that control measures will change only the constants of the estimating equations rather than the trends themselves.

### All Basins (Table 33)

When considering all basins, total phosphorus is associated most strongly with large urban (r = +0.56). Because point sources vary considerably in terms of phosphorus output and are not systematically related to occurrence of other independent variables (e.g., phosphorus effluent from a canning industry in the absence of urban conditions), variance in the trend between TP and Large Urban is expected to be large. It should also be noted that while the Large Urban category is always associated with diffuse urban drainage (where sewage is not discharged into the basin), it may also reflect the presence of municipal wastewater plants and industrial outfalls.

### NPS Basins (Table 34)

In the absence of point sources, the relationship with Large Urban becomes very significant (r = +0.86). The increased correlation with Population (from +0.21 to +0.69) reflects an improved relationship between Population and Large Urban (+0.37 to +0.64). It may be significant that the best trend is between Large Urban rather than with Population, suggesting that diffuse urban drainage (a spatial variable) is more directly related to phosphorus concentration than is population *per se.* One must approach such trends cautiously however, because deletion of point source basins has resulted in elimination of most moderate to high values of Large Urban in the NPS group, leaving very few observations to establish the trend.

One can partially separate the effect of agriculture and that of diffuse urban runoff by selecting basins with Large Urban = 0.0% and those of Large Urban >0.0%. In examining the rural subset (Large Urban = 0.0%) one finds a moderately strong positive correlation (r $^{o}$ +0.63) with Cropland (Figure 12). Considering the range of crop types and management practices on varying soil and slope types, considerable variance is to be expected in the trend between TP concentration and Cropland. Total phosphorus for the urban + rural (Large Urban >0.0%) subset has virtually no relationship (r = -0.13) with Cropland, but is very highly correlated with Large Urban (r = +0.90). As illustrated in Figure 13 the concentration-Large Urban trend is largely established by three data points and must, therefore, be approached with caution.

The urban plus rural subset (Large Urban >0.0%) not only has a much higher mean concentration than the rural subset (.558 mg/L versus .191 mg/L) but, as noted above, Cropland becomes totally unimportant. Bearing in mind the limited data for moderate to large values of the Large Urban category, it appears that while Cropland is positively related to TP production, the presence of urban conditions within a basin totally obscures the effect of Cropland, and causes a substantial increase in TP concentration. This observation is consistent with data tabulated in Uttormark et al. (1974). A test of means between the two data subsets is not significant, an observation which is hardly surprising in view of the large variance associated with the urban plus rural (Large Urban >0.0%) subset.

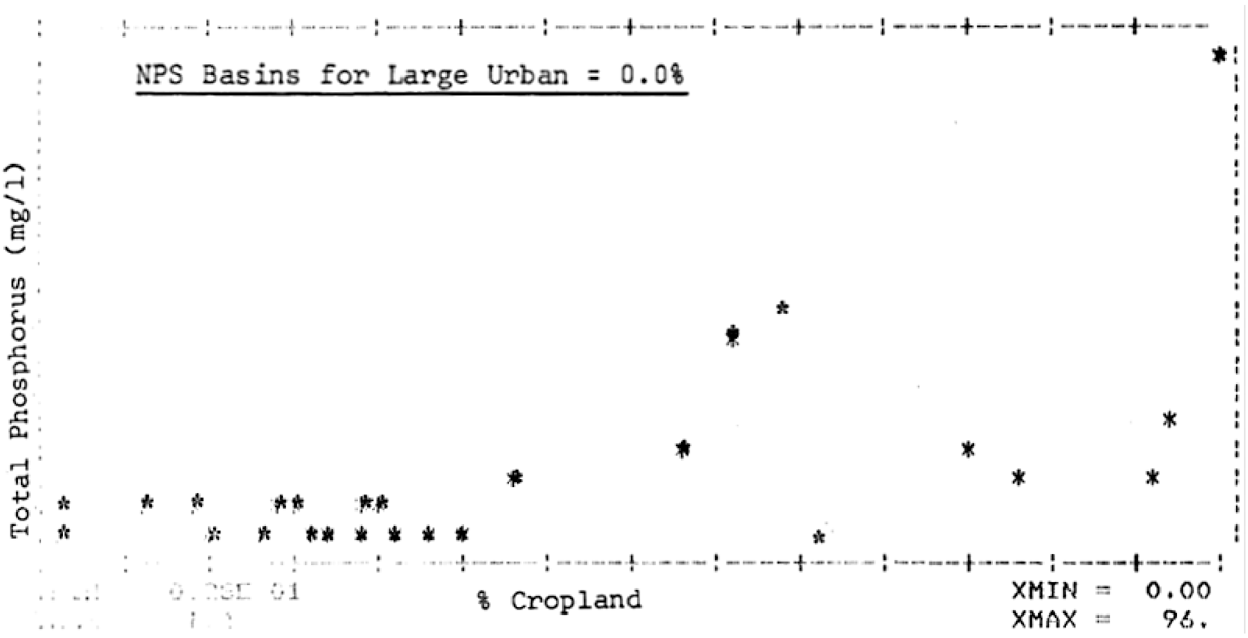

**Figure 12:** Total Phosphorus concentration vs. % Cropland

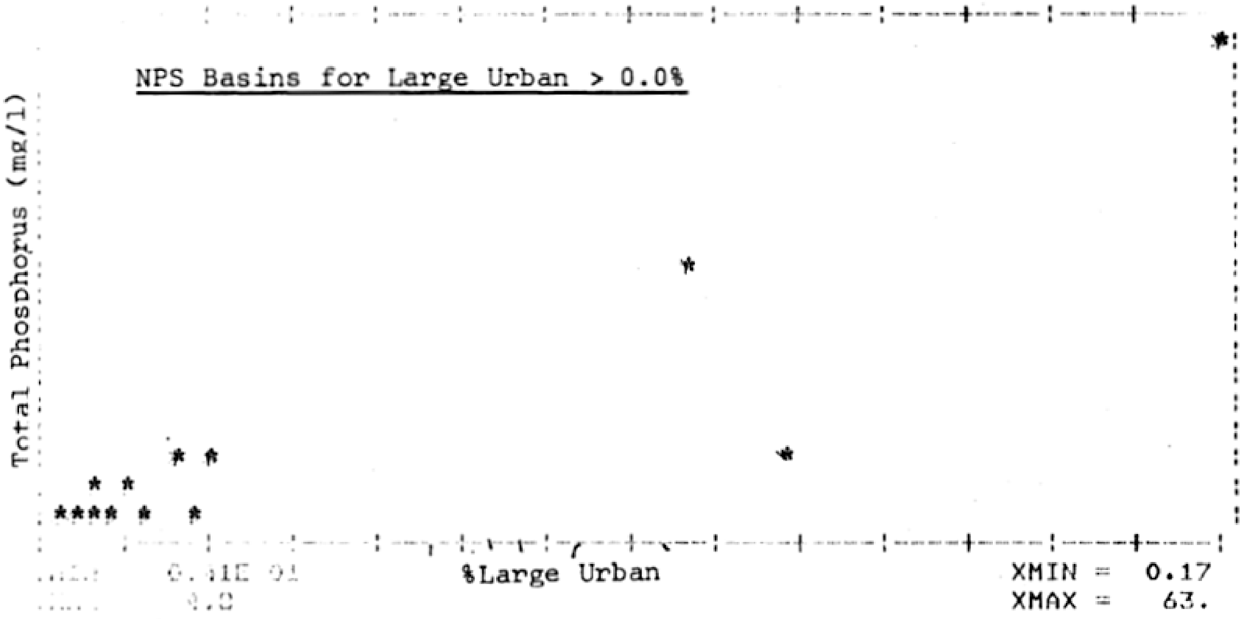

**Figure 13:** Total Phosphorus concentration vs % Large Urban.

Of note is the observation that where strong positive correlation exists between TP concentration and Cropland, there also exists a strong negative correlation with Improved Pasture. Similarly, a weak correlation with one is accompanied by a weak correlation with the other. Where strong correlations exist with TP concentration, Cropland and Improved Pasture are strongly and negatively correlated together, whereas they are weakly correlated together where associations with TP concentration are poor. Although one might expect covariance between Cropland and Improved Pasture when Large Urban is eliminated (Large Urban = 0.0%) as a land use, the explanation lies in the observation that the variance for Woodland, the only other major rural land use category used here, is considerably less for either Cropland or Improved Pasture. It follows that for the rural basin subset, variation in one of Cropland or Improved Pasture is accompanied by commensurate change in the other.

## CHLORIDE (Tables 35 and 36)

Comparison of point and nonpoint basin matrices indicates the role of point sources in chloride concentrations. Whereas the correlation coefficient between chloride concentration and Large Urban for all basins is +0.67, it falls to +0.37 when only NPS basins are considered. The mean concentration and variance is larger for all (101) basins than for the NPS basin group. Because there is no trend associated with other land use categories, it is not surprising that selection of basins having Large Urban >0.0% (elimination of wholly rural basins) both for all basins and the NPS basin group does not increase the relationship between chloride concentration and Large Urban.

## **TABLE 29:** CORRELATION COEFFICIENT MATRIX - TOTAL NITROGEN

### ALL BASINS

### OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9662 | 0.2081 | -0.1352 | 0.4830 | -0.0441 | -0.4795 | 0.2309 | -0.3810 |
| Y | 0.9662 | 1.0000 | 0.1869 | -0.1372 | 0.4132 | 0.0111 | -0.4599 | 0.2652 | -0.3476 |
| Pop. | 0.2081 | 0.1869 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.1352 | 0.1372 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.4830 | 0.4132 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | -0.0441 | 0.0111 | -0.2742 | -0.1989 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.4795 | 0.4599 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.2309 | 0.2652 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.3810 | -0.3476 | --0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

### OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9662 | -0.1352 | -0.2285 | -0.0385 | 0.2326 | -0.4795 | 0.2309 | -0.3810 |
| Y | 0.9662 | 1.0000 | -0.1372 | -0.2451 | -0.0898 | 0.2844 | -0.4599 | 0.2652 | -0.3476 |
| Area | -0.1352 | -0.1372 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.2285 | -0.2451 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | -0.0385 | -0.0898 | 0.0390 | -0.1512 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.2326 | 0.2844 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.4795 | -0.4599 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.2309 | 0.2652 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.3810 | -0.3476 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

## **TABLE 30:** CORRELATION COEFFICIENT MATRIX - TOTAL NITROGEN

### NONPOINT SOURCE BASINS

OVERVIEW

| | TN. C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop | I.Pas |
|---|---|---|---|---|---|---|---|---|---|
| TN.C | 1.0000 | 0.9712 | 0.5286 | -0.1998 | 0.5902 | 0.0526 | -0.5755 | 0.4751 | -0.5728 |
| Y | 0.9712 | 1.0000 | 0.4603 | 0.1830 | 0.5715 | 0.0441 | -0.5338 | 0.4754 | -0.5271 |
| Pop. | 0.5286 | 0.4603 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | -0.1998 | -0.1830 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.5902 | 0.5715 | 0.6400 | -0.1120 | 1.0000 | -0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | 0.0526 | 0.0441 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.5755 | -0.5338 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.4751 | 0.4754 | -0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.5728 | -0.5271 | -0.0315 | 0.2872 | -0.2767 | -0.1390 | 0.4455 | -0.6921 | 1.0000 |

OVERBURDEN - RURAL

| | TN.C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| TN.C | 1.0000 | 0.9578 | -0.3106 | -0.2569 | 0.0015 | 0.3708 | -0.6997 | 0.7393 | -0.6492 |
| Y | 0.9578 | 1.0000 | -0.2783 | -0.2934 | 0.0207 | 0.3831 | -0.6642 | 0.7198 | -0.5668 |
| Area | -0.3106 | -0.2783 | 1.0000 | -0.1181 | 0.0879 | -0.0301 | 0.5135 | -0.3888 | 0.2728 |
| Σ Sand | -0.2569 | -0.2934 | -0.1181 | 1.0000 | -0.3171 | -0.5109 | 0.4084 | -0.0372 | -0.2209 |
| Σ Loam | 0.0015 | 0.0207 | 0.0879 | -0.3171 | 1.0000 | -0.5447 | 0.1557 | -0.3829 | 0.3748 |
| Σ Clay | 0.3708 | 0.3831 | -0.0801 | -0.5109 | -0.5447 | 1.0000 | -0.6753 | 0.5238 | -0.2350 |
| Wood. | -0.6997 | -0.6642 | 0.5135 | 0.4084 | 0.1557 | -0.6753 | 1.0000 | -0.7307 | 0.3710 |
| Crop. | 0.7393 | 0.7198 | -0.3888 | -0.0372 | -0.3829 | 0.5238 | -0.7307 | 1.0000 | -0.8215 |
| I.Pas. | -0.6492 | -0.5668 | 0.2728 | -0.2209 | 0.3748 | -0.2350 | 0.3710 | -0.8215 | 1.0000 |

## **TABLE 31:** CORRECTION COEFFICIENT MATRIX - NITRATE NITROGEN

### ALL BASINS

### OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas.. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9321 | 0.0324 | -0.0746 | 0.1067 | 0.3296 | -0.5065 | 0.5093 | -0.3961 |
| Y | 0.9321 | 1.0000 | 0.0200 | -0.0378 | 0.0325 | 0.3472 | -0.4304 | 0.4978 | -0.3090 |
| Pop. | 0.0324 | 0.0200 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.0746 | -0.0378 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.1067 | 0.0325 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | 0.3296 | 0.3472 | -0.2742 | -0.1989 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.5065 | -0.4304 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.5093 | 0.4978 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.3961 | -0.3090 | -0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

### OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9321 | -0.0746 | -0.1803 | -0.2563 | 0.4139 | -0.5065 | 0.5093 | -0.3961 |
| Y | 0.9321 | 1.0000 | -0.0378 | -0.1869 | -0.2771 | 0.4324 | -0.4304 | 0.4978 | -0.3090 |
| Area | -0.0746 | -0.0378 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.1803 | -0.1869 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | -0.2563 | -0.2771 | 0.0390 | -0.1517 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.4139 | 0.4324 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.5065 | -0.4304 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.5093 | 0.4978 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.3961 | -0.3090 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

## **TABLE 32:** CORRELATION COEFFICIENT MATRIX - NITRATE NITROGEN

### NONPOINT SOURCE BASINS

### OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9605 | 0.2610 | -0.1773 | 0.1040 | 0.3449 | -0.5031 | 0.5204 | -0.4889 |
| Y | 0.9605 | 1.0000 | 0.1534 | -0.1460 | 0.0365 | 0.3907 | -0.4776 | 0.5547 | -0.4647 |
| Pop. | 0.2610 | 0.1534 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | -0.1773 | -0.1460 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.1040 | 0.0365 | 0.6400 | -0.1120 | 1.0000 | 0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | 0.3449 | 0.3907 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.5031 | -0.4776 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.5204 | 0.5547 | 0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.4889 | -0.4647 | -0.0315 | 0.2872 | -0.2767 | 0.1390 | 0.4455 | -0.6921 | 1.0000 |

### OVERBURDEN- RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9605 | -0.1773 | -0.1169 | -0.1407 | 0.2524 | -0.5031 | 0.5204 | -0.4889 |
| Y | 0.9605 | 1.0000 | -0.1460 | -0.1221 | -0.1622 | 0.2687 | -0.4776 | 0.5547 | -0.4647 |
| Area | -0.1773 | -0.1460 | 1.0000 | 0.0085 | 0.1136 | -0.1960 | 0.5283 | -0.3259 | 0.2872 |
| Σ Sand | -0.1169 | -0.1221 | 0.0085 | 1.0000 | -0.1363 | -0.5209 | 0.4471 | -0.0226 | -0.0860 |
| Σ Loam | -0.1407 | -0.1622 | 0.1136 | -0.1363 | 1.0000 | -0.7054 | 0.1534 | -0.3832 | 0.1721 |
| Σ Clay | 0.2524 | 0.2687 | -0.1960 | -0.5209 | -0.7054 | 1.0000 | -0.5666 | 0.4223 | -0.1549 |
| Wood. | -0.5031 | -0.4776 | 0.5283 | 0.4471 | 0.1534 | -0.5666 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.5204 | 0.5547 | -0.3259 | -0.0226 | -0.3832 | 0.4223 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.4889 | -0.4647 | 0.2872 | -0.0860 | 0.1721 | -0.1549 | 0.4455 | -0.6921 | 1.0000 |

## **TABLE 33:** CORRELATION COEFFICIENT MATRIX - TOTAL PHOSPHORUS

### ALL BASINS

### (A) OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9851 | 0.2116 | -0.1016 | 0.5566 | -0.2116 | -0.3393 | 0.0220 | -0.2436 |
| Y | 0.9851 | 1.0000 | 0.1939 | -0.1108 | 0.5302 | -0.1857 | -0.3392 | 0.0430 | -0.2502 |
| Pop. | 0.2116 | 0.1939 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.1016 | -0.1108 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.5566 | 0.5302 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | -0.2116 | -0.1857 | -0.2742 | -0.1989 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.3393 | -0.3392 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop: | 0.0220 | 0.0430 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.2436 | -0.2502 | -0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

### (B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9851 | -0.1016 | -0.1518 | 0.0319 | 0.1137 | -0.3393 | 0.0220 | -0.2436 |
| Y | 0.9851 | 1.0000 | -0.1108 | -0.1459 | -0.0034 | 0.1362 | -0.3392 | 0.0430 | -0.2502 |
| Area | 0.1016 | -0.1108 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.1518 | -0.1459 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | 0.0319 | -0.0034 | 0.0390 | -0.1512 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.1137 | 0.1362 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.3393 | -0.3392 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | 0.0220 | 0.0430 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.2436 | -0.2502 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

**TABLE 34:** CORRELATION COEFFICIENT MATRIX - TOTAL PHOSPHORUS

NONPOINT SOURCE BASINS

OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9880 | 0.6888 | -0.1099 | 0.8613 | -0.3117 | -0.3796 | 0.0852 | -0.3325 |
| y | 0.9880 | 1.0000 | 0.5951 | -0.1183 | 0.8399 | -0.3122 | -0.3566 | 0.0948 | -0.3381 |
| Pop. | 0.6886 | 0.5951 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | -0.1099 | -0.1183 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.8613 | 0.8399 | 0.6400 | -0.1120 | 1.0000 | -0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | -0.3117 | -0.3122 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.3796 | -0.3566 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.0852 | 0.0948 | -0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.3325 | -0.3381 | -0.0315 | 0.2872 | -0.2767 | -0.1390 | 0.4455 | -0.6921 | 1.0000 |

OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9880 | -0.1099 | -0.1252 | -0.0204 | 0.1563 | -0.3796 | 0.0852 | -0.3325 |
| Y | 0.9880 | 1.0000 | -0.1183 | -0.1187 | -0.0023 | 0.1369 | -0.3566 | 0.0948 | -0.3381 |
| Area | -0.1099 | -0.1183 | 1.0000 | 0.0085 | 0.1136 | -0.1960 | 0.5283 | -0.3259 | 0.2872 |
| Σ Sand | -0.1252 | -0.1187 | 0.0085 | 1.0000 | -0.1363 | -0.5209 | 0.4471 | -0.0226 | -0.0860 |
| Σ Loam | -0.0204 | -0.0023 | 0.1136 | -0.1363 | 1.0000 | -0.7054 | 0.1534 | -0.3832 | 0.1721 |
| ΣClay | 0.1563 | 0.1369 | -0.1960 | -0.5209 | -0.7054 | 1.0000 | -0.5666 | 0.4223 | -0.1549 |
| Wood. | -0.3796 | -0.3566 | 0.5283 | 0.4471 | 0.1534 | -0.5666 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.0852 | 0.0948 | -0.3259 | -0.0226 | -0.3832 | 0.4223 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | - 0.3325 | -0.3381 | 0.2872 | -0.0860 | 0.1721 | -0.1549 | 0.4455 | -0.6921 | 1.0000 |

## **TABLE 35:** CORRELATION COEFFICIENT MATRIX - CHLORIDE

### ALL BASINS

OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9489 | 0.3787 | -0.1388 | 0.6747 | -0.2164 | -0.4789 | -0.0548 | -0.2197 |
| Y | 0.9489 | 1.0000 | 0.4146 | -0.1724 | 0.6810 | -0.2564 | -0.4383 | -0.0748 | -0.2369 |
| Pop. | 0.3787 | 0.4146 | 1.0000 | 0.4461 | 0.3715 | -0.2742 | -0.0918 | -0.1087 | -0.0380 |
| Area | -0.1388 | -0.1724 | 0.4461 | 1.0000 | -0.1196 | -0.1989 | 0.3524 | -0.1246 | 0.0755 |
| Lg.Urb | 0.6747 | 0.6810 | 0.3715 | -0.1196 | 1.0000 | -0.5918 | -0.4000 | -0.2540 | -0.2864 |
| Agr. | -0.2164 | -0.2564 | -0.2742 | -0.1989 | -0.5918 | 1.0000 | -0.4883 | 0.6330 | 0.0683 |
| Wood. | -0.4789 | -0.4383 | -0.0918 | 0.3524 | -0.4000 | -0.4883 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | -0.0548 | -0.0748 | -0.1087 | -0.1246 | -0.2540 | 0.6330 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.2197 | -0.2369 | -0.0380 | 0.0755 | -0.2864 | 0.0683 | 0.2715 | -0.5802 | 1.0000 |

OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9489 | -0.1388 | -0.2515 | -0.0621 | 0.2709 | -0.4789 | -0.0548 | -0.2197 |
| Y | 0.9489 | 1.0000 | -0.1724 | -0.2597 | -0.0872 | 0.2898 | -0.4383 | -0.0748 | -0.2369 |
| Area | -0.1388 | -0.1724 | 1.0000 | 0.0906 | 0.0390 | -0.1426 | 0.3524 | -0.1246 | 0.0755 |
| Σ Sand | -0.2515 | -0.2597 | 0.0906 | 1.0000 | -0.1512 | -0.5796 | 0.4805 | 0.0245 | -0.1788 |
| Σ Loam | -0.0621 | -0.0872 | 0.0390 | -0.1512 | 1.0000 | -0.6375 | 0.1589 | -0.4190 | 0.2478 |
| Σ Clay | 0.2709 | 0.2898 | -0.1426 | -0.5796 | -0.6375 | 1.0000 | -0.5897 | 0.4096 | -0.1190 |
| Wood. | -0.4789 | -0.4383 | 0.3524 | 0.4805 | 0.1589 | -0.5897 | 1.0000 | -0.4514 | 0.2715 |
| Crop. | -0.0548 | -0.0748 | -0.1246 | 0.0245 | -0.4190 | 0.4096 | -0.4514 | 1.0000 | -0.5802 |
| I.Pas. | -0.2197 | -0.2369 | 0.0755 | -0.1788 | 0.2478 | -0.1190 | 0.2715 | -0.5802 | 1.0000 |

## **TABLE 36:** CORRELATION COEFFICIENT MATRIX - CHLORIDE

### NONPOINT SOURCE BASINS

### (A) OVERVIEW

| | C | Y | Pop. | Area | Lg.Urb | Agr. | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9729 | 0.4411 | -0.2774 | 0.3666 | 0.1309 | -0.4613 | 0.1632 | -0.3170 |
| Y | 0.9729 | 1.0000 | 0.3772 | -0.2982 | 0.3497 | 0.1133 | -0.4196 | 0.1420 | -0.2790 |
| Pop. | 0.4411 | 0.3772 | 1.0000 | 0.0982 | 0.6400 | -0.2850 | -0.2103 | -0.0926 | -0.0315 |
| Area | -0.2774 | -0.2982 | 0.0982 | 1.0000 | -0.1120 | -0.4227 | 0.5283 | -0.3259 | 0.2872 |
| Lg.Urb | 0.3666 | 0.3497 | 0.6400 | -0.1120 | 1.0000 | -0.4557 | -0.3701 | -0.1358 | -0.2767 |
| Agr. | 0.1309 | 0.1133 | -0.2850 | -0.4227 | -0.4557 | 1.0000 | -0.6361 | 0.6334 | -0.1390 |
| Wood. | -0.4613 | -0.4196 | -0.2103 | 0.5283 | -0.3701 | -0.6361 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.1632 | 0.1420 | -0.0926 | -0.3259 | -0.1358 | 0.6334 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.3170 | -0.2790 | -0.0315 | 0.2872 | -0.2767 | -0.1390 | 0.4455 | -0.6921 | 1.0000 |

### (B) OVERBURDEN - RURAL

| | C | Y | Area | Σ Sand | Σ Loam | Σ Clay | Wood. | Crop. | I.Pas. |
|---|---|---|---|---|---|---|---|---|---|
| C | 1.0000 | 0.9729 | -0.2774 | -0.2422 | -0.2696 | 0.4065 | -0.4613 | 0.1632 | -0.3170 |
| Y | 0.9729 | 1.0000 | -0.2982 | -0.2660 | -0.2823 | 0.4310 | -0.4196 | 0.1420 | -0.2790 |
| Area | -0.2774 | -0.2982 | 1.0000 | 0.0085 | 0.1136 | -0.1960 | 0.5283 | -0.3259 | 0.2872 |
| Σ Sand | -0.2422 | -0.2660 | 0.0085 | 1.0000 | -0.1363 | -0.5209 | 0.4471 | -0.0226 | -0.0860 |
| Σ Loam | -0.2696 | -0.2823 | 0.1136 | -0.1363 | 1.0000 | -0.7054 | 0.1534 | -0.3832 | 0.1721 |
| ΣClay | 0.4065 | 0.4310 | -0.1960 | -0.5209 | -0.7054 | 1.0000 | -0.5666 | 0.4223 | -0.1549 |
| Wood. | -0.4613 | -0.4196 | 0.5283 | 0.4471 | 0.1534 | -0.5666 | 1.0000 | -0.5672 | 0.4455 |
| Crop. | 0.1632 | 0.1420 | -0.3259 | -0.0226 | -0.3832 | 0.4223 | -0.5672 | 1.0000 | -0.6921 |
| I.Pas. | -0.3170 | -0.2790 | 0.2872 | -0.0860 | 0.1721 | -0.1549 | 0.4455 | -0.6921 | 1.0000 |

# REFERENCES CITED

Acres Consulting Services Ltd. 1977. Atmospheric Loadings of the Lower Great Lakes and the Great Lakes Drainage Basin. Report prepared for the Canada Centre For Inland Waters, 70p.

Baker, D.B., and J.W. Kramer. 1973. Phosphorus sources and transport in an agricultural river basin of Lake Erie. Proceedings, 16th Conference on Great Lakes Research, International Association for Great Lakes Research, pp.858-871.

Baker, D.B., and J.W. Kramer. 1975. Effects of advanced waste treatment and flow augmentation on water quality during low stream flows, in Baker, D.B.

J.W. Kramer and B.L. Prater (editors), Sandusky River Basin Symposium, Proceedings. International Joint Commission, 1975, pp.124-142.

Casey, D.J., and S.E. Salbach. 1974. IFYGL stream material balance study(IFYGL). Proceedings, 17th Conference on Great Lakes Research, International Association for Great Lakes Research, pp.668-687.

Chapra, S.C. 1977. Total phosphorus model for the Great Lakes. Journal of the Environmental Engineering Division, ASCE, pp.147-161.

Dillon, P.J., and W.B. Kirchner. 1975. The effects of geology and land use on the export of phosphorus from watersheds. Water Research (9), pp.135-148.

Environment Canada. 1976. The Canada Water Act Annual Report, 1975-1976. Ottawa.

Hopson, N.E. 1975. Phosphorus removal by legislation. Water Resources Bulletin(11), pp.356-364.

King, D.E. 1975. Overview of Laboratory Services Branch involvement in studies of the Great Lakes system nearshore waters since 1965. in-house memorandum, Ontario Ministry of the Environment, 8p.

Logan, T.J. 1977. Chemical and mineralogical indices of sediment transformation during fluvial transport. Proceedings of the 5th Guelph Symposium on Geomorphology (forthcoming).

Omernik, J.M. 1976. The influence of land use on stream nutrient levels, U.S. EPA-600/3-76-014, Corvallis.

Ongley, E.D. 1974. Hydrophysical characteristics of Great Lakes Tributary Drainage, Canada . Report in 5 volumes prepared for (Canadian) Task D of the Pollution From Land Use Activities Reference Group, International Joint Commission.

Ongley, E.A. 1976. Sediment yields and nutrient loadings from Canadian watersheds tributary to Lake Erie: an overview . Journal of the Fisheries Research Board of Canada (33), pp.471-484.

Ongley, E.D. 1977a. Study of statistical evaluation, and update of data bank with discharge and water quality data for the Great Lakes Interconnecting channels and associated tributaries. Report prepared for (Canadian) Task D of the Pollution from Land Use Activities Group, International Joint Commission.

Ongley, E.D. 1977b. Land use, water quality and river-mouth loadings; a regional perspective for Ontario. Draft report prepared for (Canadian) Task D of the Pollution from Land Use Activities Reference Group, International Joint Commission.

Ongley, E.D., Ralston, J.G., and R.L. Thomas. 1977. Sediment and nutrient loadings to Lake Ontario: Methodological arguments. Canadian Journal of Earth Sciences (14), pp.1555-1565.

Sweeney, R.A. 1973. Impact of detergent phosphate reductions on water quality. Proceedings , 16th Conference on Great Lakes Research, International Association for Great Lakes Research, pp.967-976.

Statistics Canada, Fertilizer Trade, Catalogue 46-207. 1969 through 1974.

U.S. Army Corps of Engineers, Buffalo District. 1975. Lake Erie Wastewater Management Study. Preliminary Feasibility Report, 3 vols. Buffalo, N.Y.

U.S. Environmental Protection Agency. 1976. Loading functions for assessment of water pollution from nonpoint sources. EPA-600/2-76-151, Washington, 445p.

Uttormark, P.D., Chapin, J.D. and K.M. Green. 1974. Estimating nutrient loadings of lakes from non-point sources. EPA-660/3-74-020. Washington.

# APPENDIX 1

## RANKED LOADINGS AND YIELD DATA FOR ALL (101) TRIBUTARY BASINS IN SOUTHERN ONTARIO, AND FOR POINT AND NONPOINT SOURCE BASINS.

(bound and circulated separately)

# APPENDIX 2

## DESCRIPTIVE STATISTICS FOR 1968-1972 DATA SET OF TABLE 1

| VARIABLE NAME(NC.) | MEAN | STD.DEV. | MINIMUM | MAXIMUM | SKEWNESS | KURTOSIS |
|---|---|---|---|---|---|---|
| TOTAL DISSOLVED SOLIDS (2) | 373.595 | 136.375 | 117.107 | 835.450 | 1.077 | 4.0 |
| TOTAL SOLIDS (17) | 414.286 | 148.508 | 125.086 | 894.806 | 0.852 | 3.5 |
| SUSPENDED SOLIDS (32) | 43.815 | 31.339 | 4.464 | 148.288 | 1.359 | 4.7 |
| NH3-N(47) | 0.556 | 1.222 | 0.037 | 6.879 | 3.552 | 15.5 |
| TOTAL KJELDAL (62) | 1.547 | 1.991 | 0.368 | 11.330 | 3.263 | 13.7 |
| NO2-N (77) | 0.047 | 0.066 | 0.004 | 0.368 | 2.695 | 10.3 |
| NO3-N (92) | 1.038 | 0.800 | 0.031 | 4.554 | 1.463 | 5.6 |
| TOTAL NITROGEN (62 + 77 + 92) | 0.476 | 1.035 | 0.027 | 5.836 | 3.694 | 16.9 |
| TOTAL PHOSPHORUS (107) | 0.229 | 0.567 | 0.006 | 3.419 | 3.023 | 18.9 |
| SOLUBLE REACTIVE P (122) | 243.571 | 60.690 | 91.233 | 458.063 | 0.287 | 4.4 |
| HARDNESS (137) | 173.983 | 40.695 | 69.971 | 260.824 | -0.629 | 2.7 |
| ALKALINITY (152) | 42.747 | 43.644 | 4.250 | 238.759 | 1.971 | 6.7 |
| CHLORIDE (167) | 1.390 | 0.894 | 0.202 | 4.811 | 1.068 | 4.4 |
| TOTAL IRON (182) | 2.632 | 2.338 | 0.403 | 12.772 | 2.416 | 9.3 |
| BASIN POPULATION (196) | 42270.871 | 107377.148 | 68.000 | 751062.000 | 4.337 | 24.2 |
| BASIN AREA (197) | 238.182 | 616.744 | 0.713 | 4472.425 | 4.438 | 26.0 |
| SAND (198 + 202 + 205) | 20.238 | 27.234 | 0.000 | 100.000 | 1.440 | 4.1 |
| LOAM-SILT (199 + 203 + 206) | 25.842 | 28.452 | 0.000 | 100.000 | 0.924 | 2.8 |
| CLAY (200 + 204 + 207) | 46.624 | 37.909 | 0.000 | 101.000 | 0.174 | 1.4 |
| URBAN >25K (212) | 6.709 | 15.388 | 0.000 | 67.529 | 2.807 | 10.2 |
| URBAN <25K (213) | 1.508 | 1.569 | 0.000 | 9.475 | 1.868 | 8.4 |
| EXTRACTIVE (214) | 0.389 | 1.036 | 0.000 | 7.877 | 4.699 | 29.8 |
| AGRICULTURE (215) | 71.352 | 17.349 | 22.049 | 97.448 | -1.020 | 3.7 |
| WOODLAND (216) | 19.071 | 15.177 | 0.000 | 69.153 | 1.025 | 3.8 |
| MARSH, SWAMP, BARREN (217) | 0.972 | 1.988 | 0.000 | 13.593 | 3.843 | 21.3 |
| LOW DENSITY COMMERICAL (218) | 0.037 | 0.252 | 0.000 | 2.392 | 8.397 | 76.5 |
| MED DENSITY COMMERICAL (219) | 0.229 | 0.880 | 0.000 | 7.111 | 5.629 | 39.7 |
| HIGH DENSITY COMMERICAL (220) | 1.385 | 4.010 | 0.000 | 24.913 | 3.631 | 17.1 |
| LOW DENSITY RESIDENTIAL (221) | 0.752 | 1.888 | 0.000 | 11.793 | 3.542 | 16.9 |
| MED DENSITY RESIDENTIAL (222) | 2.411 | 6.333 | 0.000 | 41.680 | 3.836 | 19.8 |
| HIGH DENSITY RESIDENTIAL (223) | 1.338 | 4.511 | 0.000 | 28.904 | 4.244 | 22.3 |
| TRANSPORTATION (224) | 0.556 | 1.543 | 0.000 | 8.165 | 3.246 | 13.5 |
| ORCHARDS, HORTICULTURE (225) | 4.326 | 10.419 | 0.000 | 49.745 | 2.876 | 10.4 |
| CROPLANDS (226) | 34.058 | 23.635 | 0.000 | 96.334 | 0.799 | 3.1 |
| IMPROVED PASTURE (227) | 24.075 | 16.602 | 0.000 | 58.141 | -0.030 | 1.9 |
| UNIMPROVED PASTURE (228) | 8.893 | 7.353 | 0.000 | 44.157 | 1.568 | 7.0 |

# APPENDIX 3

## DESCRIPTIVE STATISTICS BY YEAR PER LAKE GROUP OF TRIBUTARY BASINS: PHOSPHORUS, CHLORIDE

# DESCRIPTIVE STATISTICS[a] BY YEAR PER LAKE GROUP OF TRIBUTARY BASINS
## PHOSPHORUS AND CHLORIDE

| | TOTAL PHOSPHORUS | | | | SOLUBLE PHOSPHORUS | | | | CHLORIDE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Standard Deviation | Minimum | Maximum | Mean | Standard Deviation | Minimum | Maximum | Mean | Standard Deviation | Minimum | Maximum |
| LAKE ONTARIO Point Source (26) | | | | | | | | | | | | |
| 1969 | .960 | 1.790 | .062 | 8.239 | .492 | 1.048 | .014 | 4.496 | 68.83 | 66.96 | 5.9 | 239.8 |
| 1970 | .807 | 1.260 | .040 | 4.608 | .367 | .622 | .006 | 2.253 | 70.60 | 63.76 | 6.0 | 241.7 |
| 1971 | .646 | 1.060 | .031 | 3.740 | .279 | .525 | .005 | 1.777 | 75.49 | 55.78 | 5.5 | 201.2 |
| 1972 | .446 | .794 | .039 | 3.527 | .177 | .321 | .004 | 1.250 | 66.41 | 51.79 | 6.5 | 159.3 |
| 1973 | .308 | .438 | .010 | 1.583 | .146 | .259 | .002 | 1.126 | 55.46 | 41.97 | 6.1 | 144.8 |
| 1974 | .367 | .573 | .035 | 2.497 | .170 | .372 | .002 | 1.579 | 86.50 | 92.27 | 5.5 | 364.9 |
| LAKE ONTARIO Nonpoint Source (17) | | | | | | | | | | | | |
| 1969 | .352 | .640 | .055 | 2.485 | .194 | .428 | .010 | 1.634 | 45.51 | 51.93 | 5.9 | 157.7 |
| 1970 | .247 | .553 | .030 | 2.364 | .134 | .358 | .008 | 1.498 | 47.39 | 50.89 | 7.3 | 187.0 |
| 1971 | .308 | .585 | .029 | 2.397 | .152 | .443 | .004 | 1.847 | 48.91 | 47.79 | 7,2 | 177.4 |
| 1972 | .168 | .243 | .027 | 1.028 | .082 | .178 | .006 | .740 | 50.02 | 56.10 | 5.4 | 227.7 |
| 1973 | .122 | .140 | .021 | .495 | .053 | .087 | .006 | .346 | 37.39 | 32.48 | 7.0 | 117.1 |
| 1974. | .149 | .176 | .023 | .571 | .053 | .079 | .007 | .265 | 52.37 | 63.73 | 6.0 | 245.9 |
| LAKE ERIE Point Source (9) | | | | | | | | | | | | |
| 1969 | .176 | .113 | .071 | .432 | .105 | .105 | .019 | .330 | 24.77 | 12.54 | 6.5 | 51.2 |
| 1970 | .251 | .308 | .056 | 1.045 | .149 | .238 | .017 | .767 | 23.25 | 9.06 | 6.7 | 33.6 |
| 1971. | .186 | .130 | .053 | .406 | .092 | .095 | .014 | .266 | 25.01 | 11.00 | 6.1 | 37.9 |
| 1972 | .229 | .176 | .054 | .644 | .100 | .113 | .014 | .379 | 25.22 | 9.48 | 7.6 | 33,9 |
| 1973 | .198 | .108 | .074 | .322 | .077 | .065 | .018 | .208 | 24.42 | 9.28 | 8.9 | 37.1 |
| 1974 | .152 | .073 | .052 | .228 | .044 | .034 | .010 | .099 | 25.74 | 10.49 | 11.5 | 45.7 |
| LAKE ERIE Nonpoint Source (5) | | | | | | | | | | | | |
| 1969 | .398 | .352 | .039 | .597 | .213 | .267 | .010 | .663 | 36.32 | 25.08 | 5.8 | 68.7 |
| 1970 | .329 | .226 | .054 | .564 | .184 | .159 | .011 | .380 | 36.10 | 24.11 | 5.7 | 65.0 |
| 1971 | .452 | .363 | .045 | .873 | .299 | .285 | .009 | .610 | 39.98 | 26.23 | 5.0 | 73.7 |
| 1972 | .312 | .182 | .137 | .521 | .194 | .157 | .027 | .384 | 44.22 | 22.18 | 6.0 | 59.6 |
| 1973 | ,172 | .079 | .084 | .281 | .071 | .044 | .012 | .114 | 30.62 | 17.99 | 7.8 | 56.6 |
| 1974 | .174 | .094 | .054 | .282 | .062 | .039 | .007 | .097 | 36.78 | 20.01 | 8.0 | 58.7 |
| CONNECTING WATERWAYS Point Source (5) | | | | | | | | | | | | |
| 1969 | .370 | .264 | .146 | .827 | .196 | .204 | .068 | .555 | 34.88 | 18.81 | 17.6 | 66.0 |
| 1970 | .711 | .978 | .113 | 2.449 | .408 | .643 | .060 | 1.552 | 36.08 | 8.73 | 25.9 | 48.2 |
| 1971 | .670 | 1.059 | .097 | 2.559 | .484 | .898 | .035 | 2.089 | 39.50 | 19.98 | 19.7 | 72.0 |
| 1972 | .688 | .751 | .155 | 1.906 | .442 | .519 | .049 | 1.231 | 54.94 | 34.52 | 20.5 | 109.4 |
| 1973 | .631 | .597 | .133 | 1.559 | .404 | .521 | .057 | 1.267 | 39.72 | 19.16 | 14.5 | 66.0 |
| 1974 | .432 | .546 | .101 | 1.395 | .211 | .304 | .041 | .748 | 43.06 | 22.79 | 15.1 | 75.5 |
| CONNECTING WATERWAYS Nonpoint Source (1) | | | | | | | | | | | | |
| 1969 | .152 | | | | .026 | | | | 60.90 | | | |
| 1970 | .046 | | | | .010 | | | | 95.50 | | | |
| 1971 | .067 | | | | .018 | | | | 73.05 | | | |
| 1972 | .061 | | | | .019 | | | | 77.75 | | | |
| 1973 | .20 | | | | .145 | | | | 62.30 | | | |
| 1974 | .212 | | | | .107 | | | | 59.55 | | | |
| LAKE HURON-GEORGIAN BAY Point Source[b], (8) | | | | | | | | | | | | |
| 1969 | .140 | .246 | .022 | .742 | .080 | .162 | .015 | .480 | 42.59 | 75.81 | 6.0 | 226.1 |
| 1970 | .127 | .208 | .023 | .638 | .068 | .128 | .004 | .364 | 35.55 | 50.15 | 6.2 | 154.9 |
| 1971 | .126 | .183 | .018 | .573 | .062 | .138 | .003 | .403 | 39.06 | 63.72 | 7.1 | 191.5 |
| 1972 | .106 | .127 | .024 | .416 | .062 | .095 | .007 | .292 | 31.58 | 39.79 | 7.5 | 119.9 |
| 1973 | .127 | .100 | .040 | .316 | .038 | .033 | .011 | .109 | 23.35 | 20.19 | 8.8 | 61.9 |
| 1974 | .095 | .069 | .016 | .240 | .022 | .019 | .003 | .057 | 27.88 | 37.81 | 7.5 | 118.6 |
| LAKE HURON-GEORGIAN BAY Nonpoint Source (6) | | | | | | | | | | | | |
| 1969 | .050 | .024 | .024 | .090 | .020 | .011 | .009 | .035 | 8.32 | 3.53 | 3.7 | 13.0 |
| 1970 | .116 | .095 | .041 | .293 | .036 | .041 | .006 | .115 | 9.06 | 3.19 | 4.1 | 12.6 |
| 1971 | .073 | .036 | .028 | .128 | .020 | .009 | .007 | .032 | 12.46 | 13.31 | 2.3 | 42.8 |
| 1972 | .038 | .008 | .028 | .050 | .010 | .002 | .007 | .014 | 10.60 | 4.88 | 4.6 | 15.3 |
| 1973 | .058 | .037 | .030 | .124 | .011 | .007 | .006 | .025 | 13.20 | 7.98 | 4.9 | 26.8 |
| 1974 | .050 | .018 | .025 | .069 | .010 | .006 | .005 | .020 | 11.40 | 4.55 | 4.8 | 17.7 |

[a] Statistics derived from annual mean concentration per basin in Lake group.

[b] Chloride is profoundly influenced by salt mine effluent in Maitland River. ( ) Number of basins in group