United States
Environmental Protection
Agency

Office of Water Regulations
and Standards
Criteria & Standards Division
Washington DC 20460

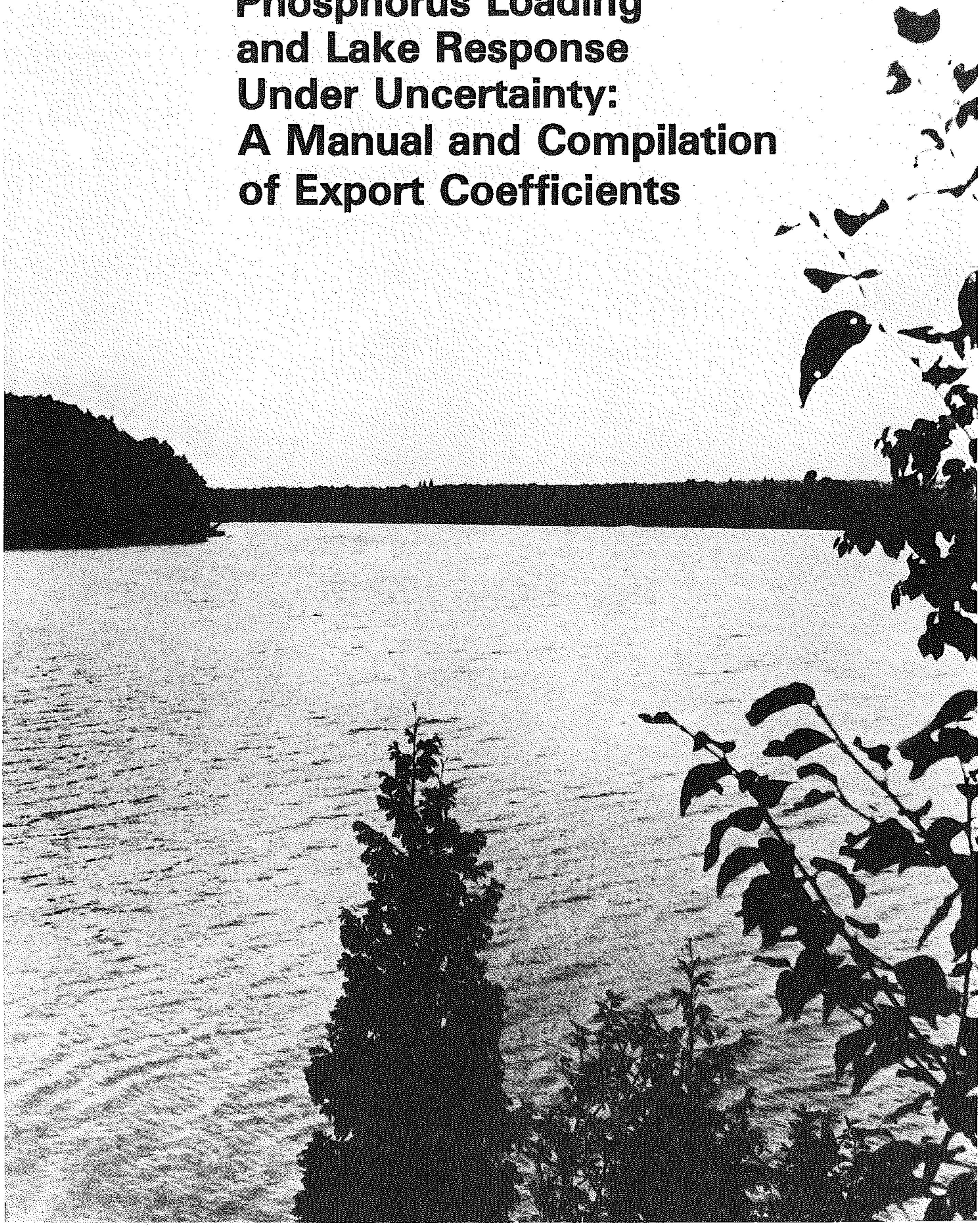
EPA 440/5-80-011

Water

D. RECKHOW

EPA

Modeling Phosphorus Loading and Lake Response Under Uncertainty: A Manual and Compilation of Export Coefficients



MODELING PHOSPHORUS LOADING AND LAKE RESPONSE UNDER UNCERTAINTY:
A MANUAL AND COMPILATION OF EXPORT COEFFICIENTS

By

Kenneth H. Reckhow
Michael N. Beaulac
Jonathan T. Simpson

Michigan State University

June, 1980

Project Manager:
Robert J. Johnson, Chief
CLEAN LAKES SECTION
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

ABSTRACT

A procedure is proposed that may be used to quantify the relationship between land use and lake trophic quality. This methodology, based on an input-output phosphorus lake model, is presented in a step-by-step manner and illustrated through example. An important part of this procedure is a section describing the estimation of nonparametric prediction intervals. These intervals quantify the total prediction uncertainty which is a measure of information value contained in a prediction.

When the methodology is employed to predict the impact of projected land use changes, it is necessary to use phosphorus export coefficients extrapolated from other points in time and/or space. These coefficients represent the mass loading of phosphorus to a surface water body per year per unit of source (e.g., per hectare of forested land). A substantial portion of this document is devoted to a presentation of carefully screened nutrient export coefficients. These values are intended for inclusion in the modeling/uncertainty analysis methodology. To that end, criteria are described that will aid the analyst in the selection of appropriate export coefficients and in the interpretation of the results of an application of this methodology.

CONTENTS

CHAPTERS

APPENDIX

FIGURES

APPENDIX FIGURES

TABLES

APPENDIX TABLES

ACKNOWLEDGMENTS

A number of people assisted in the development and preparation of this document. Appreciation is extended to Robert J. Johnson, Chief of the Clean Lakes Section, for providing the opportunity to prepare this report. Thanks are also due to Janine Niemer, Patsy Johnston, and Barb Visser for typing the document, to Paul Schneider for graphics work, and to David Lee, Ralph Ancil, and Robert Montgomery for editorial assistance and proofreading.

Chapter 1

INTRODUCTION

The planner or engineer concerned with the issue of lake trophic quality management must be conversant across several disciplines. Eutrophication is fundamentally a problem in chemical and biological limnology, so the analyst must be familiar with the natural sciences. Methods and processes for managing watershed characteristics and human activities that determine water quality have an engineering basis with strong economic and sociologic features. Finally, techniques that project the link between the engineered management strategies and the limnological water quality are in the planning arena yet require a good foundation in statistics and mathematics. It is this latter set of planning methods, with a firm basis in the statistical understanding and description of empirical relationships, that is the focus of this manual.

Lake eutrophication is both a natural and culturally-induced phenomenon. Natural eutrophication is a slow, largely irreversible process associated with the gradual accumulation of organic matter and sediments in lake basins. Cultural eutrophication is an often rapid, possibly reversible process of nutrient enrichment and high biomass production stimulated by cultural activities causing nutrient transport to lakes. Eutrophication is a complex process, and hence the reference is made above to the importance of chemical and limnological knowledge for proper lake management. Each lake is unique, and a study and understanding of the unique features are essential for good planning. Yet there are also characteristics of watershed and lake behavior that are, if not universal, certainly shared by many lake systems. The planner can exploit this commonality in management studies, providing he/she is aware of where the commonality ends and the uniqueness begins. In fact, to the degree that planning studies must surely depend upon the efficient use of resources, all planning is probably a compromise between unique and common features.

This manual is based on the aforementioned premise of similar trophic behavior among lakes. This is both a strength and a weakness. Its strength lies in the fact that the methodologies are not necessarily lake-specific, so that models and data are transferrable, keeping analysis costs low. Its weakness is that the ease of application of the methodology and statistics can foster inappropriate use of the techniques described herein and incomplete study of the unique features of a lake. This lowers planning costs but increases risks associated with poor planning decisions. Again, the lesson is to know the limitations of the general methodology.

The techniques presented in this manual for lake trophic management planning are based on control of the nutrient phosphorus. There are two reasons for this reliance on phosphorus:

1. Phosphorus is often the major nutrient in shortest supply relative to the nutritional needs of algae and aquatic plants. This means that the concentration of phosphorus is frequently a prime determinant of the total biomass in a lake.
2. Of the major nutrients, phosphorus is the most effectively controlled using existing engineering technology and land use management.

In general terms, with phosphorus as the controlling mechanism, the reasons suggest that proper management of human activities in the watershed can often be effective in the maintenance of desirable biomass levels in the lake.

Figure 1 presents a schematic diagram of the watershed and lake ecosystem as viewed from the perspective of phosphorus movement. Human activities (including land use), watershed characteristics, and climate are the general determinants of phosphorus mass transport to lakes. The phosphorus loading to a lake is empirically related to the phosphorus concentration in a lake as a function of the hydrologic and geomorphologic characteristics of the lake. Phosphorus concentration, in turn, is causally linked with biomass levels, water clarity, dissolved oxygen concentrations, and fish populations, which are all empirically interrelated.

The methodology presented herein is based on the phosphorus flow schematic in Figure 1. Historically, it is derived from the work of Vollenweider (1968, 1975) on the phosphorus loading concept. Vollenweider's contribution to this field was the recognition of the similiarities among lakes in trophic response to nutrient input. He defined nutrient loading criteria for lakes as a function of selected hydrologic and geomorphologic characteristics (e.g., mean depth or areal water loading). Vollenweider and others (Dillon and Rigler, 1975; Larsen and Mercier, 1976; Chapra, 1977; Walker, 1977; and Reckhow, 1979c) modified Vollenweider's initial approach and empirically derived simple input-output models for phosphorus. One of these cross-sectional empirical phosphorus lake models (Reckhow, 1979d) is incorporated in the procedure described in the next chapter.

The empirical phosphorus lake model mathematically describes the sections of Figure 1 from phosphorus loading to lake phosphorus concentration. The appropriate method for determination of the phosphorus loading depends in part upon whether the application is primarily descriptive or predictive. Descriptive use of the model implies an assessment of current lake and watershed conditions. Direct measurement of the phosphorus loading would therefore be possible. Alternatively, predictive use of the model suggests the estimation of the impact of projected land use on lake water quality.

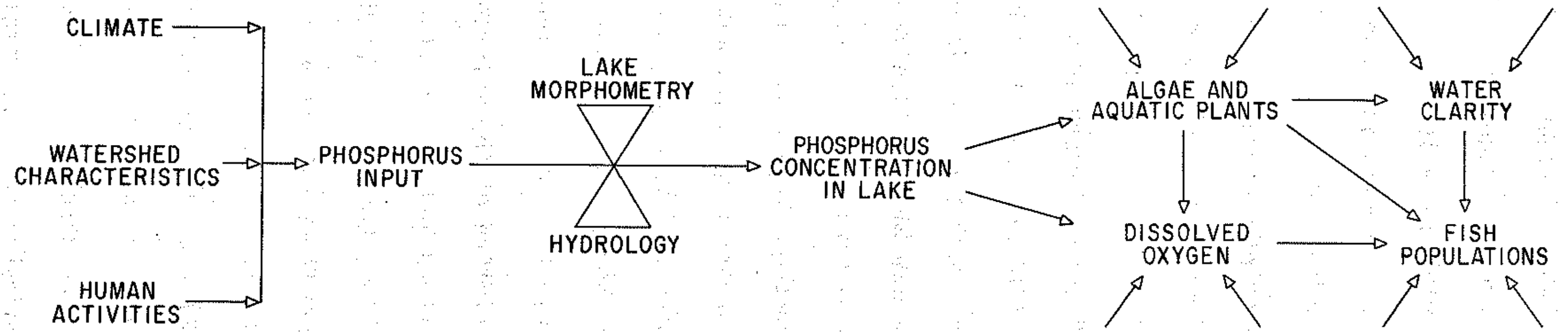


Figure 1. A Schematic Illustrating Phosphorus Loading Determinants and Lake Response

Here direct measurement is obviously impossible, so phosphorus loading information must be extrapolated from other, similar watersheds. It is for the predictive function of empirical phosphorus lake models that this manual is designed, although the methodology is equally applicable to descriptive applications (with perhaps an increase in planning risks, see Reckhow and Chapra, 1980, Chapter 1).

A significant portion of this manual is devoted to a discussion and presentation of the phosphorus export coefficients that are used to estimate phosphorus loading for the predictive mode of the modeling methodology. The phosphorus export coefficients presented are taken largely from a comprehensive study (Beaulac, 1980) of the literature on the phosphorus mass transported to surface water bodies from various land uses. An earlier survey of this topic (Uttormark et al., 1974) has been frequently cited in lake nutrient loading studies. However, the paucity of well-designed phosphorus export studies pre-dating Uttormark's thorough compilation resulted in the unavoidable inclusion of inaccurate export values in Uttormark's work. Cognizant of this problem, Beaulac screened the now considerable literature on phosphorus export to water bodies for accurate and reliable sampling designs. The result is a set of phosphorus export coefficients that are generally representative of the watershed conditions described. They are presented in Chapter 3 and in the Appendix with watershed descriptions (location, precipitation, soil type, fertilizer application, etc.) designed to facilitate selection of the appropriate export coefficients for lake phosphorus loading.

The concern expressed above for representative export coefficients is grounded in the emphasis in this manual on uncertainty. For planning to be effective, it must be based upon reliable information. Predictive models provide information for planners, so it is vital to the planning process that the reliability of this predictive information be estimated. Without a reliability measure, the planner has no basis for weighing model predictive information against other planning information. Inefficient and unpopular decisions can be the result.

Fortunately, it is possible to incorporate an uncertainty analysis into the modeling methodology. This is presented in a step-by-step manner, along with the phosphorus lake model, in Chapter 2. The end product is an estimate of total prediction uncertainty, which should be extremely useful to a planner as a measure of the value of the information provided by the model.

It must be underscored that the estimation of uncertainty does not obviate the need for consideration of the limitations of the modeling/error analysis methodology and for care in the selection of the phosphorus export coefficients. Geomorphologic and climatic constraints on the phosphorus lake model are mentioned in Chapter 2. These limitations are associated with the general rule that empirical models are developed for only a subpopulation of lakes represented by the model development data set. Application of the methodology to lakes not belonging to this subpopulation can increase uncertainty and prediction bias. Since there is no mechanism for the inclusion of this additional error term in the existing methodology, hidden planning risks may result.

The description of the modeling/uncertainty analysis methodology in Chapter 2 offers guidance on the selection of phosphorus export coefficients. Failure to adhere to the criteria for export coefficient choice or failure to carefully match the application lake watershed characteristics with the candidate export coefficient watershed characteristics (described in the tables in Chapter 3 and the Appendix) can again lead to hidden error and bias. This, too, may increase planning risks.

This manual is organized around the modeling/error analysis methodology in Chapter 2. The phosphorus export coefficients are presented in Chapter 3 and in the Appendix. This is accompanied by criteria used by Beaulac (1980) in the selection of published coefficients and criteria to be employed by users of this methodology when the export coefficients are selected and applied. At the end, some concluding thoughts are offered in Chapter 4 on the use of this manual for lake trophic management planning.

Chapter 2

THE PHOSPHORUS LAKE MODELING AND UNCERTAINTY ANALYSIS METHODOLOGY

2. 1 Introduction

The lake modeling/uncertainty analysis procedure presented below is a variation of the methodology developed by Reckhow and Simpson (1980). Because of the emphasis attached to uncertainty, the analyst is urged to follow the directions carefully. Further, it is suggested that the analyst read Chapter 1 (and possibly Chapter 4) before beginning the modeling, and read Chapter 4 before preparing a final report documenting the methodology and the application. These chapters offer valuable guidance on modeling, error analysis, and the interpretation of results for lake quality management planning.

2. 2 The Phosphorus Lake Model

In order to plan for the management of phosphorus in a lake watershed, mathematical models describing phosphorus loading and lake trophic response can be quite useful. Given the state of knowledge regarding phosphorus cycles, and the limited funds available to most planning agencies, often the most practical mathematical model for phosphorus management is the simple input/output or "black box" empirical model (Reckhow and Chapra, 1980). This type of model contains terms for the input, the output, and the settling (to the lake bottom) of phosphorus, but it does not explicitly include any biological or chemical reactions. The term "black box" is employed because the model treats the lake like a magician's black box; one is aware only of objects that enter and exit the box, as the contents and internal processes remain a mystery.

The left side of Figure 1 (Chapter 1) is a schematic of a "black box" phosphorus model and represents the conceptual foundation for the mathematical model used in this procedure. The figure shows that phosphorus input (loading) to a lake is a result of climate, watershed characteristics, and human activities. This input is modified by environmental factors and yields an output: the lake's average phosphorus concentration.

The mathematical model proposed herein was developed by Reckhow (1979d) from 47 north temperate lakes included in the Environmental Protection Agency's National Eutrophication Survey. This model expresses phosphorus concentration (P, in mg/l) as a function of phosphorus loading (L, in g/m^2-yr), areal water loading (q_s, in m/yr), and apparent phosphorus settling velocity (v_s, in m/yr) in the form:

$$P = \frac{L}{v_s + q_s} \qquad (1)$$

Using least squares regression, it was found that the apparent settling velocity could be fit using a weak function of q_s. This resulted in the fitted model:

$$P = \frac{L}{11.6 + 1.2q_s} \qquad (2)$$

A few limitations on the use of this model should be mentioned now. Since the model was constructed only from lakes within the north temperate climatic zone, it should be applied only to lakes within this zone. Furthermore, the model should not be applied to a lake with variable values more than the maximum values, or less than the minimum values, specified in Table 2. This is because an empirical model should not be used on lakes different from those used to develop the model, without prior testing. Finally, the model may be used to predict the average phosphorus concentration throughout a lake during the growing season. It cannot be used, as developed, to predict nearshore or short-term concentrations.

An important yet often overlooked aspect of the application of models is the fact that the model itself is a simplification of the real world, and thus the prediction from a model is inherently uncertain. Therefore, quantification of the prediction uncertainty should be a required step when a mathematical model is applied. This estimate of the prediction uncertainty could then be used by a modeler or planner as a weight indicating the value of the information contained in the prediction.

Uncertainty, or error, in this modeling exercise may arise from three primary sources: the model, the model parameters, and the model variables. Errors in the model and model parameters are derived from the procedure (regression analysis) used to empirically fit the model. For the phosphorus lake model (Equation 2), the model error ($s_{m\log}$) for the log transformed model is .128; parameter error was found to be quite small for most applications so it was ignored. Model variables uncertainty is estimated for the application lake, and Steps 2 and 4 in the procedure presented herein illustrate the necessary calculations. The separate error terms are combined for an estimate of total prediction uncertainty in Step 4.

Once phosphorus concentration is predicted through the application of the empirical model, it is useful to interpret this prediction in the context of expected water quality characteristics in the lake of interest. One example of a trophic state ranking scheme or index was proposed by Chapra and Reckhow (1979) based on average phosphorus concentration. In an earlier work, Dillon and Rigler (1975) also devised a trophic classification scheme which related general water quality and lake use features to the traditional trophic states. These two indices are combined in Table 1 to link phosphorus concentration to potential lake use.

Table 1: Proposed relationships among phosphorus concentration, trophic state, and lake use for north temperate lakes. (Adapted from Chapra and Reckhow, 1979; Dillon and Rigler, 1975).

Phosphorus Concentration (mg/l)	Trophic State	Lake Use
< 0.010	Oligotrophic	Suitable for water based recreation and propagation of cold water fisheries, such as trout. Very high clarity and aesthetically pleasing.
0.010 - 0.020	Mesotrophic	Suitable for water-based recreation but often not for cold water fisheries. Clarity less than oligotrophic lake.
0.020 - 0.050	Eutrophic	Reduction in aesthetic properties diminishes enjoyment from body contact recreation. Generally very productive for warm water fisheries.
> 0.050	Hypereutrophic	A typical "old aged" lake in advanced succession. Some fisheries, but high levels of sedimentation and algae or macrophyte growth may be diminishing open water surface area.

2. 3 The Modeling/Uncertainty Analysis Procedure

The method proposed herein has as its basis a procedure developed by Dillon and Rigler (1975) that may be used to calculate the capacity of a lake for development based upon the relationship between phosphorus input and water quality. The procedure presented below has two major improvements over that of Dillon and Rigler. First, the most important improvement is the addition of an error estimation procedure. This permits the quantification of prediction uncertainty, and it indicates to the user how valuable (certain) the information is that is provided by the model. Second, the phosphorus lake model imbedded in this procedure has a wider range of applicability than does the Dillon-Rigler model. Dillon and Rigler derived their model from a highly homogeneous set of lakes; the model development data set for the model presented herein includes a fairly wide range of lake types (see Table 2).

In order to facilitate understanding of the impact assessment procedure, it is presented in conjunction with an application to Higgins Lake in Michigan. It should not be inferred from this that the procedure, as presented, is applicable only to Higgins Lake. On the contrary, the procedure is quite general and can be easily applied to most lakes (subject to north temperate location and data set constraints such as those in Table 2) simply by substituting the appropriate lake data for the Higgins Lake data in the example.

Higgins Lake, located in the northern, lower peninsula of Michigan, is a deep, cool, oligotrophic lake with a well-oxygenated hypolimnion. The lake has a maximum depth of 41 meters and a mean depth of 15 meters. Some agricultural activity occurs in the watershed but most of the area is forested.

For the sake of this example, assume that an estimate of average lake phosphorus concentration in Higgins Lake, along with an assessment of water quality and recreation potential, is needed for a planning study. A measure of prediction uncertainty is also needed to evaluate the results and to compare them with alternative studies.

The method presented below will be used to solve this problem. Recall that the model is:

$$P = \frac{L}{11.6 + 1.2q_s}$$

This analysis is structured so that the variables are estimated in the following order: Step 1) areal water loading (q_s); Step 2) areal phosphorus loading (L); Step 3) lake phosphorus concentration (P); and Step 4) phosphorus prediction uncertainty (s_T).

Step 1: Estimation of q_s (areal water loading)

The estimation of q_s involves the solution of two equations:

$$Q = (A_d \times r) + (A_o \times Pr) \tag{3}$$

$$q_s = \frac{Q}{A_o} \tag{4}$$

where:

q_s = Areal water loading (m/yr)
Q = Inflow water volume to lake (m^3/yr)
A_d = Watershed area (land surface) (m^2)
A_o = Lake surface area (m^2)
r = Total annual unit runoff (m/yr)
Pr = Mean annual net precipitation (m/yr)

Ideally, Q should be determined from direct measurement of inflow or outflow, since use of any equation like Equation 3 will result in uncertainty in the predicted variable. When data for Q are not available, it becomes necessary to estimate A_d, A_o, r, and Pr and substitute them into Equation 3 to find Q.

Step 1A: Estimation of A_d (area of the watershed)

The highest points of lake and the lake outlet bound the watershed. In many situtations, all the precipitation that falls on the watershed, and is not evapotranspired, runs off or becomes groundwater and eventually reaches the lake. A topographical map enables one to locate the highest points of land surrounding a lake. Topographical maps are printed by the United States Geological Survey and must be ordered by quadrangle number or name at a U.S.G.S. office. The highest points of land may be outlined and A_d calculated by planimetry. Equation 3 requires that A_d be expressed as m^2 which may require adjustment of the units.

Step 1B: Estimation of r (annual unit runoff)

Average annual areal runoff has been mapped for many regions and again, the U.S.G.S. is a valuable source of information. Note that r must be expressed in m/yr (m^3/m^2/yr) , and note that it does not include ground water.

When A_d and r are multiplied together an estimate of the average inflow of water from surface runoff is obtained.

Step 1C: Estimation of A_o (area of lake)

The estimation of lake area requires the use of a good map or areal photograph of a known scale. The most accurate method for calculating this area is by planimetry. Note that A_o must also be expressed in m^2.

Step 1D: Estimation of Pr (precipitation)

An estimation of the average annual net precipitation (taking into account losses by evaporation) is also needed for Equation 3. This information can be obtained from the U.S.G.S. or the U.S. Weather Service. Note that Pr must be expressed in m/yr.

The statistics required to solve Equation 3 and 4 for the Higgins Lake example are presented in Table 3. Therefore, the necessary variables are:

i) Total annual inflow volume of water to Higgins Lake:

$$Q = (A_d \times r) + (A_o \times Pr)$$
$$= 30.863\ 10^6 m^3/yr$$

ii) The areal water loading:

$$q_s = \frac{Q}{A_o}$$
$$= 0.804\ m/yr$$

Step 2: Estimation of L (areal phosphorus loading)

Every watershed has a unique pattern of land use within its boundaries and each use makes a unique contribution, by way of diffuse sources, to the phosphorus loading of a lake. Technical, financial and practical constraints prohibit most water quality endeavors from conducting "in situ" studies. Therefore, many quantitative investigations rely on the application of phosphorus export coefficients derived from other studies. A compiled survey of coefficients screened according to acceptable criteria (see Chapter 3 and the Appendix) is located in Tables 6 through 12.

Table 2: Minimum and maximum value for the data set used to develop the phosphorus model (from Reckhow, 1979d)

Variable	Minimum	Maximum
P	.004 mg/l	.135 mg/l
L	.07 g/m^2-yr	31.4 g/m^2-yr
q_s	0.75 m/yr	187. m/yr

Table 3: Higgins Lake data necessary for the estimation of q_s

Variable	Estimate
A_d = Watershed area	87.41 $10^6 m^2$
r = Total annual unit runoff	0.2415 m/yr
A_o = Lake surface area	38.4 $10^6 m^2$
Pr = Mean annual net precipitation	.254 m/yr

Table 4: Land use areas in the Higgins Lake Watershed (Liebeskind et al., 1978)

Land Use	Area (hectares)	Area ($10^6 m^2$)
Agriculture	16	0.16
Forest	8347	83.47
Urban	378	3.78

In practical applications it is recommended that high, most likely, and low export coefficients be selected for some of the phosphorus source categories. This allows the calculation of high, most likely, and low total loading estimates, which ultimately represent the uncertainty that the analyst has in his/her estimates of phosphorus loading. The high and low loading estimates represent the additional phosphorus loading error that must be added to the model error for the calculation of total prediction uncertainty. It is important that the high and low loadings represent only those characteristics described in the steps below. This is because to a great extent, the error in the phosphorus loading estimates is already contained in the model error. Additional loading error for an application lake must be included only when the loading is estimated (using the procedure herein) in a different (and less precise) manner than it was estimated for the model development data set. These differences are described in the following steps.

The selection of appropriate phosphorus export coefficients is a difficult task. Since a critical aspect of this modeling exercise is the estimation of prediction errors, the analyst should realize that poor choice of export values contributes to an increase in error. This contribution may be explicit or implicit in the analysis, depending upon whether or not the analyst is aware of all of the uncertainty introduced by his/her choice of phosphorus export coefficients. Clearly, experience in the application of this modeling approach is a valuable attribute.

The estimates of phosphorus loading error are based on high, most likely, and low phosphorus export coefficients selected by the analyst. Loading uncertainty may be caused by either variability or bias. Variability may result from natural fluctuations inherent in a characteristic (e.g., natural variations in streamflow or stream phosphorus concentration), or from uncertainty inherent in a statistic summarizing a set of data. Bias may result from a number of causes, all associated with the fact that the estimate may not be representative of the characteristic that it was selected to estimate. For example, some uncertainty due to possible bias is appropriate to situations where phosphorus export coefficients generated in one watershed are applied in another watershed. As the analyst becomes more uneasy about a selected export coefficient, he/she should express this uneasiness through increased uncertainty and a greater range between high and low export coefficients.

At different points in the procedure presented below, the analyst is alerted to possible sources of bias. These result from the difference between conditions in the model development data set lakes and application lakes. When the characteristics of these two lake groups differ substantially, it is possible that a procedure appropriate for analysis of one group is inappropriate for analysis of the other. This is the justification for Table 2, which presents the limitations on the basic model variables, as defined by the model development data set. Now, when the allocation of phosphorus loading sources differs between the two lake groups, there is probably no need to restrict the use of the model, which is insensitive to the source of phosphorus. However, the error analysis (but not the mean prediction) may be affected because the loading estimation errors vary from source to source. Therefore, warnings of possible bias are stated herein when the application lake phosphorus loading allocation differs substantially from that for the model development data set. These warnings should be addressed, when appropriate, by the inclusion of a bias uncertainty addition to the high and/or low loading estimates.

The total annual mass flow of phosphorus to a lake is estimated by summing the annual phosphorus contribution from each of the nonpoint sources plus any additional point source input within the watershed. Total mass loading (M) may be expressed as (in kg/yr):

$$M = (Ec_f \times Area_f) + (Ec_{ag} \times Area_{ag}) + (Ec_u \times Area_u) + (Ec_a \times A_o) + (Ec_{st} \times \text{\# of capita-years} \times (1 - S.R.)) + PSI \qquad (5)$$

where:

Ec_f = Export coefficient for forest land (kg/ha/yr)

Ec_{ag} = Export coefficient for agricultural land (kg/ha/yr)

Ec_u = Export coefficient for urban area (kg/ha/yr)

Ec_a = Export coefficient for atmospheric input (kg/ha/yr)

Ec_{st} = Export coefficient to septic tank systems impacting the lake (kg/(capita - yr) - yr)

$Area_f$ = Area of forest land (ha)

$Area_{ag}$ = Area of agricultural land (ha)

$Area_u$ = Area of urban land (ha)

A_o = Area of lake (ha)

\# of capita-years = # of capita-years in watershed serviced by septic tank/tile field systems impacting the lake

S.R. = Soil retention coefficient (dimensionless)

PSI = Point source input (kg/year)

In order to facilitate the understanding and estimation of the variables contained in Equation 5, Step 2 has been broken into 7 sub-steps.

Step 2A: Estimation of $Area_f$, $Area_{ag}$, and $Area_u$ (watershed areas).

Recall that the watershed area was determined in Step 1. This area must now be subdivided into agricultural, forest, and urban lands. The area for each must be determined and expressed in hectares. Table 4 identifies the existing land uses and areas in the Higgins Lake watershed. When the objective of this analysis is the projection of future conditions, high and low area estimates are needed to reflect the uncertainty in the land use projections.

Step 2B: Estimation of Ec_f, Ec_{ag}, Ec_u, Ec_a and Ec_{st} (export coefficients)

This substep requires that an export coefficient (Ec_x) be chosen for each of the phosphorus source categories found in the watershed. Candidate export coefficients with a respective summary of watershed characteristics (e.g., soil type, % impervious surface, etc.) are found in Tables 6 through 12. Values for atmospheric loading are located in Table 13. These coefficients represent the expected annual phosphorus input to a lake or stream per unit of source.

It is important to understand that Ec_{st} differs from the other export coefficients in that it represents the expected annual amount of phosphorus transported not to the lake, but from households to on-site septic systems. A range of export coefficients that describes per capita export of phosphorus from households to septic systems is presented in Table 14.

After the analyst estimates the amount of phosphorus received by septic systems the next logical step is to determine how much of that phosphorus is being retained in the tile field soils (i.e., how much is exported to the lake). Phosphorus retention is addressed in substep 2C.

The high and low export coefficients selected should reflect the modeler's confidence in the extrapolation of literature export values to the application lake watershed. For example, in cases where the modeler knows that the "most likely" export coefficient chosen was determined under a good sampling program on a watershed quite similar to the application lake watershed, the high and low values should be selected to represent little uncertainty. A single precipitation loading figure (the "most likely" value) is probably adequate for most conditions. Only when the precipitation loading is deemed substantial (perhaps 25% of the total loading) should it be necessary to include possible precipitation loading error bias.

When the objective is future projection, the area estimates and export coefficients should be combined according to: high area with high export, most likely area with most likely export, and low area with low export. This calculation is to be made in Step 2F.

For the Higgins Lake example, various documents were consulted, including a study conducted by the U.S. Environmental Protection Agency (USEPA, 1975), in order to familiarize the investigators with the watershed. Phosphorus flux reports from other area lakes were also surveyed. Despite the existence of pertinent literature, the selection of application phosphorus export coefficients is still an unavoidably subjective task. This, of course, is the nature of the technique and consequently affirms the importance of the associated uncertainty analysis.

The forestland within the Higgins Lake watershed consists primarily of coniferous species with some deciduous trees and constitutes the major land use. Agriculture is rather limited and consists chiefly of grazing and pasture. The urban areas are mainly residential/recreational and all units are serviced by septic systems.

For demonstration purposes, high, most likely, and low export coefficients were chosen (based on the phosphorus export coefficients presented in Chapter 3) to reflect phosphorus source conditions found in the Higgins Lake watershed. The selected coefficients are presented in Table 5 . Note that the high and low values selected for Higgins Lake are not as high or

Table 5: High, most likely, and low export coefficients selected for Higgins lake

Source	High	Most Likely	Low
Forest	.30 kg/ha/yr	.20 kg/ha/yr	.10 kg/ha/yr
Agriculture (pasture and grazing land)	1.30 kg/ha/yr	.40 kg/ha/yr	.20 kg/ha/yr
Urban (residential)	2.70 kg/ha/yr	.90 kg/ha/yr	.35 kg/ha/yr
Precipitation	.50 kg/ha/yr	.30 kg/ha/yr	.15 kg/ha/yr
Input to septic tanks	1.0 kg/capita/yr	0.6 kg/capita/yr	0.3 kg/capita/yr

as low as some of the candidate export coefficients presented in Chapter 3. This is because conditions in the Higgins Lake watershed were judged to be not equivalent to the extreme conditions that are represented by the ranges in candidate coefficients.

The Ec_{st} export coefficients were selected to take into account Michigan's ban on the sale of phosphorus-based detergents. Thus, the selected coefficients are on the lower side of the range exhibited in Table 14. Likewise, note that Ec_a is also on the lower side of the presented atmospheric export coefficient range. This is because little agricultural and industrial activity take place in the Higgins Lake area, which probably results in small quantities of air-born phosphorus.

Step 2C: Estimation of S.R. (soil retention coefficient)

On-site septic tank-tile field systems may or may not be effective in trapping phosphorus and preventing it from entering a lake via groundwater transport. The soil retention coefficient is an estimate of how well the systems immobilize phosphorus. This coefficient may range from 0 to 1.0. For example, if it is assumed that all phosphorus transported to septic systems eventually reaches the lake, then a soil retention coefficient value of 0 would be selected. If it is assumed that no phosphorus reaches the lake, then S.R. = 1.0.

Rodiek (1979) notes that effective tile drainage fields involve both physical and chemical processes. Chemical fixation reactions require effluent-to-soil contact of sufficient time length for chemical reactions or adsorption to occur. There are four major aspects of watershed soils (within the lake impacting zone) that influence contact duration time and phosphorus immobilizing capabilities and thus should be considered when selecting S.R. These factors are: 1) phosphorus adsorption capacity;

2) natural drainage; 3) permeability and; 4) slope. As one might expect, all of these factors are closely related and of a dynamic nature. They are discussed in depth in Chapter 3.

In addition to the above soil characteristics, there are four general mechanisms of phosphate removal in soils: 1) rapid removal or adsorption; 2) slow mineralization and insolubilization; 3) plant uptake; and 4) biological immobilization (Tofflemire and Chen, 1977). The most important of the phosphorus immobilization mechanisms in septic systems are the formation of insoluble iron and aluminum phosphate compounds and the adsordtion of phosphate ions onto clay lattice structures (Tilstra et al., 1972).

Assessment of the factors discussed above is useful in determining S.R. However, because of the complexities involved, the modeler's estimation of S.R. still must be based on his/her knowledge of the soil conditions present in the application watershed, past experience with similar watersheds and his/her professional intuition. When the model is used to predict future conditions, it is often sufficient to use a single estimated soil retention coefficient. Only when the estimated loading from septic systems is thought to be substantial (perhaps 25% of the total loading), should it be necessary to employ low, most likely, and high soil retention coefficients. It is possible, however, that the error analysis may be biased when the septic tank loading becomes a sizeable fraction of the total loading.

For the Higgins Lake example, it was found that sandy/gravel soils of moraines and till plains predominate in the watershed, which tend to permit rapid infiltration and transmission of water. Nearby Houghton Lake is surrounded by various soils posessing moderate to poor phosphorus adsorbing capacities (Ellis and Childs, 1973). Based on this evidence, soil retention of phosphorus was estimated to be on the poor side. A "most likely" S.R. coefficient of .25 a "low" coefficient of .50 and a "high" coefficient of

.05 were selected to represent the soils surrounding Higgins Lake. Since evidence (USEPA, 1975) suggests that phosphorus loading from septic systems may be a substantial fraction of the total loading, three (not one) soil retention coefficients were chosen for Higgins Lake, as specified in the instructions above.

Step 2D: Estimation of # of capita-years

The number of persons contributing to septic systems that impact a lake must be estimated and expressed in capita-years. To ascertain this figure, the analyst must first determine the size of the impact zone. Often this is a strip, perhaps 20-200 meters wide, surrounding the lake. Sometimes the analyst may include border strips along tributary streams when conditions suggest that these remote areas may be important. Conditions that dictate the size and location of the impact zone include drainage patterns, water tables, and slopes.

When the model is used to assess current conditions, population surveys are quite useful for the estimation of the phosphorus loading from septic tanks. When the goal is the prediction of future conditions, population projections must be consulted. For most lakes, the high and low loading estimates for septic systems should then be based solely on the uncertainty in the population projections (the source of possible bias). The total number of capita-years may be calculated by adding together permanent resident capita-years and seasonal resident capita-years. This is described in Equation 6 below.

Permanent capita-year

$$\text{Total \# of capita-years} = \text{average \# of persons per living unit} \times \frac{\text{\# days spent at unit per year}}{365} \times \text{\# of living units}$$

$$+ \tag{6}$$

Seasonal capita-year

$$\text{average \# of persons per living unit} \times \frac{\text{\# days spent at unit per year}}{365} \times \text{\# of living units}$$

In this particular example, it was assumed that septic systems of only lakeside dwellings impact Higgins Lake. According to the EPA National Eutrophication Survey (USEPA, 1975) there are an estimated 1,000 seasonal dwellings on the Higgins Lake shoreline and all are served by septic systems. A facilities plan study estimated that each seasonal unit is occupied by an average of 3.5 people who spend 60 days a year at their residence (Progressive Engineering Consultants, 1976). This information may be inserted in Equation 6 to estimate the number of capita-years impacting the lake.

$$\text{Total \# of capita-years} = \underbrace{0}_{\text{Permanent}} + \underbrace{3.5 \times \frac{60}{365} \times 1000}_{\text{Seasonal}}$$

$$= 575.3$$

Step 2E: Estimation of PSI (point source input)

If the effluent from an industry, sewage treatment facility or other point source is deposited within the watershed, the impact must be assessed and expressed in kg phosphorus/yr.

At the present time there are no known point sources of phosphorus in the Higgins Lake watershed. Thus, PSI = 0 kg/yr. However, if point sources exist in the application lake watershed, or are projected for the future, then the uncertainty in the phosphorus loading from this source is represented by the uncertainty in the size of the projected population to be served. An example of phosphorus loads from sewage treatment plants is presented in Table 15.

As a final note on the phosphorus loading sources, it is possible that the lake sediments may be a non-negligible source. Internal phosphorus loading is most probable in shallow lakes that possess anoxic bottom waters. This condition can promote an appreciable rate of phosphorus transport from the sediment/water interface to overlying waters. In shallow lakes this sediment phosphorus may reach the photic zone and be used by the aquatic plants. If this is thought to be so for an application lake, then high, most likely, and low loading estimates should be used for the prediction and prediction error (see Reckhow, 1979b for suggested values), to reflect this source of possible bias in the uncertainty analysis.

Step 2F: Calculation of M (total phosphorus mass loading)

When Steps 2A through 2E are complete, Equation 5 may be solved to yield high, most likely, and low phosphorus mass loading estimates based on high, most likely, and low phosphorus export and soil retention coefficients.

Thus, for the Higgins Lake example:

high

$M_{(high)} = (.30 \times 8347) + (1.30 \times 16) + (2.7 \times 378) + (.50 \times 3840) +$

$(1.0 \times 575.3 \times (1 - 0.05)) + 0$

$= 6012.04$ kg/yr

most likely

$M_{(ml)} = (.20 \times 8347) + (.40 \times 16) + (.90 \times 378) + (.30 \times 3840) +$

$(0.6 \times 575.3 \times (1 - 0.25)) + 0$

$= 3426.9$ kg/yr

low

$M_{(low)} = (.1 \times 8347) + (.20 \times 16) + (.35 \times 378) + (.15 \times 3840) +$

$(0.3 \times 575.3 \times (1 - .50)) + 0$

$= 1632.5$ kg/yr

Step 2G: Calculation of L (annual areal phosphorus loading)

In order to be used in this model, annual phosphorus input must be expressed as a loading per unit lake surface area. This is accomplished by dividing M by the lake surface area, A_o:

$$L = \frac{M}{A_o} \tag{7}$$

The units are then converted so that this areal phosphorus loading term is expressed in grams per square meter of lake surface area per year. Thus, for Higgins Lake:

high

$$L_{(high)} = \frac{6012.0\ kg/yr}{38.4 \times 10^6 m^2} = 157 \times 10^{-6}\ kg/m^2/yr = .157\ g/m^2/yr$$

most likely

$$L_{(ml)} = \frac{3426.9\ kg/yr}{38.4 \times 10^6 m^2} = 89 \times 10^{-6}\ kg/m^2/yr = .089\ g/m^2/yr$$

low

$$L_{(low)} = \frac{1632.5\ kg/yr}{38.4 \times 10^6 m^2} = 43 \times 10^{-6}\ kg/m^2/yr = .043\ g/m^2/yr$$

Step 3: Calculation of P (lake phosphorus concentration)

The model may now be solved for high, most likely, and low phosphorus concentrations by substituting in values of q_s and L (high, ml, and low).

$$P_{(high)} = \frac{L\ (high)}{11.6 + 1.2q_s}$$

$$P_{(ml)} = \frac{L\ (ml)}{11.6 + 1.2q_s}$$

$$P_{(low)} = \frac{L\ (low)}{11.6 + 1.2q_s}$$

For Higgins Lake:

high

$$P_{(high)} = \frac{0.157}{11.6 + 1.2\ (0.804)} = 0.0125\ mg/l$$

most likely

$$P_{(ml)} = \frac{0.089}{11.6 + 1.2\ (0.804)} = 0.0071\ mg/l$$

low

$$P_{(low)} = \frac{0.043}{11.6 + 1.2\ (0.804)} = 0.0034\ mg/l$$

Step 4: Estimation of s_T (prediction uncertainty)

In order to estimate the uncertainty associated with a prediction calculated using the phosphorus model, estimates are needed for the error, or uncertainty, in all terms in the model, and in the model itself. However, it has been shown by Reckhow (1979d) that for most applications of this model, the error in the parameter v_s is small. Further, error in q_s is primarily a function of flow measurement error and hydrologic variability, which also affect L. Since L and q_s are in the numerator and denominator, respectively, in the model, the errors affecting both tend to cancel when they are combined to yield the resultant error in P. In addition, hydrologic variability is unimportant in lakes with low flushing rates. Therefore, it is assumed here that the prediction error is a function only of model error and of aspects of phosphorus loading uncertainty that are identified in Step 2. If the application lake flushes rapidly and is subject to great variations in year-to-year precipitation, then the modeler is urged to include hydrologic variation in the error analysis using the error propagation equation (see the Appendix for instructions).

The model error is represented by $s_{m\log}$ in the equations below and is expressed in logarithmic units of phosphorus concentration error. The loading error, s_L, on the other hand, is expressed in untransformed units of phosphorus loading error. Therefore, to combine these two values for an estimate of total prediction uncertainty, some calculations are necessary.

The procedure presented below is based on first order error analysis (Benjamin and Cornell, 1970). In this particular application, three assumptions are of some importance:

1. Model error, expressed in log-transformed concentration units, is appropriately combined with variable error terms after the transformation is removed.

2. The "range" ("high" minus "low"), for phosphorus loading error, is approximately two times the standard deviation. This is based loosely on the characteristics of the Chebyshev inequality identified below, where about 90% of the distribution is contained within ± 2 standard deviations of the mean.
3. The individual error components are adequately described by their variances (standard deviations).

In order to relax a previously imposed (Reckhow, 1979a) yet tenuous normality assumption, the confidence intervals constructed below are based on a modification of the Chebyshev inequality (Benjamin and Cornell, 1970). Therefore, it is no longer required that the total error term be normally distributed. Instead its distribution must only be unimodel and have "high order contact" with the abscissa in the distribution tails. These are achievable assumptions under almost all conditions, and it is recommended (Reckhow and Chapra, 1979) that this type of nonparametric approach be adopted until the distributions have been adequately studied and characterized.

Step 4A: Calculation of log $P_{(ml)}$

Take the logarithm of the most likely phosphorus concentration, $P_{(ml)}$.

For Higgins Lake:

$$\begin{aligned}\log P_{(ml)} &= \log 0.0071\\ &= -2.149\end{aligned}$$

Step 4B: Estimation of s_m+ ("positive" model error)

The model error, (s_{mlog}), was determined to be 0.128. Add s_{mlog} to $\log P_{(ml)}$ and take the antilog of this value. Now calculate the difference between this antilog value and $P_{(ml)}$. Label this difference s_m+, it represents the "positive" model error.

$$s_m+ = \text{antilog} [\log P_{(ml)} + s_{mlog}] - P_{(ml)} \tag{8}$$

For Higgins Lake:

$$\begin{aligned} s_m+ &= \text{antilog} (-2.149 + 0.128) - 0.0071 \\ &= 0.0024 \text{ mg/l} \end{aligned}$$

Step 4C: Estimation of s_m- ("negative" model error)

Subtract s_{mlog} from $\log P_{(ml)}$ and take the antilog of this value. Now calculate the difference between this antilog and $P_{(ml)}$, and label this difference s_m-.

$$s_m- = \text{antilog} [\log P_{(ml)} - s_{mlog}] - P_{(ml)} \tag{9}$$

For Higgins Lake:

$$\begin{aligned} s_m- &= \text{antilog} (-2.149 - 0.128) - 0.0068 \\ &= 0.0015 \text{ mg/l} \end{aligned}$$

Step 4D: Estimation of s_L+ ("positive" loading error)

Now, one must convert the loading error estimate into units compatible with the model error. Use the $P_{(high)}$ concentration estimated in Step 3 and calculate the difference between $P_{(high)}$ and $P_{(ml)}$; then divide this difference by 2. Label this value s_L+; it represents the "positive" loading error contribution.

$$s_L+ = \frac{P_{(high)} - P_{(ml)}}{2} \qquad (10)$$

For Higgins Lake:

$$s_L+ = 0.0027 \text{ mg/l}$$

Step 4E: Estimation of s_L- ("negative" loading error)

Repeat Step 4D substituting the low concentration value $P_{(low)}$ for $P_{(high)}$. Label the resultant value s_L-; it represents the "negative" loading error contribution.

$$s_L- = \frac{P_{(ml)} - P_{(low)}}{2} \qquad (11)$$

For Higgins Lake:

$$s_L- = 0.0019 \text{ mg/l}$$

Step 4F: Estimation of s_T+ (total "positive" uncertainty)

Total positive prediction uncertainty is calculated using the equation:

$$(s_T+)^2 = (s_m+)^2 + (s_L+)^2 \tag{12}$$

or:

$$s_T+ = \sqrt{(s_m+)^2 + (s_L+)^2} \tag{13}$$

For Higgins Lake:

$$= \sqrt{(0.0024)^2 + (0.0027)^2}$$

$$s_T+ = 0.0036 \text{ mg/l}$$

Step 4G: Estimation of s_T- (total "negative" uncertainty)

Total negative prediction uncertainty is calculated using the equation:

$$(s_T-)^2 = (s_m-)^2 + (s_L-)^2 \tag{14}$$

or:

$$s_T- = \sqrt{(s_m-)^2 + (s_L-)^2} \tag{15}$$

For Higgins Lake:

$$= \sqrt{(-0.0015)^2 + (0.0019)^2}$$

$$s_T- = 0.0024 \text{ mg/l}$$

Step 4H: Calculation of confidence limits

The prediction uncertainty may be expressed in terms of "confidence limits" which represent the prediction plus or minus the prediction uncertainty.

Confidence limits have a definite meaning in classical statistical inference; they define a region in which the true value will lie a pre-specified percentage of the time.

Using the modification of the Chebyshev inequality (Benjamin and Cornell, 1970), the confidence limits may be written as:

$$\text{Prob}\ [(P_{(ml)} - hs_T^{-}) \leq P \leq (P_{(ml)} + hs_T^{+})] \geq 1 - \frac{1}{2.25h^2} \qquad (16)$$

Equation 16 states that the probability that the true phosphorus concentration lies within certain bounds, defined by a multiple, h, of the prediction error, is greater than or equal to $1 - 1/2.25h^2$. (This relationship loses its significance as h drops much below one.) Substituting values for h into Equation 16 reveals that a value of one for h corresponds to a probability of about 55% (.556 to be exact), and a value of two for h corresponds to a probability of about 90% (.889 to be exact). Thus the 55% confidence limits are:

$$\text{Prob}\ [(P_{(ml)} - s_T^{-}) \leq P \leq (P_{(ml)} + s_T^{+})] \geq .55 \qquad (17)$$

Substituting the Higgins Lake data this becomes:

$$\text{Prob}\ [(0.0071 - 0.0024) \leq P \leq (0.0071 + 0.0036)] \geq .55$$

$$\text{Prob}\ [0.0047\ \text{mg/l} \leq P \leq 0.0107\ \text{mg/l}] \geq .55$$

Now that specific values for the prediction error have been inserted into the confidence limits expression, its interpretation changes somewhat. It is: "about 55% of the time (that confidence limits are estimated), one can expect that the actual average phosphorus concentration will lie within the bounds defined by the prediction plus or minus the prediction uncertainty." This same interpretation format applies when the confidence limits are widened to the 90% level (h = 2), and the Higgins Lake data are inserted:

$$\text{Prob}\ [(P_{(ml)} - 2s_{T}-) \leq P \leq (P_{(ml)} + 2s_{T}+)] \geq .90 \qquad (18)$$

Inserting the data yields:

$$\text{Prob}\ [(0.0071 - (2)\ (0.0024)) \leq P \leq (0.0071 + (2)\ (0.0036))] \leq .90$$

$$\text{Prob}\ [0.0023\ \text{mg/l} \leq P \leq 0.0143\ \text{mg/l}] \geq .90$$

2. 4 Application Summary

The application of the technique and model to Higgins Lake resulted in a "most likely" phosphorus concentration of 0.0071 mg/l (Step 3), with 55% confidence limits bounding the "true" phosphorus concentration between 0.0047 mg/l and 0.0107 mg/l.

Relating back to the trophic classification in Table 2, Higgins Lake is probably:

1) oligotrophic
2) clear, and suitable for water-based recreation and a cold water fishery.

These predicted trophic conditions in fact describe the present observed conditions in Higgins Lake. A median phosphorus concentration value of .006 mg/l was determined for Higgins Lake by the Environmental Protection Agency's National Eutrophication Survey (USEPA, 1975).