The Size Distribution of the Limnoplankton

Martin Ahrens

Department of Biology

McGill University

Montreal, Quebec

February 1989

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

The hypothesis that aquatic biomass is uniformly distributed over logarithmic size classes was evaluated with samples from epilimnetic plankton communities. Although the hypothesis is upheld for oligotrophic lakes, biomass increases between successive size classes in mesotrophic and eutrophic lakes. The abundance of organisms in logarithmic size intervals is strongly negatively correlated with size. The relationship between organism size and physiological performance in mixed communities was examined by testing the hypothesis that limnoplankton respiration rates are predictable functions of mean body size. The equation describing this relationship was found to be similar to those obtained with laboratory cultures. The total epilimnetic phosphorus concentration is correlated with both biovolume and respiration rate. Many limnological relationships, including those established in this thesis, are based on linear regressions between log-transformed variables. The rules for the correct use of backtransformed predictions are elucidated. A theorem is proven, which sets limits to the relationship between the coefficients of determination on the original and transformed scales. Simulated data and empirical results are used to illustrate the applications and limitations of the theoretical results.

RÉSUMÉ

L'hypothèse que la biomasse aquatique est distribuée uniformément parmi des groupes de taille logarithmiques a été évaluée à partir d'échantillons des communautés du plancton épilimnétique. L'hypothèse est fondée dans les lacs oligotrophes. Par ailleurs, dans les lacs eutrophes et mésotrophes, la biomasse augmente entre les classes successives. L'abondance des organismes dans ces intervalles de taille logarithmiques a une corrélation inverse très forte avec la taille. La relation entre la taille et le comportement physiologique du plancton en communautés mixtes a été étudiée en évaluant l'hypothèse que le taux de respiration est une fonction predictive de la taille des organismes. L'équation de cette relation s'est avérée semblable a celle obtenue avec les organismes cultives en laboratoire. La concentration épilimnétique totale en phosphore est en corrélation avec la biovolume et le taux respiratoire. Plusieurs relations en limnologie, y compris celles de cette thèse, sont fondées sur des regressions lineaires apres transformations logarithmiques. Les regles qui gouvernent l'utilisation exacte des prédictions sont elucidees. Un théorème est prouvé, qui établit les limites à la relation entre les coefficients de détermination avant et après transformation. Des données de simulation et des resultats empiriques sont utilisés pour illustrer les applications et les limites des résultats théoriques.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

CHAPTER I:

CHAPTER II:

CHAPTER III:

LIST OF TABLES

CHAPTER I:

CHAPTER II:

CHAPTER III:

PREFACE

Remarks on style and authorship

As required by the <u>Guidelines Concerning Thesis Preparation</u>, the following article is quoted:

> "The Candidate has the option, subject to the approval of the Department, of including as part of the thesis the text, or duplicated published text (see below), of an original paper, or papers. In this case the thesis must still conform to all other requirements explained in Guidelines Concerning Thesis Preparation. Additional material (procedural and design data as well as descriptions of equipment) must be provided in sufficient detail (e.g. in appendices) to allow a clear and precise judgment to be made of the importance and originality of the research reported. The thesis should be more than a mere collection of manuscripts published or to be published. <u>It must include a general abstract, a full introduction and literature review and a final overall conclusion</u>. Connecting texts which provide logical bridges between different manuscripts are usually desirable in the interests of cohesion.
>
> "It is acceptable for theses to include as chapters authentic copies of papers already

> published, provided these are duplicated clearly on regulation thesis stationery and bound as an integral part of the thesis. Photographs or other materials which do not duplicate well must be included in their original form. _In such instances, connecting texts are mandatory_ and supplementary explanatory material is almost always necessary.
>
> "The inclusion of manuscripts co-authored by the candidate and others is acceptable, but the candidate is required to make an explicit statement on who contributed to such work and to what extent, and supervisors must attest to the accuracy of the claims, e.g. before the Oral Committee. Since the task of the Examiners is made more difficult in these cases, it is in the candidate's interest to make the responsibilities of authors perfectly clear. Candidates following this option must inform the Department before it submits the thesis for review."

Each chapter of this thesis is based on the text of manuscripts which have been submitted to learned journals for publication. This format has produced some redundancy, for which I apologize to the reader.

Chapters I and II are based on Ahrens and Peters (1989a and 1989b). Dr. Peters' contributions were limited to the provision of financial resources and equipment for field and laboratory research, and editorial advice on the manuscripts. During the development of the research program, he of course provided the usual advice and guidance offered by a PhD supervisor. Chapter III is based on Ahrens (1989), which is entirely my own work. Diccon Bancroft provided some editorial advice on this chapter. Marika Kurer and Lesley Pope provided technical assistance with routine field and laboratory work. I developed the procedures and conventions for size distribution measurements on the inverted microscope, which were done under my supervision by Marika Kurer and Ashley Evans. All of the data analysis programs in Appendix 8 are entirely my own work.

Contributions to original knowledge

I believe that each of the chapters of this thesis contributes to original knowledge.

Chapter I shows that plankton biomass is uniformly distributed over logarithmic size classes in only the most oligotrophic lakes. As phosphorus concentrations increase, biomass per class tends to increase with size. Although total biomass tends to be correlated with total phosphorus, phosphorus cannot be used as an index of biomass without reference to the size of the organisms involved. This is because the mean phosphorus concentration per organism is size dependent.

Chapter II shows that an allometric relationship between mean organism size and respiration rate holds in natural plankton communities. The total epilimnetic phosphorus concentration is correlated with the total community respiration rate. Gravity screening is an acceptable procedure for separating size fractions of plankton for respiration rate determination, since the sum of rates measured on these fractions is not significantly different from the total community rate.

In Chapter III a formula for the correlation between the logarithms of multivariate normal random variables is developed. This correlation is usually less, and never greater, than the correlation between the original variables. Data from Chapter II and a simulation are used to show that the formula does not apply when data are not multivariate normal. The correction factor for backtransformation, which has been inaccurately presented in the literature, is clarified here.

Acknowledgements

Although the embryo of this research was once my supervisor's personal project, Rob Peters was wise enough to provide the freedom to allow me to do it my way. Nevertheless, on those few occasions when I sought his advice, it was offered

abundantly. While I developed and carried out my research program, and wrote this thesis, his insistence on clarity of exposition and demands that I "finish the thought" had considerable impact: My research program became more coherent and rigorous, this document became more readable, and my thought became more effective than they would have been without his influence.

The Limnology research group of the Biology Department usually provided a congenial environment for my PhD program. Yet I was particularly impressed by the role played by Sara Griesbach. She is faced with ever-changing demands from a succession of strong personalities, each convinced of the importance of his or her own project. Yet I always found her calm, impartial, unpretentious, and willing to deal with whatever urgent request I came up with.

As with all publications of the Limnology research group, this thesis would not have been possible without the assistance of seasonal technicians. In exchange for terrible wages and no prestige, these people are reliable, hard-working, and cooperative. This was certainly true of Ashley Evans, Marika Kurer, and Lesley Pope.

During the preparation of my thesis, I have been privileged to use the facilities of my employer, Consumers Union of United

States, Inc. For this, as well as for amicable encouragement and advice, thanks are due to my supervisor, Diccon Bancroft.

My entire research program was funded by grants to Dr. Peters and the McGill Limnology Research Centre from NSERC, FCAR, and the Donner Foundation. My personal stipends were provided by Dr. Peters' research funds and Teaching Assistantships in the Biology Department.

My warmest thanks are due to my family. My daughter, Heidi, had to forego a lot of attention and material advantages when her father decided to become a student again. Yet she has always been supportive and encouraging. But without Marika to share my life, this thesis would never have been written. You have both trusted me, and taken some big risks along with me. This thesis is dedicated to both of you!

GENERAL INTRODUCTION

The aquatic ecologist's mandate might be broadly described as the elaboration of models to predict the distribution of biomass, production, and kinds of aquatic organisms (Peters 1980). Like most ecological models, this thesis is premised on the further assumption that the variables measured are important, due either to their predictive power or to their utility. When one refers to "kinds" of organisms, one often thinks of "species". However, the research presented here explores models in which organism size, rather than taxonomy, is of fundamental importance. This introduction explains this choice.

Most previous research on plankton communities has focused on spatial and temporal patterns in the distribution of species (e.g. Hutchinson 1967). Ecological theorists have taken several approaches to the development of quantitative models, founded upon the impressive body of accumulated knowledge about plankton species. Community models based on the environmental physiology of individual taxa have not been successfully applied to natural communities, due to the great spatial and temporal diversity of the plankton. In an effort to overcome this problem, predictive models of community structure have been developed (e.g. Sommer et al. 1986). But such models are difficult to test, primarily qualitative, and difficult for impossible) to relate to quantitative variables, such as biomass, uptake rates, and productivity. At a still higher level of generalization, species abundance, species-area relations, and the size distribution of species have received attention (e.g. May 1975, 1981). But the species concept remains largely qualitative and categorical, and so predictive relationships with quantitative variables are scarce. In particular, the biomass, uptake rates, and productivity of an individual species vary considerably over space and time. Hence, the limited utility of the species concept has provided impetus for the search for a more quantitative understanding of "kinds of organisms".

This thesis contributes to a body of aquatic modelling which focuses on the distribution of organisms by size, deemphasizing their traditional taxonomic categorization. In this spirit, Sheldon et al. (1972, 1977) sought and found regularities in the size distribution of oceanic seston; Harris et al. (1983), Sprules et al. (1983, 1986), and Bailey-Watts (1986) extended this approach to the limnoplankton; Schwinghamer (1981, 1983) and Warwick (1984) studied the size distribution of marine benthos; and Strayer (1986) examined the size distribution of freshwater benthos. These empirical observations of size distributions have lead to theoretical models of plankton community physiology (e.g. Kerr 1974; Platt & Denman 1978; Silvert & Platt 1980: Borgmann 1982; Griesbach et al. 1982) founded upon apparent regularities in the size structure, as well as on the known size-dependence of production, respiration, and other physiological processes (e.g. Hemmingsen 1960; Banse 1976; Banse & Mosher 1980; Blueweiss et al. 1978).

The choice between species and size implies a choice in theoretical emphasis. A predictive model which emphasizes taxonomy implicitly attributes importance to the role played by phylogenetic determinism in structuring the community. The current community composition is explained as a consequence of the evolutionary ecology of the populations present. Abiotic variables (e.g. geological history, wind, and water course alterations) are invoked when necessary, but the primary emphasis is on the physiological and ecological requirements of the component taxa. Predictions of future system behaviour would be founded upon the known characteristics of these taxa. On the other hand, size-based models emphasize effects which act independently of taxonomy. For example, the impact of hydrography through entrainment and sinking is primarily a function of surface area and volume of particles, hence of their size. Surface area and average radius also help determine potential uptake and excretion rates, and hence the organism's reaction to a given nutrient environment. The total biomass and relative numbers of organisms at size are hypothesized to be constrained primarily by abiotic variables, and this is testable. The difference between the two types of models is a matter of choice in emphasis. Scientifically, one might justify a particular choice by demonstrating that predictions satisfy pre-established criteria of statistical accuracy.

Possibly a hybrid model would maximize predictive strength. In a given lake, taxonomic variations within size classes are due to recent phylogenetic history. Hence, such a model would incorporate some rules allowing prediction of the size-class specific species structure for the system of interest. Then there would be equations to predict size distribution parameters from abiotic variables, and further equations to predict physiological variables from the size and species composition. This type of model would recognize explicitly that not only do physiological variables depend upon size, but that predictive precision increases with taxonomic homogeneity (Banse & Mosher 1980). Nevertheless, the hybrid model would suffer from the major disadvantage of the species-based model: the enormous amount of information required to predict the species structure for a given system. Hence, we must hope that size structure alone will furnish a sufficient basis for accurate prediction of physiological variables of interest.

Chapter I first establishes that there are predictable regularities in the size structure of the limnoplankton. Its antecedents include the work of Sheldon et al. (1972), who found that pelagic oceanic waters had either a roughly constant total particle volume in successive logarithmic size classes (subtropical waters) or an approximately unimodal distribution (elsewhere). When Sprules et al. (1983) examined limnoplankton size spectra at a finer scale, they found evidence of bimodal distributions. In an effort to quantify the differences in distributions among lakes, I have attempted to estimate the parameters of uni-, bi-, and trimodal models, as well as "normalized spectra" (Appendix 1). I have also dealt with the prediction of these system parameters from other easy-to-measure variables.

It is well known that many metabolic and physiological rates of both individuals and communities are size dependent (e.g. reviews in Calder 1984 and Peters 1983a). Examples include respiration, production, longevity, growth, assimilation, and fecundity. It is thus reasonable to hypothesize that energetic and material flows within the plankton community may be predicted from the overall size composition. However, allometric relations for aquatic organisms have been founded upon laboratory studies of isolated taxa. So, before community size spectra can be used for physiological predictions, it must be shown that allometric relations also hold for natural communities composed of many taxa. Chapter II addresses this hypothesis, where respiration rate is used as the physiological "response" variable. Since respiration, or metabolic rate, and other physiological rates have a known common size dependence, a demonstration of metabolic allometry in nature should provide support for the hypothesis that other rates of interest may be predicted from the size structure. Alternatively, measurement of the metabolic rate of size classes could be used to predict values of other variables.

Several statistically significant relationships between logarithmically transformed variables are established in the first two chapters. In Chapter III, I deal with the statistical problems of backtransforming the predictions from these models. I present an original result concerning the coefficient of determination, and correct the existing ecological literature with respect to the "correction factor" for backtransformation. The impact of this work on ecological models is examined through a simulation exercise and through a discussion of two relationships established in a previous chapter.

CHAPTER I

LIMNOPLANKTON SIZE SPECTRA

ABSTRACT

The hypothesis that aquatic biomass is uniformly distributed over logarithmic size classes was evaluated with samples from epilimnetic plankton communities at 15 temperate lake sites in southern Quebec. Over the size range from 0.2 μm to 1600 μm equivalent spherical diameter (ESD), biomass tends to increase between log size classes at a median rate of 7%, in a data set with a median total phosphorus concentration (TP) of 17 mg m^{-3}. The slope of the normalized biomass spectrum (reflecting overall trends in the distribution) becomes significantly steeper with decreasing TP. Hence more oligotrophic systems have a more uniform biomass distribution. Over the observed size range, most samples were dominated by the phytoplankton mode between 20 and 50 μm ESD. Total plankton biomass was positively correlated with TP. The abundance of organisms in logarithmic size intervals was strongly negatively correlated with size. Many sample distributions did not differ significantly from unimodal lognormal distributions. Efforts to fit bimodal and trimodal distributions met with limited success, since only 39 size classes were used.

Introduction

As a result of their initial empirical observations with electronic particle counts in the ocean, Sheldon et al. (1972) hypothesized that aquatic biomass is roughly uniformly distributed over logarithmic size classes. Later evidence from both marine (Sheldon et al. 1972) and freshwater (Sprules et al. 1983, 1986; Mazumder et al. 1988) systems showed that this is approximately true, especially in more oligotrophic environments. This work also showed that local modes occur, both in pelagic distributions and in those for benthos (Warwick 1984; Schwinghamer 1981, 1983). These modes appear to correspond to the major ecophysiological groups (viz. protists, phytoplankton, zooplankton ...).

These empirical observations of size distributions have lead to theoretical models of plankton community physiology (e.g. Platt and Denman 1978; Borgmann 1982; Griesbach et al. 1982) founded upon apparent regularities in the size structure, and the well-established size dependence of production, respiration, and other physiological processes (e.g. Hemmingsen 1960; Banse 1976; Banse and Mosher 1980; Blueweiss et al. 1978). However the development of practical size-based models has been hampered by the paucity of data and the small number of size classes enumerated. The "normalized spectrum" (Platt and Denman 1978; Sprules and Munawar 1986) uses only two parameters to characterize size distributions. But substantial differences in the detailed size structure may by indistinguishable at this level of generalization. Peters (1983b) and Mazumder et al. (1988) used cumulative phosphorus concentration to develop linear models based on an easily measured system variable. These analyses assume that phosphorus is a good correlate of biomass at the level of individual size classes. None of these authors has tested quantitative models with more than two parameters. More detailed models may be required before size structure analysis becomes a useful tool in applications such as the prediction of the fate of contaminants in aquatic systems, of fishery yields, or of the impact of nutrients on the abundance of organisms of a given size.

In this paper, the uniform distribution hypothesis is tested with limnoplankton samples from 15 Quebec lakes. The number, relative sizes, and positions of local modes are also examined. In particular, the parameters of the distributions are tested for variability, since insufficient variability would obviate the ability to use variations in size distribution to predict variations in community physiology. To the extent that parameters are variable, several predictive relationships are examined. Predictive power is evaluated using an independent data set.

Methods

Sampling -- Samples were collected during the day throughout the growing season in 1986 and 1987 from 15 temperate lake sites in southern Quebec. These sites represent a broad range of trophic conditions, as evidenced by phosphorus and chlorophyll concentrations (Table 1.1). Integrated epilimnetic water samples (at least 20 L per site) were collected during the day through a 2.5 cm diameter tube connected to a piston pump.

Although Pace (1986) found that this apparatus collects both phytoplankton and zooplankton effectively, six parallel samples were collected with both the tube and a vertical haul with a conical plankton net (25 cm diameter, 120 µm mesh), for subsequent comparison of their efficiencies of collection of zooplankton (Table 1.2). Five of the six counts were similar, and showed no consistent difference between the sampling devices with respect to the number of organisms, the total volume, or the mean size. The one discrepant count represents a dense concentration of small (<500 µm length) Daphnia, and is assumed to represent a swarm which was not otherwise sampled.

Water samples were poured into 1 and 4 L bottles, and kept in the shade during transit to the laboratory refrigerator (maximum duration 1 h). Samples for microscopic examination were preserved with either Lugol's iodine (for the phytoplankton size range--1 to 100 μm) or 5% formalin (for the bacteria--0.2 to 1.5 μm, and zooplankton--80 to 1500 μm).

Size Composition -- The size distribution of plankton throughout the range 0.2 to 1500 μm equivalent spherical diameter (ESD) was determined by direct microscopic examination.

To test for comparability with studies using electronic particle counts, 7 samples were compared over the size range 4 to 31 μm using both microscopic counts and Coulter counts with a 200 μm aperture (Table 1.3). Two of the seven distributions differed significantly ($P<0.05$; Kolmogoroff-Smirnoff test). However, total counts estimated by the Coulter counter were much higher (1.7X to 18.6X) in 6 of 7 comparisons. These differences were primarily due to the two smallest size intervals, where abiotic particles are more abundant (Lal 1977). Excluding these two size classes, the counts were similar, except in the case of a sample from Lake Waterloo, where the microscopic count was higher, and a sample from Baldwin Pond, where the Coulter count was higher. Since Lake Waterloo has the highest phosphorus concentrations (>100 mg m^{-3}) of all lakes studied, while Baldwin Pond has one of the lowest (<10 mg m^{-3}), the microscopic counts are consistent with the expected trophic response, whereas the

Table 1.1. Southern Quebec lakes sampled in 1987, in order of increasing total phosphorus. Concentrations are for total epilimnetic samples (mg m^{-3}, with SD). Dates are Julian dates. Central, North, South, and Newport are basins of Lake Memphremagog.

Lake	Latitude	Longitude	No. of samples	Range of dates	Total phosphorus	Total chlorophyll
Bowker	45° 25′	72° 15′	1	183	3.50	1.39
Stukely	45° 20′	72° 15′	2	193-230	5.53 (1.53)	1.23 (1.40)
Orford	45° 15′	72° 20′	5	128-237	7.61 (2.90)	1.19 (0.50)
Lyster	45° 04′	71° 55′	3	149-227	10.16 (5.01)	1.44 (0.91)
Baldwin	45° 04′	71° 54′	3	149-227	14.23 (1.57)	2.10 (0.73)
O'Malley	45° 13′	72° 20′	1	177	14.41	4.01
Central	45° 10′	72° 15′	8	114-228	14.46 (3.03)	1.97 (0.89)
Cerises	45° 18′	72° 15′	2	193-236	14.74 (2.59)	3.82 (3.56)
North	45° 15′	72° 15′	5	120-226	15.03 (2.08)	2.18 (1.32)
Massawippi	45° 15′	72° 05′	1	177	15.52	2.55
Lovering	45° 10′	72° 10′	2	128-177	16.14 (3.01)	1.42 (0.44)
Argent	45° 15′	72° 25′	2	142-183	16.57 (0.71)	3.02 (2.01)
Trousers	45° 10′	72° 25′	1	181	17.77	3.38
South	45° 00′	72° 15′	6	114-235	19.27 (2.19)	3.87 (0.74)
Newport	44° 55′	72° 15′	4	128-204	20.70 (3.42)	3.16 (0.70)
Brome	45° 15′	72° 30′	1	142	20.82	3.15
Pond	45° 10′	72° 14′	3	183-224	24.59 (2.16)	2.54 (1.80)
Magog	45° 20′	72° 05′	4	128-223	32.61 (6.87)	5.01 (2.98)
Waterloo	45° 20′	72° 30′	4	142-229	81.72 (35.1)	17.64 (11.0)

Table 1.2. The densities and total volumes of organisms in the size range from 125 to 1500 μm ESD, sampled with a plankton net and a pump and tube Samples are in order of increasing total phosphorus concentration (measured on the tube sample).

Lake	Date	Density (L^{-1})		Volume (ppm)		Mean size ($\mu m^3 x 10^{-6}$)	
		Net	Tube	Net	Tube	Net	Tube
Stukely	230	3	14	0.6	0.3	295 8	37.8
Orford	203	2	3	0.1	0.1	38.5	22.3
South	188	539	73	21.0	1.0	48.5	14 0
Newport	204	54	86	1.7	2.8	71.9	38.5
Pond	193	66	65	2.5	2.7	115.2	103.1
Magog	223	91	94	4.2	2.9	57.2	60.1

Table 1 3 Comparative counts (ml^{-1}) from the electronic Coulter counter and inverted microscope. Interval headings are the base 2 logarithms of particle volume (μm^3). Note the differing totals (7-13) from the two methods for lakes Baldwin and Waterloo.

Lake	5	6	7	8	9	10	11	12	13	Totals	
										5-13	7-13
Coulter counter											
Baldwin	16208	10419	4168	1033	368	138	56	16	3	32409	5782
Omalley	5062	3221	1134	431	183	64	20	4	1	10120	1837
Argent	3730	1750	1156	551	191	59	13	4	1	7455	1975
Waterloo	3056	1083	835	519	247	223	95	40	6	6104	1965
Bowker	2632	1640	629	236	84	28	8	3	1	5261	989
Central	1663	735	462	260	132	47	18	4	1	3322	924
Central	1294	574	355	202	88	43	20	5	1	2582	714
Microscope											
Waterloo	870	1739	1217	1044	1261	739	1174	478	130	8652	6043
Omalley	291	558	655	485	218	279	146	121	73	2826	1977
Bowker	364	625	523	352	239	205	80	46	57	2491	1502
Argent	225	556	460	182	193	267	246	150	75	2354	1573
Central	396	499	484	132	103	176	103	29	29	1951	1056
Baldwin	242	339	339	194	194	230	121	49	36	1744	1163
Central	182	148	159	114	114	136	46	23	23	945	615

Coulter counts are not. Hence the Coulter counter is not recommended for the enumeration of natural samples of small freshwater plankton (less than 50 μm ESD), where abiotic particles may be numerous.

The abundance and approximate diameters of bacteria were determined by epifluorescence microscopy with DAPI stain (Porter and Feig 1980) of samples preserved in 2% formaldehyde. Larger organismsin the picoplankton fraction were measured at 1250X on an inverted microscope, using samples preserved in Lugol's iodine solution. Nannoplankton were measured at 1000X and 400X (Lund et al. 1958). Zooplankton were measured at 100X on the inverted microscope and at 40X under a dissecting microscope, using both Lugol's samples and others preserved in 2% formalin. Individual volumes of organisms larger than 5 μm greatest axial linear dimension (GALD) were estimated by measuring length and width and taking the volume of similar regular geometric shapes as approximations. Smaller organisms were counted in nine diameter intervals between 0.2 and 5 μm.

Chemical analyses -- Epilimnetic chlorophyll and phosphorus concentrations are easily measured variables known to be correlates of the abundance and activity of some planktonic organisms (e.g. Smith 1979; Elser et al. 1986). Hence these were measured, along with epilimnetic dry weight, as potential predictors of biomass. The total phosphorus concentrations were measured in triplicate in several size fractions, separated by stainless steel screens (130, 80, 40 μm porosity) and Nuclepore membrane filters (5 and 0.4 μm porosity), using the ascorbic acid modification of the molybdenum blue technique (Strickland and Parsons 1968) after digestion with potassium persulfate under pressure (Menzel and Corwin 1965). Whole sample chlorophyll a concentrations were also measured in triplicate (Strickland and Parsons 1968). Epilimnetic dry weights were determined by pouring a known volume of lake water through predried (60° C) weighed GFC filters, and reweighing the filters after they had dried for 24 h to determine the weight difference due to seston particles and plankton.

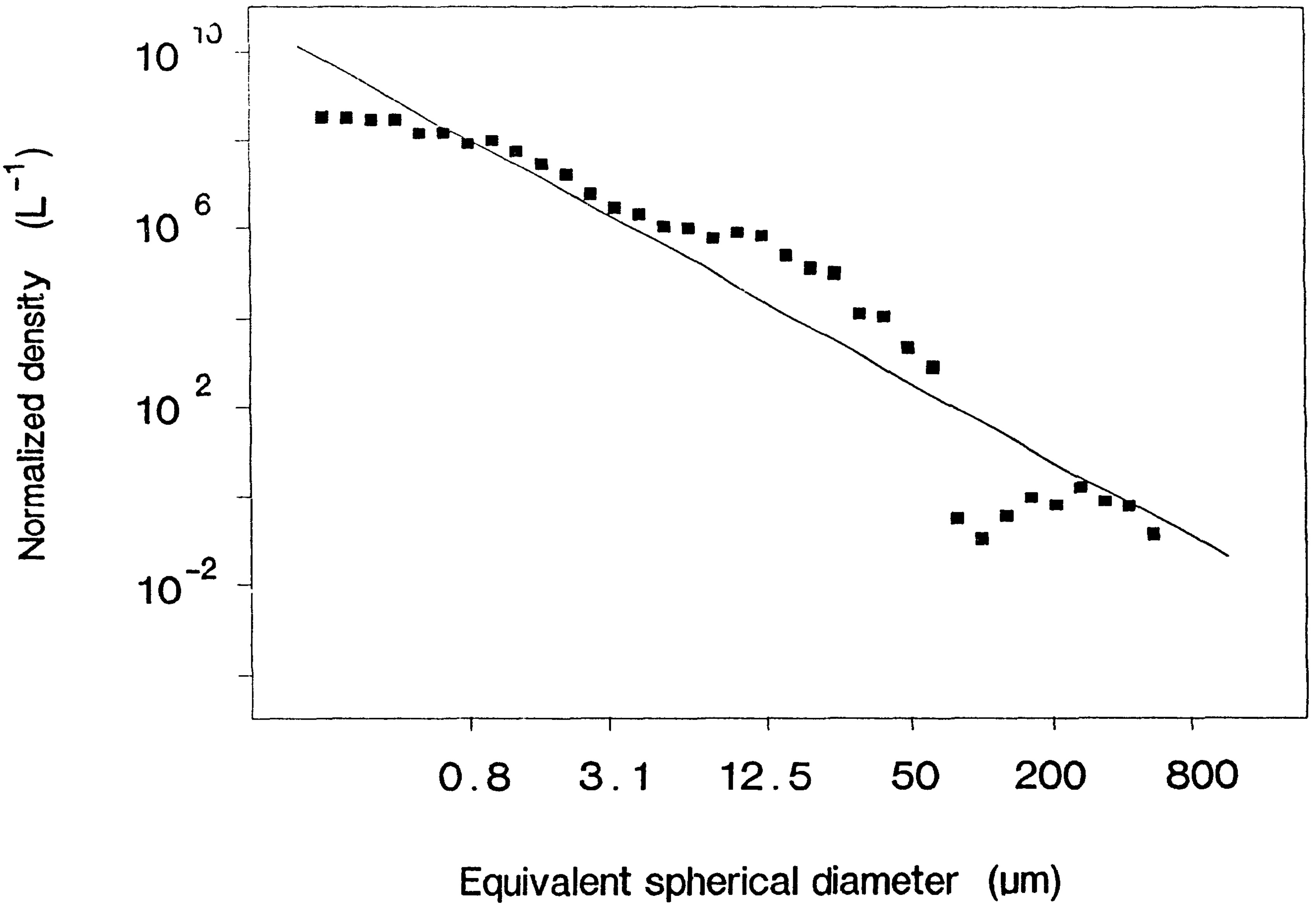
Data analyses: Size structure - At any magnification, there are maximum and minimum size limits to the organisms which can be measured. Larger organisms are too big for the visual field of the microscope, or may be substantially out of focus if the field depth is too shallow. As organism size decreases, and well before it is as small as the theoretical detection limits at a given magnification, the probability of detection decreases. Small organisms may be hidden by larger ones, confused with debris, or simply overlooked in a busy visual field. Thus one might expect an approximately bell-shaped curve of visual selectivity with a microscope, just as there is a size selection curve for a plankton net. The five magnifications used in this study allowed sufficient overlap between successive selection curves that, for many sizes, a choice was required between the counts at the two magnifications. In every case, the highest count was selected, on the assumption that lower counts were the result of partial selectivity. Computer programs were written to take length, width and shape data, magnifications, and water volumes to compute individual organism volumes, and total counts and total volumes in intervals of volume doubling.

The geometric mean volume of organisms in the smallest interval was 0.0055 μm^3, corresponding to equivalent spherical diameters from 0.2 to 0.25 μm. The mean for the next interval was 0.0110 μm^3, corresponding to ESD 0.25 to 0.31 μm. In the largest interval counted (the 39th), the mean volume of 1.5×10^9 μm^3 includes organisms with ESD between 1250 and 1600 μm.

Curve fitting - Size distributions can only be effectively modelled when there are counts in most of the 39 size intervals. In the case of these collections, 25 of the 58 samples analyzed were amenable to formal curve fitting procedures. Some samples could not be fully counted due to the presence of excessive abiotic particles. Others presented highly "accidented" distributions with empty intervals followed by large single interval peaks, due to the periodic abundance of single phyla. Finally, some oligotrophic lakes yielded too many zero counts in the zooplankton range. Some of these problems could be addressed; for example, rare organisms would be discovered if sampling were sufficiently intensive. But many of these problems cannot be dealt with after the fact. For the purpose of curve fitting, a total of 25 samples with "good coverage" were identified, where good coverage is defined as non-zero counts in at least 90% of the size classes, including at least some size classes above 600 μm ESD. Only samples meeting these criteria were used to estimate the parameters of size distribution models, described in the following paragraphs.

The spectrum of volume concentrations in base 2 logarithmic size intervals was used to compute the "normalized spectrum" (Platt and Denman 1978; Sprules and Munawar 1986; Figure 1.1) for each sample. Such spectra plot log size on the abscissa, and the ordinate is the log of the standardized abundances per interval. The latter is calculated as the volume concentration in the interval divided by the change in modal volume between intervals. The parameters of the straight line fitted to these points may be used to compare samples. Integration over any range of sizes provides a smoothed estimate of biovolume or biomass over that range. The intercept of the line provides an estimate of relative abundance at one mass (or volume) unit. The slope reflects the overall trend in mass or volume change from interval to interval. In particular, a slope of -1 would

FIGURE 1.1. An example of a normalized spectrum (Lake Orford, July 22, 1987). The logarithms of organism size (on the abscissa) are plotted against the normalized density (on the ordinate). The latter is the logarithm of the ratio between the total volume in the interval and the difference in organism size between that and the subsequent interval - hence, a density estimate. The straight line is the least squares regression fitted to these data. See Appendix 1 for further information.


10^10
10^6
10^2
10^-2
Normalized density (L^-1)
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (µm)

indicate an approximately even distribution of mass over size classes. Steeper slopes (with larger absolute values) would reflect declining biomass with increasing size, whereas shallower slopes would reflect the reverse. The existence of any significant slope would confirm that there is an overall trend with size, in spite of the existence of one or more local peaks or troughs in the size spectrum. Further details are provided in Appendix 1.

There are usually patterns in the residuals from a fitted normalized spectrum, and hence the diagnostic statistics associated with the regression (r^2, F, SEE) will probably not be accurate. Nevertheless, if the spectra to be compared are computed over similar size ranges and have similar residual patterns, then the parameters may provide a basis for comparison of general trends in biomass distribution. If the normalized spectrum is an accurate representation of the data, then numerical integration under the fitted curve (multiplying together back-transformed ordinates and abscissas and summing) should provide an estimate of biomass similar to that observed over the same range. To test this hypothesis, the predicted integrated biovolumes from normalized spectra were regressed against observed total biovolumes. Furthermore, the median computed total biovolume from all samples was compared with the observed median value using the modified t test for samples with unequal variances (Snedecor and Cochran 1967). These tests provided a measure of confidence in the normalized fits, as well as a basis of adjustment between predicted and observed values. The results were also compared with those of Sprules and Munawar (1986) for Ontario lakes.

The finer structure of distributions was examined in three stages. We first examined the possibility that plankton volume might be lognormally distributed over size. Since our size distributions are grouped by logarithmic (base 2) volume intervals, they should look like a normal distribution when graphed (log size on the x axis and linear volume on the y axis). A Kolmogoroff-Smirnoff test (Hollander & Wolfe 1973) was used to test an overall lognormal (unimodal) fit to these data. We then tested the hypothesis that the distributions were bimodal, corresponding to 2 lognormal curves. If this hypothesis were true, then each distribution would be characterized by 5 independent parameters: 2 means, 2 variances, and a parameter between 0 and 1 describing the proportion of the total distribution in one of the modes. These parameters of the mixture density function were sought by a maximum likelihood method (Appendix 1; Clarke 1984). Finally, a linearization technique for estimating the parameters of a non-linear system (Appendix 1; Draper and Smith 1981) was used to estimate the 8 independent parameters of trimodal fits.

Variability of parameters - The heterogeneity among estimated parameters was evaluated in several ways. All pairwise comparisons were made with the Least Significant Difference (95% criterion) using Tukey's test based on the Studentized range rather than the t distribution. Use of the latter would increase the chance of finding more significant differences than is warranted (Type I error; John 1971). Joint confidence regions for the two parameters of the normalized spectra were compared graphically (Appendix 1). When it was necessary to compute statistics for comparing parameter sets with two parameters, the models were first tested for homogeneity of variance by Bartlett's test, and if they passed this test, they were compared using an F test for the coincidence of two or more straight line regressions (Seber 1977).

Predictions - Predictive linear regression models were used to test the ability of various chemical and morphometric variables to predict the parameters of size distributions, to evaluate the correspondence between predictions and observations, and to identify trends between predictions and residuals. Models were evaluated using a data set (1986) independent from that used for their development (1987). In order to evaluate the potential use of phosphorus concentration as an index of size specific biovolume, the phosphorus:volume ratio was compared among size classes by analysis of variance (Seber 1977).

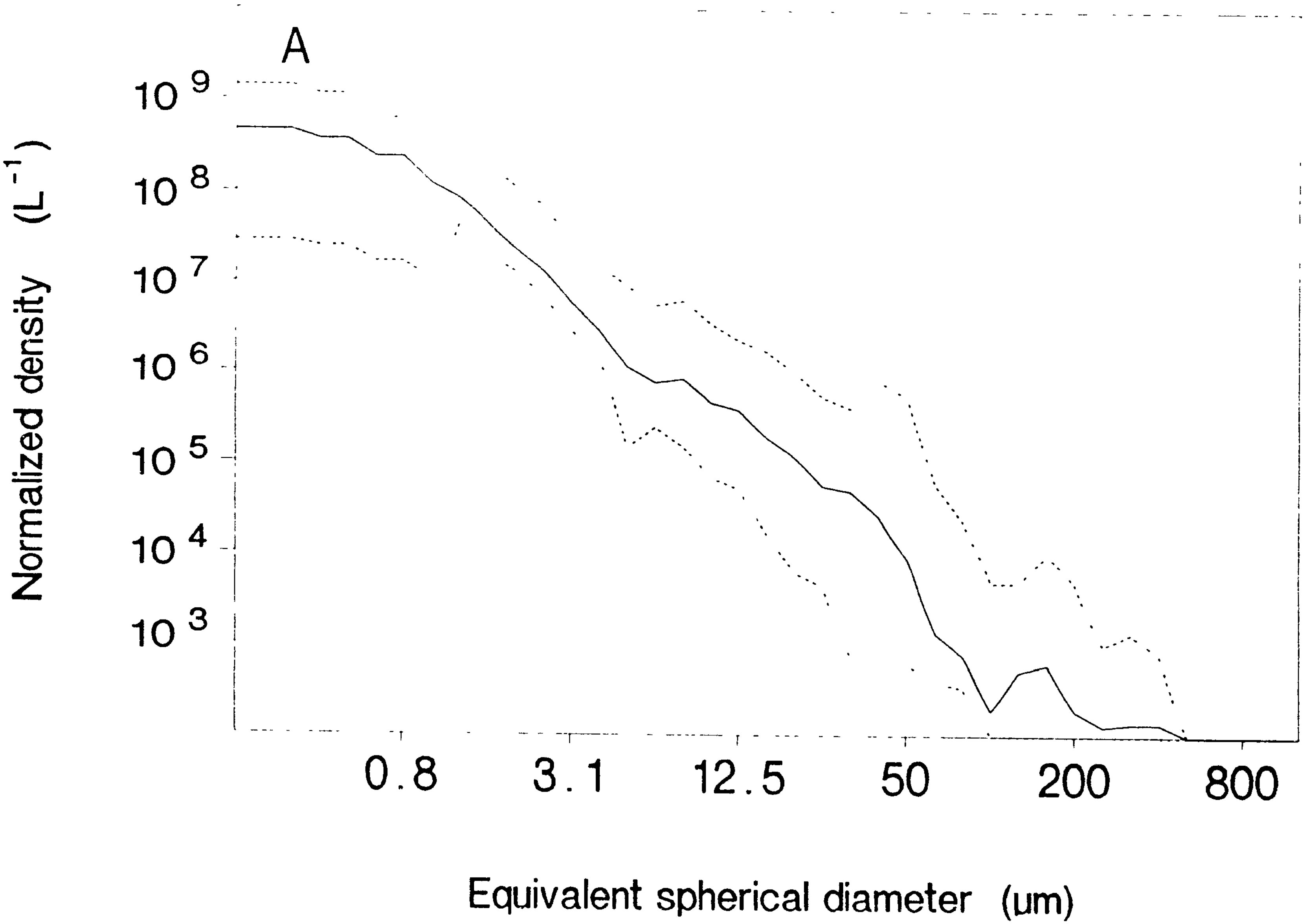
Results and Discussion

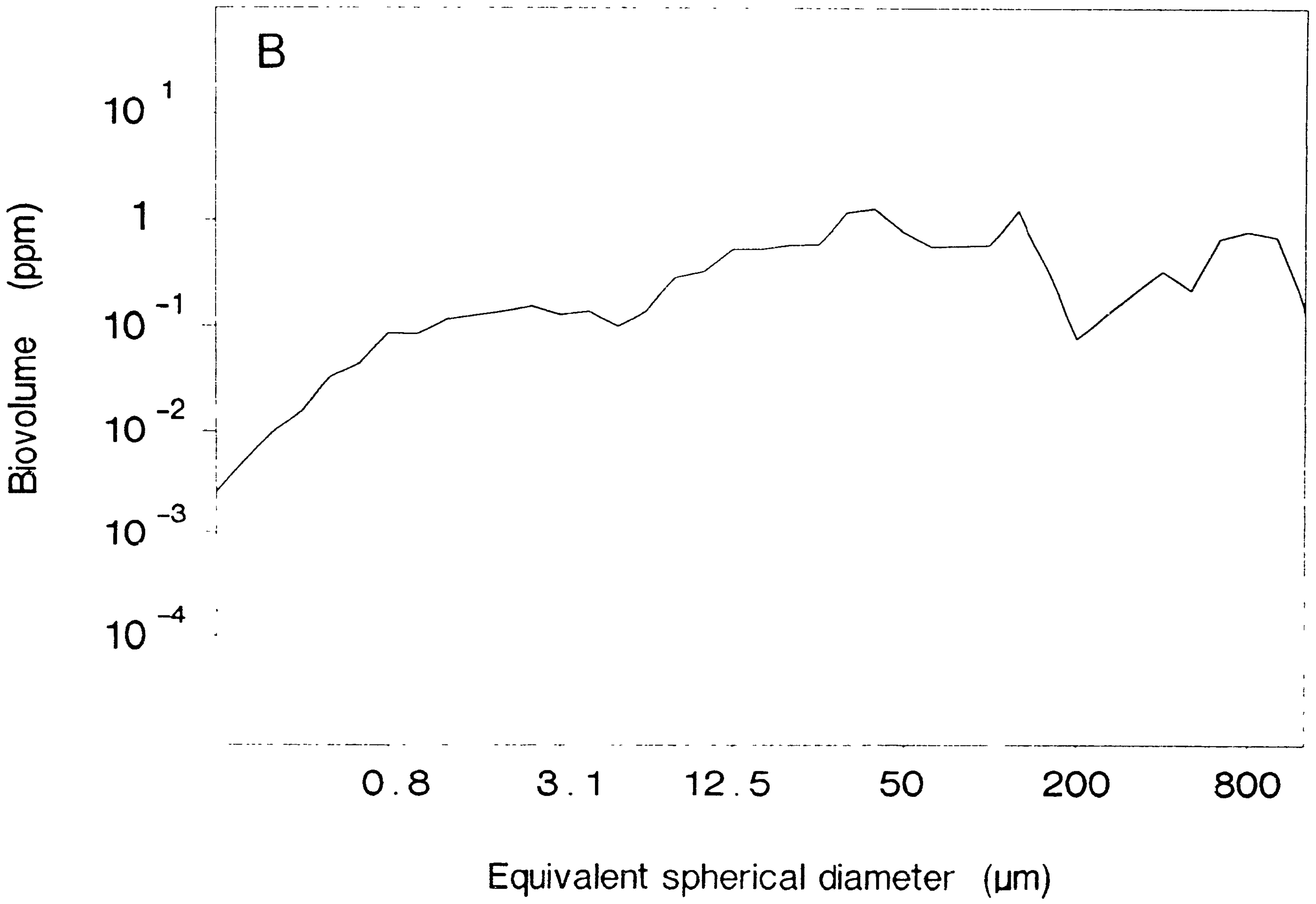
1. Detailed Size Structure

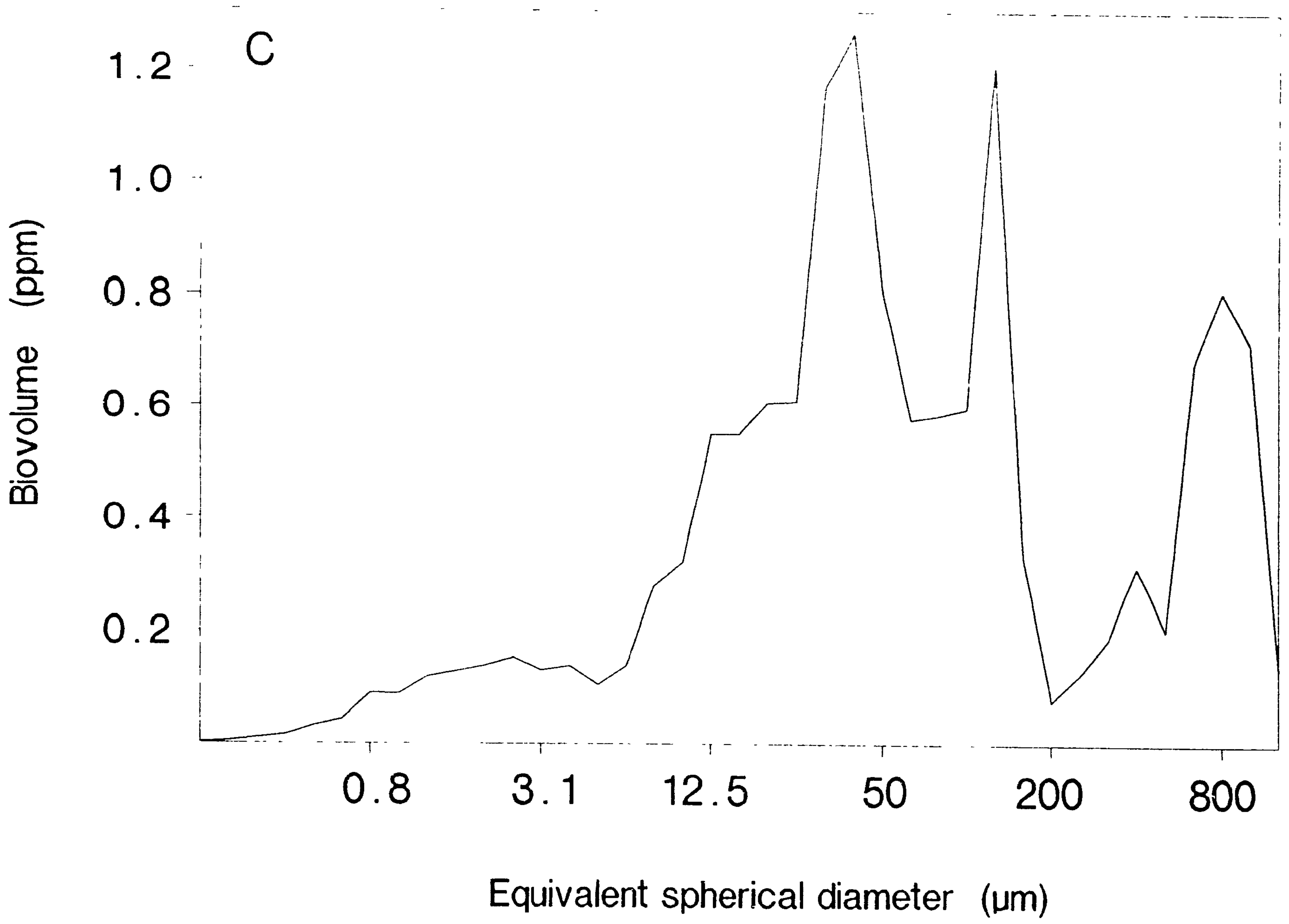
The distribution of organism density over logarithmic size classes showed a fairly uniform decline with increasing size (Figure 1.2A). The "median distribution of biomass", composed of the medians of all observed values for each size interval (Figures 1.2B & 1.2C), shows a rapid increase to about 1 μm ESD, a more gradual increase to about 30 μm, and an erratic distribution beyond that point. On average, 50% of the community biomass is less than 45 μm ESD, and 90% less than 700 μm (Figure 1.2D). The summed biomass in this median distribution is 14.1 mg L^{-1}. In contrast, the 12 samples with lowest phosphorus concentrations (median = 6.6 μg L^{-1}, versus an overall median of 17.4) provide a mean biomass of 6.2 mg L^{-1}. The peak in the distribution for these oligotrophic lakes occurs at 30 μm (Figure 1.2E), and 50% of the biomass is in particles smaller than this. 90% of the biomass is attained by 175 μm. Hence these oligotrophic lakes are clearly dominated by smaller organisms and have lower total plankton biomass concentrations than do lakes with more phosphorus in the epilimnion.

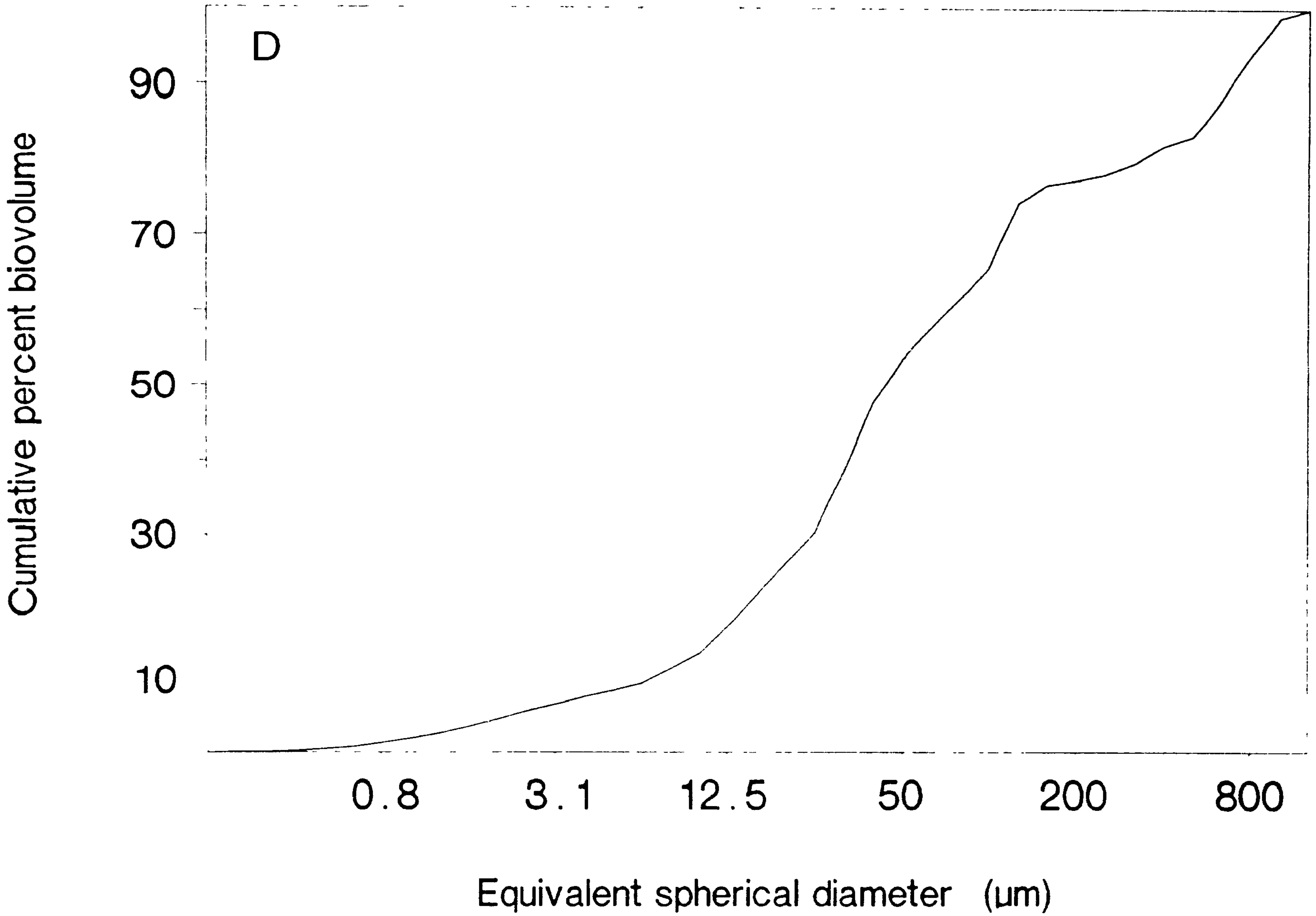
Two lakes were sampled often enough throughout the season to examine temporal trends in the phytoplankton/zooplankton ratio

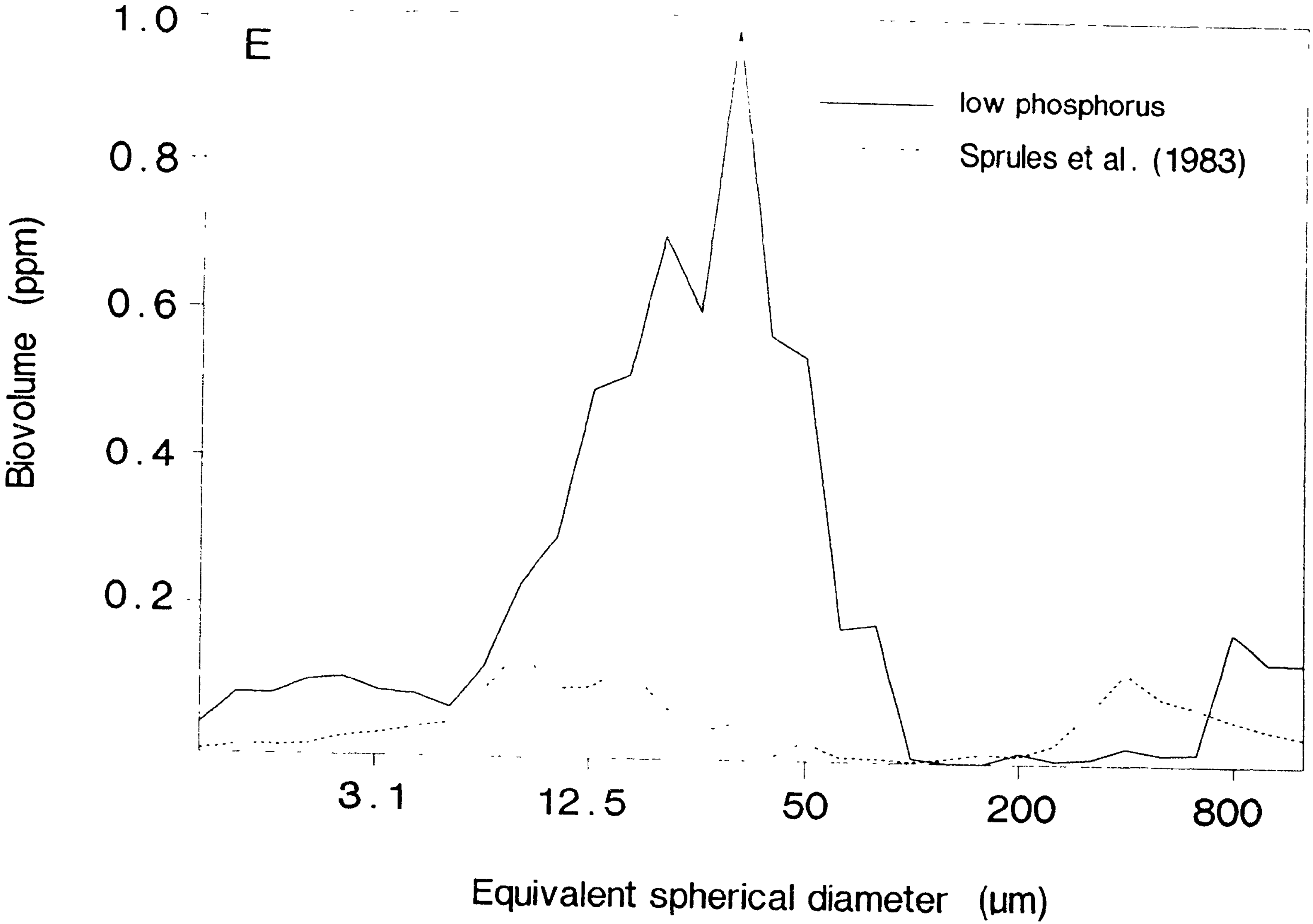
FIGURE 1.2. The median distributions (n=58) over logarithmic size classes of (A) organism density, (B) total organism volume on a logarithmic scale, (C) total organism volume on a linear scale, (D) cumulative total organism volume, and (E) low phosphorus total organism volume. The 95% confidence intervals are shown on the first two panels. Parts per million (ppm) = μm^3 L^{-1} x 10^{-9}. The mean distribution obtained by Sprules et al. (1983) for low phosphorus Ontario lakes (-------) is included in panel (E).


A
10 9
10 8
10 7
10 6
10 5
10 4
10 3
Normalized density (L -1)
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (um)


B
Biovolume (ppm)
10 1
1
10 -1
10 -2
10 -3
10 -4
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (µm)


C
1.2
1.0
0.8
0.6
0.4
0.2
Biovolume (ppm)
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (µm)


D
Cumulative percent biovolume
90
70
50
30
10
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (µm)


E
low phosphorus
Sprules et al. (1983)
Biovolume (ppm)
1.0
0.8
0.6
0.4
0.2
3.1
12.5
50
200
800
Equivalent spherical diameter (µm)

(Table 1.4). There is some indication of increasing biomass, and in the more eutrophic lake, an increasing proportion of the total biomass is in the zooplankton size range. Using all samples with good coverage of the size spectrum from all lakes, a regression on Julian dates to predict the proportion of total biomass below 80 μm shows a significant ($P<0.001$) decreasing trend (slope = -.00299), explaining 33% of the variation. Thus there is a general trend to increasing size during the period from April to September in these lakes.

Trophic response models, using the total phosphorus concentration as a trophic index, provide a crude basis of comparison for our total biomass results. Peters (1986) summarized several power equations to predict components of plankton biomass from phosphorus. Using these equations, the predicted biomasses at 7 and 17 μg phosphorus L^{-1} are 1.9 and 4.4 mg L^{-1} respectively. These predictions are somewhat lower than our observed median values of 6.2 and 14.0 mg L^{1}. However, the predictions are based on backtransformed logarithmic regressions, and must therefore be increased by a correction factor (Sprugel 1983). Neither Peters (1986) nor the original references provided sufficient statistical information to estimate this factor accurately, but an increase of 25% would not be atypical. Considering that the predictions are based on 5 independent data sets in a variety of lakes, our results are perhaps not too unusual.

Table 1.4. Temporal trends in planktonic biomass distribution in 2 lakes in 1987. The first data column shows the approximate size of an organism (μm ESD) at the 50% point in the distribution. The second column shows the percentage of the total biomass which is due to organisms smaller than 80 μm. The third column shows the total biomass concentration (ppm).

Lake Magog

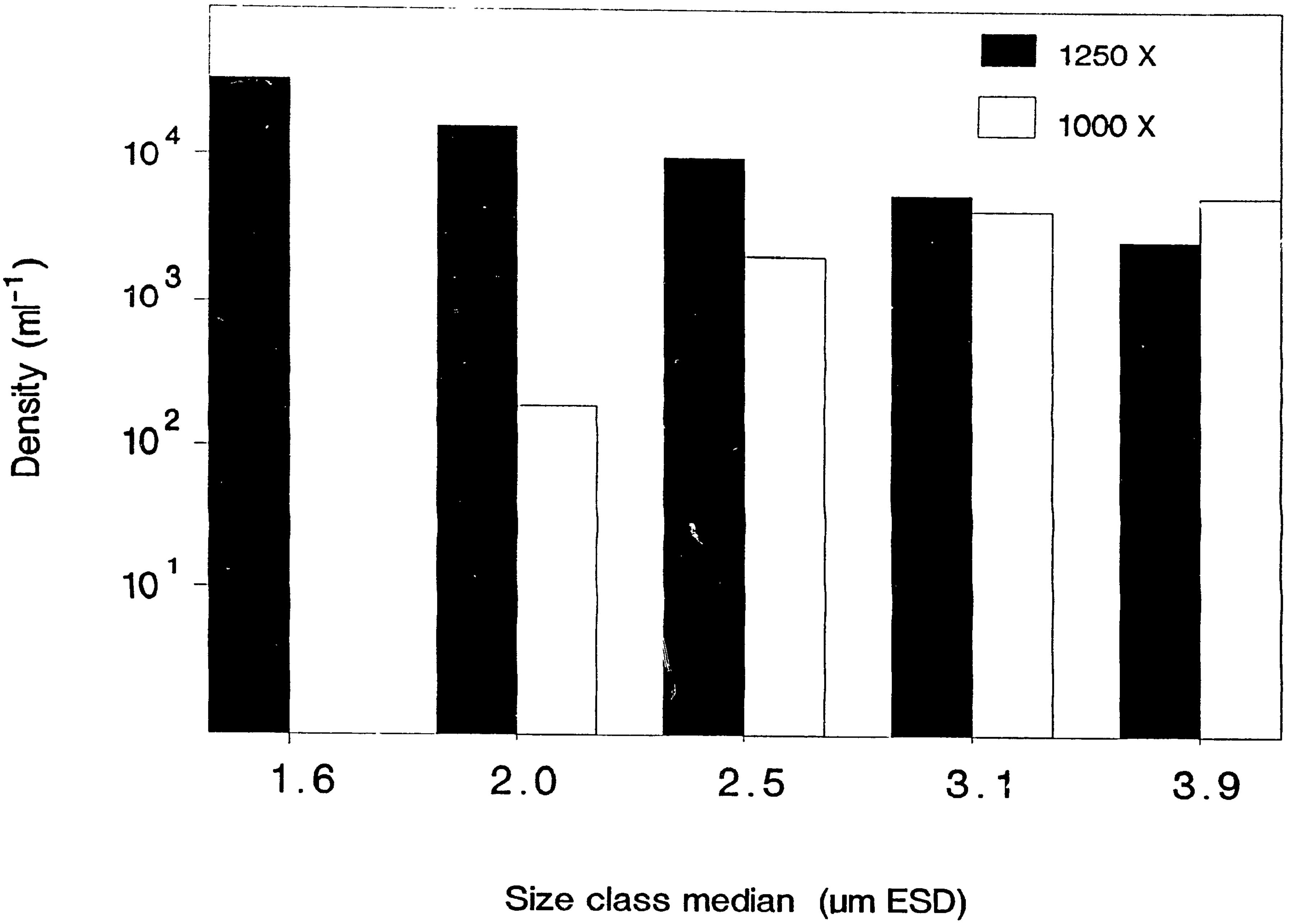
Julian date	Size at 50%	Percentage at 80 μm	Total
128	8	97	3477
181	14	97	18072
193	28	74	18328
223	50	58	18854
Lake Orford			
128	4	98	2086
177	28	92	6958
199	10	93	6466
203	18	94	8285
237	40	65	20140

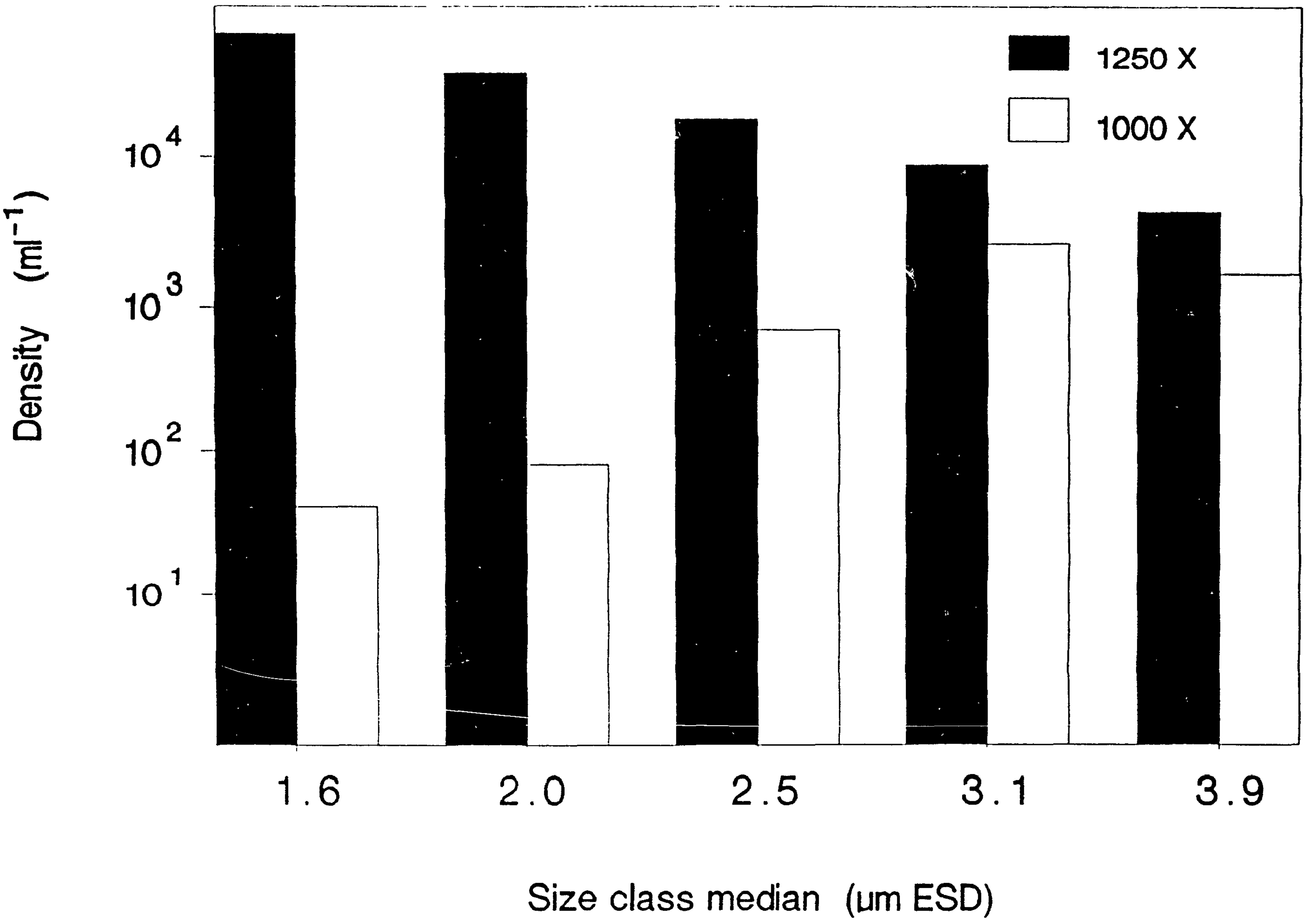
Sprules et al. (1983) determined the planktonic size distribution in several small Ontario lakes, with lower phosphorus concentrations than most of ours (Tables 1.1 and 1.5). Whereas their mean distribution shows a phytoplankton peak at about 8 μm, our low phosphorus median distribution peaks at about 30 μm (Figure 1.2E). About half of the total volume is in each of their two modes, whereas usually less than 10% of ours is in the second mode. Over the same size range ($1<\text{ESD}<1500$ μm), their integrated biomass of 1.5 mg L^{-1} is substantially lower than our 6.2. Although our lakes are larger, deeper, and richer in phosphorus (Table 1.5), the differences in observed biomasses may be due to methodology. The lower end of the observable size distribution depends upon the magnification used. Although they do not specify magnification, detection rates normally decline as size decreases. We consistently found that, when magnification was increased, counts of organisms in the smallest size ranges detected by the previous magnification were considerably higher under the new power (Figure 1.3). Our higher picoplankton counts may be due to our use of 1250X magnification for all organisms smaller than 5 μm ESD, and epifluorescence microscopy for bacterial sizes (<1.5 μm). Our bacterial counts are similar to those obtained by other authors using similar methods (Hobbie et al. 1977). Hence we are confident that our higher biomasses of bacteria and picoplankton are not a methodological artefact.

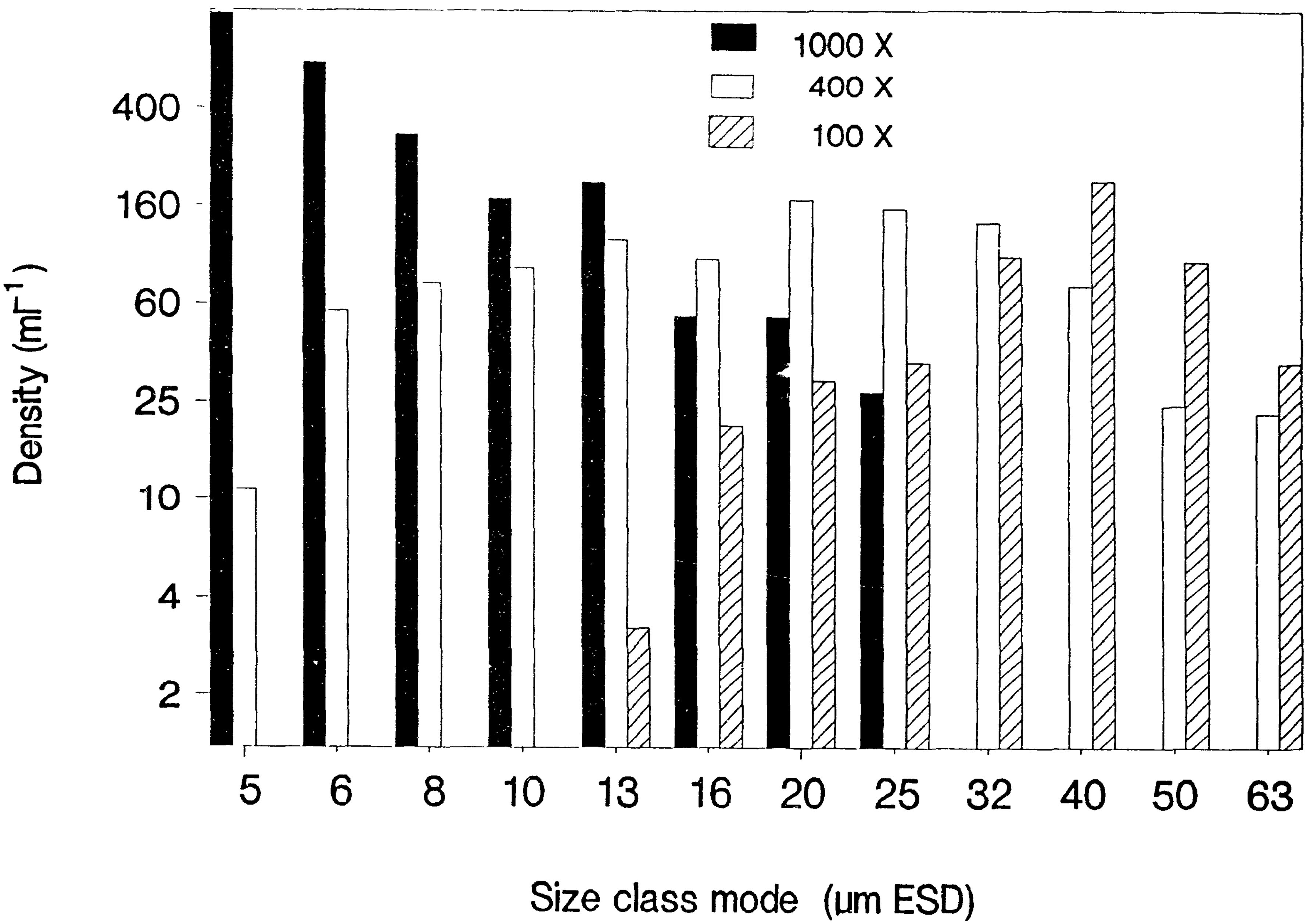
Table 1 5 Comparisons between our lakes and those studied by Sprules et al. (1983) The range and median are shown.

	All samples	Low phosphorus	Sprules et al.
Total phosphorus ($mg\ m^{-3}$)	3 - 130 (17)	3 - 12 (7)	0.5 - 27 (4)
Lake area (ha)	20 - 4000 (200)	120 - 4000 (1/0)	29 - 1142 (149)
Mean depth (m)	1 - 70 (10)	14 - 70 (23)	1 - 20 (8)

FIGURE 1.3. Examples of the impact of magnification used in microscopic analyses on observed densities. Samples are from South (A) and Central (B) basins and Lake Magog (C).


1250 X
1000 X
Density (ml^{-1})
10^4
10^3
10^2
10^1
1.6
2.0
2.5
3.1
3.9
Size class median (um ESD)


1250 X
1000 X
Density (ml^{-1})
10^{4}
10^{3}
10^{2}
10^{1}
1.6
2.0
2.5
3.1
3.9
Size class median (µm ESD)


1000 X
400 X
100 X
Density (ml⁻¹)
400
160
60
25
10
4
2
5
6
8
10
13
16
20
25
32
40
50
63
Size class mode (µm ESD)

The large difference between our results and those of Sprules et al. is primarily due to the higher biomasses we observed in the 25 to 125 μm size range. The low abundance of larger phytoplankton in their samples could also reflect shallower sampling with a small diameter sampling tube. We used a deeper range of depth integration (8 metres as opposed to 4), and our sampling tube has 32 times the cross-sectional area of theirs. Counts from this tube were used for all size classes up to 130 μm ESD. Nevertheless the basis of the differences in the distributions will remain speculation until parallel samples from both sampling protocols in both environments have been compared.

Kolmogoroff-Smirnoff tests showed that 46 of 58 volume distributions did not differ significantly from a unimodal lognormal distribution ($P<0.05$; Table 1.6; Figure 1.4). Modes ranged from 19 to 75 μm ESD, with a median value of 48 μm for the 17 samples with both good coverage of the size range and significant fits. A unimodal distribution is a good first approximation to the plankton community size distribution, apparently because the phytoplankton mode so dominates the community.

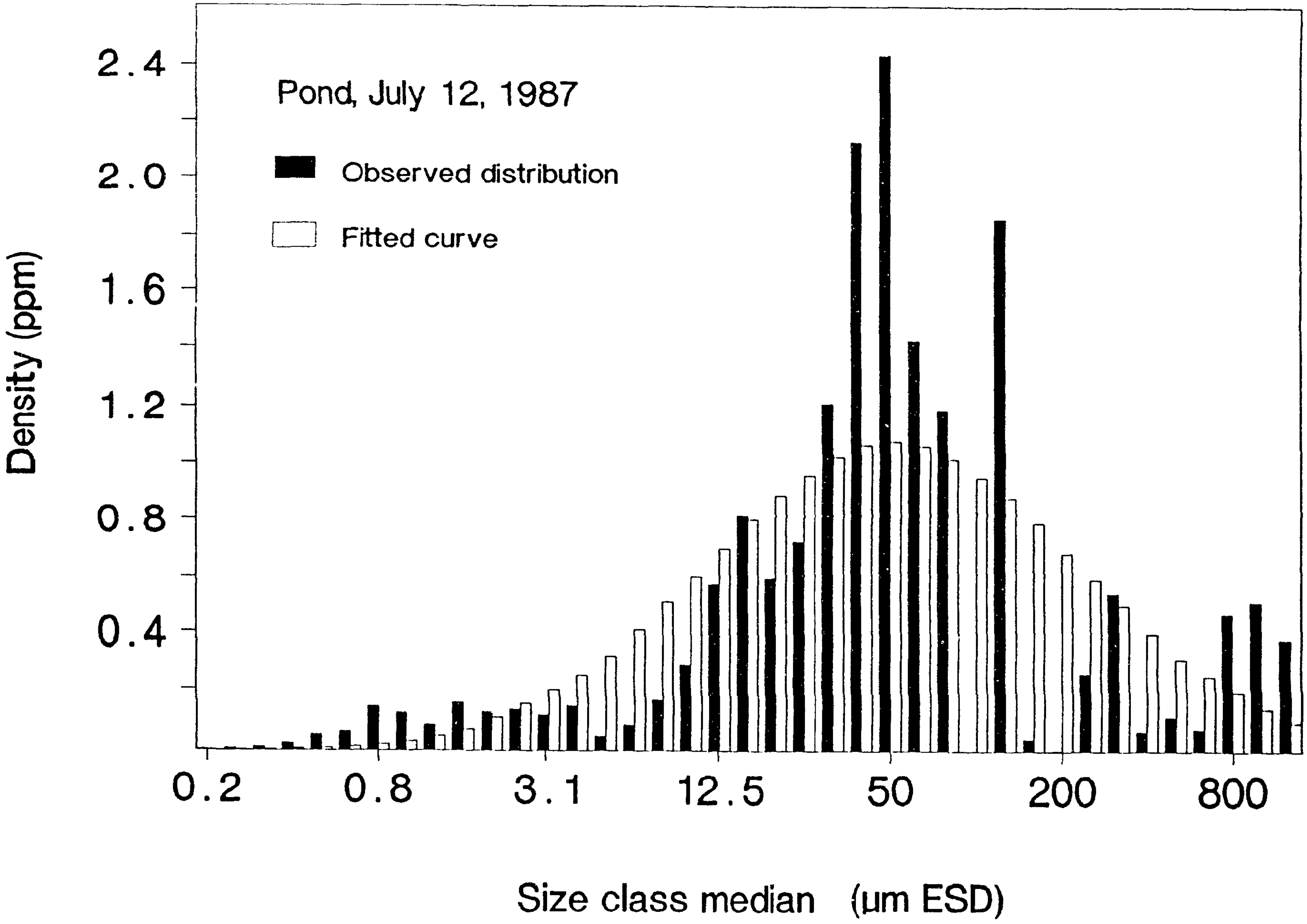
Bimodal and trimodal distributions were fit to the data, but met with limited success (Table 1.7). Only 7 samples out of 25 with good coverage did not differ significantly from a

Table 1.6. Parameters of all unimodal distributions successfully fitted to samples with good coverage of the size range. The mean is the mean base 2 logarithm of organism volume (μm^3). Samples are in order of decreasing total phosphorus concentration. The volume in a given size interval (on a base 2 log scale) may be estimated by multiplying the corresponding normal density* by the total volume.

Lake	Julian date	Mean (μ)	SD (σ)	Total volume (ppm)
Waterloo	229	16.45	4.20	192.7
Magog	223	15.77	6.56	29.5
Magog	193	15.92	6.93	18.3
Pond	193	15.96	6.48	17.3
Newport	204	17.04	7.13	14.7
South	235	15.49	5.40	29.9
Cerises	236	12.45	5.31	12.0
Central	228	17.53	7.64	35.9
Central	197a	17.66	8.22	19.6
Central	197b	17.78	8.13	19.4
North	226	12.51	5.18	9.4
Stukely	193	15.83	7.29	6.9
Orford	199	12.58	6.21	6.4
Lyster	227	13.62	5.43	11.9
Orford	203	11.82	4.31	8.3
Stukely	230	12.90	5.20	11.3
Orford	237	13.62	3.89	19.5

* normal density $(2\pi\sigma^2)^{1/2} \exp[-1/2((x-\mu)/\sigma)^2]$

FIGURE 1.4. Examples of (A) unimodal and (B) bimodal fits to the volume distribution. Note the apparent third mode below about three μm ESD in both panels.


Pond, July 12, 1987
■ Observed distribution
□ Fitted curve
Density (ppm)
2.4
2.0
1.6
1.2
0.8
0.4
0.2
0.8
3.1
12.5
50
200
800
Size class median (µm ESD)

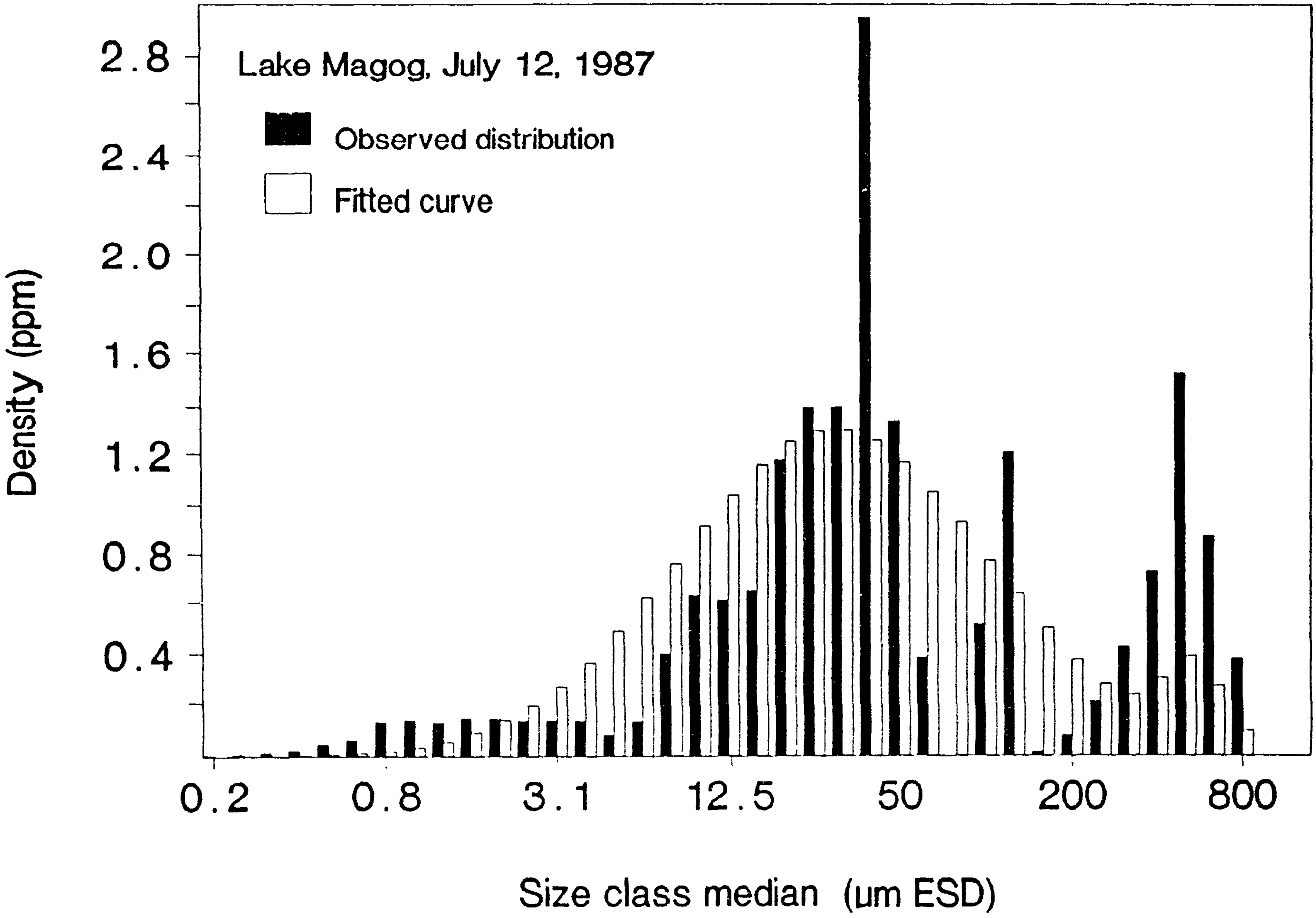

Lake Magog, July 12, 1987
Observed distribution
Fitted curve
Density (ppm)
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.2
0.8
3.1
12.5
50
200
800
Size class median (um ESD)

Table 1.7. Parameters of bimodal and trimodal curves fitted to the distributions of plankton biovolume. Units are base 2 logarithms of volume (μm^3). Samples are listed in order of decreasing total phosphorus concentration. To estimate the volume in a given interval, the contributions of each mode must be weighted by the proportion in that mode and summed. The contribution is itself estimated as in the unimodal case (Table 1.6).

(a) Bimodal:

Lake	Date	First mode		Second mode		Proportion of
		Mean	SD	Mean	SD	total in first mode
Magog	193	13.5	5.39	26.1	1.04	0.81
Pond	224	17.2	7.47	22.5	0.98	0.46
South	235	7.4	8.21	16.2	4.41	0.08
Cerises	236	12.5	5.19	18.8	2.47	1.00
North	226	7.1	4.57	14.3	3.23	0.22
Baldwin	225	11.1	5.81	28.3	0.88	0.61
Orford	203	8.8	6.85	12.8	2.28	0.25

(b) Trimodal

Lake	Date	First mode			Second mode			Third mode		
		Mean	SD	Prop	Mean	SD	Prop	Mean	SD	Prop
Magog	193	2.7	6.29	0.13	14.3	2.51	0.63	26.0	1.30	0.24
Pond	193	6.0	8.71	0.13	15.8	2.65	0.71	28.6	3.18	0.16
Newport	204	1.9	2.27	0.11	13.9	4.74	0.53	20.4	1.00	0.36
Brome	144	3.1	3.41	0.49	11.1	1.00	0.29	21.7	2.75	0.22
South	186	-1.2	1.00	0.33	15.1	1.00	0.14	25.4	2.19	0.53
Baldwin	225	2.5	3.20	0.07	13.2	2.37	0.41	29.0	1.11	0.52
North	226	-4.3	1.59	0.02	11.5	3.55	0.51	15.2	1.00	0.47
Central	130	1.7	3.76	0.51	12.8	5.58	0.30	15.3	1.64	0.19
Lyster	227	9.1	5.18	0.25	14.2	1.99	0.66	27.2	1.00	0.09
Orford	203	2.2	2.65	0.09	12.7	2.63	0.90	25.5	1.32	0.01
Stukely	230	3.8	6.13	0.14	12.3	2.23	0.76	22.1	1.00	0.10
Orford	199	6.0	4.25	0.19	13.0	1.95	0.66	28.9	1.12	0.15

mixture of two normal distributions. Ten samples did not differ significantly from trimodal distributions.

The poor success rate with the bimodal and trimodal fits may not be entirely due to the underlying structure of the community. These models have more parameters to estimate than the simpler unimodal case, and hence require either more data points (size intervals) or more "perfect" data. For example, a trimodal model requires the estimation of 8 independent parameters, and our data are grouped into only 39 intervals. Using narrower, and hence more numerous, intervals would simply generate more noise due to sampling variability. It is also possible that the iterative algorithms used for estimation led to a "dead-end" in the parameter space which is not necessarily the optimal solution. Even when the parameters cannot be estimated analytically, two or three modes are sometimes evident in the distribution (Figure 1.4).

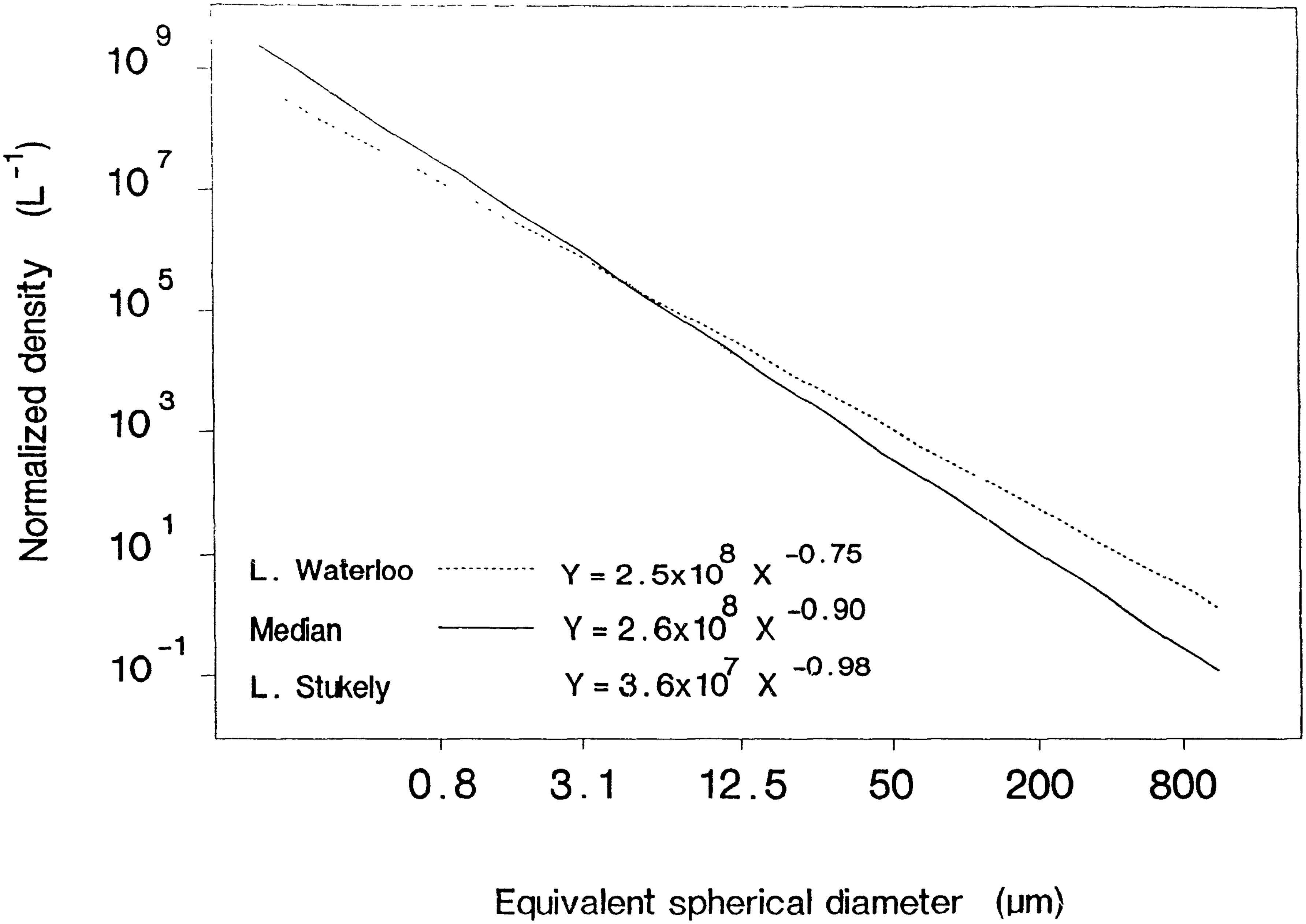
2. Normalized Spectrum

Normalized spectra (Sprules & Munawar 1986) for the 25 samples with good coverage of the size range were computed using base 2 logarithms on both axes (Table 1.8, Figure 1.5). Intercepts varied from 5.16 to 7.98, with a mean of 6.05 (95% C.I. of mean = 5.80 to 6.30). Since all of the regression slopes were highly significant (P<.001), there is indeed a constant trend

Table 1.8 Parameters of the normalized spectra for samples with good coverage of the size range, arranged in order of declining intercept. The independent variable is the base 2 logarithm of organism volume (μm^3) The dependent variable is the base 2 logarithm of the normalized density ($x10^{-6}$ L^{-1}).

Lake	Date	Intercept	SE	Slope	SE	F	R^2	SEE
Waterloo	229	7.981	0.430	-0.752	0.033	516	0.94	1.90
Lovering	177	7.066	0.910	-1.002	0.069	214	0.88	3.91
South	235	6.679	0.734	-0.951	0.048	387	0.92	3.27
South	186	6.499	0.532	-0.829	0.036	520	0.94	2.36
Central	228	6.327	0.458	-0.870	0.031	783	0.96	2.03
O'Malley	177	6.304	0.734	-0.797	0.065	152	0.84	3.22
Cerises	236	6.272	0.521	-0.904	0.043	443	0.94	2.29
Magog	193	6.247	0.516	-0.877	0.035	618	0.95	2.29
Pond	193	6.216	0.578	-0.915	0.037	622	0.94	2.58
Central	197a	6.079	0.425	-0.867	0.029	903	0.97	1.89
Magog	223	6.068	0.384	-0.853	0.027	1011	0.97	1.70
Pond	224	6.020	0.520	-0.872	0.034	649	0.95	2.31
Central	197b	5.960	0.429	-0.862	0.029	878	0.97	1.90
Central	166	5.957	0.533	-0.764	0.047	265	0.90	2.34
Cerises	193	5.952	0.532	-0.760	0.047	263	0.90	2.33
North	226	5.870	0.588	-0.938	0.045	441	0.93	2.60
Baldwin	225	5.827	0.599	-0.944	0.040	552	0.94	2.67
Stukely	230	5.700	0.655	-0.942	0.047	410	0.92	2.90
Central	186	5.636	0.703	-0.894	0.057	250	0.89	3.09
Lyster	227	5.608	0.734	-0.941	0.052	326	0.91	3.25
Orford	237	5.542	0.650	-0.837	0.055	236	0.89	2.86
Orford	203	5.483	0.800	-1.005	0.057	312	0.90	3.54
Newport	204	5.416	0.460	-0.866	0.033	682	0.95	2.04
Orford	199	5.373	0.627	-0.963	0.044	483	0.94	2.78
Stukely	193	5.162	0.683	-0.975	0.043	504	0.93	3.05

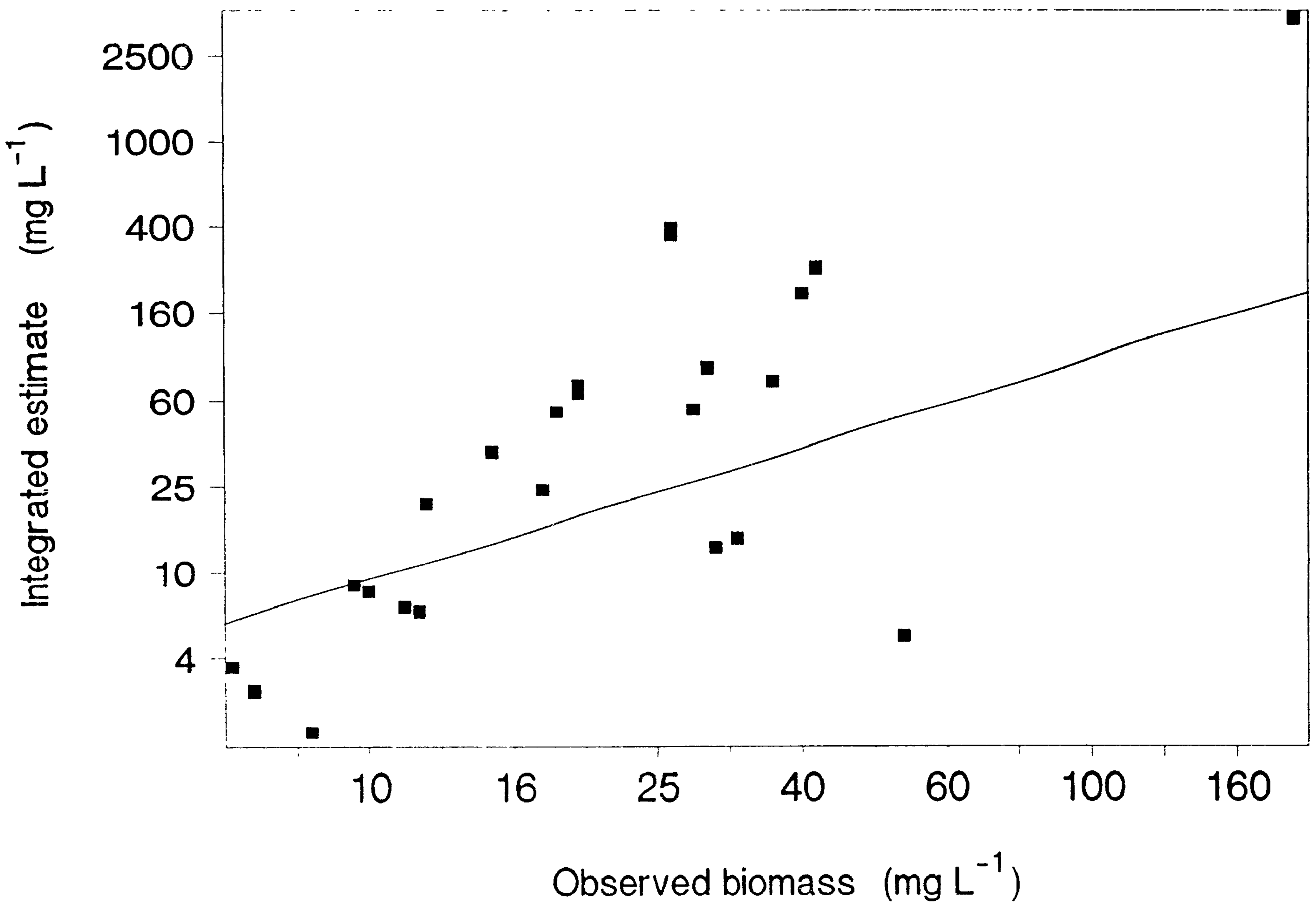
FIGURE 1.5. The mean normalized curve and the two extremes in this data set (n = 25).


Normalized density (L^{-1})
10^{9}
10^{7}
10^{5}
10^{3}
10^{1}
10^{-1}
L. Waterloo
$Y = 2.5x10^{8} X^{-0.75}$
Median
$Y = 2.6x10^{8} X^{-0.90}$
L. Stukely
$Y = 3.6x10^{7} X^{-0.98}$
0.8
3.1
12.5
50
200
800
Equivalent spherical diameter (μm)

in biomass distribution over logarithmic size classes. The slopes varied from -0.75 to -1.01, with a mean of -0.89 (95% C.I. of mean = -0.87 to -0.91), which is highly significantly shallower ($P<.001$) than the slope of a spectrum with approximately equal biomass per class. Thus the distribution is not uniform , but shows a small increase in biomass between successive size classes. The mean slope corresponds to a spectrum in which biomass tends to increase by about 7% between successive base 2 logarithmic size classes.

Over the range of sizes from 10^{-6} to 10^{3} μg, the approximate median parameters (6 and -0.9) generated an integrated biomass estimate of 7.6 mg L^{-1}, after correction for backtransformation from a logarithmic scale (Sprugel 1983). This was not significantly different from the median (14.1 mg L^{-1}) of observed biomasses. Observed biomasses varied from 6.4 to 193 mg L-1, whereas integrated estimates varied from 0.8 to 172 mg L^{-1}. The integrated estimates were also significantly correlated ($P<.001$, $F = 31$, $r^2 = 0.57$) with the observed total biomasses, where all values were log transformed (base 10) to stabilize variance (Figure 1.6). Although the parameters of this relationship (-3.36, 1.83) were significantly different from 0 and 1 respectively ($P<.025$), this regression establishes that the spectrum may be used as a basis for comparison of samples.

FIGURE 1.6. The relationship between observed total planktonic biomasses and the estimates obtained from integration of the normalized curves fitted to these data. The 1:1 line is shown.


2500
1000
400
160
60
25
10
4
Integrated estimate (mg L^{-1})
10
16
25
40
60
100
160
Observed biomass (mg L^{-1})

Sprules and Munawar (1986) reported normalized spectra for several Ontario lakes (Table 1.9). Their median slope values from replicated observations on individual lakes tended to be closer to -1 than ours. Their median intercepts (reported in base 10 logarithms) were smaller. Those authors did not provide sufficient information about the distribution of their values to undertake a parametric comparison, but numeric integration of their curves over the same range as ours yielded values in the range 0.7 to 0.8 mg L^{-1}, compared to 0.8 to 172 mg L^{-1} (median = 7.6) for our integrated totals. Their results spanned a smaller range since each spectrum was based on several samples, whereas we have computed one spectrum per sample. Some of our samples had much higher phosphorus concentrations than those of Sprules and Munawar. Nevertheless our normalized spectra usually reflect higher total plankton biovolumes than do theirs.

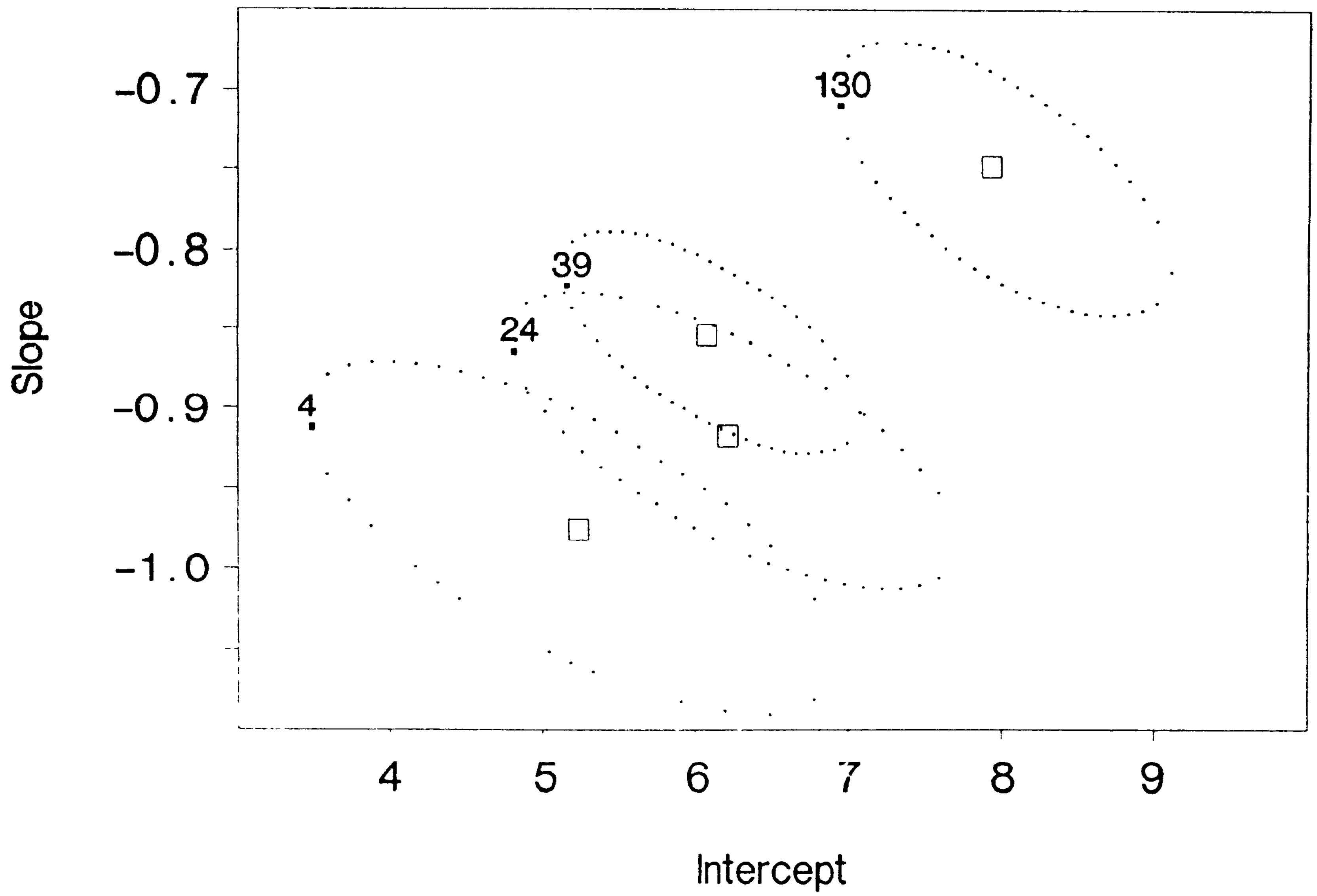
3. Variability of parameters

The elliptic confidence regions about the parameters of normalized spectra from 4 samples are illustrated in Figure 1.7. The figure shows the variability in the size of the confidence regions, as well as the extent to which different spectra may correspond. Thus ellipses with substantial overlap correspond to spectra which do not differ significantly, whereas disconnected ellipses are significantly different. The

Table 1.9. Parameters (range and median) of normalized spectra for Ontario lakes computed by Sprules and Munawar (1986). For purposes of comparison, the values for our spectra computed over a similar size range (n = 25) are included, after conversion to base 10 logarithms

Lake	Slope	Intercept
Inland lakes	-0.92 to -1.05 (-0.98)	0.92 to 1.93 (1.36)
St. Clair	-0.76 to -1.05 (-0.90)	1.11 to 1.83 (1.46)
Erie	-0.77 to -1.24 (-0.99)	1.79 to 2.41 (1.99)
Ontario	-0.90 to -1.04 (-0.97)	1.35 to 2.16 (1.82)
Huron	-0.90 to -1.18 (-1.02)	0.64 to 1.65 (0.90)
Superior	-1.00 to -1.15 (-1.10)	0.41 to 0.73 (0.61)
This study	-0.75 to -1.01 (-0.90)	1.70 to 3.89 (2.41)

FIGURE 1.7. 95% confidence regions around the parameters of some normalized curves. The total phosphorus concentrations are shown, in order to illustrate the general trend to decreasing slope and increasing intercept with increasing phosphorus concentration. Larger ellipses reflect less precise parameter estimates. Greater overlap between two ellipses reflects less significant difference between the distributions.


130
39
24
4
-0.7
-0.8
-0.9
-1.0
Slope
4
5
6
7
8
9
Intercept

mean squared errors (variance estimates) of the 25 best spectra were compared using Bartlett's test. The test statistic showed very significant variability (P<0.005) among the residuals. This variability is reflected in the relative sizes of the confidence ellipses.

The variability (among the spectra) of each of the parameters was tested with two pairwise tests. A t-test showed a significant difference (P<0.01) between the smallest and largest intercepts of the 25 normalized spectra. The first and the 24th intercepts were still different (P<0.05). The second and third were also different from the 25th. Otherwise, there were no significant differences among the intercepts. However, bivariate comparisons like this generally increase the number of Type I errors among all possible comparisons in a parameter set. Tukey's LSD test, which is more appropriate, showed no significant differences whatsoever among slopes or among intercepts of the normalized spectra. When Tukey's test was applied to the unimodal fits, four overlapping homogeneous sets of means were identified: 1 to 7, 5 to 10, 6 to 12, and 8 to 17 (where the numbers reflect the sequential order of the means from lowest to highest).

The low variability of the parameters of these two models reflects the underlying similarity of size spectra among lakes studied (Figure 1.5). This presents a problem for comparisons among lakes, and for the development of models to predict parameter differences between lakes. The sampling distribution from a single lake necessarily shows variability, due to short term changes and sampling error. But if the "true" difference between the parameters of two different systems is low, as we have found, then efforts to predict this difference will be confounded with sampling variation.

4. Predictions

These data are the largest available set describing the size distribution of plankton communities in lakes of differing trophy. This section develops correlations between the parameters of size distributions and other lake characteristics, using data collected in 1987 (Table 1.10). Where possible, the predictive power of these relations was tested with data collected in 1986. None of the relationships were significantly improved when the following morphometric variables were tested as additional regressors: lake area, lake volume, mean depth, maximum depth, and a categorical variable depending upon whether or not the lake stratifies in summer.

As a partial check on the validity of some of the data collected, and for a comparison with other published work, the relationship between mean summer chlorophyll concentrations and

Table 1 10 Predictive regressions for planktonic size distributions In general, the amount of explained variation in these analyses is moderate relative to other models (e.g Peters 1986), because the regressions use point estimates, rather than seasonal means. The F statistics do not include the intercept * = 0 05>P>0.01, ** = 0 01>P. Abbreviations are defined in the footnotes

Response	Intercept	Predictors	n	SE$_{est}$	R^2	F
$\log_{10}$ CHL	-1.22	1.29 $\log_{10}$ STP	15	0.202	0.53	15^{**}
$\log_{10}$ TVOL	3 27	0.85 $\log_{10}$ TP	25	0.237	0 55	28^{**}
UNIMODE	-6.37	2.58 $\log_{10}$ TP + 2.47 $\log_{10}$ CHL - 6.67 $\log_{10}$ DW	17	1.488	0.58	6^{**}
NSLOPE	-0.965	0.00351 TP	17	0.041	0.38	9^{**}
NINT	5 467	0 0200 TP	47	0 562	0 32	21^{**}
$\log_2$ N	5 756	- 0 793 $\log_2$ SIZE	1647	1.900	0 94	24624^{**}
$\log_2$ N	3.601	- 0 792 $\log_2$ SIZE + 0 533 $\log_2$ TP	1647	1.834	0.94	13266^{**}
$\log_{10}$ PICOVOL	2.85	0 38 $\log_{10}$ TP + 0.00047 (DATE x $\log_{10}$ DW)	58	0 151	0 45	23^{**}
$\log_{10}$ NANNOVOL	3 84	0.006 DATE + 0 499 $\log_{10}$ DW	55	0 270	0 45	22^{**}
ln NETVOL	9.07	- 13 97 ln TP + 24 08 lnPNET - 0 84 (ln DW x ln TP) - 0 038 (DATE x ln TP) - 0 078 (DATE x ln PNET) + 0 89 (ln DW x ln PNET)	17	0.259	0 96	35^{**}
$\log_{10}$ ZOOW	1 465	1 252 $\log_{10}$ TP	25	0 451	0 42	76^{**}
ln NANNO/NET	9 37	1 19 ln DW - 0 81 ln TP - 0 60 ln CHL	17	0 466	0.81	19^{**}

Footnotes overleaf

Table 1.10 footnotes:

TVOL = volume density of plankton (μm^3 x 10^{-6} L^{-1})
PICOVOL = volume density of plankton of ESD < 5 μm (μm^3 x 10^{-6} L^{-1})
NANNOVOL = volume density of plankton of 5 μm < ESD < 80 μm
NETVOL = volume density of plankton of ESD > 80 μm
TP = total epilimnetic phosphorus concentration (mg m^{-3})
DATE = Julian date
DW = epilimnetic dry mass concentration (μg L^{-1})
PNET = phosphorus concentration retained on 80 μm filter (mg m^{-3})
ZOODW = zooplankton dry mass concentration (μg L^{-1})
CHL = chlorophyll-a concentration (mg m^{-3})
N = numeric density of plankton in a size interval (x10^{-6} L^{-1})
SIZE = volume of a single organism (μm^3)
UNIMODE = mean of normal distribution of volume over $\log_2$ sizes
NSLOPE = slope of the normalized spectrum
NINT = intercept of the normalized spectrum
STP = total spring epilimnetic phosphorus concentration (mg m^{-3})

spring total phosphorus concentration was calculated. The relationship explained 53% (n = 15) of the variation in the base 10 log chlorophyll. Our parameters (-1.22, 1.29) were similar to those computed by Dillon and Rigler (1974) using a much larger data set (-1.14, 1.45), and the predictions from the two models were very highly correlated. Hence, there is reason to hypothesize that the relations built on data from these lakes may reflect more general regularities.

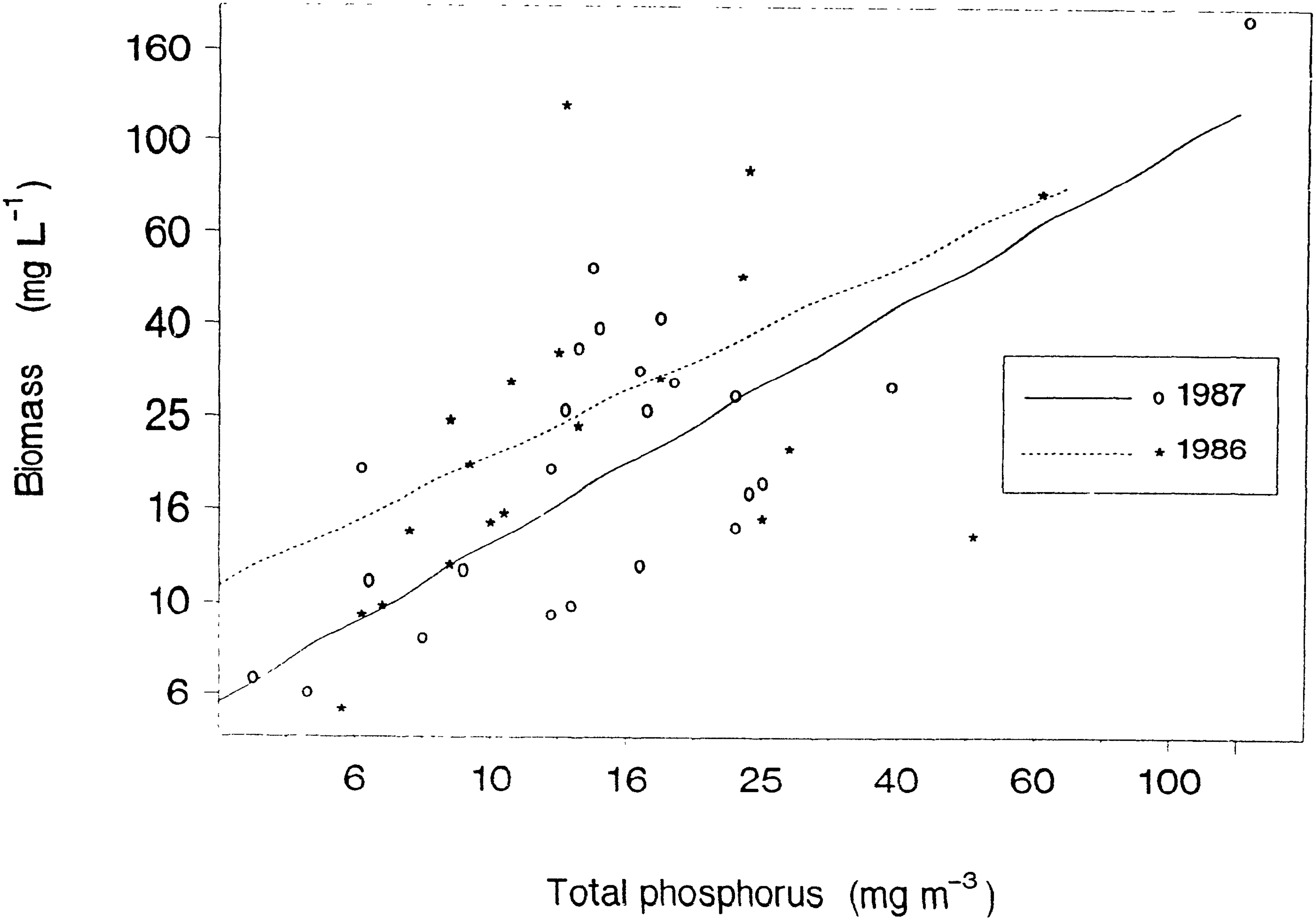
The total epilimnetic phosphorus concentration (TP) is the most useful and reliable predictor of size distribution parameters. As TP increases, the mean of the unimodal distribution increases, indicating the relatively greater number of larger organisms. This is further reflected in the shallower slope and higher intercept of the normalized spectrum. The unimodal mean is best predicted from the logs of the concentrations of TP, chlorophyll, and sestonic dry weight ($R^2 = 0.58$). The most reliable predictions of the normalized slope and intercept are based on TP alone ($R^2 = 0.38$ and 0.32 respectively).

Plankton abundance was very negatively correlated with size. Thus 94% of the base 2 log of the numerical density in a size interval was explained by a regression on the log of mean size in that interval (Table 1.10). Furthermore, this model based on 1987 data explained 82% of the variation in the independent 1986 abundance data. Incorporation of log TP as a predictor slightly increased the variation explained in the original data. This augmented model explained 87% of the variation in the independent data set.

Although only 55% of the variation in the logaritnm of the total planktonic biovolume is explained by a regression on the logarithm of the total phosphorus concentration, this relationship is not significantly improved by the inclusion of other likely regressors (e.g. chlorophyll) in the model. When applied to an independent 1986 data set, there was a significant correlation between predictions and observations. A separate regression between total volume and total phosphorus was fitted to the 1986 data. The mean squared errors of the two models were sufficiently similar (Bartlett's test, P>0.25) to allow a direct F test comparison of the two parameter sets, which were not significantly different from one another (Figure 1.8).

Regressions to predict the biomass of different functional groups of plankton explained less than half of the observed variation in the total volume of picoplankton, nannoplankton, and zooplankton as functions of Julian date and some index of trophy. Netplankton volume was more effectively described (R^2 = 0.96) by a seven parameter model, but because all terms in this relationship explained similar amounts of variation, all simpler models were less effective. For each of these models,

FIGURE 1.8. The relationship between observed total plankton biomass and total phosphorus concentration in 1986 and 1987. There is no significant difference between the two independently fitted relationships.



Biomass (mg L^{-1})
160
100
60
40
25
16
10
6
6
10
16
25
40
60
100
Total phosphorus (mg m^{-3})
o 1987
* 1986

the trend between model predictions and the original data did not differ significantly from a 1:1 relationship. However, in those cases (nannoplankton and zooplankton) where the predictions of relationships built on 1987 data were compared with data collected in 1986, there was significant deviation from a 1:1 trend.

The ratio nannoplankton/netplankton decreases as phosphorus and chlorophyll concentrations increase. This is consistent with the observed increase in average size as these trophic variables increase.

5. Use of Phosphorus as an Index of Biomass

Peters (1983b) found a roughly linear increase in particulate phosphorus concentration per logarithmic size class over 6 classes spanning roughly the range from 1 to 100 μm ESD. The slopes of these relationships increased with total phosphorus concentration. If the relationships are in fact linear, then the implication is that the phosphorus concentration per size class is constant among classes in the same lake. Increases in slopes and intercepts with total phosphorus imply that each size class contains more phosphorus. Peters' interpretation of these phosphorus distributions as biomass distributions relies upon the assumption that the ratio of phosphorus to biomass does not change with either size or total phosphorus concentration. In order to test this assumption, we conducted a two-way analysis of variance of the log-transformed mean phosphorus/biomass ratios for 3 size classes (0.4-5, 5-80, >80 μm) in 4 ranges of total phosphorus concentrations (<9, 9-21, 21-52, >52 mg m^{-3}). These intervals correspond to intervals of equal range in log-transformed phosphorus concentration values. The model explained 71% of the variation in the data (n = 121, F = 39.2, $P<0.001$, SEE = 0.1584). The predicted mean values (Table 1.11) show a trend to decreasing ratios with increasing size ($P<0.01$) and increasing ratios with increasing total phosphorus concentration ($P>0.05$; n.s.). The mean values in the table suggest that the phosphorus concentration in picoplankton is about 8 times higher than that in netplankton. Assuming that dry weight is about 30% of wet weight (Peters & Downing 1983), the mean P content of algae from Vinogradov's (1953) data (0.69% of dry mass) converts to a wet phosphorus/biomass ratio of 0.0023. This value is intermediate between those of the small and medium size classes at all phosphorus levels predicted from our analysis. Peters (1983b) noted that Vinogradov's (1953) data showed some tendency to declining phosphorus concentration with increasing organism size, and Shuter (1978) showed that the phosphorus concentration in phytoplankton cells declines with cell size. Our analysis shows a similar trend. Thus the phosphorus content of a size class is at best an approximate index of biomass, which should be adjusted by a size-specific correction

Table 1 11 Results of the two-way analysis of variance of the phosphorus/biomass ratio. Columns represent three size classes, and rows represent four levels of total phosphorus concentration. The four values in each cell are the number of observations, the mean predicted $\log_{10}$ ratio, the standard error of this estimate, and the back-transformed ratio corrected for transformational bias (Sprugel 1983)

	Picoplankton	Nannoplankton	Netplankton
TP<9	8	7	5
	-2 77	-3.63	-3 68
	0.0992	0.1020	0.1140
	0.0026	0.0004	0.0003
9<TP<21	35	24	11
	-2.63	-3.49	-3 54
	0.0599	0.0693	0 0925
	0 0036	0.0005	0 0004
21<TP<52	12	9	5
	-2.45	-3.30	-3 36
	0.0873	0.0927	0 1090
	0.0054	0 0008	0.0007
TP>52	3	1	1
	-2.39	-3.25	-3 30
	0 1800	0.1880	0 1940
	0 0062	0 0009	0 0008

factor.

The size dependence of the phosphorus/biomass ratio has important implications for the use of phosphorus as an index of biomass. First, differences in total phosphorus concentration tend to be greater than differences in biomass. This is reflected by the fractional exponent (0.85) of the regression between total biovolume and total phosphorus (Table 1.10). Second, the high concentration in the smallest size class overestimates small organism biomass compared to that of larger organisms. Hence, the linear trends in the relationship between cumulative phosphorus and organism size reported by Peters and the decline in phosphorus concentration with size imply an increase in biomass with size. This confirms the tendency indicated by our results from the normalized spectrum.

6. Extrapolations

The mean slope and intercept of the normalized spectrum may be used to predict the abundance of organisms at larger sizes. The projected 7% increase in biomass per log 2 size class is necessarily constrained by the more rapid increase in organism size (doubling between classes) and the maximum size of the habitat. Our largest lake (Memphremagog) has a volume of about 1.5 km^3. If at least 500 individuals are required for a subsistence population, then the rarest species in Lake

Memphremagog might have as few as 330 km^{-3}. The model projects an abundance this low for the interval containing organisms of about 4 metres ESD (18 t). If we apply the additional constraint that the rarest organism has a mass lower than 30 kg (corresponding to a large lake sturgeon), then the largest size interval spans the range 18 to 36 kg. The geometric mean weight of fish in this size range is 25 kg. Using our mean parameters from the normalized spectrum, we predict a total abundance of about 150,000 organisms of this size km^{-3}, or over 200,000 in a lake the size of Lake Memphremagog. This seems unreasonable, but the 95% confidence interval is very broad (between 12,000 and 1,880,000). In light of these unrealistic predictions of the abundance of larger organisms, it is convenient to invoke the rule that predictions beyond the range of the variables used to build a regression are statistically unjustified. Nevertheless, the implication that the size spectrum does not have a constant trend throughout the range of all aquatic organisms needs to be examined. Current theories of energy balance in size-structured aquatic systems (e.g. Platt and Denman 1978; Platt and Silvert 1981) rely essentially upon pelagic analyses. However, beyond the size range of organisms included in our models, interactions with the hypolimnion and the benthos become important. For example, larger zooplankton and fish may conduct diurnal vertical migrations, and some fish spend part of their life cycle near the bottom and part as pelagic feeders. In oceanic systems, the epilimnion and the benthos are relatively isolated. But in most lakes, the much smaller spatial scales produce more frequent interactions, at least for larger organisms. Hence, accurate predictive models for fish abundance and production may need to build upon the size distributions of both the benthos and the plankton (cf. Sprules & Munawar 1986).

CHAPTER II

PLANKTON COMMUNITY RESPIRATION: RELATIONSHIPS WITH SIZE DISTRIBUTION AND LAKE TROPHY

ABSTRACT

We tested the hypothesis that limnoplankton community respiration rates are predictable functions of mean body size, without reference to taxonomic composition. Over a broad range of trophic conditions ($6.5 \leq [TP] \leq 130\ \mu g\ L^{-1}$; $1.2 \leq [chl\text{-}a] \leq 29\ \mu g\ L^{-1}$), the mean respiration rate per organism for picoplankton, nannoplankton, and netplankton assemblages is a power function of mean organism size, with an exponent of 0.73. The total epilimnetic phosphorus concentration is correlated with both the biovolume and respiration rate of the plankton community, as well as with the respiration rates of the three plankton size classes. The summed respiration rates of the three screened fractions are not significantly different from the mean total community rate. When respiration (R) and biovolume (B) are standardized to equivalent carbon units, the R/B ratio is a power function of mean organism size, with an exponent of -0.30. These results provide empirical support for the contention that size distributions may be used to construct comprehensive models of community physiology.

Introduction

The respiration rate is a fundamental measure of biological activity for all organisms. It is correlated with other physiological rates, such as growth rate (Banse 1976), and the rates of excretion of ammonia and phosphorus (Ikeda 1985). Since longterm production is proportional to respiration for a wide range of populations (Humphreys 1979), the respiration rate of an entire community might well provide an easily-measured index of production in the community. Furthermore, the bioaccumulation of contaminants by aquatic organisms is affected by the metabolic rate (Neely 1979), and contaminant fluxes show size effects (Jorgensen 1979).

Direct measurement of plankton community respiration rates must be done in situ or very rapidly after removal of the sample from the lake (Lampert 1984). In addition to being time-consuming, the procedures themselves may introduce bias into the measurements. Not all procedures measure the same thing: whereas measurement of the activity of the electron transport system (ETS) represents the maximum potential oxygen demand of the community, the measurement of oxygen uptake represents the extent to which this potential is being realized (Packard 1971, 1985; Devol 1975). In spite of these methodological biases and inconsistencies, some general trends may be identified: Plankton community respiration rates tend to increase with biomass, although assemblages of larger organisms respire at a lower rate than similar biomasses of smaller organisms (Williams 1984). But at present there are no models to predict the respiration rate of planktonic communities, although two approaches hold promise.

One approach would extend the existing knowledge of community responses to trophic variables. A number of relationships exist, describing both the static and dynamic properties of plankton communities as functions of the total phosphorus concentration (see Peters 1986 for a review). For example, the relationships between phosphorus concentration and photosynthetic production (Smith 1979; Gelin and Ripl 1978; Elser et al. 1986), zooplankton abundance (Pace 1986), and relative abundances of netplankton and nannoplankton (Kalff and Knoechel 1978) are known. If the trophic response of respiration rate were known, then one could predict total community respiration from the phosphorus concentration. The second approach is an extension of the known size-dependence of respiration rates of individual taxa, observed in laboratory studies. It would provide a test of the assumption (e.g. Griesbach et al. 1982) that relationships observed on isolated taxa also apply in the field. Rather than concentrating on the allometric response of individual organisms, one could seek size-based trends within whole communities. Thus one could measure the respiration rate of plankton assemblages delimited by size (rather than by taxon), in an effort to determine the size dependence of the community rate. Since the exponents of taxonomically-specific relationships tend to be similar (Lavigne 1982; Banse 1979), one might hypothesize that community regressions would follow a similar trend.

In this paper, we contribute to models of community physiology based upon the community size structure and a trophic variable (the total phosphorus concentration). We examine the hypotheses that respiration rates are predictable functions of the total epilimnetic phosphorus concentration, and of the abundance and size of plankton.

Materials and Methods

Sampling -- Samples were collected during July and August, 1987 from 13 lake sites in southern Quebec. These sites represented a broad range of trophic conditions, as evidenced by phosphorus and chlorophyll concentrations (Table 2.1). Integrated epilimnetic samples were collected during the day through a 2.5 cm. diameter tube connected to a piston pump. This apparatus has been shown to collect both phytoplankton and zooplankton effectively (Pace 1986). Water samples were stored in dark brown Nalgene bottles in a cooler containing ice packs during transit to the laboratory (maximum duration of one hour).

Table 2.1. Concentrations of phosphorus (mg m^{-3}) and chlorophyll (mg m^{-3}) in integrated epilimnetic samples used for respiration observations. Newport, North, Central, and South are basins of Lake Memphremagog. Concentrations refer to the total filtrate of the indicated filter, including both the filtrate and retentate of smaller filters.

	Porosity of screen or filter (μm)						
	0.4	5		40		TOTAL	
Lake & Julian date	P	P	Chl	P	Chl	P	Chl
Orford 237	3.7	5.3	0.6	5.3	1.3	6.5	1.6
Stukely 230	4.1	5.4	0.3	5.7	1.4	6.6	2.2
Orford 203	3.7	5.7	-	6.6	-	7.9	-
Lyster 227	4.4	8.4	0.3	7.6	1.1	9.1	1.6
North 226	7.9	9.7	0.5	10.3	0.8	12.2	1.6
Baldwin 225	7.2	10.8	0.7	12.0	1.0	13.1	2.0
Central 228	7.8	10.3	0.0	11.7	0.2	13.5	1.2
Cerises 236	8.0	12.2	-	15.5	4.2	16.6	6.3
South 235	6.9	11.9	1.0	12.8	1.8	18.6	3.5
Newport 204	11.7	16.5	-	17.9	-	22.7	4.1
Pond 224	15.8	20.1	1.4	21.8	1.6	22.9	3.8
Magog 223	14.5	24.1	0.8	29.5	5.5	39.3	9.0
Waterloo 229	19.7	34.6	4.5	57.3	16.2	130.1	28.9

Oxygen uptake -- For oxygen uptake experiments, samples were divided upon arrival at the lakeside laboratory into picoplankton, nannoplankton, and netplankton, using 40 μm square mesh stainless steel screens and 5 μm Nuclepore filters. In order to obtain measurable and reliable rates from individual size fractions, the subsamples were then concentrated on glass fibre filters (0.45 μm nominal porosity). Cornett and Rigler (1986) have shown that this concentration procedure does not significantly alter the rate of oxygen consumption of seston samples. For each subsample except one, a total of one litre of water was used (895 ml filtered through the glass fibre filter and 105 ml added to the incubation bottle). The remaining sample, from a highly eutrophic lake, received only 400 ml of source water. The fractionated subsamples, along with whole community samples, were then incubated in the dark for twenty four hours at in situ temperatures (18 - 22°C). The whole community sample provided a check on the accuracy of the fractionation procedure. Initial and final oxygen concentrations were determined using the sodium azide modification of the Winkler technique (APHA 1971).

Size Composition -- The size distribution of plankton throughout the range 0.2 to 1500 μm equivalent spherical diameter (ESD) was determined by direct microscopic examination. The abundance and approximate diameters of bacteria were determined by epifluorescence microscopy with DAPI stain (Porter and Feig 1980) of samples preserved in 2% formaldehyde. Larger organisms in the picoplankton fraction were measured at 1250X on an inverted microscope, using samples preserved in Lugol's iodine solution. Nannoplankton were measured at 1000X and 400X (Lund et al. 1958). Netplankton were measured at 100X on the inverted microscope and at 40X under a dissecting microscope, using both Lugol's samples and others preserved in 2% formalin. Individual volumes of organisms larger than 5 μm greatest axial linear dimension (GALD) were estimated by measuring length and width and taking the volume of similar regular geometric shapes as approximations. Smaller organisms were counted in nine diameter intervals between 0.2 and 5 μm. In each of the three size fractions (pico-, nanno-, and netplankton), at least 400 organisms were counted and measured.

Chemical analyses -- Epilimnetic phosphorus concentrations were measured as an indicator of lake trophic state. The total phosphorus concentrations in each size fraction and in the total sample were determined in triplicate, using the ascorbic acid modification of the molybdenum blue technique (Strickland and Parsons 1968) after digestion with potassium persulfate under pressure (Menzel and Corwin 1965). As further confirmation of trophic differences between lakes, chlorophyll a concentrations were also measured in triplicate (Strickland

and Parsons 1968), and corrected for phaeophytin.

Data analyses -- The biovolume (ppm) in each size class was defined as the sum of the volumes of all organisms in the size interval (1 ppm = 10^9 μm^3 L^{-1}). Individual organism volumes are expressed in μm^3. When conversions among units were required, all organisms were assumed to have a density of 1 $g.cm^{-3}$, implying that 10^6 μm^3 of biovolume (.001 ppm) is equivalent to 1 μg of biomass; 1 μg of oxygen respired is equivalent to 0.375 μg of carbon (Parsons et al. 1984); picoplankton were assumed to contain 0.0963 pg μm^{-3} of carbon (Simon 1987); nannoplankton and netplankton volumes were converted to carbon equivalents using the empirical formula for phytoplankton from Mullin et al. (1966): $\log_{10} C = -0.29 + 0.76 \log_{10} V$.

Results and discussion

Respiration rates ranged from 94 mg to 1.3 g O_2 m^{-3} d^{-1} (Table 2.2). The highest rate is similar to the 1.2 g O_2 m^{-3} d^{-1} measured in natural Anabaena collections (Gessner & Pannier 1958), but lower than the 6.8 g O_2 m^{-3} d^{-1} in a fertilized Georgia pond (Welch 1968). In mesotrophic Lake Washington (summer chlorophyll 5 mg m^{-3}), Devol and Packard (1978) found a summer average respiration rate of 180 mg O_2 m^{-3} d^{-1}, similar to our intermediate values. In Findlay Lake (Devol 1979), the

Table 2.2 Rates of oxygen uptake (mg O_2 m^{-3} d^{-1}) in dark-bottle incubated samples. SE = standard error of mean (n = 3).

Lake & Julian date	Size Fraction						Sum of Fractions	Measured Total	
	<0.5 µm		5-40 µm		>40 µm				
	mean	SE	mean	SE	mean	SE		mean	SE
Orford 237	74	5.3	1	0.5	20	2.0	95	109	5.1
Stukely 230	18	3.5	88	21.5	1	0.9	106	94	9.4
Orford 203	42	12.2	1	0.5	7	1.6	50	59	5.8
Lyster 227	105	8.6	7	2.6	70	4.4	182	157	13.0
North 226	126	10.8	21	5.2	32	3.0	179	194	8.4
Baldwin 225	60	13.8	14	7.0	126	8.6	200	187	6.3
Central 228	207	13.8	14	3.5	60	4.7	280	295	24.4
Cerises 236	130	5.6	53	15.7	11	1.6	193	183	15.0
South 235	47	9.8	40	7.7	42	13.8	130	155	12.2
Newport 204	182	9.3	70	11.1	14	3.2	266	234	10.4
Pond 224	228	13.4	35	16.1	98	9.7	361	320	20.8
Magog 223	123	3.9	70	6.8	189	10.8	382	408	16.9
Waterloo 229	819	19.2	459	37.7	158	10.1	1435	1287	124.7

depth-averaged maximum annual respiration rate of 20 mg O_2 m^{-3} d^{-1} was lower than our measurements. But this lake was more oligotrophic (summer PO_4-P = 1 mg m^{-3}) than any of ours (Table 2.1). Hence our respiration measurements are consistent with previous results.

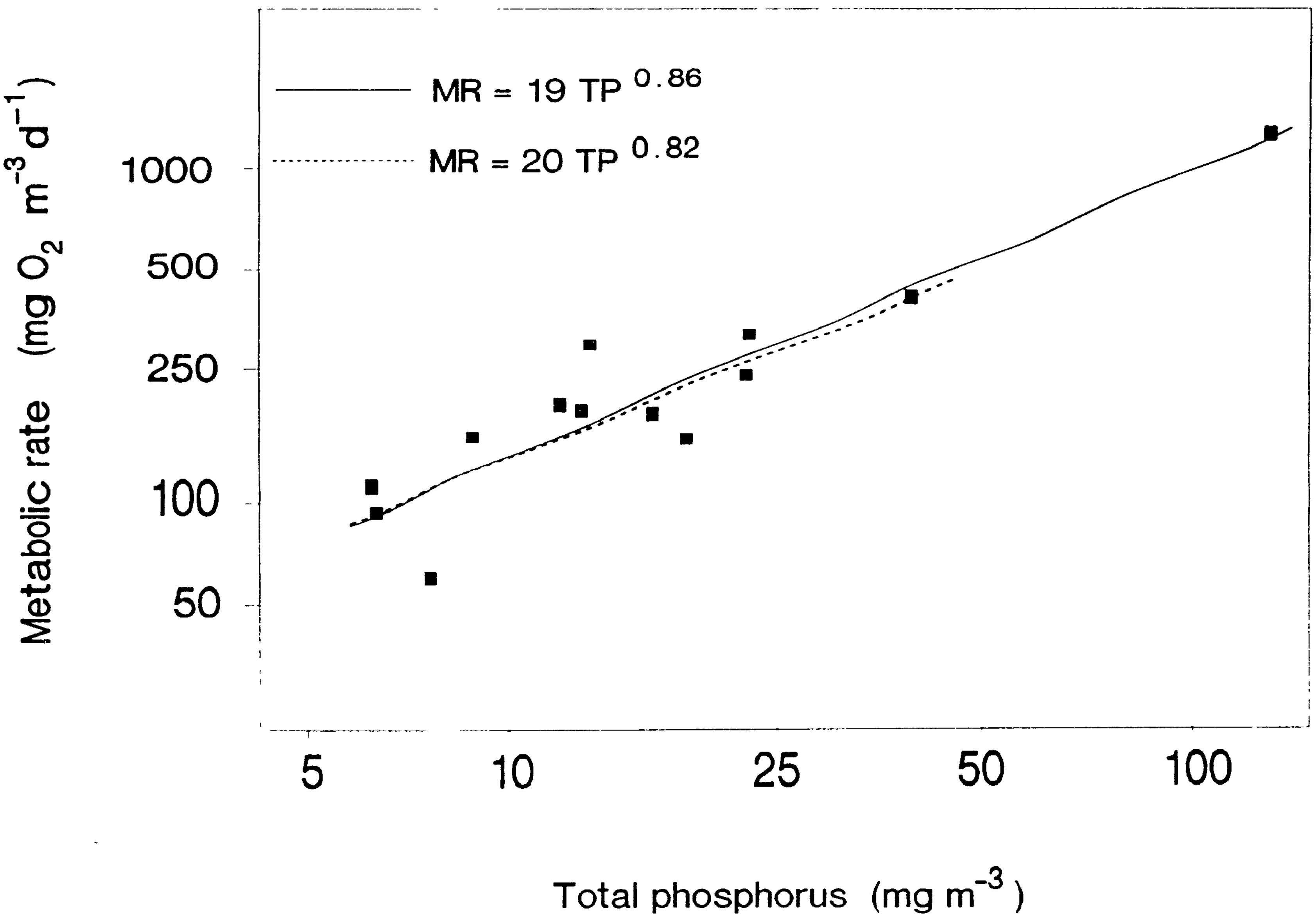
Relationships with phosphorus:

Over 85% of the variation in the log transformed total epilimnetic respiration rate is explained by variation in the total phosphorus concentration. Potential bias due to the high value for Lake Waterloo was verified by calculating a separate relationship for the remaining points. The resulting parameters (3.017, 0.824) were not significantly different (P<0.01) from those for all thirteen points (Figure 2.1). Total phosphorus (Table 2.1) also predicts a significant portion of the variation in individual size class respiration rates (Table 2.2; Figure 2.2).

Phosphorus is strongly correlated with the chlorophyll concentration, an indicator of algal biomass (Table 2.3). The parameters of this relationship between base 10 logarithms (-0.72, 1.00) describe a line intermediate between those for Florida lakes (-0.15, 0.74; Canfield 1983) and for spring turnover phosphorus versus summer mean chlorophyll (-1.14, 1.45; Dillon & Rigler 1974). Our parameters are individually

FIGURE 2.1. The relationship between the total epilimnetic phosphorus concentration and the plankton community metabolic rate. The regression lines with and without Lake Waterloo (n = 12 and 13) are not statistically different from one another ($P>0.05$).

FIGURE 2.2. The relationships between the total epilimnetic phosphorus concentration and the respiration rates of three plankton size classes. In each case, n = 13.


MR = 19 TP ^0.86
MR = 20 TP ^0.82
1000
500
250
100
50
Metabolic rate (mg O2 m-3 d-1)
5
10
25
50
100
Total phosphorus (mg m-3)

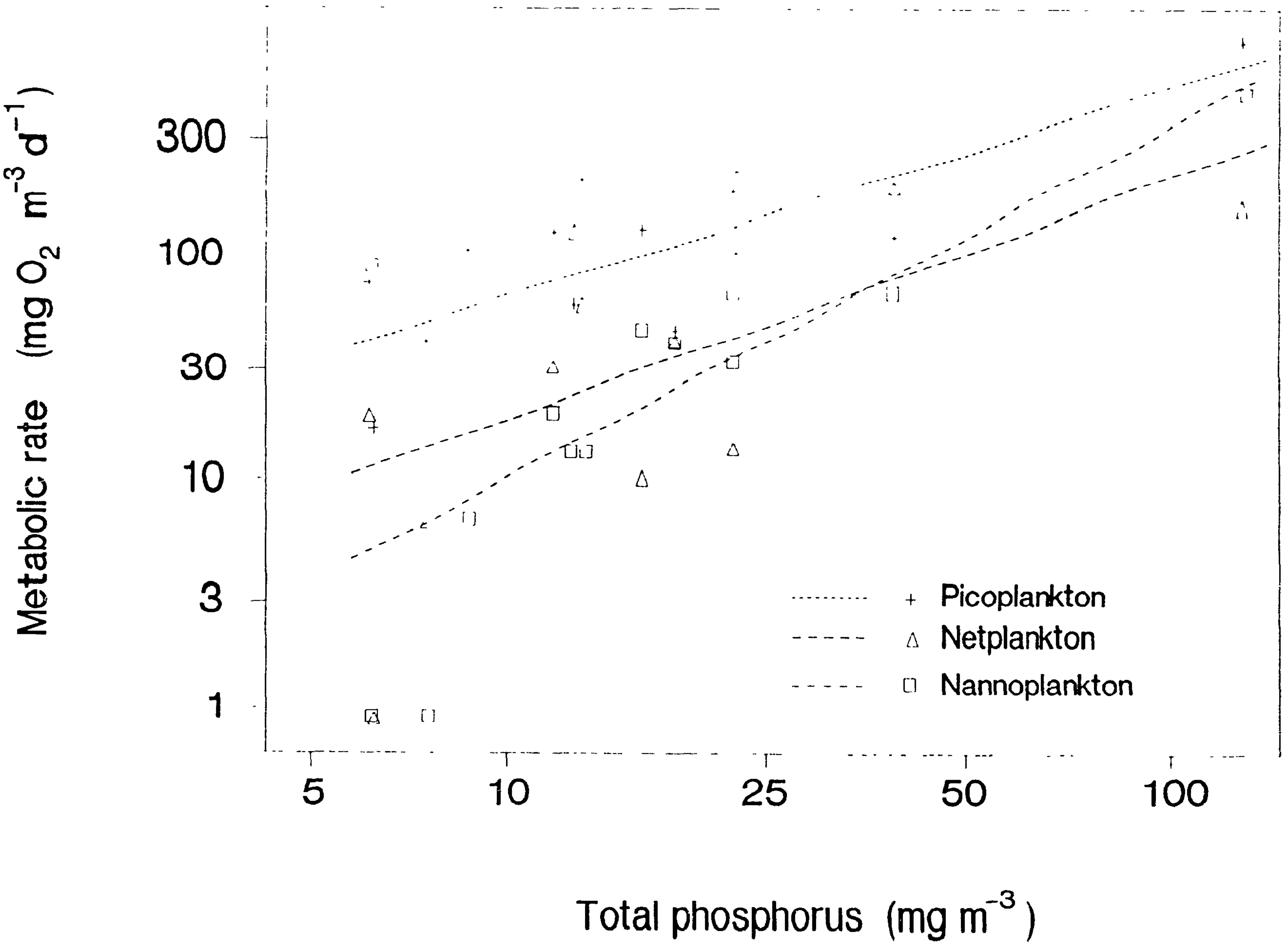

Metabolic rate (mg O_2 m^{-3} d^{-1})
300
100
30
10
3
1
5
10
25
50
100
Total phosphorus (mg m^{-3})
Picoplankton
Netplankton
Nannoplankton

Table 2.3. Parameters of the Model I predictive regressions between logarithms of the variables measured. Natural logs were used in every case, except #4, where base 2 was used. In every case $P < .005$, except #10, where $P < .05$. Abbreviations are defined in the footnotes.

	Variables Dep.	Ind.	Intercept b_0	SE_{b0}	Slope b_1	SE_{b1}	SE_{est}	R^2	F	n
1.	SR	TR	-0.31	0.242	1.059	0.045	0.119	0.980	551	13
2.	MR	MV	-16.50	0.324	0.728	0.041	1.37	0.894	313	39
3.	TR	SV	3.34	0.529	0.754	0.194	0.515	0.579	15	13
4.	N	MV	5.756	0.058	-0.793	0.005	1.900	0.937	24624	1647
5.	R/B	MC	-1.871	0.262	-0.295	0.041	1.364	0.589	53	39
6.	TR	TP	2.932	0.310	0.858	0.107	0.304	0.854	64	13
7.	SV	TP	0.700	0.553	0.692	0.191	0.541	0.544	13	13
8.	R_s	TP	2.133	0.629	0.909	0.217	0.616	0.614	18	13
9.	R_m	TP	-1.278	1.230	1.579	0.425	1.204	0.557	14	13
10.	R_l	TP	0.423	1.260	1.095	0.435	1.233	0.365	6	13
11.	MP	MC	-4.269	0.194	0.795	0.030	1.010	0.950	703	39
12.	CH	TP	-1.650	0.447	1.003	0.151	0.171	0.815	44	12
13.	PR	TP	-0.729	0.287	1.694	0.084	0.665	0.863	410	67
14.	PR	TR	-2.687	1.196	1.537	0.222	0.626	0.81	48	13

SR = summed respiration rate of all three size classes (mg O_2 m^{-3} d^{-1})
TR = total community respiration rate (mg O_2 m^{-3} d^{-1})
R_s, R_m, R_l = respiration rates of small, medium, and large size classes (mg O_2 m^{-3} d^{-1})
SV = summed volume of all three size classes (ppm)
MR = mean respiration rate of an individual organism (µg O_2 d^{-1})
MV = mean volume of an individual organism (μm^3)
TP = total epilimnetic phosphorus concentration (mg m^{-3})
N = density of organisms in a size class ($x10^{-3}$ mL^{-1})
MP = mean phosphorus content per organism (pg)
MC = mean carbon content per organism (pg)
R/B= ratio between size class respiration and size class biomass (d^{-1})
CH = epilimnetic chlorophyll-a concentration (mg m^{-3})
PR = estimated primary production (mg C m^{-3} d^{-1})

not significantly different from those calculated for the same lakes by Pace (-0.53, 1.05; 1984), although their joint distribution is significantly different ($P<0.01$).

The total phosphorus concentration and the total biovolume are also highly significantly correlated, as are total respiration and total biovolume (Table 2.3). Some coefficients of determination in Table 2.3 are lower than those reported elsewhere for similar relationships (e.g. Peters 1986) because they are based on point values rather than seasonal means.

The trophic response of total respiration may be compared with that of total production. It is well known that production is higher in lakes with higher phosphorus concentrations (e.g. Gelin and Ripl 1978, Elser et al. 1986). Smith (1979) showed that growing season mean primary production (mg C m^{-3} d^{-1}) is linearly related to the total phosphorus concentration. Transforming his data for a direct comparison with our results, we found the following relationship between mean production and total phosphorus concentration:

$$\ln PROD = -0.729 + 1.694 \ln TP$$

The parameters of this relationship may be compared with those (1.951, 0.858) for the trophic response of respiration (mg C m^{-3} d^{-1}). At the lowest observed level of phosphorus (4 mg m^{-3}), production is 5 mg C m^{-3} d^{-1}, whereas the predicted respiration rate is 23 mg C m^{-3} d^{-1}. But as phosphorus increases, production increases faster than respiration. Production equals respiration at about 25 mg TP m^{-3}. In eutrophic systems beyond this point production exceeds respiration. Only 3 of our 13 lakes exceeded this threshold.

The ratio between production and respiration is a characteristic of the ecosystem, and is not size-dependent (Humphreys 1979; Banse & Mosher 1980; Schwinghamer et al. 1986). Hence a regression of production vs respiration over systems with roughly the same ratio should have a slope not significantly different from unity. A steeper slope would indicate a trend to an increasing ratio within the sample. The comparison in the preceding paragraph suggests that the production/respiration ratio is not constant in our data set. A regression between respiration and estimated production reveals an increasing trend in the P:R ratio:

$$\ln PROD = -2.687 + 1.537 \ln RESP$$

The slope is significantly greater than unity ($P>0.05$).

Because these comparisons involve average trends from quite different water bodies, the discrepancy between production and respiration cannot be interpreted very closely. Nevertheless, the magnitude of the P:R ratio in oligotrophy (0.22) requires some explanation. Because the respiration rates are consistent with other estimates for plankton respiration and with allometric estimates of respiration for the different size classes (see below), one cannot easily dismiss the low P:R ratios as an artefact. Instead, the low ratio suggests some source of fixed carbon other than primary production. In oligotrophic Lake Almind, Denmark, bacterial uptake of dissolved organic carbon (DOC) accounted for 75% of daily carbon fixation (Sondergaard et al. 1988). A similar supplement would restore the carbon balance in our oligotrophic lakes.

The DOC pool in these lakes, calculated from lake colour, varies between 2 and 4 mg C L^{-1} (Rasmussen, unpublished). These levels are larger than those estimated for Lake Almind, and if similar mechanisms are at work, they would be sufficient to explain the excess of respiration over production in some of our lakes.

If we are correct in suggesting that DOC plays so important a role in oxygen metabolism in the surface waters of oligotrophic lakes, then these lakes are dependent upon energy subsidies from the watershed. As lakes become progressively more eutrophic, this subsidy becomes less important. As a result, metabolism rises more slowly than primary production with eutrophy. In eutrophic lakes, here as elsewhere, production exceeds respiration.

Allometric relationships:

In order to test the applicability of allometric relationships based upon laboratory studies to our field samples, we needed to estimate mean organism size and corresponding mean individual respiration rates. The mean organism size in each size fraction was estimated by dividing the class biovolume by the number of organisms (Table 2.4). The individual respiration rates were similarly calculated by dividing the respiration rate of the size class by the number of organisms in the class. This procedure is valid only if the sum of the size class rates approximates the observed community total. Table 2.2 shows that this is so. A regression between the mean sizes and the mean respiration rates per organism measures the allometric response of respiration to body size in mixed natural plankton communities. This relationship (Figure 2.3) is highly significant ($R^2 = 0.89$, $P<.005$), and has a slope of 0.73.

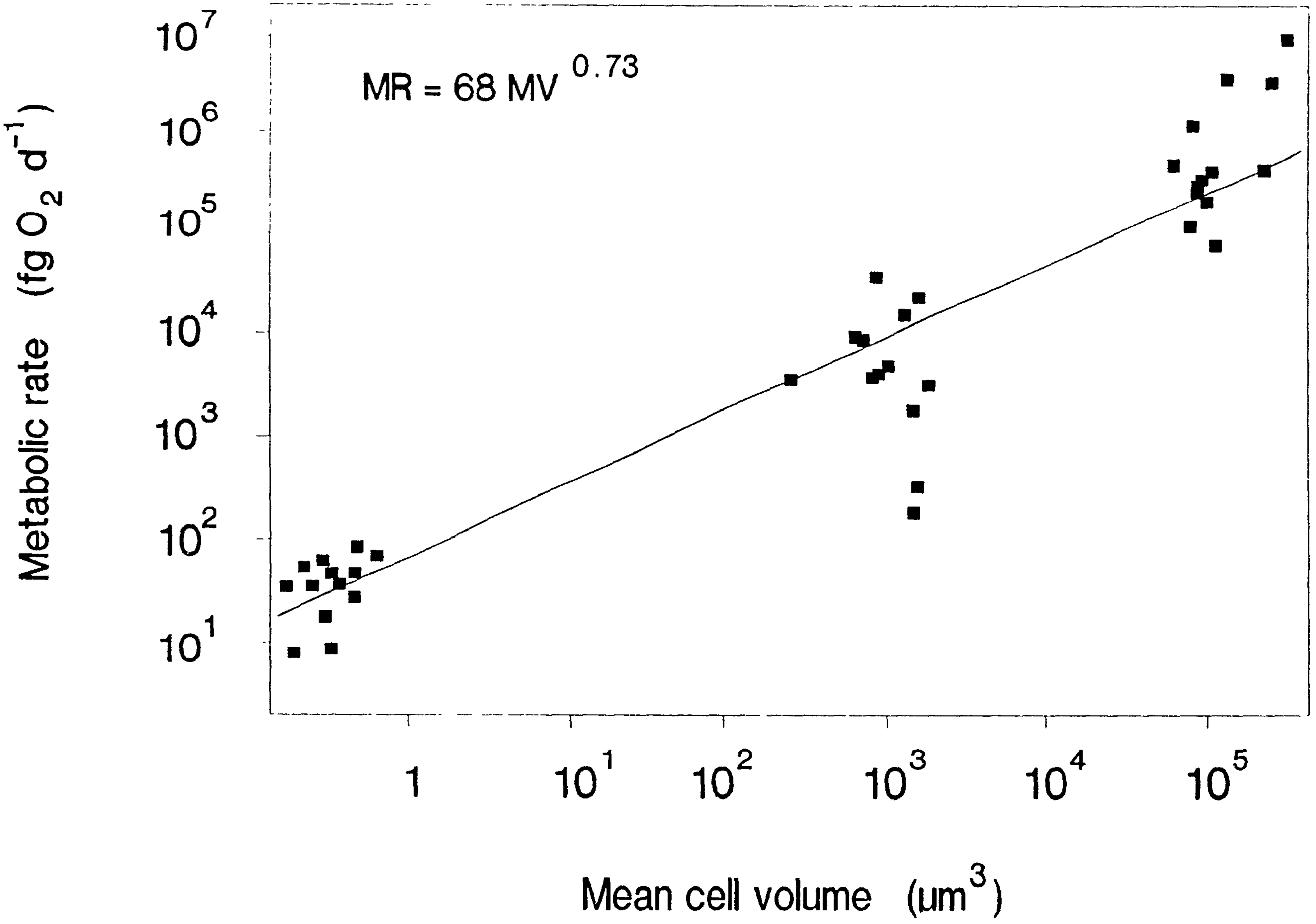
Several relationships between metabolic rate and body size have been determined for planktonic organisms (Table 2.5, Figure 2.4). These relationships differ methodologically from the regression calculated in this paper in that cultured organisms from specific taxa were used. We used natural communities divided into three size classes, measuring the true mean size of the organisms in each class for each sample. The

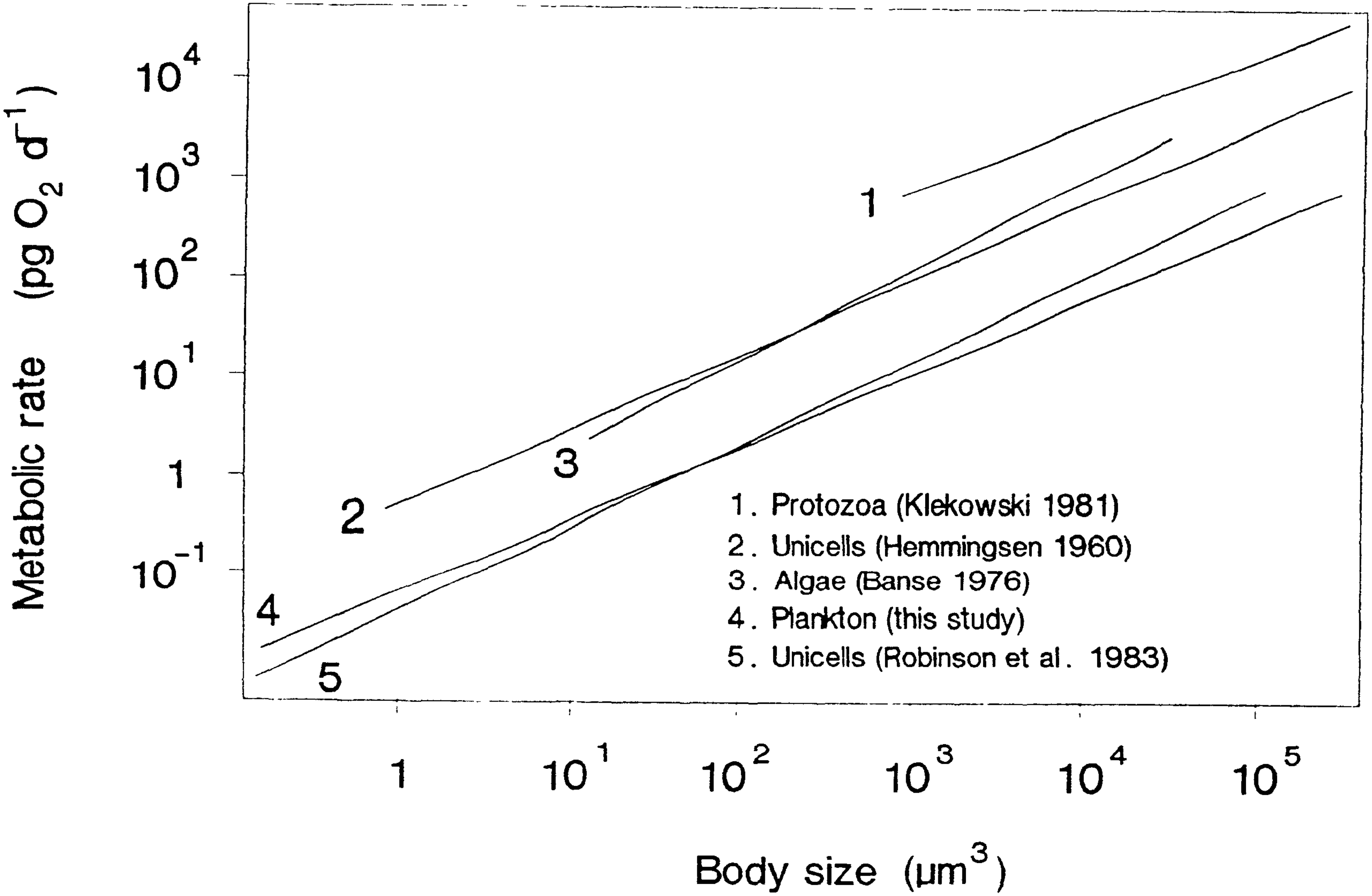
Table 2.4. Numbers, mean sizes, and total volumes of organisms in the three size classes. MV = mean volume (μm^3). Volume in ppm = $\mu m^3 \times 10^{-6}$ mL^{-1}.

		<0.5 μm ESD			5 - 40 μm ESD			>40 μm ESD			
Lake Julian	& date	N ml^{-1} ($x10^{-3}$)	MV	Volume (ppm)	N ml^{-1}	MV	Volume (ppm)	N ml^{-1}	MV	Volume (ppm)	Total (ppm)
Orford	237	1522	0.46	0.70	5220	1460	7.62	70	89282	6.25	14.6
Stukely	230	2030	0.33	0.66	5520	1295	7.15	12	115737	1.39	9.2
Orford	203	1516	0.46	0.70	3000	1550	4.65	18	95544	1.72	7.1
Lyster	227	1434	0.64	0.91	3840	1456	5.59	57	84006	4.79	11.3
North	226	2615	0.33	0.85	5350	813	4.35	60	64145	3.85	9.1
Baldwin	225	3302	0.30	0.97	2710	1022	2.77	15	324450	4.87	8.6
Central	228	5858	0.17	1.00	4260	1822	7.76	129	110870	14.31	23.1
Cerises	236	3474	0.37	1.30	5680	711	4.04	49	103040	5.05	10.4
South	235	5794	0.19	1.08	9300	887	8.25	125	89884	11.24	20.6
Newport	204	2764	0.29	0.81	1780	860	1.53	29	234445	6.82	9.2
Pond	224	4063	0.22	0.89	9360	255	2.39	29	258342	7.51	10.8
Magog	223	3438	0.25	0.87	7000	641	4.49	52	136687	7.12	12.5
Waterloo	229	9143	0.48	4.43	18890	1589	30.02	1259	81541	102.7	137.1

FIGURE 2.3. The relationship between the mean size of plankton in screen-separated assemblages and the mean metabolic rate per organism.

FIGURE 2.4. A comparison of five regressions between body size and metabolic rate for planktonic organisms. Only regression #4 (this study) is based on natural community samples separated by size only.


MR = 68 MV $^{0.73}$
Metabolic rate (fg O_2 d^{-1})
10^7
10^6
10^5
10^4
10^3
10^2
10^1
1
10^1
10^2
10^3
10^4
10^5
Mean cell volume (um^3)



Metabolic rate (pg O_2 d^{-1})
10^4
10^3
10^2
10^1
1
10^{-1}
1
2
3
4
5
1. Protozoa (Klekowski 1981)
2. Unicells (Hemmingsen 1960)
3. Algae (Banse 1976)
4. Plankton (this study)
5. Unicells (Robinson et al. 1983)
1
10^1
10^2
10^3
10^4
10^5
Body size (μm^3)

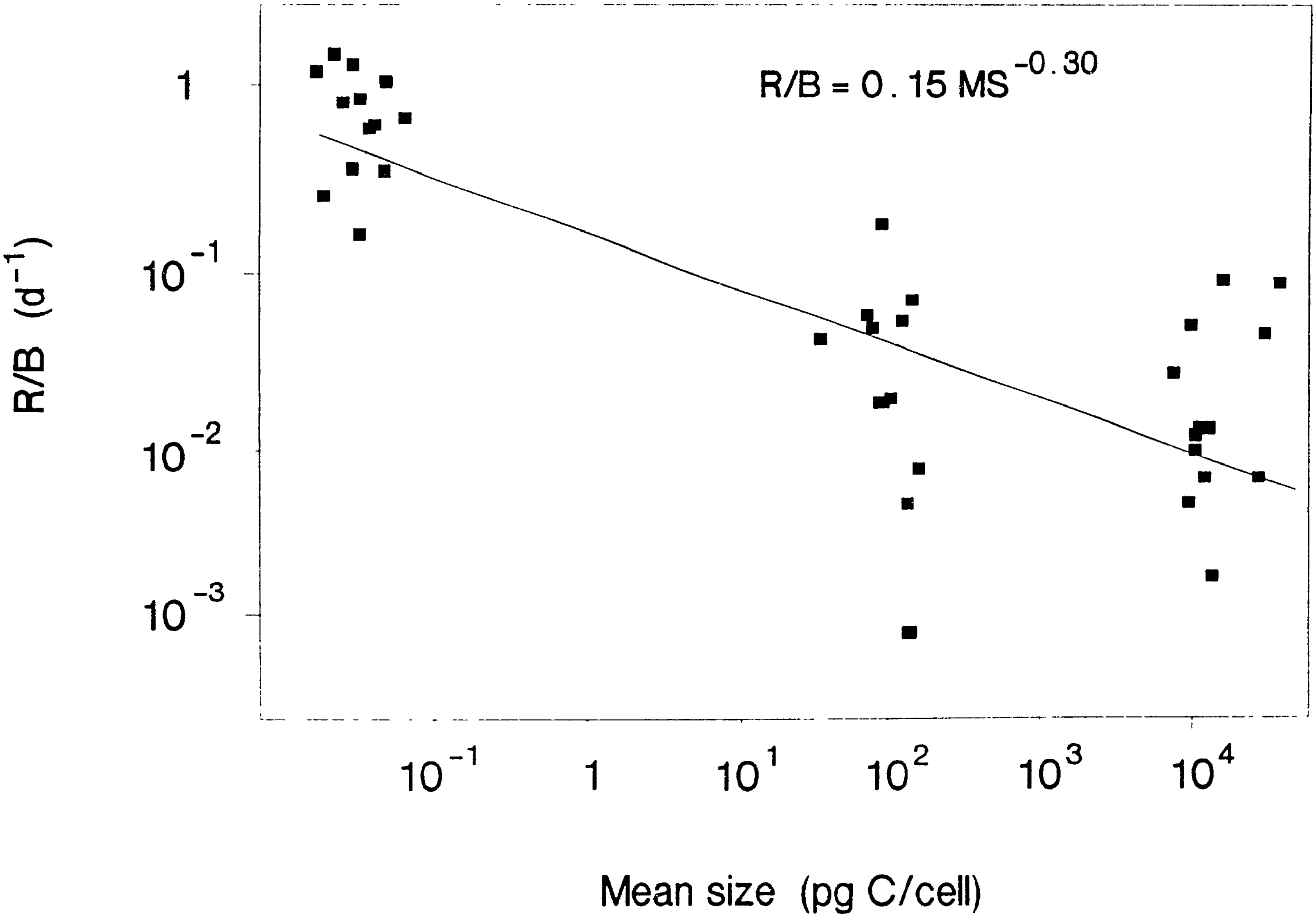
Table 2.5. Parameters of linear regressions between natural logarithms of metabolic rate (pg O_2 d^{-1}) of individual organisms and body size (μm^3). Most relationships had to be transformed to these common units. The unknown precision of these transformations precludes the setting of confidence limits on the intercept. Although the slope parameter does not change under transformation of units, its standard error may be affected. So, the CI shown here may be too narrow. The intercepts shown have been backtransformed from the log scale to the metabolic rate scale.

Organisms	Size range (ESD, µm)	Backtransformed Intercept	Slope	95% CI of Slope	Source
Algae	3 - 40	0.242	0.90	0.79 - 1.02	Banse (1976)
Unicells	0.6 - 60	0.045	0.83	0.72 - 0.94	Robinson et al(1983)
Unicells	1.2 - 125	0.507	0.76	0.72 - 0.80	Hemmingsen (1960)
Protozoa	12 - 270	6.596	0.68		Klekowski (1981)
Euk. unicells	130 - 250	0.628	0.74	0.66 - 0.82	Banse (1982)
Rotifers	100 - 230	19.936	0.52	0.21 - 0.82	Banse (1982)
Zooplankton	300 - 14.4×10^3	0.979	0.84	0.82 - 0.85	Ikeda (1985)
Crustaceans	250 - 17×10^3	2.943	0.78	0.77 - 0.80	Ivleva (1980)
Poikilotherms	780 $5\ 7\times10^5$	1.623	0.74	0.72 - 0.76	Hemmingsen (1960)
Plankton	0 7 - 85[a]	0.068	0.73	0.65 - 0.81	This paper

[a] Median sizes per class.

transformations required to standardize the units and permit comparisons would affect the confidence limits on the parameters to an unknown and variable degree. However we were able to compute 95% confidence limits for our own parameters. Five of the nine slope estimates in Table 2.5 are not significantly different from ours. Furthermore, the intercepts (at a body size of 1 μm^3) from those three regressions which include organisms as small as the picoplankton are similar to ours. Finally, our predicted respiration rates for picoplankton are only about 50% higher than those predicted by the Robinson et al (1983) unicell regression, and about 50% lower in the upper size range of their data set. The mean respiration:biomass ratios (R:B) for picoplankton, nannoplankton, and netplankton respectively were 0.727, 0.038, and 0.027 per day. The ratio declined as a power function of mean organism size, with an exponent of -0.30 (Figure 2.5). The slopes of the relationships of both P:B and R:B to body size tend to decrease with increasing ranges of body sizes in taxonomically homogeneous laboratory samples (Banse and Mosher 1980, Dickie et al. 1987). This is also true for at least some field samples from mixed communities: for example Schwinghamer et al. (1986) fitted slopes of -.304 and -.337 to the relationships between R:B and size of marine benthic meiofauna and macrofauna respectively, whereas the slope of the regression for both size groups combined was only -.21. Both Banse and Mosher (1980) and Dickie et al. (1987) have

FIGURE 2.5. The relationship between the respiration/biomass ratio and the mean size of plankton in screen-separated assemblages from natural plankton communities.

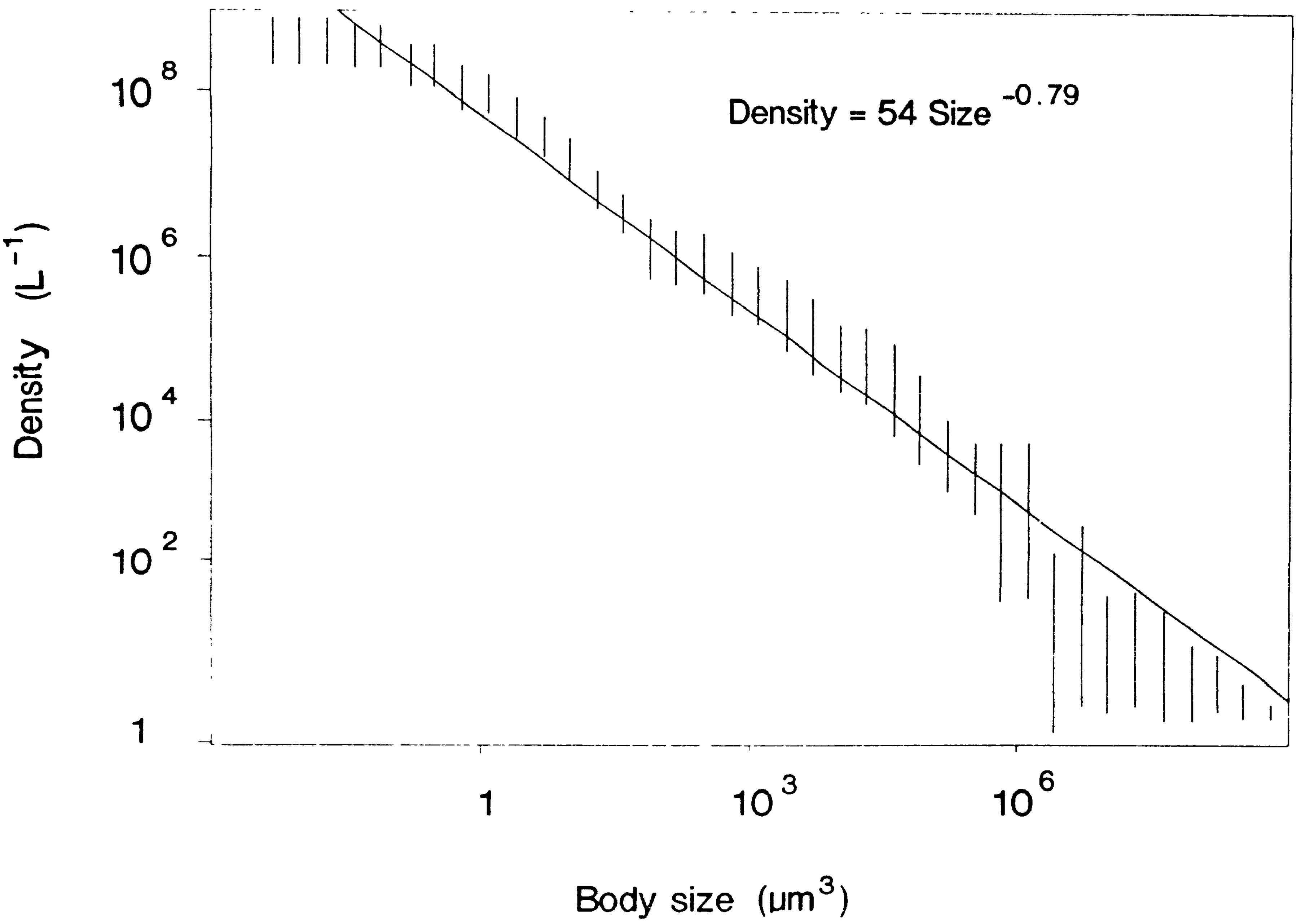

R/B = 0 . 15 MS$^{-0.30}$
R/B (d^{-1})
1
10^{-1}
10^{-2}
10^{-3}
10^{-1}
1
10^{1}
10^{2}
10^{3}
10^{4}
Mean size (pg C/cell)

speculated that the steeper slope within more homogeneous groups (e.g. meiofauna, fish) reflects an ecological scaling factor, compounding the overall general power relationship between metabolic rate and body mass. Our R:B versus body size regression has a slope (-0.30) which is intermediate between the two extremes (-0.21 for the large scale regression, -0.37 for more homogeneous groups). This might be expected, since our size range is relatively small (compared to the range from bacteria to whales), yet we have a mixture of taxonomic and ecological types (prokaryotes and eukaryotes; autotrophs and heterotrophs; unicells and small metazoans, etc.). However, there is the alternate possibility that the proposed ecological scaling factor is in part a statistical artefact, since the probability of obtaining a steeper regression slope increases as the range of the independent variable decreases (Peters 1988).

Dickie et al. (1987) argued that the ecologically-realized respiration rate for individual organisms, termed an "ecological food requirement", should be proportional to the 0.67 power of body size. Because we did not measure individual rates, but have estimated them by calculating means, our regression is a crude test of this hypothesis with respect to natural plankton communities. However, our exponent of 0.73 is not significantly different from 0.67. These authors based their conclusion on parameter estimates for herbivorous mammals. They noted that density was proportional to the -0.75 power of body mass, and that the ratio B:R is proportional to the 0.33 power. We tested these hypotheses with respect to the plankton. Using a large set of size distribution data (n = 1647) collected in our lakes throughout the summer of 1987, we found that density was proportional to the -0.79 power of body mass (Figure 2.6), which does not differ significantly from the mammalian result. B:R was proportional to the 0.30 power of body mass, which is not significantly different from 0.33. Our results confirm that the mean respiration rate per organism in lakes, as well as in fields, scales to the 2/3 to 3/4 power of body size. Our value is intermediate in this range. Of course, even if some particular data set were to yield values significantly different from one extreme or the other, this statistical observation in itself could not establish the conclusion that specific ecological or physiological mechanisms are at work. But our allometric relation for field metabolism is particularly significant for its demonstration that laboratory-based respiration rates and allometric relations can be extrapolated to the field. This is often assumed in limnology, but rarely demonstrated.

It has often been noted (Hemmingsen 1960; Banse 1976) that the common size dependence of growth, respiration, and photosynthesis would suggest that numerous underlying physiological processes are size dependent. In the last ten

FIGURE 2.6. The relationship between body size and density in natural plankton communities. Estimated densities were based on microscopic counting of abundance in as many as 39 size intervals in 58 plankton samples collected throughout the summer of 1987 at 15 lake sites in southern Quebec. Organism size doubles from one interval to the next (log two scale).


Density = 54 Size $^{-0.79}$
Density (L^{-1})
10^8
10^6
10^4
10^2
1
1
10^3
10^6
Body size (μm^3)

years new data and re-analyses of existing data have borne out this contention (e.g. Blueweiss et al. 1978, Schlesinger et al 1981, Knoechel and Holtby 1986, Ikeda 1985, Fenchel 1974, Peters 1983, Calder 1984). The exponents of the allometric dependences of production, growth, and turnover rates on body size may be predicted from the exponent of the equation for respiration rate (e.g. Dickie et al. 1987). For example, if growth and respiration rate are related to body size by a common exponent of 0.75, then the instantaneous rate of increase (r) is a power function of body size with an exponent of -0.25. Platt and Silvert (1981) have proposed that the respiration rate exponent for aquatic organisms is 0.67, and that it is 0.75 for terrestrial organisms. Confidence intervals about the exponents found in most studies do not permit a test of this hypothesis (Table 2.5), but some aquatic results are significantly higher than 0.67 (e.g. Banse 1976), and others are even higher than 0.75 (e.g. Ivleva 1980, Ikeda 1985). Perhaps more remarkable than the question as to whether a particular rate has an exponent closer to 0.67 or to 0.75 is the observation that many studies using different methodologies, very different organisms, and different environments yield fairly similar results. We have shown that the relationship between metabolic rate and body size in field communities is not different from that which would be predicted from more restricted laboratory relationships.

CHAPTER III

LOGARITHMIC TRANSFORMATION IN PREDICTIVE ECOLOGY

ABSTRACT

Ecological data are frequently log transformed prior to the development of linear regression models, which are then used for prediction on the original untransformed scale. In this paper, the rules for backtransformation are elucidated. A theorem is presented, which sets limits to the relationship between the coefficients of determination on the original and transformed scales. Simulated data and empirical results from limnology are used to illustrate the applications and limitations of the theoretical part of the paper.

Introduction

Predictive ecologists use regression analyses of log transformed data extensively (e.g. Peters 1986). Regression establishes quantitative relationships between variables, without addressing issues of mechanism or causality. Each regression is in itself a predictive model applicable to new data drawn from the same populations as the original sample data used in the regression. Since its proponents eschew questions of mechanism (Peters 1983), predictive ecology often resembles a specialized branch of applied regression analysis. At present, the theory and techniques of linear regression are far more advanced than those of non-linear models (e.g. Draper & Smith 1981; Seber 1977). Hence ecologists are frequently obliged to transform non-linear phenomena for analysis in a linear framework.

The conventional linear regression model requires that the variance of the response variable be constant at all levels of the regressor variable(s), that the deviations of repeat observations of the response at a given value of the regressor(s) have an expected value of zero, and that these deviations be uncorrelated (Draper & Smith 1981). Logarithmic transformation of original data following a "power" or exponential curve will often generate new variables which satisfy these requirements. The further requirement that the above deviations be normally distributed about zero is not essential to regression analysis, but it is essential to the use of the t and F distributions. If the data are normal, then these distributions may be used to establish confidence limits, test the significance of the regression results, and evaluate the precision of predictions, using the original or other suitable data. In practice, if the other conditions of regression analysis are met, then ecologists usually assume that the normality assumption is also met. In most cases the data are not remarkably deviant, and this assumption is a minor source of error. If the results of the regression analysis are satisfactory (high R^2 and low MSE relative to the mean value of the response), then one can draw descriptive conclusions about the relationship between the variables.

This paper addresses the problems which arise when a regression on log transformed data is intended for predictions, beyond the simple description of a trend. These problems concern the correct computation of statistics and predictions for the original variables, when the regression was performed on transformed data. Rules for backtransformation will be elucidated, since these have been ambiguously represented in the ecological literature (Sprugel 1983). The variance and confidence limits on predictions will be related to the broader question of what it is we want to predict. The coefficient of determination (R^2) expresses the proportion of variation in the dependent variable explained by the regression on the independent variable(s), but this value does not apply to backtransformed values, if the regression was performed on log transformed data. A formula for the calculation of the backtransformed R^2 appropriate in many cases is presented in a theorem. Other, less idealized cases, are treated in a simulation exercise. Finally, these statistical results are compared with the analysis of some results in predictive limnology, relating plankton respiration rates to body size and phosphorus concentrations.

Problems of Backtransformation

1. The Correction Factor:

The regression between two log transformed variables produces a series of predicted mean values y_i, each with its own standard error

$$s_i = \mathrm{SEE}[1/n + (x_i - \mu)^2 / \Sigma(x - \mu)^2]$$

where SEE is the standard error of the estimate, or square root of the mean squared error (MSE) of the regression, μ is the mean value of the independent variable, and x_i is the value of x yielding the prediction y_i (Draper and Smith 1981). This formula is simply the straight line case of the more general multiple regression formula

$$s_i = \mathrm{SEE}[\mathbf{x}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i']$$

where $\mathbf{x}_i$ is the vector of values of the independent variables yielding the prediction y_i, and X is the matrix of observations (Seber 1977). Each such prediction is expected to be normally distributed, but it is sufficient that it be unimodal and free from skewness and kurtosis. The antilogarithm of each prediction y_i corresponds to a prediction on the scale of the original data, prior to the logarithmic transformation. However, on this scale, the errors are lognormally distributed, skewed to the right. Hence the antilogarithm of the mean y_i yields the median on the original scale, but not the mean of the antilogs.

The correction factor required to estimate the mean of a backtransformed prediction may be derived as follows. For simplicity, and without loss of generality, we may assume that natural logarithms were used in our regression. Now the moment generating function (Hogg and Craig 1978) of a normally distributed random variable y with mean μ and variance σ^2 is

$$E[e^{tY}] = \exp[\mu t + \sigma^2 t^2/2] \text{ , for all real t.}$$

Hence, the mean prediction of exp[y] at x_i is

$$E[e^{y_i}] = \exp[y_i + s_i^2/2] \text{ ,}$$

which yields the correction factor $\exp[s_i^2/2]$.

This correction factor has been presented several times in the recent ecological literature (Baskerville 1972; Sprugel 1983; Lehman 1988; Welsh et al. 1988). However, none of these references make clear what formula is to be used for the standard error, when regression estimates are backtransformed. In fact, Sprugel (1983) incorrectly states that the regression SEE should be used in computing the above correction factor. But it is essential to use the standard error of the prediction y_i here, rather than the overall regression standard error, as was incorrectly recommended by Sprugel. Hence the correction factor is not constant for all predictions from a single regression, but increases with distance from the overall mean of the logged observations.

2. Confidence and Predictions:

The variance of the predictions on the scale of the original data may be calculated from the variance of the predictions from the regression using the formula

$$\mathrm{Var}(\exp[y]) = \exp[2y_i + s_i^2](\exp[s_i^2]-1)$$

However, this result cannot be used to calculate confidence limits in the usual fashion, based on the t distribution, since this assumes normality. The simplest procedure to calculate confidence limits on a corrected backtransformed mean prediction is to calculate the confidence limits on the regression prediction (still on a natural log scale), followed by backtransformation and correction using the same correction factor as was used for the mean. This results in an appropriately asymmetric confidence interval.

This result is pertinent to the choice of what we want to predict. If we choose to predict the mean, then this will correspond to the average of a long run of observations. But this predicted value will be higher than most individual observations. A smaller number of very large values will account for the expectation of eventually attaining the mean. If a correction factor is not used, then we have chosen the median by default. Our predictor will be lower than the long term average. These facts, combined with the asymmetry of the confidence interval, justify the choice of the median (no correction factor) to predict an individual value. A greater proportion of the observations will be closer to this value than to the corrected mean. However, the mean should be used to predict the average of a set of observations of the dependent variable.

3. The Coefficient of Determination:

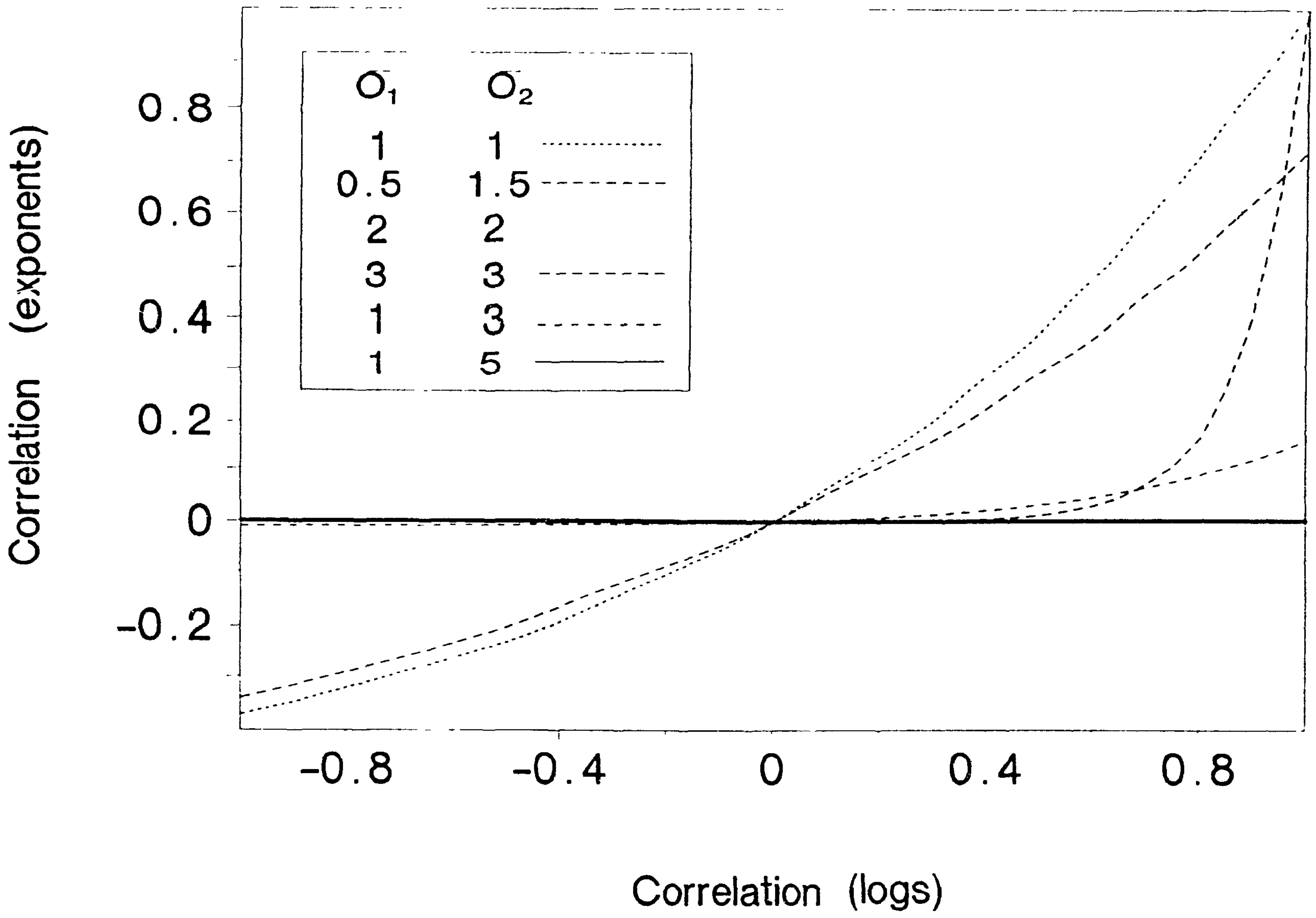
The coefficient of determination, equal to the square of the correlation coefficient, is the proportion of the variation in the observed values of the dependent variable which is accounted for by the regression model. But if the data undergo a logarithmic transformation prior to analysis, then the coefficient of determination does not describe the variation of the original observations accounted for. It applies only to the log transformed values. Appendix 2 shows that if the two variables, y_1 and y_2 , in a straight line regression have a bivariate normal distribution, then the coefficient of coefficient of determination between $\exp[y_1]$ and $\exp[y_2]$ is

$$r_z^2 = \frac{(\exp[r\sigma_1\sigma_2] - 1)^2}{(\exp[\sigma_1^2] - 1)(\exp[\sigma_2^2 - 1)}$$

and this value is less than or equal to r^2 between y_1 and y_2, with equality occurring only at r = 1 or 0 (Appendix 2). As Figure 3.1 shows, negative correlations are reduced more by exponentiation than are positive ones, and the difference between coefficients of correlation on the logarithmic and exponential scales increases rapidly with increasing variance.

In order to meet the requirements for linear regression, a set of data does not have to have a bivariate or multivariate normal distribution. This is the "ideal" regression situation, where an elliptical cloud of points has a major axis corresponding to the regression line. But even when the predictor(s) and response are normally distributed, they are not necessarily bi- or multivariate normal. For example, there might be two or more clusters of data aligned on the regression line. Under these circumstances, the above formula for the coefficient of determination between the exponentiated variables does not apply, and the correlation on the log scale may sometimes be lower than that on the exponential scale. However, if the response and the predictor(s) are jointly multivariate normal, then the regression predictions and the observed values of the response are expected to be samples from

FIGURE 3.1. Relationship between the correlation of bivariate normal variables and the correlation of their exponents. Trend lines are shown for six different combinations of standard deviations.



σ_1
σ_2
1 1
0.5 1.5
2 2
3 3
1 3
1 5
0.8
0.6
0.4
0.2
0
-0.2
Correlation (exponents)
-0.8
-0.4
0
0.4
0.8
Correlation (logs)

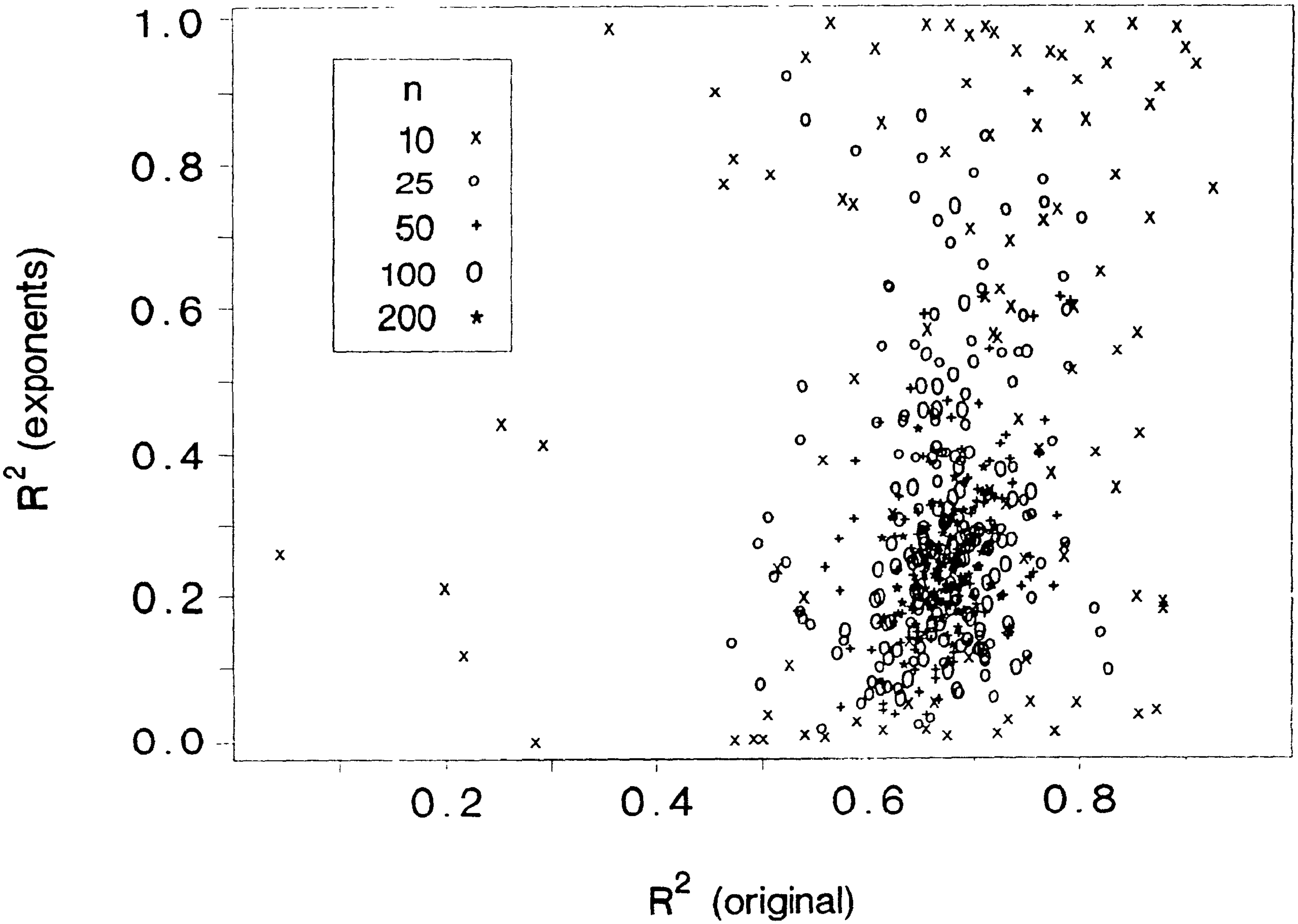
a bivariate normal distribution, and the above formula will apply. In this case, s_1 and s_2 are the respective standard errors of the observed and predicted values of the response variable. Caution must be exercised in interpreting the coefficient of determination. It always expresses the proportion of variation in the response variable explained by a regression on the predictor variable(s). When the conditions of linear regression are met, and especially when all variables concerned have a multivariate normal distribution, the explained and residual variance are distributed randomly among the data. But when the data are exponentiated, distributions are lognormal, and the explained variation is systematically distributed, with unexplained variation due primarily to the skewed pattern in the data. Under these circumstances, the coefficient of determination when correctly calculated with the data on the exponential scale expresses the proportion of variation explained, but this may be an operationally useless observation. In the extreme case, with a cloud of data at one end of the graph and one point at the other, r^2 may depend entirely upon that one point, and provide no information about the rest of the data.

Simulation

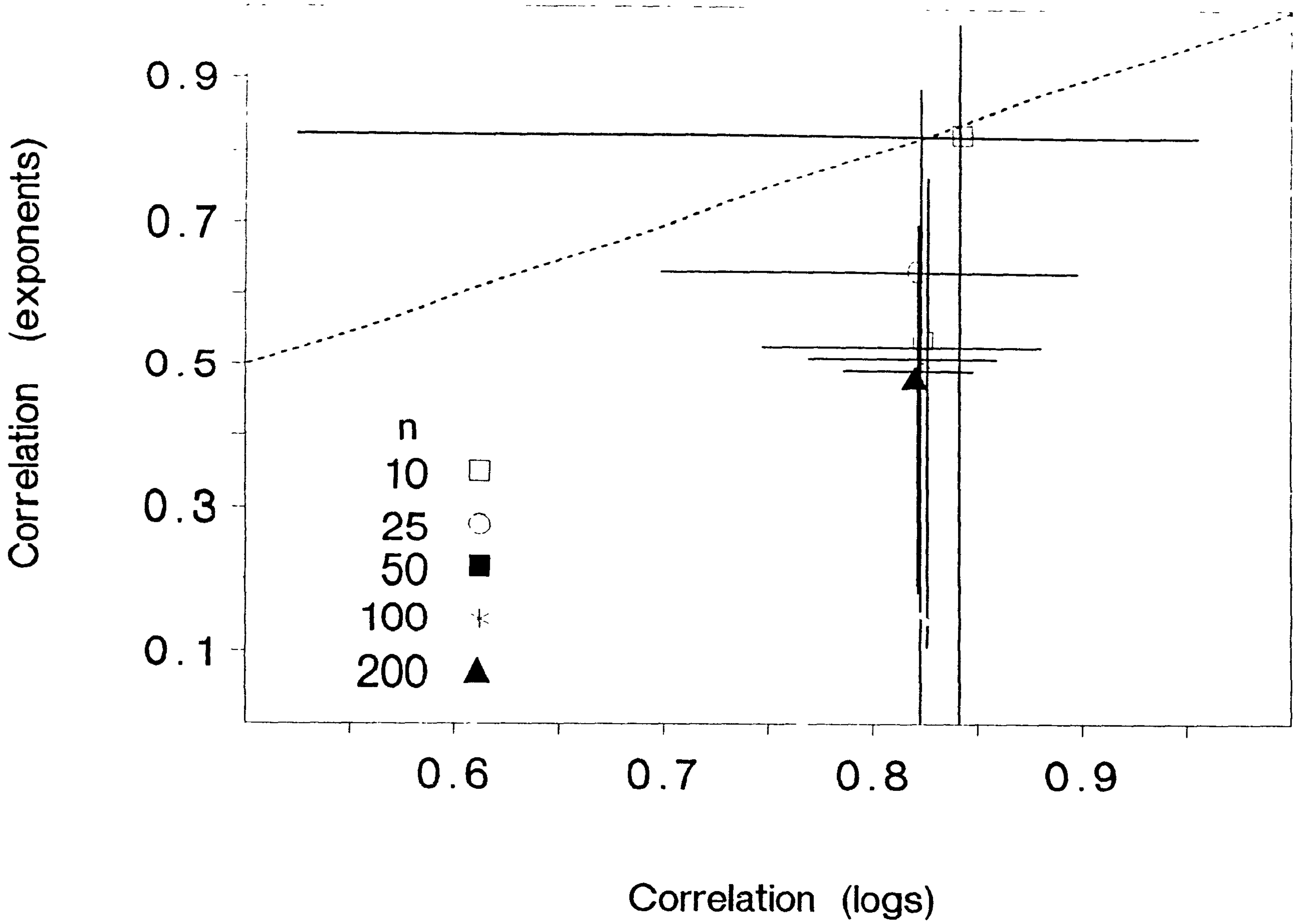
In order to observe the behaviour of r^2 under exponentiation when the requirements of multivariate normality were not met, simulation data were generated: at 100 values of the independent variable (0.1 to 10, by intervals of 0.1), the dependent variable values were calculated in the following manner. A vector of 500 "observations" was constructed, with 5 values at each X spaced symmetrically around the X value. Thus at X = 5 for example, the five Y values were 2, 4, 5, 6, and 8. The same deviations were set at each X value, to ensure homogeneity of variances. Then five different random samples were selected from this population (without replacement), with sample sizes of 10, 25, 50, 100, and 200. The coefficient of determination (squared correlation coefficient) was calculated for each sample, as well as for the antilogs of both the X and the Y values. These transformed values are analogous to the skewed backtransformed estimates obtained after linear regression on data which has undergone a logarithmic transformation. This procedure was repeated 100 times.

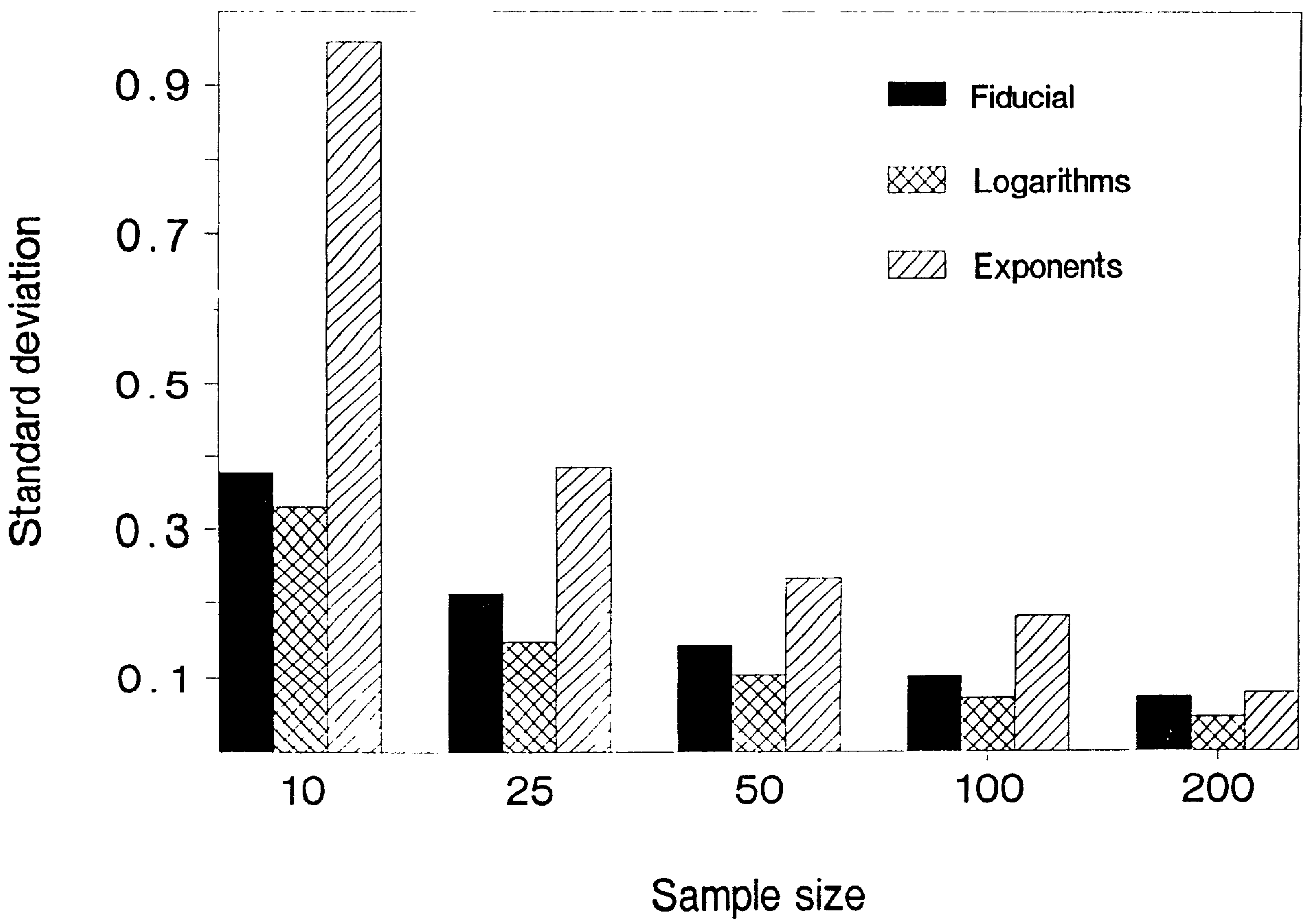
Almost all coefficients of determination for the data with homogeneous variance (original samples = "logarithmic" scale) were higher than those for the antilogged data (Figure 3.2A). Median correlations on the log scale were invariably higher than on the antilogged ("exponential") scale (Figure 3.2B). Median coefficients for the original samples did not differ significantly with sample size. However, the coefficients of antilog data actually decreased with increasing sample size. This is due to the increasing numbers of repeat observations

FIGURE 3.2. (A) Comparison between the coefficients of determination for the simulation data before and after transformation. The X-axis coefficient is for the original data, where the dependent variable has homogeneous variance throughout the range of the independent variable. The Y-axis gives the coefficient for the relationship after both variables have been exponentiated. There are 500 points, 100 for each of five sample sizes. (B) The median correlations for each of the five sample sizes plotted in Figure 3.2A. The 1:1 line and 95% confidence intervals are shown. (C) Standard deviations of the correlation coefficients calculated in the simulation. The fiducial standard deviation is based on a formula which only applies if the underlying data are normally distributed. The "logarithms" are data with homogeneous variance of the dependent variable, whereas the third bar is for the exponents of these same data.



1.0
0.8
0.6
0.4
0.2
0.0
R^2 (exponents)
0.2
0.4
0.6
0.8
R^2 (original)
n
10 x
25 o
50 +
100 0
200 *


Correlation (exponents)
0.9
0.7
0.5
0.3
0.1
0.6
0.7
0.8
0.9
Correlation (logs)
n
10 □
25 ○
50 ■
100 ✳
200 ▲


Fiducial
Logarithms
Exponents
Standard deviation
0.9
0.7
0.5
0.3
0.1
10
25
50
100
200
Sample size

at individual X values. In any data set, when different repeat observations of the dependent variable are available at given values of the independent variable, then the maximum possible r^2 is less than one. The increasing number of repeat observations had a greater impact on the antilog data since the individual observations were not symmetrically distributed about the median values, and the total sum of squares is potentially greater. The resulting median values of the coefficient of determination are summarized in the following table.

Sample size	Original data with stable variance r^2	Between exponents r^2
10	0.711	0.699
25	0.677	0.413
50	0.682	0.288
100	0.674	0.263
200	0.675	0.242

With a small sample size (10) and few repeats, a regression between these log transformed values would be expected to explain about 71% of the variation, corresponding to about 70% of the variation in the backtransformed data. At a larger sample size (200) with more repeats, the log:log regression would explain about 68% of the variation, corresponding to only 24%-41% of the backtransformed values.

Confidence limits on the correlation coefficients were calculated by first computing the empirical standard error of the z transformation of the coefficient, calculating the upper and lower 95% limits on z, and then backtransforming to the r scale. The empirical results for the log data were similar to fiducial standard errors, equal to $(n - 3)^{-0.5}$ (Snedecor & Cochran 1967), which assume normality of the underlying data (Figure 3.2C). All standard errors decreased with increasing sample size (Figure 3.2C). At all sample sizes, the correlation coefficients of the log data had substantially lower standard errors than those for the exponentiated data (Figure 3.2C), resulting in narrower confidence regions (Figure 3.2B). These limits express the confidence one has in the statement that variation in one variable accounts for a certain percentage of the variation in the other variable. This modelling exercise has shown that, for this data set at least, regression with homogeneous variance of the response accounts for a higher percentage of the total variation than regression between the antilogs of the same variables, and that our confidence in this proportion must be higher.

As a general rule, if the variables of interest have been log transformed prior to the examination of predictive relationships in linear regression, then statements of confidence should refer to backtransformed (real world) results, using a correction factor if mean values are of interest. Statements about the "percentage variation explained" should not be taken directly from the regression analysis at all. Rather, the squared correlation coefficient between the original observations and backtransformed predictions may be separately calculated. The validity of this statistic does not depend upon normality, and the result is a legitimate expression of the extent to which variability in the independent variable accounts for variability in the dependent variable.

Examples in Limnology

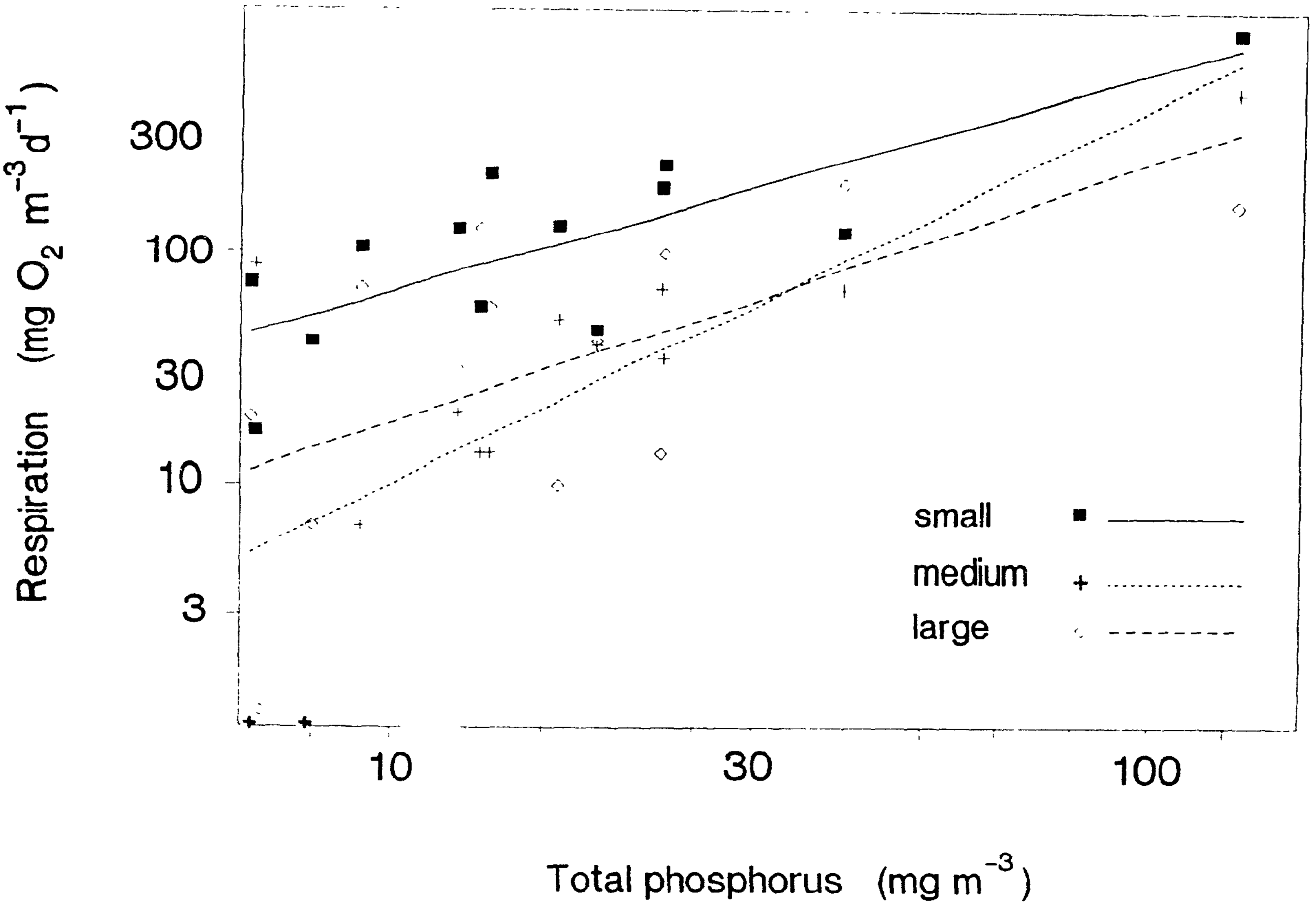
This section will follow up on two empirical relationships in plankton communities, based on a small number of samples collected in Canadian lakes. The respiration rate of plankton assemblages ($\mu g\ O_2\ L^{-1}\ d^{-1}$) was found to be significantly related to the phosphorus concentration measured in the same water samples (Chapter 2 above). After measuring the abundance of organisms and total volume in these same samples, the calculated mean respiration rate of individual plankton was found to be significantly dependent upon the size of the organisms. In both cases, logarithmic transformation was required prior to regression analysis. So they will be treated as examples in which the implications of the first statistical section of this chapter are explored.

1. Total phosphorus and plankton respiration

The relationships between the total phosphorus concentration and the respiration rates of small (<5 μm), medium (5 - 40 μm), and large (>40 μm) plankton are shown in Figure 3.3 and Table 3.1. In the first two cases, r^2 is significantly higher between the backtransformed values than between the log transformed data. The presence of an outlier at the high end of both log transformed scales makes it clear that these data do not correspond to bivariate normal distributions (Figure 3.3B). Hence the theorem in Appendix 1 does not apply. The regression lines are clearly strongly determined by one point which accounts for most of the variance in the data set. Hence, the "percentage variation explained" by these relationships is very high, even though it is not very relevant to variation among most of the points. In the third data set, the linear fit is poor (although significant), and on either scale only a small portion of the variation is explained by the regression on the independent variable.

In Figure 3.4, the predicted log values are compared with the input log values. For ease of reference, the axes are labelled with backtransformed (median) units. The statistics of

FIGURE 3.3. (A) The relationship between the total phosphorus concentration and plankton respiration rates, for assemblages of three sizes of organisms collected in Canadian lakes. The least squares regression lines between the logarithms of the data are shown. For each line, n = 13. (B) The data and trend lines from 3a on a linear scale.


Respiration (mg O_2 m^{-3} d^{-1})
300
100
30
10
3
10
30
100
Total phosphorus (mg m^{-3})
small
medium
large

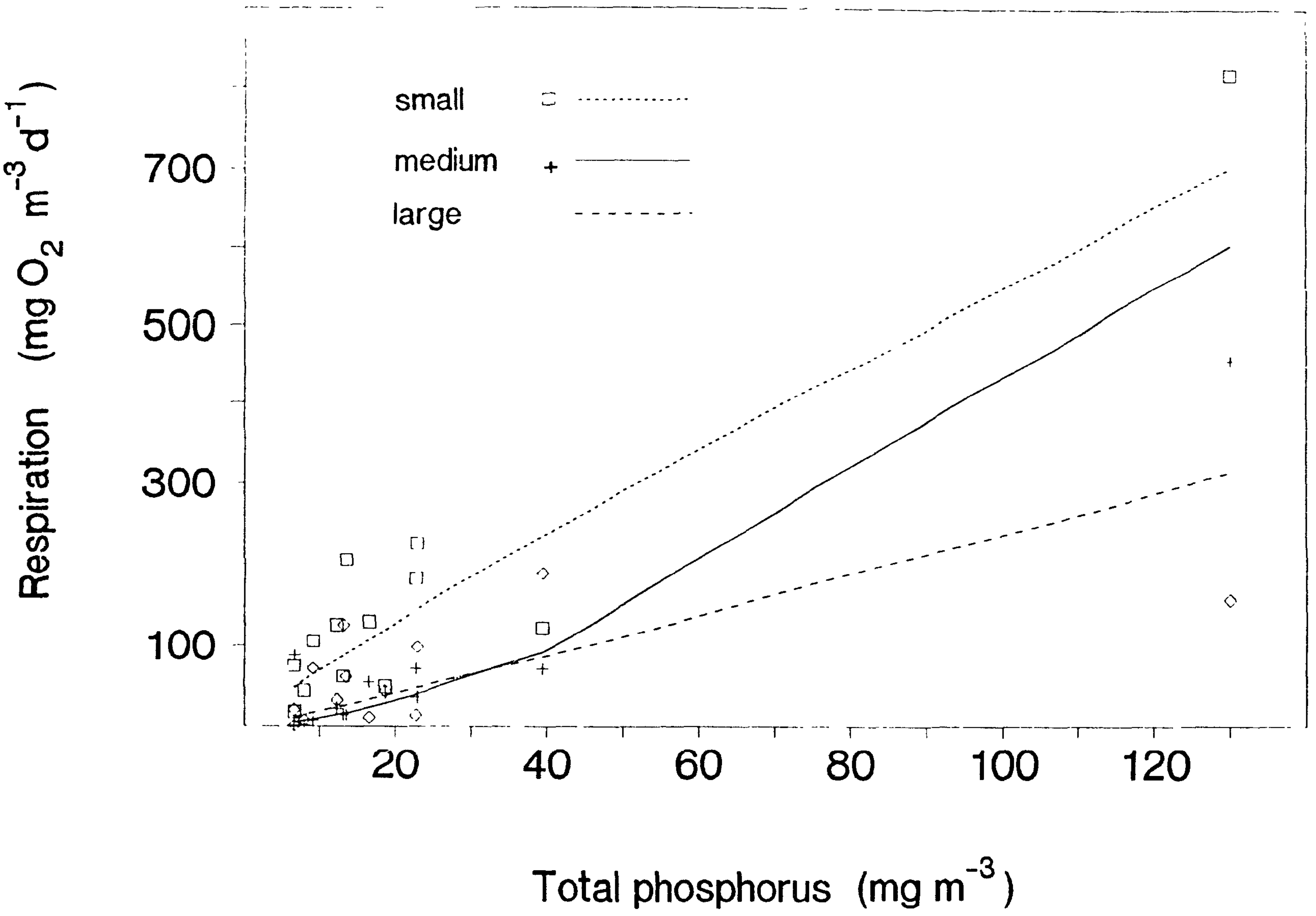

small
medium
large
700
500
300
100
20
40
60
80
100
120
Respiration (mg O_2 m^{-3} d^{-1})
Total phosphorus (mg m^{-3})

Table 3.1. Statistics of the regressions between the base 10 logarithms of the total phosphorus concentration (TP) and the respiration rates of small, medium, and large size plankton. In each case, n = 13. The first value of the coefficient of determination is the proportion of the variation in the response variable explained by the predictor variable on a log scale (TP). The second value was calculated between the original variables prior to logarithmic transformation. The third value was calculated between the observed response and the predicted response after correction for backtransformation. The lower part of the table presents statistics of the relationships between the log transformed observations and the predictions of the regressions on TP

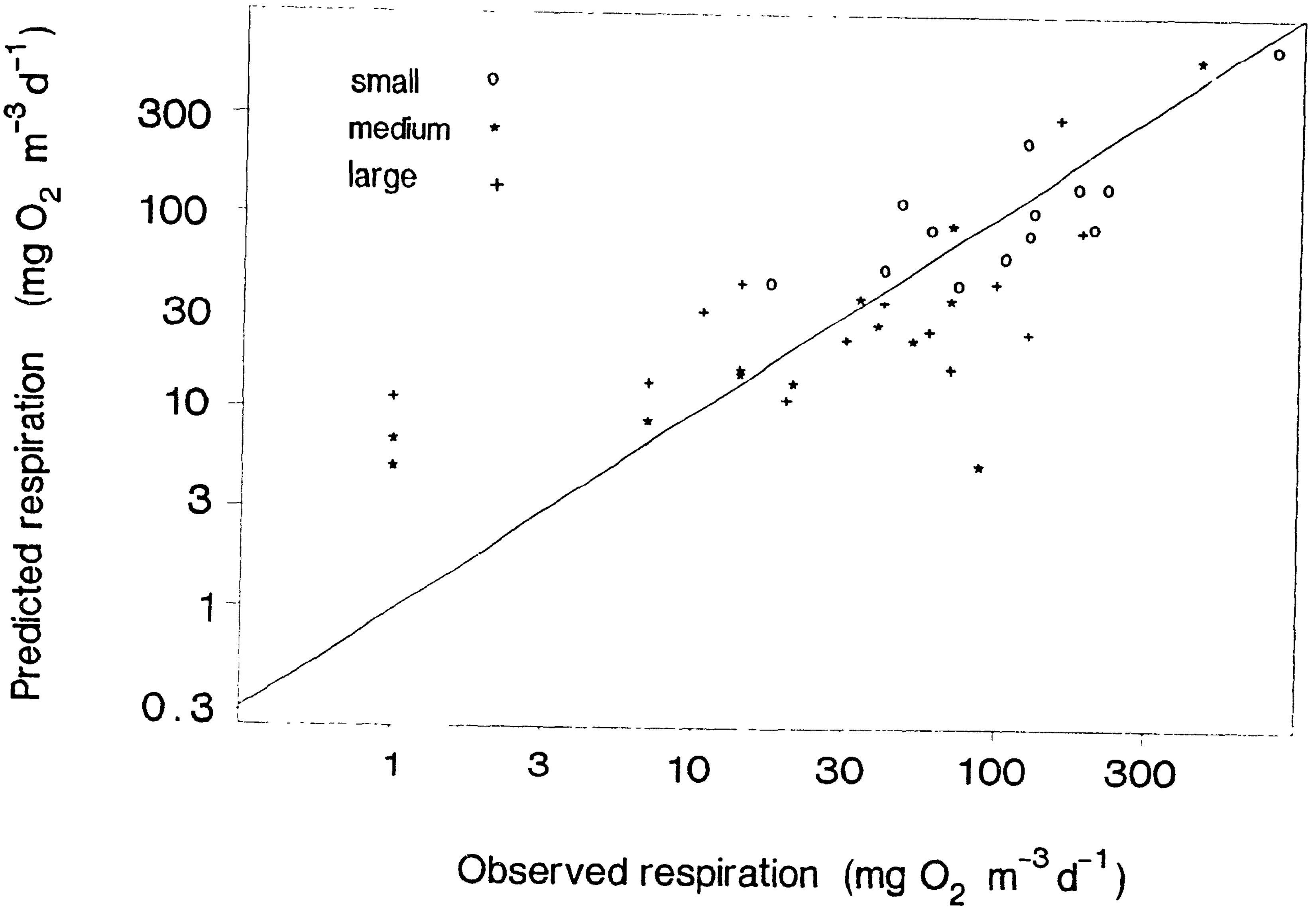
	Parameters		MSE	F	Coefficients of determination		
	a	b			1	2	3
Small	.926	.909	.0715	17.5	.614	.906	.906
Medium	-.555	1.579	.2735	13.8	.557	.936	.958
Large	.184	1.095	.2869	6.3	.365	.371	.313

Observations vs predictions (log scale):

	Parameters		F*	
	a	b		
Small	.782	.614	3.5	P>.05
Medium	.601	.557	4.4	P<.05
Large	.958	.365	9.6	P<.01

* This F test compares the joint parameters with the pair (0,1)

FIGURE 3.4. The logarithms of measured respiration rates of plankton assemblages compared with the predicted log rates from the three size-specific linear regressions on the log of the total phosphorus concentration. The 1:1 line is shown.



small
medium
large
Predicted respiration (mg O_2 m^{-3} d^{-1})
300
100
30
10
3
1
0.3
1
3
10
30
100
300
Observed respiration (mg O_2 m^{-3} d^{-1})

regressions between these predicted and observed values are presented at the bottom of Table 3.1. Ideally, these lines would have intercepts of 0 and slopes of 1, with equal distributions of points on either side of the line. Only the first fit (for small plankton) is not significantly different from this ideal (P>.05). Nevertheless, one is usually more interested in the match between the model results and empirical observations than in the match between log transformed numbers.

In order to examine the success of these models at the scale of observable data, predictions were backtransformed and compared with observed input values (Table 3.2). Backtransformed predictions were evaluated with and without correction. Backtransformation without correction should provide an estimate of median values, whereas use of a correction factor based on the specific standard error results in a mean estimate (as discussed in the first part of this paper). A further estimate, based on a uniform correction using the MSE, was also computed. In general, the correspondence between predicted values and observations reflects the relative precision of these three regressions (Table 3.3, Figure 3.5). Hence, in the case of small plankton, all of the three sets of backtransformed estimates are similar to the observed values, and follow the same trend. In the cases of the other two regressions, all backtransformations deviate significantly from the observed values, both in

Table 3.2. Observed and predicted respiration rates ($\mu g\ O_2\ L^{-1}\ d^{-1}$) for three plankton size classes. Predicted rates are backtransformed from linear regressions between logarithmic transformations of observed rates and total phosphorus concentrations. Median predictions are uncorrected. Mean predictions are corrected for backtransformation with the standard error specific to each prediction. MSE-corrected predictions are corrected with a single correction factor for all predictions from the same regression

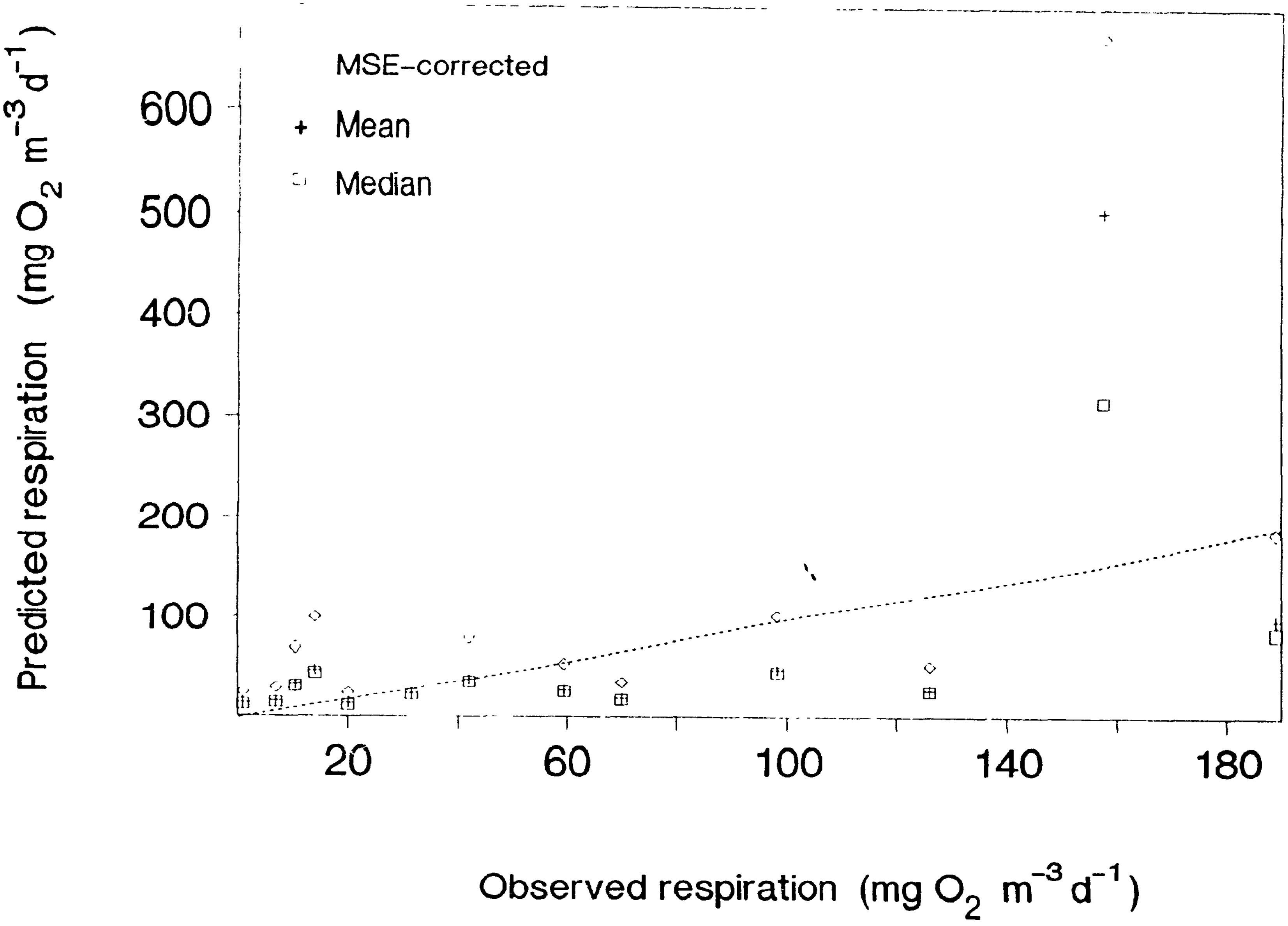
Observed rate	Predicted rates		
	Median	Mean	MSE-corrected
Small size class (organisms smaller than 5 μm ESD)			
73	46	48	56
17	47	49	57
42	55	57	66
105	63	64	76
126	82	83	99
59	87	89	106
206	90	91	108
130	108	110	131
47	121	122	146
182	144	147	175
227	146	148	176
123	238	246	288
819	706	793	853
Medium size class (organisms between 5 and 40 μm ESD)			
1	5	6	11
88	5	6	11
1	7	8	15
7	9	10	19
21	14	15	30
14	16	17	33
14	17	18	35
53	23	25	48
40	28	30	58
70	39	41	80
35	39	42	81
70	92	104	189
459	606	947	1251
Large size class (organisms larger than 40 μm ESD)			
20	12	14	25
1	12	14	26
7	15	16	31
70	17	19	37
32	24	25	51
126	25	27	54
59	26	28	56
11	33	35	71
42	38	40	80
14	47	50	100
98	47	51	101
189	85	97	182
158	316	504	675

Table 3.3. Statistics of the regressions between observed and predicted respiration rates (Table 3 2). The second F value refers to a test of the hypothesis that the parameter set is equivalent to (0,1). Significant values correspond to biased fits between predictions and observations.

Small size class

	a	b	MSE	r^2	F	F (vs 0,1)	
Median	-0 088	1.117	4633	900	100	0.96	
Mean	10.038	0.990	4354	906	107	0.11	
MSE-corr	-0 088	0.924	4633	.900	100	0.60	
Medium size class							
Median	16.62	0 728	664	.958	254	17.8	P< 01
Mean	21.97	0.462	666	.958	253	180.9	P<.01
MSE-corr	16.62	0 352	664	.958	254	485.3	P<.01
Large size class							
Median	39 29	0 453	2660	357	6	4.7	P<.05
Mean	45 10	0 261	2840	313	5	20.3	P<.01
MSE-corr	39 29	0 212	2660	357	6	48.7	P<.01

FIGURE 3.5. Observed plankton respiration rates compared with the backtransformed predictions from a linear regression between the logs of observed rates and total phosphorus. The predictions are shown with and without correction for backtransformation. The 1:1 line is also shown. Separate panels are shown for small, medium, and large plankton.


MSE-corrected
+ Mean
□ Median
Predicted respiration (mg O2 m-3 d-1)
600
500
400
300
200
100
20
60
100
140
180
Observed respiration (mg O2 m-3 d-1)

individual precision and in trend. Hence, for predictive purposes in these examples, it makes no significant difference whether one chooses the mean or median predictions. This is clear from the examination of confidence limits on the predicted mean values. Ninety-five percent confidence limits were first calculated for regression predictions. These were then backtransformed and corrected to produce the confidence intervals (Table 3.4). At every point the median values (Table 3.2) are well within the 95% confidence intervals for the means.

2. Plankton size and respiration

There is a significant statistical relationship between the mean size and respiration rate of plankton. If the size distribution of a plankton community is known, then predictions from this relationship might be used to estimate the community respiration rate. Using the same respiration data as above, the log of mean organism size explained 89% of the variation in the log of the mean respiration rate, and predictions did not differ significantly from observations (Table 3.5). When the predictions were backtransformed, and compared with the observations, the median and MSE-corrected values performed about as well as the mean values, although the variation in the predicted values explained only 53% of the variation in the observations (Table 3.5, Figure 3.6). Hence, in this case the

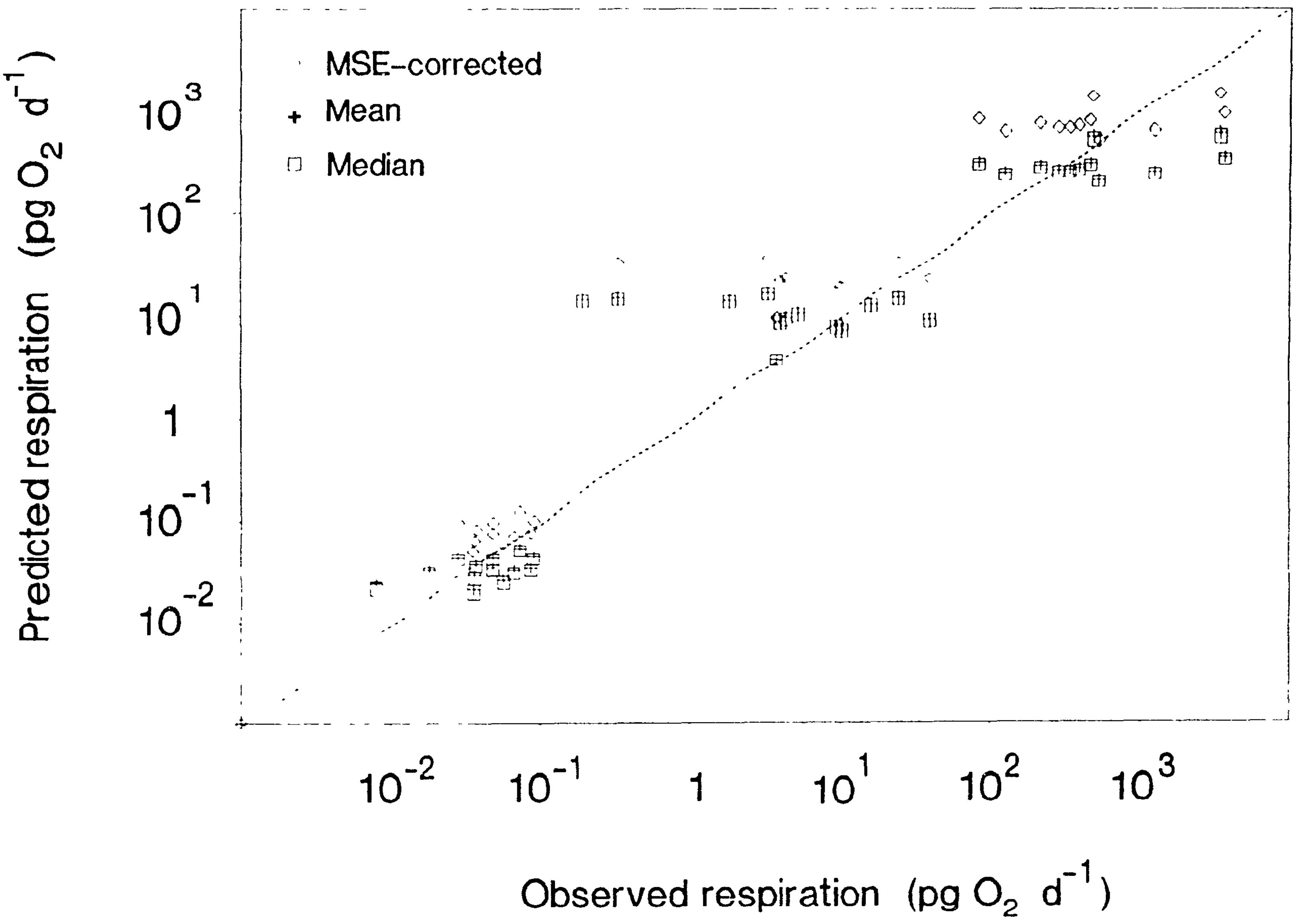
Table 3.4 Predicted mean size class respiration rates (μg O_2 L^{-1} d^{-1}), with upper (U) and lower (L) 95% confidence limits.

TP	Small			Medium			Large		
	L	Mean	U	L	Mean	U	L	Mean	U
6.52	27	48	85	2	6	19	4	14	43
6.61	27	48	86	2	6	19	4	14	43
7.85	34	57	94	3	8	22	6	16	45
9.12	40	64	103	4	10	25	7	19	48
12.22	56	83	125	7	15	34	11	25	56
13.06	60	89	131	8	17	37	12	27	59
13.46	62	91	134	8	18	38	13	28	61
16.56	76	110	160	12	25	52	16	35	74
18.62	84	122	179	14	30	63	19	40	86
22.70	98	147	221	19	41	92	22	50	113
22.91	98	148	223	19	42	93	22	51	115
39.36	140	246	433	34	104	315	31	97	302
130.02	274	793	2299	118	947	7587	60	504	4247

Table 3.5. Statistics of the relationship between the natural logarithms of plankton organism size and respiration rate (n = 39) At the bottom of the table, the total observed respiration rate in three size classes in thirteen different lakes (n = 39) is compared with predicted rates estimated by applying the results of the regression between size and respiration to the known detailed size composition of these samples

	a	b	MSE	r^2	F	F (vs 0,1)
ln(size) vs ln(resp)	-16.502	0.728	1.8656	.894	313	
observed vs pred.	-1.296	0.894	1.6684	894	313	2 2
Observed vs predicted (after backtransformation)						
Median	.0000792	0.0915	<0.0001	.527	41	2186
Mean	.0000833	0 0974	<0.0001	.531	42	1932
MSE-corrected	000201	0.2325	<0.0001	527	41	230
Size class respiration (predicted vs observed)·						
Median	25.184	0.351	3264	456	31	69
Mean	26.221	0 369	3449	467	32	55

FIGURE 3.6. Observed mean respiration rates of individual plankton organisms compared with mean rates predicted from a linear regression between the logarithms of mean size and mean rate. Observations were taken on three broad size intervals. Predicted rates are shown with and without correction for backtransformation. This figure is shown on a logarithmic scale to facilitate examination of the results. The 1:1 line is shown.


MSE-corrected
+ Mean
Median
Predicted respiration (pg O_2 d^{-1})
10^3
10^2
10^1
1
10^{-1}
10^{-2}
10^{-2}
10^{-1}
1
10^1
10^2
10^3
Observed respiration (pg O_2 d^{-1})

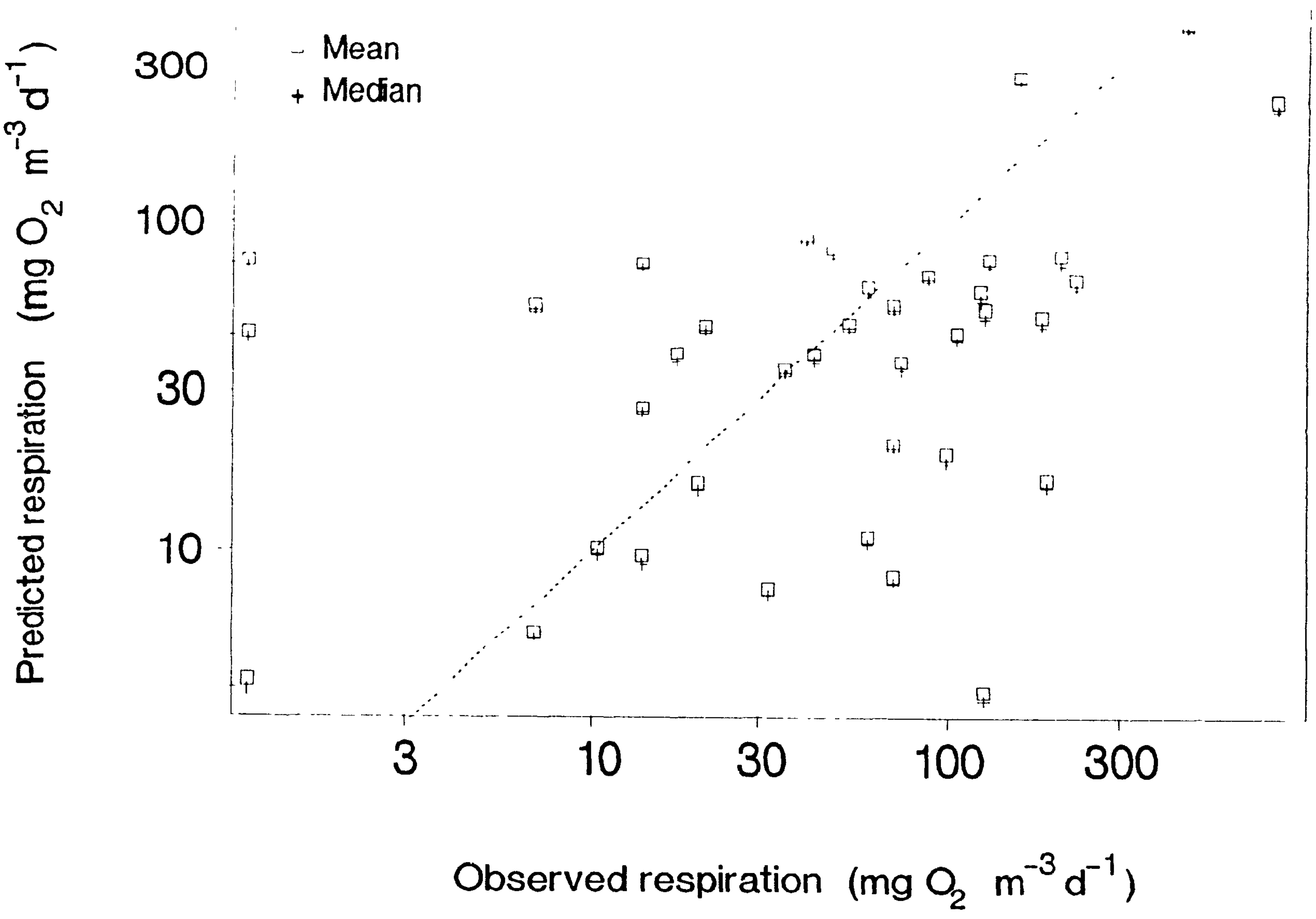
coefficient of determination was lower after backtransformation. If the log-transformed data had been bivariate normal, which they clearly are not (Figure 3.6), then the theorem in Appendix 1 would predict a maximum r^2 after backtransformation of 15.1%, significantly lower ($P<.05$) than the empirical result of 59%.

The relationship examined in the previous paragraph was based on mean size and respiration measurements on plankton communities screened into three broad size classes. When this relationship between size and oxygen consumption rates is applied to the known finer size distribution of these same samples (39 size classes), and the finer size class predicted respiration is recombined to generate predictions for the three coarser size intervals, the mean and the median predictions performed equally well, explaining about 47% of the variation in the observations (Table 3.5, Figure 3.7). This suggests that the relationship between size and oxygen consumption is a real one, even at very fine scales of observation.

Conclusions

This paper has discussed three related problems encountered in the interpretation of log-transformed predictive regressions. The juxtaposition of theoretical results with a simulation and some empirical regressions from limnology justifies the

FIGURE 3.7. Observed respiration rates of 39 plankton assemblages (three size intervals from each of 13 lakes), compared with predicted rates from a regression between mean size and mean individual rate, combined with detailed data on the size composition of each size interval.


Mean
Median
300
100
30
10
3
10
30
100
300
Predicted respiration (mg O2 m-3 d-1)
Observed respiration (mg O2 m-3 d-1)
]

following conclusions:

1. If regression results are to be used to predict mean values of the response variable, then the appropriate correction factor must be used after backtransformation. However, in actual use the magnitude of this adjustment may be very small compared to the mean itself or the width of the confidence interval.

2. If the purpose of the regression is prediction on the scale of the original observations, then it is essential to give statistics to estimate the confidence interval on the backtransformed (observable) scale. When backtransformed to the original scale, 95% confidence intervals for predictions from quite strong regressions may still span a range as large as the mean value.

3. The theorem proved here states that, when logged random variables are multivariate normal, the correlation in the original scale can be no greater than the correlation between the logs. As variance increases, the original scale correlation becomes substantially less. The homoscedastic simulation data analyzed here usually gave a lower r after backtransformation. But this was not always true of the small empirical data sets with less perfect distributions. In the latter case, there is no general formula for predicting the

backtransformed r from the regression r. Hence, the former should be computed directly, and the implication that the proportion of the variation explained by the regression is indicative of expectations for the backtransformed scale should be avoided. Nevertheless, confidence limits on both r and the predictions are often so broad with ecological survey data, that the pretention that one can predict individual values with a "practical" level of precision should usually be avoided. One can be more confident in predictions of mean values. For example, one could predict the total oxygen consumption of small plankton from the measured phosphorus concentration within a 95% confidence range from about one half the predicted mean to about 170% of the mean (near the mid-point of the data distribution used to construct the regression). But predictions of the mean value in a series of lakes with that phosphorus level, or a series of samples from a single lake when the level was reasonably stable (or recurring), would be more precise (from 85% to 125% of the mean prediction).

CONCLUSION

This thesis has dealt with the distribution of biomass and oxygen uptake among planktonic organisms ranging in size from 0.2 μm to 1.5 mm ESD. Some general patterns were found. Thus, the abundance of organisms in logarithmic size intervals is strongly negatively correlated with size, and biomass per logarithmic size class tends to remain constant or increase slightly as size increases. The normalized spectrum is a standardized expression of the logarithmic decline in abundance in a given lake. All lakes are dominated by a strong phytoplankton mode, and lesser modes for unicellular plankton and zooplankton are usually observed. The mean respiration rate per organism and the ratio between respiration and biovolume are power functions of mean organism size. This collection of statements provides a general coherent description of part of the limnetic ecosystem.

However, this thesis deals more with prediction than with description alone. To what extent can these generalizations be used for predictions of abundance, biomass, or oxygen requirements of specific size intervals in specific lakes? There is a lot of "noise" around these general trends. The relationship between abundance and size applies to all lakes, although some lakes contain two orders of magnitude more plankton than others. Seasonal variation within one lake may be of the same order as variation among a large number of lakes. It thus appears that useful predictive models built upon many lakes, but intended for single lake predictions, would need to incorporate an ancillary variable, such as the phosphorus concentration, which is correlated with both biovolume and respiration rates. This is fortunate because phosphorus is easy to measure.

The broader promise of size-based models relies upon the laboratory-based correlations among physiological variables. Since this thesis shows that community respiration/size relationships are much as would be expected from lab results, then other physiological rates should be similarly predictable. For example, there is reason here to expect that one should be able to measure the phosphorus concentration a few times throughout the growing season, and hence predict the production of zooplankton.

The final chapter stands as a caveat to all of the rest. Coefficients of determination should always be verified on the scale of the original data. Backtransformations for predictions within a restricted range of the data should be appropriately corrected, depending upon the type of prediction desired. The operator should always be aware of both the assymetry and breadth of the confidence region about the backtransformed predictions.

BIBLIOGRAPHY

AMERICAN PUBLIC HEALTH ASSOCIATION, AMERICAN WATER WORKS ASSOCIATION, AND WATER POLLUTION CONTROL FEDERATION. 1971. Standard methods for the examination of water and wastewater. 13th ed. Washington, DC. 1193 p.

BAILEY-WATTS, A.E. 1986. Seasonal variation in size spectra of phytoplankton assemblages in Loch Leven, Scotland. Hydrobiologia 138: 25-42.

BANSE, K. 1976. Rates of growth, respiration, and photosynthesis of unicellular algae as related to cell size - a review. J. Phycol. 12: 135-140.

BANSE, K. 1979. On weight dependence of net growth efficiency and specific respiration rates among field populations of invertebrates. Oecologia (Berl.) 38: 111-126.

BANSE, K. 1982. Mass-scaled rates of respiration and intrinsic growth in very small invertebrates. Mar. Ecol.-Prog. Ser. 9: 281-297.

BANSE, K., AND S. MOSHER. 1980. Adult body mass and annual production/biomass relationships of field populations. Ecol. Monogr. 50: 355-379.

BASKERVILLE, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. Canadian Journal of Forest Research 2: 49-53.

BLUEWEISS, L., H. FOX, V. KUDZMA, D. NAKASHIMA, R. PETERS, AND S. SAMS. 1978. Relationships between body size and some life history parameters. Oecologia (Berl.) 37: 257-272.

BORGMANN, U. 1982. Particle-size conversion efficiency and total animal production in pelagic ecosystems. Can. J. Fish. Aquat. Sci. 39: 668-674.

BORGMANN, U. 1985. Predicting the effect of toxic substances on pelagic ecosystems. The Science of the Total Environment 44: 111-121.

CALDER, W.A. III. 1984. Size, Function, and Life History. Harvard University Press, Cambridge, Mass.

CLARKE, K.R. 1984. Parameter estimation for mixture distributions of species sizes, appendix p. 40-41. <u>In</u> R.M. Warwick. Species size distributions in marine benthic communities. Oecologia (Berl.) 61: 32-41.

CORNETT, R.J., AND R.H. RIGLER. 1986. Simple method of measuring seston respiration in oligotrophic lakes. Can. J. Fish. Aquat. Sci. 43: 1660-1663.

DEVOL, A.H. 1975. Biological oxidations in oxic and anoxic marine environments: rates and processes. PhD thesis, Univ. Washington, Seattle, Wash. 208 p.

DEVOL, A.H. 1979. Zooplankton respiration and its relation to plankton dynamics in two lakes of contrasting trophic state. Limnol. Oceanogr. 24: 893-905.

DEVOL, A.H. AND T.T. PACKARD. 1978. Productivity, chlorophyll, and respiration in Lake Washington. Limnol. Oceanogr. 23: 104-111.

DICKIE, L.M., S.R. KERR, AND P.R. BOUDREAU. 1987. Size-dependent processes underlying regularities in ecosystem structure. Ecol. Monogr. 57: 233-250.

DRAPER, N.R. AND H. SMITH. 1981. Applied Regression Analysis. 2nd. ed. John Wiley and Sons, New York, NY. 709 p.

ELSER, J.J., M.M. ELSER, AND S.R. CARPENTER. 1986. Size fractionation of algal chlorophyll, carbon fixation, and phosphatase activity: relationships with species-specific size distributions and zooplankton community structure.

J. Plankton Res. 8: 365-383.

FENCHEL, T. 1974. Intrinsic rate of natural increase: the relationship with body size. Oecologia 14: 317-326.

GELIN, C., AND W. RIPL. 1978. Nutrient decrease and response of various phytoplankton size fractions following the restoration of Lake Trummen, Sweden. Arch. Hydrobiol. 81: 339-367.

GESSNER, F. AND F. PANNIER. 1958. Influence of oxygen tension on respiration of phytoplankton. Limnol. Oceanogr. 3: 478-480.

GRIESBACH, S., R.H. PETERS, AND S. YOUAKIM. 1982. An allometric model for pesticide bioaccumulation. Can. J. Fish. Aquat. Sci. 39: 727-735.

HANSON, J.M. AND R.H. PETERS. 1984. Empirical prediction of zooplankton and profundal macrobenthos biomass in lakes. Can. J. Fish. Aquat. Sci. 41: 439-445.

HARRIS, G.P., B.B. PICCININ, AND J. VAN RYN. 1983. Physical variability and phytoplankton communities. V. Cell size, niche diversification, and the role of competition. Arch. Hydrobiol. 98: 215-239.

HEMMINGSEN, A.M. 1960. Energy metabolism as related to body size and respiratory area, and its evolution. Rep. Steno. Hosp. Copenh. 9: 1-110.

HOBBIE, J.E., R.J. DALEY, AND S. JASPERS. 1977. Use of Nuclepore filters for counting bacteria by fluorescence microscopy. Appl. Environ. Microbiol. 33: 1225-1228.

HOGG, R.V. AND A.T. CRAIG. 1978. Introduction to Mathematical Statistics. 4th ed. Macmillan, New York. 438 p.

HOLLANDER, M. AND D. WOLFE. 1973. Nonparametric Statistical Methods. John Wiley and Sons, New York.

HUMPHREYS, W.F. 1979. Production and respiration in animal populations. J. Animal Ecol. 48: 427-453.

HUTCHINSON, G.E. 1967. A Treatise in Limnology, vol. 2. John Wiley, New York, NY. 1115 p.

IKEDA, T. 1985. Metabolic rates of epipelagic marine zooplankton as a function of body mass and temperature. Mar. Biol. 85: 1-11.

JOHN, P.W.M. 1971. Statistical Design and Analysis of

Experiments. Macmillan, New York, NY. 356 p.

JORGENSEN, S.E. 1979. Modelling the distribution and effect of heavy metals in aquatic ecosystems. J. Ecol. Model. 6: 199-223.

KALFF, J. AND R. KNOECHEL. 1978. Phytoplankton and their dynamics in oligotrophic and eutrophic lakes. Ann. Rev. Ecol. Syst. 9: 475-495.

KERR, S.R. 1974. Theory of size distributions in ecological communities. J. Fish. Res. Board Can. 31: 1859-1862.

KLEKOWSKI, R.Z. 1981. Size dependence of metabolism in protozoans. Verh. int. Verein. Theor. Angewan. Limnol.21: 1498-1502.

KNOECHEL, R. AND L.B. HOLTBY. 1986. Construction and validation of a body-length-based model for the prediction of Cladoceran community filtering rates. Limnol. Oceanogr. 31: 1-16.

LAL, D. 1977. The oceanic microcosm of particles. Science 198: 997-1009.

LAMPERT, W. 1984. The measurement of respiration. In J.A.

Downing and F.H. Rigler [eds]. A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters. IBP Hdbk. 17, Blackwell.

LAVIGNE, D.M. 1982. Similarity in energy budgets of animal populations. J. An. Ecol. 51: 195-206.

LEHMAN, J.T. 1988. Ecological principles affecting community structure and secondary production by zooplankton in marine and freshwater environments. Limnol. Oceanogr. 33: 931-945.

LUND, J.W., C. KIPLING, AND E.D. LECREN. 1958. The inverted microscope method of estimating algal cell numbers and the statistical basis of estimation by counting. Hydrobiologia 11: 143-170.

MAZUMDER, A., D.J. MCQUEEN, W.D. TAYLOR, AND D.R.S. LEAN. 1988. Effects of fertilization and planktivorous fish (yellow perch) predation on size distribution of particulate phosphorus and assimilated phosphate: Large enclosure experiments. Limnol. Oceanogr. 33: 421-430.

MENZEL, D.W. AND N. CORWIN. 1965. The measurement of total phosphorus in seawater based on the liberation of organically bound fractions by persulfate oxidation.

Limnol. Oceanogr. 10: 280-282.

MULLIN, M.M., P.R. SLOAN, AND R.W. EPPLEY. 1966. Relationship between carbon content, cell volume, and area in phytoplankton. Limnol. Oceanogr. 11: 307-311.

NEELY, W.B. 1979. Estimating rate constants for the uptake and clearance of chemicals by fish. Environ. Sci. Technol. 13: 1506-1510.

PACE, M.L. 1984. Zooplankton community structure, but not biomass, influences the phosphorus-chlorophyll a relationship. Can. J. Fish. Aquat. Sci. 41: 1089-1096.

PACE, M.L. 1986. An empirical analysis of zooplankton community size structure across lake trophic gradients. Limnol. Oceanogr. 31: 45-55.

PACKARD, T.T. 1971. The measurement of respiratory electron transport activity in marine phytoplankton. J. Mar. Res. 29: 235-244.

PACKARD, T.T. 1985. Measurement of electron transport activity of microplankton. Adv. Aquat. Microbiol. 3: 207-261.

PARSONS, T.R., Y. MAITA, AND C.M. LALLI. 1984. A Manual of

Chemical and Biological Methods for Seawater Analysis. Pergamon Press.

PETERS, R.H. 1983a. The Ecological Implications of Body Size. Cambridge University Press, 329 p.

PETERS, R.H. 1983b. Size structure of the plankton community along the trophic gradient of Lake Memphremagog. Can. J. Fish. Aquat. Sci. 40: 1770-1778.

PETERS, R.H. 1986. The role of prediction in limnology. Limnol. Oceanogr. 31: 1143-1159.

PETERS, R.H. 1988. The relevance of allometric comparisons to growth, reproduction, and nutrition in primates and man. p. 1-19 _in_ K. Blaxter and I. Macdonald [eds] Comparative Nutrition. John Libbey, London.

PETERS, R.H. AND J.A. DOWNING. 1984. Empirical analysis of zooplankton filtering and feeding rates. Limnol. Oceanogr. 29: 763-784.

PLATT, T. AND K. DENMAN. 1978. The structure of pelagic marine ecosystems. Rapp. P.-v. Réun. Cons. int. Explor. Mer 173: 60-65.

PLATT, T. AND W. SILVERT. 1981. Ecology, physiology, allometry, and dimensionality. J. Theor. Biol. 93: 855-860.

PORTER, K.G. AND Y.S. FEIG. 1980. The use of DAPI for identifying and counting aquatic microflora. Limnol. Oceanogr. 25: 943-948.

ROBINSON, W.R., R.H. PETERS, AND J. ZIMMERMAN. 1983. The effects of body size and temperature on metabolic rate of organisms. Can. J. Zool. 61: 281-288.

SCHLESINGER, D.A., L.A. MOLOT, AND B. SHUTER. 1981. Specific growth rates of freshwater algae in relation to cell size and light intensity. Can. J. Fish. Aquat. Sci. 38: 1052-1058.

SCHWINGHAMER, P. 1981. Characteristic size distributions of integral benthic communities. Can. J. Fish. Aq. Sci. 38: 1255-1269.

SCHWINGHAMER, P. 1983. Generating ecological hypotheses from biomass spectra using causal analysis: a benthic example. Mar. Ecol. - Prog. Ser. 13: 151-166.

SCHWINGHAMER, P., B. HARGRAVE, D. PEER, AND C.M. HAWKINS. 1986. Partitioning of production and respiration among size groups of organisms in an intertidal benthic community. Mar. Ecol.-Progr. Ser. 31: 131-142.

SEBER, G.A.F. 1977. Linear Regression Analysis. John Wiley and Sons, New York, NY. 465p.

SHELDON, R.W., A. PRAKASH, AND W.H. SUTCLIFFE. 1972. The size distribution of particles in the ocean. Limnol. Oceanogr. 17: 327-340.

SHELDON, R.W., W.H. SUTCLIFFE, AND M.A. PARAJNAPE. 1977. Structure of pelagic food chains and relationship between plankton and fish production. J. Fish. Res. Board Can. 34: 2344-2353.

SHUTER, B.J. 1978. Size dependence of phosphorus and nitrogen subsistence quotas in unicellular microorganisms. Limnol. Oceanogr. 23: 1248-1255.

SILVERT, W. AND T. PLATT. 1980. Dynamic energy-flow model of the particle size distribution in pelagic ecosystems, p. 754-763. In W.C. Kerfoot [ed.] Evolution and Ecology of Zooplankton communities. Univ. Press of New England, Boston.

SIMON, M. 1987. Biomass and production of small and large freeliving and attached bacteria in Lake Constance. Limnol. Oceanogr. 32: 591-607.

SMITH, V.H. 1979. Nutrient dependence of primary productivity in lakes. Limnol. Oceanogr. 24: 1051-1064.

SNEDECOR, G.W. AND W.C. COCHRAN. 1967. Statistical Methods. 6th ed. Iowa State Univ. Press, Ames, Iowa. 593 p.

SOMMER, U., Z.M. GLIWICZ, W. LAMPERT, AND A. DUNCAN. 1986. The PEG-model of seasonal succession of planktonic events in fresh water. Archiv f. Hydrobiol. 106: 433-471.

SONDERGAARD, M., B. RIEMANN, L. MOLLER JENSEN, N.O.G. JORGENSEN, P.K. BJORNSEN, M. OLESEN, J.B. LARSEN, O. GEERTZ-HENSEN, J. HANSEN, K. CHRISTOFFERSEN, A.-M. JESPERSEN, F. ANDERSEN, AND S. BOSSELMANN. 1988. Pelagic food web processes in an oligotrophic lake. Hydrobiologia 164: 271-286.

SPRUGEL, D.G. 1983. Correcting for bias in log-transformed allometric equations. Ecology 64: 209-210.

SPRULES, W.G. AND M. MUNAWAR. 1986. Plankton size spectra in relation to ecosystem productivity, size, and perturbation. Can. J. Fish. Aq. Sci. 43: 1789-1794.

SPRULES, W.G., J.M. CASSELMAN, AND B.J. SHUTER. 1983. Size distribution of pelagic particles in lakes. Can. J. Fish. Aq. Sci. 40: 1761-1769.

STRAYER, D. 1986. The size structure of a lacustrine zoobenthic community. Oecologia (Berl.) 69: 513-516.

STRICKLAND, J.D.H. AND T.R. PARSONS. 1968. A Practical Handbook of Seawater Analysis. 2nd. ed. Bull. Fish. Res. Board Can. 167.

VEZINA, F. 1986. Body size and mass flow in freshwater plankton: models and tests. J. Plankton Res. 8: 939-956.

VINOGRADOV, A.P. 1953. The elementary chemical composition of marine organisms. Mem. no. 2, Sears Found. New Haven, CT.

WARWICK, R.M. 1984. Species size distributions in marine benthic communities. Oecologia (Berl.) 61: 32-41.

WELCH, H.E. 1968. Use of modified diurnal curves for the measurement of metabolism in standing water. Limnol. Oceanogr. 13: 679-687.

WELSH, A.H., A.T. PETERSON, AND S.A. ALTMANN. 1988. The fallacy of averages. Am. Nat. 132: 277-288.

WILLIAMS, P.J.L. 1984. A review of measurements of respiration rates of marine plankton populations, p. 357-389. In J.E. Hobbie and P.J.L. Williams [eds.] Heterotrophic Activity in the Sea. Plenum.

APPENDICES

APPENDIX 1. Analytical procedures used in Chapter I.

This appendix provides supplementary technical information about the normalized spectrum, estimation procedures for bimodal and trimodal distributions, and joint confidence regions.

1. The Normalized Spectrum

The normalized spectrum is a straight line least squares regression between the logarithms of body size and normalized abundance (Platt and Denman 1978). The latter is the biomass (or biovolume) density in a size interval divided by the nominal size of organisms in the interval. If logarithms are to base 2, then the size doubles between successive intervals, and nominal size is equal to the interval "width".

The spectrum is clearly a transformation of a histogram representation of biomass per log size class. Whereas the biomass in the system is equal to the <u>sum</u> of histogram ordinates, it may also be estimated by the integral of the normalized function over mass units. This is evident from the following considerations: using N_i, B_i, and m_i to represent the abundance, biomass, and size respectively of organisms in interval i,

$$N_i = B_i/m_i$$

or in the limit, $N = dB/dm$

The normalized spectrum corresponds to the model

$$\log_2 N = b_0 + b_1 \log_2 m$$

and hence $dB/dm = 2^{b0} m^{b1}$, approximately.

Therefore, $B = \int_{m1}^{m2} 2^{b0} m^{b1} dm = \left. \frac{2^{b0} m^{b1+1}}{b_1 + 1} \right|_{m_1}^{m_2}$ μg is an estimate of the biomass density over the range of sizes from m_1 to m_2 .

Hence, the two parameters of the normalized spectrum provide a standardized basis for comparing estimated biomasses between systems or between size ranges. Furthermore, each of these parameters may be interpreted in terms of observable system characteristics. The intercept b_0 is an estimate of log abundance in the size class corresponding to one size unit (volume or mass). In this paper, the spectra were fitted in volume units (μm^3). So the intercept size class corresponds to organisms of about 1 μm^3 (ESD = 1.24 μm).

The slope b_1 is interpretable as a measure of the rate of abundance change with size. For example, if the ordered abscissas are represented by m_i , then clearly $m_{i+1} = 2m_i$ and the normalized abundance in size class i+1 is

$$\log (B_{i+1}/2m_i) = b_0 + b_1 \log (2m_i)$$

Hence $$\frac{B_{i+1}}{2m_i} = 2^{b0} (2m_i)^{b1}$$

and $$B_{i+1} = 2^{b0+b1+1} m_i^{b1+1}$$

Similarly, $$B_i = 2^{b0} m_i^{b1+1}$$

Therefore, in general $$B_{i+1} = 2^{b1+1} B_i .$$

In particular, if $b_1 = -1$, then $B_{i+1} = B_i$, and biomass is uniformly distributed over logarithmic size intervals. This is Sheldon's hypothesis (Sheldon et al. 1972). Our slopes were almost always "shallower" than -1, and our modal slope of 0.9 implies that $B_{i+1} \doteq 1.07 B_i$. In other words, over the range of observed sizes, biomass increases by about 7% between successive log size intervals.

Our parameter estimates (Table 1.8) are based on calculations in μm^3. If we assume an organism mass density of 1 g cm^{-3} , then 1 $\mu g = 10^6 \mu m^3$. Hence our volumetric range from 2^{-8} to 2^{30} μm^3 corresponds to a mass range from 2^{-28} to 2^{10} μg. The slope does not change between scales. But if b_0^* is the volumetric intercept (Table 1.8), then the corresponding mass scale intercept is $b_0 = b_0^* + 19.932 (1+b_1)$, and the estimated biomass over the full range of observed sizes is

$$B = \int_{2^{-28}}^{2^{10}} 2^{b0} m^{b1} dm = \frac{2^{b0} m^{b1+1}}{b_1+1} \Bigg|_{2^{-28}}^{2^{10}} \mu g.$$

2. Bimodal Parameter Estimates

Clarke (1984) presented algorithms for maximum likelihood estimation of the five independent parameters of the mixed normal density function $\Phi(x) = p_1\Phi_1(x) + p_2\Phi_2(x)$ where Φ_i is distributed $N(\mu_i,\sigma_i^2)$ and $p_2 = |1-p_1|$ is clearly not independent of $0 \leq p_1 \leq 1$. The vector $\mathbf{V} = (\mu_1,\mu_2,\sigma_1,\sigma_2,p_1)$ is estimated by maximizing the log-likelihood function $L(\mathbf{V}) = \Sigma_i\ f_i\ \log[(2\pi)^{1/2}\Phi(x_i)]$ where f_i represents the observed biomass at size x_i, and the summation is over all size classes. After an initial estimate is obtained by inspection, it is improved by iterating the equation

$$\mathbf{V}_{i+1} = \mathbf{V}_i - [L''(\mathbf{V}_i)]^{-1}L'(\mathbf{V}_i)$$

where $L'(\mathbf{V})$ is the 5x1 vector of first derivatives of L with respect to $\mathbf{V}$, and $L''(\mathbf{V})$ is the 5x5 matrix of second derivatives. The distribution corresponding to the final parameter estimates may be compared with the observed distribution by a non-parametric test, such as the Kolmogoroff-Smirnoff, which we used.

3. Trimodal Parameter Estimates

The distribution of organism volume over log size classes was used to estimate the eight independent parameters of the trimodal mixed normal density function

$$f(x,\mathbf{V}) = p_1\Phi_1(x) + p_2\Phi_2(x) + p_3\Phi_3(x)$$

where $x = \log_2$ (size) is the abscissa,

the parameter vector $\mathbf{V} = (\mu_1,\mu_2,\mu_3,\sigma_1,\sigma_2,\sigma_3,p_1,p_2)$

and $\Phi_j(x)$ is distributed $N(\mu_j,\sigma_j^2)$.

Again, $p_3 = |1 - (p_1+p_2)|$ is not independent of $0 \le p_1+p_2 \le 1$. This function is nonlinear in the parameters μ_j and σ_j, and hence a linearization technique for nonlinear systems was used (Draper & Smith 1981):

$\mathbf{Y} = (y_1,\ldots,y_{39})'$ is the vector of observed biomasses at the 39 log sizes $\mathbf{X} = (x_1,\ldots,x_{39})'$, and

$$y_i = f(x_i,\mathbf{V}) + \epsilon_i$$

By analogy to the typical linear model, the error sum of squares is $S(\mathbf{V}) = \Sigma_i\,[\,y_i - f(x_i,\mathbf{V})]^2$

If $\mathbf{V}_0$ is an estimate of $\mathbf{V}$ close to $\mathbf{V}_0$, then the Taylor series expansion of f, truncated after the first derivative term, allows the following approximation

$$f(x_i,\mathbf{V}) \approx f(x_i,\mathbf{V}_0) + \sum_{j=1}^{8}\left[\frac{\delta f(x_i,\mathbf{V})}{\delta \mathbf{V}_j}\right]_{\mathbf{V}=\mathbf{V}0}(\mathbf{V}_j-\mathbf{V}_{j0})$$

Defining $\mathbf{Z}$ as the 39x8 matrix of partial derivatives and $\beta = \mathbf{V}-\mathbf{V}_0$, we have $y_i - f(x_i,\mathbf{V}_0) = \mathbf{Z}_i'\beta + \epsilon_i$

Then the least squares estimate of β is

$$\beta = (\mathbf{Z}'\mathbf{Z})^{1}\mathbf{Z}'(\mathbf{Y}-f(\mathbf{X},\mathbf{V}_0))$$

The solution to this equation results in a new improved estimate $\mathbf{V} = \mathbf{b} + \mathbf{V}_0$, which may be substituted for $\mathbf{V}_0$ in $f(X,\mathbf{V}_0)$ and the whole process is repeated until S converges to an acceptably low asymptotic level.

Most of the data sets tested in this paper did not support stable asymptotic convergence. This may have been due to one or more of high error sums of squares, insufficient data, or lack of true correspondence to a trimodal mixed normal distribution in the underlying population.

4. Joint Confidence Regions

The joint 100(1−α)% confidence region for the estimated parameters β of a linear model is defined by

$$(\beta-\mathbf{b})'\mathbf{X}'\mathbf{X}(\beta-\mathbf{b}) \leq ps^2F_{p,n-p,1-\alpha}$$

where n = number of observations

p = number of parameters in β

s^2 = estimated mean squared error

$\mathbf{X}$ = matrix of regressors

Any vector **b** satisfying this equation defines a point within the p-dimensional confidence ellipsoid.

If the model has only two parameters (a slope and an intercept), then the confidence region is an ellipse. In this case, a simple algebraic solution is available:

Given $$\mathbf{A} = \mathbf{X}'\mathbf{X} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and $L = ps^2F$,

then $L = a_{11}C_1^2 + 2a_{12}C_1C_2 + a_{22}C_2^2$ defines the contour of the confidence region. Substituting values of C_i falling within the single parameter confidence region for C_i, we may solve the resulting quadratic equation for the two corresponding values of C_j $(i \neq j)$, located on the confidence contour.

More generally, the dimensions and orientation of the p-dimensional confidence ellipsoid may be estimated from the eigenvalues ϵ_i of the positive definite matrix $\mathbf{A} = \mathbf{X'X}$, where

$$\sum_{i=1}^{p} \epsilon_i = \mathrm{tr}\mathbf{A} \qquad \text{and} \qquad \mathrm{Prod}\ \epsilon_i = |\mathbf{A}|$$

The eigenvalues determine the lengths of the half-axes, $l_i = \epsilon_i^{1/2}$, whereas the direction of these axes is determined by the corresponding eigenvectors. For a given eigenvalue ϵ_i, the eigenvector $\mathbf{x}_i$ is a solution to the equations $\mathbf{A}^*\mathbf{x}_i = 0$, where $\mathbf{A}^*$ is equivalent to $\mathbf{A}$ with ϵ_i subtracted from each value on the main diagonal. This method was used to estimate the contours of the confidence ellipsoids in this paper.

APPENDIX 2. The correlation between the exponents of bivariate normal variables.

Theorem: Let $(X_1,X_2)'$ be bivariate normal $N(\boldsymbol{\mu},\Sigma)$ where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Then the correlation coefficient between $Z_1 = \exp[X_1]$ and $Z_2 = \exp[X_2]$ is

$$r_z = \frac{\exp[r\sigma_1\sigma_2] - 1}{(\exp[\sigma_1^2] - 1)^{½}(\exp[\sigma_2^2] - 1)^{½}}$$

and $$|r_z| \leq |r| \quad .$$

Proof:

The moment generating function of **X** is

$$E(e^{t'X}) = \exp[\mathbf{t'}\mu + ½\mathbf{t'\Sigma t}]$$

so that $E(\mathbf{Z'}) = (\exp[\mu_1 + \sigma_1^2/2], \exp[\mu_2 + \sigma_2^2/2])$

$Y = X_1 + X_2$ is univariate normal since it is a linear combination of the elements of a bivariate normal variable, and

the variance of Y is

$$\mathbf{1'\Sigma 1} = \sigma_1^2 + \sigma_2^2 + 2r\sigma_1\sigma_2$$

Hence $$E(Z_1Z_2) = E(\exp[X_1 + X_2]) = E(e^Y)$$

$$= \exp[\mu_1 + \mu_2 + (\sigma_1^2 + \sigma_2^2)/2 + r\sigma_1\sigma_2]$$

Therefore $\mathrm{Cov}(Z_1,Z_2) = E(Z_1Z_2) - E(Z_1)E(Z_2)$

$$= E(Z_1Z_2) - \exp[\mu_1 + \mu_2 + (\sigma_1^2 + \sigma_2^2)/2]$$

$$= \exp[\mu_1 + \mu_2 + (\sigma_1^2 + \sigma_2^2)/2](\exp[r\sigma_1\sigma_2] - 1)$$

and $\qquad \mathrm{Var}(Z_1) = \exp[2\mu_1 + \sigma_1^2](\exp[\sigma_1^2] - 1)$

Hence the correlation between Z_1 and Z_2 is

$$r_Z = \frac{\exp[r\sigma_1\sigma_2] - 1}{(\exp[\sigma_1^2]-1)^{1/2}(\exp[\sigma_2^2]-1)^{1/2}}$$

$$= \frac{\exp[r\sigma_1\sigma_2] - 1}{\exp[\sigma_1\sigma_2] - 1} \cdot \frac{\exp[\sigma_1\sigma_2] - 1}{(\exp[\sigma_1^2]-1)^{1/2}(\exp[\sigma_2^2]-1)^{1/2}}$$

The last factor is the correlation when $r = 1$, and so this last factor lies between 0 and 1. Hence, r_Z is less than or equal to the first factor. So, to show that $|r_Z| \le |r|$, let $\sigma = (\sigma_1\sigma_2)^{1/2}$.

Then

$$r_Z \le \frac{\exp[r\sigma^2] - 1}{\exp[\sigma^2] - 1}$$

Consider the Taylor series expansions

$$\exp[r\sigma^2] - 1 = r\sigma^2 + \frac{r^2\sigma^4}{2!} + \frac{r^3\sigma^6}{3!} + \ldots.$$

$$(\exp[\sigma^2] - 1)\ r = r\ \sigma^2 + \frac{r\ \sigma^4}{2!} + \frac{r\ \sigma^6}{3!} + \ldots.$$

But $|r| \ge r^i$, $i \ge 1$.

Hence $\quad (\exp[\sigma^2] - 1)\ r \ge \exp[r\sigma^2] - 1 \qquad$ for $\quad r \le 1$

and $\quad |r_Z| \le |r|$

QED

APPENDIX 3. Morphometric information on the lakes sampled in southern Quebec, in order of increasing total phosphorus. Central, North, South, and Newport are basins of Lake Memphremagog.

Lake	Latitude	Longitude	Area km^2	Mean Depth m	Maximum Depth m
Bowker	45° 25'	72° 15'	2.3	24.0	59.0
Stukely	45° 20'	72° 15'	3.9	13.6	32.2
Orford	45° 15'	72° 20'	1.3	17.9	48.0
Lyster	45° 04'	71° 55'	1.7	20.0	50.0
Baldwin	45° 04'	71° 54'	0.3	3.0	7.6
O'Malley	45° 13'	72° 20'	0.2	4.0	10.4
Central	45° 10'	72° 15'	24.6	44.3	117.0
Cerises	45° 10'	72° 15'	2.0	2.0	3.0
North	45° 15'	72° 15'	20.4	14.7	33.5
Massawippi	45° 15'	72° 05'	17.9	40.2	85.7
Lovering	45° 10'	72° 10'	4.6	9.7	24.9
Argent	45° 15'	72° 25'	1.0	4.6	15.5
Trousers	45° 10'	72° 25'	2.0	5.2	10.1
South	45° 00'	72° 15'	56.1	6.2	12.8
Newport	44° 55'	72° 15'	2.0	3.4	9.5
Brome	45° 15'	72° 30'	14.5	5.8	12.8
Pond	45° 10'	72° 14'	0.2	0.9	1.5
Magog	45° 20'	72° 05'	10.8	9.8	19.2
Waterloo	45° 20'	72° 30'	1.5	2.9	4.9

Appendix 4. Chemical data for the samples collected in 1986. Chlorophyll and phosphorus data are in mg m^{-3}. Dry weight is in mg L^{-1}. Phosphorus fractions refer to all phosphorus passing through a screen or filter of the stated porosity (microns).

APPENDIX 5. Chemical data for the samples collected in 1987 Chlorophyll and phosphorus data are in mg m^{-3} Dry weight is in mg L^{-1} Phosphorus fractions refer to all phosphorus passing through a screen or filter of the stated porosity (μm)

Lake	Date	Chlor			Phosphorus					
			DW	SE_{DW}	Total	<120	<80	<40	<5	<0.4
South	114	3 879			16.87	17.67	16.62	16.27	11.96	
Central	114	0 824			12.11	9.91	10.56	11.61	4.55	
South	120	2.440			33.63	34.38	22.27	30 13	46.35	
Central	120	0 456			18 32	15.57	15 42	18.07	16.57	
North	120	0 666			29 98	25.33	24.27	24.37	24.67	
Magog	128	3.893			28.38	29.43	24 37	26.13	20.17	
Orford	128	1 828			12.51	11.41	7 66	6.41	3.90	
Lovering	128	1 731			18.27	17 82	14.36	15.67	12.56	
Newport	128	2 334			16 67	16.52	15.42	15.02	11.91	
South	128	4 036			18 02	15.82	15.22	15.57	11.31	
Central	128	1 358			10 56	16.62	16.02	13.71	10.86	
South	142	4 294			17 32	17.42	15.77	17 22	12 01	
Central	142	2 652	2 24	0 25	19 72	13.31	13.11	11.76	9.46	
North	142	4.522	2.20	0 38	17 27	16.07	7.36	6.71	6.66	
Waterloo	142	6 126	8 92	0.44	48 25	38.64	38.24	26.28	17.32	
Brome	142	3.154	3.11	0 26	20.82	17 57	20.67	17.92	14 01	
Argent	142	1 596			17.07	15.22	18.77	12 26	9.96	
Lyster	149	0 457	3.48	0 26	15 62	14.91	13 76	13 76	9.81	
Baldwin	149	1 438	2 68	0 15	16 02	15.82	14.86	16.77	12 76	
South	149	2 517	2 63	0 12	21.67	20.32	16.22	16.37	11.71	
South	166		1 40	0 17	22 37	14 51	13 11	11.41	13.16	
Newport	166		1 92	0 26	19 17	18.27	19.32	16.72	13 91	
Central	166		1 30	0 18	16 92	15.12	17.52	14.11	12 71	
North	166		1 22	0.20	16 82	13.96	15 47	14.26	10 46	
Lovering	177	1 106	2 20	0 11	14 01	11.81	12.41	10.91	9.46	
Massawippi	177	2 546	3 22	0 23	15 52	12 91	12 86	11 56	10 36	
Orford	177	0 885	1 52	0 07	5 86	4 80	5 16	4.05	4.15	
Omalley	177	[illegible] 008	3 63	0 01	14 41	13 96	13 01	11.26	8 06	
Magog	181	5 101	4 40	0 10	37 49	29.88	28.53	25.53	23 32	
Trouser	181	3 381	4 60	0 16	17 77	16 77	16 07	15 42	13 76	
Pond	183	3 310	2 08	0.24	27 03	18.62	24 32	22 42	20.52	
Bowker	183	1 385	2 02	0.12	3 50	5 81	5.21	4 90	4 20	
Waterloo	183	24 901	16.40	0 27	81 68	71 47	58.36	44 44	28 88	
Argent	183	[illegible] 442	2 13	0 21	16 07	15.27	14.51	14 06	11.61	
Newport	186	3 115	2 78	0 31	24 22	17.12	16 32	16.52	13.61	
South	186	[illegible] 419	2 62	0.08	17 62	13 81	14.31	13.21	10 81	
Central	186	1.041	1 08	0.09	16 67	14.81	10.76	9.91	8 86	
North	186	1.478	1 72	0.09	13 96	11.11	10.41	8.86	8.11	
Lyster	191	2 255	0 90	0 01	5 76	5.01	5.36	4.85	4 40	
Baldwin	191	[illegible] 884	2 55	0 09	13 61	13 21	12.61	12 16	10 51	
Central	192	1 371			17 97					
Pond	193	0 480	0 88	0 07	23 82	22 37	21 87	21 42	19 02	
Cerises	193	1 303	2 00	0.10	12 91	12 06	11 41	11.06	9 51	
Stukely	193	0 234	0 68	0 07	4 45	3 80	3 80	3 45	2.80	
Magog	193	1 998	2 42	0 11	25 23	24 12	21 32	20.57	17 42	
Waterloo	195	10 6[illegible]7	9 00	0 10	[illegible]6 82	59 16	53 80	52 75	[illegible]9 84	

Lake	Date	Chlor.	DW	SE_{DW}	Phosphorus Total	<120	<80	<40	<5	<0.4
Central	197	3.002	1.90	0.15	12.41	10.46	10.16	9.66	8.61	
North	197	1.841	1.33	0.22	14.91	12.66	12.21	11.86	10.36	
Central	198	2.472	2.03	0.13	13.51	9.76	9.36	9.16	8.56	
Orford	199	0.796	2.15	0.33	5.31	4.95	4.45	4.15	3.65	
Orford	203		2.97	0.23	7.86	7.46	6.76	6.56	5.71	3.7
Newport	204	4.056	3.83	0.40	22.72	19.02	18.37	17.87	16.47	11.[illegible]
Magog	223	9.039	4.78	0.21	39.34	36.64	35.04	29.53	24.12	14.5
Pond	224	3.834	1.08	0.19	22.92	23.02	21.42	21.82	20.12	15.8
Baldwin	225	1.976	2.30	0.41	13.06	13.71	12.51	12.01	10.76	7.2
North	226	1.568	4.17	0.21	12.21	12.26	10.81	10.26	9.71	7.9
Lyster	227	1.592	3.33	0.29	9.11	9.01	7.21	7.61	8.41	4.4
Central	228	1.195	1.47	0.12	13.46	13.06	12.91	11.66	10.31	7.8
Waterloo	229	28.862	32.96	0.24	130.13	100.00	94.09	57.31	34.63	19.[illegible]
Stukely	230	2.216	3.70	0.10	6.61	6.76	6.16	5.71	5.41	4.1
South	235	3.515	2.30	0.15	18.62	13.21	12.91	12.81	11.91	6.9
Cerises	236	6.343	2.50	0.32	16.57	16.27	15.42	15.47	12.21	8.0
Orford	237	1.623	2.57	0.12	6.51	5.86	5.56	5.31	[illegible].31	3.7

APPENDIX 6 Size distributions of the samples collected in 1986. The values in the first column are the base 2 logarithms of organism volume (μm^3), at the lower limit of the size interval Samples are identified by the name of the lake and the Julian date N = abundance (ml^{-1})
V - biovolume in the interval (μm^3 ml^{-1}) A zero value is an approximation, given the number of significant digits shown. Sampling and measuring procedures are described in the Methods section of Chapter I.

Lakes are in alphabetical order The basins of Lake Memphremagog are grouped under the name of the lake

	Lac d'Argent		Baldwin Pond		Lake Bowker		Lake Brome	
	214		212		230		214	
	N	V	N	V	N	V	N	V
6	556	53869	339	31837	625	58405	550	52988
7	460	81957	339	61849	523	93638	636	118473
8	182	66232	194	64496	352	119954	182	61376
9	193	143034	194	130179	239	156296	205	149210
10	267	393205	230	337925	205	290849	182	257373
11	246	647847	121	399587	80	198397	91	253909
12	86	467857	71	445031	39	217827	18	111095
13	27	346277	67	802301	32	396960	9	110575
14	18	418938	22	515682	9	222586	14	260747
15	14	579108	9	421568	11	540252	9	386196
16	5	556835	5	328135	9	930373	23	2094198
17	5	773381	0	0	7	1361956	5	981024
18	3	1038095	0	0	1	223225	2	825251
19	1	1073981	1	368894	0	0	3	1971080
20	1	986892	0	325375	0	0	3	5367199
21	2	7815208	0 4	1156253	0	0	2	6995798
22	0	0	0 5	3927336	0	0	0 6	4305701
23	0	0	0	0	0	0	0	33503
24	0	0	0	78656	0	50969	0	49944
25	0	317891	0	137725	0	56325	0	0
26	0	0	0	0	0	0	0	250029
27	0	0	0	0	0	296405	0	651809
28	0	0	0	0	0	663504	0	0

LAKE MEMPHREMAGOG

	North Basin		Central Basin		Central Basin		Central Basin	
	151		188		209		225	
	N	V	N	V	N	V	N	V
6	557	59522	568	54115	148	13834	499	48703
7	375	72076	76	17924	159	28815	484	81955
8	159	55955	189	85530	114	41798	132	53423
9	114	74498	114	74337	114	78379	103	77808
10	68	98436	38	52796	136	202579	176	240216
11	34	79931	38	118791	46	133569	103	276184
12	46	272471	141	828290	27	167446	46	260632
13	9	101616	101	1287212	32	346673	64	730273
14	23	499114	172	4606913	32	749496	9	188086
15	23	1002261	131	6000936	23	961369	32	1572599
16	14	1335138	71	7491141	18	2195129	14	1342194
17	5	718688	0	0	27	4948923	0	0
18	14	5183638	111	42835179	14	4713048	0	0
19	2	1482548	7	5401857	5	3459616	4	2738711
20	1	632233	2	3010000	0	0	1	919566
21	0	0	1	2810487	1	1737128	0	0
22	0.3	2267697	7	44454458	0	0	0.3	1970698
23	0	170551	0	2325864	0	73855	0	0
24	0	167082	0	217635	0	102712	0	0
25	0	392067	0	0	0	0	0	147660
26	0	316672	0	0	0	0	0	741918
27	0	0	0	0	0	0	0	508938
28	0	0	0	0	0	0	0	0

LAKE MEMPHREMAGOG (continued)

	South Basin		South Basin		Newport Basin		Newport Basin	
	149		195		225		255	
	N	V	N	V	N	V	N	V
6	1193	115171	487	47010	1406	134865	1337	127640
7	653	127954	414	72021	970	167886	1248	231663
8	256	89611	333	113879	584	213123	832	289291
9	426	315528	268	190882	564	413384	475	328740
10	199	317975	203	275098	396	550408	535	551830
11	85	247343	162	473217	218	618797	287	827797
12	55	320019	18	105717	91	551826	105	600909
13	77	942942	9	115288	50	664124	32	363799
14	109	2472189	17	418610	14	418313	9	182886
15	32	1535428	9	396909	14	557531	14	572203
16	59	5700159	9	942184	18	1729090	9	1035119
17	50	8790735	2	305512	9	1445780	27	5276695
18	27	11354493	1	320216	0	0	5	2375827
19	15	11613661	0	0	8	6327054	3	2006436
20	4	5101328	0	0	3	5626251	1	554346
21	1	1649021	0	1279923	0	0	0	0
22	4	26778838	3	19501767	0	0	0	0
23	0	0	0 6	5682860	1 1	10523486	0	8581
24	0	289529	0	257481	0	84446	0	0
25	0	0	0	283738	0	166479	0	36994
26	0	0	0	0	0	0	0	0
27	0	1227714	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0

	Lake Lovering		Lake Lyster		Lake Magog		Lake Magog	
	170		212		167		185	
	N	V	N	V	N	V	N	V
6	247	22925	261	23777	1165	108234	398	37788
7	364	70162	239	42228	767	135354	455	85885
8	156	52422	216	83036	256	82017	199	71041
9	39	28845	216	154761	199	155043	341	241317
10	78	97373	102	128453	142	204935	483	700709
11	65	179631	102	285709	85	204915	114	292422
12	68	404103	48	273379	171	1032183	62	364969
13	14	168783	34	411844	159	1904950	20	245946
14	14	373958	23	532725	141	3309124	58	1412733
15	9	482632	5	189819	77	3513377	34	14980[illegible]
16	14	1407898	9	1113669	25	2372738	12	1064200
17	18	3895806	18	3423287	23	3723574	4	661002
18	3	2285315	2	653904	16	5423234	12	4417833
19	25	17960478	3	2335495	20	13370695	3	2517898
20	2	2010802	0	347647	3	5049147	1	531590
21	1	3807031	0	0	3	11197149	0	0
22	0	0	0	0	0	0	0	0
23	0	0	0	0	0	94412	0	54733
24	0	161006	0	0	0	0	0	0
25	0	296968	0	0	0	0	0	0
26	0	336628	0	0	0	0	0	0
27	0	1212681	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0

	Lake Massawippi		Lake O'Malley		Lake Orford		Lake Waterloo	
	226		214		185		214	
	N	V	N	V	N	V	N	V
6	909	86327	558	52331	307	30379	1739	157502
7	568	105684	655	123560	250	45497	1217	205448
8	308	109828	485	181205	205	71966	1044	356212
9	308	211595	218	157932	102	66428	1261	947564
10	260	364268	279	373168	125	181569	239	1052702
11	138	421821	146	409082	91	238478	1124	3503783
12	64	372196	55	336180	55	316163	110	611178
13	91	1125684	41	515951	39	415610	12	173056
14	32	779049	59	1390775	18	451135	12	220832
15	5	250558	23	1007598	2	100230	550	28680687
16	14	1406935	9	752345	14	1418444	0	0
17	0	0	9	2068708	21	3329499	0	0
18	5	1435792	5	1527954	7	2378179	6	1976144
19	19	12511881	3	2202122	1	485210	9	8654789
20	0	0	0	259843	7	9111680	28	41544594
21	1	1597012	0	706322	0	0	0	0
22	1	3934277	0	1492304	0	0	0	0
23	0	0	0	0	0	0	0	83623
24	0	0	0	0	0	0	0	91670
25	0	0	0	101965	0	0	0	178762
26	0	0	0	772366	0	0	0	289885
27	0	0	0	0	0	3489000	0	462671
28	0	0	0	0	0	0	0	0

APPENDIX [illegible] Size distributions of the samples collected in 1987

The first column is the base 2 logarithm of organism size (μm^3), which defines the lower bound of each size interval. Each sample is identified by the name of the lake and the Julian date. N = abundance (ml^{-1}). V = biovolume ($\mu m^3\ ml^{-1}$). A zero entry is an approximate value, given the number of significant digits shown. A blank entry indicates that no organisms of this size were counted. Sampling and counting procedures are described in the Methods section of Chapter [illegible].

Lakes are listed in alphabetical order. All basins of Lake Memphremagog are grouped under the name of the lake.

	Lac d'Argent		Lac d'Argent		Lake Bowker		Lake Brome	
	144		183		183		144	
	N	V	N	V	N	V	N	V
-8	102663	565	336328	1850	146232	804	462697	2545
-7	102663	1129	336328	3700	146232	1609	462697	5090
-6	102663	2259	336328	7399	146232	3217	462697	10179
-5	82960	3650	298958	13154	97488	4289	365587	16086
-4	82960	7300	298958	26308	97488	8579	365587	32172
-3	49257	8718	142005	25135	57884	10245	298468	52829
-2	49257	17437	142005	50270	57884	20491	298468	105658
-1	25925	18329	93424	66051	33850	23932	142808	100965
0	152954	215665	150404	212070	70782	99803	159011	224206
1	117076	331325	86343	244351	27089	76662	60856	172222
2	64203	363389	65105	368494	18569	105101	37544	212499
3	22660	256375	37601	425418	8848	100106	22085	249870
4	15107	341871	13926	315145	4369	98870	10797	244336
5	5665	456398	7659	346646	1748	79114	5399	244359
6	2242 4	203108	2818 0	243567	692 0	60876	697 0	58590
7	1030 3	176244	3000 0	416228	667 0	112182	824 0	123383
8	1151 5	415342	1727 0	559951	690 0	224014	255 0	95496
9	582 0	412960	727 0	493702	370 0	251863	315 0	236341
10	424 2	589231	563 6		270 0	349532	230 0	325894
11	145 0	408221	472 7		120 0	317407	158 0	490118
12	91.0	430000	309 1		140 0	702041	61 0	321423
13	53.3	589683	90 9	1389739	30.0	336443	36 0	402901
14	25 5	624109	87 9	1836088	31 0	633564	12 0	301136
15	15.8	755004	67 2	2797477	12 4	572709	1 7	72755
16	2 4	245666	34 5	2968781	20 0	1982134	0 0	0
17	1.2	306138	10 3	189326	1 7	268082	0 0	0
18	1 2	467710	5 2	1571517	0 8	280268	0 9	394821
19			0.800	667227				
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								

	Baldwin Pond		Baldwin Pond		Baldwin Pond		Lake Hertel	
	150		191		225		277	
	N	V	N	V	N	V	N	V
-8	274151	1508	542349	2983	542349	2983	447909	2463
-7	274151	3016	542349	5966	542349	5966	447909	4927
-6	274151	6031	542349	11932	542349	11932	447909	9854
-5	182768	8042	428523	18855	428523	18855	364963	16058
-4	182768	16084	428523	37710	428523	37710	364963	32117
-3	120576	21342	254436	45035	254436	45035	236397	41842
-2	120576	42684	254436	90070	254436	90070	236397	83685
-1	63461	44867	167392	118346	167392	118346	103683	73304
0	87493	123365	113819	160485	69770	98376		
1	44647	126351	65340	184912	32043	90682		
2	33665	190544	35832	202809	21965	124322		
3	17823	201649	14227	160964	11628	131559		
4	6481	146665	7026	158998	4651	105252		
5	3240	146642	4216	190816	2067	93552		
6			2428.6	207729	921.0	77977	2400.0	219914
7			1642.9	283363	579.0	101822	1466.7	245191
8			928.6	325071	447.0	154312	1181.8	403643
9			571.4	401357	290.0	200314	654.5	423309
10			449.3	615690	233.3	317207	509.1	667431
11			362.3	959050	94.4	253003	345.4	972408
12	36.2	215510	260.9	1421031	116.7	655920	218.2	1133867
13	77.6	832237	159.4	1726508	78.9	850470	133.3	1469806
14	84.5	1839600	47.8	1042597	22.2	514197	70.1	1597378
15	36.2	1621436	26.6	1132282	8.9	402465	50.6	2325701
16	20.7	1680137	9.1	811071	2.9	240060	54.5	5488377
17	5.2	990980	3.6	569959	1.7	304172	18.2	3181476
18	5.2	1273503	1.2	402936	1.0	336010	19.5	6259897
19			0.000	0			3.900	2200364
20			0.000	0			2.600	3893809
21			0.600	2433213	0.001	3116	1.300	2620852
22					0.003	15580		
23					0.000	0		
24					0.000	0		
25					0.001	49855		
26					0.001	112174		
27					0.004	715109		
28					0.002	810925		
29					0.002	1885146		
30								

	Lac Cerises 193		Lac Cerises 236		Lake Lovering 129		Lake Lovering 177	
	N	V	N	V	N	V	N	V
-8	398348	2191	521762	2870	275393	1515	1147665	6312
-7	398348	4382	521762	5739	275393	3029	1147665	12624
-6	398348	8764	521762	11479	275393	6059	1147665	25249
-5	295073	12983	463789	20407	244794	10771	741925	32645
-4	295073	25966	463789	40813	244794	21542	741925	65289
-3	192719	34111	309796	54834	118920	21049	440518	77972
-2	192719	68223	309796	109668	118920	42098	440518	155943
-1	92210	65192	181167	128085	69544	49168	289814	204898
0	99022	139621	83479	117705	131229	185033	207140	292067
1	71057	201091	42598	120552	50224	142134	118914	336527
2	46760	264662	26281	148750	37870	214344	79702	451113
3	22692	256737	17005	192395	20049	226834	46031	520795
4	7335	165991	7558	171038	9721	219986	13639	308651
5	4584	207472	3779	171038	4860	219964	8524	385796
6	2214 0	196656	2313.0	199523	154.0	15394		
7	1429 0	232603	1375 0	222273	359.0	63942		
8	1071 0	322158	688.0	225859	177 0	67773	1236 0	398464
9	643 0	406258	531.0	366019	77 0	59185	600 0	397935
10	347 8	451782	250.0	339248	91.0	139716	491 0	666549
11	289.9	738683	281.0	738622	50.0	146580	255 0	726868
12	173.9	845528	166 7	823795	18 0	111713	164 0	776829
13	71.4	973981	52 1	593761	4.5	61871	36 4	416746
14	89.1	1837695	67.5	1407088	0.7	11775	57.9	1194313
15	32.5	1423617	29.9	1235964	4.5	250328	36 4	1990511
16	24.7	2098722	10.4	788928	0.7	65414	18 2	1218556
17	9.1	1726868	6.5	1036121			6.3	1062117
18	3.6	1462267	1.3	545913			1 6	595739
19	2.600	1699015	1 300	696318				
20	1.300	1782574					0 800	1378709
21			0 003	15179				
22	1.300	8703872	0 002	11875				
23			0.002	22054				
24			0 001	30536				
25							0 800	40271926
26								
27								
28								
29								
30								

	Lake Lyster		Lake Lyster		Lake Lyster		Lake Massawippi	
	150		191		227		177	
	N	V	N	V	N	V	N	V
-8	285507	1570	292969	1611	208798	1148	872428	4798
-7	285507	3141	292969	3223	208798	2297	872428	9597
-6	285507	6281	292969	6445	208798	4594	872428	19193
-5	304540	13400	358073	15755	185598	8166	704992	31020
-4	304540	26800	358073	31510	185598	16333	704992	62039
-3	180821	32005	193278	34210	110199	19505	334871	59272
-2	180821	64011	193278	68420	110199	39010	334871	118544
-1	79307	56070	101725	71920	48333	34171	220310	155759
0	127624	179950			81409	114787	295667	416890
1	87919	248811			41542	117564	190952	540394
2	36160	204666			22781	128940	93080	526833
3	17017	192530			12061	136458	67757	766603
4	7563	171151			6700	151621	21901	495620
5	4254	192536			3015	136459	13688	619519
6	1273.0	102556	1344.8	113864	758.0	68504	2818.0	247857
7	2273.0	346200	827.6	144746	727.0	122700	1273.0	223616
8	1177.0	386464	1034.5	350097	818.0	289302	1732.0	547602
9	455.0	276270	545.5	364361	505.6	350175	455.0	304490
10	182.0	217390	515.2	698658	359.6	493598	659.0	949785
11	182.0	452778	242.4	677187	292.1	755299	268.0	705350
12	182.0	1047958	137.9	835677	179.8	973304	97.6	429095
13	182.0	1842457	30.3	273160	151.5	1609794	48.8	510828
14	48.1	1014676	47.0	257282	73.0	1515424	81.0	1647300
15	18.2	759803	69.0	3126025	36.5	1619764	47.6	2157866
16	16.1	1370139	4.2	346070	17.4	1422934	6.3	474242
17	1.3	267386	1.2	162799	2.6	417742	2.8	459534
18					0.9	233115	1.4	406003
19							1.600	980415
20								
21					0.002	6786		
22					0.006	33929		
23					0.004	44108		
24					0.008	173039		
25					0.000	0		
26					0.000	0		
27					0.001	162860		
28					0.000	0		
29					0.001	678583		
30								

	Lake Magog		Lake Magog		Lake Magog		Lake Magog	
	129		181		193		223	
	N	V	N	V	N	V	N	V
-8	348520	1917	597574	3287	670160	3686	453952	2497
-7	348520	3834	597574	6573	670160	7372	453952	4993
-6	348520	7667	597574	13147	670160	14744	453952	9987
-5	309796	13631	482888	21247	536128	23590	605269	26632
-4	309796	27262	482888	42494	536128	47179	605269	53264
-3	183941	32558	229372	40599	353696	62604	299482	53008
-2	183941	65115	229372	81198	353696	125208	299482	106017
-1	121014	85557	150903	106688	186156	131612	157622	111439
0	115338	162627	133850	188729	85974	121223	48111	67837
1	49659	140535	51226	144970	49355	139675	30381	85978
2	36310	205515	31603	178873	24359	137872	18175	102871
3	12815	144989	14872	168262	11463	129692	6415	72579
4	7832	177238	7436	168277	5732	129715	4277	96789
5	3560	161126	4957	224354	2866	129715	1604	72597
6	614.6	56259	2214 0	187508	921 6	80275	2313 0	193453
7	593 7	105751	1929.0	329739	705.9	127797	1563 0	276392
8	312.5	108630	1896.0	605532	1236.4	404755	1375.0	460836
9	156.3	108673	792 0	557895	890.9	634021	750 0	527719
10	104.2	146026	417.0	568104	472 7	613999	563 0	848615
11	27.0	91515	214 0	582121	236.4	652008	290 0	815390
12	15.2	93626	125 0	541052	218 2	1181245	130 4	683804
13	24.2	299747	71.0	885438	127.3	1399658	29 0	330246
14	22.7	528840	85.5	1848361	54.5	1403133	44 3	947640
15	4 5	197714	187.5	7206534	61.7	2941083	36 5	1592066
16	0.0	0	20 0	1639305	18.2	1338584	6 1	519380
17	0.0	0	7 3	1123021	2 6	387903	4 3	716130
18	1.5	510230	1 8	609278	0.0	0	14 5	4468039
19					0.870	512853		
20					0.870	1208469		
21					0.005	14678	0 023	69439
22					0 014	78991	0 013	71149
23					0 018	207047	0 023	232625
24					0 018	428680	0 013	338464
25					0 014	731180	0 013	638574
26					0 016	1523372	0 003	206120
27					0 005	869286	0 003	488580
28					0 001	385141	0 003	844266
29								
30								

LAKE MEMPHREMAGOG

	Central Basin		Central Basin		Central Basin		Central Basin	
	130		143		166		186	
	N	V	N	V	N	V	N	V
-8	464850	2557	254136	1398	433101	2382	462903	2546
-7	464850	5113	254136	2795	433101	4764	462903	5092
-6	464850	10227	254136	5591	433101	9528	462903	10184
-5	550933	24241	301198	13253	279984	12319	336657	14813
-4	550933	48482	301198	26505	279984	24639	336657	29626
-3	218078	38600	163933	29016	249361	44137	244310	43243
-2	218078	77200	163933	58032	249361	88274	244310	86486
-1	143472	101435	78437	55455	109369	77324	116895	82645
0	58833	82955	49304	69519	107541	151633	59118	83356
1	32775	95583	28304	80100	67349	190598	38180	108049
2	20205	114360	19402	109815	36933	209041	20938	118509
3	8023	90772	12326	139456	17924	202792	13548	153282
4	4358	98622	5478	123967	8690	196655	5474	123877
5	1783	80699	2511	113648	4345	196655	3011	136278
6	327 0	30523	660.0	60602	1240.0	104313	889.0	75935
7	316.0	58247	400.0	74192	727 0	121097	389.0	71726
8	156 0	57718	320 0	121888	1255.0	416210	383.0	130198
9	84 0	60499	186 7	137141	436.0	309464	176.0	127834
10	51.0	81511	166 7	256387	418.0	556118	124.0	160389
11	55 0	150623	106.7	290996	291.0	730720	83.3	261768
12	12 6	68123	26 7	175053	164 0	874923	52 0	258359
13	17 4	206705	6 7		109 0	1390982	30 4	347835
14	5 5	119978	8 4	194389	83 1	1731546	39 2	816267
15	1.6	76411	7 0	340480	32 5	1423617	22.3	956611
16	1.6	147893	2 8	206872	24.7	2098722	10.1	866189
17					9 1	1726868	1.4	259841
18					2.6	1155888	1.4	597751
19					2.600	1699015	0 700	365172
20					1.300	1782574		
21					0 000	0	0 700	1494486
22					1.300	8703972		
23								
24							1 400	24344775
25								
26								
27								
28								
29								
30								

LAKE MEMPHREMAGOG (continued)

	Central Basin		Central Basin		Central Basin		Central Basin	
	197a		197b		198		228	
	N	V	N	V	N	V	N	V
-8	587467	3231	675836	3717	691572	3804	991650	5454
-7	587467	6462	675836	7434	691572	7607	991650	10908
-6	587467	12924	675836	14868	691572	15215	991650	21816
-5	580215	25529	533994	23496	558846	24589	881466	38785
-4	580215	51059	533994	46991	558846	49178	881466	77569
-3	275602	48782	317059	56119	331815	58731	380633	67372
-2	275602	97563	317059	112239	331815	117463	380633	134744
-1	181317	128191	208591	147474	174639	123470	250417	177045
0	60945	85932	46710	68681	66837	94240	57305	80800
1	29156	82511	34954	98920	28777	81439	24673	69825
2	15989	90498	19168	108491	15781	89320	13530	76580
3	8465	95773	10148	114814	8355	94528	7959	90048
4	3010	68116	3608	81649	4126	93371	3537	80042
5	1505	68116	2481	112290	2475	112019	1415	64043
6			1484.8	130046	1131.6	98765	954.5	76868
7			727.3	119329	421.1	70703	1000.0	172839
8			484.8	166243	342.1	114175	645.2	210519
9			454.5	330732	236.8	167309	532.3	579018
10			463.6	635495	236.8	332873	435.5	549271
11			242.4	651976	184.2	549258	193.5	488061
12			118.2	576922	54.5	321277	225.8	1103137
13			81.8	1005047	36.4	393789	112.9	1243124
14			121.1	2970684	23.5	517903	161.3	3539120
15			54.9	2240907	11.4	501144	96.8	4371472
16			7.0	598088	3.0	263208	32.3	2537594
17								
18								
19								
20								
21	0.004	13572					0.004	12215
22	0.011	67010					0.014	90082
23	0.004	43684					0.012	125199
24	0.008	192972					0.014	345060
25	0.007	317662					0.002	68707
26	0.007	760438					0.004	320631
27	0.011	2176556					0.010	1932945
28	0.002	1017875					0.008	2748263
29	0.002	1587885					0.002	1758888
30								

LAKE MEMPHREMAGOG (continued)

	Newport Basin		Newport Basin		Newport Basin		Newport Basin	
	130		166		186		204	
	N	V	N	V	N	V	N	V
-8	612629	3369	741272	4077	309727	1703	458658	2523
-7	612629	6739	741272	8154	309727	3407	458658	5045
-6	612629	13478	741272	16308	309727	6814	458658	10090
-5	435647	19168	658908	28992	250285	11013	339747	14949
-4	435647	38337	658908	57984	250285	22025	339747	29898
-3	387998	68676	391227	69247	133746	23673	242070	42846
-2	387998	137351	391227	138494	133746	47346	242070	85693
-1	170175	120314	228788	161753	78214	55297	106171	75063
0	143991	203027	147197	207548	123540	174191	54793	77258
1	92994	263173	63376	179354	59101	167256	28310	80117
2	62329	352782	34755	196713	29169	165097	20700	117162
3	26998	305455	22488	254429	13727	155307	8219	92990
4	14666	331892	8178	185068	8388	189820	4059	91855
5	7999	362035	4543	205616	3050	138043	1826	82645
6	400 0	35355	636 0	54495	1818.0	166992	390 0	34436
7	409 0	78099	727 0	122635	2000.0	379277	377 0	62840
8	212 0	74765	841.0	276336	6091.0	2057934	416.0	158462
9	150 0	104273	696.0	501077	3727 0	2715168	371.0	242591
10	119 0	181195	435 0	586215	2455 0	3124900	208.0	303490
11	36 0	96106	203 0	499498	1455.0	3636746	66 7	175511
12	7 8	46777	58.0	296251	546.0	3149966	52.4	282524
13	5 8	22131	53.2	644681	454 5	4109579	29.0	338683
14	5 0	132236	157.4	3316806	272.7	7491832	25.6	597685
15	5 0	315956	76 6	3525593	181.8	8270950	14.8	672945
16	1 9	191115	29 0	2018585	90 9	6692919	8 8	796675
17	1 3	188704	19.1	3175920			0.7	107978
18	0 6	284726	2 1	739223				
19			0.000	0			4 000	2456388
20			2.100	3212425				
21							0 008	24241
22							0 013	74602
23							0.021	218232
24							0 015	339688
25							0 018	945841
26							0 008	803652
27							0 003	380529
28								
29								
30								

LAKE MEMPHREMAGOG (continued)

	South Basin		South Basin		South Basin		South Basin	
	130		143		151		166	
	N	V	N	V	N	V	N	V
-8	450066	2475	198609	1092	539042	2965	708373	3896
-7	450066	4951	198609	2185	539042	5929	708373	7792
-6	450066	9901	198609	4369	539042	11859	708373	15584
-5	327320	14402	215773	9494	522708	22999	629665	27705
-4	327320	28804	215773	18988	522708	45998	629665	55411
-3	237535	42044	93175	16492	258631	45778	456944	80879
-2	237535	84087	93175	32984	258631	91555	456944	161758
-1	113653	80353	61299	43338	136122	96238	218634	154574
0	117400	165534	83125	117206	69398	97851	134772	190029
1	67396	190731	47720	135048	32596	92247	92843	262746
2	36959	209188	23790	134651	15889	89932	46671	264158
3	21741	245978	13854	156744	9463	107064	20216	228724
4	11595	262395	5038	114010	5608	126909	10981	248500
5	3865	174930	2799	126683	2570	116318	4993	225938
6			786.0	69577	2363 6	186687		
7			452 0	81657	1454 5	233260		
8			371 0	138321	1145.5	372875		
9			242.0	166906	654.5	492221		
10			210 0	347999	690.9	927805		
11			56.0	164304	90.9	222843		
12			56 0	294536	90.9	534995		
13			32.0	417928	72 7	893811		
14			7.6	148817	90 9	1949690		
15			8 0	491771	4 5	209592		
16			4 5	344022	0 0	0		
17					0 0	0		
18					18 2	6092781		
19					0 000	0		
20					2 273	2388370		
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								

LAKE MEMPHREMAGOG (continued)

	South Basin		South Basin		North Basin		North Basin	
	186		235		143		166	
	N	V	N	V	N	V	N	V
-8	775158	4263	1000658	5504	515098	2833	311147	1711
-7	775158	8527	1000658	11007	515098	5666	311147	3423
-6	775158	17053	1000658	22014	515098	11332	311147	6845
-5	612471	26949	741228	32614	416241	18315	345718	15212
-4	612471	53897	741228	65228	416241	36629	345718	30423
-3	363654	64367	484114	85688	271857	48119	225797	39966
-2	363654	128734	484114	171376	271857	96237	225797	79932
-1	239246	169147	231634	163765	130075	91963	108037	76382
0	149958	211441	52376	73850	73410	103508	56640	79862
1	68869	194899	27061	76583	30649	86737	35767	101221
2	51930	293924	16489	93328	18909	107025	14265	80740
3	27492	311044	6983	79006	12235	138427	8496	96124
4	11108	251374	4268	96585	4943	111860	4196	94955
5	5554	251374	2328	105365	1977	89479	2308	104460
6	1315 8	108561	2750 0	242216	513 0	52785	1357 0	120639
7	1105 3	187877	1750.0	304442	333 0	53059	893 0	155641
8	1345 5	450005	1188.0	405209	205 0	63376	691 0	233228
9	909 1	656727	1537 0	1055269	128 0	88351	250 0	186662
10	454 5	583231	1268 0	1727005	72 0	116653	200.0	260629
11	200 0	513621	561.0	1443803	42.0	133315	250.0	667219
12	272 7	1235617	125.0	611630	19.3	95978	54.5	297937
13	105 3	1056978	73.2	772371	21 6	265833	63.6	645225
14	50 9	1224439	73.2	1818547	17.0	385070	18.2	505092
15	82 4	3878189	66.7	3254750	3.4	144034	27.3	1264861
16	18 2	1589727	47 3	3802734	2 3	164712	4 7	411285
17	7 3	1303739	9 7	1403829			3.4	576150
18	1 2	8490	24 4	9494038			0 8	318064
19	0 002	1367	0 001	660			0 400	327151
20	0 009	16798	0 000	0				
21	0 023	4198	0 003	11482				
22	1 200	5615107	0.008	50690				
23	0 131	1519752	0 012	151777				
24	0 142	3406180	0 008	181092				
25	0 070	3274810	0 009	403127				
26	0 077	6996824	0 003	285819				
27	0 021	3850045	0 000	0				
28	0 005	1460743	0 002	827533				
29			0 001	580189				
30								

LAKE MEMPHREMAGOG (continued)

	North Basin		North Basin		North Basin		Lake O'Malley	
	186		197		226		177	
	N	V	N	V	N	V	N	V
-8	763584	4200	816585	4491	379652	2088	443491	2439
-7	763584	8399	816585	8982	379652	4176	443491	4878
-6	763584	16799	816585	17965	379652	8352	443491	9757
-5	848427	37331	967804	42583	337468	14849	328512	14455
-4	848427	74662	967804	85167	337468	29697	328512	28909
-3	403003	71332	430975	76283	275511	48765	175548	31072
-2	403003	142663	430975	152565	275511	97531	175548	62144
-1	265133	187449	252032	178187	131823	93199	102660	72581
0	86282	121658	70508	99416	60273	84985	141244	199154
1	45029	127432	44975	127279	25164	71214	89193	252416
2	19755	111813	27130	153556	17250	97635	40019	226508
3	15688	177494	15669	177279	9132	103319	25895	292976
4	5810	131480	5223	118196	3653	82667	8370	189413
5	2905	131480	3192	144470	2435	110208	5754	260426
6	1136.4	98005	634.1	51493	1591 0	135201	2625 0	232113
7	590.9	109475	292 7	49040	1455 0	253907	2125 0	411775
8	831.3	276558	878 0	289535	1136 0	376943	3000 0	1073940
9	241.0	171028	1414 6	935454	463 4	324598	1294 0	1009674
10	227.3	283095	658.5	910956	353 7	500175	529 0	709875
11	325 3	901750	473 7	1335559	158 5	411216	1000 0	2639188
12	144.6	740860	414 6	2312210	122 0	607272	1000 0	5885503
13	24.1	289077	210 5	1914051	61 0	610132	86 8	970128
14	47.3	1108321	76.8	1615604	61.0	1171350	113 2	2396510
15	78 5	3752953	42 7	1806752	39 1	1758165	42 1	1832837
16	29 0	2483907	21 9	1785258	20 0	1584250	26 3	2246248
17	3 2	627255	4 9	795093	0 9	119355	10 5	1896330
18	1 1	345913					5 3	1568987
19								
20					0 002	2777	2 600	2844495
21					0 001	2314		
22					0 007	43072	2 600	12399514
23					0 001	9643		
24					0 004	83239		
25					0 003	151717		
26					0 001	92573		
27								
28								
29								
30								

	Lake Orford		Lake Orford		Lake Orford		Lake Orford	
	129		177		199		203	
	N	V	N	V	N	V	N	V
-8	202849	1116	30332	167	234617	1290	234617	1290
-7	202849	2231	30332	334	234617	2581	234617	2581
-6	202849	4463	30332	667	234617	5162	234617	5162
-5	160276	7052	26962	1186	208549	9176	208549	9176
-4	160276	14104	26962	2373	208549	18352	208549	18352
-3	142746	25266	18010	3188	103188	18264	103188	18264
-2	142746	50532	18010	6376	103188	36529	103188	36529
-1	62608	44264	10532	7446	54310	38397	54310	38397
0	81121	114381	96730	136389	46981	66243	66508	93776
1	37255	105432	66636	188580	32964	93288	34999	99047
2	25538	144545	36542	206828	14790	83711	17448	98756
3	10816	122372	17734	200642	10440	118118	10161	114962
4	5408	122383	5732	129715	3867	87510	3695	83618
5	3305	149584	3224	145918	1547	70017	1847	83595
6	312.0	29389	1167.0	53920	1090.9	95796	1318.0	121470
7	425.0	77805	592.0	101855	757.6	132528	772.7	126303
8	173.0	64207	829.0	279737	606.1	209802	590.9	229900
9	67.0	48574	421.0	292499	292.9	207914	363.6	261694
10	67.0	93646	224.0	312090	285.7	378618	467.8	732688
11	17.0	42732	158.0	423729	192.9	513666	455.0	1254001
12	6.0	43952	100.0	507569	185.7	967932	136.4	854803
13	9.9	122190	26.0	269176	57.1	599595	90.9	919201
14	12.9	292159	34.6	688387	50.0	981433	45.5	1425496
15	2.3	125635	5.1	219978			8.4	403659
16	0.6	63820	3.3	243278			10.4	693310
17	0.0	0	13.0	2031782			1.4	252171
18	0.6	174125	0.9	253551			0.7	196242
19			0.5	250544			0.000	35
20								
21					0.001	3393	0.000	555
22					0.002	13572	0.001	3573
23					0.002	23750	0.002	
24					0.002	40715	0.001	25847
25					0.000	0	0.001	
26					0.002	169646	0.000	36151
27					0.000	0	0.000	16029
28					0.001	343533		
29								
30								

	Lake Orford		Pond		Pond		Pond	
	237		183		193		224	
	N	V	N	V	N	V	N	V
-8	246793	1357	534275	2939	593210	3263	593210	3263
-7	246793	2715	534275	5877	593210	6525	593210	6525
-6	246793	5429	534275	11754	593210	13051	593210	13051
-5	219372	9652	522402	22986	632757	27841	632757	27841
-4	219372	19305	522402	45971	632757	55683	632757	55683
-3	86835	15370	253780	44919	375700	66499	375700	66499
-2	86835	30740	253780	89838	375700	132998	375700	132998
-1	57128	40389	148410	104926	164781	116500	164781	116500
0	50524	71239	72031	101564	59928	84498	43058	60712
1	29005	82084	27567	78015	51605	146042	30211	85497
2	15906	90028	15118	85568	21225	120134	16568	93775
3	9473	107178	10004	113185	11237	127135	7176	81189
4	4678	105863	4891	110683	4994	113014	3544	80201
5	2573	116454	2445	110661	3052	138134	1595	72190
6	678 6	57169	1393 9	117593	590 0	51695	4750 0	382182
7	975 6	169245	939.4	160021	487 0	87216	3750 0	586666
8	1439 0	473278	1167 0	354124	461.5	155571	500 0	157548
9	878 0	622393	145 8	91019	409 1	299047	140 5	101132
10	463 4	651202	114.6	156337	409 1	577443	86 5	113851
11	341 5	873709	60 6	133878	307 7	816565	54 1	146043
12	195.1	975068	60.6	329498	109 1	599530	54 1	307570
13	122 0	1234160	17 8	195066	76 9	729294	24 7	313919
14	124 2	2560593	30 2	604867	55 5	1216607	22 7	506675
15	73 2	2697170	5 5	203255	51 8	2141487	10 8	455921
16	27 3	2116226	1 1	91245	29 1	2446861	8 1	581449
17	3.0	467926	0 0	0	8.2	1431072	7 1	1285505
18	4 5	1228305	0 4	168453	3 6	1193210	4 5	1590885
19	1.500	935851			0.000	0	1 000	725660
20					0 900	1371702	0 500	566922
21					0 014	43898	2 000	5368710
22					0 000	1585	0 003	19849
23	0 002	18044			0 025	564214	1 000	11526150
24	0 000	0			0 022	242786	0 012	281697
25	0 001	43375			0 002	74231	0 009	432851
26					0 001	117977		
27					0 000	82310	0 004	821425
28					0 001	471482	0 001	268719
29					0 001	209710	0 001	714548
30					0 000	392403		

	Lake Stukely		Lake Stukely		Lake Trousers	
	193		230		181	
	N	V	N	V	N	V
-8	376595	2071	376595	2071	134083	737
-7	376595	4143	376595	4143	134083	1475
-6	376595	8285	376595	8285	134083	2950
-5	223168	9819	223168	9819	143022	6293
-4	223168	19639	223168	19639	143022	12586
-3	132506	23454	132506	23454	63689	11273
-2	132506	46907	132506	46907	63689	22546
-1	87175	61633	87175	61633	37245	26332
0	59784	84295	47423	66866	146482	206540
1	22880	64750	21780	61637	77084	218148
2	18821	106527	17915	101399	30743	174005
3	6643	[illegible]	9485	107313	20345	230183
4	4428	100206	3162	71556	8138	184163
5	1845	83505	1756	79477	4521	204620
6	393.9	33746	727.0	65803	3364.0	288827
7	257.6	[illegible]	727.0	138120	2909.0	477529
8	363.6	127445	1546.0	591277	2429.0	785112
9	223.5	165876	1102.0	713460	1429.0	1114668
10	147.1	200470	510.0	733122	1071.0	1451974
11	70.6	194568	346.9	928675	727.0	2240974
12	53.0	295088	508.2	2246216	393.0	1986682
13	47.1	500289	81.6	712606	214.0	2365442
14	34.4	706999	90.9	1924279	138.8	2955203
15	15.8	695495	6.1	259875	59.2	2939428
16	6.1	543111	3.6	304832	35.7	2393593
17	1.2	165842	0.9	175472	2.0	280119
18	0.0	0	0.6	181159		
19	0.000	0	0.200	148945		
20	0.000	129	0.000	0		
21	0.000	1057	0.001	3214		
22	0.400	1875491	0.200	1169814		
23	0.001	10058	0.005	54644		
24	0.000	8078	0.002	54644		
25	0.000	2185	0.001	43394		
26	0.000	13726	0.000	0		
27	0.001	109633	0.001	144645		
28	0.000	174791				
29	0.000	138380				
30	0.000	136866				

	Lake Waterloo		Lake Waterloo		Lake Waterloo		Lake Waterloo	
	144		183		195		229	
	N	V	N	V	N	V	N	V
-8	373291	2053	662835	3646	368092	2025	1469619	8083
-7	373291	4106	662835	7291	368092	4049	1469619	16166
-6	373291	8212	662835	14582	368092	8098	1469619	32332
-5	368683	16222	707024	31109	392632	17276	1175695	51731
-4	368683	32444	707024	62218	392632	34552	1175695	103461
-3	262686	46495	349830	61920	194271	34386	620506	109830
-2	262686	92991	349830	123840	194271	68772	620506	219659
-1	115213	81456	184121	130174	102248	72289	408228	288617
0	205677	290005	215027	303188	332338	468597	322454	454660
1	177111	501224	100997	285822	152629	431940	205681	582077
2	80938	458109	67694	383148	83700	473742	112293	638408
3	51419	581755	32580	368610	49851	564014	47771	540481
4	22853	517163	15928	360451	19694	445675	29193	660638
5	9522	430966	5792	262146	12309	557105	15924	720720
6	2667 0	238802	7636 0	666218	9091 0	830739	5909 0	519600
7	1424 0	258556	3818 0	614344	5455 0	996906	5227 0	863898
8	1000.0	372875	2909 0	957687	3263 0	1095248	4546 0	1532514
9	576.0	415905	2364 0	1645147	2579.0	1907884	2471 0	1751108
10	518.0	784656	1273 0	1696035	1842 0	2741230	2177 0	3031799
11	82.0	250772	909 0	2510420	1212 0	3748606	1823 5	4645036
12	82 0	457492	182.0	1146418	210 5	975326	941 2	5996372
13	47.0	510832	40 0	347435	157 9	1547646	529 4	5598189
14	35 0	870656			235 7	5120373	411 8	10776516
15	24 0	1152975			157 9	5672058	823 5	37373833
16	12.0	1459234			28 6	2136999	529 4	45274376
17					7 1	1225519	58 8	8426853
18							10 0	2680824
19							[illegible] 000	4203531
20							[illegible] 000	6176617
21							10 000	24663576
22							[illegible] 000	21618160
23							0 087	928811
24							0 073	1781282
25							0 013	473317
26								
27								
28								
29								
30								

APPENDIX 8 Principal original APL functions used

VOLUMES - computes the volumes of particles Input data include the shape, length, width and height

DVOLS - computes the volumes of particles composed of several identical subunits Input data include the shape, length, width, height, and number of subunits

INTERVALLS computes the frequency distribution of the output from the two preceding functions The user may specify whether logarithmic or linear units are desired, and the base of the logs Output includes both number and total volume for each interval

KOLSMIRBAT performs a series of Kolmogoroff-Smirnoff tests on batch input

MIXDIST computes the parameters of the best bimodal fit.

TRIMODAL computes the parameters of the best trimodal fit

CONFELLIPSE computes 100 points on the contour of a two parameter confidence region, using the statistics of the regression used to estimate the parameters

MLINREG versatile linear regression function, with detailed diagnostic statistics and hypothesis testing Suitable for all multiple linear models.

ERRORSUM - optional routine called up by MLINREG. When repeat observations are available to compute the F ratio for lack of fit

```
S MLINREG T;P;RSS;DF;TSS;SSREG;R2;SSB;TS;F1;ANS;EST;YM;YSD;A;G;C;
⎕TCFF
Q←⌹((⍉S)+.×S)
B←⍉⍉(Q+.×(⍉S)+.×T)       ⍝VECTOR OF REGRESSION COEFFICIENTS
N←1↑⍴S ⋄ P←(¯1)↑⍴S
YHAT←S+.×B
RESIDUALS←T-YHAT
RSS←,RSS←(⍉RESIDUALS)+.×RESIDUALS
S2←RSS-DF←N-P
TSS←+/,TSS←(T-YM←(+/,T)-N)*2
SSREG←TSS-RSS
R2←SSREG-TSS
TS←TSS+SSB←N×YM×YM
F1←SSREG-((P-1)×,S2)
'DO YOU WANT TO TEST FOR LACK OF FIT?'
ANS←⍞
→(ANS='Y')/LOF
NORMAL: ⎕TCFF
'                               ANOVA '
'_____________________________________________________________________      '
' SOURCE         DF            SS            MS          F RATIO'
'_____________________________________________________________________      '
⎕←'REGRESSION     ',(5 0⍕(P-1)),(15 4⍕SSREG)
⎕←'RESIDUAL       ',(5 0⍕DF),(15 4⍕RSS),(13 4⍕S2),(10 2⍕,F1)
'_____________________________________________________________________      '
→(ANS≠'Y')/TAB2
⎕←'LACK OF FIT    ',(5 0⍕LDF),(15 4⍕LSS),(13 4⍕LI)
⎕←'PURE ERROR     ',(5 0⍕TOTDF),(15 4⍕TT),(13 4⍕PEMS),(10 2⍕F)
'_____________________________________________________________________      '
TAB2:
⎕←'TOTAL (CORR.) ',(5 0⍕(DF+1)),(15 4⍕TSS)
⎕←'INTERCEPT      ',(5 0⍕1),(15 4⍕SSB)
⎕←'TOTAL          ',(5 0⍕(DF+2)),(15 4⍕TS)
⎕TCLF
⎕←'MEAN RESPONSE = ',(3⍕YM),'    R2 = ',(5 3⍕,R2)
⎕TCLF
⎕←'REGRESSION PARAMETERS:' ⋄ ⎕TCLF
⎕←'     ',(⍕,B) ⋄ ⎕TCLF
'ANY KEY TO CONTINUE' ⋄ ANS←⍞
⎕TCFF
⎕←'COVARIANCE MATRIX OF PARAMETERS:' ⋄ ⎕TCLF
⎕←(12 6⍕Cm←Q×S2)
⎕TCLF
⎕←'CORRELATION MATRIX:' ⋄ ⎕TCLF
X←(-G⍴(+⌿A*2)*.5)×A←S-(G←N,P-1)⍴+⌿(S←S[;1+⍳(P-1)])-N
 X←X,[2](T-(N,1)⍴YM)-TSS*0.5
```

```
                                      ⍝ X IS CENTRED/SCALED REGRESSORS AND
                                      ⍝                        RESPONSE.
 ⎕←8 3⍕CORRMAT←(⍉X)+.×X
 ⎕TCLF
 ⎕←'VARIANCE INFLATIONARY FACTORS (DIAG OF ⌹CORRMAT):'
 ⎕TCLF
 ⎕←8 3⍕1 1⍉⌹CORRMAT
⎕TCLF ⋄ 'ANY KEY TO CONTINUE' ⋄ ANS←⍞
⎕TCFF
'TYPE A VECTOR OF VALUES FOR THE PREDICTOR VARIABLES'
'STARTING WITH 1, TO OBTAIN THE ESTIMATED RESPONSE'
'WITH ITS STANDARD DEVIATION.  TYPE 999 TO BYPASS.'
ESTIMATE: EST←⎕
→(EST=999)/CONT
YH←(⍉EST)+.×B
YSD←((⍉EST)+.×Q+.×EST×S2)*0.5
⎕←' EST. Y = ',(⍕YH),'     SD = ',(⍕YSD)
'REPEAT FOR ANOTHER ESTIMATE.' ⋄ →ESTIMATE
CONT:
'DO YOU WANT TO TEST FOR SIGNIFICANCE OF PARAMETERS (Y/N)?'
ANS←⍞
→(ANS='N')/FIN
AGAIN: 'ENTER SELECTION VECTOR OR MATRIX'
A←⎕
A←(G←((⍴A)÷P),P)⍴A
'ENTER HYPOTHETICAL VALUES'
C←⎕
C←(G←G[1],1)⍴C
F2←(⍉((A+.×B)-C))+.×(⌹(A+.×Q+.×⍉A))+.×((A+.×B)-C)÷(S2×G[1])
'F FOR HYPOTHESIS = ',(⍕,F2) ⋄ ⎕TCLF
'ANOTHER PARAMETER TEST (Y/N)?'
ANS←⍞
→(ANS='Y')/AGAIN
→FIN
LOF: ERRORSUM
→NORMAL
FIN:
```

```
ERRORSUM
' '
'THE FUNCTION 'ERRORSUM', CALLED UP FROM MLINREG, COMPUTES THE '
'ERROR SUM OF SQUARES FROM REPEAT OBSERVATIONS, AND HENCE THE '
'F RATIO FOR LACK OF FIT.  DRAPER AND SMITH, P. 33'
' '
X←S ◊ Y←T
⍞←'PERCENTAGE TOLERANCE (+/-) FOR X INTERVALS?' ◊ PCT←⎕
I1←(100-PCT)-100 ◊ I2←(100+PCT)-100
I←1 ◊ R←J←2 ◊ SIG←Y[1;1]
SS←SIG*2 ◊ TOT←1 ◊ TOTDF←TOTERRSQ←0 ◊ VEC←⍳0
NEXT: →((X[J;R]≥(X[I;R]×I1))∧(X[J;R]≤(X[I;R]×I2)))/AUG
⍞←(⍕UPTO←N-(⍴VEC)),' REMAINING'
R←2
SKIP: J←J+1
→(J∊VEC)/SKIP
→(J≤N)/NEXT
→COMPUTE
INC: I←I+1
→(I∊VEC)/INC
→(I≥N)/FINALE
J←I+1
SIG←Y[I;1]
SS←Y[I;1]*2
TOT←1
→(J≤N)/NEXT
→FINALE
AUG: R←R+1
→(R≤P)/NEXT
VEC←VEC,J
SIG←SIG+Y[J;1]
SS←SS+Y[J;1]*2
TOT←TOT+1 ◊ R←2
→SKIP
COMPUTE: PARTERRSQ←SS-(SIG*2)-TOT
TT←TOTERRSQ←TOTERRSQ+PARTERRSQ
TOTDF←TOTDF+TOT-1
→INC
FINALE: ' '
→(TOTDF=0)/FIN1
  PEMS←TOTERRSQ-TOTDF
LF←((LSS←RSS-TOTERRSQ)-LDF←DF-TOTDF)
F←LF-PEMS
→
FIN1: 'NO REPEAT OBSERVATIONS AVAILABLE'
```