

MUL 09 1997

ORNL/TM-9838

ornl

**OAK RIDGE
NATIONAL
LABORATORY**

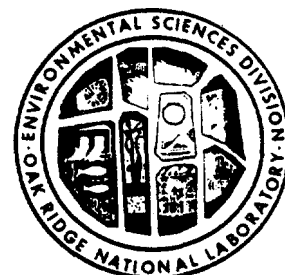
MARTIN MARIETTA

**Empirical Relationships Between
Watershed Attributes and
Headwater Lake Chemistry in the
Adirondack Region**

C. T. Hunsaker
S. W. Christensen
J. J. Beauchamp
R. J. Olson
R. S. Turner
J. L. Malanchuk

Environmental Sciences Division
Publication No. 2884

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY



Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes-Printed Copy: A07 Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ENVIRONMENTAL SCIENCES DIVISION

**EMPIRICAL RELATIONSHIPS BETWEEN WATERSHED ATTRIBUTES AND
HEADWATER LAKE CHEMISTRY IN THE ADIRONDACK REGION**

**C. T. Hunsaker, S. W. Christensen, J. J. Beauchamp,¹
R. J. Olson, R. S. Turner, and J. L. Malanchuk²**

**Environmental Sciences Division
Publication No. 2884**

¹Engineering Physics and Mathematics Division, P.O. Box Y, Oak Ridge,
Tennessee 37831

²Acid Deposition Assessment Staff, U.S. Environmental Protection
Agency, Washington, D.C.

Date Published: December 1986

**Prepared for the
U.S. Environmental Protection Agency**

**Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
operated by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-84OR21400**

CONTENTS

	<u>Page</u>
LIST OF FIGURES	v
LIST OF TABLES	vii
ABSTRACT	xi
1. INTRODUCTION	1
2. ADIRONDACK WATERSHED DATA BASE	6
2.1 WATERSHED SELECTION	6
2.2 VARIABLES	8
2.3 DATA UNCERTAINTY	20
2.4 DATA SUBSETS	22
2.5 DATA BASE MANAGEMENT	24
3. STATISTICAL ANALYSES	25
3.1 ANALYSES INVOLVING SINGLE PREDICTOR VARIABLES	25
3.1.1 Methods	25
3.1.2 Results	26
3.1.2.1 Watershed and Lake Morphology, Physiography, and Hydrology	30
3.1.2.2 Atmospheric Inputs	31
3.1.2.3 Watershed Soils	35
3.1.2.4 Watershed Geology	37
3.1.2.5 Watershed Vegetation	38
3.1.2.6 Watershed Disturbance	43
3.1.2.7 Lake pH, Color, and Dissolved Organic Carbon	46

	<u>Page</u>
3.2 MULTIPLE LINEAR REGRESSION	47
3.2.1 Methods	49
3.2.1.1 Data Transformations	50
3.2.1.2 Collinearity Diagnostics and Model Development	54
3.2.1.3 Model Verification	59
3.2.2 Results	61
3.2.2.1 Model Development and Verification . . .	62
3.2.2.2 Model Predictions	67
3.3 DISCRIMINANT ANALYSIS	84
3.3.1 Methods	85
3.3.2 Results	86
4. DISCUSSION AND CONCLUSIONS	89
R E F E R E N C E S	101
APPENDIX A: PROCEDURES USED IN DETERMINING WHICH COLLINEAR VARIABLE TO ELIMINATE	109

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1 Watershed boundaries for 463 selected headwater watersheds in Adirondack region	7
2 Contours of annual H^+ concentration in precipitation . . .	14
3 Contours of annual H^+ wet deposition rates	15
4 Contours of annual NO_3^- concentration in precipitation . .	16
5 Frequency distributions of vegetation for 463 watersheds in Adirondack region: (a) dominant forest cover type, (b) wetland type as percentage of watershed, and (c) wetland type as percentage of lake perimeter	17
6 Frequency distributions of soil and geologic attributes for 463 watersheds in Adirondack region: (a) percentage of watershed with steep slopes, (b) dominant hydrologic type, and (c) bedrock buffering capacity	18
7 Frequency distributions of watershed disturbance for 463 watersheds in Adirondack region: (a) beaver activity index, (b) vegetation disturbance, and (c) cabins	19
8 Model development procedure	48
9 Plots of residuals from MLR models for pH and ANC	66
10 Frequency of predicted and observed mean pH and ANC for selected MLR models	72
11 Frequency distributions of observed and predicted pH and ANC values for lakes in calibration subsets	73
12 Frequency distributions of observed and predicted pH and ANC values for lakes in combined calibration and verification subsets	74
13 Spatial pattern of observed summer mean pH for headwater lakes in Adirondack region	82
14 Spatial pattern of observed summer mean ANC for headwater lakes in Adirondack region	83
15 Cumulative R^2 values for stepwise MLR models for pH and ANC	94

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Overlay	
Spatial pattern of predicted summer mean pH for headwater lakes in Adirondack region	Inside back cover
Spatial pattern of predicted summer mean ANC for headwater lakes in Adirondack region	Inside back cover

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 Data sources used to compile the Adirondack Watershed Data Base	9
2 Variable names for watershed attributes and their units of measure	10
3 Subsetting codes for 463 headwater lakes in Adirondacks . .	23
4 Spearman correlations between mean surface water pH and ANC and watershed attributes	27
5 Spearman correlations between NSWs pH and ANC and watershed attributes	32
6 Wilcoxon test results for lake pH and ANC for wetlands . .	40
7 Variables used in MLR and discriminant models	51
8 Variable transformations	53
9 Candidate variables listed in order of their elimination because of collinearity	56
10 Indices for pH model verification	64
11 Indices for ANC model verification	65
12 Estimated coefficients and their standard errors for selected MLR models and reduced stepwise model for pH . . .	68
13 Estimated coefficients and their standard errors for selected MLR models and reduced stepwise model for ANC . .	70
14 Comparison of percentages of lakes in pH categories for observed, MLR-prediction, and adjusted MLR-prediction values	77
15 Comparison of percentages of lakes in ANC categories for observed, MLR-prediction, and adjusted MLR-prediction values	78
16 Estimated number of Adirondack headwater lakes in pH categories	79
17 Estimated number of Adirondack headwater lakes in ANC categories	80

<u>Table</u>	<u>Page</u>
18 Results of discriminant analysis for pH	87
19 Results of discriminant analysis for ANC	87
A-1 Groupings and other information used in process of selecting variables for elimination in multicollinearity analysis	111
A-2 Priority groups for wetland variables used in process of selecting variables for elimination in multicollinearity analysis	116

ACKNOWLEDGMENTS

We thank the following individuals for contributing to the development of the Adirondack Watershed Data Base.

Environmental Sciences Division, ORNL North Carolina State University

A. E. Rosen

J. P. Baker

**Computing and Telecommunications
Division, ORNL**

Adirondack Park Agency

**R. C. Durfee
P. R. Coleman
D. L. Wilson
F. E. Latham**

R. P. Curran

State University of New York/Plattsburgh

**G. K. Gruendling
D. J. Bogucki
K. S. Adams
E. S. Allen**

Within the Environmental Sciences Division, we thank J. A. Solomon and R. S. Cook for reviewing this report and L. J. Allison, L. K. McDonald, D. D. Rhew, and G. R. Carter for technical support. This research was sponsored as part of the interagency National Acid Precipitation Assessment Program (NAPAP) by EPA, with the support of the Office of Environmental Analysis, U.S. Department of Energy, under EPA Interagency Agreement No. 40-1488-84 and under Contract No. DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc. The research described in this report has not been subjected to EPA's or NAPAP's required peer and policy review and does not necessarily reflect the views of these organizations, therefore no official endorsement should be inferred.

ABSTRACT

HUNSAKER, C. T., S. W. CHRISTENSEN, J. 3. BEAUCHAMP,
R. 3. OLSON, R. S. TURNER, and J. L. MALANCHUK. 1986.
Empirical relationships between watershed attributes
and headwater lake chemistry in the Adirondack region.
ORNL/TM 9838. Oak Ridge National Laboratory, Oak Ridge,
Tennessee. 135 pp.

Surface water acidification may be caused or influenced by both natural watershed processes and anthropogenic actions. Empirical models and observational data can be useful for identifying watershed attributes or processes that require further research or that should be considered in the development of process models. This study focuses on the Adirondack region of New York and has two purposes: to (1) develop empirical models that can be used to assess the chemical status of lakes for which no chemistry data exist and (2) determine on a regional scale watershed attributes that account for variability in lake pH and acid-neutralizing capacity (ANC). Headwater lakes, rather than lakes linked to upstream lakes, were selected for initial analysis. The Adirondacks Watershed Data Base (AWDB), part of the Acid Deposition Data Network maintained at Oak Ridge National Laboratory (ORNL), integrates data on physiography, bedrock, soils, land cover, wetlands, disturbances, beaver activity, land use, and atmospheric deposition with the water chemistry and morphology for the watersheds of 463 headwater lakes. The AWDB facilitates both geographic display and statistical analysis of the data. The report, An Adirondack Watershed Data Base: Attribute and Mapping Information for Regional Acidic Deposition Studies (ORNL/TM-10144), describes the AWDB.

Both bivariate (correlations and Wilcoxon and Kruskal-Wallis tests) and multivariate analyses were performed. Fifty-seven watershed attributes were selected as input variables to multiple linear regression and discriminant analysis. For model development -200 lakes for which pH and ANC data exist were randomly subdivided into a specification and a verification data set. Several indices were used to select models for predicting lake pH (31 variables) and ANC (27 variables). Twenty-five variables are common to the pH and ANC models: four lake morphology, nine soil/geology, eight land cover, three disturbance, and one watershed aspect. An atmospheric input variable (H^+ or NO_3^-) explains the greatest amount of variation in the dependent variable (pH and ANC) for both models. The percentage of watershed in conifers is the next strongest predictor variable. For all headwater lakes in the Adirondacks, -60% of the lakes are estimated to have an ANC ≤ 50 $\mu\text{eq/L}$, and 40% of the lakes have a pH ≤ 5.5 , levels believed to be detrimental to some fish species.

1. INTRODUCTION

For this study, a set of headwater lakes within the Adirondack Park of New York was selected for developing an empirical model to evaluate alternative hypotheses concerning factors contributing to acidification of surface waters and to predict the pH and acid neutralizing capacity (ANC) of lake water. The Adirondacks are a logical area for a regional study of lake water quality because a large number of lakes have been monitored over the past several decades and lakes in the region appear to be undergoing acidification. Water chemistry within the Adirondacks has been studied extensively (Schofield 1976a, 1976b; Colquhoun et al. 1984), and relationships between water chemistry and fish status also have been studied (Baker and Harvey 1984, Reckhow et al. 1985). However, only limited studies relating watershed characteristics to lake chemistry have been performed in the Adirondacks. The Integrated Watershed Acidification Study/Regionally Integrated Watershed Acidification Study (ILWAS/RILWAS) projects (Goldstein 1983) monitored and studied three Adirondack watersheds extensively over several years to develop and test a watershed model. Regional assessments (Schnoor et al. 1985, Nair 1984) have used a limited number of variables obtained from small-scale regional maps for model input.

The primary objective of the present study is to examine on a regional scale watershed attributes that may account for variability and change in water chemistry in the Adirondacks. A secondary objective is to use the empirical relationships developed through the

statistical analyses to assess the status of additional headwater lakes in the Adirondacks for which no water chemistry data exist. This study differs from other studies of lakes in the Adirondacks by including a large number of lakes, more watershed attributes, increased spatial resolution of the data used in the analysis, and more-extensive statistical analysis.

The organization of this report is presented to help the reader identify areas of interest. The introduction (Chap. 1) presents the background and rationale for the analysis, followed by chapters discussing the data base (Chap. 2), the analyses (Chap. 3), and the conclusions (Chap. 4). The data base chapter describes the population of lakes that were used and the watershed variables, including the sources of data. More details on the development of the Adirondack Watershed Data Base (AWDB) are given in Rosen et al. (1986). The statistical-analyses chapter discusses the types of analyses and presents results. The chapter is detailed because of the desire to apply multiple tests to help verify the overall conclusions. That is, the analysis is based on observational data obtained from a variety of sources, and comparable results in the relationships between watershed attributes and lake chemistry were obtained from the independent statistical approaches. A detailed statistical discussion to confirm the interpretation of results presented in the final chapter is included. The casual reader may wish to concentrate on the discussion of the "selected" best models (highlighted by bold type in tables).

To determine the causes of lake acidification, one must determine whether observed lake acidity can be attributed to atmospherically

deposited acids or naturally derived acids based on the type of acids found in the lake waters and on the relative importance of the potential sources and sinks of different types of acids in the watersheds. This study identifies potential causes of lake acidification. Simple correlations and multivariate relationships between pH and ANC and those lake or watershed factors that could contribute to natural or anthropogenic acidification of Adirondack headwater lakes are evaluated. The watershed factors examined are lake morphology, water chemistry, wetlands, land cover, land use, soil associations, precipitation, and beaver activity. Anthropogenic factors that could cause changes in lake chemistry include increased atmospheric deposition of pollutants, development around lakes, and land disturbances. Natural factors that could also contribute to changes in lake chemistry include areal extent of wetlands, coniferous forests, bedrock, depth to bedrock, and acidic soils. Many hypotheses to determine which of these factors were significantly associated with pH and ANC for headwater lakes were evaluated in this study.

Turner et al. (1986a); Schnoor and Stumm (1985); Johnson et al. (1985); and Mason and Seip (1985) have recently summarized the state of knowledge on factors controlling surface water chemistry, including (1) atmospheric inputs; (2) canopy interactions; (3) anion mobility, cation exchange, and weathering in soils and bedrock as mediated by hydrologic contact; (4) uptake and redistribution of chemicals within the ecosystem by vegetation; and (5) in-stream/in-lake processes.

Various processes in the terrestrial and aquatic environment can neutralize or enhance acidic precipitation after it enters the

watershed. Denitrification, sulfate reduction, and chemical weathering decrease acidity; photosynthetic assimilation and nitrification increase the acidity of waters. Two pools of bases are in soils--a small pool of exchangeable bases with relatively rapid kinetics and a large pool of mineral bases with the slow kinetics of chemical weathering (Schnoor and Stumm 1985). Thus, we were interested in quantifying rock outcrops, soil cation exchange capacity, soil base saturation, and soil pH. Acidic lakes occur where the residence time of acidic precipitation in soils and the watershed is relatively short (i.e., soils are thin) and where lakes and their watersheds are small (Schnoor and Stumm 1985). Sensitive watershed attributes were characterized by using (1) measures of soil infiltration rates; (2) depths to bedrock, to a root restrictive zone, and to a low permeability horizon; (3) soil steepness; and (4) lake morphology and hydrology.

The release of H^+ by aggrading vegetation may exceed the rate of H^+ consumption by weathering and cause progressive acidification in noncalcareous soils. Also, in some wetlands, aggrading humus and net production of base-neutralizing capacity can cause acidic conditions and release humic or fulvic acids to the water (Gorham et al. 1985). Therefore, data on land cover and percentage of wetlands in each watershed were developed. Correspondingly, disturbance processes (e.g., fire, logging, and human development) reduce vegetative growth, thus producing an alkalinity generating process (i.e., the ashes of trees are alkaline). Denitrification (NO_3^- reduction) and SO_4^{2-} reduction induced by decomposition of organic matter (oxidation) cause an increase in ANC (Schindler et al. 1986).

Sediments in a water-sediment system are usually highly reducing environments. Beaver dams trap sediments and add organic matter to waters, thus providing more sites for microbial colonization which create reducing environments for NO_3^- and SO_4^{2-} (Driscoll et al., in press; Francis et al. 1985). Beaver activity also floods soils and may result in humic acid inputs (Salzer 1935, Adams 1953, Call 1966, Naiman et al. 1986). Watershed variables, such as the beaver activity index, the percentage of watershed disturbed, and the percentage of watershed in a wetland type, were developed to capture these watershed attributes.

The results of this study need to be interpreted in light of the sources of data and statistical methods. The AVDB data are observational (not collected under statistically designed conditions to test specific hypotheses of interest for this study). Associations between factors can be determined, but cause and effect relationships cannot be proven.

2. ADIRONDACK WATERSHED DATA BASE

2.1 WATERSHED SELECTION

The subset of watersheds selected for this study includes 463 headwater lake systems within the Adirondack Ecological Zone (AEZ) (Fig. 1). The development of the (AWDB) including lake selection, data sources, and computational algorithms, is described in detail in Rosen et al. (1986). Headwater lake systems consist of a single headwater lake and its watershed, in contrast with lakes that are fed by wetlands or by streams draining other lakes (the latter are considered to be complex lake systems). The identification of watershed characteristics that influence lake chemistry should be easier for headwater lakes than for nonheadwater lakes because a headwater lake is only affected by processes in its immediate watershed. The AEZ, as defined by the -300-m elevation contour surrounding the Adirondack Park, contains 2759 lakes (Colquhoun et al. 1984). The selection of lakes was restricted to an area having wetland maps (-63% of the AEZ). Watershed boundaries were outlined using 1:62,500 or 1:24,000 scale United States Geological Survey (USGS) topographic maps with the aid of 1:20,000 aerial photographs (Gruendling et al. 1985). This process resulted in the majority of headwater lakes in the AEZ being included in the AWDB. However, some headwater lakes were excluded from this study for one of the following potentially confounding reasons: lack of a pond number assigned by the New York State Department of Environmental Conservation (NYDEC), man-made lakes (reservoirs, quarries, tailings lakes, etc.), lakes adjacent to roads or railroad embankments (potential changes in hydrologic flow), lakes with

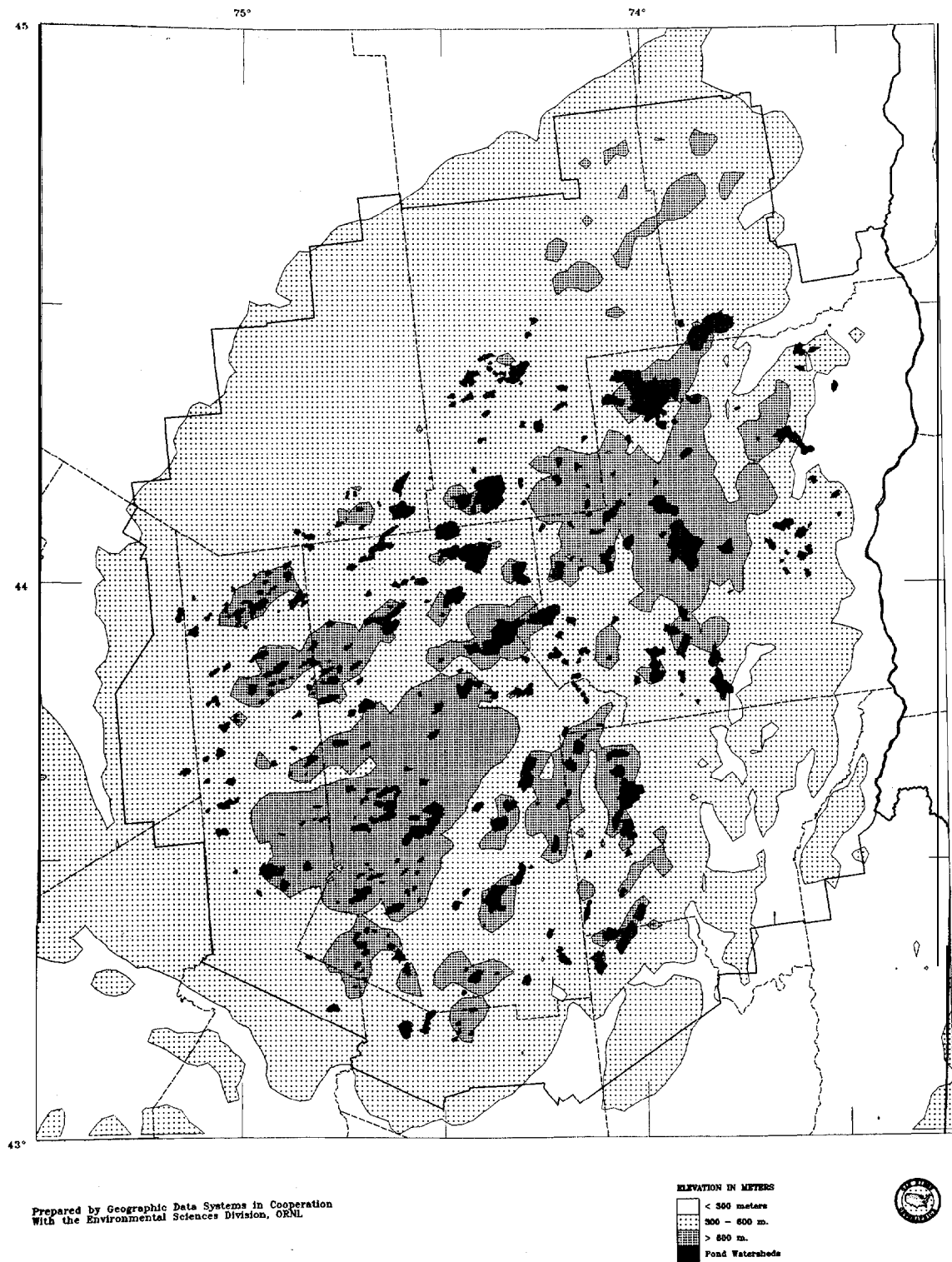


Fig. 1. Watershed boundaries for 463 selected headwater watersheds in Adirondack region.

significant size changes caused by beavers (as seen in aerial photos), and lakes in flat areas with watershed boundaries that were difficult to define.

Researchers have suggested that small, high-elevation lakes are most susceptible to acidification from acidic deposition. Schofield (1976b) defined lakes above 610 m as high-elevation lakes in the Adirondack region. Colquhoun et al. (1984) also refer to high-elevation lakes as those above 610 m and define small lakes as those with an area <40 ha. The 463 AVDB headwater lakes are generally smaller and occur at higher elevations than the average lake within AEZ. The average size of the 2759 lakes in AEZ is 41 ha; the AVDB lakes average 18 ha with a median lake size of 10 ha. The average elevation of lakes in the Adirondacks is 499 m, the AVDB lakes average 587 m. Thus, the AVDB headwater lakes, selected so that watershed influences would not be confounded by upstream lake processes, are a subset of lakes atypical of all lakes in the AEZ.

2.2 VARIABLES

Lake and watershed attributes thought to influence lake acidification were compiled from a variety of sources (Table 1) into AVDB for analysis. Attributes include lake morphology, water chemistry, health of fish populations, bedrock type, soils, hydrology, vegetation, wetlands, beaver activity, fire and logging disturbances, land use, climate, and atmospheric deposition (Table 2). AVDB includes data sources, data manipulations, and data base characteristics and is documented by Rosen et al. (1986). Watershed data were compiled primarily from extant sources, such as maps and aerial photographs.

Table 1. Data sources used to compile the Adirondack Watershed Data Base

Data type	Source	Compiled by^a
Morphology	USGS topographic maps (1:62,500 to 1:24,000), NYDEC records, aerial photos (1:20,000, 1968, B/W)	SUNY/P
Physiography	DMA TOPOCOM digital representation of USGS 1:250,000 topographic maps	ORNL
Bedrock	1982 geologic map^b	ORNL
Soils	1974 SCS Mesoscale maps, SCS SOILS-5, geoecology chemistry	APA/ORNL
Land cover	1978 Landsat imagery	APA/ORNL
Wetlands	1982 wetland map	SUNY/P
Cabins	1978-1983 aerial photos	SUNY/P
Fire, logging	1916 NY state map (1:125,720) of wildfires and timber harvesting	SUNY/P
Beavers	1978-1981 aerial photos (1:24,000)	SUNY/P
Deposition	1951-1980 Precipitation norms 1980-1982 Deposition monitoring	ORNL
Land use	APA Park plan	APA
Water chemistry	FIN-assembled from several sources	NCSU
Fish status	FIN-assembled from several sources	NCSU

^aSUNY/P - State University of New York at Plattsburgh (Gruendling et al. 1985); ORNL - Oak Ridge National Laboratory; APA - Adirondack Park Agency (R. Curran, personal communication); NCSU - North Carolina State University (Baker et al. 1984).

^bS. A. Norton et al. 1982; National Atmospheric Deposition Program and Association of State Agricultural Experiment Stations of the North Central Region n.d.

Table 2. Variable names for watershed attributes and their units of measure

Variable name for watershed attributes	Watershed attributes	Units of measure
<u>Morphologic and Physiographic</u>		
LAKE_A	Lake area	ha
WTRSHD_R	Watershed to lake area ratio	
DRAIN_A	Watershed area	ha
LAKE_DEV	Lake development ratio ^a	
LAKE_E	Lake elevation	m above mean sea level
ASPECT_S	Southern aspect	% watershed area
ASPECT_N	Northern aspect	% watershed area
LAKE_V	Lake volume	10 ⁶ •m ³
<u>Hydrologic</u>		
RUNOFF	Annual runoff	cm
HYDTYP1	Seepage lake (no inlets or outlets)	
HYDTYP2	Spring lake (outlets, no inlets)	
HYDTYP3	Drainage lake (both inlets and outlets)	
<u>Atmospheric</u>		
PPT	Average annual precipitation	cm
H_WET	Average annual hydrogen wet deposition	g/m ²
NO3_WET	Average annual nitrate wet deposition	g/m ²
SO4_WET	Average annual sulfate wet deposition	g/m ²
SO4_NO3	Average annual combined sulfate and nitrate wet deposition	g/m ²
H_CONC	Average annual hydrogen wet concentration	mg/L
NO3_CONC	Average annual nitrate wet concentration	mg/L
SO4_CONC	Average annual sulfate wet concentration	mg/L
<u>Physical Soil Type</u>		
RELIEF_R	Relief (maximum elevation-lake elevation) to square root (watershed area) ratio	
STONEYP	Stoney soils	% watershed area
ROCK_P	Rock outcrops	% watershed area
HYDRO_A	High infiltration rate	% watershed area
HYDRO_B	Moderate infiltration rate	% watershed area
HYDRO_C	Slow infiltration rate	% watershed area
HYDRO_D	Very slow infiltration rate	% watershed area
STEEPM_P	Moderately steep soils	% watershed area
STEEPV_P	Very steep soils	% watershed area

Table 2. (continued)

Variable name for watershed attributes	Watershed attributes	Units of measure
<u>Physical Soil Type (continued)</u>		
SHL2_B_P	Depth to bedrock ≤ 100 cm	% watershed area
SHL2_P_P	Depth to low-permeability horizon ≤ 100 cm	% watershed area
SHL2_R_P	Depth to root restrictive zone ≤ 100 cm	% watershed area
SHL1_B_P	Depth to bedrock ≤ 50 cm	% watershed area
SHL1_P_P	Depth to low-permeability horizon ≤ 50 cm	% watershed area
SHL1_R_P	Depth to root ≤ 50 cm	% watershed area
SHL1_Z_P	Shallow soils ≤ 50 cm	% watershed area
SHL2_Z_P	Shallow soils ≤ 100 cm	% watershed area
DPTH_B_U	Mean depth to bedrock – upper	cm
DPTH_P_U	Mean depth to low permeability – upper	cm
DPTH_R_U	Mean depth to root restrictive zone – upper	cm
<u>Geology</u>		
ROCK12_P	Medium to no acid-neutralizing capacity	% watershed area
ROCK1_P	Low to no acid-neutralizing capacity	% watershed area
ROCK2_P	Medium to low acid-neutralizing capacity	% watershed area
ROCK3_P	High to medium acid-neutralizing capacity	% watershed area
ROCK4_P	Infinite acid-neutralizing capacity	% watershed area
<u>Chemical Soil Type</u>		
ACID_P	Extractable acidity > 20 meq/100 g	% watershed area
BSA_L_P	Base saturation $\leq 20\%$	% watershed area
BSA_M_P	Base saturation (NH ₄ OAC) 20–60%	% watershed area
BSC_L_P	Base saturation (sum) $\leq 20\%$	% watershed area
BSC_M_P	Base saturation (sum) 20–60%	% watershed area
CECS_L_P	Cation exchange capacity ≤ 20 meq/100 g (Sum of cations)	% watershed area
CEC_L_P	Cation exchange capacity ≤ 10 meq/100 g	% watershed area
OM_H_P	Organic matter content $\geq 2\%$	% watershed area
PH_VL_P	Soil pH (H ₂ O) ≤ 4.5	% watershed area
PHC_L_P	Soil pH (CaCl ₂) ≤ 5.0	% watershed area
PHC_VL_P	Soil pH (CaCl ₂) ≤ 4.5	% watershed area
ACID_EX	Mean extractable acidity	meq/100 g
CEC	Mean cation exchange capacity	meq/100 g
ORG_MAT	Mean organic matter content	%

Table 2. (continued)

Variable name for watershed attributes	Watershed attributes	Units of measure
<u>Forest Cover</u>		
CONFR2_P	Area in coniferous forest	% watershed area
HRDWD2_P	Area in deciduous forest	% watershed area
NONFR2_P	Area not in forest	% watershed area
MIXED2_P	Area in mixed forest	% watershed area
<u>Wetland Type</u>		
WTLND_PP	All wetland types	% lake perimeter
VACID_PP	Very acid wetland type	% lake perimeter
NACID_PP	Nonacid wetland type	% lake perimeter
MACID_PP	Moderately acid wetland type	% lake perimeter
OTHER_PP	Other wetland type	% lake perimeter
WTLND_PW	All wetland types	% watershed area
VACID_PW	Very acid wetland type	% watershed area
NACID_PW	Nonacid wetland type	% watershed area
MACID_PW	Moderately acid wetland type	% watershed area
OTHER_PW	Other wetland type	% watershed area
WTLND_PL	All wetland types	% lake area
VACID_PL	Very acid wetland type	% lake area
NACID_PL	Nonacid wetland type	% lake area
MACID_PL	Moderately acid wetland type	% lake area
OTHER_PL	Other wetland type	% lake area
<u>Disturbance</u>		
DISTRB_P	Sum of logged, burned, denuded area	% watershed area
BVRINDEX	Beaver activity index	
CABN78_R	Number of 1978 cabins to lake area ratio	
BURNED_P	Burned area	% watershed area
DENUDE_P	Denuded area	% watershed area
LOG_SH_P	Logged softwood and hardwood area	% watershed area

^aThe perimeter of the lake divided by the perimeter of a circle with the same area as that of the lake (Wetzel, R. G. 1975. Limnology. W. B. Sanders Co., Philadelphia, PA.).

Water chemistry and fish data were obtained from the Fish Information Network (FIN) data base (Baker et al. 1984) and from the Eastern Lake Survey-Phase I (Linthurst et al. 1986). Chemistry data are available for about one-half of the AWDB lakes. Every attempt was made to use watershed data from the same time period as the FIN water chemistry data (1974-1983).

Atmospheric deposition has been suggested as a principal candidate in the acidification of Adirondack lakes (Altshuller and Linthurst 1984). Annual average wet deposition rates for sulfate, nitrate, and total hydrogen ion were calculated for watersheds based on the years 1980-1982 (Rosen et al. 1986). The concentration of ions in precipitation (interpolated between monitoring sites) was multiplied by precipitation amounts (also interpolated between the more numerous weather stations) to calculate total wet deposition rates. The patterns for hydrogen ions (Figs. 2 and 3), nitrate (Fig. 4), and sulfate are all similar, showing higher levels in the western Adirondacks.

Deciduous vegetation dominated the landscape (Fig. 5a). The majority of watersheds contained wetlands, and based on either percentage of wetland area in the watershed or percentage of wetlands in contact with the shoreline, the majority of wetlands were classified as very acid, a condition thought to produce organic acids (Figs. 5b and 5c). The majority of watersheds have slow infiltration and very steep slopes (Fig. 6). Most watersheds are underlain by bedrock with low to moderate buffering capacity (Fig. 6). Eighty-five percent of the lakes do not have cabins near them and only one-half of the lakes have beaver activity (Fig. 7). Based on the Adirondack land management plan,

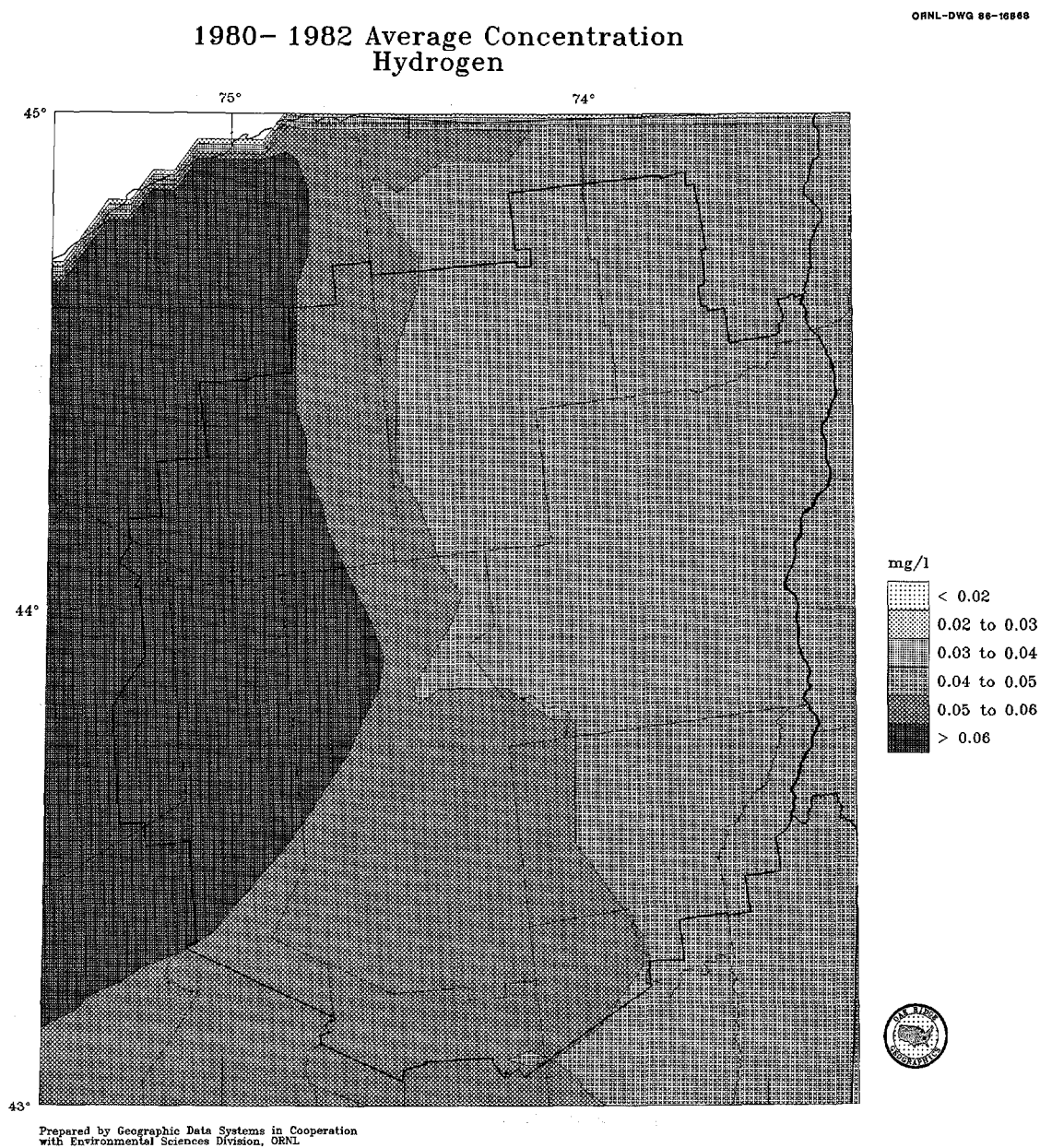


Fig. 2. Contours of annual H^+ concentration in precipitation (overlays inside back cover).

1980-1982 Average Annual Hydrogen Ion Wet Deposition

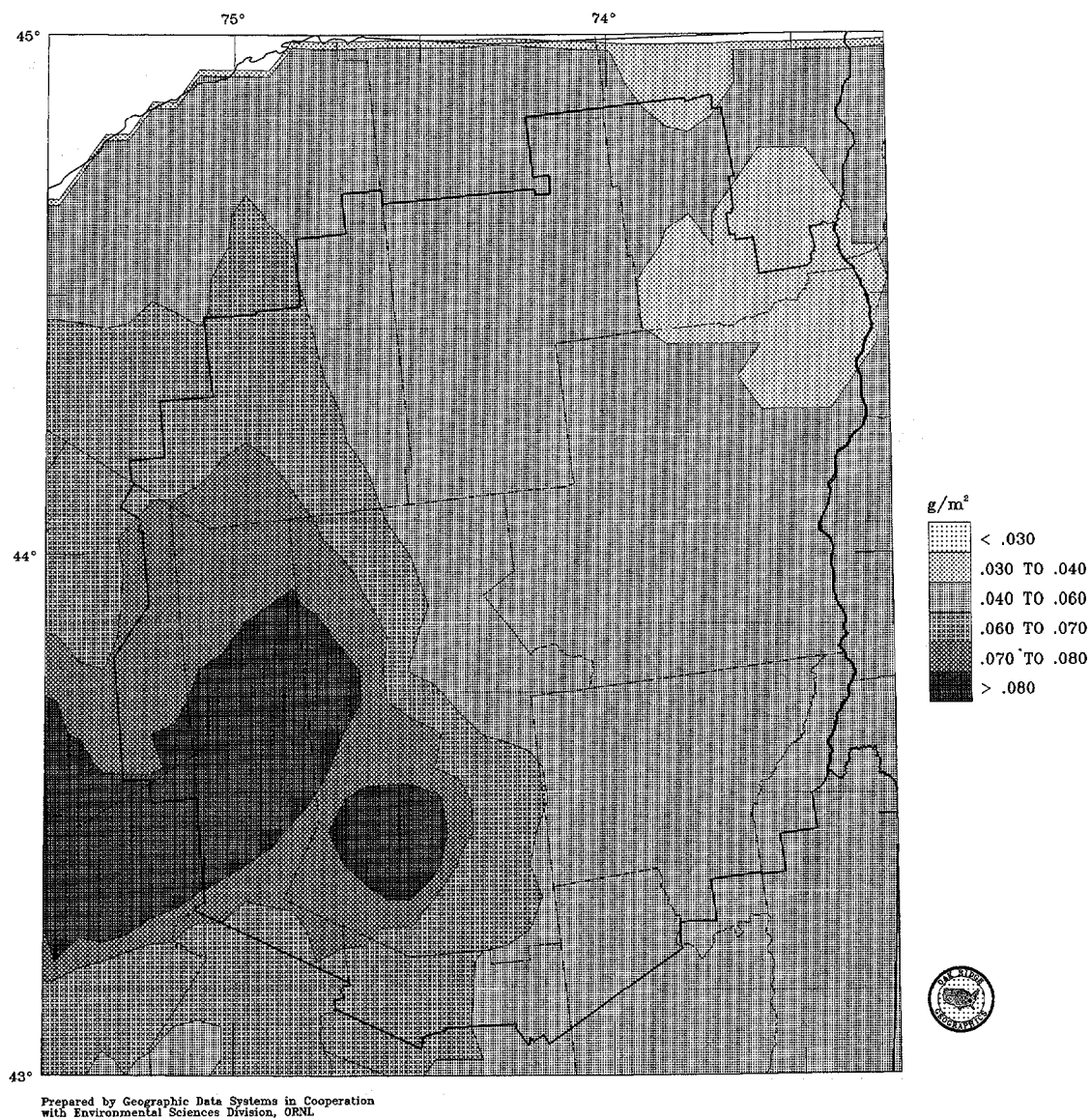


Fig. 3. Contours of annual H^+ wet deposition rates (overlays inside back cover).

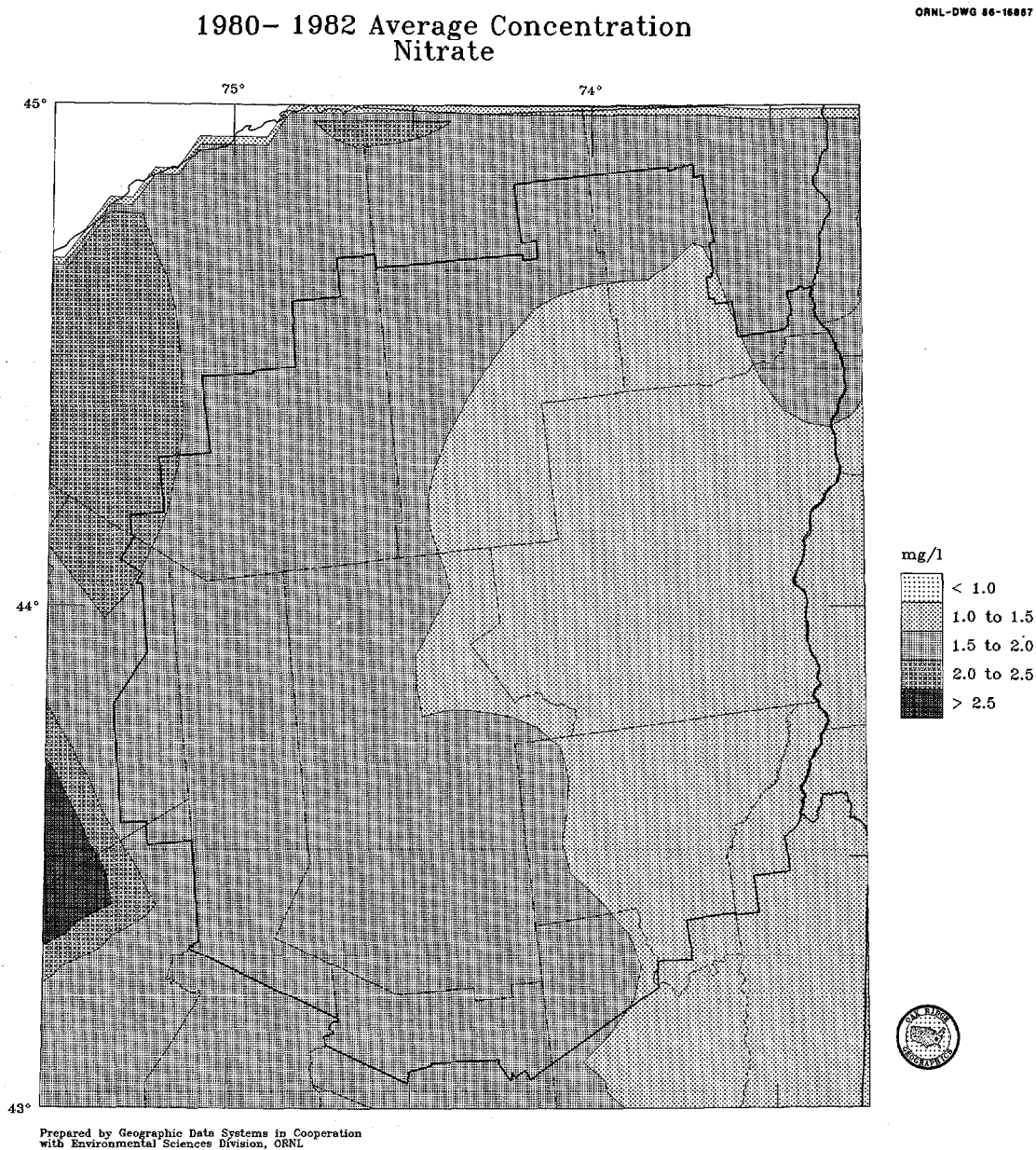


Fig. 4. Contours of annual NO_3^- concentration in precipitation (overlays inside back cover).

ORNL-DWG 86C-16807

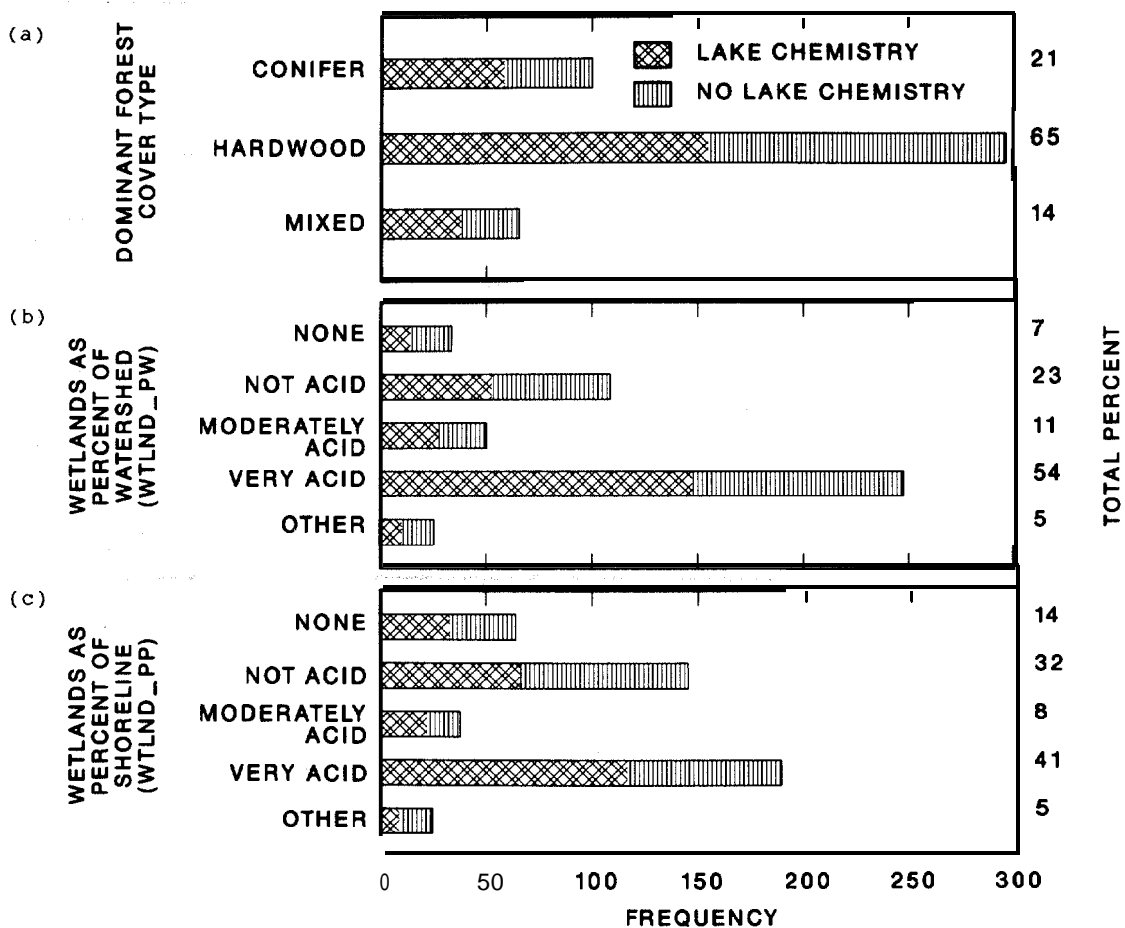


Fig. 5. Frequency distributions of vegetation for 463 watersheds in Adirondack region: (a) dominant forest cover type, (b) wetland type as percentage of watershed, and (c) wetland type as percentage of lake perimeter.

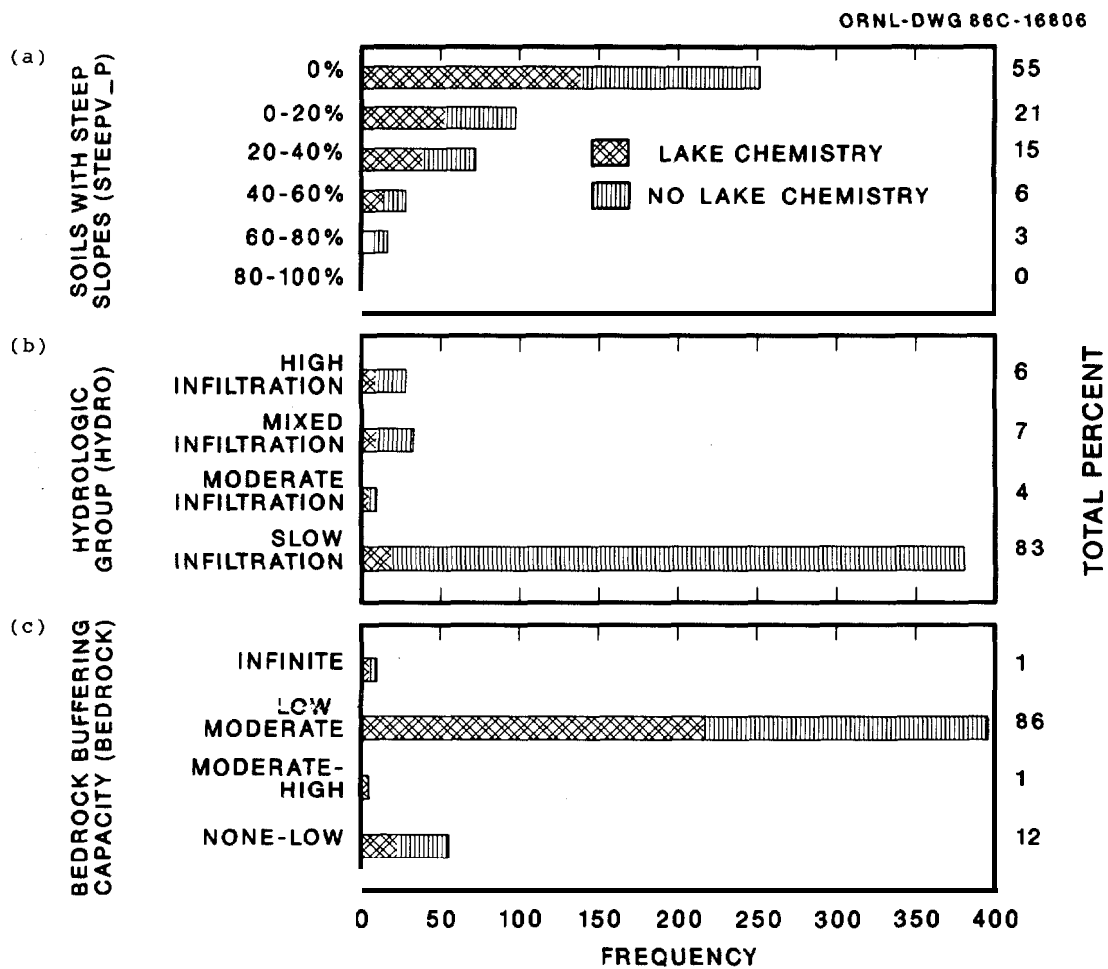


Fig. 6. Frequency distributions of soil and geologic attributes for 463 watersheds in Adirondack region: (a) percentage of watershed with steep slopes, (b) dominant hydrologic type, and (c) bedrock buffering capacity.

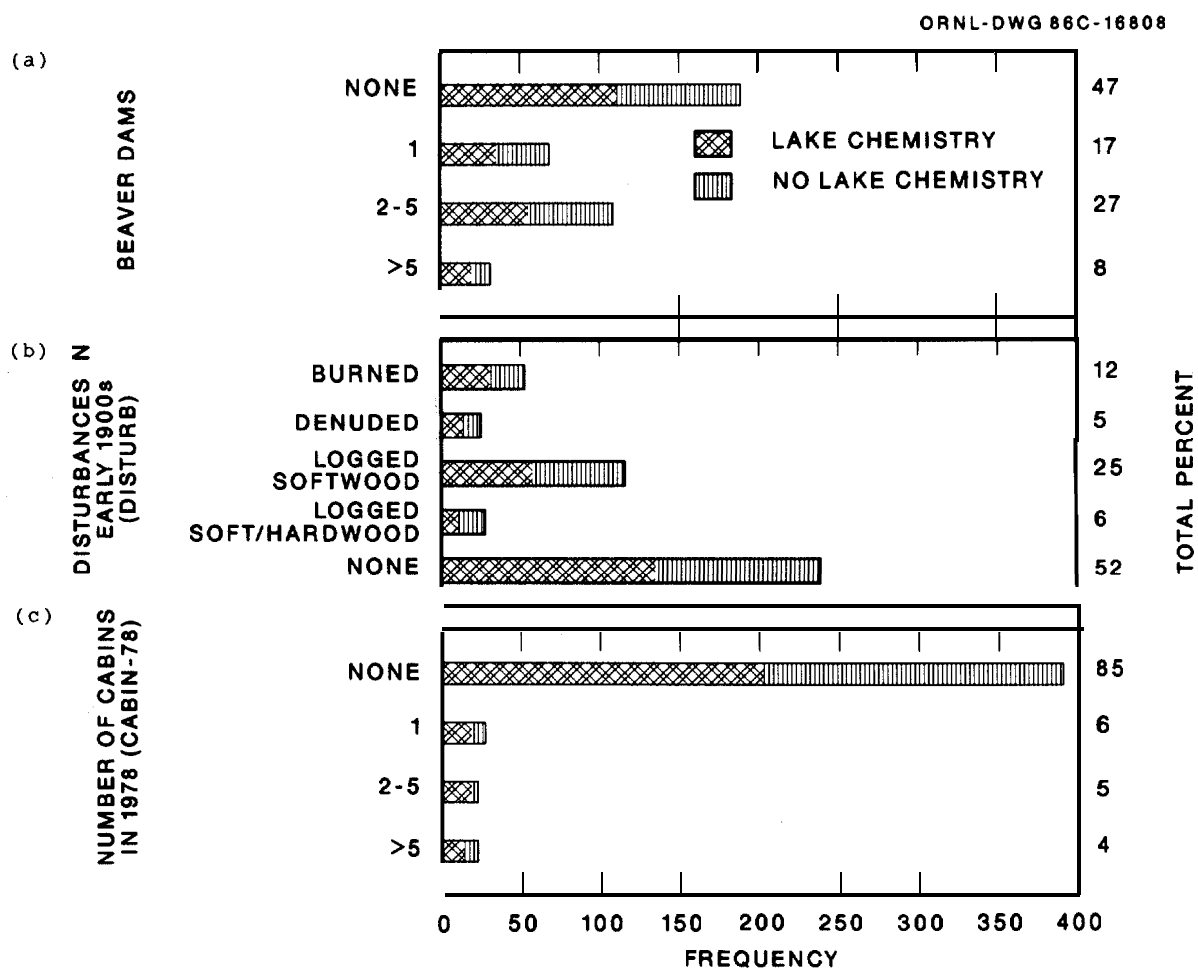


Fig. 7. Frequency distributions of watershed disturbance for 463 watersheds in Adirondack region: (a) beaver activity index, (b) vegetation disturbance, and (c) cabins.

most watersheds are located in areas designated as primitive unmanaged forest or in areas with some type of resource management. As a final example of the distribution of watershed characteristics, almost one-half of the watersheds had some sort of historic logging, wildfire, or other disturbance, based on a 1916 map of the region (Fig. 7).

2.3 DATA UNCERTAINTY

Watershed attributes were compiled from a variety of different source materials, including remote imagery, aerial photographs, maps of various scale, and sparse regional monitoring networks (Table 1). Uncertainty of the data relates to the coarse and different spatial scale of some source materials, interpretation errors (e.g., boundary delineation and remote-sensed data, and association of mapping units with parameters used in the analysis. Because of the small size of watersheds and the small scale of some source maps, individual watersheds may be assigned incorrect attributes. For example, land-cover data involved the unsupervised classification of Landsat scenes. Four Landsat scenes with four different dates were required to cover the Adirondack region and to obtain cloud-free scenes, resulting in pattern changes at boundaries between adjacent scenes. Based on a working knowledge of the park, the Adirondack Park Agency has verified the overall correctness of the data (Curran, personal communication). The regional coverage and large number of watersheds should minimize the effects of individual watershed misclassification.

Soil mapping units were assigned chemical properties by merging soil chemistry data with each soil series identified in a mapping unit. Occasionally, data were not available for a soil series, or

mapping units (A and E soil horizons) included a "miscellaneous" category. In all cases, the soil chemistry values for mapping units were derived by prorating the available data for soil series according to their relative abundance within a mapping unit (Turner et al. 1986aj). The uncertainty or variability of the soil chemistry data is unknown because often only single measurements on typical soil series profiles are available. These problems are being addressed by Oak Ridge National Laboratory (ORNL) staff in collaboration with the National Soils Laboratory of the Soil Conservation Service (SCS) and also by the Environmental Protection Agency soils survey projects.

The wet deposition data contain uncertainty related to interpolating from monitoring stations to the individual watersheds. Deposition contours were derived from the nonuniformly distributed monitoring sites by generating a Thiessen polygon network between the sites, interpolating a regularly spaced grid, and calculating contours (Rosen et al. 1986). This rigorous mathematical approach defines a smooth deposition between the irregularly spaced monitoring sites; however, it does not explicitly account for possible orographic factors.

Water chemistry data within FIN were collected by many investigators for different purposes, using a variety of analytical techniques. As a result, the data are often not ideally suited for use in statistical analyses. Two pervasive problems are the representativeness of the sample and variations in data quality. The issues discussed above are common to environmental data for regional studies. Despite these imperfections, the results of this study show that analysis of the existing data base can contribute significantly to

understanding the acidity status of lakes in the Adirondacks and the relationships between watershed attributes and lake chemistry.

The National Surface Water Survey (NSWS) measured pH, ANC, color, dissolved organic carbon, and other parameters in the fall of 1984 for 46 of the Adirondack headwater lakes considered in this study (Linthurst et al. 1986). Although measurements on each lake were made only once during the autumn overturn, extensive precautions were taken to minimize any variability associated with sample collection and handling, laboratory bias in analysis, and data entry (-30% of data were collected for quality assurance checking). Lakes were selected to be regionally representative by using a stratified systematic sampling scheme based on alkalinity and geographic region. The overall uncertainty of NSWS chemistry values should be less than that for FIN chemistry data. For the 46 headwater lakes in both FIN and NSWS, the pH and ANC values are very similar.

2.4 DATA SUBSETS

The FIN lakes were divided into separate data sets for model calibration (parameter estimation) and several types of verification. A variable designated "SUBSET" was assigned a value from 1 to 9 for each lake, identifying the use to be made of that lake in the model development process (Table 3). One lake lacked a predictor variable and was excluded from the analysis (SUBSET=1). The headwater lakes that were also included in Phase I of NSWS were set aside for use in a secondary verification (SUBSET=2). Two hundred lakes lacking both pH and ANC measurements were assigned values of SUBSET=3; these represented lakes for which both pH and ANC needed to be predicted. The remaining

Table 3. Subsetting codes for 463 headwater lakes in Adirondacks

Condition	Number of lakes	Subset code^a
Lacks one or more predictor variables	1	1
Exclusive of (1) (i.e., has all predictor variables) and also included in NSW Phase I (secondary verification)	46	2
Exclusive of (1) and (2) and lacks both pH and ANC	200	3
Exclusive of (1) and (2), and has pH but not ANC	10	4
one-third reserved for verification	21	7
two-thirds available for calibration		
Exclusive of (1) and (2), and has both pH and ANC	57	5
one-third for verification	114	8
two-thirds for calibration		
Exclusive of (1) and (2), and has ANC but not pH	4	6
one-third for verification	10	9
two-thirds for calibration		
Total	463	

^aDefinition of calibration subsets and FIN and NSW verification subsets:

FIN calibration subsets

pH: Codes 7 & 8 (n = 135)
 ANC: Codes 8 & 9 (n = 124)

FIN verification subsets

pH: Codes 4 & 5 (n = 67)
 ANC: Codes 5 & 6 (n = 61)

NSW secondary verification subset
 (i.e., chemistry data from NSW)

pH: Code 2 (n = 46)
 ANC: Code 2 (n = 46)

Lakes without chemistry data

pH: Codes 1, 3, 6, & 9 (n = 215)
 ANC: Codes 1, 3, 4, & 7 (n = 232)

216 lakes formed the pool of lakes available either for model calibration (specification) or for primary verification (testing) of the fitted model. Two-thirds of these available lakes were randomly selected for model calibration; the remaining one-third was reserved for primary verification. The selection (Table 3) was done separately for lakes having measurements of pH only (SUBSET=4 and 7), of both pH and ANC (SUBSET=5 or 8), or of ANC only (SUBSET=6 or 9). To maintain as much overlap as possible between the sets of lakes used for pH and for ANC, each lake in subsets 5 or 8 (having measurements of both pH and ANC) was assigned to either the calibration or the primary verification subset for statistical analyses. An algorithm for drawing an exact-size random sample without replacement was used (SAS 1983). After a random number was assigned to each lake and lakes were sorted by this random number, subsetting was done using the algorithm

2.5 DATA BASE MANAGEMENT

AVDB consists of digital data (watershed boundaries, topography, soils, landcover, etc.) within a geographic information system (Durfee and the Geographic Data Systems Section 1986) and watershed/lake attribute data (mean water chemistry, lake size, average slope, total wet deposition, etc.) within a statistical data management system (Rosen et al. 1986). The combined systems provide the capability to extract data from maps, perform statistical analyses or run models, map attributes, and display results of analyses. Watershed attributes were entered into an SAS (1985) data base, and SAS was used for data management, statistical analysis (SAS 1985), and display. The attribute data are available as SAS-formatted data sets by request from R. 3. Olson (ORNL).

3. STATISTICAL ANALYSES

This study used numerous statistical procedures [bivariate analyses, multiple linear regression (MLR), and discriminant analysis] to identify watershed attributes that might influence lake chemistry in the Adirondacks. Before application of these procedures, several steps, involving selection of variables from the complete AWDB, transforming some variables, and creating subsets of the data for specific analyses, were performed. Variables and their units used in this study are listed in Table 2.

AWDB contains observational data (not collected under statistically designed conditions to test specific hypotheses); therefore, very little control over the representativeness of the data for variables of interest existed. To verify the MLR and discriminant analyses, duplicate analyses were performed using both a subset of the FIN chemistry data and the set of lakes having independent chemistry data from NSWs. Results from these analyses were quantitatively compared with results from the principal analyses by using the calibration subset.

3.1 ANALYSES INVOLVING SINGLE PREDICTOR VARIABLES

3.1.1 Methods

Analyses using single predictor variables included the nonparametric Spearman rank correlations (nonparametric procedures are based on ranks rather than actual observed values of the random variables), the Kruskal-Wallis test for more than two samples, and the Wilcoxon two-sample test (Conover 1980). Spearman correlations were

performed for pH and ANC with each of 84 watershed attributes for all of the FIN data, the calibration subset of FIN, and the NSW6 subset (subsets defined in Sect. 3.2.1.1 and Table 3). Wilcoxon two-sample tests were performed for data on forest cover and wetlands, and Kruskal-Wallis tests were performed on beaver data. These tests were used to evaluate hypotheses about individual watershed attributes that might influence lake acidification.

The nonparametric tests compare the mean ranks of the dependent variable in each class to determine if significant differences exist among the classes. When the results of the Kruskal-Wallis test indicated significant differences among the classes, a multiple comparison was performed to determine which pairwise combinations of the four classes differ significantly [i.e., which class showed a higher or lower mean value when compared with the others (Conover 1980)]. Some parametric procedures were also used to substantiate results of nonparametric procedures.

3.1.2 Results

In this section, hypotheses about the influence of individual watershed characteristics on the chemistry of headwater lakes are examined. To simplify discussion of the results, watershed attributes are grouped into the following categories: morphology, physiography, and hydrology; atmospheric input; soil; geology; vegetation; and disturbances.

Spearman correlation results from the calibration and full FIN subsets are presented in Table 4 for the relationship between pH and ANC and the various watershed attributes. The a priori expected

Table 4. Spearman correlations between mean (1974-1983) surface water pH and ANC ($\mu\text{eq/L}$) and watershed attributes for 463 headwater lakes in Adirondacks (FIN data)

Variable name	Candidate ^a variable	pH ^b				ANC ^b				Expected direction of relationship (for candidate variables only) ^c
		All data		Calibration data		All data		Calibration data		
		r	p	r	p	r	p	r	p	
<u>Morphologic and Physiographic</u>										
LAKE_Ad	Y	0.35	≤0.01	0.27	≤0.01	0.20	≤0.01	0.12	0.18	+
WTRSHD_Rd	Y	-0.02	0.78	0.04	0.64	0.04	0.59	0.04	0.69	+
DRAIN_Ad	Y	0.32	≤0.01	0.30	≤0.01	0.20	≤0.01	0.16	0.07	+
LAKE_DEV	Y	0.06	0.40	-0.05	0.57	-0.06	0.41	-0.15	0.09	?
LAKE_E	Y	-0.49	≤0.01	-0.40	≤0.01	-0.42	≤0.01	-0.31	≤0.01	?
ASPECT_S	Y	-0.01	0.86	-0.01	0.87	-0.07	0.34	-0.03	0.72	?
ASPECT_N	N	0.06	0.37	0.05	0.58	0.12	0.09	0.07	0.42	
LAKE_Ve	N	0.31	≤0.01	0.36	≤0.01	0.11	0.22	0.17	0.15	
<u>Hydrologic</u>										
RUNOFF	Y	-0.48	≤0.01	-0.42	≤0.01	-0.44	≤0.01	-0.41	≤0.01	
<u>Atmospheric</u>										
PPT	Y	-0.51	≤0.01	-0.45	≤0.01	-0.49	≤0.01	-0.45	≤0.01	-
H_WET	Y	-0.56	≤0.01	-0.53	≤0.01	-0.52	≤0.01	-0.52	≤0.01	-
NO3_WET	Y	-0.55	≤0.01	-0.52	≤0.01	-0.51	≤0.01	-0.51	≤0.01	-
SO4_WET	Y	-0.52	≤0.01	-0.46	≤0.01	-0.50	≤0.01	-0.47	≤0.01	-
SO4_NO3	N	-0.53	≤0.01	-0.49	≤0.01	-0.50	≤0.01	-0.48	≤0.01	-
H_CONC	Y	-0.53	≤0.01	-0.53	≤0.01	-0.54	≤0.01	-0.58	≤0.01	-
NO3_CONC	Y	-0.56	≤0.01	-0.57	≤0.01	-0.57	≤0.01	-0.61	≤0.01	-
SO4_CONC	Y	-0.55	≤0.01	-0.53	≤0.01	-0.57	≤0.01	-0.57	≤0.01	-
<u>Physical Soil Type</u>										
RELIEF_R	Y	0.11	0.09	0.13	0.14	0.19	≤0.01	0.19	0.03	
STONEYP	Y	0.12	0.08	0.10	0.25	0.12	0.08	0.11	0.24	either
ROCK_P	Y	0.06	0.38	0.07	0.45	0.08	0.27	0.08	0.37	
HYDRO_A	Y	0.26	≤0.01	0.23	≤0.01	0.25	≤0.01	0.25	≤0.01	+
HYDRO_B	Y	0.24	≤0.01	0.29	≤0.01	0.23	≤0.01	0.28	≤0.01	+
HYDRO_C	Y	-0.14	0.03	-0.12	0.16	-0.15	0.03	-0.17	0.06	
HYDRO_D	N	0.06	0.38	0.11	0.19	0.07	0.34	0.09	0.30	
STEEPM_P	Y	-0.26	≤0.01	-0.21	0.02	-0.28	≤0.01	-0.21	0.02	
STEEPV_P	Y	0.07	0.28	0.05	0.55	0.06	0.39	0.04	0.67	
SHL2_BP	Y	-0.07	0.28	-0.05	0.60	-0.05	0.50	-0.03	0.71	?

Table 4. (continued)

Variable name	Candidatea variable	pH ^b				ANC ^b				Expected direction of relationship (for candidate variables only) ^c
		All data		Calibration data		All data		Calibration data		
		r	p	r	p	r	p	r	p	
<u>Physical Soil Type (continued)</u>										
SHL2_P_P	Y	-0.27	≤0.01	-0.26	≤0.01	-0.28	≤0.01	-0.28	≤0.01	?
SHL2_R_P	Y	-0.20	≤0.01	-0.17	0.05	-0.17	≤0.01	-0.17	0.06	?
SHL1_B_P	N	-0.07	0.28	-0.05	0.60	-0.05	0.50	-0.03	0.71	
SHL1_P_P	N	-0.08	0.21	-0.06	0.48	-0.05	0.51	-0.02	0.79	
SHL1_R_P	N	-0.09	0.15	-0.07	0.40	-0.06	0.37	-0.05	0.57	
SHL1_Z_P	N	-0.08	0.21	-0.06	0.48	-0.05	0.51	-0.02	0.79	
SHL2_Z_P	N	-0.27	≤0.01	-0.26	≤0.01	-0.28	≤0.01	-0.28	≤0.01	
DPTH_B_U	Y	0.06	0.36	0.04	0.67	0.03	0.63	0.02	0.80	?
DPTH_P_U	Y	0.15	0.02	0.12	0.16	0.11	0.11	0.10	0.28	?
DPTH_R_U	Y	0.15	0.02	0.13	0.13	0.13	0.07	0.12	0.19	?
<u>Geology</u>										
ROCK12_P	Y	-0.18	≤0.01	-0.14	0.10	-0.22	≤0.01	-0.15	0.10	-
ROCK1_P	N	0.16	≤0.01	0.16	0.07	0.15	0.04	0.18	0.04	
ROCK2_P	N	-0.23	≤0.01	-0.20	0.02	-0.25	≤0.01	-0.23	≤0.01	
ROCK3_P	N	0.18	≤0.01	0.19	0.03	0.20	≤0.01	0.20	0.03	
ROCK4_P	N	0.08	0.21	0.02	0.78	0.14	0.04	0.04	0.63	
<u>Chemical Soil Type</u>										
ACID_P	Y	-0.20	≤0.01	-0.14	0.10	-0.16	0.02	-0.13	0.16	?
BSA_L_P	Y	0.17	≤0.01	0.11	0.22	0.14	0.05	0.09	0.32	?
BSA_M_P	N	-0.09	0.18	0.01	0.93	-0.07	0.31	0.01	0.89	
BSC_L_P	Y	-0.04	0.56	-0.06	0.52	-0.05	0.48	-0.05	0.54	-
BSC_M_P	N	0.12	0.07	0.18	0.04	0.10	0.14	0.14	0.12	
CECS_L_P	Y	0.34	≤0.01	0.28	≤0.01	0.30	≤0.01	0.25	≤0.01	?
CEC_L_P	N	0.17	≤0.01	0.23	≤0.01	0.15	0.03	0.18	0.05	
OM_H_P	Y	-0.22	≤0.01	-0.18	0.03	-0.22	≤0.01	-0.22	≤0.01	-
PH_VL_P	Y	-0.16	0.02	-0.12	0.16	-0.13	0.07	-0.11	0.23	
PHC_L_P	N	0.00	1.00	-0.01	0.88	-0.02	0.80	-0.02	0.79	
PHC_VL_P	Y	0.19	≤0.01	0.14	0.10	0.15	0.04	0.12	0.17	-
ACID_EX	Y	-0.21	≤0.01	-0.18	0.04	-0.19	≤0.01	-0.18	0.05	
CEC	Y	-0.18	≤0.01	-0.14	0.10	-0.14	0.04	-0.12	0.18	
ORG_MAT	Y	-0.23	≤0.01	-0.18	0.04	-0.16	0.02	-0.14	0.13	

Table 4. (continued)

Variable name	Candidate variable	pH ^b				ANC ^b				Expected direction of relationship (for candidate variables only) ^c
		All data		Calibration data		All data		Calibration data		
		r	p	r	p	r	p	r	p	
<u>Forest Cover</u>										
CONFR2_P	Y	-0.42	≤0.01	-0.41	≤0.01	-0.39	≤0.01	-0.34	≤0.01	+ + either
HRDWD2_P	Y	0.25	≤0.01	0.26	≤0.01	0.17	≤0.01	0.17	0.06	
NONFR2_P ^d	Y	0.29	≤0.01	0.26	≤0.01	0.24	≤0.01	0.20	0.03	
MIXED2_P	N	0.15	0.02	0.13	0.14	0.20	≤0.01	0.17	0.06	
<u>Wetland Type</u>										
WTLND_PP	Y	-0.20	≤0.01	-0.12	0.17	-0.17	0.02	-0.11	0.21	either
VACID_PP	Y	-0.15	0.03	-0.19	0.03	-0.15	0.03	-0.21	0.02	either
NACID_PP	Y	-0.03	0.69	0.17	0.05	-0.03	0.62	0.12	0.18	
MACID_PP	N	0.21	≤0.01	0.22	≤0.01	0.19	≤0.01	0.18	0.04	?
OTHER_PP	N	0.12	0.07	0.10	0.25	0.07	0.32	0.06	0.51	
WTLND_PW ^d	Y	-0.13	0.04	-0.04	0.65	-0.10	0.15	-0.01	0.91	either
VACID_PW ^d	Y	-0.12	0.08	-0.11	0.21	-0.07	0.28	-0.06	0.48	either
NACID_PW ^d	Y	0.03	0.67	0.17	0.05	0.03	0.69	0.16	0.08	
MACID_PW	N	0.20	≤0.01	0.27	≤0.01	0.15	0.04	0.21	0.02	?
OTHER_PW	N	0.14	0.03	0.16	0.06	0.07	0.29	0.07	0.44	
WTLND_PL ^d	Y	-0.14	0.04	-0.02	0.81	-0.08	0.23	0.00	1.00	either
VACID_PL ^d	Y	-0.11	0.11	-0.10	0.26	-0.05	0.51	-0.06	0.54	either
NACID_PL ^d	Y	0.02	0.71	0.17	0.05	0.05	0.47	0.17	0.07	
MACID_PL	N	0.18	≤0.01	0.28	≤0.01	0.13	0.05	0.21	0.02	?
OTHER_PL	N	0.13	0.05	0.16	0.07	0.06	0.35	0.07	0.42	
<u>Disturbance</u>										
DISTRB_P	Y	0.13	0.04	0.06	0.47	0.10	0.14	0.01	0.91	+
BVRINDEX ^d	Y	-0.29	≤0.01	-0.21	0.02	-0.27	≤0.01	-0.24	≤0.01	either
CABN78_R	Y	0.23	≤0.01	0.17	0.05	0.16	0.02	0.12	0.17	+
BURNED_P	N	0.16	≤0.01	0.08	0.35	0.14	0.05	0.01	0.89	+
DENUDE_P	N	0.04	0.59	-0.01	0.94	0.00	0.98	-0.06	0.53	+
LOG_SH_P	N	0.13	0.05	0.16	0.06	0.12	0.08	0.10	0.27	+

^aCandidate variables were 57 variables selected as input variables to the MLR analysis.

Y means yes this variable was included, and N means no it was not.

^bUnless otherwise indicated: for pH, n = 234 for all data and n = 135 for the calibration data; for ANC, n = 208 for all data and n = 124 for the calibration data.

^cA priori expectations are provided to aid the reader unfamiliar with hypotheses about lake acidification in the literature. "Either" means arguments could be made to support both positive or negative correlations; "?" means we did not have an expectation.

^dThe variable is log₁₀ transformed for all statistical analyses.

^en = 136 for pH, and n = 124 for ANC.

direction (positive or negative) of the relationship between many of the watershed attributes and lake pH and ANC are also listed in this table. For comparison, the correlations for the same watershed attributes using the NSW subset of 46 lakes are given in Table 4.

3.1.2.1 Watershed and Lake Morphology, Physiography, and Hydrology

Generally, physiography affects the amount of water and accompanying acids that move along various hydrologic pathways to the streams and lakes. Lakes at higher elevations receive more precipitation and acidic deposition as a result of orographic effects, and as expected, lake elevation (LAKE-E) was strongly correlated (inversely) with lake ANC ($r = -0.42$, $p \leq 0.01$) and with lake pH ($r = -0.49$, $p \leq 0.01$) for all headwater lakes with chemistry data. These values were very close to the correlation coefficients for runoff, precipitation, and wet deposition. Watershed drainage area (DRAIN-A) was positively correlated with lake ANC ($r = 0.20$, $p \leq 0.01$) and lake pH ($r = 0.32$, $p \leq 0.01$). An explanation for this association may be that large watersheds with longer hydrologic pathways for water flowing into the lakes have a greater contact time between water and soil and, thus, a greater capacity to neutralize atmospherically deposited acids. However, the ratio of watershed to lake area (WTRSHD_R) was not significantly correlated to lake pH or ANC.

In-lake processes, such as sulfate reduction and primary productivity, can increase lake ANC and pH. Processes occurring in the littoral zone may generate alkalinity or net acidity, depending on the vegetation type. The relationships between lake chemistry and several characteristics that may be surrogates for in-lake processes of

Adirondack headwater lakes were examined: lake volume type (LAKE-V), lake area (LAKE-A), and lake development ratio (LAKE_DEV) (Tables 4 and 5).

The influence of in-lake processes on lake pH and ANC does not appear to be strong for the Adirondack headwater lakes analyzed below. A significant positive correlation existed between lake area and lake chemistry (Tables 4 and 5). The positive correlation between pH and lake volume was expected because a larger lake volume may reflect a slower flushing rate and, thus, a greater residence time of water, fostering internal production of alkalinity. The absence of a significant correlation between ANC and lake volume, however, indicates the need for caution in interpreting the pH results. The lake development ratio, defined as the perimeter of the lake divided by the perimeter of a circle with the same area as that of the lake (Hutchinson 1957), did not have significant correlations with lake chemistry.

The dominant slope aspect of each watershed might be related to surface water chemistry because of potentially greater wet and dry deposition on slopes facing the prevailing wind direction or perhaps because of differences in hydrology, snowmelt, soils, and vegetation types on slopes with different aspects. Significant correlations between dominant watershed aspect (ASPECT-N and ASPECTS) and lake chemistry were not found.

3.1.2.2 Atmospheric Inputs

The patterns of atmospheric inputs expressed as wet deposition rate and concentration of hydrogen ion (H^+) (Figs. 2 and 3), nitrate anion (NO_3^-) (Fig. 4), and sulfate (SO_4^{2-}) anion are similar

Table 5. Spearman correlations between National Surface Water Survey measurements of pH and ANC ($\mu\text{eq/L}$) and watershed attributes for 46 headwater lakes in Adirondacks (NSWS data)

Variable name	pH ^a		ANC ^a	
	r	p	r	p
<u>Morphologic and Physiographic</u>				
LAKE-A	0.38	≤ 0.01	0.39	≤ 0.01
WTRSHD_R	-0.02	0.88	-0.03	0.85
DRAIN-A	0.32	0.03	0.33	0.02
LAKE-DEV	0.00	0.99	0.01	0.97
LAKE-E	-0.67	≤ 0.01	-0.66	≤ 0.01
ASPECT-S	-0.07	0.62	-0.06	0.71
ASPECT_N	0.09	0.56	0.08	0.61
LAKE_V ^b	0.20	0.30	0.19	0.31
<u>Hydrologic</u>				
RUNOFF	-0.65	≤ 0.01	-0.62	≤ 0.01
<u>Atmospheric</u>				
PPT	-0.66	≤ 0.01	-0.64	≤ 0.01
H-WET	-0.67	≤ 0.01	-0.65	≤ 0.01
NO3_WET	-0.67	≤ 0.01	-0.65	≤ 0.01
SO4_WET	-0.67	≤ 0.01	-0.66	≤ 0.01
SO4_NO3	-0.69	≤ 0.01	-0.67	≤ 0.01
H_CONC	-0.63	≤ 0.01	-0.62	≤ 0.01
NO3_CONC	-0.60	≤ 0.01	-0.59	≤ 0.01
SO4_CONC	-0.62	≤ 0.01	-0.62	≤ 0.01
<u>Physical Soil Type</u>				
RELIEF-R	0.23	0.12	0.24	0.11
STONEYP	0.15	0.32	0.15	0.32
ROCK-P	0.00	1.00	0.01	0.92
HYDRO_A	0.34	0.02	0.36	≤ 0.01
HYDRO_B	0.33	0.02	0.33	0.03
HYDRO_C	-0.12	0.44	-0.13	0.39
HYDRO_D	-0.03	0.83	-0.03	0.84
STEEPM_P	-0.17	0.27	-0.17	0.27
STEEPV_P	0.21	0.17	0.22	0.13
SHL2_B_P	-0.11	0.45	-0.12	0.43
SHL2_P_P	-0.27	0.07	-0.28	0.06
SHL2_R_P	-0.21	0.16	-0.20	0.18
SHL1_B_P	-0.11	0.45	-0.12	0.43
SHL1_P_P	-0.13	0.38	-0.13	0.39

Table 5. (continued)

Variable Name	pH ^a		ANC ^a	
	r	P	r	P
<u>Physical Soil Type (continued)</u>				
SHL1_R_P	-0.14	0.35	-0.14	0.36
SHL1_Z_P	-0.13	0.38	-0.13	0.39
SHL2_Z_P	-0.27	0.07	-0.28	0.06
DPTH_B_U	0.12	0.44	0.12	0.43
DPTH_P_U	0.20	0.17	0.20	0.18
DPTH_R_U	0.16	0.28	0.16	0.30
<u>Geology</u>				
ROCK12_P	-0.26	0.09	-0.21	0.15
ROCK1_P	0.24	0.11	0.26	0.08
ROCK2_P	-0.34	0.02	-0.33	0.02
ROCK3_P	0.13	0.39	0.06	0.68
ROCK4_P	0.23	0.12	0.24	0.11
<u>Chemical Soil Type</u>				
ACID-P	-0.24	0.11	-0.26	0.09
BSA_L_P	0.11	0.47	0.09	0.57
BSA_M_P	-0.00	0.99	-0.01	0.97
BSC_L_P	-0.01	0.93	-0.04	0.81
BSC_M_P	-0.02	0.89	-0.03	0.84
CECS_L_P	0.26	0.08	0.29	0.05
CEC_L_P	0.12	0.41	0.14	0.35
OM_H_P	-0.16	0.28	-0.17	0.26
PH_VL_P	-0.19	0.21	-0.21	0.16
PHC_L_P	0.02	0.91	-0.01	0.97
PHC_VL_P	0.13	0.39	0.10	0.49
ACID-EX	-0.23	0.12	-0.25	0.09
CEC	-0.16	0.29	-0.16	0.28
ORG_MAT	-0.24	0.10	-0.24	0.10
<u>Forest Cover</u>				
CONFR2_P	-0.39	≤0.01	-0.40	≤0.01
HRDWD2_P	0.24	0.11	0.27	0.07
NONFR2_P	0.17	0.25	0.18	0.24
MIXED2_P	0.18	0.23	0.15	0.32
<u>Wetland Type</u>				
WTLND_PP	-0.40	<0.01	-0.40	≤0.01
VACID_PP	-0.23	-0.12	-0.24	0.10
NACID_PP	-0.36	≤0.01	-0.37	≤0.01

Table 5. (continued)

Variable name	pH ^a		ANC ^a	
	r	P	r	P
Wetland Type (continued)				
MACID_PP	0.01	0.95	0.04	0.79
OTHER_PP	0.13	0.37	0.13	0.39
WTLND_PW	-0.33	0.03	-0.32	0.03
VACID_PW	-0.27	0.07	-0.27	0.07
NACID_PW	-0.15	0.32	-0.14	0.36
MACID_PW	0.04	0.81	0.06	0.70
OTHER_PW	-0.02	0.90	-0.00	0.98
WTLND_PL	-0.29	0.05	-0.27	0.07
VACID_PL	-0.26	0.08	-0.26	0.08
NACID_PL	-0.15	0.31	-0.14	0.34
MACID_PL	0.01	0.97	0.03	0.85
OTHER_PL	0.01	0.97	0.02	0.90
Disturbance				
DISTRB_P	0.24	0.10	0.28	0.06
BVRINDEX	-0.36	≤0.01	-0.35	0.02
CABN78_R	0.11	0.48	0.11	0.46
BURNED_P	0.32	0.03	0.32	0.03
DENUDE_P	0.06	0.68	0.09	0.56
LOG_SH_P	0.22	0.14	0.22	0.14

^a_n = 46 unless otherwise indicated.

^b_n = 29 for LAKE-V.

for the Adirondack region. Spearman correlation coefficients between the mean annual concentration or wet deposition of, on the one hand, H^+ , NO_3^- , SO_4^{2-} , and $\text{SO}_4^{2-} + \text{NO}_3^-$ and, on the other hand, lake chemistry are also very similar, ranging from -0.50 to -0.57 ($p \leq 0.01$) (Tables 4 and 5). Correlation coefficients for precipitation with lake chemistry are slightly lower, -0.49 to -0.51 ($p \leq 0.01$). All of the atmospheric input variables are very highly intercorrelated (Sect. 3.2.1.2). Atmospheric input variables had the highest correlations with water chemistry that were found in this study, a finding consistent with the hypothesis that the chemistry of the atmospheric input to individual watersheds is important in regulating lake pH and ANC. Unfortunately, data for dry deposition are not available. The variation in wet plus dry acidic inputs might account for a larger proportion of variation in lake chemistry than wet deposition alone and might help clarify the importance of atmospheric deposition in lake acidification.

3.1.2.3 Watershed Soils

Several physical and chemical soil characteristics were investigated to test for relationships between watershed soil properties and lake chemistry. The soil hydrologic group, a Soil Conservation Service interpretation assigned to each soil series to designate the potential of that series to generate surface runoff, is derived from soil properties such as permeability, slope, and depth to bedrock or to impermeable soil horizon. Lakes in watersheds with large areas having low runoff potential [high infiltration potential (hydrologic groups A and B)] would be expected to have relatively high ANC and pH as a

result of acid neutralization as drainage waters follow relatively deep subsurface pathways to streams and lakes. Conversely, lakes in watersheds with large areas of soils with high runoff potential (hydrologic groups C and D) would be expected to have low ANC and pH as a result of direct, rapid runoff of precipitation into the surface waters. This was the pattern observed (Tables 4 and 5). Correlation coefficients for HYDRO_A and HYDRO_B were significant and ranged from 0.23 to 0.36; coefficients for HYDRO_C and HYDRO_D were around -0.14 and often were not significant. Thus, the soil hydrologic group appears, from this limited data set, to be a reasonable indicator of watershed hydrologic behavior and neutralizing capacity.

Slope categories (as assigned to soil series by the Soil Conservation Service) from the soils data set showed the expected correlations: steep slopes were negatively correlated with pH and ANC. The percentage of a watershed with slopes >15% was inversely correlated with lake ANC ($r = -0.28$, $p \leq 0.01$) and lake pH ($r = -0.26$, $p \leq 0.01$).

Soil stoniness and percent of the watershed in rock outcrop were not significantly correlated with lake ANC or pH. Significant negative correlations ($r = -0.17$ to -0.28 , $p \leq 0.01$) were obtained for lake chemistry and depth to a low-permeability horizon (SHL2_P_P) and depth to a root-restrictive zone (SHL2_R_P). Correlations with depth to bedrock were not significant (Tables 4 and 5). These last three variables should have indicated the extent of shallow soils.

Soil chemical characteristics investigated included pH, base saturation, cation exchange capacity, exchangeable bases, and percent of organic matter. Significant inverse correlations were found between

headwater lake pH and ANC and the percentage of watersheds having soils with high extractable acidity (ACID-P and ACID-EX), high organic matter content (OM_H_P and ORG_MAT), and a cation exchange capacity ≤ 20 meq/100 g (CECS_L_P). In other words, watersheds with large areas having high extractable acidity, high organic matter, a high cation exchange capacity, or a combination of these factors also had lakes with low pH and ANC. Significant correlations of base saturation, soil pH, or exchangeable bases with lake chemistry were not found.

3.1.2.4 Watershed Geology

The frequency distribution of selected Adirondack watersheds based on dominant bedrock sensitivity categories (Norton et al. 1982) is shown in Fig. 6. The New York state geologic map was classified into four acid-neutralizing-capacity groups, ranging from the low to none category (granite and quartz sandstone types) to the infinite category (highly fossiliferous sediments and limestone or dolostone types). The percentage of watershed area having low to medium buffering capacity bedrock (ROCK2_P, category 2) was inversely correlated with lake ANC ($r = -0.25$, $p \leq 0.01$) and lake pH ($r = -0.23$, $p \leq 0.01$), i.e., the more area a watershed has of a bedrock with low to medium buffering capacity, the lower the lake ANC and pH. Similarly, the percentage of watershed area having medium to high buffering capacity bedrock (ROCK3_P, category 3) was positively correlated with lake ANC ($r = 0.20$, $p \leq 0.01$) and lake pH ($r = 0.18$, $p \leq 0.01$). i.e., the more area a watershed has with medium to high buffering capacity bedrock, the higher the lake ANC and pH. Thus, bedrock type seemed to be related to

surface water chemistry. Higher correlations might be obtained if data were available on thickness and buffering capacity of the surficial deposits overlying the bedrock.

3.1.2.5 Watershed Vegetation

In this section, the influence of both forest cover and wetland type on lake chemistry is discussed. The canopy of a deciduous forest surrounding a lake may help neutralize acidic deposition; rain percolating through a coniferous forest, however, often increases in acidity. Humic matter can acidify water percolating through it and contribute hydrogen ions to the surface waters (Viro 1974, Marcus et al. 1983). Therefore, land-cover types were grouped into deciduous, conifer, mixed, and nonforested classes for this analysis. Deciduous and conifer are the dominant land-cover types. The percentage of a land-cover class in a watershed was determined based on the land area (lake area not included). For Adirondack headwater lakes, the null hypothesis examined was that there is no significant association between pH or ANC and percentage of watershed in a land-cover class.

Almost all land-cover types had significant and large correlations with ANC (Tables 4 and 5). The magnitude of the correlations for conifers (CONFR2_P) as among the largest values ($r = -0.39$, $p \leq 0.01$ for ANC; $r = -0.42$, $p \leq 0.01$ for pH). The signs of the correlation coefficients support the hypothesis that coniferous vegetation in a lake's watershed is associated with lower pH and ANC values in the lake and that nonconiferous vegetation is associated with higher pH and ANC values. Therefore, the null hypothesis was rejected, and the conclusion was that the percentage of the watershed in certain

land-cover types was significantly associated with the water chemistry of the Adirondack headwater lakes.

Another hypothesis tested was that no significant difference exists between the pH and ANC in the headwater lakes of watersheds defined as coniferous vs those classified as nonconiferous. Because it is believed that conifers located near the lakes and a relatively small proportion of the watershed in conifers could affect lake chemistry, a coniferous watershed was defined as one having at least 33% of the watershed in coniferous vegetation (Table 6). The nonparametric Wilcoxon test showed a significant difference between the pH and ANC in lakes of coniferous watersheds and nonconiferous watersheds (i.e., <33% of watershed in conifers). The lakes in coniferous watersheds were acidic with a median pH of 5.0 ($n = 55$) and a median ANC of $-3.2 \mu\text{eq/L}$ ($n = 54$). Headwater lakes in nonconiferous watersheds had a median pH of 6.2 ($n = 179$) and a median ANC of $33.5 \mu\text{eq/L}$ ($n = 154$), thus indicating lakes less sensitive to acidification.

High concentrations of organic acids in lakes are indicative of one type of natural acidification process that is typically associated with naturally acidic wetlands or acidic and humic-rich soils within a watershed. Dark-water lakes are often associated with bogs and bog forests and may be naturally acidified by these adjacent wetlands (Bogucki and Gruendling 1982, Gorham et al. 1985). Therefore, we evaluated if a significant association existed between, on the one hand, pH and ANC or concentration of dissolved organic matter in Adirondack headwater lakes (as indicated by mean visual color) and, on the other hand, percentage of watershed in wetland or type of wetland in a watershed.

Table 6. Summary of Wilcoxon test results for lake chemistry with wetland vegetation classes (significant at $p \leq 0.05$)

	Median values for wetland classes ^a					
	$\geq 20\%$	$< 20\%$	$\geq 33\%$	$< 33\%$	$\geq 50\%$	$< 50\%$
pH						
WTLND_PP	5.4 ^b (n = 132)	6.1 (n = 102)	5.3 (n = 100)	6.1 (n = 134)	5.1 (n = 78)	6.1 (n = 156)
MACID_PP	6.4 (n = 32)	5.7 (n = 202)				
WTLND_PW	5.2 (n = 50)	6.0 (n = 184)	4.8 (n = 15)	6.0 (n = 219)	4.8 (n = 5)	6.0 (n = 229)
MACID_PW	6.4 ^c (n = 36)	5.7 (n = 198)	6.6 (n = 34)	5.7 (n = 200)	6.5 (n = 29)	5.9 (n = 205)
ANC						
WTLND_PP					-0.5 (n = 74)	33.5 (n = 134)
VACID_PP			6.3 (n = 107)	34.0 (n = 101)	6.0 (n = 97)	36.0 (n = 111)
MACID_PP	73.5 (n = 24)	10.1 (n = 184)				
WTLND_PW	-4.5 (n = 44)	19.5 (n = 164)	-9.5 (n = 12)	14.0 (n = 196)		
MACID_PW	69.0 (n = 27)	9.0 (n = 181)	73.5 (n = 26)	9.8 (n = 182)	69.0 (n = 23)	10.6 (n = 185)
Color						
WTLND_PW	35.0 (n = 11)	15.0 (n = 35)				

^aPercentage of watershed (WTLND_PP and WTLND_PW) or percentage of total wetlands (MACID_PP, MACID_PW, VACID_PP, and VACID_PW).

^bMedian and sample size for the class where the perimeter of the lake in all wetland types is $\geq 20\%$ of the lake's perimeter.

^cMedian and sample size for the class where the moderately acid wetland type represents $\geq 20\%$ of the lake's watershed.

Three groups of variables were created to test hypotheses about wetlands. One variable expresses the area of wetlands as a percentage of the total watershed area (identified by the suffix _PW for "percent of watershed"), another variable expresses the length of wetland in direct contact with the shoreline of a lake as a percentage of the total shoreline (identified by _PP for "percent of perimeter"), and a third variable expresses the wetland area in the watershed as a percentage of lake area (identified by _PL for "percent of lake"). The National Wetlands Inventory cover types were grouped into four categories--very acid, moderately acid, nonacid, and "other." The very acid cover type was predominantly needle-leaved evergreen forest and scrub/shrub vegetation, including bog mats; the nonacid cover type was predominantly broad-leaved deciduous forest and scrub/shrub vegetation. The moderately acid cover type included a vegetation mixture of needle-leaved evergreen, broad-leaved deciduous forest, and scrub/shrub vegetation. Persistent emergent vegetation, dead forest, open water, etc., were included in the "other" category. Percentages for these wetland types and total wetland within a watershed were calculated and used in analyses.

For wetlands the null hypothesis tested was that there is no significant correlation between pH or total ANC in Adirondack headwater lakes and the percentage of wetland vegetation. The results of the analyses are variable. Total wetland percentages generally had significant correlations ($p \leq 0.05$) with lake pH and ANC measurements (Tables 4 and 5). A negative correlation coefficient indicated that as the amount of wetland increased, the lake pH and ANC decreased.

The very acid wetland type expressed as a percentage of lake perimeter (VACID_PP) was significantly and inversely correlated with lake ANC and pH ($r = -0.15$, $p = 0.03$). However, when this wetland type was expressed as a percentage of watershed or lake area, correlations were not significant. Unexpectedly, the moderately acid wetland type consistently had a significant positive correlation with lake pH and ANC.

The nonparametric Wilcoxon test was used to make additional tests of the association between wetlands and lake chemistry for the headwater lakes. Tests were made to see if there is a significant difference between the pH or ANC for a lake that has a high percentage of total wetland (_PP, _PW), very acid wetland type (VACID), moderately acid wetland type (MACID), and those that do not. Because there was no a priori knowledge about how to define "a high percentage of wetland," several different definitions were tested. Significant results ($p \leq 0.05$) are reported for tests in which the lakes with a high percentage of wetland are those with total wetland or wetland type $\geq 20\%$, $\geq 33\%$, and $\geq 50\%$ (Table 6). Twice these tests indicated a significantly lower ANC for the very acid wetland class when compared with the "other" wetland class; however, the median ANC for both of these classes was always $< 50 \mu\text{eq/L}$. Although significant correlations exist between lake chemistry and the amount of wetland or the amount of a wetland type, a strong association cannot be made between the very acid wetland type and low pH and ANC measurements in associated lakes. Stronger relationships existed for total wetlands and moderately acid wetlands. Again, moderately acid wetlands indicated

less-acidic lakes (pH difference ≥ 0.6), and a high percentage of total wetlands indicated more-acidic lakes (pH difference ≥ 0.7).

The average percentage of total wetlands in the watersheds of Adirondack headwater lakes is somewhat higher for colored lakes (15%) when compared with clear lakes (11%). However, the percentages for total wetland and wetland cover types were seldom significantly correlated with the visual color measurement available for 101 of the Adirondack headwater lakes. The Wilcoxon test showed that Adirondack headwater lakes with a visual color value >2.5 have significantly ($p \leq 0.05$) more of the very acid wetland type in their watersheds than clear lakes. However, the difference in the average percentage of the watershed in wetlands is only 4% for the two color classes. Color (PCU) measured in the NSW is significantly correlated with dissolved organic carbon ($r = 0.74$, $p \leq 0.01$, $n = 46$), which would be expected to be abundant in lakes with a high percentage of wetlands. Lakes with $>20\%$ of their watershed in wetlands are darker in color by 20 PCU. Although the mean visual color measurements ($n = 10$) from FIN were not significantly correlated with color or dissolved organic carbon measurements from NSW, a relationship between these data does seem to exist based on observation of plotted data.

3.1.2.6 Watershed Disturbance

Several hypotheses that would link lake acidification, or acid neutralization, to both natural and man-induced disturbances, such as fire, logging, tree blowdown, beaver activity, and number of buildings, have been proposed (Krug and Fink 1983, Schnoor and Stumm 1985, Retzsch

et al. 1982, Rush et al. 1985). Human settlements might be associated with sewage discharges, sites of land clearing, and road building. Several hypotheses about watershed disturbance were examined for relationships to lake chemistry, with extensive analyses performed on beaver activity.

Duhaine et al. (1983) suggested evaluating the effects of suppression of fires and restriction of timber harvesting on acidification of aquatic ecosystems. Historical (ca. 1905 to 1915) data on fires and logging for Adirondack headwater watersheds were compiled. Since protection of the park was ensured in 1894, no logging and very few fires have occurred. Correlations between lake pH and ANC with the percentage of watershed burned (BURNED-P), denuded (DENUDE-P), logged (LOG_SH_P), and disturbed (DISTRB-P) are relatively low (0.13 to 0.16) (Tables 4 and 5). The single variable (DISTRB-P) represented total watershed disturbance (the sum of logged, burned, and denuded area). Most of the headwater watersheds had no indication of fire or logging influence; therefore, the correlations are based on only ~50% of the watersheds (Fig. 7). The percentage of a watershed that had been burned had the most association with lake chemistry, especially pH, of the land-cover disturbance variables.

The presence of roads, human settlements, and domestic animals, such as horses and cows, in a lake's watershed may contribute to increased alkalinity (Schnoor et al. 1985). Although very few headwater lakes in the Adirondacks have human settlements, the number of cabins in the watershed in 1978 (CABN78_R) was significantly correlated with lake pH and ANC (Table 4). This relationship was not

significant for the smaller NSW data (Table 5). These data support the belief that human settlements and associated watershed disturbance may be associated with an increase in a lake's ANC.

Relationships between lake chemistry (pH, ANC, temperature, color, dissolved oxygen, sulfate, and nitrate) and beaver activity were examined. Beaver activity was defined by the number of beaver dams present in 1968 and 1978 (data for 1978 weighted by a factor of two) normalized by lake area because beaver were expected to affect a small lake more than a large lake. Data from 1978 were given more weight because the chemistry data were from 1974 to 1984.

The total number of beaver dams on an outlet (TOTOUT_A) and of downstream dams (OUT-A) and the sum of all dams (TOTAL-A) were correlated with lake pH (r ranging from -0.27 to -0.29, $p \leq 0.01$). The Kruskal-Wallis test with a multiple comparison showed that lakes with dams present anywhere on the outlet (ANYOUT) and those with dams present on the lake at the outlet (OUTLAKE) had significantly lower pH values than lakes with dams on inlets. For this test, lakes were divided into four independent classes: those with beaver activity (1) on the inlet only, (2) on the outlet only, and (3) on the inlet and outlet and those (4) with no activity. Spearman correlations and Kruskal-Wallis test results for ANC were consistent with those for pH.

A MLR using only four beaver variables (lakes with dams only upstream, lakes with dams only on the lake at the inlet, lakes with dams only on the lake at the outlet, and lakes with dams only downstream) was performed to substantiate results of correlations and Kruskal-Wallis tests and evaluate several variables for beaver activity

at one time. These variables explained 8.7% of a lake's pH; dams on the outlet had a significant and negative correlation with pH. Beaver dam variables also explained 8% of the variation in a lake's ANC; downstream dams were the most explanatory.

Spearman correlations, Kruskal-Wallis tests, and MLR results for beaver activity and lake temperature showed no consistent pattern. Analyses of dissolved oxygen and sulfate data were not significant. Analyses on nitrate data gave conflicting results for FIN and NSW data. Spearman correlations for FIN data were often significant with consistently negative values for the number of beaver dams and lake nitrate (r ranging from -0.22 to -0.34). Significant results for NSW were positive with r values ranging from 0.32 to 0.45. The MLR model developed for the FIN nitrate data explained 10% of the variation in nitrate with dams at the outlet of a lake being most explanatory and having a negative relationship. Spearman correlations indicated that the number of beaver dams on the inlet of a lake are positively correlated with lakes having higher color values ($r = 0.28$, $p \leq 0.01$); however, MLR results were not significant.

3.1.2.7 Lake pH Color, and Dissolved Organic Carbon

The null hypothesis that current pH levels in lakes with low organic content (clear water lakes, visual color <1.5) are not significantly different from current pH levels in lakes with high concentrations of dissolved organic content (lakes colored light brown or brown, visual color >2.5) was tested by Baker and Harvey (1984) for Adirondack lakes. They found that the data do not support the

contention that the presence of organic acids is the major factor controlling lake acidity. This hypothesis was examined for the Adirondack headwater lakes with information on color, which are a subset of the Adirondack lakes examined by Baker and Harvey (1984). The nonparametric Wilcoxon two-sample test indicated that pH levels in clear and colored lakes are not significantly different; therefore, the null hypothesis was not rejected. For the subset of the NSW lakes in AWDB, there was a significant correlation between mean visual color and dissolved organic carbon ($r = 0.58$, $p = 0.01$, $n = 18$) and between color (PCU) and dissolved organic carbon ($r = 0.87$, $p < 0.0001$, $n = 46$). Correlations between pH and visual water color in FIN and for the subset of NSW lakes in AWDB ($n = 46$) are not significant at the $p < 0.05$ level. There is agreement with Baker and Harvey that the data do not support the idea that the presence of organic acids is the major factor controlling lake acidity.

3.2 MULTIPLE LINEAR REGRESSION

A MLR model makes it possible to determine how much of the between-lake variance for a selected lake chemistry variable (the dependent variable) can be explained by a selected set of watershed attributes ("explanatory" or "predictor" variables). We used regression models and discriminant functions to evaluate the relative importance of the anthropogenic and natural watershed variables to lake chemistry (Fig. 8).

MLR and multivariate techniques, such as discriminant analysis (see Sect. 3.3) are applied for two reasons in this study:

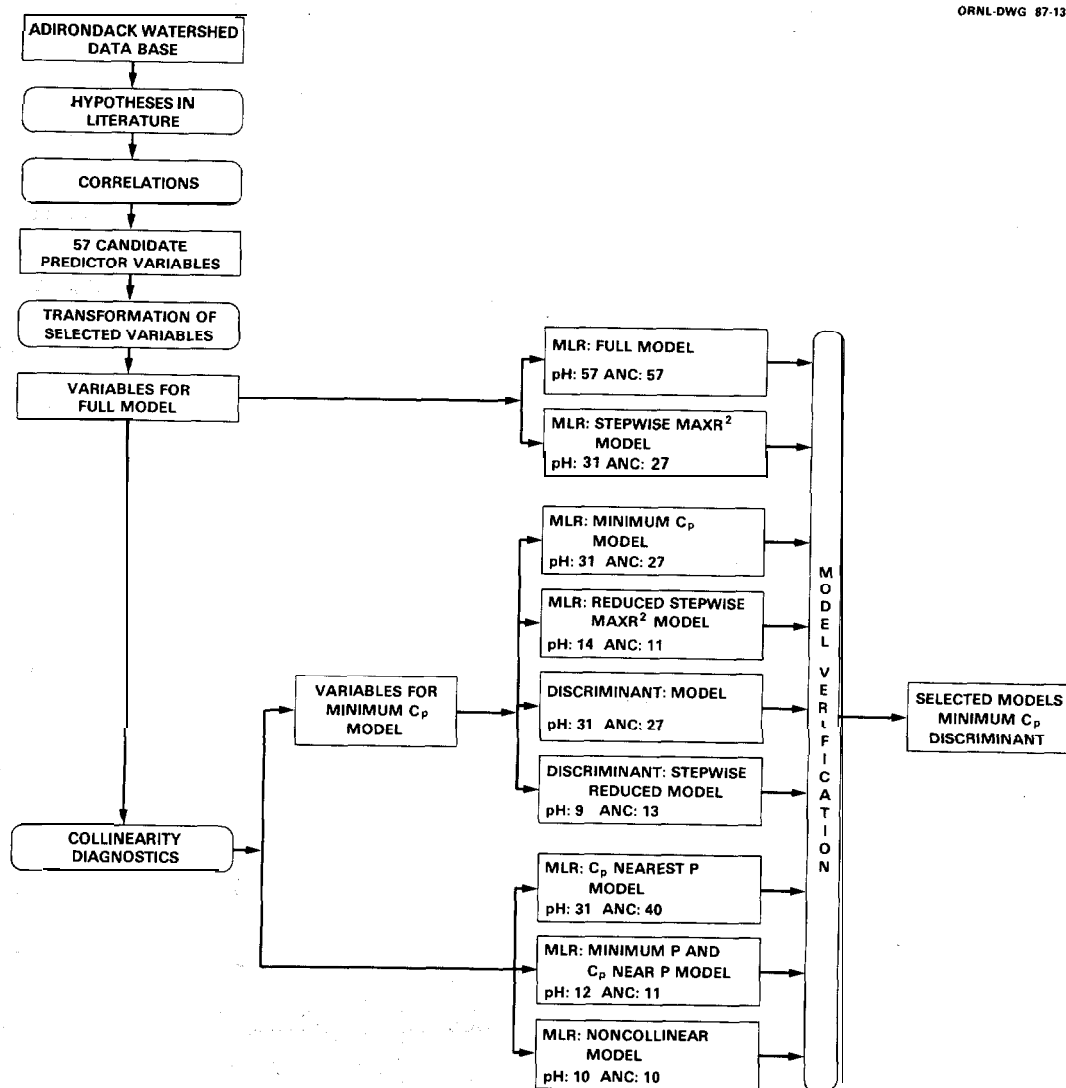


Fig. 8. Model development procedure.

1. A MLR model or a discriminant function can consider the combined effects of a number of potentially causal variables and can, therefore, provide stronger support for hypotheses about relationships involving these variables than can analyses involving only two variables at a time.
2. These models can be used to predict values of the response variables (pH or ANC) for lakes where measurements of the response variables are unavailable. For example, pH or ANC measurements are available for only slightly more than one-half of the 463 headwater lakes (248 and 231 lakes, respectively, considering both FIN and NSW data). Predictions of pH and ANC values for the other lakes by using MLR should be considerably more reliable than simple extrapolation (i.e., assuming that the percentage of lakes with pH (5 for the group of lakes without pH measurements is the same as that observed in the group of lakes having pH measurements).

3.2.1 Methods

MLR is a statistical technique in which a collection of predictor variables are used to estimate the value of a response variable. In the context of this analysis, the response variable was either pH or ANC. The predictor variables were a subset of the available watershed attributes. Only lakes having data for the complete subset of predictor variables were included in the analyses. The lakes were randomly divided into subsets to establish separate data sets for model calibration (parameter estimation) and several types of verification (Sect. 2.4).

MLR was applied to the headwater lake data after collinearity diagnosis and variable transformation. Separate models were developed using pH and ANC as dependent (response) variables (Fig. 8). The 57 candidate predictor variables for model development were selected based on knowledge of watershed and lake processes, quality of data, and results from the Spearman rank correlations discussed in Sect. 3.1 (Table 7). Some variables were not included as candidate variables because missing values would have substantially reduced the number of lakes available for analysis (e.g., lake volume). A complete discussion of methods used here is provided; however, for simplicity the reader may wish to focus on the "selected" models (minimum C_p) and reduced stepwise $MAXR^2$ models.

3.2.1.1 Data Transformations

The decision to transform variables was based on the Kolomogorov D statistic (SAS 1985) for normality of the residuals with regressions fitted by using the full model (all 57 variables). Residuals, studentized residuals, and studentized residuals with the current observation deleted (RSTUDENT) were all examined; emphasis was placed on the studentized residuals (Draper and Smith 1981). Transformation decisions for predictor variables were based on fits using the model-calibration subset. Without transformations, there were suggestions of nonnormality of the residuals for pH. Based on an examination of the horizontal pattern of dispersion in residual plots for each predictor variable, the variables shown in Table 8 were selected for transformation. When it was necessary to add constants to a variable to prevent taking the logarithm of a negative number, the

Table 7. Variables used in MLR and discriminant models for lake ANC and pH^a

Candidate variable	MLR model					Discriminant model reduced stepwise	
	Noncollinear	Minimum p and Cp near p	Selected minimum Cp	Cp nearest p ^b	Stepwise MAXR ²		Reduced stepwise MAXR ²
<u>Morphologic and Physiographic</u>							
LAKE_A							
WTRSHD_R							
DRAIN_A			ANC, pH	ANC	ANC		
LAKE_DEV			ANC, pH	ANC	ANC, pH	ANC	
LAKE_E			pH	ANC	ANC, pH	pH	pH
ASPECT_S			ANC, pH	ANC	pH		
<u>Hydrologic</u>							
RUNOFF			pH	ANC	ANC, pH	pH	
HYDTYP1	ANC, pH	ANC, pH	ANC, pH	ANC			
HYDTYP2			ANC, pH	ANC	ANC, pH		ANC
HYDTYP3	pH	pH	pH	ANC			
<u>Atmospheric</u>							
PPT					ANC, pH		
H_WET					pH		
NO3_WET				ANC	ANC		
SO4_WET					ANC, pH		
H_CONC	pH	pH	pH		ANC, pH	pH	pH
NO3_CONC	ANC	ANC	ANC	ANC	ANC	ANC	ANC
SO4_NO3					ANC, pH		
<u>Physical Soil Type</u>							
RELIEF_R			ANC, pH	ANC	ANC	pH	pH
STONEYP			ANC, pH	ANC	pH	pH	ANC
ROCK_P					ANC		
HYDRO_A	ANC, pH	ANC, pH	ANC, pH	ANC	ANC, pH	pH	pH
HYDRO_B	ANC, pH	ANC, pH	ANC, pH	ANC	ANC, pH	ANC	
HYDRO_C				ANC			
STEEPP_P			ANC, pH	ANC	ANC, pH	ANC, pH	pH
STEEPP_P	ANC, pH	ANC, pH	ANC, pH	ANC	ANC, pH	pH	
SHL2_B_P					pH		
SHL2_P_P				ANC			
SHL2_R_P							
DPTH_B_U			ANC	ANC		ANC	
DPTH_P_U							
DPTH_R_U							

Table 7. (continued)

Candidate variable	MLR model				Stepwise MAXR ²	Reduced stepwise MAXR ²	Discriminant model reduced stepwise
	Noncollinear	Minimum p and Cp near p	Selected minimum Cp	Cp nearest p ^b			
<u>Geology</u>							
ROCK12_P			ANC, pH	ANC	ANC, pH		
<u>Chemical Soil Type</u>							
ACID_P							
BSA_L_P				ANC	pH		
BSC_L_P							
CECS_L_P			ANC, pH	ANC			
OM_H_P			pH	ANC	pH	pH	
PH_VL_P					pH		
PHC_VL_P			pH	ANC	pH		
ACID_EX				ANC			
CEC			ANC, pH	ANC		pH	
ORG_MAT					pH		
<u>Forest Cover</u>							
CONFR2_P		ANC	ANC, pH	ANC	pH	ANC, pH	ANC, pH
HRDWD2_P			ANC, pH	ANC	ANC, pH	ANC	ANC, pH
NONFR2_P	ANC, pH	ANC, pH	ANC, pH	ANC	ANC	ANC	ANC
<u>Wetland Type</u>							
WTLND_PP		pH	ANC, pH	ANC		ANC	ANC, pH
VACID_PP			ANC, pH	ANC	ANC, pH	pH	
NACID_PP			ANC, pH	ANC			ANC
WTLND_PW			ANC, pH	ANC		ANC, pH	ANC, pH
VACID_PW				ANC	ANC, pH		
NACID_PW	ANC	ANC, pH	ANC, pH	ANC	ANC, pH		ANC
WTLND_PL				ANC			
VACID_PL					ANC, pH		
NACID_PL				ANC	ANC, pH		
<u>Disturbance</u>							
DISTRB_P	ANC, pH	ANC, pH	ANC, pH	ANC			ANC
BVRINDEX	ANC, pH	ANC, pH	ANC, pH	ANC	ANC, pH	ANC, pH	ANC
CABN78_R	ANC, pH	ANC, pH	ANC, pH	ANC	ANC		ANC

^aAn entry of pH or ANC in the table indicates that the variable was included in that particular model.

^bThe Cp nearest p model was the same as the minimum Cp model for pH.

**Table 8. Variable transformations for MLR
and discriminant analysis procedures**

Untransformed variable name	Transformation
LAKE- A	$\log_{10} (\text{LAKE_A})$
WTRSHD- R	$\log_{10} (\text{WTRSHD- R})$
DRAIN- A	$\log_{10} (\text{DRAIN- A})$
NONFR2_ P	$\log_{10} (\text{NONFR2_P} + 1)$
WFLND- PW	$\log_{10} (\text{WFLND- PW} + 5)$
VACID- PW	$\log_{10} (\text{VACID- PW} + 5)$
NACID- PW	$\log_{10} (\text{NACID- PW} + 1)$
WFLND- PL	$\log_{10} (\text{WFLND- PL} + 100)$
VACID- PL	$\log_{10} (\text{VACID- PL} + 100)$
N A C I D- P L	$\log_{10} (\text{NACID- PL} + 50)$
BVRINDEX	$\log_{10} (\text{BVRINDEX} + 1)$

constants were chosen to provide a ratio of maximum to minimum of about ten for the untransformed variable. The results of these transformations were to decrease the significance of all D statistics to $p > 0.15$, indicating the residuals are more nearly normal.

For all analyses, ANC was log transformed after adding 100 to the observed ANC value. The log transform of ANC, a dependent variable, was clearly justified because the significance of all D values decreased from $p < 0.01$ to $p > 0.15$ with this transform. This decrease was true whether or not any predictor variables were transformed. The pH measure is a log transformation of H^+ concentration. For the dependent variable pH, the transformations of predictor variables resulted in a decrease in significance of D from $p < 0.02$ to $p > 0.15$ for all types of residuals. The transformed variables (Table 8) were chosen for the regressions involving both pH and ANC to simplify model interpretation, simplify the collinearity analysis, and foster comparability between the MLR analysis and the subsequent discriminant analysis.

3.2.1.2 Collinearity Diagnostics and Model Development

For each of the two regressions (pH and ANC), the REG procedure in SAS (1985) was used to identify and eliminate excessively collinear candidate predictor variables. When a predictor variable is nearly a linear combination of other predictor variables in a model (i.e., is collinear), the affected parameter estimates are unstable and may have large standard errors (Draper and Smith 1981). Although significant correlation between two variables implies significant collinearity, nonsignificant correlation between two variables does not necessarily

indicate the absence of collinearity, because there may be some linear combination of two or more variables that, taken as a whole, would create a condition of collinearity. In selecting the candidate variables, we excluded combinations of variables that would obviously be collinear (e.g., percentages that sum to 100%). However, it was considered desirable to keep some candidate variables (i.e., the various wet deposition rates, concentration, and precipitation variables) that were believed to be collinear and allow an established protocol to select variables for removal.

A protocol was designed to identify and remove collinear variables from the set of candidate variables. This procedure is described in detail in Appendix A. Briefly, the collinearity option in the REG (SAS 1985) procedure was applied successively. At each step (until the maximum condition index was ≤ 30) with each model, one of the identified collinear variables was eliminated; this was not necessarily the same variable for both the pH and ANC regressions (Table 9). A sequence of priorities was developed to determine which variable to eliminate. The intercept was always retained. The C_p statistic (Mallows 1973, Draper and Smith 1981) was calculated for each of the reduced models in which one of the potentially collinear variables was omitted. In the C_p statistic, p represents the total number of parameters, including the intercept, in the model. If the C_p statistic differed substantially, then the model with the lower C_p statistic determined the variable to be omitted. Otherwise, within a step successive preference was given to keeping a collinear variable that (see Appendix A for more detail).

Table 9. Candidate variables listed in order of their elimination because of collinearity with other variables^a

Model	Dependent variable	Predictor variables eliminated and order eliminated
Cp nearest p	pH	SHL2_B_P, LAKE_A, NO3_CONC, SHL2_R_P, DPTH_P_V, PPT, SO4_WET, ROCK_P, HYDRO_C, WTLND_PL, BSC_L_P, ORG_MAT, NO3_WET, DPTH_B_V, PH_VL_P, SHL2_P_P, BSA_L_P, SO4_CONC, DPTH_R_V, NACID_PL, VACID_PW, ACID_P, WTRSHD_R, VACID_PL, H_WET
	ANC	SHL2_B_P, LAKE_A, H_WET, PPT, ROCK_P, SO4_CONC, DPTH_P_V, SHL2_R_P, ORG_MAT, VACID_PL, DPTH_R_V, BSC_L_P, SO4_WET, ACID_P, PH_VL_P, H_CONC, WTRSHD_R
Selected models, minimum Cp	pH	No additional variables removed
	ANC	Variables listed above, plus NACID_PL, ACID_EX, BSA_L_P, SHL2_P_P, VACID_PW, OM_H_P, WTLND_PL, NO3_WET, PHC_VL_P, HYDTYP3, HYDRO_C
Minimum p and Cp near p	pH	Variables listed above, plus ACID_EX, RUNOFF, OM_H_P, HYDTYP2, ELEV, WTLND_PW, CECS_L_P, ROCK12_P, LAKE_DEV, DRAIN_A, PH_C_VL_P, VACID_PP, HRDWD2_P, RELIEF_R, CEC, STEEPM_P, STONEY_P, NACID_PP, ASPECT_S, CONFR2_P
	ANC	Variables listed above, plus ELEV, WTLND_PW, DRAIN_A, CECS_L_P, ROCK12_P, LAKE_DEV, DPTH_B_V, VACID_PP, HRDWD2_P, CEC, RELIEF_R, STEEPM_P, STONEY_P, HYDTYP2, NACID_PP, WTLND_PP, ASPECT_S
Noncollinear	pH	Variables listed above, plus WTLND_PP, NACID_PW
	ANC	Variables listed above, plus CONFR2_P

^aSee Table 7 for a list of the variables in each model.

1. was originally the only variable in its "group" or was the last remaining variable in its group, all others having been eliminated in previous steps;
2. was more mechanistic, i.e., potentially causal, than other collinear variables;
3. was considered a more reliable measure than other collinear variables;
4. was easier to obtain in other data sets to which the model might be applied;
5. was the only collinear variable in its group, with at least two other collinear variables belonging to a single different group;
6. was significantly correlated with the dependent variable;
7. avoided misinterpretation of the importance of sulfate atmospheric inputs over nitrate inputs or vice-versa;
8. maintained diversity of wetland variables within its respective group; or
9. had higher variance-decomposition proportions than other collinear variables.

Application of this procedure should have helped ensure that the remaining variables are not substantially collinear. After elimination of potentially collinear variables, the following models were available for subsequent comparisons: the minimum C_p model, the model with C_p nearest p , the model with minimum p and C_p near p , and the noncollinear model.

In addition to reduction of collinearity, there are a number of other criteria for selecting a potentially useful model. For comparison of model performance, a stepwise procedure based on the maximum R^2 improvement technique (MAXR²) (SAS 1985) was applied to the 57 candidate variables without concern for collinearity among these variables and to the variables remaining in the selected minimum C_p models resulting from the collinearity procedure. The resulting models are called the stepwise MAXR² and the reduced stepwise MAXR² models, respectively. The maximum R^2 improvement technique finds the one-variable model producing the highest R^2 . Then another variable, which yields the greatest increase in R^2 , is added. Once the two-variable model is obtained, each of the variables in the model is compared with each variable not in the model. After comparing all possible combinations, the one that produces the largest increase in R^2 is used, and comparisons begin again. The comparing and switching process is repeated in a stepwise manner to find the "best" model for the maximum number of predictor variables that significantly improve the R^2 value. The stepwise MAXR² models developed for pH and ANC from all 57 candidate variables contained the same number of predictor variables as the minimum C_p models for pH and ANC, but the predictor variables were selected using the maximum R^2 improvement technique. The reduced stepwise MAXR² models were developed by applying this technique to the more limited set of predictor variables present in the minimum C_p models for pH and ANC and then applying an F-test to determine the significance of each variable, based on the reduction in unexplained variance as a result of the addition of each variable. A

level of significance, $p \leq 0.25$, was chosen for evaluating the minimum number of variables to include in the model. This value is conservative because it ensures that potentially important variables will be included.

3.2.1.3 Model Verification

Each model was applied to an independent verification data set to see how it would perform for predictive purposes. The original data set was randomly divided into two sets in which the ratio of lakes in the two sets would be 2:1. The larger set, designated as set A or the calibration data set, was used for developing the model, and the second set, designated as set B, was used for the verification study.

Before the different verification measures are described, some notation should be set up. Let $N_A(N_B)$ be the number of observations in data set A(B). For data set B, let $Y_{Bi}, i=1,2,\dots,N_B$, be the observed values of the dependent variable (pH or ANC). Let $\hat{Y}_{Bi}, i=1,2,\dots,N_B$, be the corresponding predicted values for these observations that were obtained from the model derived from data set A. One measure of the predictive capability of the model is the mean square error (MSE):

$$MSE = \frac{1}{N_B} \sum_{i=1}^{N_B} (Y_{Bi} - \hat{Y}_{Bi})^2 ,$$

which is a general measure of how large the differences are between the predictions and observations (i.e., smaller values of this quantity would indicate better predictions than larger values). Because MSE is

a general measure of model predictive ability, note the "systematic error" or bias and "random error" or precision component of this measure. Let

$$d_{Bi} = y_{Bi} - \hat{y}_{Bi} ,$$

for $i=1,2,\dots,N_B$. Then the model bias may be estimated by

$$\text{BIAS} = \frac{1}{N_B} \sum_{i=1}^{N_B} (y_{Bi} - \hat{y}_{Bi}) = \frac{1}{N_B} \sum_{i=1}^{N_B} d_{Bi} = \bar{d}_B ,$$

and the "random error" component may be estimated by

$$\text{VAR} = \frac{1}{N_B} \sum (d_{Bi} - \bar{d}_B)^2 .$$

It can be shown mathematically that

$$\text{MSE} = (\text{BIAS})^2 + \text{VAR} .$$

Therefore, what portion of the differences between the model predictions and observations that may be attributed to bias or random error may be determined.

Also, the value of MSE may be compared with the error mean square for the model when the N_A observations in data set A are used to estimate the necessary parameters. Another interesting comparison would be the square root of MSE with the anticipated measurement error of the dependent variable if such a quantity is available.

Another measure of the predictive capability of the derived model is related to the \tilde{R}^2 , the square of the multiple correlation coefficient. Let

$$\tilde{R}^2 = 1 - \left[\frac{\sum_{i=1}^{N_B} (Y_{Bi} - \hat{Y}_{Bi})^2}{\sum_{i=1}^{N_B} (Y_{Bi} - \bar{Y})^2} \right]$$

where \bar{Y} could be the mean of the N_B observations in data set B--in which case \tilde{R}^2 could be thought of as the amount of variation explained by using the model relative to the sample mean (\tilde{R}_B^2)--or \bar{Y} could be the mean of the N_A observations in data set A--in which case \tilde{R}^2 could be thought of as the reduction in error achieved by using the derived model vs a model based only on the mean from data set A (\tilde{R}_A^2). Both measures of \tilde{R}^2 relate to predictions for the verification data set and can be used to assist in evaluating how well the model does in predicting for an independent data set. Note that in either case the maximum value \tilde{R}^2 can achieve is 1 but \tilde{R}^2 can be less than zero. Negative values of \tilde{R}^2 would indicate very poor predictions.

3.2.2 Results

The purpose of the MLR analysis was to develop models that have relatively few parameters so that they are informative about which watershed attributes might influence lake acidification. The verification procedure was developed to evaluate the predictive ability of the resulting models for an independent data set. The results from the MLR analysis are discussed in two parts. The variables used in the different models and the indices of performance used to evaluate the models are discussed in Sect. 3.2.2.1. For the selected pH and ANC models, the results from predictions are discussed in Sect. 3.2.2.2.

3.2.2.1 ~~Model Development and Verification~~

Seven models each were developed for pH and ANC (Fig. 8). Four of these resulted from the collinearity diagnostics, and two were developed using the maximum R^2 improvement technique. The seventh model, the "full" model contained the same 57 predictor variables for pH and ANC. [For simplicity the reader may wish to focus on the "selected" (minimum C_p models) and the reduced stepwise $MAXR^2$ models. Selected models are highlighted in bold type in tables.] The 57 candidate predictor variables were selected because they were thought to be potentially important in the lake acidification process and/or they had significant and reasonably large correlation coefficients compared with other possible candidate variables. The "minimum C_p ", " C_p nearest p," "minimum p and C_p near p," and "noncollinear" models resulted from the collinearity diagnostic procedure in which collinear variables were removed in a stepwise manner. The order in which candidate variables were eliminated can be seen in Table 9. The noncollinear model contained ten variables for both pH and ANC. The minimum p and C_p near p model contained 12 variables for pH and 11 variables for ANC. The minimum C_p and C_p nearest p models for pH were the same with 31 variables; the ANC C_p nearest p model contained 40 variables. The general progression of variables eliminated for the pH and ANC models was similar.

The maximum R^2 improvement technique provided stepwise models ($MAXR^2$) that did not take collinearity into consideration yet performed reasonably well with any number of variables up to the full 57. The "reduced" stepwise $MAXR^2$ models did take collinearity

into consideration because only the variables in the selected minimum C_p models (27 for ANC and 31 for pH) were used as input to the reduced $MAXR^2$ models. The reduced $MAXR^2$ models were developed to determine if a smaller model, developed to maximize R^2 rather than further reduce collinearity, was reasonably robust compared with other models. The predictor variables used in each of the MLR models for pH and ANC are indicated in Table 7.

The MSE and \tilde{R}^2 measures defined in Sect. 3.2.1.3 were the indices used to evaluate performance of the various models (Tables 10 and 11). Based on the verification measures, the minimum C_p models for pH and ANC (developed using the collinearity diagnostics) were chosen as the "selected" best models, which have the lowest MSE and low-bias components (6 to 9%). Although the $MAXR^2$ model for pH and the C_p nearest p and reduced $MAXR^2$ models for ANC have smaller bias components than the selected minimum C_p models, they also have higher MSEs. A plot of model residuals (observed value - predicted value) against predicted values also helps evaluate model bias (Fig. 9). For an ideal model, residuals would be randomly distributed about the zero residual line. The residual plots for the selected pH and ANC models have no distinct patterns, thus indicating a lack of bias in these models (Fig. 9).

Another way to evaluate a model's predictive capability is to compare the R^2 of the calibration subset with \tilde{R}^2 , which can be thought of as the reduction in error achieved by using the derived model vs a model based only on the mean from the calibration (\tilde{R}_C^2) or verification (\tilde{R}_V^2) subsets. The minimum C_p models have the largest \tilde{R}^2 values, indicating that the models are a substantial

Table 10. Indices for pH model verification

Model	Verification data					Calibration data	
	$\sqrt{\text{MSE}}^a$	Bias	Random error	\tilde{R}_C^2	\tilde{R}_V^2	R^2	Error mean square
Noncollinear 10 variables	0.75 (0.57) ^b	-0.31	0.47 (83%) ^c	0.38	0.39	0.37	0.55
Minimum p/Cp near p 12 variables	0.77 (0.60)	-0.32	0.50 (83%)	0.35	0.36	0.38	0.54
Selected model minimum Cp and Cp nearest p 31 variables	0.69 (0.48)	-0.17	0.45 (94%)	0.48	0.49	0.62	0.40
Full 57 variables	0.92 (0.86)	-0.25	0.79 (93%)	0.06	0.09	0.72	0.39
Stepwise MAXR ² 31 variables	0.81 (0.65)	-0.16	0.63 (96%)	0.28	0.30	0.69	0.33
Reduced stepwise MAXR ² 14 variables	0.74 (0.54)	-0.26	0.48 (88%)	0.41	0.42	0.57	0.38

^aMSE is the mean square error; the $\sqrt{\text{MSE}}$ is in the same units as the measured values.

^bMSE in parentheses.

^cPercentage of MSE that is random error.

Table 11. Indices for ANC model verification

Model	Verification data					Calibration data	
	$\sqrt{\text{MSE}}^a$	Bias	Random error	\tilde{R}_C^2	\tilde{R}_V^2	R^2	Error mean square
Noncollinear 10 variables	0.18 (0.034) ^b	-0.065	0.030 (88%) ^c	0.12	0.15	0.39	0.026
Minimum p/Cp near p 11 variables	0.18 (0.031)	-0.069	0.026 (85%)	0.19	0.22	0.45	0.023
Selected model minimum Cp 27 variables	0.17 (0.030)	-0.053	0.027 (91%)	0.21	0.25	0.58	0.021
Cp nearest p 40 variables	0.19 (0.034)	-0.048	0.032 (93%)	0.10	0.14	0.66	0.019
Full 57 variables	0.24 (0.055)	-0.090	0.047 (85%)	-0.44	-0.38	0.74	0.018
Stepwise MAXR ² 27 variables	0.20 (0.040)	-0.063	0.037 (90%)	-0.06	-0.01	0.66	0.017
Reduced Stepwise MAXR ² 11 variables	0.18 (0.033)	-0.052	0.030 (92%)	0.14	0.18	0.52	0.021

^aMSE is the mean square error; the $\sqrt{\text{MSE}}$ is in the same units as the measured values.

^bMSE in parentheses.

^cPercentage of MSE that is random error.

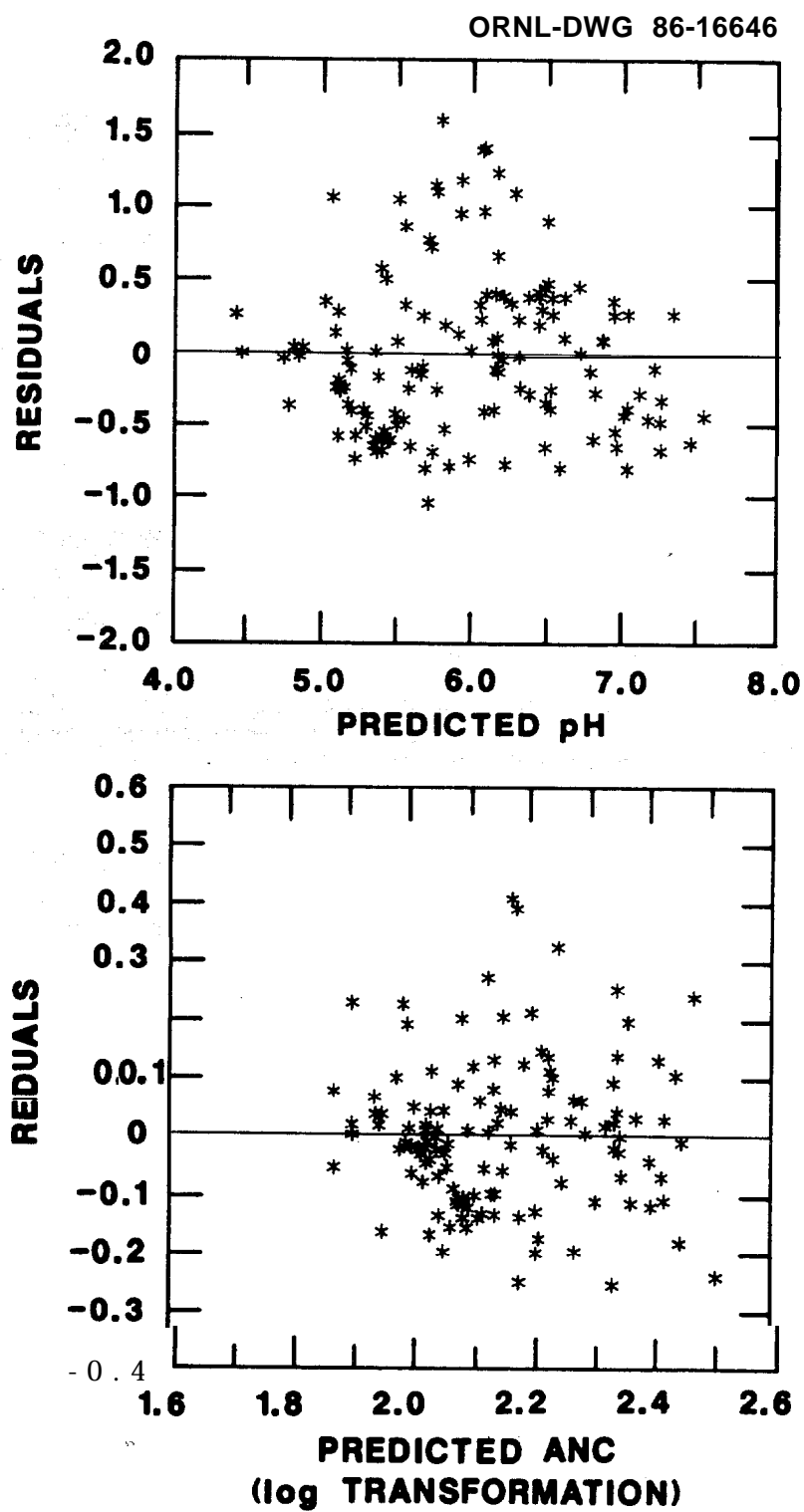


Fig. 9. Plots of residuals from MLR models for pH and ANC.

improvement over simply using a mean. The full and MAXR^2 models for ANC have negative \tilde{R}^2 values, which indicate very poor model performance. Overall, the minimum C_p models are the most robust of our models based on the MSE and \tilde{R}^2 indices, and they also explain a reasonably large amount of variation in the dependent variables ($R^2 = 0.62$ for pH and $R^2 = 0.58$ for ANC). The reduced MAXR^2 models, which use about one-half the number of variables as the minimum C_p models, account for 5 to 6% less variation in the dependent variables but still perform quite well, based on the indices (Tables 10 and 11). Therefore, the estimated coefficients and standard error of estimates for only the selected best models (minimum C_p) and the reduced MAXR^2 models are listed in Tables 12 and 13).

3.2.2.2 Model Predictions

A regression model predicts estimates of the mean value of the response variable, given a particular vector of predictor variables. Projections near the extremes of the distribution, therefore, would be expected to be underrepresented, and they were. This is graphically shown in Fig. 10 for the pH and ANC calibration data sets. For the most acidic classes, the models correctly predicted only 7 of 37 lakes into the <5.0 pH class and 16 of 43 lakes into the <0 $\mu\text{eq/L}$ class. A similar underprediction of the frequency of lakes with relatively high ANC values also occurred. The distributions for the observed and predicted values for the combined data in the calibration and verification subsets are shown in Figs. 11 and 12.

To minimize this underrepresentation of extreme values, a technique to include information on the variance of predicted means

Table 12. Estimated coefficients and their standard errors for selected MLR model and reduced stepwise model for pH^a

Variable name	Selected minimum Cpb (R ² = 0.62)		Reduced stepwise ^c (R ² = 0.57)	
	Estimated coefficient	Standard error of estimate	Estimated coefficient	Standard error of estimate
<u>Morphologic and Physiographic</u>				
DRAIN- A	-0.025	0.18		
LAKE- DEV	-0.20	0.20		
LAKE- E	-0.0016	0.00078	-0.0014	0.00066
ASPECTS	-0.0023	0.0023		
<u>Hydrologic</u>				
RUNOFF	0.020	0.0094	0.019	0.0078
HYDTYP1	-0.55	0.68		
HYDTYP2	-0.30	0.52		
HYDTYP3	-0.17	0.50		
<u>Atmospheric</u>				
H_CONC	-71.09	14.29	-70.86	11.15
<u>Physical Soil Type</u>				
RELIEF- R	0.016	0.015	0.018	0.012
STONEYP	0.0062	0.0037	0.0069	0.0029
HYDRO_A	-0.030	0.010	-0.028	0.0064
HYDRO_B	-0.0092	0.013		
STEEPM_P	-0.0071	0.0034	-0.0046	0.0025
STEEPV_P	-0.011	0.0055	-0.011	0.0048
<u>Geology</u>				
ROCK12_P	-0.0060	0.0039		
<u>Chemical Soil Type</u>				
CECS_L_P	-0.0029	0.010		
OM_H_P	-0.027	0.0079	-0.023	0.0062
PHC_VL_P	0.0049	0.0046		
CEC	0.036	0.017	0.033	0.012

Table 12. (continued)

Variable name	Selected minimum Cp ^b (R ² = 0.62)		Reduced stepwise ^c (R ² = 0.57)	
	Estimated coefficient	Standard error of estimate	Estimated coefficient	Standard error of estimate
<u>Forest Cover</u>				
CONFR2_P	-0.013	0.0055	-0.017	0.0038
HRDWD2_P	0.0037	0.0044		
NONFR2_P	0.39	0.77		
<u>Wetland Type</u>				
WTLND_PP	-0.0051	0.0050		
VACID_PP	-0.0035	0.0050	-0.0060	0.0024
NACID_PP	0.0058	0.0067		
WTLND_PW	0.96	0.39	0.61	0.30
NACID_PW	-0.42	0.31		
<u>Disturbance</u>				
DISTRB_P	-0.0011	0.0017		
BVRINDEX	-0.55	0.36	-0.54	0.30
CABN78_R	-0.018	0.018		

^aValues reported to two significant digits.

^bSelected or best overall model.

^cReduced version of selected model.

Table 13. Estimated coefficients and their standard errors for the selected MLR model and reduced stepwise model for ANC^a

Variable name	Selected minimum Cp ^b (R ² = 0.58)		Reduced stepwise ^c (R ² = 0.52)	
	Estimated coefficient	Standard error of estimate	Estimated coefficient	Standard error of estimate
<u>Morphologic and Physiographic</u>				
DRAIN-A	-0.026	0.042		
LAKE-DEV	-0.082	0.046	-0.093	0.040
ASPECT-S	0.00024	0.00050		
<u>Hydrologic</u>				
HYDTYP1	-0.14	0.094		
HYDTYP2	-0.048	0.043		
<u>Atmospheric</u>				
NO3_CONC	-0.40	0.092	-0.42	0.071
<u>Physical Soil Type</u>				
RELIEF-R	0.0029	0.0032		
STONEY_P	0.00097	0.00083		
HYDRO_A	-0.0012	0.0023		
HYDRO_B	0.0022	0.0028	0.0026	0.0012
STEEPM_P	-0.0015	0.00074	-0.0015	0.00060
STEEPV_P	-0.0020	0.0011		
DPTH_B_U	-0.0019	0.0018	-0.00097	0.00040
<u>Geology</u>				
ROCK12_P	-0.00090	0.00091		
<u>Chemical Soil Type</u>				
CECS_L_P	-0.0015	0.0024		
CEC	0.0018	0.0034		
<u>Forest Cover</u>				
CONFR2_P	-0.0017	0.0012	-0.0022	0.0011
HRWD2_P	0.0015	0.00099	0.0016	0.00087
NONFR2_P	-0.25	0.15	-0.23	0.14

Table 13. (continued)

Variable name	Selected minimum Cp ^b (R ² = 0.58)		Reduced stepwise ^c (R ² = 0.52)	
	Estimated coefficient	Standard error of estimate	Estimated coefficient	Standard error of estimate
<u>Wetland Type</u>				
WTLND_PP	-0.00059	0.0011	-0.0014	0.00057
VACID_PP	-0.0016	0.0011		
NACID_PP	0.0010	0.0017		
WTLND_PW	0.28	0.093	0.24	0.07
NACID_PW	-0.12	0.069		
<u>Disturbance</u>				
DISTRB_P	-0.00033	0.00041		
BVRINDEX	-0.25	0.085	-0.17	0.073
CABN78_R	-0.0056	0.0040		

^aValues reported to two significant digits.

^bSelected or best overall model.

^cReduced version of selected model.

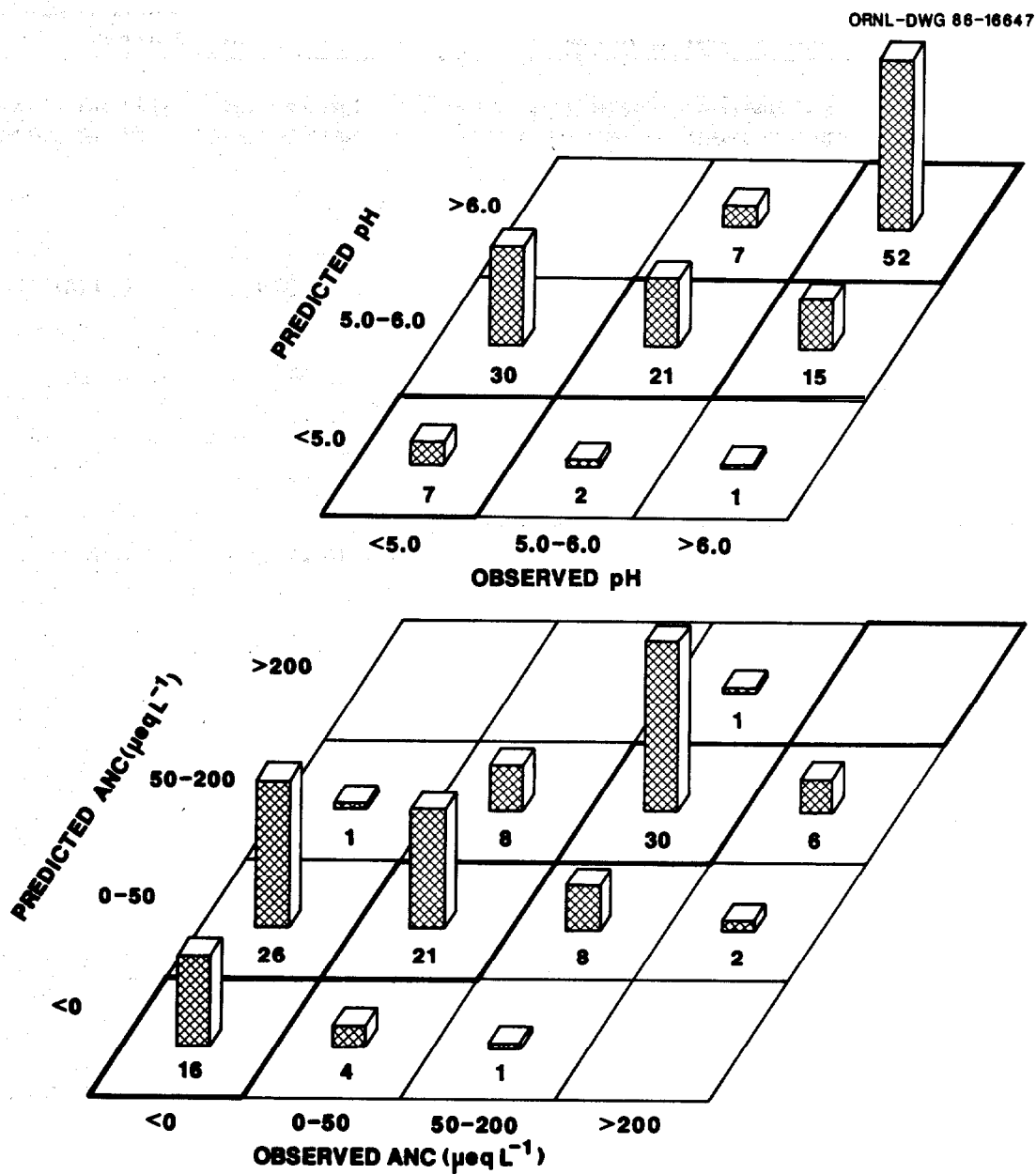


Fig. 10. Frequency of predicted and observed mean pH and ANC for selected MLR models.

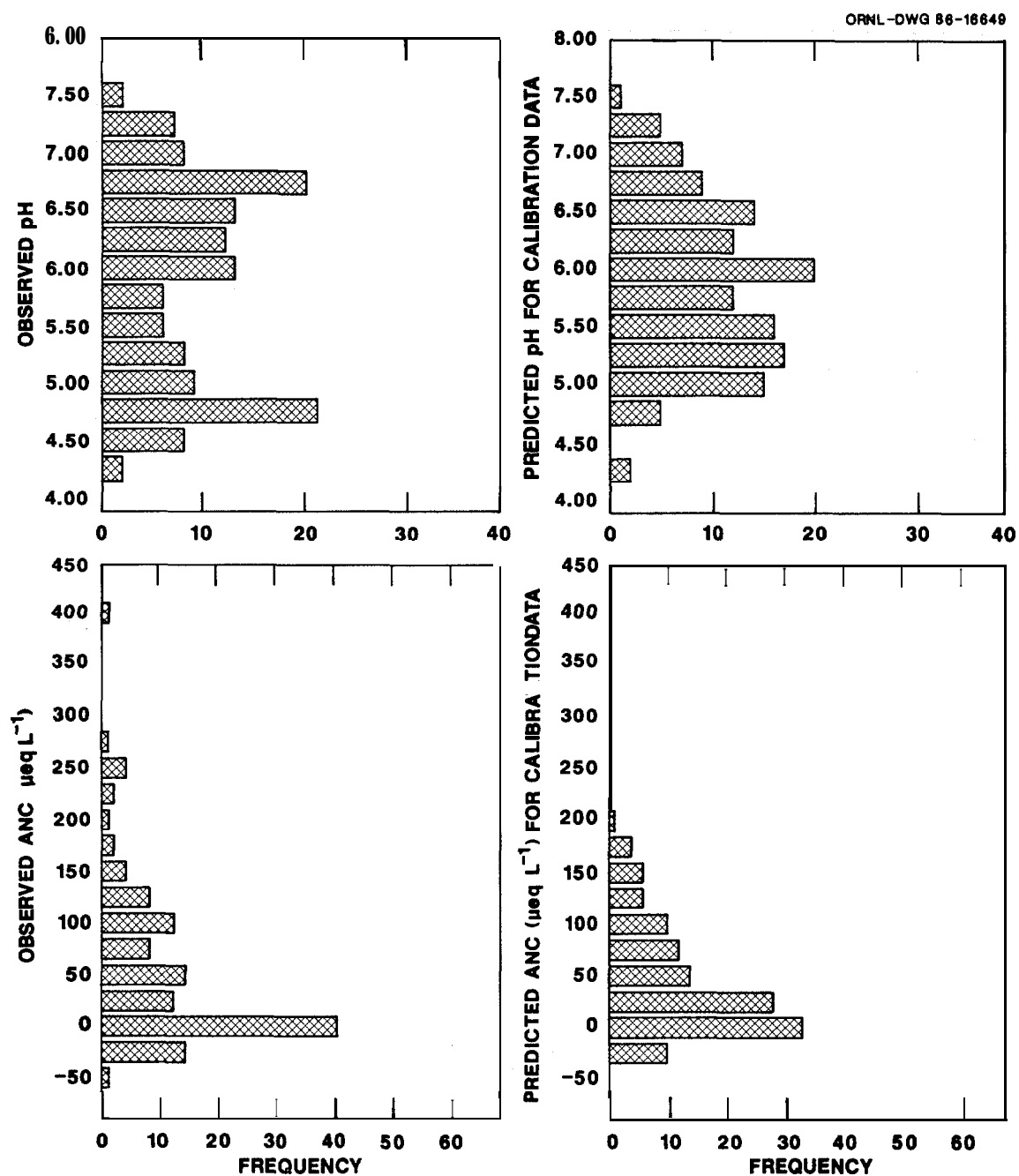


Fig. 11. Frequency distributions of observed and predicted pH and ANC values for lakes in calibration subsets.

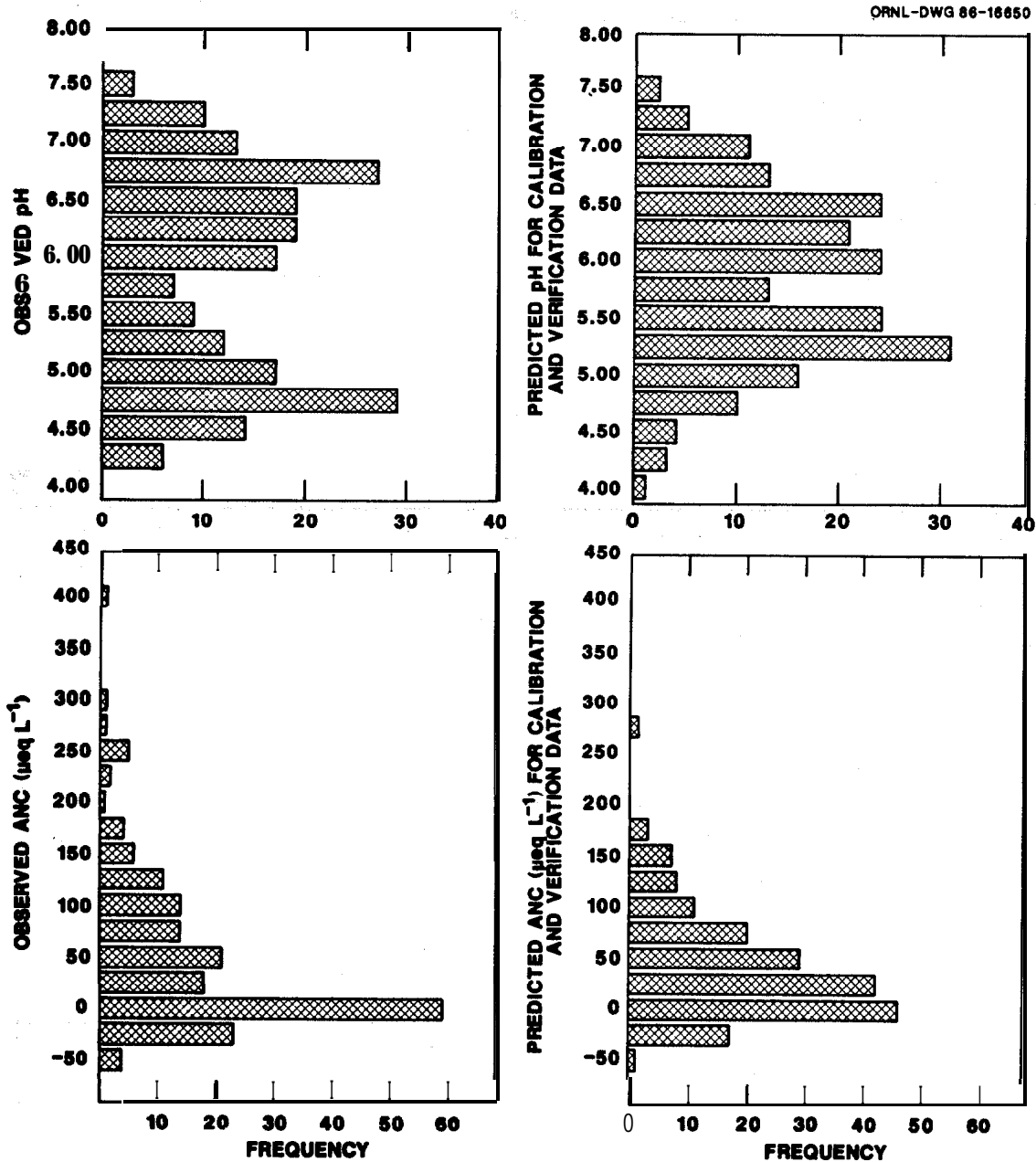


Fig. 12. Frequency distributions of observed and predicted pH and ANC values for lakes in combined calibration and verification subsets.

from the MLR was developed. Thus, in our final estimates we calculated the probability of pH or ANC being less than the specified value for individual predicted observations, according to the following procedure.

Assume there exists a data set of n_s observations (or lakes) that have been used to relate lake pH or ANC to a collection of p predictor variables x_1, \dots, x_p . This data set is called the calibration or estimation data set. The assumed model is

$$y_i = B_0 + B_1 x_1 + \dots + B_p x_{pi} + \epsilon_i ,$$

where $i = 1, 2, \dots, n_s$ with the usual linear regression model assumptions. The n_s observations have been used to estimate the unknown B 's to produce the estimation or prediction equation

$$\hat{y}_i = b_0 + b_1 x_i + \dots + b_p x_{pi}$$

for the n_s observations in the calibration data or any other choice of values for the x 's.

The above equation should help predict the number of lakes with pH or ANC values (y) less than some specification limit L . Because our model assumes a distribution of pH or ANC values for any given combination of the predictor variables, it is proposed to incorporate this distribution into the estimation process. Let

$$\begin{aligned} I(y) &= 1 \text{ if } y \leq L , \\ &= 0 \text{ if } y > L , \end{aligned}$$

and n_p equal the number of observations for which predictions are desired. The quantity

$$\psi = \sum_{j=1}^{np} P \ y \leq L/x_j = \sum_{j=1}^{np} \left(\Phi \frac{L - \beta_0 - \beta_1 x_{1j} - \dots - \beta_p x_{pj}}{\sigma} \right),$$

can be considered as the expected value of the number of lakes with values $\leq L$. However, only estimates of these quantities, not the actual B 's or σ , are available. Therefore, the estimate of ψ is given by

$$\hat{\psi} = \sum_{j=1}^{np} \Phi \frac{L - \hat{b}_0 - \hat{b}_1 x_{1j} - \dots - \hat{b}_p x_{pj}}{s},$$

where s^2 is the error mean square from the fit of this model to the calibration data set and $\Phi(\cdot)$ is standard normal cumulative distribution function. The quantity $\hat{\psi}$ is referred to as the "adjusted predictions." Because the adjusted predictions were almost always closer to the observed values, all final estimates for the number of lakes in a pH or ANC category used this adjustment (Tables 14 and 15).

Estimates of the current number of headwater lakes in specified pH and ANC categories are a combination of actual (observed) values and predicted values, using the selected MLR models (Tables 16 and 17). For actual measurements NSWs data were used when available (46 lakes); otherwise, FIN data were used. If no measurement was available, a value was predicted using the selected MLR models. Even with the adjusted predictions, estimates for the acidic lake categories are

Table 14. Comparison of percentages of lakes in pH categories for observed, MLR-prediction, and adjusted MLR-prediction values (using selected minimum Cp model)

pH category and data subset	Number of lakes (%)		
	Observed	Adjusted estimate	MLR prediction
(5.0			
FIN calibration	27.4	18.1	7.4
FIN verification	31.3	17.6	10.4
NSWS verification	26.1	17.2	4.3
≤5.5			
FIN calibration	39.2	35.5	35.6
FIN verification	49.2	33.6	29.8
NSWS verification	43.5	36.3	34.8
≤6.0			
FIN calibration	49.6	55.6	56.3
FIN verification	55.2	53.7	53.7
NSWS verification	54.3	58.0	65.2

Table 15. Comparison of percentages of lakes in ANC categories for observed, MLR-prediction, and adjusted MLR-prediction values (using selected minimum Cp model)

ANC category and data subset	Number of lakes (%)		
	Observed	Adjusted estimate	MLR prediction
$\leq 0 \mu\text{eq L}^{-1}$			
FIN calibration	34.7	26.4	16.9
FIN verification	39.3	23.9	14.8
NSWS verification	28.3	24.8	10.9
$\leq 50 \mu\text{eq L}^{-1}$			
FIN calibration	58.9	53.1	57.3
FIN verification	67.2	50.2	50.8
NSWS verification	69.6	53.9	60.9
$\leq 200 \mu\text{eq L}^{-1}$			
FIN calibration	93.5	94.3	99.2
FIN verification	96.7	93.6	95.1
NSWS verification	93.5	95.4	97.8

Table 16. Estimated number of Adirondack headwater lakes in pH categories^a

Analysis technique and basis	Lakes in pH category ^b (%)			Total number of lakes
	≤5	≤5.5	≤6	
Measured	28.2 (n = 70)	42.7 (n = 106)	52.0 (n = 129)	248
Predicted				
MR model	16.1 (f = 34.5)	31.7 (f = 67.8)	50.8 (f = 108.7)	214
Discriminant model	ne ^c	34.6 (n = 74)	ne ^c	214
Combined (measured and predicted)				
MR model	22.6 (f = 104.5)	37.6 (f = 173.8)	51.4 (f = 237.7)	462
Discriminant model	ne ^c	39.0 (n = 180)	ne ^c	462

^aBased on available measurements (NSWS data if available, otherwise FIN data) or, if no measurement is available, on prediction using MLR or discriminant analysis. Selected models used.

^bf = expected cumulative frequency.

^cNot estimated.

Table 17. Estimated number of Adirondack headwater lakes in ANC categories^a

Data source	Lakes in ANC category ^b			Total number of lakes
	≤0	≤50	≤200	
Measured	34.6 (n = 80)	63.2 (n = 146)	94.4 (n = 218)	231
Predicted				
MLR model	22.3 (f = 51.5)	48.1 (f = 111.2)	92.6 (f = 213.9)	231
Discriminant model	ne ^c	52.0 (n = 120)	ne ^c	231
Combined (measured and predicted)				
MLR model	28.5 (f = 131.5)	55.7 (f = 257.2)	93.5 (f = 438.9)	462
Discriminant model	ne ^c	57.6 (n = 266)	ne ^c	462

^aBased on available measurements (NSWS data if available, otherwise FIN data) or, if no measurement is available, on prediction using MLR discriminant analysis. Selected models used.

^bf = expected cumulative frequency.

^cNot estimated.

probably low ($\text{pH} \leq 5$; $\text{ANC} \leq 0 \mu\text{eq L}^{-1}$), as shown in Fig. 10 and Tables 14 and 15. The spatial patterns of observed (1974-1983 summer means) pH and ANC values for headwater lakes in the Adirondack Region are shown in Figs. 13 and 14 and can be compared with predicted mean values shown on overlays inside the back cover.

The estimates for the unknown lakes ("predicted" in Tables 16 and 17) indicate that 16 to 22% of these lakes are in the low pH and ANC categories. These estimates may have greater uncertainty than might be suggested by this relatively narrow range because both the models and the measurements used in the predictions are subject to uncertainty that has not been explicitly treated. However, a smaller proportion of the "unknown" lakes are in the low pH and ANC categories than is the case for lakes for which measurements exist. This may reflect, in part, a tendency for chemistry data to be available more frequently for lakes where acidification was a concern.

A number of factors act to create uncertainty both in the fits of MLR models to data and particularly in predictions based on such models. Measurement uncertainty exists in both the predictor and the response variables. For some predictor variables, this uncertainty is substantial. The inability of the model to reproduce the bimodal distribution for pH might indicate a missing predictor variable. Improvement of the R^2 values might be possible through inclusion of additional important variables or selected interaction terms in the model. The use of selected regression diagnostics to examine influential data points might also improve the R^2 values. Quite possibly, a relatively small number of lakes with aberrant

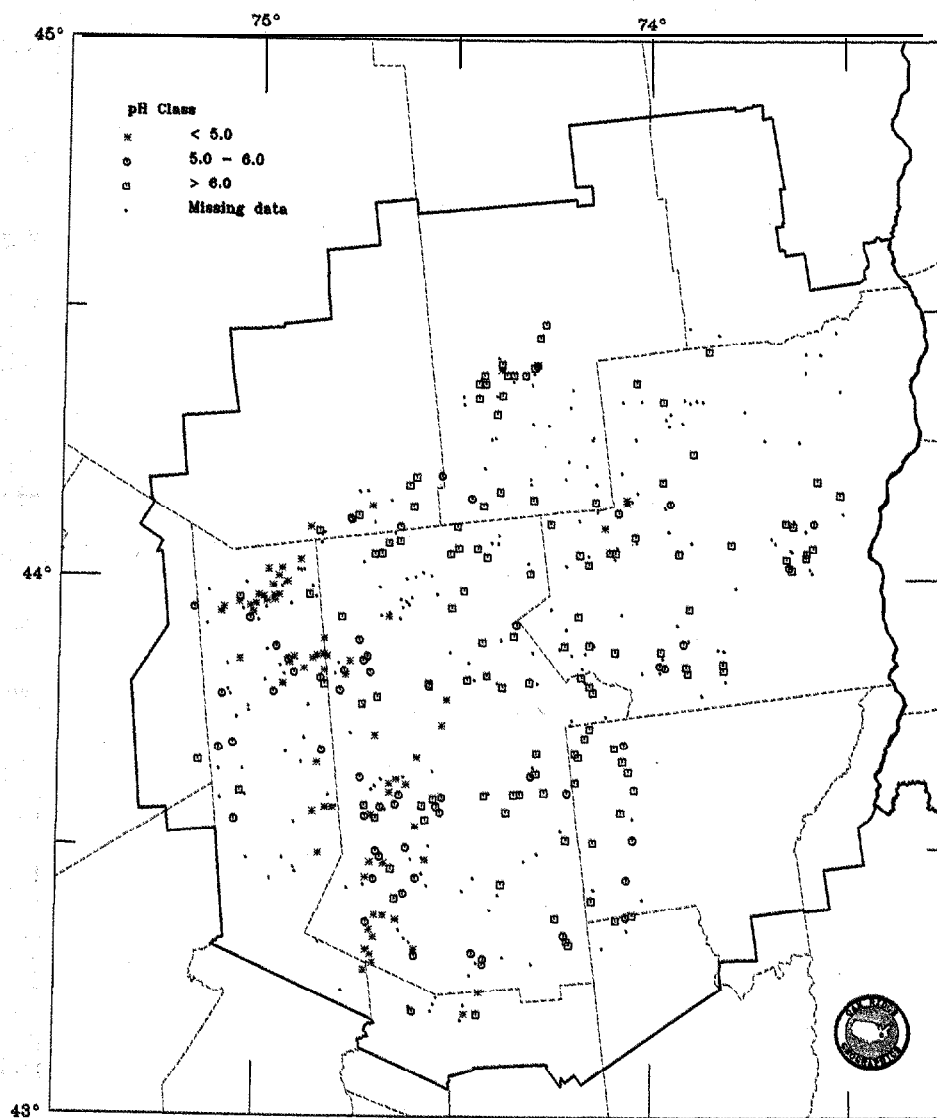


Fig. 13. Spatial pattern of observed summer mean pH for headwater lakes in Adirondack region (overlays inside back cover).

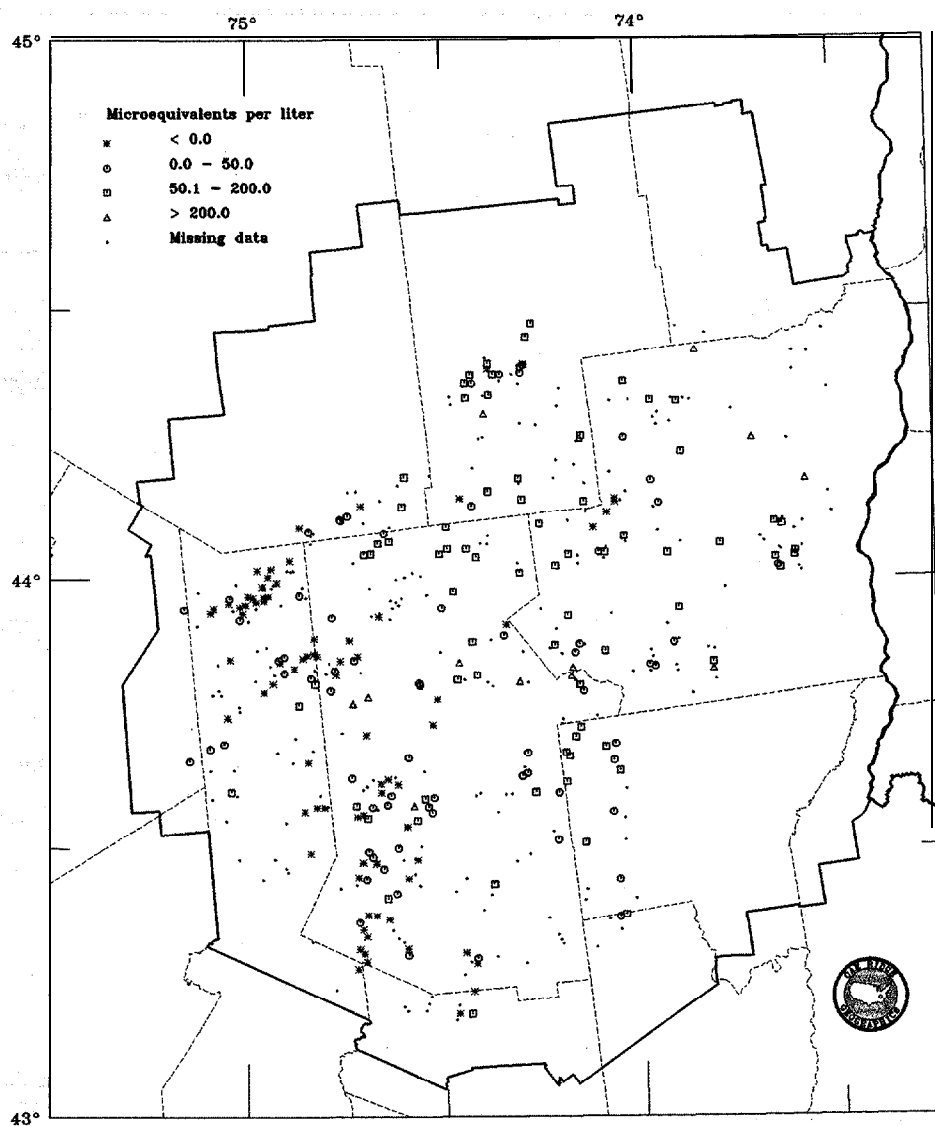


Fig. 14. Spatial pattern of observed summer mean ANC for headwater lakes in Adirondack region (overlays inside back cover).

characteristics or measurements are significantly influencing the estimated regression coefficients.

MLR can lead to inappropriate conclusions under some circumstances (McFadden 1963, Barnthouse 1982, Goodyear and Christensen 1984). In our applications, however, the control of collinearity (Table 10) and the consistency of selection of particular explanatory variables (Table 9) lend support to the general validity of our models. Still, the inability of the models to reproduce the pattern observed for the response variables indicates an inaccuracy with respect to estimating cumulative frequencies, thus, of course, increasing uncertainty.

MLR, a useful tool that provides better means of projecting pH or ANC values than many other nonmechanistic techniques (e.g., proportional estimates), should be viewed as complementing, but not necessarily replacing or substituting for, more-mechanistic models (Schnoor 1986, Wright 1983). In addition, considerable uncertainty may be associated with predictions from MLR, although such uncertainty also applies to other available methods.

3.3 DISCRIMINANT ANALYSIS

Discriminant analysis was used on the same data as the MLR to see if a multivariate analysis technique would select similar predictor variables and make similar estimates of the number of lakes in prespecified pH or ANC categories. In essence, discriminant analysis was used as a check on the MLR analysis.

3.3.1 Methods

Discriminant analysis procedures analyze data with one classification variable, in this case pH or ANC, and several explanatory variables (watershed attributes). A stepwise discriminant procedure, STEPDISC (SAS 1985), was used to find a subset of variables that best reveals differences between two defined classes for ANC and for pH in the calibration subset. Because the frequency distributions for pH appeared to be bimodal, a discriminant function was determined to be best developed on only two classes for pH and ANC (which will be referred to as "acidic" and "basic") (Figs. 11 and 12). The values for determining these classes are ≤ 5.5 (acidic) and > 5.5 (basic) for pH and $\leq 50.0 \mu\text{eq L}^{-1}$ (acidic) and $> 50.0 \mu\text{eq L}^{-1}$ (basic) for ANC. Reasons for choosing these values for class limits include frequency distributions of the class variables, consistency with NSW classes, and relevance to observed effects on fish (Baker 1984). Two sets of discriminant models were developed. A 27-variable model and a 31-variable model (for ANC and pH, respectively) were developed to be compared with the selected MLR models, which are referred to as the "selected" discriminant models and use the same variables as the MLR models. Then the 27 and 31 variables from the selected ANC and pH MLR models were used as candidate variables for the STEPDISC procedure, just as they were for the reduced MAXR^2 models. STEPDISC uses a stepwise selection procedure to identify the most important explanatory variables.

After STEPDISC was used to identify 9 and 13 significant explanatory watershed variables for pH and ANC, respectively ($p \leq 0.25$),

a discriminant procedure (DISCRIM, SAS 1985) was used to determine linear discriminant functions that classify lakes into either the acidic or basic class. DISCRIM develops a discriminant function by using a generalized squared distance measure, assuming that the distribution of the observed predictor variables within each class follows a multivariate normal distribution. The classification criterion was based on the individual within-class covariance matrices and prior probabilities for the classes. The "priors" represent the probability that a lake comes from one of the classes, and priors were proportional to the frequency distributions of the acidic and basic classes for the calibration subset. Each observation is placed in the class from which it has the smallest generalized squared distance.

3.3.2 Results

The results of the discriminant analysis procedure (DISCRIM, SAS 1985) for the FIN calibration subset and the FIN and NSW verification subsets are given in Tables 18 and 19. The 31-variable model for pH and 27-variable model for ANC correctly classify 93 and 90%, respectively, of headwater lakes into the acidic ($\text{pH} \leq 5.5$ or $\text{ANC} (50 \mu\text{eq L}^{-1})$) and basic classes. These models use the same variables as the selected minimum C_p models from the MLR procedure. In the verification exercise, these models correctly classify between 72 and 76% of the lakes. The equations for the discriminant models are not presented because discriminant analysis was only used as a check on MLR.

The reduced discriminant models perform almost as well as the larger models (Tables 18 and 19). The g-variable pH model correctly classifies 81% of the calibration data, and the 13-variable ANC model

Table 18. Results of discriminant analysis for pH

Model and subset	Lakes misclassified (%)		Total
	Acid to basic	Basic to acid	
Selected model			
31 variables			
FIN calibration (n = 135)	6	8	7
FIN verification (n = 67)	39	9	24
NSWS verification (n = 46)	33	15	28
Stepwise reduced model			
9 variables			
FIN calibration	21	17	19
FIN verification	24	9	16
NSWS verification	36	15	30

Table 19. Results of discriminant analysis for ANC

Mdel and subset	Lakes misclassified (%)		
	Acid to basic	Basic to Acid	Total
Selected model			
27 variables			
FIN Calibration (n = 124)	4	20	10
FIN Verification (n = 61)	27	25	26
NSWS Verification (n = 46)	2 8	3 3	28
Stepwise Reduced Model			
13 variables			
FIN Calibration	7	20	12
FIN Verification	36	25	33
NSWS Verification	26	33	26

correctly classifies 88%. As expected, verification subsets had somewhat higher percentages of misclassifications compared with the larger models, with 67 to 84% of the lakes correctly classified.

The reduced discriminant and MLR models have the following eight variables in common for pH: lake elevation, H^+ concentration, watershed relief, soils with high infiltration rates, moderately steep soils, percentage of watershed in conifers, percentage of watershed in hardwoods, and percentage of watershed in wetlands (Table 7). For ANC the reduced discriminant and MLR models have seven variables in common: NO_3^- concentration; percentage of watershed in conifers, in hardwoods, not in forest, and in wetlands; percentage of lake perimeter in wetlands; and beaver activity (Table 7). The first five variables selected in the stepwise reduced-discriminant-model exercise for pH are the same as those selected for the stepwise reduced MLR model. For ANC the first two variables are the same (NO3_CONC and CONFR2_P) for the stepwise discriminant and MLR models, but bedrock acid-neutralizing capacity and stony soils have more discriminating power for the discriminant model than the MLR model.

As with the MLR models, the discriminant model tends to underestimate the number of lakes in the low pH class; however, the ANC discriminant model misclassifies about the same proportion of lakes in the acidic and basic classes. The selected discriminant models predict 39% of Adirondack headwater lakes to have a pH ≤ 5.5 and 58% to have an ANC $\leq 50 \mu eq L^{-1}$. These results are similar to the adjusted predictions from the MLR models (Tables 16 and 17).

4. DISCUSSION AND CONCLUSIONS

The MLR model and discriminant analysis are techniques that allow evaluation of interactions among predictor variables and between predictor and dependent variables. The MLR model and the discriminant analysis were used to develop robust models for predicting lake pH and ANC. The regression model was the primary predictive tool, and the discriminant analysis served as corroborative evidence. In other words, did a different statistical technique give similar predictions about the number of lakes in a pH or ANC category? The two models do give similar predictions (Tables 16 and 17).

The candidate variables used in the multivariate analyses (Table 7) were selected to include as many potential controlling factors as possible for each watershed, either as direct measurements of the above factors or as surrogates for those factors. For example, H^+ , SO_4^{2-} , and NO_3^- concentrations in precipitation represent atmospheric inputs. Cation exchange capacity of the soils is a direct measurement taken from soil chemistry data. Slope, SCS soil hydrologic group, and watershed drainage area are potential surrogates for hydrologic contact with soil and bedrock or hydrologic behavior of the watersheds. Lake area may be a surrogate for in-lake processes. The expected signs for each variable's relationship with pH and ANC (Table 4) are based on the conceptual framework of how watershed attributes control surface water chemistry (Turner et al. 1986b) and are provided for the reader unfamiliar with the literature.

Correlation coefficients and the amount of variance explained in the multivariate analyses were not high, possibly because of the nature

of observational data and the complexity of watershed systems. Our highest correlation coefficient was 0.61, and the MLR models accounted for -60% of the variation in lake chemistry. These values are similar to, or better than, those obtained by researchers performing similar analyses (Rapp et al. 1985, Nair 1984). For the bivariate analyses, a few variables consistently had significant ($p \leq 0.01$) and high correlation coefficients for our data (r values >0.3); many other variables had correlation coefficients in the range of 0.2 to 0.3. The variables for wet deposition rates and precipitation concentrations had the highest correlation coefficients, ranging from -0.46 to -0.61. Coefficients for precipitation and runoff were slightly lower in the -0.4 to -0.5 range. Lake elevation and the percentage of watershed in conifers had correlation coefficients in the -0.3 to -0.5 range. These results support the hypothesis that throughfall and soils in coniferous forests contribute to water acidity.

Many watershed attribute variables had correlation coefficients in the range of 0.2 to 0.3 with pH and ANC. Such variables include those representing the percentage of a lake's watershed having

- (1) soils with high or moderate infiltration rates,
- (2) moderately steep soils,
- (3) soils with a depth to low-permeability horizon or shallow soils ≤ 100 cm
- (4) bedrock with low to medium ANC,
- (5) soils with a cation exchange capacity ≤ 20 meq/100 g,
- (6) soils with high organic matter content,
- (7) hardwood or no forest, and
- (8) moderately acid wetlands.

Moderately acid wetlands expressed as a percentage of lake perimeter or area also had relatively high correlation coefficients; although this wetland type was expected to indicate acidic lakes, it seems to consistently indicate nonacidic lakes. Correlation coefficients for lake area, drainage area, and lake volume were also relatively high for pH but not for ANC. For the other variables previously mentioned, correlation coefficients were similar for pH and ANC. Several other variables had highly significant coefficients, but the magnitude was so low ($r < 0.2$) that these values were not considered to indicate important relationships.

A large correlation coefficient for a watershed variable with lake pH or ANC does not necessarily mean that variable will be important in a MLR model or discriminant analysis (e.g., lake area or organic matter content of soil). In a multivariate analysis more than one predictor variable can account for the same variation in the dependent variable (i.e., collinearity), thus making some predictor variables less explanatory or completely unnecessary for an acceptable model. Because of this, details concerning how and why variables were eliminated have been given. Also, a variable with a small correlation coefficient may be useful for developing a good predictive model and understanding factors contributing to variation in the dependent variable (e.g., percentage of watershed in wetland, percentage of watershed disturbed, and percentage of stony soils).

Our estimates indicate that 23 to 29% of Adirondack headwater lakes are acidic ($\text{pH} \leq 5.0$ or $\text{ANC} \leq 0 \mu\text{eq L}^{-1}$). Also, 38 to 56% of headwater lakes may be in danger of becoming more acidic and are

already at levels detrimental to some fish species- ($\text{pH} \leq 5.5$ or $\text{ANC} \leq 50 \mu\text{eq L}^{-1}$). We believe these estimates are reasonable because they include observed data for about one-half of the lakes; predictive models were verified with independent subsets of data; and, if anything, the predictions underestimate the number of acidic headwater lakes. These estimates should be representative of all headwater lakes in the Adirondacks because the models are based on data for -63% of the AEZ (an area defined by the availability of wetland maps) and the lakes are scattered throughout the AEZ (Fig. 1).

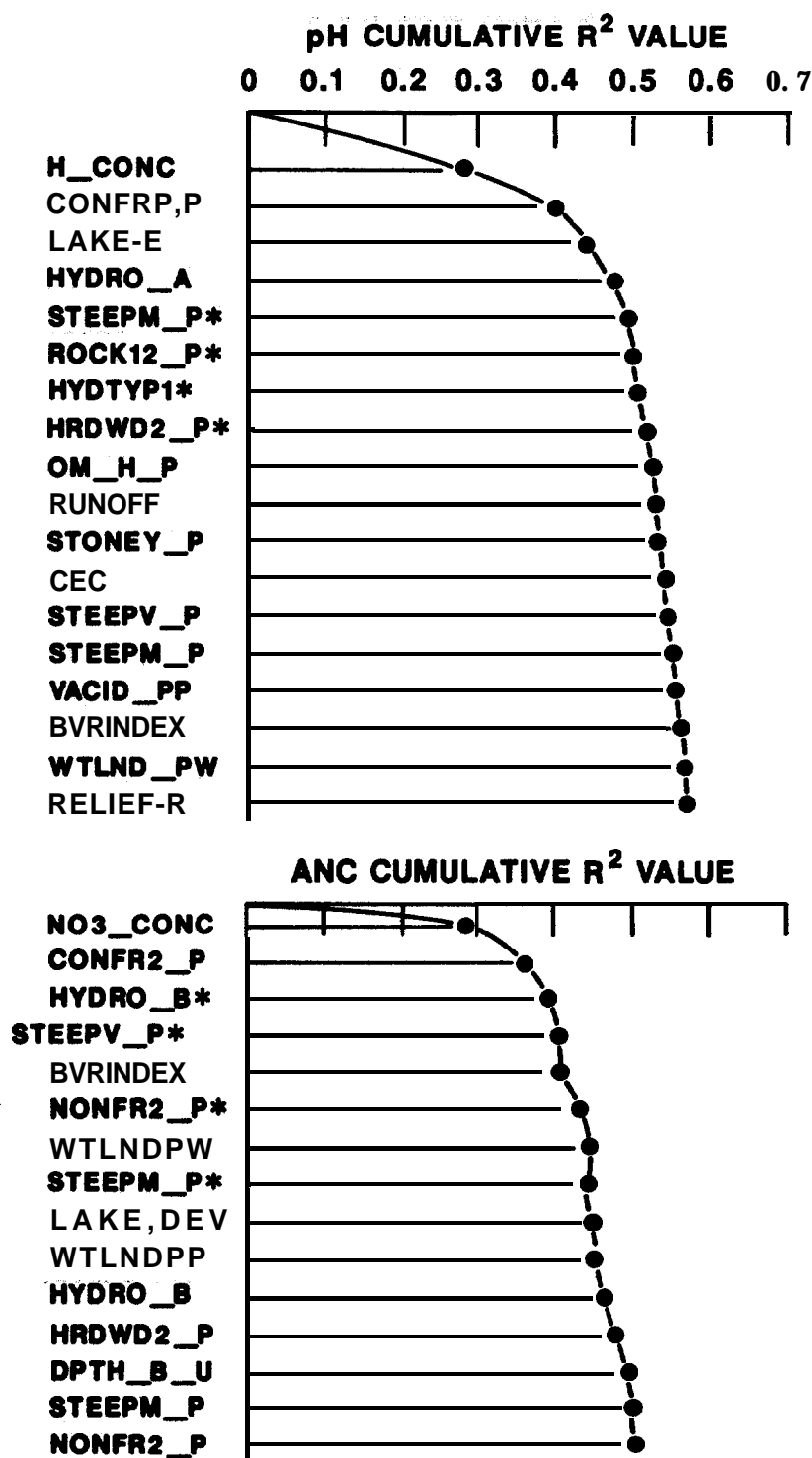
Baker and Harvey (1984) estimated that 20 to 25% of Adirondack lakes have a $\text{pH} < 5.0$. They used the distribution of lakes with known measurements, stratified by lake area and elevation, and performed a logistic multiple regression using lake area and elevation. Schofield (1976) found that 51% of Adirondack high-elevation lakes ($> 610 \text{ m}$) had a $\text{pH} < 5.0$. Colquhoun et al. (1984) found that 19% of the most recent pH measurements were < 5.0 for lakes monitored by the NYDEC (1047 waters with 1 or more measurements of summer surface pH from 1974 to the present). Schnoor et al. (1985) estimated that 19% of the Adirondack lakes have an $\text{ANC} \leq 40 \mu\text{eq L}^{-1}$; this estimate is based on the "Trickle-Down" model, coupled with MLR model results. Our estimates are in general agreement with other researchers' estimates, considering that each estimate was made with different data and different methods. Because headwater lakes tend to be at higher elevations, the percentage of headwater lakes having low pH and ANC is greater than that of all Adirondack lakes. However, our empirical models tend to underpredict the number of lakes in low pH and ANC categories, so these factors tend to balance each other out.

Perhaps even more important than the ability of a model to predict lake chemistry for lakes without chemistry data is the insight gained from a model about the possible importance of watershed attributes or the environmental processes they represent. For simplification, variables present in the following three models for pH, ANC, or both are considered to have strong relationships with lake chemistry: the selected minimum C_p models, the reduced $MAXR^2$ models, and the reduced discriminant models. Such variables include lake elevation; atmospheric inputs of H^+ and strong acid anions; watershed relief; soil stoniness, steepness, and infiltration rate; forest cover; wetland type; and beaver activity. Additional watershed attributes that appear to be related to a lake's acidity are the shape and hydrologic type, runoff, depth of bedrock, soil cation exchange capacity, land-cover disturbance, and human development within the watershed. Many of these variables also have significant and large correlation coefficients.

Significant statistical relationships among observational data do not prove cause and effect relationships; however, such analyses are useful for evaluating existing hypotheses about ecological processes (e.g., lake acidification). Watershed variables having strong relationships with lake pH or ANC will now be discussed concerning the conceptual framework of how watershed factors control surface water chemistry (Schnoor and Stumm 1985, Turner et al. 1986b). A few variables account for most of the variation explained (Fig. 15).

The strongest predictor of lake pH and ANC, of all the variables examined, is wet acidic input from the atmosphere. The H^+ and NO_3^- concentrations in precipitation were used in the models, but

ORNL-DWG 86-16648



*VARIABLES REPLACED BY ANOTHER VARIABLE

Fig. 15. Cumulative R^2 values for stepwise MLR models for pH and ANC.

wet deposition rates or SO_4^{2-} concentration in precipitation would give similar results because they are all collinear. Like the deposition variables, precipitation amount alone is significantly correlated (inversely) with lake pH and ANC. Dry deposition rate data are not available for the Adirondack watersheds. Given the strength of wet deposition rates as a predictor, a significantly higher R^2 could possibly be obtained if dry deposition or a suitable surrogate for it were available.

The next best predictor of lake pH and ANC to appear in all of the models is percentage of watershed in coniferous vegetation. Conifers intercept more cloud and fog moisture than deciduous trees and may indicate higher relative levels of dry deposition as well. They generally have a much larger total leaf area than deciduous trees, retain their foliage year round, and may scavenge more dry, deposited acids from the atmosphere. Conifers may also contribute to natural acidification of soils and soil water more than hardwoods (Brady 1974, Russell 1954). However, available data do not allow assessment of the relative importance of each of these individual processes.

Lake elevation is the third best predictor variable in the pH model, but it does not appear to be important for ANC. Lake elevation per se would not be expected to exert a mechanistic control on lake chemistry. Lake elevation, however, is strongly correlated with wet deposition and with many other variables, such as soil characteristics, hydrologic characteristics, and vegetation type, that are expected to be mechanistically associated with lake chemistry. For the Adirondacks, at least, elevation seems to be a readily available integrator of many other controlling variables.

Percentage of watershed with very steep slopes (>25% slope) and moderately steep slopes (>15% slope) appears to be important. Watersheds with steep slopes would be likely to have rapid, near-surface runoff through acidic, low base saturation and low SO_4^{2-} -retaining soil horizons, causing the water to have little chance for neutralization by cation exchange or weathering. The water might become more acidic in soils with very low base saturation (because of the salt effect) if the anions present in the soil solution are strong acid anions (SO_4^{2-} , NO_3^- , Cl^-) (Reuss and Johnson 1985). In contrast, if drainage water encounters easily weathered minerals or soils with relatively high base saturation before reaching surface waters, the hydrogen ions will be consumed or exchanged for base cations. Watersheds with a large percentage of soils with low to moderate runoff potential (soil hydrologic groups A and B) allow more infiltration of water than group C soils and, therefore, more contact with weatherable materials or higher base-saturation soil horizons. The percentage of watersheds having soil hydrologic groups A and B also appears to be important, showing a positive relationship with pH and ANC.

Two additional variables account for some variation in lake ANC: beaver activity and percentage of watershed that is not in forests (barren, developed, etc.). As expected, the percentage of watershed not in forests and the percentage of hardwoods are positively correlated with ANC. Beaver activity is negatively correlated with lake pH and ANC and is also present in the selected MLR models for pH

and ANC, although it does not appear to be as explanatory for pH as for ANC. Lower pH values in beaver ponds have been documented by several investigators (Rupp 1955, Salyer 1935, Adams 1953, Gard 1961, Leighton 1950, and Needham 1938). One factor that may account for the lower pH values is increased humic acid content of waters because of submerged vegetation and soils. Beaver activities may not be causing acidification of surface waters but may reflect the results of another watershed attribute not in the data set. Driscoll et al. (in press) reported increased ANC below a beaver dam

Watershed attributes selected as good predictor variables are similar for the pH and ANC selected models. In the reduced models, the variables differ somewhat -- physical and chemical soil characteristics are important for pH and forest and wetland types are important for ANC. Wet acidic inputs are the single characteristic explaining the most variability in lake chemistry. Lake elevation and the percentage of watershed in coniferous forest are also important factors. The extent to which lake elevation and percentage of watershed in coniferous vegetation are surrogates for dry deposition inputs is uncertain (e.g., high elevation and high proportion of conifers may indicate higher levels of dry deposition). To the extent that the distribution of conifers is correlated with the distribution of higher deposition, acidic deposition amounts (wet plus dry) may explain most of the variability observed in the pH and ANC of Adirondack headwater lakes. To the extent that lake elevation is an integrator (a surrogate) of other watershed processes and that coniferous vegetation modifies natural soil and water acidification by means other than dry deposition

or fog interception, watershed characteristics other than atmospheric deposition inputs may be important in controlling lake chemistry.

In the Upper Midwest, statistical analyses using cluster and discriminant techniques show that lake hydrologic type (e.g., drainage and seepage) and lake surface area are characteristically associated with the variability in lake alkalinity (Eilers et al. 1983, Schnoor et al. 1985, Schnoor et al. 1986). Although Eilers et al. (1983) did not find surficial geology or soil type to be related to lake alkalinity in Wisconsin, Schnoor and Nikolaidis (1983) found that bedrock geology, soil pH, and runoff helped discriminate between seepage/inflow lakes and drainage/spring lakes in Michigan. The shoreline development factor, percentage of watershed not in forest, and bedrock geology helped discriminate between seepage and inflow lakes. In Minnesota Rapp et al. (1985) found that bedrock type, forest type, lake morphology, lake hydrologic type, and deposition were associated with the variability in lake alkalinity.

In the Northeast Schnoor et al. (1985) used MLR with alkalinity as the dependent variable to identify watershed characteristics related to the variability in lake chemistry. They found, as we did, that precipitation or deposition is strongly associated with lake water quality. Schnoor et al. (1985) found that bedrock type is also a good predictor variable for the Northeast and that soil pH is important in the Adirondacks. In contrast, we found physical soil characteristics (i.e., percentage of watershed having steep soils and moderate to high infiltration rates) and the percentage of watershed in conifers to be consistent predictors for both pH and alkalinity in the Adirondacks.

These results, though not markedly different from earlier results (Hunsaker et al. 1986), are the outcome of statistical tests performed on a refined data set. Measurements of the size of small lakes were refined, and additional variables were added (beaver index, runoff, and depth to bedrock). Complex lakes are being added to the data base and will be examined later. Data for several additional variables, such as dry deposition, lake retention time, forest age, dissolved organic carbon concentrations, and complete anion data in lake water, will still be lacking. These attributes are needed to better resolve the importance of watershed disturbance, forest regrowth, and in-lake processes in controlling surface water pH and to quantify the relative importance of organic anions and strong acid anions in the lake water. The ongoing Adirondack Lake Survey by NYDEC will supply some of these data and provide an additional data set for validation of the models.

Empirical models have several appealing features for regional assessments: they are relatively easy, inexpensive, and rapid to use; there are diagnostic techniques for identifying questionable data or bias; and the models can be verified. Empirical models are useful for exploratory research such as this study because a large number of variables can be screened uniformly and important associations identified. Such associations, however, do not imply cause and effect relationships. These models can be developed from either experimental or observational data, but reliable multivariate data for large regions are rarely available. As shown by this study, empirical models are useful for evaluating resources at risk for large geographic areas; however, the uncertainty associated with individual lake predictions

can be large, and process models using experimental data may be more appropriate for individual lake predictions or for future predictions of lake acidification. However, the results from empirical models can provide important information about variables or watershed processes that should be considered during process model development.

REFERENCES

- Adams, A. K. 1953. Some physico-chemical effects of beaver dams upon Michigan trout streams in the Watersmeet area. Ph.D. thesis. University of Michigan, Ann Arbor.**
- Altshuller, A., and R. Linthurst (eds.). 1984. The acidic deposition phenomenon and its effects: Critical assessment review papers, Vol. II. EPA-600/8-83-016BF, U.S. Environmental Protection Agency, Washington, D.C.**
- Baker, J. P. 1984. Fish. pp. 5-74 to 5-133. IN A. Altshuller and R. Linthurst, (eds.), The Acidic Deposition Phenomenon and Its Effects: Critical Assessment Review Papers, Vol. II. EPA-600/8-83-016BF, U.S. Environmental Protection Agency, Washington, D.C.**
- Baker, J. P., T. B. Harvey, and J. P. Nicolette. 1984. Compilation of available data on the status of fish populations in regions of the northeastern U.S. susceptible to acid deposition. Draft final report to the U.S. Environmental Protection Agency. Acid Deposition Program North Carolina State University, Raleigh.**
- Baker, J. P., and T. B. Harvey. 1984. Critique of acid lakes and fish population status in the Adirondack Region of New York state, draft. Acid Deposition Program North Carolina State University, Raleigh, N.C. (draft).**

- Barnthouse, L. W 1982. Density-dependent growth: A critique of the utilities' "empirical evidence." pp. V-1--V-40. IN**
- L. W Barnthouse, W Van Winkle, J. Columbek, G. F. Cada, C. P. Goodyear, S. W Christensen, J. B. Cannon, and D. W Lee, Impingement impact analyses, evaluation of alternative screening devices, and critiques of utility analyses relating to density-dependent growth, the age structure of the Hudson River striped bass population, and the LMS Real-Time Life-Cycle Model, . Vol. II of The Impact of Entrainment and Impingement of Fish Populations in the Hudson River Estuary. ORNL/NUREG/TM-385/V2 (NUREG/CR-2220/V2).**
- Bogucki, D. J., and G. K. Gruending. 1982. Wetland interpretation and mapping project for the Adirondack Park. Final report to Adirondack Park Agency, Ray Brook, N.Y.**
- Brady, N. C. 1974. The nature and properties of soils, 8th ed. MacMillan Publishing Co., New York.**
- Call, M W 1966. Beaver pond ecology and beaver-trout relationships in southeast Wyoming. University of Wyoming and Wyoming Game and Fish Commission, Laramie.**
- Colquhoun, J., W Kretser, and M Pfeiffer. 1984. Acidity status update of lakes and streams in New York state. New York State Department of Environmental Conservation, Albany.**
- Conover, W J. 1980. Practical Nonparametric Statistics. John Wiley and Sons, New York.**
- Curran, R. Personal communication. Adirondack Park Agency, New York.**

- Draper, N. R. and H. Smith. 1981. Applied Regression Analysis, 2nd ed. John Wiley and Sons, New York.**
- Driscoll, C. T., B. J. Wiskowski, C. C. Cosentini, and M. E. Smith. Processes regulating temporal and longitudinal variations in the chemistry of a low-order woodland stream in the Adirondack region of New York. Biogeochemistry (in press).**
- Duhaine, P. F., A. G. Everett, and W. C. Retzsch. 1983. Adirondack land use: A commentary on past and present impacts on terrestrial and aquatic ecosystems. Report to the American Petroleum Institute, Washington, D.C.**
- Durfee, R. C., and the Geographic Data Systems Section. 1986. Review of geographic processing techniques applicable to regional analysis. ORNL/CSD/TM 226 (in press).**
- Eilers, J. M., G. E. Glass, and K. E. Webster. 1983. Relationships between susceptibility of lakes to acidification and factors controlling lake water quality: Hydrology as a key factor. Can. J. Fish. Aquat. Sci. 40:1896-1904.**
- Francis, M. M., R. J. Naiman, and J. M. Melillo. 1985. Nitrogen fixation in subarctic streams influenced by beaver (*Castor canadensis*). Hydrobiologia 121:193-202.**
- Gard, R. 1961. Effects of beaver on trout in Sagehen Creek, California. J. Wild. Mgt. 25(3):221-242.**
- Goldstein, R. A. 1983. The integrated lake-watershed acidification study. Vol. I: Model principles and application procedures. EPRI EA-321. Electric Power Research Institute.**
-

- Goodyear, C. P. and S. W. Christensen. 1984. On the ability to detect the influence of spawning stock on recruitment. *North Am J. Fish. Manage.* 4:186-193.
- Gorham E., S. J. Eisenreich, J. Ford, and M. V. Santelmann. 1985. The chemistry of bog waters. pp. 339-363. IN W. Stumm (ed.), *Chemical Processes in Lakes*. John Wiley & Sons, New York.
- Gruendling, G. K., D. J. Bogucki, and K. B. Adams. 1985. Data collection for testing alternative hypotheses to increased acidification and fish population declines in Adirondack surface waters. Final report to Oak Ridge Nat. Lab. for Subcontract No. 19X-89684. State University of New York, Plattsburgh.
- Hunsaker, C. T., J. L. Malanchuk, R. J. Olson, S. W. Christensen, and R. S. Turner. 1986. Adirondack headwater lake chemistry relationships with watershed characteristics. *Water, Air, Soil Pollut.* 31:79-88.
- Hutchinson, G. E. 1957. A treatise on limnology. *Geography, Physics, and Chemistry*, Vol. 1. John Wiley & Sons, New York.
- Johnson, D. W., I. S. Nilsson, J. O. Reuss, H. M. Seip, and R. S. Turner. 1985. Predicting soil and water acidification: Proceedings of a workshop. ORNL/TM 9258.
- Krug, E. C. and C. R. Fink. 1983. Acid rain on acid soil. A new perspective. *Science* 221:520-525.
- Leighton, R. 1950. Beaver survey in New Hampshire. New Hampshire Fish and Game Department, Quart. Prog. Rept., Proj. 9-R-4, (unpubl.).

- Linthurst, R. A., D. H. Landers, J. M. Eilers, D. F. Brakke,
W. S. Overton, E. P. Meier, and R. E. Crowe. 1986.
Characteristics of lakes in the eastern United States. Vol. I:
Population descriptions and physicochemical relationships.
EPA-600/4-86-007A. U.S. Environmental Protection Agency,
Washington, D.C.
- Mallows, C. L. 1973. Some comments on C_p . *Technometrics* 15:661-675.
- Marcus, M. D., B. R. Parkhurst, and F. E. Payne. 1983. An assessment
of the relationship among acidifying depositions, surface water
acidification, and fish populations in North America, Vol. 1.
EPRI-EA-3127. Electric Power Research Institute, Palo Alto, Calif.
- Mason, J. and H. M. Seip. 1985. The current state of knowledge on
acidification of surface waters and guidelines for research.
Ambio 14:45-51.
- McFadden, J. T. 1963. An example of inaccuracies inherent in
interpretation of ecological field data. *Am Nat.* 97:99-116.
- Naiman, R. J., J. M. Melillo, and J. E. Hobbie. 1986. Ecosystem
alteration of boreal forest streams by beaver (Castor canadensis).
Ecology 67(5):1254-1269.
- Nair, D. R. 1984. Multiple regression analyses of factors affecting
alkalinity of lakes in northeastern United States. M.S. thesis.
University of Iowa, Iowa City.
- National Atmospheric Deposition Program and Association of State
Agricultural Experiment Stations of the North Central Region.
No date. Distribution of surface waters sensitive to acidic
precipitation: A state-level atlas. NC-141.

Needham, P. R. 1938. Trout streams -- conditions that determine their productivity and suggestions for stream and lake management.

Cornstock Publ. Co., Ithaca, New York.

Norton, S. A. 1980. Geologic factors controlling the sensitivity of aquatic ecosystems to acidic precipitation. pp. 521-531.

**IN D. S. Shriner, C. R. Richmond, and S. E. Lindberg (eds.),
Atmospheric Sulfur Deposition: Environmental Impact and Health
Effects Ann Arbor Science, Ann Arbor, Mich.**

**Norton, S. A., J. J. Akielaszek, T. A. Haines, K. L. Stronberg, and
J. R. Longcore. 1982. Bedrock geologic control of sensitivity of
acidic ecosystems in the United States to acidic deposition.
National Atmospheric Deposition Program North Carolina State
University, Raleigh.**

**Rapp, G., J. D. Allert, B. W. Liukkonen, J. A. Ilse, O. L. Loucks, and
G. E. Glass. 1985. Acid deposition and watershed characteristics
in relation to lake chemistry in northeastern Minnesota. Environ.
Int. 11:425-440.**

**Reckhow, K. H., R. W. Black, T. B. Stockton, Jr., J. D. Vogt, and
J. G. Wood. 1985. Empirical models of fish response to lake
acidification. Draft final report to the U.S. Environmental
Protection Agency. Duke University, Durham N.C.**

**Retzsch, W. C., A. G. Everett, P. F. Duhaine, and R. Notwanger. 1982.
Alternative explanations for aquatic ecosystems effects attributed
to acidic deposition. Report to the Utility Air Regulatory Group.**
**Reuss, J. O., and D. W. Johnson. 1985. Effect of soil processes on the
acidification of water by acid deposition. J. Environ. Qual.
14:26-31.**

- Rosen, A. E., R. J. Olson, G. K. Gruendling, D. J. Bogucki,
J. L. Malanchuk, R. C. Durfee, R. S. Turner, K. B. Adams,
D. L. Wilson, P. R. Coleman, C. C. Brandt, and C. T. Hunsaker.
1986. An Adirondack Watershed Data Base: Attribute and mapping
information for regional acidic deposition studies. ORNL/TM 10144.
- Rupp, R. S. 1955. Beaver-trout relationships in the headwaters of
Sunkhaze stream. *Maine Trans. Am. Fish. Soc.* 84:75-85.
- Rush, R. M., R. B. Honea, E. C. Krug, R. W. Peplies, J. E. Dobson, and
F. P. Baxter. 1985. An investigation of landscape and lake
acidification relationships. ORNL/TM 9754.
- Russell, E. J. 1954. Soil Conditions and Plant Growth. Longmans,
Green, and Co., New York.
- Salyer, J. C. 1935. Preliminary report on the beaver-trout
investigation. *Am. Game* 24(6):13-15.
- SAS Institute, Inc. 1983. SAS Views: SAS Processing. Cary, N.C.
- SAS Institute, Inc. 1985. SAS User's Guide: Statistics, Version 5
Edition. Cary, N.C.
- Schindler, D. W., M. A. Turner, M. P. Stainton, and G. A. Linsey.
1986. Natural sources of acid neutralizing capacity in low
alkalinity lakes of the Precambrian Shield. *Science*
232(4752):844-847.
- Schofield, C. 1976a. Acid precipitation: Effects on fish. *Ambio*
5:228-230.
- Schofield, C. 1976b. Dynamics and management of Adirondack fish
populations. Project report April 1, 1975-March 31, 1976,
No. F-28-R. New York Department of Environmental Conservation,
Albany, N.Y.

- Schnoor, J. L., and W Stumm 1985. Acidification of aquatic and terrestrial systems. pp. 311-338. IN W Stumm (ed.), Chemical Processes in Lakes. John Wiley & Sons, New York.**
- Schnoor, J. L., D. R. Nair, M A. Kuhns, S. J. Lee, and N. P. Nikolaidis. 1985. Lake resources at risk due to acid deposition in the northeastern United States, draft final report to U.S. Environmental Protection Agency. Department of Civil and Environmental Engineering, University of Iowa, Iowa City.**
- Schnoor, J. L., N. P. Nikolaidis, and G. E. Glass. 1986. Lake resources at risk to acidic deposition in the upper Midwest. J. Water Pollut. Control Fed. 58(2):139-148.**
- Turner, R. S., R. J. Olson, and C. C. Brandt. 1986a. Areas having soil characteristics that may indicate sensitivity to acidic deposition under alternative forest damage hypotheses. ORNL/TM 9917.**
- Turner, R. S., D. W Johnson, J. W Elwood, W Van Winkle, R. B. Clapp, and J. O. Reuss. 1986b. Factors affecting response of surface waters to acidic deposition. ORNL/TM 9787.**
- Viro, P. J. 1974. Effects of forest fire on soil. IN T. T. Kozlowski and C. E. Ahlgren (eds.), Fire and Ecosystems. Academic Press, New York.**
- Wright, R. F. 1983. Predicting acidification of North American lakes. Report o-81036. Norwegian Institute for Water Research, Oslo.**

Appendix A

PROCEDURES USED IN DETERMINING WHICH COLLINEAR VARIABLE TO ELIMINATE

This appendix describes in more detail the procedures used to determine which variable to eliminate at each step of the collinearity analysis. A summary description of these procedures is provided in Sect. 3.2.1 of the main text.

The SAS procedure REG (SAS 1985) was applied repeatedly, using the COLLIN option. For each run, the collinearity diagnostics were examined. If the largest condition index (CI) was >30 , 2 or more parameters corresponding to particular variables usually also had a variance-decomposition proportion (VP) > 0.5 . In such a case (i.e., $CI > 30$), the corresponding variables with a VP > 0.5 were "marked" for possible elimination. If <2 variables (i.e., not including the intercept) had a VP > 0.5 and the next largest CI was also >30 , the sums of the VPs for the 2 (or 3, if necessary to obtain at least 2 variables) largest CIs were examined, and all variables with combined VPs >0.5 were marked. Once the marked variables were identified, they were examined in relation to the following priority list, and some of them were unmarked until only one marked variable remained. This sole remaining marked variable was eliminated from all remaining analyses. This procedure was followed iteratively until the largest CI no longer exceeded 30.

PRIORITY LIST

1. The intercept always was judged appropriate for inclusion in the model; it was not marked, and, therefore, it was never eliminated.
2. The Cp statistic was calculated for all models, omitting marked variables one at a time until only one marked variable remained. This variable was eliminated from further consideration. The Cp values were examined for amount of difference and for clusters of like values. If the values differed by more than -1 to 2%, one or more remaining marked variables were identified as having lower Cp values, and all other remaining marked variables were unmarked.
3. For this step only, all remaining variables (i.e., not eliminated by prior passes through this priority list) were considered, whether marked or unmarked. If one or more remaining marked variables were the only variables remaining in their group (with groups being defined by letter codes in Table A-1), these were unmarked, provided that this did not result in unmarking all of the marked variables. The intent of this was to retain in the final model, where possible, at least one variable in each group.
4. All variables had been classified according to the degree to which they represented direct measures of potentially mechanistic processes (Table A-1). If the remaining marked variables differed with respect to this class, all remaining marked variables except those in the least mechanistic (largest number) class were unmarked (e.g., a soils variable would be unmarked, and elevation or lake area would remain marked for possible elimination if all were collinear).

Table A-1. Groupings and other information used in process of selecting variables for elimination in multicollinearity analysis

(See key at end of table for explanation of column headings)^a

Variable name by group	Mechanistic potential	Reliability	Availability		Correlation ^d	
			NSWS ^b	WDNR ^c	pH	ANC
Group A						
LAKE-A	3	1	1	1	** +	
WTRSHD_R	2	1	1	1		
DRAIN-A	2	1	1	1	** +	
Group B						
LAKE-DEV	2	1	4	1		
Group C						
HYDTYP1	2	2	1	1		
HYDTYP2	2	2	1	1	** -	** -
HYDTYP3	2	2	1	1	* +	
Group D						
ELEV	3	1	1	1	** -	** -
Group E						
RELIEF-R	2	2	1	1		* +
STONEY_P	1	1	2	2		
ROCK-P	1	1	2	2		
HYDRO_A	1	1	2	2	** +	** +
HYDRO_B	1	1	2	2	** +	** +
HYDRO_C	1	1	2	2		
STEEPM_P	1	1	2	2	* -	* -
STEEPV_P	1	1	2	2		
SHL2_B_P	1	1	2	2		
SHL2_P_P	1	1	2	2	** -	** -
SHL2_R_P	1	1	2	2	* -	
DPTH_B_U	1	1	2	1		
DPTH_P_U	1	1	2	2		
DPTH_R_U	1	1	2	2		
Group F						
ASPECTS	2	2	2	2		
Group G						
ROCK12_P	1	1	3	3		
Group H						
ACID-P	1	1	2	2		
BSA_L_P	1	1	2	2		
BSC_L_P	1	1	2	2		
CECS_L_P	1	1	2	2	** +	** +
OM_H_P	1	1	2	2	* -	* -
PH_VL_P	1	1	2	2		
PHC_VL_P	1	1	2	2		
ACID-EX	1	1	2	2	* -	* -
CEC	1	1	2	1		
ORG_MAT	1	1	2	2	* -	

Table A-1. (continued)

Variable name by group	Mechanistic potential	Reliability	Availability		Correlation ^d			
			NSWS ^b	WDNR ^c	pH	A	N	C
Group I								
CONFR2_P	2	1	1	1	** -	** -		
HRDWD2_P	2	1	1	1	** +			
NONFR2_P	2	1	1	1	** +	* +		
Group J								
WTLND_PP	2	1	4	4				
VACID_PP	2	2	4	4	* -	* -		
NACID_PP	2	2	4	4	* +			
WTLND_PW	2	1	1	1				
VACID_PW	2	2	4	4				
NACID_PW	2	2	4	4	* +			
WTLND_PL	2	1	1	1				
VACID_PL	2	2	4	4				
NACID_PL	2	2	4	4				
Group K								
DISTRB_P	2	1	1	1				
Group L								
BVRINDEX	2	1	4	4	* -	** -		
Group M								
CABN78_R	1	1	4	1	* +			
Group N								
H_WET	1	1	1	2	** -	** -		
NO3_WET	1	1	1	2	** -	** -		
SO4_WET	1	1	1	2	** -	** -		
H_CONC	1	1	1	2	** -	** -		
NO3_CONC	1	1	1	2	** -	** -		
SO4_CONC	1	1	1	2	** -	** -		
Group O								
RUNOFF	1	2	1	2	** -	** -		
Group P								
PPT	1	1	1	2	** -	** -		

^aRefer also to Table A-2.

^bNational Surface Water Survey data (Linthurst et al., Kanciruk et al. 1986).

^cWisconsin Department of Natural Resources data (Webster et al. 1983).

^d* indicates $p \leq 0.05$, ** indicates $p \leq 0.01$, and the direction of the correlation is indicated by a + or -.

Key to Table A-1

Mechanistic Potential:

- 1: Direct measure of potentially mechanistic process. These are the most mechanistic variables and are best from the point of view of mechanisms.
- 2: Indirect measure, conflicting direction of mechanistic effect possible, or both.
- 3: Surrogate variable. These are the least mechanistic variables and are worst from this point of view.

Reliability:

- 1: Higher reliability.
- 2: Lower reliability.

Availability (NSWS or Eilers):

- 1: Available as needed or able to be calculated with minimal effort.
- 2: Available in digital form but not extracted for lakes in the data set. Could be developed with moderate effort.
- 3: Could be developed with considerable effort.
- 4: Not available, not likely to become available. This is not good.

Relatedness:

Letter codes are used to define groups of variables considered to be related strongly to each other in terms of type of mechanistic potential. Inherently nonmechanistic variables are, of course, not related to any other variables in terms of this criterion.

Correlations (with pH or ANC):

Results are indicated for the calibration data set; levels of significance would be expected to become greater with the full data set in many cases. The direction of the correlation is included for the convenience of the reader, but this information was not used in the collinearity procedure.

- ** -: Spearman correlation significant at $P < 0.01$, negative correlation.
- ** +: Spearman correlation significant at $P < 0.01$, positive correlation.
- * -: Spearman correlation significant at $P < 0.05$, negative correlation.
- * +: Spearman correlation significant at $P < 0.05$, positive correlation.

5. If the remaining marked variables differed with respect to reliability (Table A-1), all of them except those in the least reliable class were unmarked. (Those in the least reliable class thus remained as candidates for elimination.)
6. The remaining marked variables were evaluated with respect to availability in the larger National Surface Water Survey (NSWS) data set outside of the Adirondacks (see Table A-1). If these marked variables differed substantially with respect to availability in this data set (i.e., did not all have the same code value), all remaining marked variables except those in the least available (largest code number) category were unmarked.
7. Step 6 was repeated with respect to the Eilers' data set.
8. If two or more of the remaining marked variables were closely related (i.e., were in the same group as defined in Table A-1) and one or more of the other remaining marked variables were unrelated, the unrelated remaining marked variable(s) were unmarked.
9. The remaining marked variables were evaluated with respect to their correlation with pH or alkalinity (whichever was appropriate) in the specification data set (CORRELATION columns in Table A-1). If some but not all remaining marked variables were significantly correlated ($p < 0.05$), these were unmarked, regardless of whether the direction of the correlation was in accordance with expectations.

10. With respect to the wet deposition variables (Group N), if one or more hydrogen ion variables remained marked, all of them for which a corresponding (i.e., concentration or wet deposition, as the case may be) sulfate or nitrate variable also remained marked were unmarked. Because all atmospheric input variables were highly correlated, inclusion of the hydrogen ion over inclusion of the sulfate or nitrate was favored. The purpose was to hinder the inference that inclusion of one of these anions in the model means it is more important than the other in controlling lake acidification.
11. If more than one wetland variable (Group J) remained marked, a separate wetland priority list (Table A-2) was examined. If the remaining marked wetland variables fell in more than one wetland group according to this list, all marked variables except those marked wetland variables falling in the lowest priority group were unmarked.
12. At this point, one or more marked variables remained. If only one marked variable remained, that variable was eliminated. If two or more marked variables remained, the one with the highest VP was eliminated. All other variables were unmarked in preparation for the next run.

**Table A-2. Priority groups for wetland variables
used in process of selecting variables for
elimination in multicollinearity analysis**

Variable	Priority group^a
WTLND_PW	1
WTLND_PP	2
WTLND_PL	3
VACID_PW	4
NACID_PW	4
VACID_PP	5
NACID-PP	5
VACID_PL	6
NACID-PP	6

a1 = highest priority to remain in model.

REFERENCE FOR APPENDIX A

- Linthurst, R. A., D. H. Landers, J. M. Eilers, D. F. Brakke,
W. S. Overton, E. P. Meier, and R. E. Crowe. 1986.
Characteristics of lakes in the eastern United States. Vol. I:
Population descriptions and physicochemical relationships.
EPA-600/4-86-007A. U.S. Environmental Protection Agency,
Washington, D.C.**
- Kanciruk, P., M. Gentry, R. McCord, L. Hook, J. Eilers, and M. D. Best.
1986. National Surface Water Survey: Eastern Lake Survey - Phase
I, Data base dictionary. ORNL/TM 10153.**
- Webster, K. E., M. D. Johnson, L. Huberty, M. Eckardt, S. Schueler,
T. Franti, T. Tellun, and J. E. Eilers. 1983. Documentation and
quality assurance summary for morphometric and terrigenous
variables for the Wisconsin 60 lake subset. Wisconsin Department
of Natural Resources, Rhinelander, Wis.**
- SAS Institute, Inc. 1985. SAS User's Guide: Statistics, Version 5
Edition. Cary, N.C.**

4

4

6

3

2

1

INTERNAL DISTRIBUTION

- | | | | |
|--------|-------------------|--------|-------------------------------|
| 1. | L. J. Allison | 38. | M. E. Mitchell |
| 2. | S. I. Auerbach | 39. | R. V. O'Neill |
| 3-7. | J. J. Beauchamp | 40-44. | R. J. Olson |
| 8. | D. E. Carpenter | 45. | S. F. Railsback |
| 9-13. | S. W. Christensen | 46. | R. M. Reed |
| 14. | R. B. Cook | 47-48. | D. E. Reichle |
| 15. | N. S. Dailey | 49. | C. R. Richmond |
| 16. | J. E. Dobson | 50. | A. E. Rosen |
| 17. | R. C. Durfee | 51. | R. M. Rush |
| 18. | J. W. Elwood | 52. | M. J. Sale |
| 19. | M. P. Farrell | 53. | D. S. Shriner |
| 20. | W. Fulkerson | 54. | R. R. Turner |
| 21. | R. H. Gardner | 55-59. | R. S. Turner |
| 22. | R. L. Graham | 60. | W. Van Winkle |
| 23. | S. G. Hildebrand | 61. | Central Research Library |
| 24. | R. B. Honea | 62-76. | ESD Library |
| 25-34. | C. T. Hunsaker | 77-78. | Laboratory Records Department |
| 35. | D. W. Johnson | 79. | Laboratory Records, RC |
| 36. | R. J. Luxmoore | 80. | ORNL Patent Office |
| 37. | R. A. McCord | 81. | ORNL Y-12 Technical Library |

EXTERNAL DISTRIBUTION

82. R. Abler, Director, Geography and Regional Science Program
National Science Foundation, 1800 G Street, N.W., Room 312,
Washington, DC 20550
83. K. B. Adams, State University of New York/Plattsburgh Center
for Earth and Environmental Science, Hudson Hall, Plattsburgh,
NY 12901
84. Richard B. Alexander, U.S. Geological Survey, Reston, VA 22092
85. R. April, Department of Geology, Colgate University, Hamilton,
NY 13346
86. Robert G. Bailey, Land Management Planning Systems, Rocky
Mountain Forest and Range Experiment Station, 3825 East
Mulberry Street, Fort Collins, CO 80524
87. J.P. Baker, Kilkelly Environmental Associates, Water Garden,
P.O. Box 31265, Raleigh, NC 27622
88. Tom Barringer, 400 West Hottel Street, #801, Philadelphia,
PA 19119
89. D. A. Bennett, Acid Deposition Assessment Staff,
U.S. Environmental Protection Agency, 401 M Street, SW RD-680,
Washington, DC 20460
90. Dan Botkin, University of California, Santa Barbara, CA 93196

91. **D. J. Bogucki, State University of New York/Plattsburgh, Center for Earth and Environmental Science, Hudson Hall, Plattsburgh, NY 12901**
92. **R. L. Burgess, Chairman, Department of Environmental and Forest Biology, SUNY College of Environmental Sciences and Forestry, Syracuse, NY 13210**
93. **M. R. Church, Corvallis Environmental Research Laboratory, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333**
94. **J. R. Colquhoun, New York State Department of Environmental Conservation, Fisheries Division, 50 Wolf Road, Albany, NY 12233**
95. **B. J. Cosby, University of Virginia, Department of Environmental Sciences, Clark Hall, Charlottesville, VA 22903**
96. **R. P. Curran, Adirondack Park Agency, P.O. Box 99, Route 86, Ray Brook, NY 12977**
97. **M. B. David, University of Illinois, Department of Forestry, 1301 W. Gregory Drive, 110 Mumford Hall, Urbana, IL 61801**
98. **C. T. Driscoll, Syracuse University, Department of Civil Engineering, 237 Hinds Hall, Syracuse, NY 13244**
99. **J. M. Eilers, Environmental Research Laboratory/Corvallis, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333**
100. **W. E. Fallon, Acid Deposition Planning Staff, U.S. Environmental Protection Agency, 401 M Street, SW RD-680, Washington, DC 20460**
101. **G. J. Foley, Office of Environmental Process and Effects Research, U.S. Environmental Protection Agency, 401 M Street, SW EN-338, Washington, DC 20460**
102. **R. M. Friedman, Office of Technology Assessment, U.S. Congress, 600 Pennsylvania Avenue, SE, Washington, DC 20003**
103. **S. A. Gherini, Tetra Tech, Inc., 3746 Mt. Diablo Boulevard, Suite 300, Lafayette, CA 94549**
104. **G. E. Glass, Environmental Research Laboratory-Duluth, U.S. Environmental Protection Agency, 6201 Congdon Blvd., Duluth, MN 55804**
105. **W. Glaze, Environmental Sciences and Engineering Program, School of Public Health, University of California, Los Angeles, CA 90024**
106. **M. Godron, Centre d'Etudes Phytosociologiques et Ecologiques, CNRS, P.P. 5051, Route de Mende, 34033 Montpellier, France**
107. **C. R. Goldman, Professor of Limnology, Director of Tahoe Research Group, Division of Environmental Studies, University of California, Davis, CA 95616**
108. **R. A. Goldstein, Ecological Studies Program Electric Power Research Institute, 3412 Hillview Avenue, P.O. Box 10412, Palo Alto, CA 94303**
109. **G. K. Gruendling, State University of New York/Plattsburgh, Center for Earth and Environmental Science, Hudson Hall, Plattsburgh, NY 12901**

110. G. R. Hendrey, Brookhaven National Laboratory, 1 South Technology Street, Building 318, Upton, NY 11973
111. W. T. Hinds, Acid Deposition Planning Staff, U.S. Environmental Protection Agency, 401 M Street, SW RD-680, Washington, DC 20460
112. Thomas Hoekstra, Rocky Mountain Forest and Range Experiment Station, 240 West Prospect Street, Fort Collins, CO 80526
113. G. R. Holdren, Northrop Services, Inc., 200 SW 35th Street, Corvallis, OR 97333
114. C. Robert Horn, Water Division, U.S. Environmental Protection Agency (WH-586), 401 M Street, S.W., Washington, DC 20460
115. J. W. Huekabee, Manager, Ecological Studies Program, Electric Power Research Institute, 3412 Hillview Avenue, P.O. Box 10412, Palo Alto, CA 94303
116. Louis R. Iverson, Illinois Natural History Survey, 607 East Peabody Drive, Champaign, IL 61820
117. George Y. Jordy, Director, Office of Program Analysis, Office of Energy Research, ER-30, G-226, U.S. Department of Energy, Washington, DC 20545
118. E. Kaplan, Brookhaven National Laboratory, 1 South Technology Street, Building 475, Upton, NY 11973
119. J. Kinsman, International Science and Technology, 11260 Roger Bacon Drive, Suite 201, Reston, VA 22090
120. Karen Klima, U.S. Environmental Protection Agency (WH 595), 401 M Street, S.W., Washington, DC 20460
121. W. A. Kretser, New York State Department of Environmental Conservation, Route 86, Ray Brook, NY 12977
122. E. C. Krug, Atmospheric Chemistry Section, Illinois State Water Survey, 2204 Griffith Drive, Champaign, IL 61820
123. J. L. Kulp, Director of Research, National Acid Precipitation Assessment Program, 722 Jackson Place, NW Washington, DC 20503
124. J. R. Krummel, Energy and Environmental System Division, Bldg. 362, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439
125. D. A. Landers, Corvallis Environmental Research Laboratory, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333
126. M. Lang, Tetra Tech, Inc., 3746 Mt. Diablo Blvd., Suite 300, Lafayette, CA 94549
127. R. A. Linthurst, U.S. Environmental Protection Agency, Environmental Systems Monitoring Laboratory, MD-39, Research Triangle Park, NC 27711
128. O. L. Loucks, Holcomb Research Institute, Butler University, Indianapolis, IN 46208
- 129-133. J. L. Malanchuk, Acid Deposition Planning Staff, U.S. Environmental Protection Agency, 401 M Street, SW RD-680, Washington, DC 20460
134. C. J. Mankin, Director, Oklahoma Geological Survey, The University of Oklahoma, 830 Van Vleet Oval, Rm 163, Norman, OK 73019

135. **D. R. Marmorek, Environmental and Social Systems Analysts, Ltd., Box 12155 Nelson Square, Suite 705, 808 Nelson Street, Vancouver, BC Canada V6Z 2H2**
136. **H. M. McCammon, Director, Ecological Research Division, Office of Health and Environmental Research, Office of Energy Research, MS-E201, ER-75, Room E-233, Department of Energy, Washington, DC 20545**
137. **D. H. McKenzie, Corvallis Environmental Research Laboratory, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333**
138. **V. Meentemeyer, Department of Geography, University of Georgia, Athens, GA 30602**
139. **J. Messer, Corvallis Environmental Research Laboratory, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333**
140. **R. E. Miller, Department of Regional Science, University of Pennsylvania, Philadelphia, PA 19104**
141. **P. A. Mundy, Acid Deposition Assessment Staff, U.S. Environmental Protection Agency, 401 M Street, SW RD-680, Washington, DC 20460**
142. **R. M. Newton, Smith College, Department of Geology, Northampton, MA 01063**
143. **N. L. Oden, Brookhaven National Laboratory, Biomedical and Environmental Assessments Division, Upton, NY 11973**
144. **H. Olem, Tennessee Valley Authority, 248 401 Building, Chattanooga, TN 37401**
145. **J. M. Omernik, Corvallis Environmental Research Laboratory, U.S. Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR 97333**
146. **W. S. Overton, Oregon State University, Department of Statistics, Corvallis, OR 97331**
147. **R. W. Peplies, Department of Geography, East Tennessee State University, Johnson City, TN 37601**
148. **D. S. Porcella, Electric Power Research Institute, 3412 Hillview Avenue, P.O. Box 10412, Palo Alto, CA 94303**
149. **W. S. Platts, Intermountain Forest and Range Experiment Station, 316 East Myrtle, Boise, ID 83702**
150. **J. C. Randolph, School of Public and Environmental Affairs, Indiana University, 400 E. Seventh Street, Bloomington, IN 47405**
151. **G. Rapp, Jr., College of Science and Engineering, University of Minnesota, Duluth, MN 55812**
152. **P. L. Ringold, Deputy Director, Interagency Task Force on Acid Precipitation, 722 Jackson Place, NW Washington, DC 20503**
153. **Paul G. Risser, Vice President for Research, The University of New Mexico, 108 Scholes Hall, Albuquerque, NM 87131**
154. **V. S. Robinson, Dept. of Geology and Geography, Hunter College, CUNY, New York, NY 10021**
155. **J. L. Schnoor, University of Iowa, Department of Civil and Environmental Engineering, Engineering Building, 1134 EB, Iowa City, IA 52242**

156. **C. L. Schofield, Cooperative Fishery Institute,
Cornell University, 118 Fernow Hall, Ithaca, NY 14853**
157. **R. K. Schreiber, Eastern Energy and Land Use Team, U.S. Fish
and Wildlife Service, Box 705, Kearneysville, WV 25430**
158. **Dan Simberloff, Florida State University, Tallahassee, FL
32306**
159. **Richard A. Smith, U.S. Geological Survey, Reston, VA 22092**
160. **R. J. Stern, Director, Office of Environmental Compliance,
MS PE-25, FORRESTAL, U.S. Department of Energy,
1000 Independence Avenue, SW Washington, DC 20585**
161. **K. W. Thornton, Ford, Thornton, Norton and Associates,
3 Innwood Circle, Suite 220, Little Rock, AR 72211**
162. **M. G. Turner, Institute of Ecology, University of Georgia,
Athens, GA 30602**
163. **L. H. Weinstein, Program Director of Environmental Biology,
Cornell University, Boyce Thompson Institute for Plant
Research, Ithaca, NY 14853**
164. **K. A. Wheeler, Soil Conservation Service, U.S. Courthouse and
Federal Building, 100 S. Clinton Street, Room 771, Syracuse,
NY 13260**
165. **Ray Whittemore, National Council of the Paper Industry for
Air and Stream Improvement, Inc., Department of Civil
Engineering, Tufts University, Medford, MA 02155**
166. **Raymond G. Wilhour, Chief, Air Pollution Effects Branch,
Corvallis Environmental Research Laboratory,
U.S. Environmental Protection Agency, 200 SW 35th Street,
Corvallis, OR 97330**
167. **T. Williams, Office of Environmental Analysis, PE-26,
Room 4G-036, Forrestal Building, U.S. Department of Energy,
Washington, DC 20585**
168. **Frank 3. Wobber, Ecological Research Division, Office of
Health and Environmental Research, Office of Energy Research,
MS-E201, Department of Energy, Washington, DC 20545**
169. **M. G. Wolman, The Johns Hopkins University, Department of
Geography and Environmental Engineering, Baltimore, MD 21218**
170. **Office of Assistant Manager for Energy Research and
Development, Oak Ridge Operations, P.O. Box E,
U.S. Department of Energy, Oak Ridge, TN 37831**
- 171-200. **Technical Information Center, Oak Ridge, TN 37831**

