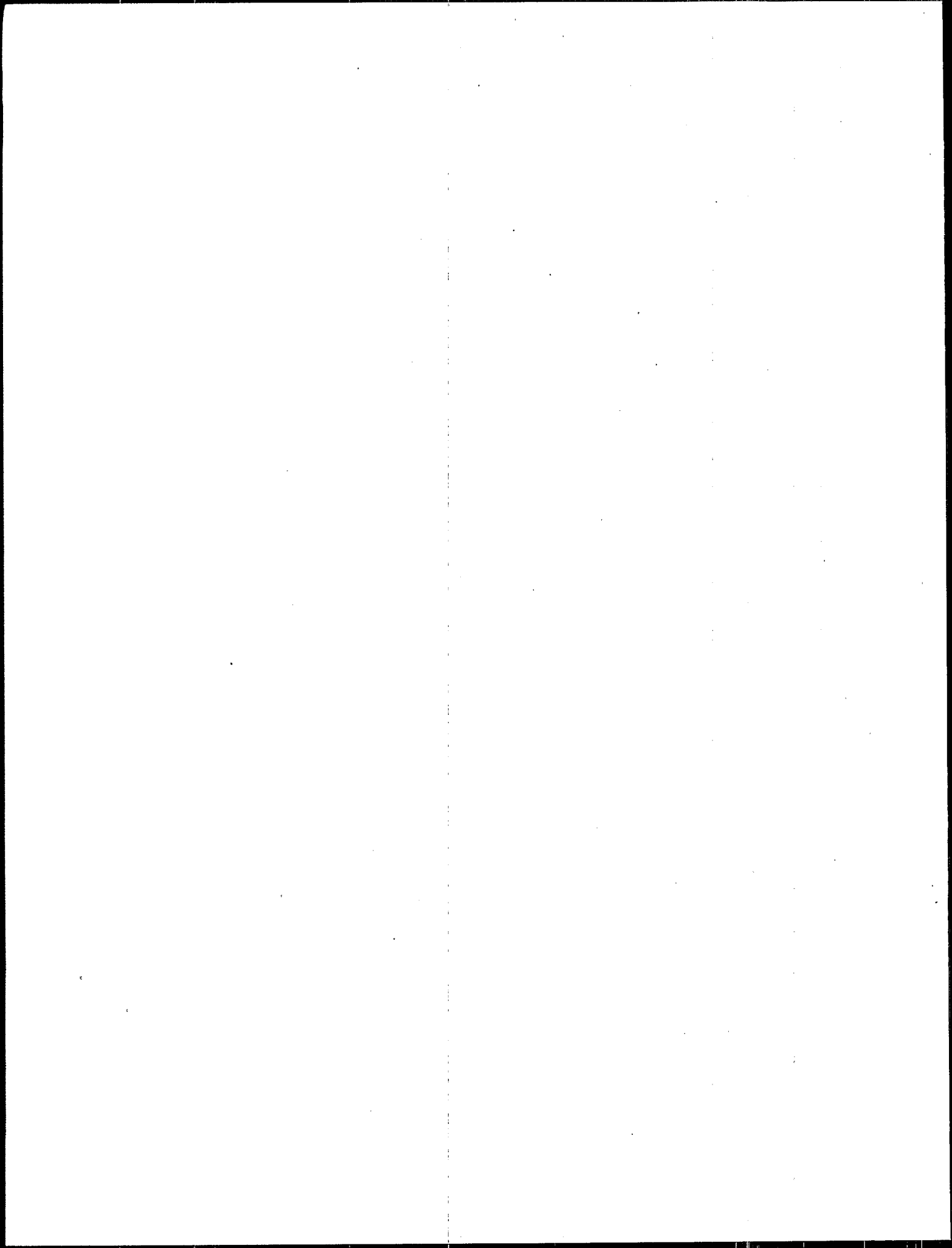**♻EPA**

User's Guide and Data Dictionary for Kenai Lakes Investigation Project

USER'S GUIDE AND DATA DICTIONARY FOR KENAI LAKES INVESTIGATION PROJECT

March 1992

By

Avis D. Newell
ManTech Environmental Technology, Inc.
U.S. EPA Environmental Research Laboratory
200 SW 35th Street
Corvallis, OR 97333

and

Mark E. Mitch
Department of Environmental Sciences
University of Virginia
Charlottesville, VA 22903

Printed on Recycled Paper

DISCLAIMER

The research described in this report has been funded by the U.S. Environmental Protection Agency. This document has been prepared at the EPA Environmental Research Laboratory in Corvallis, Oregon, through Contract No. 68-C8-0006 to ManTech Environmental Technology, Inc. It has been subjected to the Agency's peer and administrative review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1. INTRODUCTION

1.1 PURPOSE AND SCOPE

In 1984, the U.S. Environmental Protection Agency (EPA) implemented the National Surface Water Survey (NSWS) (Linthurst et al., 1986; Landers et al., 1987; Kaufmann et al., 1988) as part of the Aquatic Effects Research Program (AERP). The AERP conducted several integrated studies in areas containing surface waters considered potentially sensitive to change as a result of acidic deposition. The AERP addressed five major environmental policy issues:

1. The present status and extent of acidic and low alkalinity surface waters in the United States

2. The extent and magnitude of past change

3. The change to be expected in the future under various rates of acidic deposition

4. The maximum rates of deposition below which further change would not be expected

5. The rate of change or recovery of aquatic ecosystems if deposition rates decreased.

The NSWS focused its assessment on lakes and streams located in the contiguous United States. Since the majority of the systems examined in the NSWS receive moderate to high levels of acidic deposition, it is difficult to evaluate the role of natural factors in controlling the chemistry of aquatic ecosystems. Therefore, the EPA implemented a project to collect data on lakes in the Kenai Peninsula of Alaska, an area expected to receive low levels of acidic deposition.

The data contained on the accompanying tape or floppy diskettes were collected during the Kenai Lakes Investigation Project (KLIP) in the late summer of 1988 (Eilers et al., in press). Like other components of the NSWS, KLIP is based on a probability sample from an explicitly defined population of surface waters. Data were collected during fall overturn, considered to be a representative "index" period. Sample information was then extrapolated to represent the target lake population in Alaska's Kenai Peninsula.

This database guide provides a brief overview of the survey and the KLIP database. Detailed information on KLIP results is found in Eilers et al. (in press). This document also summarizes the sampling (Section 2.1) and analytical methods (Section 2.3), sources of geographic information (Section 2.5), and precision and accuracy results from quality assurance (QA) analysis (Section 2.4). The datasets are described in Section 3 and their formats in Section 6. The variables are defined in Section 5, and Appendix A contains a list of the lakes and their chemistry. Appendix B provides reference values taken from the Long Range Transport of Airborne Pollutants (LRTAP) Project audit samples.

The results of this study showed that over 90% of the target population lakes are seepage systems, having no surficial inlets or outlets and only a small direct terrestrial drainage area. The median lake surface area was 8.3 ha and the median watershed area was 79.3 ha. Watershed processes probably have a minimal influence on these systems, except in cases of extreme disturbance. In turn, these systems are expected to be highly responsive to groundwater inputs. When such inputs are small, these systems are likely to be responsive to atmospheric inputs of acidic materials due to lack of base cations from the weathering of soils and tills.

1.2 OBJECTIVES

From a regional perspective, little is known about the chemical character of Alaskan lakes, and the KLIP survey provided an opportunity to collect such information in one part of Alaska. We selected the Kenai Peninsula lakes in part because the surficial morphology of the region is similar to that of the midwest, yet the Kenai receives considerably less acidic deposition. The area contains many seepage lakes in the relatively flat topography characteristic of glacial till deposits. We hoped that these lakes would provide a good comparison to the morphologically similar but moderately impacted Upper Midwest lakes. In addition, we were curious about the possible impacts from the Nikiski Industrial complex, located on the northwest corner of the area studied, and thus have included the distance from each lake to this area. Results from this survey are available in Eilers et al. (in press).

1.3 KLIP TARGET POPULATION

KLIP focused on a population of 902 lakes portrayed on 1:250,000-scale U.S. Geological Survey (USGS) topographic maps of the Kenai Peninsula in an area delineated by the eastern and southern extent of glacial deposits on the lowland area of the peninsula (Figure 1). The sampling frame was generated by assigning each lake a unique identification number. Lakes were numbered consecutively, starting in the northwest corner of the peninsula and working east. This procedure was repeated from north to south until all lakes were numbered.

A subsample of over 60 lakes was then systematically selected for field visitation, using a random start. Once lakes were identified for sampling, detailed geographic data were recorded for each selected lake (watershed area, lake area, proximity to roads, etc.) and used to determine access for sampling. During pre-sampling field reconnaissance, four of the selected lakes were determined to be dry, and one lake had nesting swans. During sampling, two sites were found not to be lakes, and one other had nesting swans. We agreed not to sample lakes with nesting
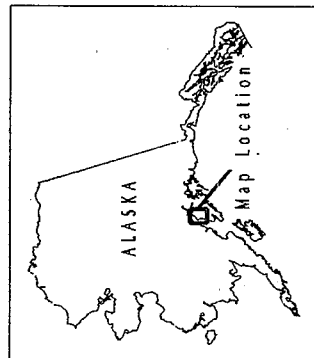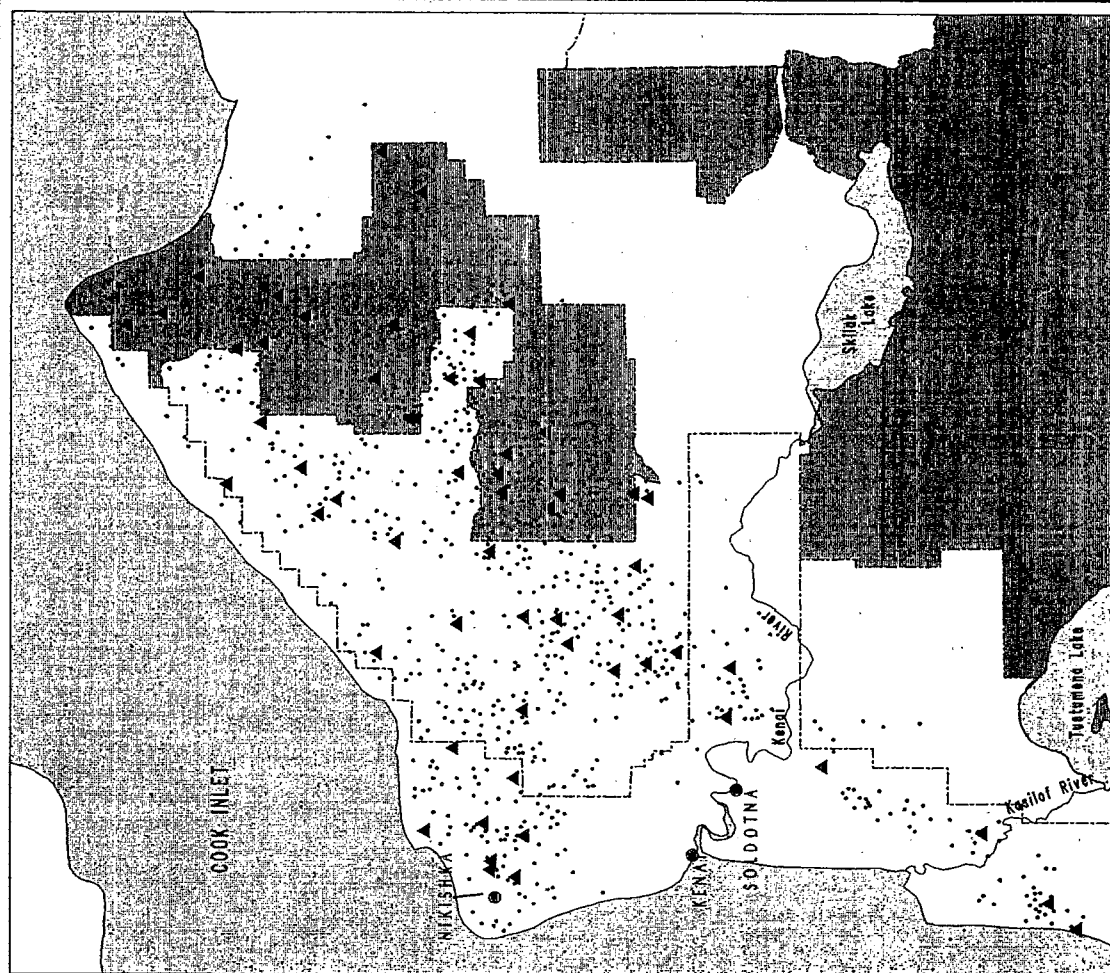
Figure 1. A map of the Kenai Peninsula showing sampled and target lakes, and Wildlife Refuge and Wilderness boundaries.

swans, following U.S. Fish and Wildlife management policy regarding disturbance of the nest sites. Thus, 59 lakes were successfully sampled, representing a target population of 819 lakes (59 lakes × 13.8769, the WEIGHT for each lake). The target population of lakes characterized by KLIP is considered to be only those represented on 1:250,000-scale topographic maps that do not have nesting swans. KLIP data should not be used to infer conditions of lakes either not depicted on these maps or outside of the sampling area.

Unlike the other regions studied by the NSWS (Linthurst et al., 1986; Landers et al., 1987), a stratified sampling design was not used in KLIP. Few existing lake chemistry data were available for the delineation of strata on the Kenai Peninsula, so the area was sampled as a single stratum. This procedure has simplified data summary calculations, as each observation in KLIP represents 13.88 lakes in the target population, as indicated by the weighting factor (the variable WEIGHT in the datset).

2. DATA COLLECTION

2.1 SAMPLE COLLECTION

Field crews sampled 59 lakes during August of 1988. Lakes were accessed via float plane, helicopter, or canoe. Secchi disk measurements were made, followed by a water temperature profile to determine if the lake was thermally stratified. A Van Dorn[TM] sampler was used to collect a 6-L water sample at a depth of 1.5 m at the deepest point in the lake. From this, a 4.5-L Cubitainer[TM] was filled and a 60-cc polypropylene syringe sample was withdrawn. Both were stored on ice until arrival at the processing laboratory. Duplicates, collected as part of the quality assurance (QA) program, consisted of field splits; two Cubitainers[TM] and two pH syringes were filled from a single Van Dorn[TM] sample. National Surface Water Survey field methods were followed (Chaloud et al., 1987).

2.2 PROCESSING LABORATORY

A processing laboratory was set up in the U.S. Fish and Wildlife facility to stabilize water samples before shipment to the analytical laboratory. All samples were processed within 6-12 hours of collection. The samples were processed into 0.45-μm filtered aliquots for cation, anion, and dissolved organic carbon (DOC) analysis. An unfiltered aliquot was prepared for acid neutralizing capacity (ANC) and laboratory conductance measurements. Cation and DOC aliquots were preserved with ultra-pure HNO_3 and H_2SO_4, respectively. All of these aliquots were shipped to

the EPA Region X laboratory in Manchester, Washington, via overnight courier, where they were later analyzed. Samples were stored on ice, or refrigerated at 4°C at all times, including shipping, with the exception of sample processing.

Measurements of specific conductance, color, and closed-cell pH were made at the processing laboratory, using methods developed as part of the NSWS (U.S. EPA, 1987). Color was determined on filtered samples using a visual comparator. Conductance was measured on an unfiltered aliquot. Syringes for closed-cell pH samples were filled directly from the Van Dorn^{TM} sampler in the field; pH was measured in the processing laboratory, on samples equilibrated to room temperature. Replicate values were averaged to obtain PHSTVL, the variable for field pH in this database.

2.3 ANALYTICAL METHODS

Inductively coupled plasma (ICP) spectroscopy was used to analyze for Ca, Mg, Na, Al, Fe, Mn, and Si. Potassium was analyzed using atomic absorption spectroscopy. Sulfate, NO_3^-, and Cl$^-$ were analyzed using ion chromatography. DOC was determined by infra-red (IR) spectroscopy after persulfate oxidation on a Dohrman^{TM} carbon analyzer. Dissolved P was determined on an autoanalyzer using ammonia-persulfate digestion followed by single-reagent ascorbic acid colorimetry (U.S. EPA, 1979; Method 365.2). Dissolved ammonia-nitrogen was determined by an automated colorimetric phenate method (U.S. EPA, 1979; Method 350.1).

ANC was determined by a double endpoint titration. Although this method is generally not considered as accurate as the Gran titration method that was used in other NSWS studies, the results of three audit samples ranging from 50 to 814 μeq L^{-1} ANC were comparable to Gran titration reference values for these audit samples.

pH was measured in the analytical laboratory. The field pH value, PHSTVL, results in better ion balances and calculated conductance values. For KLIP, the PHSTVL measurements were made within 12 hours of collection, whereas the laboratory pH values (PH_LAB) were measured after shipping; holding times on these ranged up to three months.

Both laboratory (U.S. EPA Manchester Laboratory) and field conductance (measured in the field laboratory) values are available. Based on the results of 9 audit samples, 10 duplicate lake samples, and calculated versus measured conductance for the lake samples, the laboratory conductance (CONDLAB) values appear to be both more accurate and more precise than field conductance (CONDFLD) measurements.

2.4   QUALITY ASSURANCE

Three types of QA samples were used during KLIP. These included 9 audit samples for assessing bias, obtained from the Long Range Transport of Airborne Pollutants (LRTAP) Interlaboratory Study in Canada (Aspila, 1989), duplicate samples from 11 of the 59 lakes sampled for assessing precision, and 12 deionized water blanks to identify possible occurrences of contamination. Of the 12 blank samples, 8 Cubitainers[TM] were filled in the field and treated as samples upon return to the laboratory; these were considered field blanks. The remaining four were filtered in the same way as lake samples and field blank samples, but had originated in the processing laboratory. Results for these three types of QA samples are presented in the tables included with the discussion in the following sections.

The following tables present average percent bias and precision (as relative standard deviation, which is the same as the coefficient of variation). These tables are intended only as a broad summary of data quality. As indicated in Section 3.2, the results for all of the QA samples are included in Dataset 3, and the reference values for the audit samples are included in Appendix B, so that more detailed analysis of the QA data by any user is possible.

2.4.1   Bias

Nine audit samples were processed and analyzed during KLIP. The samples were provided by the LRTAP Interlaboratory Study in Canada (Aspila, 1989, study number 16). These LRTAP samples, consisting of natural lake, stream, or rain water, had been sent to over 50 laboratories for analysis. The median values of this large number of analyses were then used as reference values to estimate percent bias. Quality assurance objectives (QAOs) were defined before the project was begun. For bias, these objectives were set to $\pm$ 10% for all variables, with the exception of conductance, which was set at $\pm$ 5%. The average percent bias is shown in Table 1. Not all of the LRTAP sample concentrations corresponded to the concentration ranges of the KLIP lake samples; only the audit samples that bounded the KLIP lake concentration range are included in Table 1.

The worst performance was for K, for which 5 of 9 audit samples were slightly high. The K concentrations of both the audit samples (< 23 μeq L^{-1}) and KLIP lakes (< 37 μeq L^{-1}) were quite low, however, and the absolute differences between audit values and laboratory results for all samples exceeding the QAO were < 3 μeq L^{-1}. This performance is within the analytical precision for K, and is thus an acceptable result, so the QAO is overly stringent.

Table 1. Average Percent Bias from Up to Nine Audit Samples Processed in the Field Laboratory and Sent on to the Analytical Laboratory[a]

Variable	n	Average % Bias	Minimum and Maximum % Bias	Concentration Range of Audit Samples	Concentration Range of Audit Samples
ANC	3	6.5%	−17.2 — −0.8%	50—814 μeq L^{-1}	20—1738 μeq L^{-1}
pH	4	−2.6%	−5.2 — −.2%	5.56—7.75 pH units	5.53—8.88 pH units
Conductance	7	1.85%	−0.5—8.5%	6.0—94 μS cm^{-1}	8.3—175 μS cm^{-1}
Ca	9	2.8%	0—6.7%	15—2138 μeq L^{-1}	11—1367 μeq L^{-1}
Mg	9	4.5%	0—6.7%	4—794 μeq L^{-1}	17.5—389 μeq L^{-1}
K	9	18.4%	−2.1—40.6%	3.5—23 μeq L^{-1}	2—37 μeq L^{-1}
Na	8	−3.3%	−12.9—5.9%	13—264 μeq L^{-1}	24—219 μeq L^{-1}
SO_4^{2-}	9	−4.2%	−14.5—4.7%	8—782 μeq L^{-1}	1[b]—45 μeq L^{-1}
Cl^-	8	2.0%	−2.8—8%	6—306 μeq L^{-1}	23—189 μeq L^{-1}
NO_3^-	6	6.7%	−44—82%	1—15 μeq L^{-1}	0.4[b]—3.4 μeq L^{-1}
NH_3	9	−32%	−72—3.4%	0.4—25 μeq L^{-1}	0.4[b]—6.2 μeq L^{-1}
Color PCU	5	138%	20—400%	1—110 PCU	5—100 PCU
DOC	9	6%	−22—35%	0.1—11 mg L^{-1}	2.3—17.5 mg L^{-1}

[a] Also shown on the table are the minimum and maximum percent bias, the concentration range of the audit samples, and the concentration range of the Kenai Lake samples.

[b] Minimum concentration is the analytical detection limit.

[c] The %bias values that exceed the quality assurance objective (>10%) for these samples represent only a 1 μeq L^{-1} absolute difference between the audit result and the reference value.

In addition to the calculation of percent bias to determine if bias was present for any of the variables in the KLIP dataset, the Wilcoxin signed rank test (Hogg and Tanis, 1983) was applied to see if there was a consistent deviation from the median value for any the variables. These results showed no consistent bias for any variables.

2.4.2 Precision

Of the 59 lakes sampled, 10 duplicate samples were taken, with 1 or 2 taken on each day of sampling. Precision values are determined by calculating the relative percent standard deviation (%RSD), which is the same as the coefficient of variation, and is equal to the standard deviation divided by the sample mean, and multiplied by 100. These results are presented in Table 2. The QAOs for precision were set at 5% RSD for most variables, with the exception of ANC (10% RSD), NO_3^- (10% RSD), and DOC at concentrations < 5 mg L^{-1} (10% RSD). The precision criterion for color was an absolute difference of 5 PCU, and was not based on relative standard deviation.

The three ANC duplicate pairs that exceeded the QAO were at low ANC concentrations, and reflected acceptable absolute differences of 10—15 μeq L^{-1} (26.5% RSD at 31 μeq L^{-1} ANC; 15.7% RSD at 36 μeq L^{-1}; and 13% RSD at 44 μeq L^{-1}). The two pairs of Ca duplicates that exceeded the QAO were also at low concentration and reflected < 5 μeq L^{-1} absolute difference within pairs (10.8% RSD at 20 μeq L^{-1} Ca, and 12.5% RSD at 17 μeq L^{-1} Ca). Thus, although some duplicate pairs for ANC and Ca exceeded their QAO, the QAO was inaccurate in defining an acceptable quality for these variables at these levels.

The two pairs of NH_3 and Mn exceeding the precision goals were at very low concentrations (< 1 μeq L^{-1} NH_3 and < 10 μg L^{-1} Mn). Two sets of duplicate values for Fe had poor precision (47% RSD at 10 μg L^{-1}, and 20% RSD at 50 μg L^{-1}). Three duplicate pairs of color were outside the acceptable limits; however, color values were generally very low (< 100 PCU); 75% of the color values were ≤ 30 PCU, the definition of low color in the NSWS. Color was the variable of least interest in the KLIP Project.

NO_3^- precision was poor, but difficult to quantify, as all three pairs with detectable NO_3^- had high %RSD values. Three of the DOC pairs differed by as much as 2.5 mg L^{-1}, and exceeded the precision goal of 10% RSD.

2.4.3 Detection Limits and Variable Tags

The analytical laboratory determined the analytical detection limits (Table 3), and did not report values for results that were below these levels. Observations with low concentrations were

Table 2. Average Percent Relative Standard Deviation (RSD) of Duplicate Samples for the Major Analytes of the Kenai Lakes Investigation Project[a]

Variable	n	Average % Bias	Minimum and Maximum % Bias	Concentration Range of Audit Samples	Concentration Range of Audit Samples
ANC	10	6.5%	0—26.5%	20—1188 μeq L^{-1}	20—1738 μeq L^{-1}
pH	10	0.6%	0—3.4%	5.53—7.96 pH units	5.53—8.88 pH units
Conductance	10	0.2%	0—1.0%	9.5—126 μS cm^{-1}	8.3—175 μS cm^{-1}
Ca	10	3.4%	0—12.5%	17—895 μeq L^{-1}	11—1367 μeq L^{-1}
Mg	10	0.2%	0—0.7%	21—305 μeq L^{-1}	17.5—389 μeq L^{-1}
K	10	2.5%	0.1—8.7%	8.4—34 μeq L^{-1}	2—37 μeq L^{-1}
Na	10	1.3%	0.5—3.3%	25—219 μeq L^{-1}	24—219 μeq L^{-1}
SO$_4^{2-}$	10	2.8%	0—8.3%	1.0—46 μeq L^{-1}	1[b]—45 μeq L^{-1}
Cl$^-$	10	1.4%	0—4.3%	22—189 μeq L^{-1}	23—189 μeq L^{-1}
NO$_3^-$	10	30%	0—133%	0.4—6 μeq L^{-1}	0.4[b]—3.4 μeq L^{-1}
NH$_3$	10	8.5%	0—33%	0.4—6 μeq L^{-1}	0.4[b]—6.2 μeq L^{-1}
Color PCU	10	12.4%	0—47%	5—125 PCU	5—100 PCU
DOC	10	7%	0.4—22%	3.7—17.5 mg L^{-1}	2.3—17.5 mg L^{-1}
Al	10	1%	0—6.4%	25—231 μg L^{-1}	25[b]—231 μg L^{-1}
Fe	10	9.4%	0—47.1%	7—521 μg L^{-1}	7[b]—830 μg L^{-1}
Mn	10	11.3%	0—78.5%	2—37 μg L^{-1}	2[b]—208 μg L^{-1}
Si	10	0.3%	0—2.7%	40—2395 μg L^{-1}	40[b]—8670 μg L^{-1}

[a] Also shown on the table are the range of %RSD observed among 10 duplicate pairs, the concentration range of the duplicate samples, and the concentration range of the Kenai Lake samples.
[b] Minimum concentration is the analytical detection limit.

Table 3.   Analytical Detection Limits for Appropriate Variables

Variable	Analytical Detection Limit
Ca (μeq L^{-1})	0.05
Mg (μeq L^{-1})	0.08
Na (μeq L^{-1})	0.9
K (μeq L^{-1})	1.5
SO_4^{2-} (μeq L^{-1})	1.0
Cl$^-$ (μeq L^{-1}	2.8
NO_3^- (μeq L^{-1})	0.4
NH_3 (μeq L^{-1})	0.4
P (μg L^{-1})	7.0
DOC (mg L^{-1})	0.1
Si (μg L^{-1})	40
Al (μg L^{-1})	25
Fe (μg L^{-1})	7
Mn (μg L^{-1})	2

reported as the detection limit value and tagged; thus, no zero or negative values occur in the dataset. Tag values assigned by the analytical laboratory for each chemical variable appear in a separate variable, and are used to qualify the values for the chemical variables. Two tag values were employed, 'U' and 'M'. The 'U' value indicates that the sample was analyzed, but no analyte was detected in the sample. A tag value of 'M' indicates that the analyte was detected, but not quantified; in other words, the analyte occurred at a trace amount.

2.4.4 Blanks

Two kinds of blanks were included in the KLIP project; deionized water was taken into the field and used to fill a Cubitainer^{TM} each day. These were not put through the Van Dorn^{TM} water sampler, as there was no means of carrying enough water into the field to rinse the sampler prior to taking a blank sample. Blanks were also filtered in the laboratory, without having been processed in the field. The results of all blank samples for Mg, Na, Si, Al, Mn, Fe, SO_4^{2-}, NO_3^-, and Cl^- were at the detection limit values for these variables. Results for the remaining variables are shown in Table 4. None of the results for these blanks were high enough to be of concern. The blank results for Ca, K, NH_3, and DOC were very close to the detection limit values, and far below values for the KLIP lakes.

Table 4.   Results of Analyses on Field and Filtered Blanks for Those Variables with Blank Results Greater than the Detection Limit for the Corresponding Analyte

	Mean	Standard Deviation	Number of Blanks	Detection Limit
ANC (μeq L^{-1})	5.99	6.38	11	n.a.[a]
Conductance (μS cm^{-1})	1.08	0.09	11	n.a.[a]
Ca (μeq L^{-1})	0.34	0.35	11	0.05
K (μeq L^{-1})	1.43	0.29	11	0.77
NH$_3$ (μeq L^{-1})	0.37	0.06	12	0.36
DOC (mg L^{-1})	0.14	0.04	12	0.1

[a] No detection limits are available for these variables.

2.4.5 Implications for Data Interpretation

Evaluation of the QA data from the KLIP Project indicates that only the duplicate values for NO_3^- were of questionable quality. Values for NO_3^- were below detection limit for so many of the KLIP lake samples that it is difficult to make many inferences concerning NO_3^-, thus the poor precision is of little concern. Percent bias for NO_3^- was within acceptable limits; so although the data may not be precise, they are not biased. Color data for KLIP were not good, but are generally of little interest due to the subjective nature of obtaining color data.

As demonstrated in this text, the KLIP data are of very high quality. Quality assurance samples that failed to meet their QAOs were generally of an acceptable level of quality to meet the study objectives. Inaccurate selection of QAOs at variable concentration levels resulted in the identification of some precision and audit data that did not meet the QA objectives, but were in fact at acceptable performance levels.

2.5 GEOGRAPHIC VARIABLES

Several geographic variables occur in the database and were obtained from 1:25,000-scale topographic maps of the region: latitude, longitude, watershed area, lake area, lake name, lake elevation, and lake hydrologic type [seepage (no inlets or outlets) or drainage (outlet present)]. Measurements of lake area, watershed area, distance to the coast, and distance to the refineries at Nikiski were determined by electronic planimetry.

3. DATABASE DEVELOPMENT

3.1 TRANSFER MEDIA

The database exists in two formats: as SAS transport datasets (SAS, 1988; version 6.04), and as ASCII files on 5¼-inch, low-density (360 Kb) floppy diskettes. The floppies contain (1) two datasets: datasets 3 and 4, KLIPDS3 and KLIPDS4, described in Section 3.2, and (2) the database descriptions in two files: CONTENTS.LIS, a listing of the variable names and definitions, and FORMAT.TXT, a description of the ASCII file columnar format. These are further described in Section 6. The SAS transport datasets have been created using PROC CPORT, and can be transferred back into SAS datasets using PROC CIMPORT (SAS, 1988).

3.2 DATASETS

One of the two KLIP datasets, KLIPDS3, includes all data collected during the KLIP project: information on the 62 lakes that were visited out of the 67 lakes selected for sampling, and all QA samples, such as audits and duplicates. The five lakes determined to be unsuitable for sampling during pre-sampling reconnaissance are not included. The other dataset, KLIPDS4, includes information only for the 59 sampled lakes used to characterize the KLIP target population. Dataset 4 is the easiest to use for characterizing the Kenai Lakes or for analyzing inter-variable relationships. Dataset 3 documents the results of all analyses performed for KILP, and may be used to characterize the data quality. Further differences between the two datasets are described in the following sections.

3.2.1 Dataset 3 (KLIPDS3)

Dataset 3 (KLIPDS3) contains verified data and includes QA data in the form of audit results and analysis of the laboratory (filter) and field blanks and the duplicate pairs. Verified data indicate that the numbers from the field forms and analytical laboratory were properly transferred to the database. All other possible errors, such as analytical errors or recording errors, are included in Dataset 3. Data found to be in error through validation are indicated as such in the comment field of KLIPDS3, but the original values remain. Validation was performed by examining ion balances, comparing calculated to measured conductance, and examining relationships between variables for outliers in much the same way as described in U.S. EPA (1992).

Field blanks are identified as such only in the sample code (SAMPCODE) variable. These samples have LAKE_ID and LAKENAME values that indicate the site at which the field blanks were taken. Filter blank samples are indicated as 'BLANK' in LAKENAME, and as 'BFIL' in SAMPCODE. Two SPLIT samples are included in KLIPDS3. These were part of a comparison for another study, and are not indicative of KLIP performance.

3.2.2 Dataset 4 (KLIPDS4)

Dataset 4 (KLIPDS4) contains only data from the lakes successfully sampled. This is the dataset used for regional characterization of the Kenai Lakes. Duplicate values have been averaged, and extreme outliers, identified during the validation process (Section 3.2.1), have been omitted or replaced. Of the three values found to be in error, two were from duplicate samples. Only the acceptable duplicate was included in KLIPDS4, not an average of the two duplicate

values. The third, an ANC value, was estimated from the relationship between ANC and the sum of nonmarine base cations minus nonmarine sulfate. The error about this relationship was used to estimate a random error component for the datum and incorporated into the estimate (Little and Rubin, 1987).

3.3 MISSING VALUES

Missing values are represented by the number –999.000 for the card-image ASCII format. Standard SAS notation for missing values, a "." for numeric values and blank space for character values, is used in all SAS files. The only missing values in KLIPDS4 occur in the variables WSHED (watershed area), LAKESIZ (lake area), and WTEMPB (water temperature 1.5 m above the lake bottom), and for the various TAG variables.

3.4 CALCULATED VARIABLES

A few variables were calculated as part of the validation process and remain in the final datasets. These are hydrogen and hydroxyl ion concentrations, the sum of anions, the sum of cations, the anion deficit, the calculated conductance, and the organic ion present as calculated from the Oliver model (Oliver et al., 1983). The equations used to calculate these variables are summarized in Section 5.

3.5 FLAGS AND TAGS

Qualifiers from the analytical laboratory are indicated in the tag variables for each analyte. Values that were determined to be in error through charge balances or other relationships among variables are identified in the comment field. Original values are included in Dataset 3, and values that are representative of the Kenai Lowland lake population (described in Section 3.2.2) are included in Dataset 4. The tag values are defined in Section 2.4.3.

4. TARGET POPULATION OF LAKES

4.1 DESCRIBING THE TARGET POPULATION

The KLIP database structure facilitates examination of the target population in a variety of ways. Estimates of subpopulations of lakes can readily be calculated and the distribution of the

entire target population of systems can be examined. For example, if one is interested in examining information only for seepage lakes, the data can be subset using the appropriate variable (in this case, LAKE_TYP='S'). The number of lakes described by each sampled lake is calculated by multiplying by the weighting factor (WEIGHT). Since KLIP was not based on a stratified sampling design, each lake is equally weighted (e.g., the variable WEIGHT is the same for each lake).

4.2 CUMULATIVE DISTRIBUTION FREQUENCY

Population distributions can be generated for measured attributes using the general estimators (e.g., number of lakes, watershed area, elevation) and are presented using the cumulative distribution function, $F(x)$, sometimes called a cumulative proportion. The distribution, $F(x)$, shown in Figure 2, is interpreted as the proportion of target lakes in the population having the attribute $X \leq x$. To read this figure, select a value of x of the attribute X, along the horizontal axis (in this example, ANC) and read the Y-axis value, $F(x)$, the cumulative proportion of lakes at this value. In this example, to determine the proportion of lakes with ANC $\leq$ 200 μeq L^{-1}, locate the ANC value of 200 along the X-axis, then read the corresponding $\hat{F}(x)$ as 0.7. The estimated number of lakes ($\hat{N}$) with ANC $\leq$ 200 μeq L^{-1} can be calculated by multiplying $\hat{F}$ by $\hat{N}_{total}$ (the total number of lakes in the population of target lakes). For our example, this is (0.7)(819) = 573 estimated lakes with ANC $\leq$ 200 μeq L^{-1}. We can also choose a value of $F(x)$, such as the median at $F(x) = 0.5$, and read off the estimated x value (in this case, 93 μeq L^{-1}). Thus we can estimate the median (or other percentile) ANC value.

Distributions are sometimes presented as an inverse or descending cumulative proportion, $1-F(x)$. For these, the distribution is read as the estimated proportion of lakes having values $\geq$ x ($X \geq x$). For the lake surveys, the inverse cumulative proportions have been used to describe distributions where there is more interest in the maximum values than the minimum values. For example, because SO_4^{2-} can be indicative of acidic inputs, we were more interested in what proportion of the population had SO_4^{2-} above a certain value. Similarly, higher elevation lakes were expected to have greater inputs of atmospheric acids, so we were more interested in estimating the proportion of lakes found at greater than a given elevation, rather than the proportion with elevation below a given value.

5. VARIABLE DEFINITIONS

Following is an alphabetical list of all the variables in the KLIP dataset, and their definitions, including units used for measure, and a brief description of how the measure was made. The
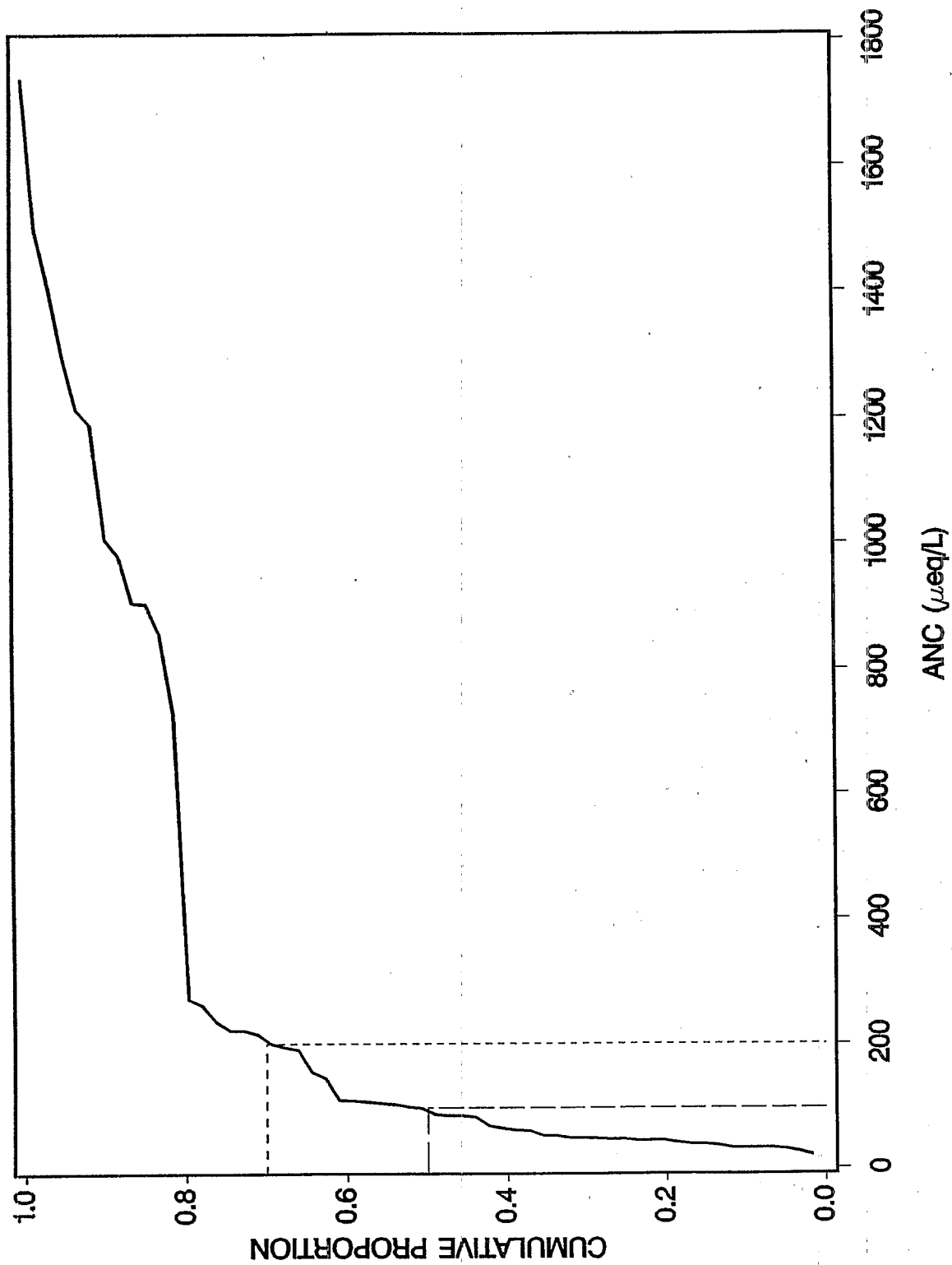
Figure 2. A cumulative frequency distribution for ANC in the Kenai Lakes population. The dotted line demonstrates how to determine the proportion of lakes with ANC below 200 μeq L^{-1}, and the dashed line depicts the determination of the median ANC value, both as described in the text.

TAG variables indicate whether the variable value is accompanied by 'U' or 'M', which are defined in Section 2.4.3.

AL	$\mu g\ L^{-1}$	Total dissolved aluminum measured using inductively coupled plasma, with a detection limit of 25 $\mu g\ L^{-1}$.
ALK_TOT	$\mu eq\ L^{-1}$	The ANC of the water measured by double endpoint titration.
AL_T		Analytical laboratory assigned tag for Al (U or M).
ANDEF	$\mu eq\ L^{-1}$	Anion deficit calculated as cations minus anions, using the equation:

$$ANDEF = CATSUM - ANSUM.$$

ANSUM	$\mu eq\ L^{-1}$	The sum of anions calculated as follows:

$$ANSUM = SO4 + NO3 + CL + SCO3 + OHMINUS,$$

where $\quad OHMINUS = (10^{pH-14})(10^6)$.

CA	$\mu eq\ L^{-1}$	Dissolved calcium measured by inductively coupled plasma. The detection limit was 0.05 $\mu eq\ L^{-1}$.
CA_T		Analytical laboratory assigned tag for Ca (U or M).
CATSUM	$\mu eq\ L^{-1}$	The sum of cations calculated as follows:

$$CATSUM = CA + MG + Na + K + NH4ION + HPLUS,$$

where $\quad NH4ION = NH3(10^{-pH})/[5.012 \times 10^{-10} + (10^{-pH})]$.

CL	$\mu eq\ L^{-1}$	Dissolved chloride concentration measured by ion chromatography. The detection limit was 2.8 $\mu eq\ L^{-1}$.
CL_T		Analytical laboratory assigned tag for Cl (U or M).
COLVAL	PCU	The true color measured with a visual comparator, in platinum-cobalt units (PCU).
COMMENT		A 40-character string for notations concerning each observation, this variable has values only in Dataset 3 (KLIPDS3).

CONDCAL $\mu M\ cm^{-1}$ Conductance calculated from the measures of the major ions and their respective contributions to conductance as follows:

CONDCAL = [(CA × 59.47) + (MG × 53.0) + (NA × 50.08) + (K × 73.48) + (NH4ION × 73.5) + (HPLUS × 349.65) + (SO4 × 80.0) + (SCO3 × 44.5) + (Cl × 76.31) + (NO3 × 71.42) + (OHMINUS × 198)]/1000,

where   NH4ION = NH3(10^{-pH})/[5.012 × 10^{-10} + (10^{-pH})]
OHMINUS = (10^{pH-14})(10^6).

CONDFLD $\mu M\ cm^{-1}$ Specific conductance measured in the field laboratory using a Fisher portable conductance meter, on samples that had equilibrated to room temperature. QA data suggest that CONDLAB was a more accurate measure of specific conductance.

CONDLAB $\mu M\ cm^{-1}$ Specific conductance measured in the analytical laboratory, also measured at room temperature. QA data suggest that this measure was more accurate than the field conductance value.

DATE   DDMMMYY   The date on which the sample was taken from the lake, in the format day, month, year (i.e., 18AUG88).

DIS_REF   km   The distance from the lake to the closest refinery, measured from 1:25,000-scale USGS topographic maps.

DOC   mg C L^{-1} Dissolved organic carbon, measured by IR spectroscopy after persulfate oxidation on a Dohrman carbon analyzer. Detection limit was 0.1 mg L^{-1}.

DOC_T   Analytical laboratory assigned tag for DOC.

DS_COAST   km   The shortest distance from the lake to the coast, measured from 1:25,000-scale USGS topographic maps.

ELEV   m   The lake elevation above mean sea level, as determined from 1:25,000-scale USGS topographic maps.

FE $\mu g\ L^{-1}$ Dissolved iron, measured by inductively coupled plasma. The detection limit was 7 $\mu g\ L^{-1}$.

FE_T   Analytical laboratory assigned tag for Fe (U or M).

HPLUS μeq L^{-1} Hydrogen ion concentration determined from PHSTVL as follows:

$$HPLUS = (10^{-pH})(10^6).$$

K . μg L^{-1} Dissolved potassium measured by atomic absorption spectroscopy. Detection limit was 1.5 μg L^{-1}.

K_T Analytical laboratory assigned tag for K (U or M).

LAKENAME The lake name taken from 1:25,000-scale USGS topographic maps.

LAKE_ACC A code indicating how the lake was accessed, where:

HE = Helicopter
FW = Fixed wing aircraft
CA = Canoe.

LAKE_ID A seven-digit code used to identify the lake. All KLIP lakes start with '7A0-'. This convention is from the NSWS, where the region is represented by the first number, the subregion by the letter, and the stratum in the third column. Here, Alaska is Region 7, the Kenai Peninsula is subregion A, and, as the sample was not stratified, the stratum is always 0.

LAKE_SIZ ha The lake area in hectares as measured from 1:25,000-scale USGS topographic maps.

LAKE_TYP A variable indicating lake type, as estimated from 1:25,000-scale USGS topographic maps. There were 10 drainage lakes in the survey.

D = Drainage lakes, with inlets and outlets present.
S = Seepage lakes, where inlets and outlets were absent.

LAT_DMS DMS The latitude of the lake in degrees, minutes, and seconds, determined from 1:25,000-scale USGS topographic maps.

LONG_DMS DMS The longitude of the lake in degrees, minutes, and seconds, determined from 1:25,000-scale USGS topographic maps.

MG μeq L^{-1} Dissolved magnesium concentration measured by inductively coupled plasma. The detection limit was 0.08 μeq L^{-1}.

MG_T   Analytical laboratory assigned tag for Mg (U or M).

MN μg L^{-1} Dissolved manganese measured by inductively coupled plasma. The detection limit was 2 μg L^{-1}.

MN_T   Analytical laboratory assigned tag for Mn (U or M).

NA μeq L^{-1} Dissolved sodium measured by inductively coupled plasma. The detection limit was 0.87 μeq L^{-1}.

NA_T   Analytical laboratory assigned tag for Na (U or M).

NH3 μeq L^{-1} Dissolved ammonia-N measured by an auto-analyzer using the phenate colorimetric method. The detection limit was 0.36 μeq L^{-1}. Ammonium ion was estimated for charge balance using:

$$NH4ION = NH3(10^{-pH})/[5.012 \times 10^{-10} + (10^{-pH})].$$

This value differed little from NH$_3$ due to the circumneutral pH of the lakes.

NH3_T   Analytical laboratory assigned tag for NH$_3$ (U or M).

NO3 μeq L^{-1} Dissolved nitrate concentration measured by ion chromatography. The detection limit was 0.36 μeq L^{-1}.

NO3_T   Analytical laboratory assigned tag for NO$_3^-$ (U or M).

ORGION μeq L^{-1} The organic ion concentration, as calculated using the Oliver model (Oliver et al., 1983) from DOC and pH using the following equation:

$$ORGION = (KONST)(DOC)(10)/(KONST+10^{-pH}).$$

where $KONST = 10^{-(.96+.9(pH)-0.039(pH^2)}$.

P μg L^{-1} Dissolved phosphorus measured by single reagent ascorbic acid automated colorimetry, after ammonia-persulfate digestion. The detection limit of this method was 7 μg L^{-1}.

PHSTVL   Closed-cell pH, measured according to U.S. EPA (1987). Samples were taken in syringes directly from the Van Dorn sampler, and delivered to pH measurement cells from the same, without intro-

ducing air to the sample. This pH measurement is more accurate than the laboratory pH, as holding times for the laboratory pH measurements ranged up to three months.

PH_LAB

pH measured by the analytical laboratory, as late as three months after sample collection. These pH measurements were taken from sample bottles in open containers, and were thus subject to gas exchange.

P_T

Analytical laboratory assigned tag for P (U or M).

SAMPCODE

Indicates whether an observation is a Routine or QA sample. KLIPDS3 contains all samples—routine, duplicate, audits, blanks, and splits. KLIPDS4 contains only routine samples, and where duplicates were available, averages of the duplicate pairs.

R	=	Routine
D	=	Duplicate
BFIL	=	Filtered Blank
BFLD	=	Field Blank
A	=	Audit, unfiltered
AFIL	=	Audit, filtered
S	=	Split

SAMPLE_N

This is a unique identifier from the analytical laboratory. In cases where the data were averaged in KLIPDS4, the number refers to only one of the duplicate pair.

SCO3 μeq L^{-1}

The concentration of carbonate and bicarbonate estimated from the ANC and pH measurements as follows:

$$SCO3 = ALK\_TOT - OHMINUS + HPLUS,$$

where $OHMINUS = (10^{PHSTVL-14})(10^6).$

(When SCO3 < 0, SCO3 was set equal to 0.)

SECCHI   m

The Secchi disk depth, taken as the mean of depths of disappearance and reappearance.

SI μg L^{-1}

Dissolved silicon concentration measured by inductively coupled plasma, with a detection limit of 40 μg L^{-1}.

SI_T   Analytical laboratory assigned tag for Si (U or M).

SITE_DP   m   The depth of the lake at the sampling site in meters. An attempt was made to sample at the deepest part of the lake located by sonar, so this measurement serves as an estimate of maximum lake depth as well.

SO4 μeq L^{-1} Dissolved sulfate concentration measured by ion chromatography.

SO4_T   Analytical laboratory assigned tag for SO_4^{2-}.

VISITED   Not all of the selected lakes were sampled. Some were dry and others had nesting swans, so we were denied access. Location information about lakes Not Sampled, or Not Visited exists in KLIPDS3 but not in KLIPDS4, although the variable VISITED occurs in both datasets.

                              VS = lake was visited and sampled.
                              NV = lake was not visited. This occurred due to nesting swans on the lake. We refrained from sampling in order not to disturb the nests.
                              VNS = lake was visited but not sampled. This occurred twice, due to wetland areas portrayed as lakes on the map.

WEIGHT   This is the statistical weighting factor to be applied to each lake when population estimates are desired. As the sample was randomly drawn from a defined population, each lake sampled represents a specified proportion of the population. The sample was not stratified, so weighting factors for all lakes are the same.

WSAREA   ha   The watershed area of the lake, in hectares, as measured on 1:25,000-scale USGS topographic maps.

WTEMPB   °C   The water temperature measured 1.5 m above the bottom of the lake.

WTEMPS   °C   The water temperature measured at 1.5-m depth. A difference of > 4°C between surface and bottom temperatures indicated thermal stratification of the lake.

6. ASCII DATASET FORMAT

This section describes the files and information contained on the KLIP data diskette. The following files are included:

KLIPDS4.ASC   ASCII listing of KLIPDS4, suitable for input to computer program.

KLIPDS3.ASC   ASCII listing of KLIPDS3, suitable for input to computer program, same format as KLIPDS4.ASC.

FORMAT.TXT   Format description for KLIPDS4.ASC and KLIPDS3.ASC, suitable for translation to various programming languages.

CONTENTS.LIS   A list of variables and their SAS labels, which are short definitions of variables, including variable units.

KLIP.DAT   A SAS transport file (Version 6.04, created using PROC CPORT) of KLIP Dataset 3 (KLIPDS3.SSD) and Dataset 4 (KLIPDS4.SSD).

FORMAT.TXT, shown on the next page, describes the format for the KLIPDS3 and KLIPDS4 ASCII files. The format of this file is the one used to input the data into a SAS data file. However, the structure is set up for easy editing, so that the file can be input into other databases as well. The listing of FORMAT.TXT is followed by CONTENTS.LIS, a list of the variables included, with brief definitions, including the units for each variable in the dataset. This is the list that would be created by running a PROC CONTENTS on the SAS datasets. PROC CONTENTS is a SAS procedure that provides a description of the dataset, including the number of observations, and the listing of the variables included, along with short definitions included in the variable label.

6.1 FORMAT.TXT

```
Format description:

1)  Record Structure
Each record of the KLIP database spans 10 'cards' or input lines
in files KLIPDS3.ASC or KLIPDS4.ASC.
Input lines are referred to as 'CARD n' in the description below.

2)  Field Structure

@50     VARNAME    11.5
 |        |          |
 |        |          Field width (numeric), decimal width.
 |        Variable Name as listed in file CONTENTS.LIS.
Starting column on CARD n

@50     VARNAME    $10.
 |        |          |
 |        |          Field width (character).
 |        Variable Name as listed in  file CONTENTS.LIS.
Starting column on CARD n
```

3) Missing values are coded as −999 for numeric variables, and as blanks for character variables in the ASCII files.

4) Card Descriptions

```
CARD 1
@0 LAKE_ID $10.
@12 LAKENAME $25.
@39 AL 11.5
@52 AL_T $2.
@56 ALK_TOT 11.5
;
CARD 2
@0 ANDEF 11.5
@13 HPLUS 11.5
@26 ANSUM 11.5
@39 CA 11.5
@52 CA_T $2.
@56 CATSUM 11.5
;
CARD 3
@0 CL 11.5
@13 CL_T $2.
@17 COLVAL 11.5
@30 CONDCAL 11.5
@43 CONDFLD 11.5
@56 CONDLAB 11.5
;
CARD 4
@0 DATE DATE7.
@13 DIS_REF 11.5
@26 DOC 11.5
@39 DOC_T $2.
@43 DS_COAST 11.5
@56 ELEV 11.5
;

CARD 5
@0 FE 11.5
@13 FE_T $2.
@17 K 11.5
@30 K_T $2.
@34 LAKE_ACC $6.
@42 LAKE_TYP $8.
@52 LAREA 11.5;
;
CARD 6
@0 LAT_DMS $10.
@12 LONG_DMS $10.
@24 MG 11.5
@37 MG_T $2.
@41 MN 11.5
@54 MN_T $2.
@58 NA 11.5
@71 NA_T $2.
;
CARD 7
@0 NH3 11.5
@13 NH3_T $2.
@17 NO3 11.5
@30 NO3_T $2.
@34 ORGION 11.5
@47 P 11.5
@60 P_T $2.
;
```

```
CARD 8
@0 PHSTVL 11.5
@13 PH_LAB 11.5
@26 SAMPCODE $4.
@30 SAMPLE_N $10.
@42 SCO3 11.5
@55 SECCHI 11.5
;
CARD 9
@0 SI 11.5
@13 SI_T $2.
@17 SITE_DP 11.5
@30 SO4 11.5
@43 SO4_T $2.
@47 VISITED $3.
@52 WEIGHT 11.5
;
CARD 10
@0 WSAREA 11.5
@13 WTEMPB 11.5
@26 WTEMPS 11.5
@39 COMMENT $60.
;
```

6.2 CONTENTS.LIS

KLIP Variable List

NUMBER OF OBSERVATIONS: KLIPDS3 95; KLIPDS4 59   NUMBER OF VARIABLES: 62

ALPHABETIC LIST OF VARIABLES AND ATTRIBUTES

VARIABLE	TYPE	LENGTH	LABEL
AL	NUM	8	Total aluminum (μg L^{-1})
ALK_TOT	NUM	8	ANC (μeq L^{-1})
AL_T	CHAR	2	Aluminum tag
ANDEF	NUM	8	Anion deficit (μeq L^{-1})
ANSUM	NUM	8	Sum of anions (μeq L^{-1})
CA	NUM	8	Ca (μeq L^{-1})
CATSUM	NUM	8	Sum of cations (μeq L^{-1})
CA_T	CHAR	2	Calcium tag
CL	NUM	8	Cl (μeq L^{-1})
CL_T	CHAR	2	Chloride tag
COLVAL	NUM	8	True color
COMMENT	CHAR	60	Comment about observation
CONDCAL	NUM	8	Calculated conductance (μM cm^{-1})
CONDFLD	NUM	8	Field conductance (μM cm^{-1})
CONDLAB	NUM	8	Laboratory conductance (μM cm^{-1})
DATE	NUM	8	Sample Date

DIS_REF	NUM	8	Distance from refinery (km)
DOC	NUM	8	DOC (mg L^{-1})
DOC_T	CHAR	2	DOC tag
DS_COAST	NUM	8	Distance from coast (km)
ELEV	NUM	8	Lake elevation (m)
FE	NUM	8	Iron (μg L^{-1})
FE_T	CHAR	2	Iron tag
HPLUS	NUM	8	Hydrogen ion concentration (μeq L^{-1})
K	NUM	8	K (μeq L^{-1})
K_T	CHAR	2	Potassium tag
LAKENAME	CHAR	25	Lake name
LAKE_ACC	CHAR	6	Lake access method
LAKE_ID	CHAR	10	Lake identifier
LAKE_SIZ	NUM	8	Lake surface area (ha)
LAKE_TYP	CHAR	8	Lake hydrologic type
LAT_DMS	CHAR	10	Latitude (dms)
LONG_DMS	CHAR	10	Longitude (dms)
MG	NUM	8	Mg (μeq L^{-1})
MG_T	CHAR	2	Magnesium tag
MN	NUM	8	Manganese (μg L^{-1})
MN_T	CHAR	2	Manganese tag
NA	NUM	8	Na (μeq L^{-1})
NA_T	CHAR	2	Sodium tag
NH3	NUM	8	NH3 (μeq L^{-1})
NH3_T	CHAR	2	NH3 tag
NO3	NUM	8	NO3 (μeq L^{-1})
NO3_T	CHAR	2	NO3 tag
ORGION	NUM	8	Organic anion (μeq L^{-1})
P	NUM	8	Phosphorous (μg L^{-1})
PHSTVL	NUM	8	Closed-cell pH
PH_LAB	NUM	8	Laboratory pH
P_T	CHAR	2	Phosphorous tag
SAMPCODE	CHAR	4	Sample type
SAMPLE_N	CHAR	10	Sample number
SCO3	NUM	8	SCO3 calc w/phstvl (μeq L^{-1})
SECCHI	NUM	8	Secchi depth (m)
SI	NUM	8	Silicon (μg L^{-1})
SITE_DP	NUM	8	Lake depth at sample site (m)
SI_T	CHAR	2	Silicon tag
SO4	NUM	8	SO4 (μeq L^{-1})
SO4_T	CHAR	2	Sulfate tag
VISITED	CHAR	3	Lake visitation status

WEIGHT	NUM	8	Population weighting factor
WSAREA	NUM	8	Watershed area (ha)
WTEMPB	NUM	8	Temperature bottom (°C)
WTEMPS	NUM	8	Temperature surface (°C)

7. REFERENCES

Aspila, K.I., ed. 1989. *A Manual for Effective Interlaboratory Quality Assurance.* NWRI Contribution 89-99. National Water Research Institute, Canada Centre for Inland Waters, P.O. Box 5050, Burlington, Ontario, Canada L7R 4A6.

Chaloud, D.J., J.M. Nicholson, B.P. Baldigo, C.A. Hagley, and D.W. Sutton. 1987. *Handbook of Methods for Acid Deposition Studies: Field Methods for Surface Water Chemistry.* EPA/600/4-89/020. U.S. Environmental Protection Agency, Washington, D.C.

Eilers, J.M., D.H. Landers, A.D. Newell, M.E. Mitch, M. Morrison, and J. Ford. In press. Major ion chemistry of lakes on the Kenai Peninsula, Alaska. *Can. J. Fish. Aquat. Sci.*

Hogg, R.V., and E.A. Tanis. 1983. *Probability and Statistical Inference.* Macmillan Publishing Co., Inc., New York.

Kaufmann, P.R., A.T. Herlihy, J.W. Elwood, M.E. Mitch, W.S. Overton, M.J. Sale, J.J. Messer, K.A. Cougan, D.V. Peck, K.H. Reckhow, A.J. Kinney, S.J. Christie, D.D. Brown, C.A. Hagley, and H.I. Jager. 1988. *Chemical Characteristics of Streams in the Mid-Atlantic and Southeastern United States. Volume I: Population Descriptions and Physico-Chemical Relationships.* EPA/600/3-88/021a. U.S. Environmental Protection Agency, Washington, D.C. 397 pp.

Landers, D.H., J.M. Eilers, D.F. Brakke, W.S. Overton, R.D. Schonbrod, R.E. Crowe, R.A. Linthurst, J.M. Omernik, S.A. Teague, and E.P. Meier. 1987. *Characteristics of Lakes in the Western United States. Volume I: Population Descriptions and Physico-Chemical Relationships.* EPA-600/3-86/054a. U.S. Environmental Protection Agency, Washington, D.C. 136 pp.

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. Volume I: Population*

Descriptions and Physico-Chemical Relationships. EPA-600/4-86/007a. U.S. Environmental Protection Agency, Washington, D.C. 275 pp.

Little, R.J., and D.B. Rubin. 1987. *Statistical Analysis with Missing Data.* John Wiley & Sons, New York. 61 pp.

Oliver, B.G., E.M. Thurman, and R.L. Malcolm. 1983. The contribution of humic substances to the acidity of colored natural waters. *Geochim. Cosmochim. Acta* 47:2031-2035.

SAS. 1988. *SAS Language Guide for Personal Computers, Version 6.03.* SAS Institute, Inc., Box 8000, Cary, North Carolina 27511-8000. 559 pp.

U.S. Environmental Protection Agency (EPA). 1979. *Methods for Chemical Analysis of Water and Wastes.* EPA/600/4-79/020. Environmental Monitoring and Support Laboratory, Cincinnati, Ohio.

U.S. Environmental Protection Agency (EPA). 1987. *Handbook of Methods for Acid Deposition Studies. Laboratory Analyses for Surface Water Chemistry.* EPA/600/4-87/026. Office of Research and Development, Washington, D.C.

U.S.Environmental Protection Agency (EPA). 1992. *Data User's Guide to the United States Environmental Protection Agency's Long-Term Monitoring Project: Quality Assurance and Data Dictionary.* EPA/600/3-91/072. U.S. EPA Environmental Research Laboratory, Corvallis, OR.

APPENDIX A

KLIP SAMPLE LAKES

This appendix contains a list of the 59 sampled lakes that represent the Kenai target population, their location, and their major ion chemistry.

Major Ion Chemistry for Lakes on the Kenai Peninsula Sampled by the EPA in August 1988 (Page 1 of 2)

Lake Name	Lake ID	Latitude DMS	Longitude DMS	Area (ha)	Elev (m)	pH	Alkalinity (µeq/L)	Cond (µS)	Si (µg/L)	Ca	Mg	Na	K	NH$_4$	SO$_4$	Cl	NO$_3$	DOC (mg/L)
No Name	51	60 27 27	151 07 30	4.4	60	5.53	20	13.0	<40	21	25	39	25	6.2	8	47	3.4	13.2
Taiga	05	60 54 41	150 26 56	41.8	56	6.56	26	11.7	<40	17	27	34	14	<0.4	26	37	<0.4	2.1
No Name	36	60 43 27	150 30 10	7.9	83	6.31	30	7.1	<40	11	18	29	14	<0.4	3	30	<0.4	4.2
No Name	39	60 40 16	150 42 33	7.1	88	6.42	32	7.2	<40	20	18	24	11	<0.4	1	26	<0.4	4.6
No Name	54	60 34 17	150 56 29	3.5	67	5.92	32	7.6	<40	19	21	25	16	1.1	2	23	<0.4	13.9
No Name	20	60 44 50	150 24 44	5.2	83	6.55	32	7.9	<40	13	19	26	23	<0.4	1	26	<0.4	4.4
No name	46	60 41 46	151 18 14	3.9	34	6.14	32	11.9	<40	18	24	49	21	<0.4	5	54	<0.4	4.0
Birch Tree	10	60 51 34	150 23 54	43.3	67	6.64	36	8.7	<40	17	23	33	11	0.5	5	36	<0.4	3.7
No Name	49	60 36 13	150 48 09	20.9	88	6.41	38	10.6	<40	22	30	35	17	<0.4	6	32	<0.4	6.8
Moon	62	60 52 47	150 21 51	47.2	59	6.71	38	13.0	<40	27	33	47	15	<0.4	28	42	<0.4	3.6
Vixen	19	60 43 54	150 25 34	55.2	82	6.81	40	8.7	<40	17	24	30	14	<0.4	3	32	<0.4	3.4
No Name	37	60 42 04	150 22 46	19.4	71	6.64	44	10.8	<40	27	25	36	17	<0.4	1	36	<0.4	5.4
Pine	27	60 42 58	151 17 32	6.0	33	6.36	44	10.9	<40	21	26	54	11	<0.4	5	51	<0.4	4.4
Aqua Linda	58	60 19 54	151 13 44	6.2	35	6.63	44	11.4	<40	24	26	42	15	0.7	3	41	2.1	7.1
No Name	43	60 41 29	151 02 10	18.7	28	6.58	46	10.5	<40	26	27	48	2	0.6	1	46	<0.4	4.1
No Name	06	60 55 13	150 40 08	13.1	38	6.59	46	12.9	<40	23	36	53	12	<0.4	14	54	<0.4	5.4
No Name	12	60 50 57	150 43 06	4.2	54	5.95	47	17.5	53	62	63	67	8	0.6	1	55	<0.4	17.5
Porcupine	35	60 42 13	150 37 15	62.0	73	6.92	48	9.1	<40	24	25	31	17	<0.4	5	33	<0.4	2.8
Laura	28	60 42 58	151 17 32	15.5	34	6.62	48	12.8	<40	26	29	47	11	0.6	10	52	<0.4	4.8
Long	38	60 40 24	150 35 12	27.3	64	6.81	52	10.4	<40	28	29	34	20	<0.4	11	31	<0.4	5.4
No Name	22	60 44 28	150 39 05	3.5	54	6.96	52	14.3	120	56	51	46	14	<0.4	2	35	<0.4	13.2
No Name	04	60 56 27	150 19 56	7.3	48	6.60	60	20.7	<40	45	65	77	22	<0.4	25	74	<0.4	6.9
No Name	23	60 44 34	150 53 49	2.9	54	6.33	61	18.0	<40	49	64	69	18	<0.4	3	53	<0.4	11.8
No Name	08	60 53 29	150 26 23	23.5	53	6.63	64	14.7	<40	46	50	46	4	0.6	1	40	1.1	10.7
No Name	13	60 48 17	150 56 37	15.8	33	6.49	68	20.6	<40	61	51	88	11	1.1	2	82	1.1	12.2
No Name	59	60 16 49	151 20 26	4.9	63	6.80	82	20.8	<40	51	52	84	21	<0.4	9	74	<0.4	6.9
No Name	17	60 46 28	150 24 44	9.7	83	6.88	84	13.6	<40	51	41	44	18	0.6	1	30	<0.4	4.8
No Name	16	60 48 17	150 29 59	4.8	78	6.67	84	14.1	<40	47	44	42	21	<0.4	1	36	<0.4	5.1
No Name	48	60 37 00	150 52 53	2.9	98	6.77	86	16.5	<40	51	42	45	33	<0.4	1	42	<0.4	6.5
No Name	52	60 35 37	150 41 34	8.8	54	6.82	96	16.8	<40	83	57	47	9	<0.4	1	31	<0.4	11.1
No Name	55	60 31 57	151 02 46	2.2	59	6.89	98	18.8	<40	56	53	66	23	<0.4	1	58	<0.4	4.8
No Name	47	60 37 11	150 58 18	1.9	68	6.94	102	15.0	<40	47	44	38	31	<0.4	1	30	<0.4	9.9

Major Ion Chemistry for Lakes on the Kenai Peninsula Sampled by the EPA in August 1988 (Page 2 of 2)

Lake Name	Lake ID	Latitude DMS	Longitude DMS	Area (ha)	Elev (m)	pH	Alkalinity (µeq/L)	Cond (µS)	Si (µg/L)	Ca	Mg	Na	K	NH₄ (µeq/L)	SO₄	Cl	NO₃	DOC (mg/L)
No Name	07	60 53 29	150 34 07	5.2	62	6.45	104	23.2	800	110	83	66	11	<0.4	9	51	<0.4	13.8
No Name	63	61 00 26	150 22 01	11.5	34	6.96	106	30.4	<40	69	81	126	25	<0.4	23	125	<0.4	5.0
No Name	64	60 42 54	151 16 43	9.2	33	6.93	108	20.7	<40	66	52	79	17	0.9	7	62	<0.4	4.3
No Name	67	60 46 04	150 11 51	.	67	7.02	109	19.3	<40	86	54	64	14	0.6	1	41	1.2	9.7
No Name	31	60 41 27	150 53 06	11.4	44	6.99	144	24.0	<40	115	64	58	20	0.6	1	51	<0.4	9.8
No Name	65	60 50 08	150 04 42	5.1	62	7.17	154	25.8	124	121	80	72	11	0.6	3	50	2.0	10.0
No Name	03	60 58 08	150 23 28	12.1	41	7.20	190	30.6	<40	110	79	85	24	0.6	5	85	<0.4	7.3
No Name	60	60 15 30	151 23 01	5.8	68	6.63	194	41.8	<40	97	142	152	31	0.7	3	146	<0.4	10.7
Anertz	32	60 43 00	150 46 45	12.9	74	7.39	200	27.2	<40	147	76	53	16	<0.4	4	36	<0.4	5.5
No Name	41	60 39 53	150 53 22	3.8	68	7.01	216	31.6	<40	150	99	70	33	<0.4	6	46	<0.4	8.6
No Name	34	60 42 38	150 39 10	10.7	63	7.14	222	28.3	171	164	83	57	16	<0.4	1	30	<0.4	7.5
Derks	56	60 31 46	150 57 59	13.3	48	6.96	222	31.7	845	126	95	80	25	0.6	2	57	<0.4	8.9
No Name	53	60 35 42	150 57 35	3.3	63	7.16	236	29.1	122	110	96	88	24	0.6	1	41	<0.4	6.9
No Name	26	60 46 01	151 13 48	6.6	47	7.04	262	40.1	<40	152	110	147	26	<0.4	3	108	<0.4	7.3
No Name	11	60 51 46	150 38 36	35.6	53	7.27	272	36.8	238	207	110	72	10	0.7	4	49	<0.4	12.4
No Name	25	60 44 44	151 05 49	16.9	14	7.65	726	74.2	1270	403	196	179	25	<0.4	4	92	<0.4	6.8
No Name	45	60 41 23	151 14 17	9.1	29	7.14	855	91.2	4880	594	223	160	21	1.6	7	74	<0.4	4.4
No Name	42	60 39 22	150 55 43	3.7	54	7.35	903	93.0	6120	584	299	146	28	<0.4	1	46	<0.4	6.3
No Name	02	60 68 48	150 24 37	14.6	31	7.67	905	112.8	2395	609	301	219	33	0.5	45	189	<0.4	6.6
Beaver Pond	50	60 36 17	150 41 08	18.6	54	7.79	979	96.7	1760	753	237	104	18	0.6	6	24	<0.4	6.6
No Name	14	60 47 25	150 45 46	16.9	67	8.09	1007	98.6	2340	669	233	119	33	0.6	31	52	<0.4	5.4
No Name	33	60 42 27	150 40 59	10.2	58	7.95	1189	121.2	1540	896	305	146	34	<0.4	11	40	<0.4	5.6
Timber Lost	29	60 41 52	151 13 03	46.5	17	7.87	1213	118.4	8670	813	318	184	25	0.6	1	63	<0.4	5.3
Moose Pasture	18	60 47 55	150 07 52	22.0	58	8.88	1297	128.0	1630	999	309	112	15	0.6	37	61	<0.4	6.0
Duckbill	40	60 39 42	150 41 08	9.3	58	7.59	1405	138.6	1920	1078	353	147	35	0.7	3	50	<0.4	6.0
Hat	15	60 46 42	150 33 46	15.7	56	7.80	1500	146.0	4990	1173	346	141	37	<0.4	1	39	<0.4	5.5
Pan	21	60 46 26	150 03 44	4.1	56	7.70	1738	167.9	7590	1367	389	154	33	0.9	1	35	<0.4	6.4

APPENDIX B
REFERENCE VALUES FOR LRTAP SAMPLES

This appendix contains the reference values for the nine Long Range Transport of Airborne Pollutants Project audit samples from Study 16, used during the Kenai Lakes Investigation Project to assess accuracy of the data.

Reference Values for LRTAP Samples

	LRTAP Sample Number								
	04	05	10	01	03	08	07	06	09
ANC μeq L^{-1}	-43.4	-40.0	0.00	-1.4	110.9	50.7	0.8	2.0	814.8
Ca μeq L^{-1}	75.8	88.8	2138.2	14.9	172.2	224.5	109.3	37.9	654.6
Cl μeq L^{-1}	26.7	5.6	2992.0	14.1	119.3	306.2	20.6	118.4	34.4
Color PCU	1.0	30.0	1.0	2.0	1.0	30.0	37.2	110.0	5.0
Conductance μS cm^{-1}	36.3	36.0	445.0	6.9	45.3	68.0	26.9	28.7	94.0
DOC mg L^{-1}	0.5	5.7	0.14	0.4	0.2	5.2	5.9	10.9	1.4
K μeq L^{-1}	3.6	3.6	23.	4.1	7.7	20.2	5.1	7.4	12.9
Mg μeq L^{-1}	28.0	33.2	794.	4.1	58.4	60.1	45.3	37.0	227.9
Na μeq L^{-1}	13.5	23.5	839.6	14.8	145.3	264.5	34.4	136.2	56.1
NH$_3$ μeq L^{-1}	25.4	0.5	0.6	0.7	2.1	0.64	0.5	1.3	0.4
NO$_3^-$ μeq L^{-1}	53.9	0.6	2.1	16.8	0.8	15.7	2.9	2.5	21.1
pH	4.40	4.42	5.56	5.45	6.99	6.41	5.20	5.11	7.75
Si mg L^{-1}	0.10	2.4	0.01	0.02	0.16	2.5	2.1	1.2	1.1
SO$_4^{2-}$ μeq L^{-1}	110.1	171.1	782.0	7.5	156.5	153.9	143.6	39.3	64.5