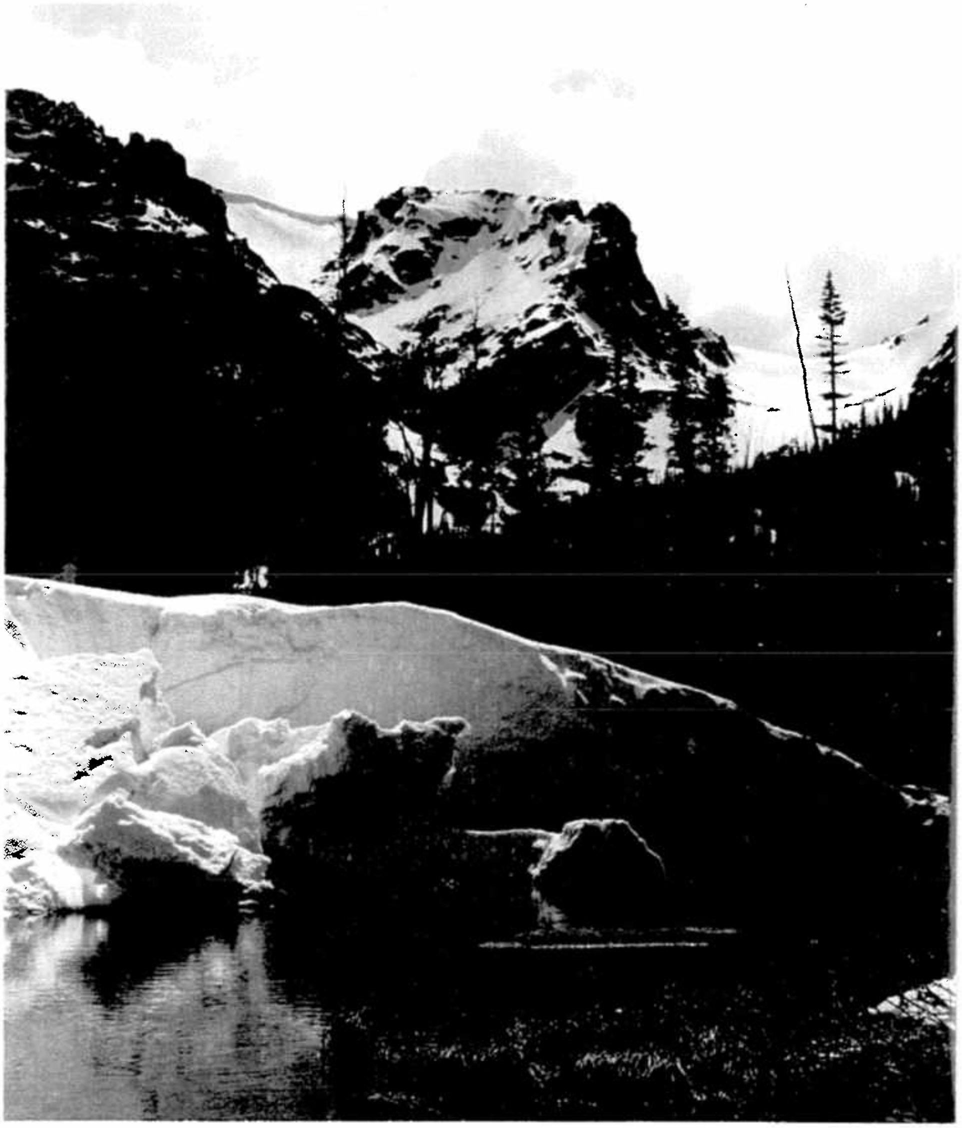
Loch Vale Watershed Project
Quality Assurance Report: 1995-1998

Open File Report 99-111



U.S. Department of the Interior
U.S. Geological Survey

USGS
science for a changing world

U.S. DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

Loch Vale Watershed Project
Quality Assurance Report: 1995-1998

By

U.S. Geological Survey

Eric J. Allstott[2], Michael A. Bashkin[2], Jill S. Baron[1]

Open-File Report 99-111

Prepared in cooperation with the Natural Resource Ecology Laboratory
and Colorado State University

This report is preliminary and has not been reviewed for conformity with the U.S. Geological Survey editorial standards. Any use of trade, product or firm names is for descriptive purposes and does not imply endorsement by the U.S. Government.

[1]Midcontinent Ecological Science Center

[2] Natural Resource Ecology Laboratory

1999

Table of Contents

List of Tables and Figures

Introduction

The Loch Vale Watershed (LVWS) project was initiated in 1980 by the National Park Service with funding from the Aquatic Effects Research Program of the National Acid Precipitation Assessment Program. Initial research objectives were to understand the processes that would either mitigate or accelerate the effects of pollution on soil and surface water chemistry, and to build a record in which long-term trends could be identified and examined.

It is important for all data collected in Loch Vale to meet the high standards of quality set forth in previous LVWS QA/QC reports and LVWS Methods Manuals. Given the ever-widening usage of data collected in Loch Vale, it is equally important to provide users of that data with a report assuring that all data are sound. Parameters covered in this report are the quality of meteorological measurements, hydrological measurements, surface water chemistry, and similarities in catch efficiency of two raingage types in Loch Vale for the period of 1995-1998.

Routine sampling of weather conditions, precipitation chemistry, and stream/lake water chemistry began in 1982. Since then, all samples and data have been analyzed according to widely accepted and published methods. Weather data have been collected, analyzed, and stored by LVWS project personnel. Methods for the handling of meteorological data are well documented (Denning 1988, Edwards 1991, Newkirk 1995,and Allstott 1995). Precipitation chemistry has always been collected according to National Atmospheric Deposition Program protocol (Bigelow 1988), and analyzed at the Central Analytical Laboratory of the Illinois State Water Survey in Champaign, IL. QA/QC procedures of the National Atmospheric Deposition Program are well documented (Aubertin 1990). Protocols for sampling surface waters are also well documented (Newkirk 1995). Analysis of surface water chemistry has been performed using standard EPA protocol at the US Forest Service's Rocky Mt. Station Biogeochemistry Laboratory since 1993.

Comparison of Alter-Shielded and Nipher-Shielded Raingages

An Alter-shielded Belfort raingage was installed in Loch Vale during September 1983 as part of the National Atmospheric Deposition Program's (NADP) CO98 network site. A Nipher-shielded Belfort raingage was added as a backup gage in October 1987. The two gages are approximately 5 meters apart, and should receive the same amount of precipitation. The gages may differ in their collection efficiency due to the physical properties of each shield type. Comparisons of catch efficiency between Nipher- and Alter-shielded Belfort raingages (from October 1987 - April 1989) revealed a slightly greater snow catch (10%) from the Nipher-shielded gage (Bigelow et al. 1990).

NADP has adopted the Alter-shielded gage as standard equipment. Nipher-shielded data were reported to NADP from the Loch Vale site instead of the Alter-shielded data from 1/1/95 - 12/11/97. The purpose of this analysis is twofold: 1) to analyze the preparation of Alter-shielded gage data from 1/1/95 - 12/11/97 for possible substitution into the NADP database, and 2) to compare the daily total and cumulative precipitation by season and precipitation type between the two shields.

Methods

Precipitation data from the Nipher-shielded gage were recorded continuously on strip charts. These have been summed by hand to yield a daily total. Error in the Nipher-shielded precipitation record may be caused by changes in temperature, windshake, snow buildup on shield, and human recording error. Precipitation data from the Alter-shielded gage were recorded every fifteen minutes and transmitted hourly to Wallops Island, Va. via the GOES West satellite network. The data are made available through the USGS Water Resources Division DOMSAT Receive Station located in the Denver Federal Center. Sources of error in the Alter-shielded data include windshake, changes in temperature, problems with satellite transmission, weather station voltage, and state of the DCP's calibration and human interpretation. Raw data from the Alter-shielded gage have many false dips and spikes due to noise in the voltage transducer. The dips and spikes were smoothed by hand to derive a cumulative record.

Various running medians were applied to smooth the data (1:3 through 1:50). No one was able to simultaneously smooth small and large oscillations. Each data point was evaluated in relation to surrounding data and corrected when necessary. For example, when data oscillated continuously, the mean was assigned to all respective data. When a large dip or spike appeared it was flattened out to conform to the surrounding data points. The smoothed data were converted to a daily total by subtracting the midnight value from the previous day's midnight value.

Results and Discussion

The smoothed daily total Alter-shielded and Nipher-shielded gage data were not statistically different (paired t-test: $p = 0.34$, $n= 975$). The data were further tested for differences by season and precipitation type using Bonferroni-adjusted paired t-tests to control for Type 1 error. There were no significant differences between winter (10/1 - 4/14), spring (4/15 - 7/14), summer (7/15 - 9/30), and rain or unknown precipitation types

(Table 1). However, the Nipher-shielded gage reported 50 mm more precipitation as snow (p = 0.0001, n = 310). This is similar to results from Bigelow et al. (1990) and Goodison et al. (1989). Winter and snow results can be different because the winter comparison uses all days (n = 460) whether or not there was precipitation, while the comparison based on snow uses only days when precipitation occurs (n = 310).

Table 1. P values for Bonferroni-adjusted paired t-test of Nipher-shielded and Alter-shielded gage data by season and precipitation type from 1/1/95 - 12/11/97.

Season / Type	*P Value*	*n**
All Data	0.34	975
winter	0.18	460
spring	0.78	230
summer	0.96	199
rain	0.48	109
unknown	0.56	60
snow	0.0001	310

*The total number of days is 1077. The total number of possible comparisons is 975 because of missing data. For comparisons based on precipitation type, only a subset of the data had precipitation type recorded.

Correcting for Missing Data

Because missing data may bias the cumulative record, we compared the record from the Alter-shielded and the Nipher-shielded gages when the data were substituted, and one when the record was incomplete. The incomplete record is the smoothed Alter-shielded gage and raw Nipher-shielded gage data. The complete record uses data from one gage when the other gage is missing points. For example, the Alter-shielded gage data were not transmitted on day 121, so the Nipher-shielded gage data were used as a proxy for that day. However, problems of auto-correlation may appear when comparing the cumulative records. The Alter-shielded gage missed 37 days out of 1077 days (3.4%) between 1/1/95 and 12/11/97. The Nipher-shielded gage missed 65 out of 1077 (6%). Given these low percentages the problem of auto-correlation is negligible.

To compare the cumulative records, 95% confidence intervals were calculated for the slopes of the incomplete/complete Alter-shielded and Nipher-shielded records. Figures 1 through 4 show the comparisons between gages and between complete and incomplete data from the same gage. None of the confidence intervals overlapped, suggesting all slopes are statistically different (Table 2). This is the opposite effect detected with the daily comparisons when using paired t-tests. However, the paired t-test is a more robust comparison because each pair is considered an independent sample. Comparison of slopes may be prone to error because an error early in the data is perpetuated through out the entire data set.

Table 2. Slopes and 95% CI for cumulative Alter & Nipher-shielded gage data.

Record	*Slope*	*Upper 95%limit*	*Lower 95% limit*
Incomplete Alter-shielded	0.137	0.138	0.136
Complete Alter-shielded	0.143	0.144	0.142
Incomplete Nipher-shielded	0.126	0.127	0.125
Complete Nipher-shielded	0.134	0.135	0.133

Figure 1. Comparison of complete and incomplete Nipher data.

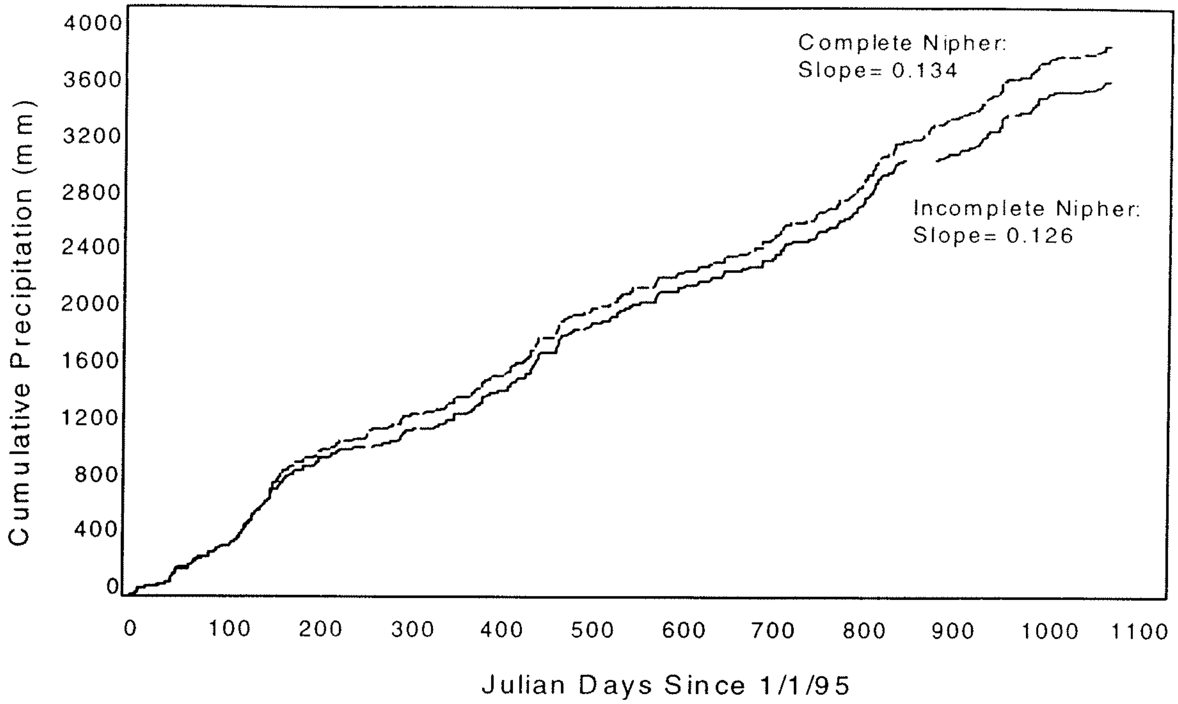


Figure 2. Comparison of complete and incomplete Alter data.

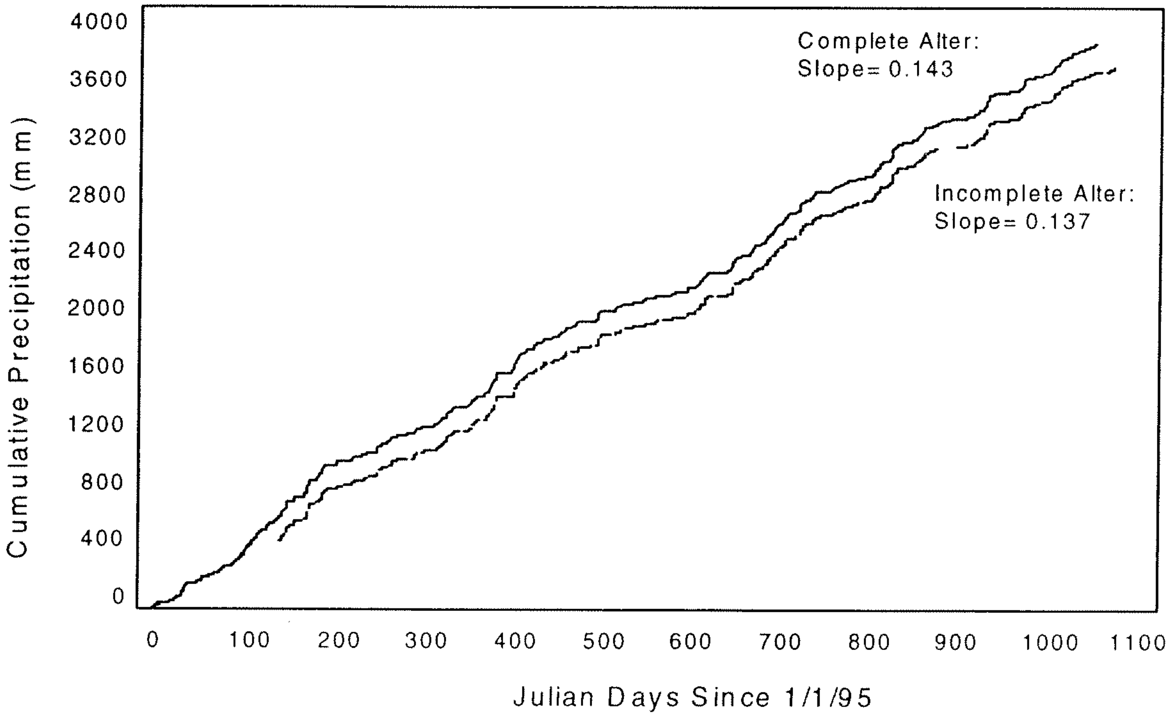


Figure 3. Comparison of incomplete Alter and Nipher data.

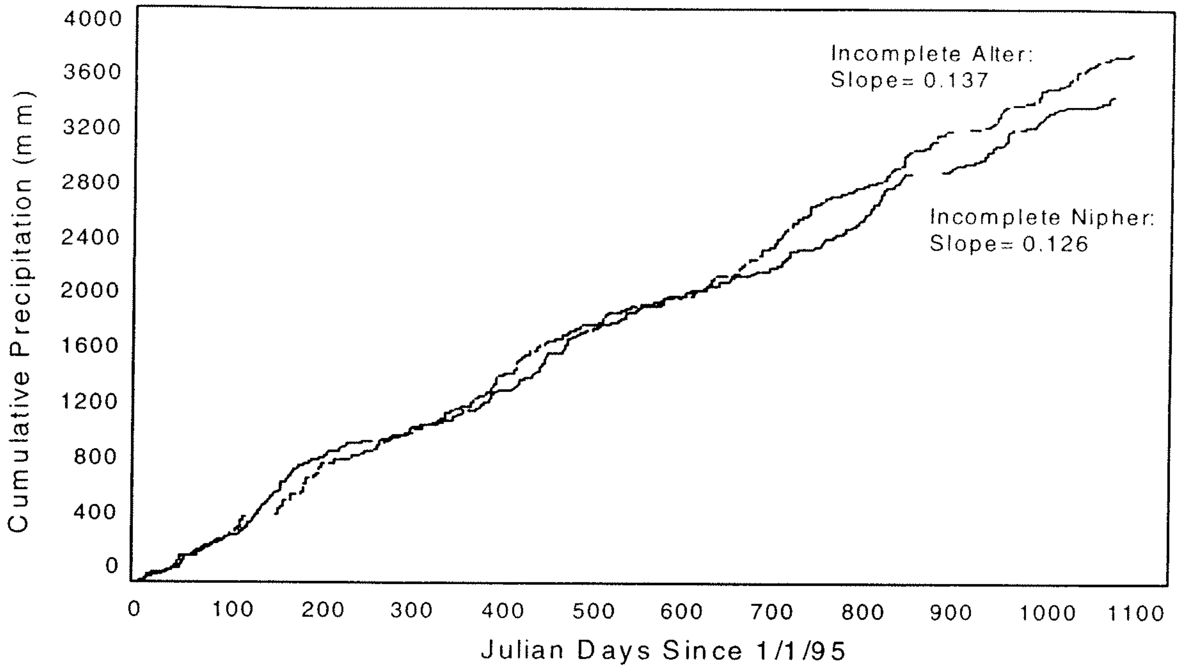
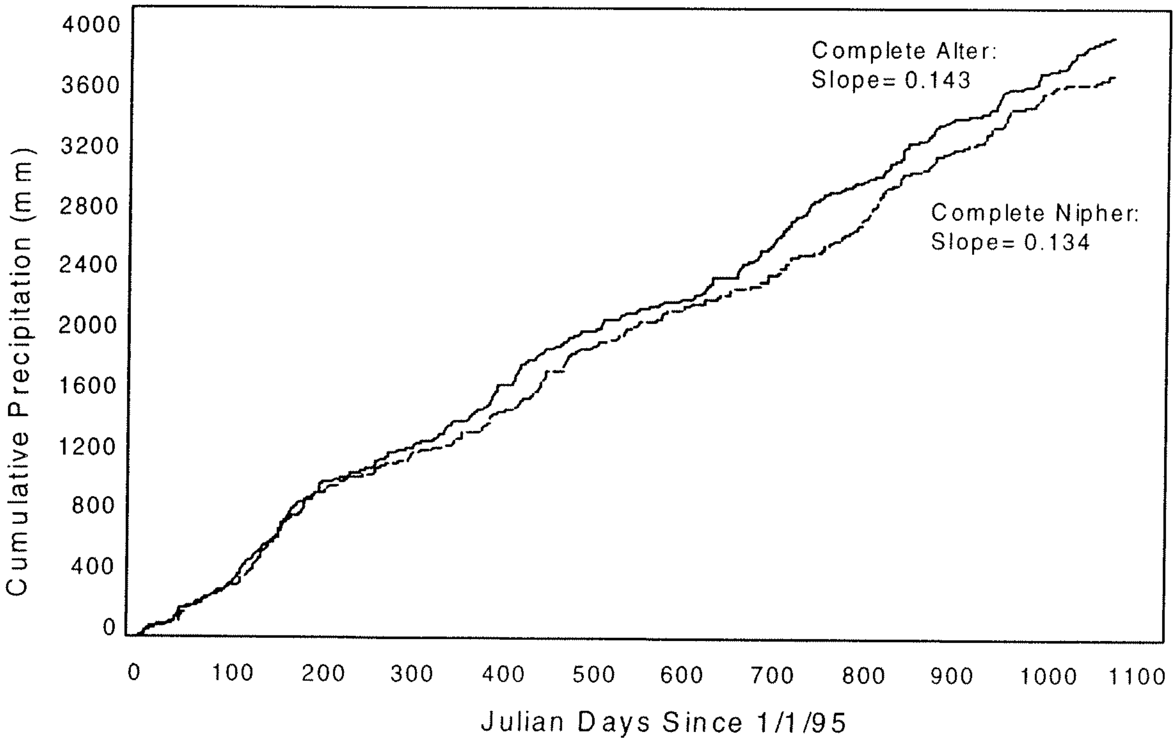


Figure 4. Comparison of complete Alter and Nipher data.



Differences in total cumulative precipitation between complete and incomplete data from both gages suggest missing data (< 6%) may slightly bias cumulative precipitation records (Table 3). The total difference between the complete and incomplete Alter-shielded and Nipher-shielded gage is 168 mm (4.5 % of total record) and 240 mm (6.9 % of total record) respectively.

Table 3. Difference in Alter-shielded and Nipher-shielded total complete and total incomplete cumulative precipitation records, by year and total period of analysis (1/1/95 - 12/13/97). Data are expressed in millimeters.

Record	*Cumulative precipitation 1995*	*Cumulative precipitation 1996*	*Cumulative precipitation 1997*	*Total Cumulative precipitation*
Incomplete Alter-shielded	1255	1345	1196	3796
Complete Alter-shielded	1414	1357	1193	3964
Incomplete Nipher-shielded	1226	1151	1099	3476
Complete Nipher-shielded	1332	1176	1208	3716

Conclusion

There was no difference between the Alter-shielded and Nipher-shielded gages when daily records were compared with a paired t-test, except when snow was tested separately. The Nipher-shielded gage reported 18% more snow than the Alter-shielded gage. This is similar to Bigelow et al. (1990) who reported 10% more snowfall from the Loch Vale Nipher -shielded gage. Goodison et al. (1989) reported 15% greater snow catch from Canadian sites. Previous studies suggest the Nipher-shielded gage is better at estimating snowfall, while the Alter-shielded gage under-represents snowfall (Goodison et al. 1983).

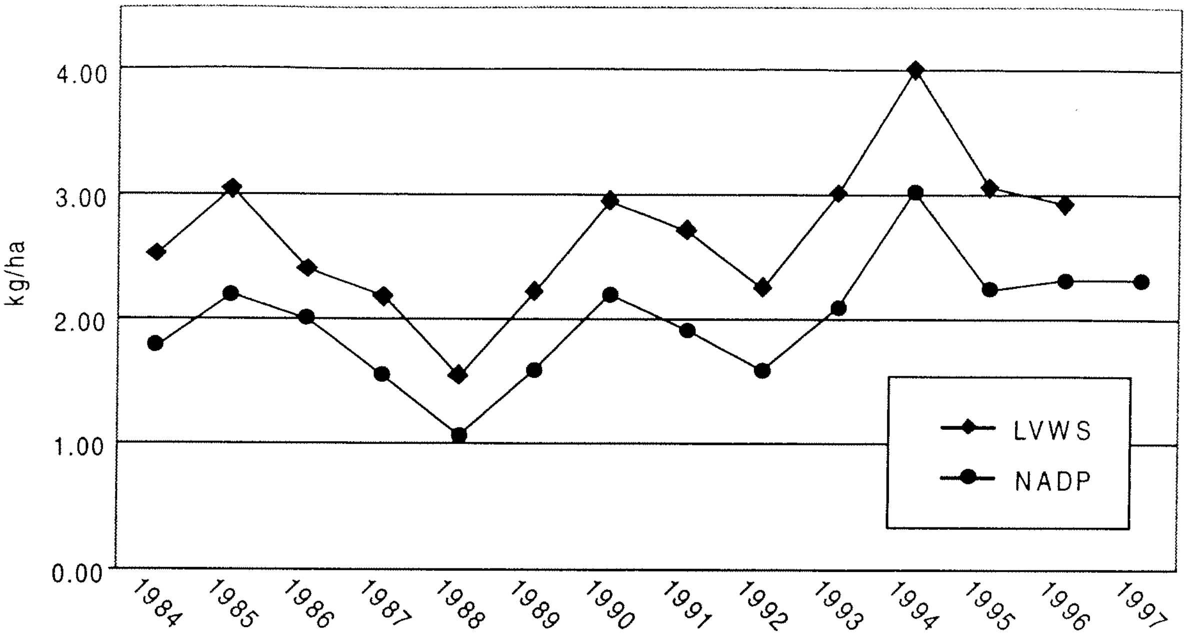
There was no overall difference between the two gages and the percentage of missing data was low (< 6%). Therefore, we suggest substituting for missing data is a sensible practice to better estimate cumulative precipitation.

Substituting for missing or corrupt data allows a more accurate calculation of annual chemical input budgets. Since under-reporting precipitation amount drives deposition down, it only makes sense to report all available precipitation data.

An amendment to modify NADP protocol is before the review committee as of this date. This amendment, if approved, would allow pre-approved sites to substitute Nipher data in the absence of Alter data. This will be an important change, if approved, because the threat of invalidation of our 1995-97 precipitation data will be lifted.

LVWS precipitation records are more complete than NADP's since Nipher data has always been included in our database. Use of LVWS precipitation records in the calculation of chemical deposition values produces more realistic estimates than when NADP records are used. NADP's conservative use of precipitation data causes a significant under calculation of chemical deposition. This is seen in the case of Nitrogen deposition where a t-test produced a P-value of 0.001- an underestimation of N deposition by NADP of 1.37 times (Figure 5).

Figure 5. Estimated nitrogen deposition at CO98: NADP vs. LVWS calculation.



Shields were exchanged during summer 1998 so the Alter shield now resides on the gage used to collect weekly precipitation amounts. The Nipher shield is now on the gage connected to the AeroChem event recorder. This gage collects data pertaining only to timing and duration of weekly precipitation- not amount. The as-of-yet-invalid shield type is not an issue because of the modified function of this gage. Function of both gages will be combined into one summer 2000. This will bring collection of CO98 data into line with the other NADP network sites.

Comparison of BRD and WRD Loch Vale Weather Station Data

There are four weather stations in LVWS. One is near the base of the Sharkstooth, and another is in Andrews Meadow. The two others, which are the focus of this section, are co-located near the confluence of Icy Brook and Andrews Creek. These two stations are ~30m apart. One station, established in 1983, was maintained by the USGS-Biological Resources Division (BRD). The second weather station, established in 1991, is maintained by the USGS-Water Resources Division (WRD).

The BRD station was shut down on Nov. 24, 1998 for several reasons. The quality of data from all measured parameters declined even after each sensor was recalibrated and the data collection platform reprogrammed. Other reasons for decommissioning the BRD station include reduction of growing equipment costs, lack of available replacement components, lack of certified repair service, and the importance of a more standard method of data collection. Data from the WRD station will be included in the LVWS database. Merging the two station's data sets into one is prudent only if data from each measured parameter proves to be statistically similar. This report compares weather data from both stations from 1/1/93 - 12/7/97.

The BRD station was a solar powered remote area weather station (RAWS). Weather parameters measured by this station were wind speed, wind direction, relative humidity, ambient air temperature, solar radiation, and barometric pressure. Parameters were measured every 15 minutes by a Handar 524 data collection platform (DCP). These data were relayed hourly to the National Oceanic Atmospheric Administration (NOAA) receive station at Wallops Island, VA via the Geostationary Operational Environmental Satellite West (GOES) satellite network.

Data were quality checked and edited by LVWS personnel before loading into the master database for two reasons: 1. the program that loaded data into the database could only process the data in one particular format; and 2. so that incomplete or erroneous transmissions would not be included in the database. These data were imported weekly the LVWS master database.

The WRD station scans every 15 seconds and calculates average hourly and daily values. The average hourly and daily values are recorded to Campbell Scientific data module that is collected and exchanged for a new module every two to three months. The data are available through an FTP link to the WRD.

Methods

The weather stations record many types of data, but only five variables are common to both stations (Table 4). The variables compared in this analysis are temperature, solar radiation, relative humidity, wind speed, and wind direction. Average daily values from both stations were used for comparison. The data were converted into the appropriate units for comparison when necessary. Data were adjusted to the nearest possible value when obviously erroneous. For example, in some cases, the BRD relative humidity was 150% or -10%; these values were adjusted to 100 and 0% respectively. Data were tested for significant differences with paired t-tests, and subsets of the data were tested for differences using Bonferroni-adjusted paired t-tests to control for Type 1 error.

Table 4. Meteorological parameters recorded by WRD- BRD weather stations.

WRD STATION VARIABLES	UNITS	BRD STATION VARIABLES	UNITS
Daily Ave. QS radiation in	Langleys/hour	Hourly incoming QS rad.	Watts/m^2
Daily Ave. QS radiation out	Langleys/hour		
Daily Ave. QA radiation in	Langleys/hour		
Daily Ave. QA radiation out	Langleys/hour		
Daily Ave. net radiation	Langleys/hour		
Minimum 6 & 2 m air temperature.	Degrees C		
Average 6 & 2 m air temperature.	Degrees C	2m air temp.	Degrees C
Maximum 6 & 2 m air temperature.	Degrees C		
Minimum 6 & 2 m relative hum.	%		
Average 6 & 2 m relative hum.	%	Relative humidity	%
Maximum 6 & 2 m relative hum.	%		
Maximum 6 & 2 m windspeed	mph		
6 & 2 m EPA windspeed	mph	6m wind speed	Meters/sec.
6 & 2 m Campbell wind speed	mph		
6 & 2 m Campbell vector magnitude	mph		
6 & 2 m EPA UT vector dir.	360 Degrees		
6 & 2 m EPA vector dir. std. dev.	360 Degrees		
6 & 2 m Campbell vector dir.	360 Degrees	6m wind dir.	360 Degrees
6 & 2 m Camp. vector dir. std. dev	360 Degrees		
		Barometric pressure	Millibars

Results and Discussion

Temperature

On day 1357 (and continuing through the data set), the BRD station temperature data became spurious (Figure 7). Before that, WRD daily average temperature was 0.17 C lower then the BRD station ($p < 0.0001$, $n = 1352$). Although the difference is statistically significant, 0.17 C is within the sensor's accuracy range (+/- 1degree F). Temperature data were divided by year to determine if differences in temperature are at a constant significance level. P-values steadily decreased from 0.2290 in 1993 to <0.0001 in 1997. The difference between stations markedly increased starting late in 1995 (Figure 6 and Table 5). Since the BRD temperature/humidity sensor was replaced with a reconditioned one in 1994, 1996, and 1997, it appears the problem was with BRD station's DCP motherboard.

Figure 6. Difference of temperature measured at BRD and WRD stations.

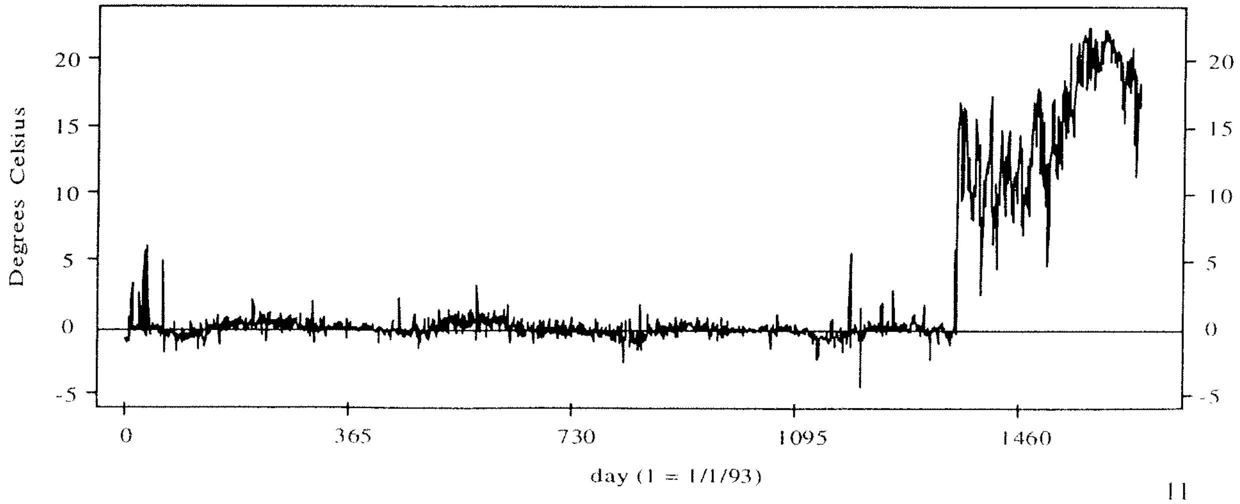


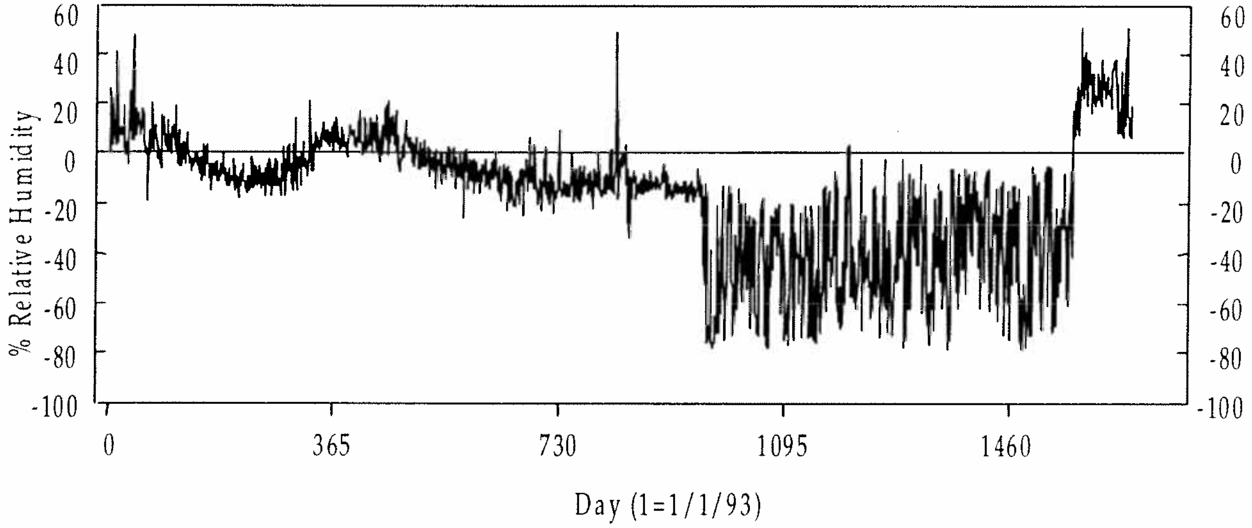
Table 5. P values and mean difference by year for temperature measured by the WRD and BRD weather stations. P values reported are Bonferroni adjusted to protect against Type 1 errors. Mean difference is WRD minus BRD.

Year	*p value*	*mean difference*
93	0.2290	-0.054
94	0.0310	-0.062
95	<0.0001	-0.309
96	<0.0001	1.513
97	<0.0001	15.574

Humidity

Relative humidity from the BRD station reported very high values, often > 100% starting day 960, and then reported low (real?) values from day 1561 to the end of the record. However, prior to this period the average difference between the two stations was 4.1% ($p < 0.0001$, $n = 960$) (Figure 7). Relative humidity is difficult to measure accurately due to glazing of water and or ice on the sensor itself.

Figure 7. Difference of relative humidity measured at BRD and WRD stations.



Solar Radiation

The BRD pyranometer appeared to malfunction on day 1323 (Figure 8). When these data were excluded from statistical comparison, the WRD station on average reported 1.9 watts/m^2 more then the BRD station ($p = 0.002$; $n = 1323$). Solar radiation data were further divided by year to determine if difference was at a constant significance level over time. The BRD pyranometer was reconditioned and calibrated on July 1, 1997, yet spurious measurements persisted. P-values decrease from 1993 to 1997, again suggesting problems with the BRD DCP motherboard (Table 6).

Figure 8. Difference of solar radiation measured at BRD and WRD stations.

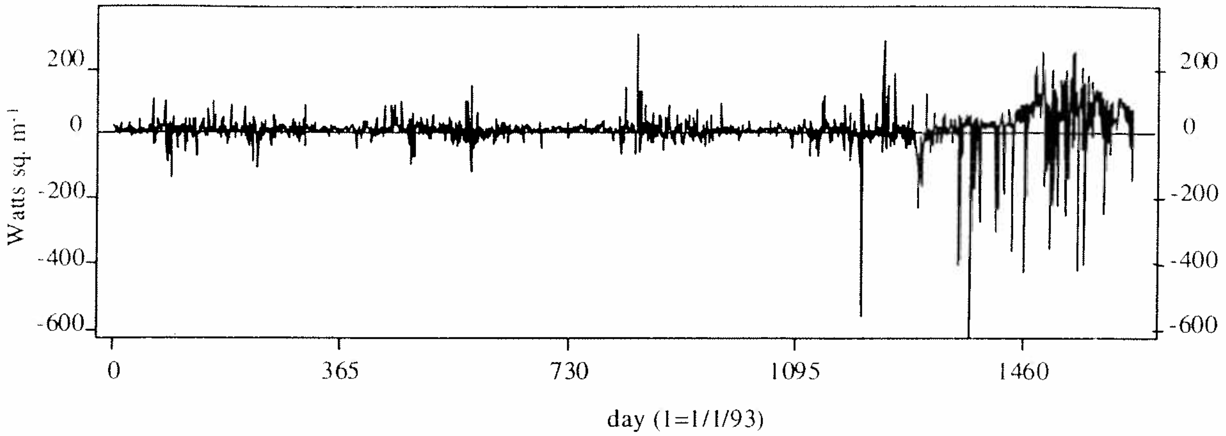


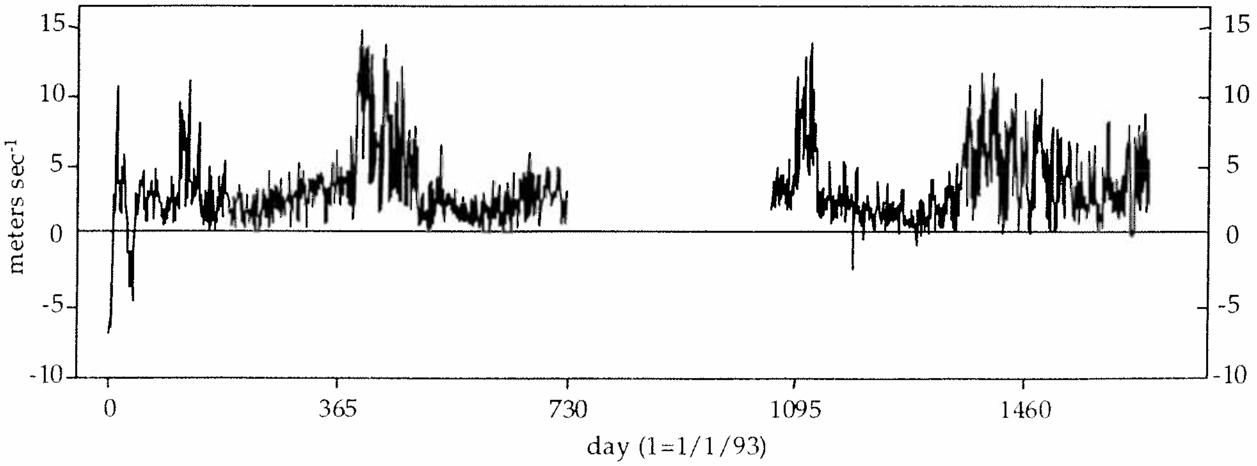
Table 6. P values and mean difference by year for solar radiation measured by the WRD and BRD weather stations. P values reported are Bonferroni adjusted to protect against Type 1 errors. Mean difference is WRD minus BRD.

Year	*p value*	*mean difference watts/m^2/day*
93	0.1870	1.68
94	0.0320	2.64
95	0.0400	3.15
96	0.0450	-8.26
97	<0.0001	35.32

Wind Speed

The BRD anemometer malfunctioned between days 728 and 1055. Excluding erroneous data, wind at the WRD station was an average of 1.7 m/s greater than at the BRD station (p <0.0001, n = 1135) (Figure 9). WRD wind speeds were greater for the majority of days, suggesting a difference in prevailing wind direction/speed or in sensor calibration.

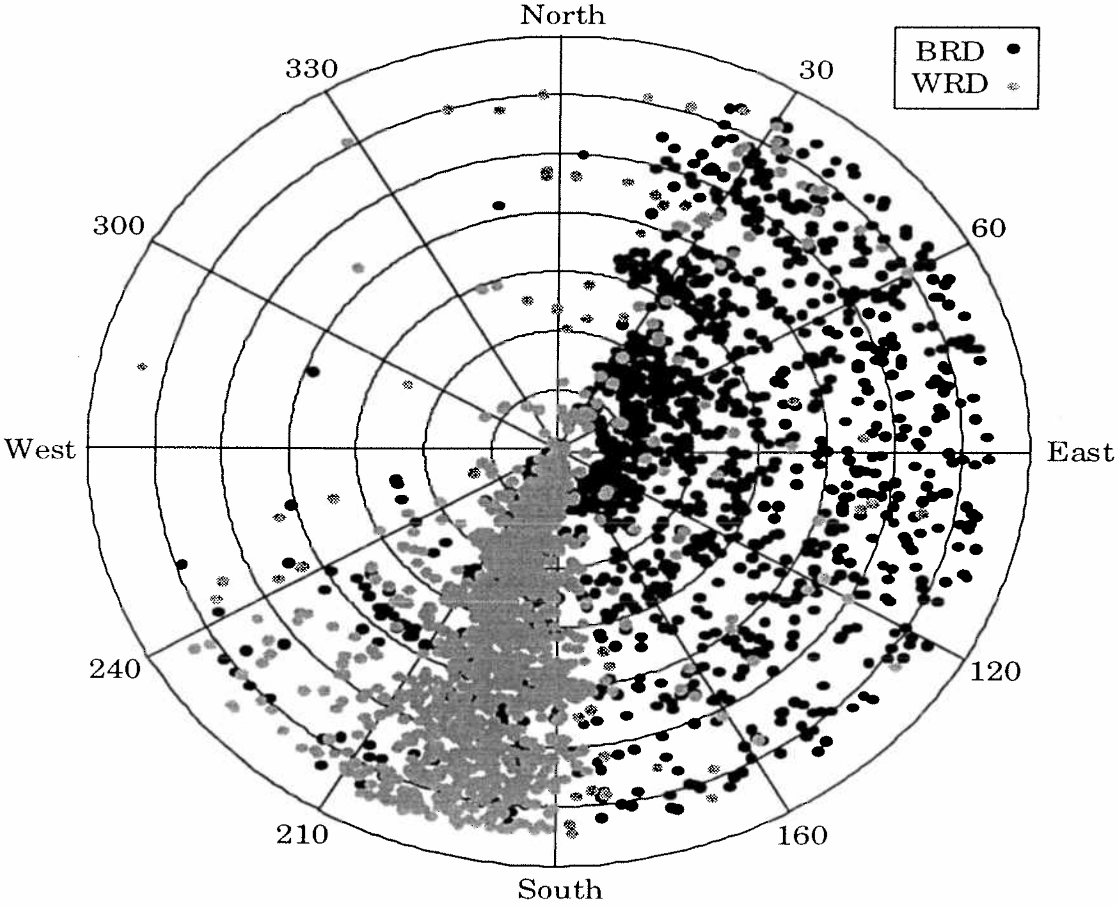
Figure 9. Difference of wind speed measured at BRD and WRD stations.



Wind Direction

Wind direction measurements are the only ones at the BRD station that don't show conclusive signs of malfunction. Wind direction was different between the two stations, with 84% of the WRD wind direction coming between 180 - 210 degrees, while the BRD station wind direction reported 81% wind direction coming from between 180 - 30 degrees (Figure 10). The BRD wind direction record showed more variability than the WRD record, perhaps due to topographic differences between the two stations. There is nothing to suggest that either atmometer was malfunctioning. Several atmometers may be needed to capture the variability of wind direction within LVWS.

Figure 10. Prevailing wind direction at BRD and WRD stations.



Conclusion

Temperature and solar radiation were similar until the BRD station began to fail. Both temperature and solar radiation became increasingly different with time, pointing to a progressive degradation of the BRD station. The BRD record shows erratic relative humidity and wind speed measurements throughout the record, suggesting intermittent station malfunctions. Microclimatological variability seems to drive differences in wind direction.

A new set of fields will be set up in the database for the inclusion of WRD station data. Requests for Loch Vale weather data will be filled with BRD data from Sept. 8, 1982 until Dec. 31, 1994 and from Jan. 1, 1995 - on from the WRD station.

Surface Water Chemistry

Surface water samples are routinely collected throughout the entire watershed. The focus of this section is on quality control/quality assurance (QA/QC) samples collected from the outlet of the Loch. Twelve percent of the total numbers of samples taken from the outlet (from 1995-1998) were QA/QC samples. All QA/QC procedures used in the preparation, collection, processing, and shipping of samples have been described in the LVWS Methods Manual (Newkirk, 1995), and by Denning (1988 QA/QC), Edwards (1991 QA/QC), and Allstott (1994 QA/QC). Small procedural changes since 1995 will be updated in the forthcoming 1999 version of the methods manual.

USFS Water Quality Laboratory

Chemical analyses of water samples were performed by Louise O'Deen at the US Forest Service's Rocky Mt. Station Biogeochemistry Laboratory, Fort Collins, CO. The USFS lab participates in a national blind audit round-robin sampling program sponsored by the USGS. Results from these tests provide laboratories with information about their analytical precision and accuracy. The USFS lab has always (with four exceptions) exceeded minimum standard requirements in the analysis of standard reference samples (USGS Open-File Reports 96-138, 96-436, 98-52, and 98-391). The cause of these exceptions is most probably because most labs that participate in the comparison report values gained from using atomic absorption methods, while the USFS laboratory reports values from their ion chromatography methods. A summary of results from 1995-1998 is listed in Table 7.

Table 7. Results of USGS Analytical Evaluation Program for Lab #2 (US Forest Service's Rocky Mt. Station Biogeochemistry Laboratory). 4=Excellent, 3=Good, 2=Satisfactory, 1=Marginal, 0=Unsatisfactory.

	Constituent	*Reported Value*	*Most Probable*	*Rating*
1995	Ca	1.84	1.67	1
	Cl	1.39	1.30	3
	F	Not Reported	Not Reported	Not Reported
	K	0.06	0.55	3
	Mg	0.40	0.35	0
	Na	1.94	1.28	0
	SO_4	2.26	2.34	4
	pH	6.42	6.52	4
1996	Ca	0.48	0.45	4
	Cl	Not Reported	Not Reported	Not Reported
	F	Not Reported	Not Reported	Not Reported
	K	0.16	0.15	3
	Mg	0.07	0.06	3
	Na	Not Reported	Not Reported	Not Reported
	SO_4	Not Reported	Not Reported	Not Reported
	pH	4.68	4.70	4
1997	Ca	1.87	1.84	4
	Cl	0.16	0.20	4
	F	0.03	0.06	3
	K	0.35	0.37	4
	Mg	0.59	0.57	3
	Na	0.80	0.66	0
	SO_4	1.05	1.10	4
	pH	6.83	6.85	4
1998	Ca	0.11	0.13	3
	Cl	0.20	0.23	4
	F	0.20	0.21	4
	K	0.14	0.14	4
	Mg	0.02	0.03	4
	Na	0.34	0.34	4
	SO_4	0.41	0.40	4
	pH	5.41	5.35	4

NREL Water Quality Laboratory

The Loch Vale water quality laboratory participates in the USGS Intersite Comparison Program in which conductivity and pH measurements are periodically tested on samples of known chemical properties. Results from past studies are listed in Table 8. We have met all accuracy goals with one exception. The measured pH in Study #41 was .02 over the upper limit of acceptable values. The source of the error was determined to be the pH probe itself, which was replaced within the week.

Table 8. USGS Intersite Comparison Program Results for CO98 (LVWS).

Comp #	*Date*	*Exp. pH*	*Meas. pH*	*Met Goal?*	*Expected Cond.*	*Measured Cond.*	*Met Goal?*
#35	May-95	4.95	5.02	yes	15.7	18.0	yes
#36	Nov-95	4.92	4.98	yes	5.3	5.2	yes
#37	May-96	4.47	4.51	yes	26.3	28.0	yes
#38	Dec-96	4.22	4.25	yes	26.8	27.0	yes
#39	May-97	4.78	4.80	yes	13.0	14.4	yes
#40	Nov-97	4.12	4.17	yes	33.9	35.0	yes
#41	May-98	4.62	4.76	no	20.1	21.0	yes
#42	Dec-98	4.38	4.39	yes	26.9	25.9	yes

The pH meter used since the mid-eighties became impossible to calibrate in January 1999. The meter was replaced within the week with a new VWR 8000 model (approved by NADP). We continue to use pH probes supplied by NADP for all samples.

Samples to be analyzed for dissolved organic carbon are collected and processed according to standard methods (LVWS Methods Manual, 1995). Filtered samples are delivered bi-weekly to George Aiken's Boulder, CO USGS laboratory.

Preliminary QA/QC Checks

Results of chemical analyses are screened by USFS laboratory personnel and LVWS staff. The USFS uses the ion percent difference (IPD = (sum of anions + sum of cations)/total ion concentration) as a general indicator of analytical accuracy. Samples with an IPD > 15% are flagged and rerun. If the IPD is still out of acceptable range after being rerun, each sample is checked ion by ion for possible contamination. If one analyte is the clear cause of the imbalance, it may be dropped from the record (as long as the rest of the analytes look normal for that time of year). Twenty-two of 342 normal samples (6.4%) fell outside the 15% IPD margin from 1995 to 1998. Nineteen of these samples were taken between May and September of 1995. There is no apparent cause for this period of charge imbalance. Contamination introduced in the field or lab is the most probable cause of these errors.

Bias and Contamination

Bias and contamination of water samples are quantified by measuring elemental concentrations in field blank samples. Field blanks are deionized water samples that are processed from start to finish just as normal samples. These samples are taken every other third week. 5.8% of the total number of samples taken at the Loch outlet were field blanks. Sources of bias and/or contamination can include contamination of DI water columns, improper bottle washing procedures, errors in handling and processing of samples, and analytical bias in the lab. Mean concentrations of each analyte from field blank samples are presented in Table 9.

Table 9. Blank Sample Mean Concentration.

ANALYTE	*UNITS*	*MEAN CONC.*	*STD DEV.*	*N*
FLDCOND	us/cm	1.31	1.25	18
LABCOND	us/cm	1.66	1.09	20
FLDPH	pH	5.84	0.31	17
LABPH	pH	5.30	0.21	20
ALK	ueq/l	-1.80	3.27	20
Ca	mg/l	0.02	0.03	20
Mg	mg/l	0.00	0.00	20
Na	mg/l	0.00	0.02	20
K	mg/l	0.00	0.01	20
NH_4	mg/l	0.00	0.00	20
SO_4	mg/l	0.01	0.03	20
NO_3	mg/l	0.01	0.02	20
PO_4	mg/l	0.00	0.00	20
Cl	mg/l	0.01	0.03	20
F	mg/l	0.00	0.01	20
SiO_2	mg/l	0.35	0.52	17
DOC	mg/l	0.83	0.73	17

The only known contamination of the DI water system since moving to the NESB building from the Grasslands Lab (in 1994) was a silica spike during the summer/fall of 1996. The only blank sample from that period was taken on 8/13/96, and measured 44.83 mg/l (this value was not included in the average presented in Table 9). On 8/19/96, the USFS lab notified us that there was a problem with DI. DI water from the USFS Laboratory was used for all general laboratory practices from 8/19/96 until 9/11/96, at which time CSU Facilities repaired the faulty deionizing columns. No samples (besides the one blank) were effected by this contamination, as all samples were drawn from LVWS using bottles that had been stored with DI poured before 8/19/96.

Still, the average value of SiO_2 in field blank samples is higher than would be expected. Because SiO_2 carries little charge, DI scrubbing columns are not able to efficiently capture all of it. Higher than expected average DOC concentrations in field blank samples are likely due to the same cause.

Precision

Analytical precision is measured by quantifying differences between duplicate paired samples. Duplicate samples are taken every other third week at the outlet of the Loch using the identical technique as for normal samples. 6.2% of the total number of samples taken at the Loch outlet were duplicate samples.

Since precision often varies with concentration, plots of a duplicate pair's mean versus its standard deviation (SD) are useful in determining deficiencies in precision (US EPA, 1980). A problem with precision would be indicated by a relationship between all paired means and their respective SDs. The lack of relationship present in all regressions indicates good precision at varying concentrations. Table 10 describes the R-square value for each regression. The only analytes that showed any relationship between mean concentration and SD were Potassium (K), Silica (SiO_2), and Phosphate (PO_4). The SD of K (.34) and SiO_2 (.19) may indicate a slight lack of analytical precision for those solutes, or could be indicative of the smaller N value for each. The SD of PO_4 (.99) is due to the extremely low concentration of PO_4 in LVWS sample waters (a mean of ~0.0 produces a SD of ~0.0, and the relationship between the two lends a very high R-square value).

Table 10. R-Square value for duplicate pair mean vs. duplicate pair standard deviation

Analyte	*R-Square Value*
Field Conductivity	0.04
Lab Conductivity	0.01
Field pH	0.00
Lab pH	0.12
ALK	0.03
Ca	0.01
Mg	0.00
Na	0.14
K	0.34
NH_4	0.12
SO_4	0.02
NO_3	0.02
PO_4	0.99
F	0.06
Cl	0.09
SiO_2	0.19
DOC	0.02

Another method for examining analytical precision is to calculate the statistical difference between normal and duplicate pairs. Paired two-tailed t-tests were used to generate the results shown in Table 11.

Table 11. T-test differences between normal and duplicate Loch.O samples.

Analyte	*Normal Mean*	*Dupe Mean*	*N*	*Pearson Correlation*	*t-test P*
FLDCOND	14.48	14.58	18	0.99	0.44
LABCOND	12.66	12.66	16	1.00	0.97
FLDpH	6.38	6.49	18	0.94	<0.0001
LABpH	6.31	6.29	16	0.83	0.42
ALK	69.45	68.56	17	1.00	0.23
Ca	1.61	1.58	17	0.98	0.31
Cl	0.14	0.13	17	0.95	0.30
DOC	1.83	1.75	15	0.97	0.31
K	0.18	0.18	17	0.99	0.66
Mg	0.27	0.27	17	0.98	0.30
Na	0.73	0.74	17	0.98	0.51
NH_4	0.02	0.01	17	0.91	0.12
NO_3	0.87	0.87	17	0.99	0.94
PO_4	0.002	0.001	16	1.00	0.33
SiO_2	2.44	2.39	9	0.87	0.70
SO_4	1.62	1.62	17	1.00	0.92

Most pairs are statistically similar. T-test P values are high with the exceptions of NH_4 and field pH. Although not significant, the difference between paired NH_4 measurements (p= .12) is most likely due to low sample concentrations that are at, or below analytical detection limits. All approaches to the determination of detection limits are based on a statistical statement of where "signal" exceeds "noise" (MacTaggart, 1998). This limit will vary between different methods and instruments. The accuracy of any measurement of concentration below this threshold is suspect; yet still useful in the fact that we know the concentration is exceedingly small.

The only statistically significant difference found is between field pH normal and duplicate samples (p= .00005). Causes of this error could be due to a faulty pH probe, a faulty pH meter, or error in the calibration/usage of the meter. However, paired lab pH measurements (splits of the same pairs used for field pH) show a much higher degree of similarity (p= .42). These results suggest that pH measurements from the USFS lab are more accurate and reproducible.

The total number of normal samples taken at each site in Loch Vale is described in Table 12.

Table 12. Number of samples taken at each site in Loch Vale.

Sample Site	*Times Sampled (1995-1998)*
Andrews Creek	13
Andrews Tarn	1
Emerald Outlet	3
Glass Outlet	13
Haiyaha Outlet	3
Husted Surface	4
Little Loch Creek	12
Loch Hypolimnion	8
Loch Inlet	13
Loch Outlet	208
Loch Surface	9
Louise Inlet	1
Louise Outlet	4
Sky Hypolimnion	5
Sky Inlet (North)	11
Sky Inlet (South)	12
Sky Outlet	12
Sky Surface	4

While reviewing time-series charts of concentration data, it became apparent that some data points were unrealistic, and likely incorrect. Forty-nine suspicious points were identified from throughout the entire LVWS data set (1982-98). Data were reviewed on a point-by-point basis. Twenty data points proved to be erroneous, and were corrected or deleted. Common instances included blank samples recorded as normal samples, samples that were out of the +/-15 percentage IPD margin, contamination errors, and data entry/transfer errors. Corrected time series data (1990-1998) are presented in Figures 11-22.

Figure 11. Alkalinity.

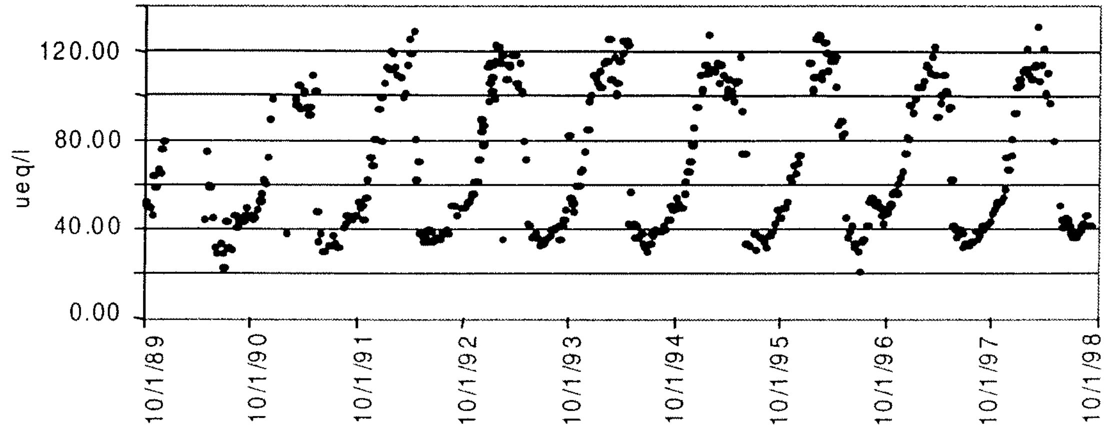


Figure 12. Calcium.

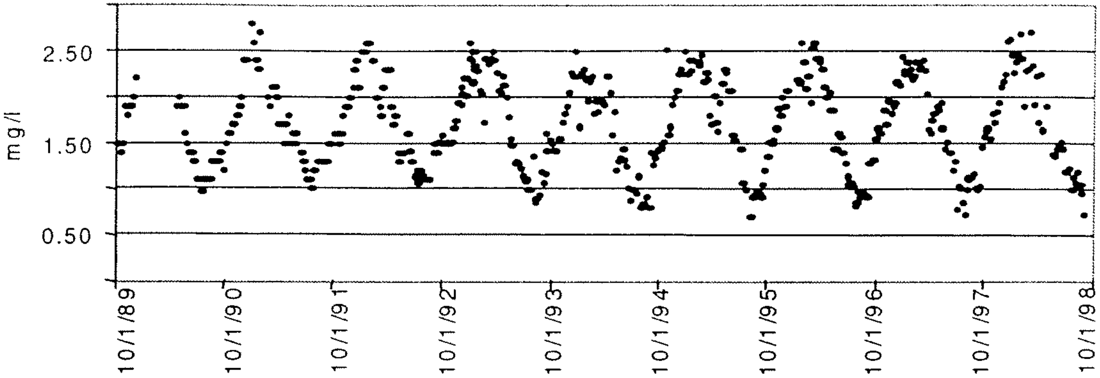


Figure 13. Magnesium.

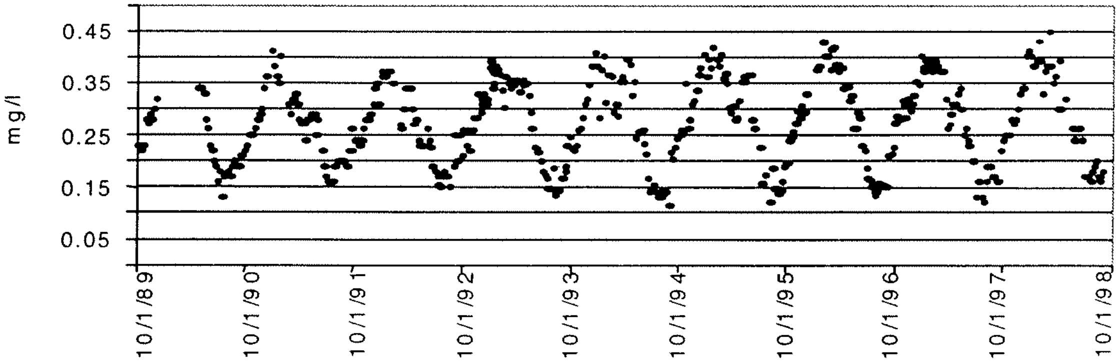


Figure 14. Potassium.

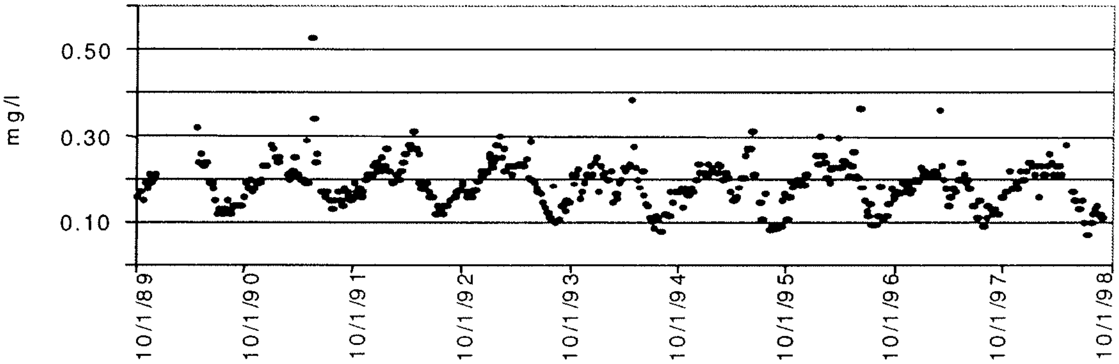


Figure 15. Sodium.

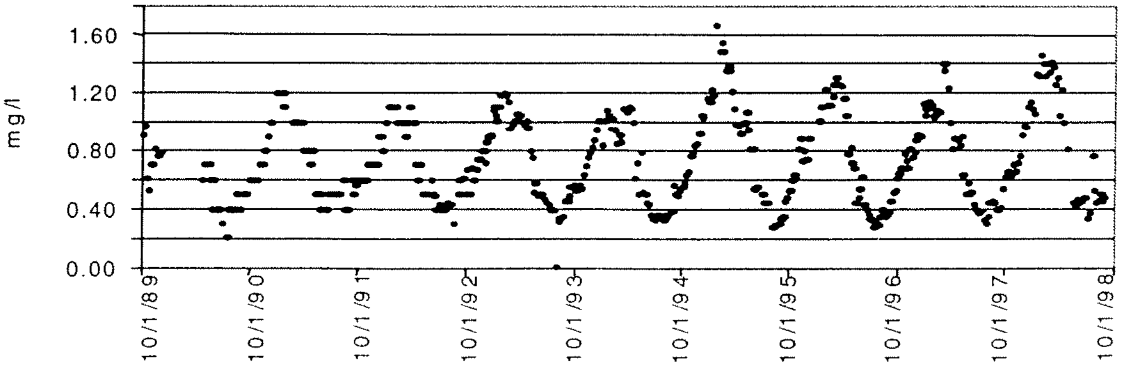


Figure 16. Ammonium.

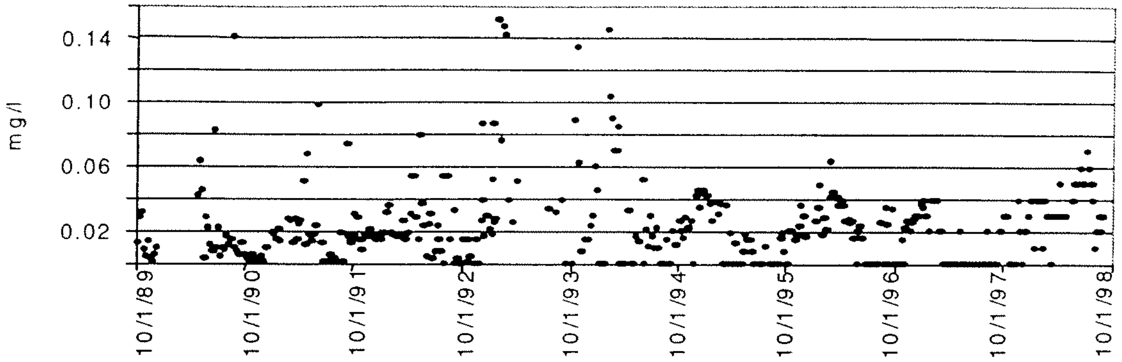


Figure 17. Sulfate.

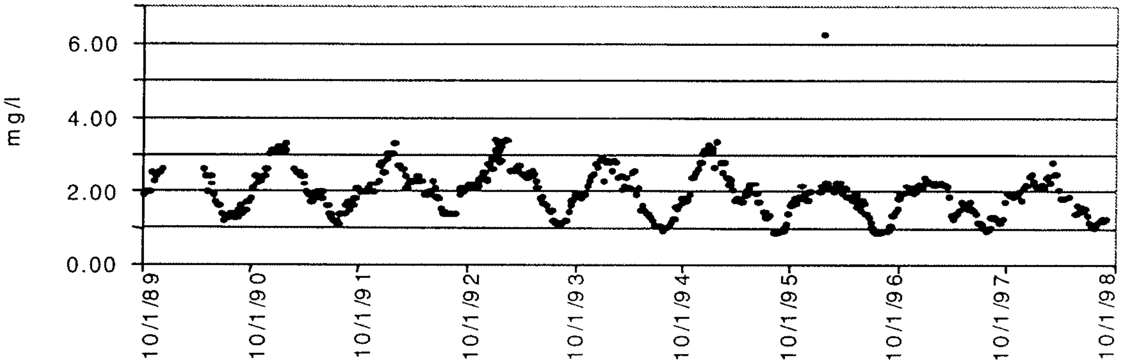


Figure 18. Nitrate.

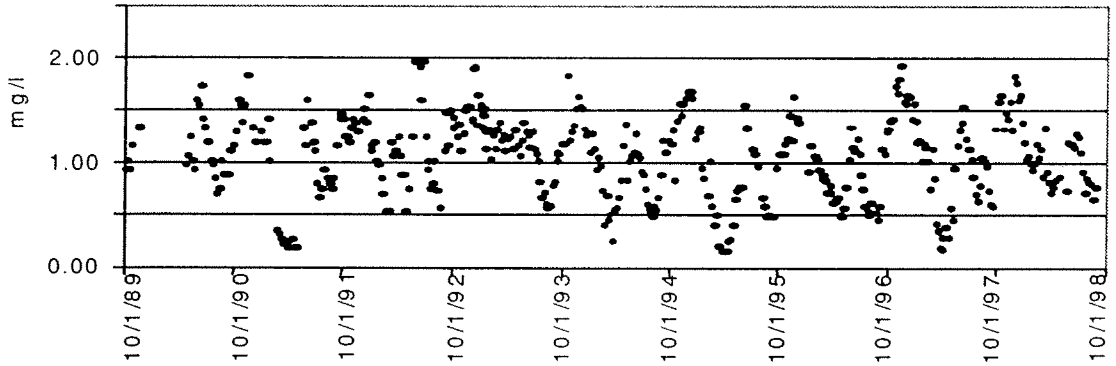


Figure 19. Chloride.

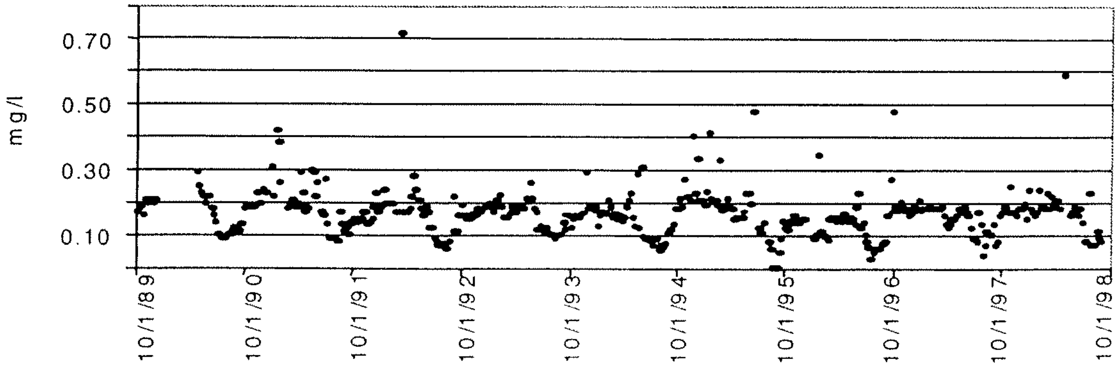


Figure 20. Phosphate.

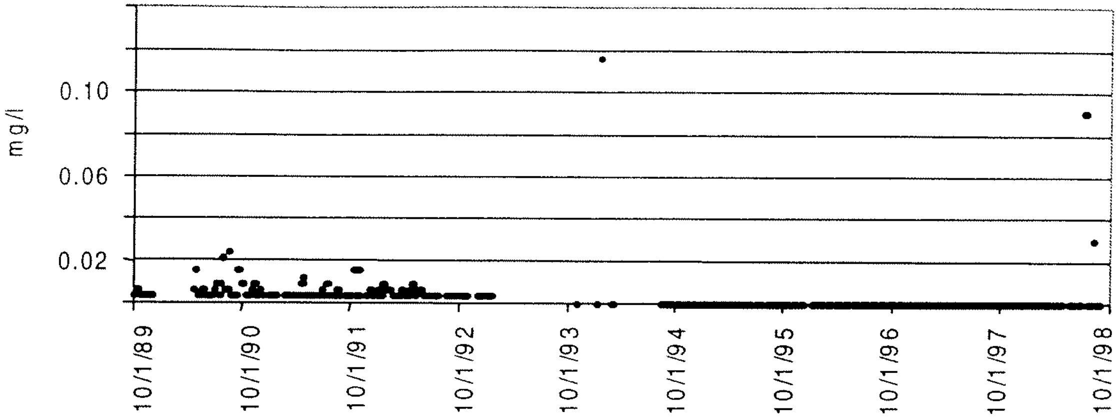


Figure 21. Silica.

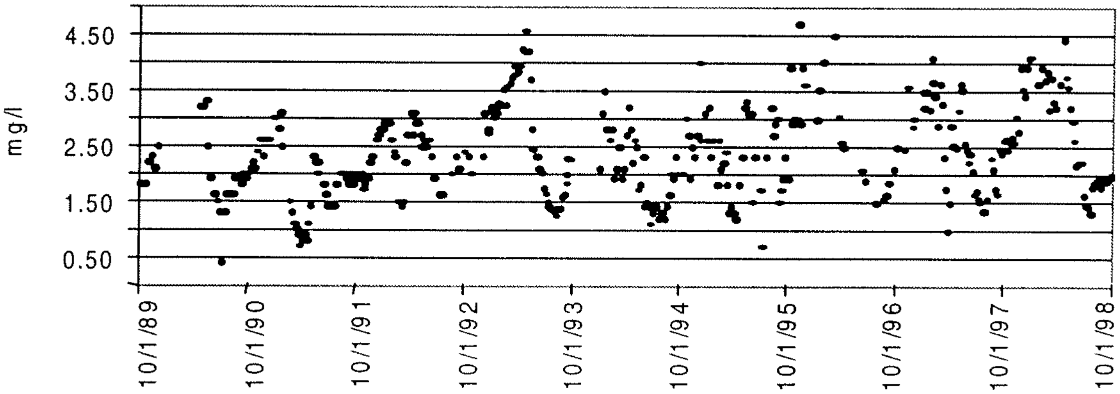
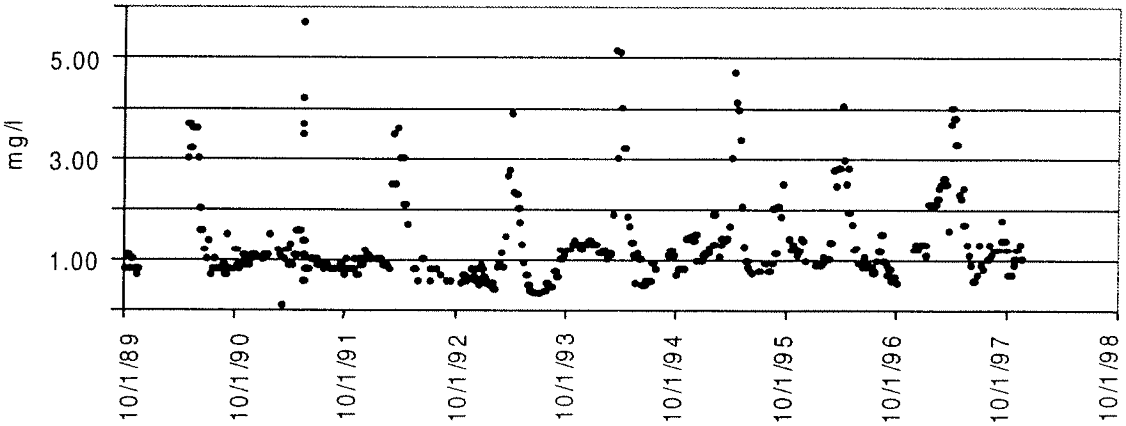


Figure 22. Dissolved Organic Carbon.



Hydrology

The outlet of the Loch is located at the northeastern edge of the basin, and has been gaged with a Parshall flume and stilling well since 1983. The accuracy of a Parshall Flume is ~ 5% (Winter, 1981). Prior to 1994, it was estimated that an additional 5% uncertainty was introduced by water flowing around or under the frame structure. Frequent examination of the flume's condition reveals that, since some minor repairs in 1994, flow around and under the flume is insignificant. A recent low-water inspection of the inside of the flume revealed that it is still of its original dimensions (extremely important when calculating flow volumes and rates), and that it is structurally sound.

Until 1997, stage height data were collected both mechanically (with a Leupold & Stevens chart recorder) and electronically (with an Omnidata DP115 datalogger). Necessary corrections to DP115 data (due to equipment and/or operator error) were executed and noted after reducing raw data from the data storage module (DSM). The data were then run through a Fortran program that calculates discharge as a flow rate of cubic meters/second (CMS) using a formula provided by the Thompson Pipe and Steel Co.

As a quality spot-check of flow data, weekly checks of actual stage height are recorded after observing a staff gage located just inside the flume as flow and snow permits. Stage height is converted to flow rate, and compared to the mechanically and electronically collected values.

Use of the Omnidata DP115 was discontinued at the end of the 1997 water year. This decision was due to a number of factors: 1) its overly cumbersome method of operation, 2) the many possible sources of error introduced in the process of reducing discharge data from the DP115's DSM, and 3) the impossibility of attaining a proper offset factor from the 1997 data.

The 1997 and 1998 hydrographs were created by reading instantaneous values at maximum and minimum flow each day from the Leupold & Stevens chart recorder stripcharts. These values were averaged into daily mean CMS.

Before accepting this method of estimating flow, and including the data in our database, its validity was put to test. Several intervals of stripchart daily averages were compared to stage data gathered via the DP115. In light of some significant differences found, the next step was to recalculate all flow from all years using the L&S stripcharts to make certain that historical data was correct. Years 1984 – 1991, 1993, and 1996 matched very well (T-test p values ranging from 0.85 to 0.93). However, there were significant differences in 1994 (p=0.008) and 1995 (p=0.026) between calculated (DP115) and observed flow (from charts and fieldbooks). These errors resulted in a serious underestimation of flow (Figures 23 and 24).

Figure 23. Comparison of 1994 DP115 to L&S stripchart data.

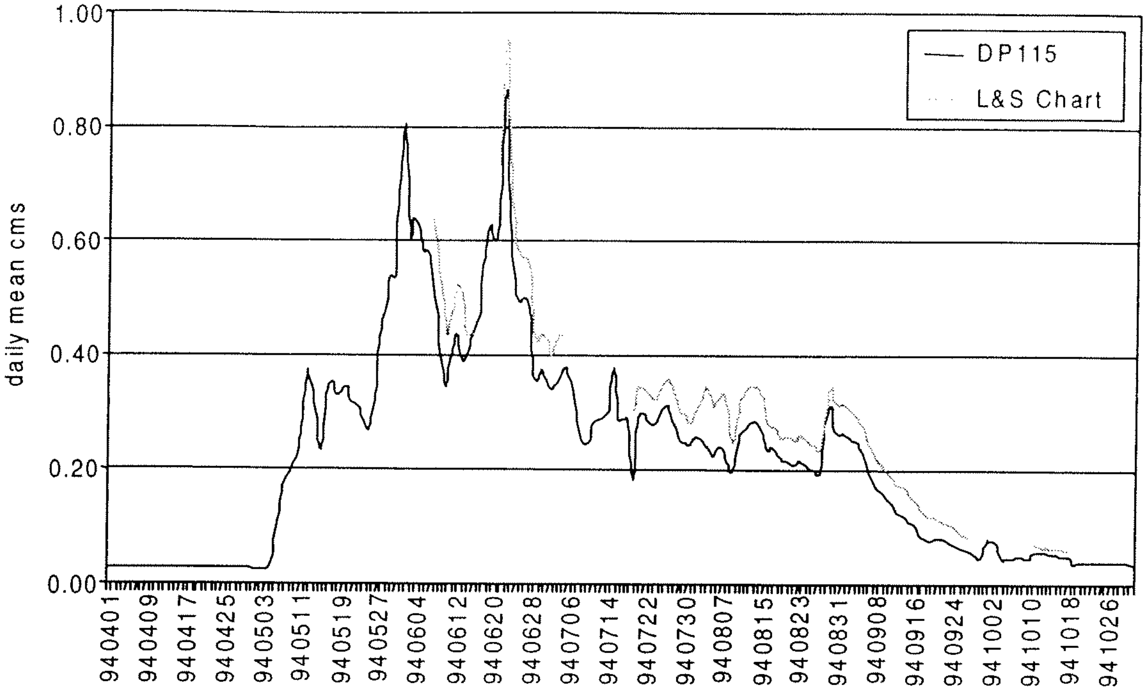
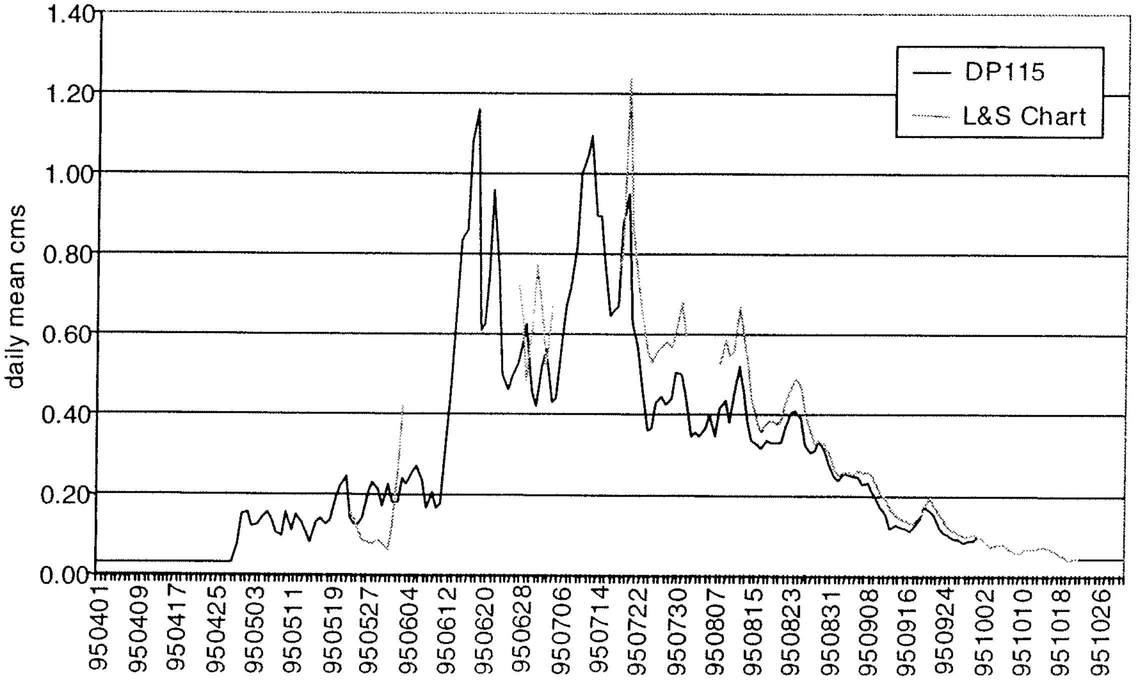


Figure 24. Comparison of 1995 DP115 to L&S stripchart data.



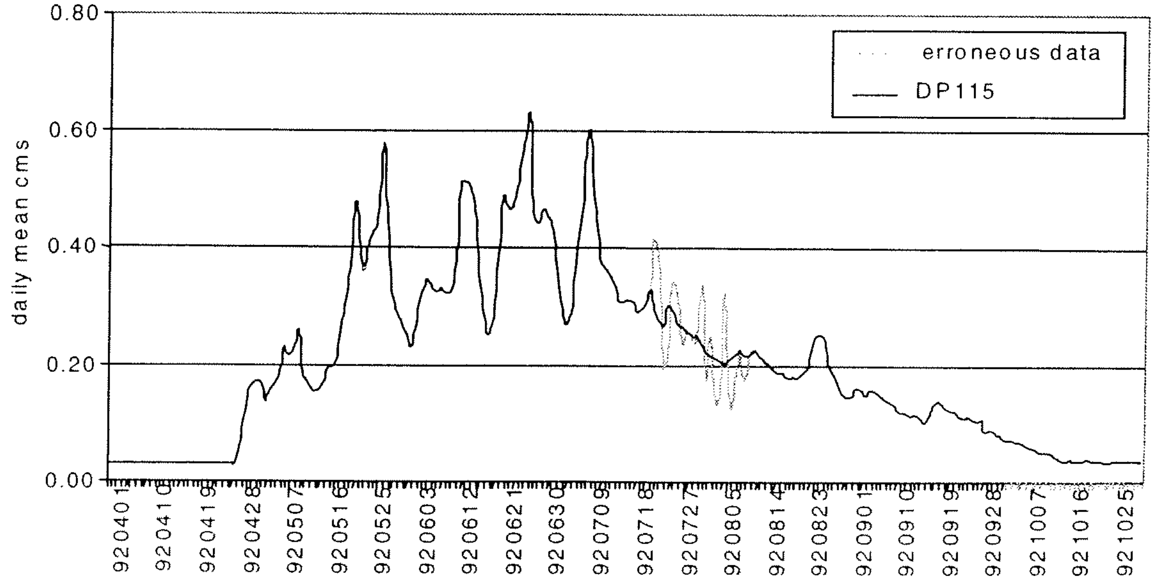
A unique adjustment factor for each erroneous period during 1994 and 1995 was calculated by subtracting DP115 flow data from L&S chart data. Each factor was then added to its respective erroneous interval. This corrected flow was then rechecked against data from field notebook and the L&S charts. A new master flow file was created using all the corrected data and imported into the Loch Vale database.

In discussion following these actions, the question was raised as to whether or not two values per day would lend an accurate daily average. Ten days of flow from 1996 (evenly distributed throughout the hydrograph) were calculated using different numbers of hourly measurements for averages. The most measurements used in a single day's average were twelve. Averages built on twelve observations per day were not significantly different from those built on two (max/min) (p= 0.83). Because maximum flow usually occurs within an hour of midnight, and minimum flow usually occurs within an hour of noon, there was no significant difference between max/min and midnight/noon averages (p=0.96). This is useful because it is easier to work through the stripcharts using midnight to noon values that it is to use maximum and minimum values.

The problems with the DP115 were caused by its inability to "remember" the proper offset adjustment. The Fortran code that reduces flow data from the raw DSM output required a single factor to act as an offset adjustment (equivalent to zeroing a balance). When the DP115 was new, one offset factor could be entered for the entire year. As the datalogger aged, a drift developed in the logger's memory that made offset change from week to week, or in some cases from day to day. This made it impossible to acquire one, or even several factors to input to the Fortran reduction code. Although there is no documentation of the actions taken, a past data manager spent several weeks with the 1996 hydrograph data. It is believed that he went through similar contortions (as previously described).

There was an erroneous period in 1992 of unknown origin. Data from this period was substituted with stripchart data (Figure 25) and entered into Oracle.

Figure 25. Spurious 1992 data.



A number of errors were identified in the flow record from 1984–1991, 1993, and 1996. Although none of these errors had significant bearing on the total estimate of yearly efflux, they were changed in the database and documented in the archive of yearly flow files.

A Campbell Scientific CR500 datalogger and an EnviroSystems shaft encoder will be installed in a new stilling well at the onset of flow in 1999. This recorder will be much easier to operate than the DP115. The CR500 will provide instantaneous, hourly, and daily flow measurements without having to manipulate any data. Extensive testing and calibration of the CR500 has been performed to ensure the quality of all data collected.

General Conclusions

In general, quality of LVWS data was good through the period of analysis. A summary of all actions taken is presented in Table 13.

Table 13. Summary of quality and actions taken.

	Quality	*Action Taken*
Met Station	Insufficient/failing.	decommissioned Nov. 24, 1998 switched to WRD station
Aerochem	Sensor not responding (November 98).	replaced regulator, added 80amp/hr battery replaced motorbox replaced sensor
Belfort Raingages	Non-standard (Nipher) wind-drop shield.	installed standard (Alter) shield
	Missing points from NADP.	Combined all precip data from FORFs, charts, and database.
Stream Gage/Flow	Underestimates/spurious points.	Recalculated discharge 1984-1997.
	Omnidata datalogger bad.	Installed new Campbell Sci. CR500.
NREL/LVWS Laboratory	Duplicate analysis- good.	N/A
	Blank analysis- good.	N/A
	Si contamination in DI (8/19/96 - 9/11/96).	Imported clean DI water from USFS lab. DI columns were replaced.
	pH probe failed.	Replaced November 1998.
	pH meter failed.	Replaced January 1999.
USFS Laboratory	Water quality- good.	N/A

In spite of some major equipment failures, no data was lost due to backup devices. Decommissioning the Omnidata stream recorder and RAWS were important steps to insure the quality of future data collected. The new systems should be faster, more reliable, and less expensive to use/maintain.

Water quality measurements were very accurate, with the exception of a short period when our DI showed signs of SiO_2 contamination. The water quality database was examined virtually point by point. Several points were either corrected or deleted. This, combined with refinements to the streamflow and precipitation records, allow us the ability to construct the most accurate chemical input/output budgets for LVWS to date.

References

Allstott, E.J. *Quality Assurance Report: 1991-1994*, Loch Vale Watershed Project. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO. (1995).

Aubertin, G.M.; Bigelow, D.S.; Malo, B.A. *Quality Assurance Plan NADP/NTN Deposition Monitoring*. National Atmospheric Deposition Program. (1990)

Bigelow, D.S.; Dossett, S.R. *Instruction Manual NADP/NTN Site Operation*. NADP/CAL, (1988- updated frequently)

Bigelow, D. S.; Denning, S. A.; Baron, J. *Differences between Nipher and Alter Shielded Universal Belfort Precipitation Gages at Two Colorado Deposition Monitoring Sites*. Environ. Sci. and Tech. 1990, 24, 758-760.

Denning, A.S. *Quality Assurance Report: 1982-1987*. Loch Vale Watershed Project. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO. (1988).

Edwards, R. *Quality Assurance Report: 1989-1990*, Loch Vale Watershed Project. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO. (1991).

Goodison, B. E.; Turner, V. R.; Metcalfe, J. R. Proceeding: *Conference on Meteorological Observation and Instrumentation*; American Meteorological Society: Boston, MA, 1983; pp. 21-26.

Goodison, B. E.; Vet, R. J. In Atmospheric Deposition; Delleur, J. W., Ed.; IAHS Publication No. 179; International Association of Hydrological Sciences: 1989; pp. 47-55.

MacTaggart, D. Guest Editorial: *Realistic detection limits*. American Environmental Laboratory, 4:10/98, 1998.

Newkirk, B.K. *LVWS Methods Manual*. Loch Vale Watershed Project. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO. (1995).

USGS Open-File Reports 96-138. Results of the US Geological Survey's Analytical Evaluation Program for Standard Reference Samples. (1996)

USGS Open-File Reports 96-436. Results of the US Geological Survey's Analytical Evaluation Program for Standard Reference Samples. (1996)

USGS Open-File Reports 98-52. Results of the US Geological Survey's Analytical Evaluation Program for Standard Reference Samples. (1998)

USGS Open-File Reports 98-391. Results of the US Geological Survey's Analytical Evaluation Program for Standard Reference Samples. (1998)

U.S. Environmental Protection Agency, *Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans*, Office of Monitoring Systems and Quality Assurance, Office of Research and Development, QAMS-005/80, Washington, DC (1980).