

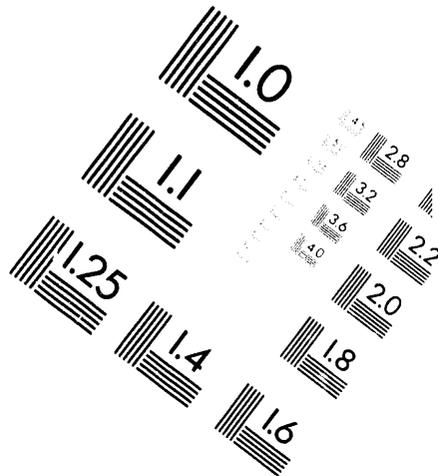
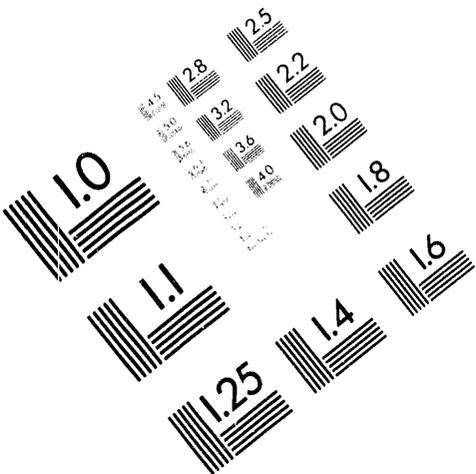


AIM

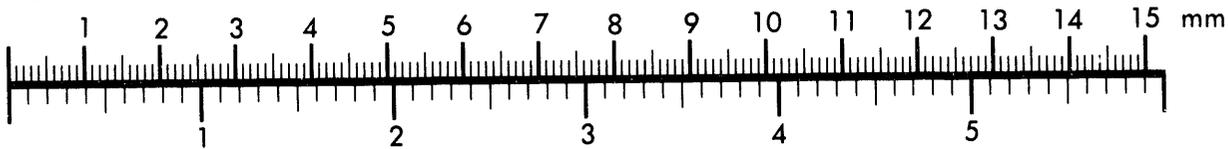
Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910

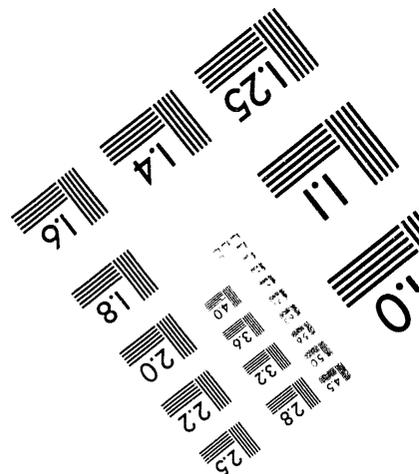
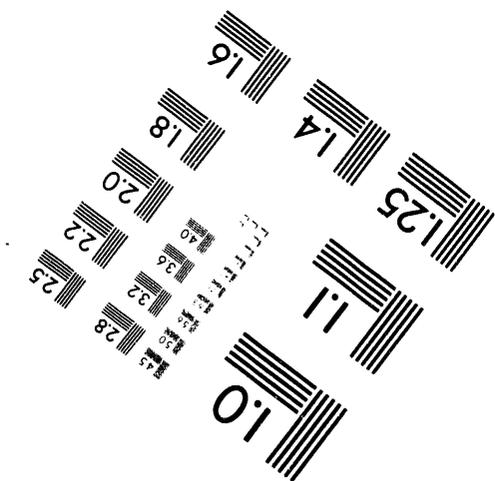
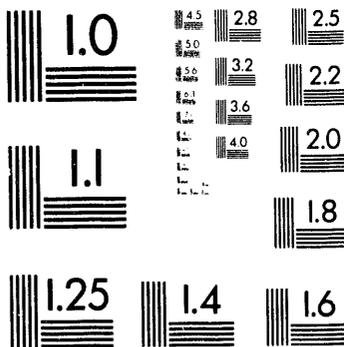
301/587-8202



Centimeter



Inches



MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

Some Statistical Aspects of Background Based Groundwater Standards at an Arid Hazardous Waste Site

C. J. Chou
F. N. Hodges
V. G. Johnson

Date Published
July 1994

To Be Presented at
Institute of Nuclear Materials
Management Annual Meeting
Naples, Florida
July 17-20, 1994

Prepared for the U.S. Department of Energy
Office of Environmental Restoration and
Waste Management



Westinghouse
Hanford Company

P.O. Box 1970
Richland, Washington

Hanford Operations and Engineering Contractor for the
U.S. Department of Energy under Contract DE-AC06-87RL10930

Copyright License By acceptance of this article, the publisher and/or recipient acknowledges the
U.S. Government's right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

Approved for Public Release

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

SOME STATISTICAL ASPECTS OF BACKGROUND BASED
GROUNDWATER STANDARDS
AT AN ARID HAZARDOUS WASTE SITE

V. G. Johnson, C. J. Chou & F. N. Hodges
Westinghouse Hanford Company
Richland, WA 99352

ABSTRACT

Statistical goodness-of-fit tests and "Box and Whisker" plots of hydrochemical data from selected contaminant-free downgradient wells, and wells located upgradient in a non-contaminated or background area show that spatially distinct sample populations do not exhibit significant differences in groundwater chemical composition within the upper unconfined aquifer. Well location dominates natural constituent variability at this arid site. Spatial coverage should be emphasized in such cases rather than sampling frequency.

INTRODUCTION

Background hydrochemical data is needed for cleanup standards or guidelines and establishment of alternative enforcement levels. However, the location of a "representative" background area at large chemical and nuclear waste sites with widespread existing groundwater contamination often complicates this objective. At the Hanford Site in south-central Washington State, widespread contaminant plumes overlap, making selection of local area background wells difficult.

More distant upgradient areas, while free of contaminant plumes, may not be representative because of time or distance-dependent variations in geochemical conditions. Thus, the

objectives of this study were 1) to determine if there is a significant difference between spatially distinct hydrochemical sample populations in the unconfined aquifer and 2) to assess the relative importance of sampling depth, frequency, and well location. Such factors are fundamental to development and/or application of background-based standards.

STUDY AREA

The Hanford Site is located in a "mid-latitudinal desert" on the Columbia Plateau (Figure 1). Aridity is due to the rain shadow effect of the Cascade Mountains. Potential evapotranspiration exceeds the average annual precipitation of 16 cm/yr. Sagebrush and cheatgrass are the dominant plants.

The unconfined aquifer occurs in glacial flood deposits and semi-cemented fluvial deposits. The average depth to groundwater is 70 m in the vicinity of and to the west of the 200 West and 200 East areas, dropping to ~20 m near the river. The average thickness of the unconfined aquifer is about 80 m.

Precipitation on basalt ridges bounding the basin is the principal source of natural recharge. The total natural recharge is uncertain, however, surface runoff in Cold Creek and Dry Creek accounts for approximately 2 cfts (Figure 1). The composition of recharge water is principally controlled by two processes: 1) reaction of CO₂

charged precipitation with basalt; and 2) variable degrees of evaporation. Chemical precipitation, ion-exchange, and absorption of CO₂ can further alter composition over time and distance. Carbon-14 in springs and unconfined aquifer monitoring wells along the western margin of the Site indicate groundwater residence times in the hundreds to a few thousand years, consistent with arid site recharge conditions. Thus, changes in the natural chemical composition of groundwater with distance was expected.

STATISTICAL ANALYSIS TECHNIQUES

Kolmogorov-Smirnov (K-S) Two-Sample Test. The null hypothesis (H_0), whether spatially distinct subsets or populations are identical within the upper unconfined aquifer, was tested using a U.S. Geological Survey data set acquired during 1979-1984. The K-S two-sample test was conducted on subsets of major and trace constituents from wells shown in Figure 1. These wells were selected from a data set consisting of approximately 100 wells located across the Hanford Site. Only data from those wells (42) drilled within the upper unconfined aquifer and with tritium concentrations of less than 1,000 pCi/l were selected for the comparisons. Three subsets were identified (Figure 1): (1) a Rattlesnake Mountain subset (10 wells) located along the western side (i.e., upgradient) of the Hanford Site; (2) an area-wide subset (32 wells) located across the Hanford Site; (3) a Gable Mountain north subset (10 wells), a subset of (2), representing an area far downgradient from the recharge location in upper Cold Creek Valley.

The K-S test is preferred over other tests such as the median test, the Mann-Whitney test, or the

parametric t-test because these tests are sensitive only to differences between the two means or medians. The K-S test is sensitive against all types of differences (e.g., differences in variances) that may exist between the two distribution functions.

The test is conducted as follows. Let $S_1(x)$ and $S_2(y)$ be the empirical distribution functions based on the random sample X_1, X_2, \dots, X_n and the other random sample Y_1, Y_2, \dots, Y_m , respectively. For a two-sided test, the test statistic is the maximum vertical distance between the plots of two empirical distribution functions. The decision rule is to reject H_0 at the level of significance α , if the test statistic exceeds its $1-\alpha$ quantile as provided in Conover (1980).

Results of comparison of the Rattlesnake Mountain subset vs area-wide and Rattlesnake Mountain vs Gable Mountain North are presented in Figure 2. The test results strongly suggest there is little, if any, difference between upgradient locations as represented by the Rattlesnake Mountain subset and downgradient locations.

Analysis of Variance (ANOVA) with the Nested Random Effects Model.

In order to assess the relative importance of sampling frequency and well location in characterizing background, the ANOVA procedure with a nested random effects model was used to analyze the sources of data variability arising from spatial, temporal, and analytical factors. The test data for this evaluation consisted of hydrochemical results for seven widely spaced wells from the Rattlesnake Mountain subset (Figure 1) that were sampled quarterly and in duplicate. The nested model is used when background wells are sampled at different

times. The general form of this model is given by:

$$Y_{ijk} = \mu + W_i + T_{(i)j} + \epsilon_{(ij)k}$$

$$i = 1, 2, \dots, a,$$

$$j = 1, 2, \dots, b_i,$$

$$k = 1, 2, \dots, n_{ij},$$

where Y_{ijk} denotes the k^{th} analysis on the j^{th} sampling time from the i^{th} upgradient well; μ is the unknown true mean level; W_i denotes the true effect of the i^{th} well and is assumed to be a random variable (rv) from a population with zero mean and with variance σ_W^2 ; $T_{(i)j}$ is the true effect of the j^{th} sampling time within the i^{th} well and is assumed to be a rv from a population with zero mean and with variance σ_T^2 ; and $\epsilon_{(ij)k}$ is the analytical error associated with the k^{th} analysis of samples collected from well i at time j . $\epsilon_{(ij)k}$ is assumed to be a rv from a population with zero mean and with variance σ_A^2 .

It is also assumed that the W_i are uncorrelated, the $T_{(i)j}$ are uncorrelated, and the $\epsilon_{(ij)k}$ are uncorrelated, and there are no correlations among the W_i , $T_{(i)j}$, and $\epsilon_{(ij)k}$. Thus, y_{ijk} is an observation from a population with mean μ and with variance $\text{Var}(Y_{ijk}) = \sigma_W^2 + \sigma_T^2 + \sigma_A^2$. It is of interest to estimate the variance components σ_W^2 , σ_T^2 , and σ_A^2 . One may use the formulas and the nested random effects ANOVA table provided in Bowen & Bennett (1988) to estimate σ_W^2 , σ_T^2 , and σ_A^2 . The results (Table 1) show that spatial variability is the most dominant factor in the total variance components while seasonal and analytical variations are minimal. In such cases, the most effective way to reduce the uncertainty in estimating background is to increase the spatial coverage by sampling more upgradient wells. To offset the cost, one can reduce the sampling frequency and reduce the number of replicate analyses,

especially when the cost associated with each chemical analysis is high.

Box and Whisker Plots. One simple method of displaying data is a box and whisker plot. In this plot, the upper (Q_3) and lower (Q_1) quartiles of the data are shown by the top and bottom of a box and the median (Q_2) is indicated by a line segment within the box. The box covers the middle 50% of the data values. The whiskers extend only to those points that are within 1.5 times the interquartile range ($Q_3 - Q_1$). When extremely large or small values occur, they are plotted as individual points.

Multiple box and whisker plots may be used to show the changes in chemical composition through time or over space. This is accomplished by displaying chemical concentration data for a particular analyte of concern over different regions. An illustration using redox sensitive species for the four data subsets previously described is shown in Figure 3. The elevated iron and depressed uranium (based on gross alpha) for region 4 (the "deep" unconfined aquifer wells located within the 200 West waste management area, Figure 1), are due to chemical reducing conditions (i.e. the absence of oxygen and negative oxidation-reduction potentials). This demonstrates the influence of sampling depth or redox status of the aquifer. Care must be taken not to "mix" data from deep and shallow wells (or data from oxic and anoxic zones) where redox sensitive analytes are concerned.

DISCUSSION

The absence of statistically significant differences in constituent concentrations between the upgradient and downgradient wells suggests groundwater chemical

composition is set early in its evolutionary history (i.e., at or near the recharge area). Subsequent hydrochemical changes are apparently minor or secondary. Thus, major portions of the unconfined aquifer across the Hanford Site should exhibit similar "natural" concentration ranges of major and trace constituents. Upgradient wells, the preferred location for background characterization purposes, should therefore be representative of more distant downgradient locations. In addition, existing wells downgradient of contaminant source areas, and outside of contaminant plume boundaries, may be used to increase spatial coverage. Since spatial variation accounts for most of the overall variability (Table 1), increasing spatial coverage results in more reliable background comparison tests. For this purpose, use of the entire concentration-distribution function (e.g., K-S test) has been recommended over the threshold value calculated from background measurements (Chou et al, 1993). In addition to use of the K-S test, as discussed in this paper, the Wilcoxon Rank Sum (WRS) test and/or Quantile test may be used (Gilbert and Simpson, 1992).

Additional sampling is currently underway to increase spatial coverage, to update analytical results and to extend the number of analytes of regulatory interest (Hoover, 1994).

CONCLUSIONS

Hydrochemical data from selected downgradient wells and wells located upgradient in a non-contaminated or background area show no significant differences in groundwater chemical composition. Quarterly time-series data suggest there are no observable seasonal variations. Well location

dominates natural constituent variability within the oxygenated zone of the unconfined aquifer at this arid site where natural recharge is slow and the unsaturated zone is relatively thick. In such cases, spatial coverage should be emphasized rather than sampling frequency, and quarterly sampling may be unnecessary to define natural background for remediation or compliance testing purposes.

REFERENCES

1. Bowen, W. M. and C. A. Bennett, 1988, *Statistical Methods for Nuclear Material Management*, NUREG/CR-4604, PNL-5849, U.S. Government Printing Office, Washington, DC., pp. 282-299.
2. Chou, C. J., V. G. Johnson, and F. N. Hodges, 1993, *Statistical Application of Groundwater Monitoring Data at the Hanford Site*, WHC-SA-1792-FP, Westinghouse Hanford Company, Richland, Washington.
3. Conover, W. J., 1980, *Practical Nonparametric Statistics*, Second Edition, John Wiley and Sons, Inc., New York, New York, pp. 369-373.
4. Gilbert, R. O. and J. C. Simpson, 1992, *Statistical Methods for Evaluating the Attainment of Cleanup Standards*, PNL-7409, Vol. 3, Rev. 1, Richland, Washington.
5. Hoover, J., 1994, *Background Activities at the Hanford Site*, 6th National Tech. Information Exchange (TIE) Workshop, May 10-12, Kennewick, Washington.

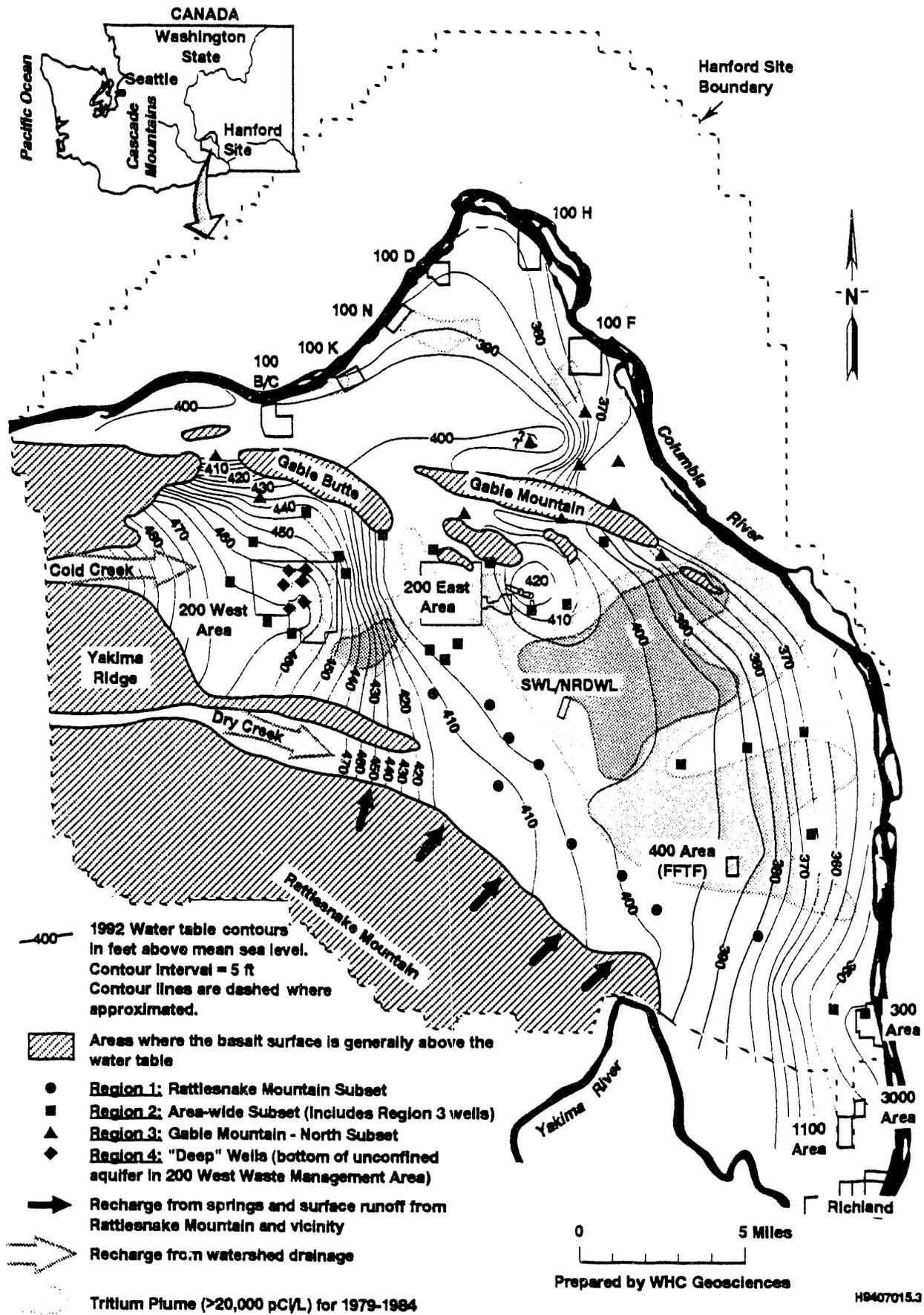


Figure 1. Site map showing location of monitoring wells in relation to nuclear waste storage and disposal sites. Generalized groundwater flow directions are perpendicular to water table elevation contours.

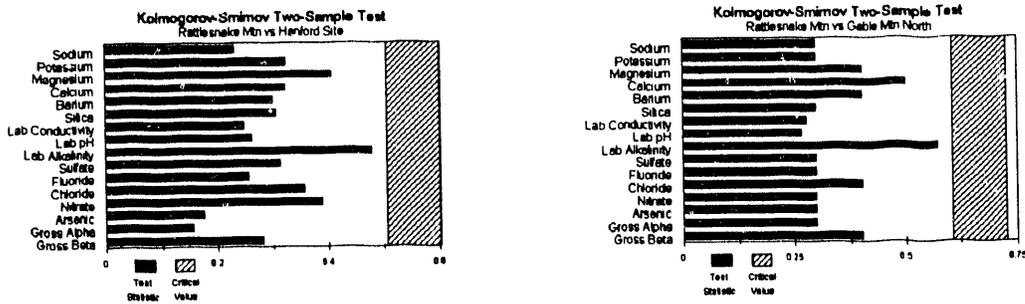


Fig. 2. Kolmogorov-Smirnov two sample test for Rattlesnake Mountain data subset vs site-wide subset and Rattlesnake Mountain vs Gable Mountain north subset. See Fig. 1 for subset well locations.



Fig. 3. Box and whisker plots for selected redox sensitive analytes from the four regions shown in Fig. 1. Gross alpha is used as an indicator for uranium, the primary naturally occurring alpha emitter in the unconfined aquifer. Wells for region 4 are completed at the bottom of the unconfined aquifer (ca. 80 m below the water table elevation). All other data are from wells completed at the top of the aquifer (0 - 10 m).

Table 1. Estimates of Variance Components in Percentages Using Data from Wells Located in the Rattlesnake Mountain Subset.

Variance Components	Sodium (%)	Potassium (%)	Magnesium (%)	Calcium (%)	Silica (%)
Spatial	99.40	97.66	96.82	98.65	98.74
Temporal	0.58	2.19	3.05	1.29	1.24
Analytical	0.02	0.15	0.13	0.06	0.02
Total	100.00	100.00	100.00	100.00	100.00
Variance Components	Lab Conductivity (%)	Lab pH (%)	Lab Alkalinity (%)	Sulfate (%)	Fluoride (%)
Spatial	97.86	81.57	98.35	99.73	81.10
Temporal	0	11.53	0	0.10	0
Analytical	2.14	6.90	1.65	0.17	18.90
Total	100.00	100.00	100.00	100.00	100.00
Variance Components	Chloride (%)	Field Alkalinity (%)	Field pH (%)	Strontium (%)	
Spatial	99.75	97.92	63.93	98.79	
Temporal	0.22	1.38	0	0.64	
Analytical	0.03	0.70	36.07	0.57	
Total	100.00	100.00	100.00	100.00	

**DATE
FILMED**

8/25/94

END

