1 of 1

2

Conf-940748--87

WSRC-MS-94-0111

Estimating an Appropriate Sampling Frequency for Monitoring Ground Water Well Contamination

by

R. C. Tuckfield

Westinghouse Savannah River Company
Savannah River Site
Aiken, South Carolina 29808

A document prepared for INTERNATIONAL NUCLEAR MATERIALS MANAGEMENT (INMM) ANNUAL MEETING 1994 at Naples from 07/17/94 - 07/20/94.

DOE Contract No. DE-AC09-89SR18035

This paper was prepared in connection with work done under the above contract number with the U. S. Department of Energy.

ESTIMATING AN APPROPRIATE SAMPLING FREQUENCY FOR MONITORING GROUND WATER WELL CONTAMINATION

R. C. Tuckfield
Westinghouse Savannah River Company
Savannah River Technology Center
Applied Statistics Group
Aiken, SC 29802

ABSTRACT

Nearly 1500 ground water wells at the Savannah River Site (SRS) are sampled quarterly to monitor contamination by radionuclides and other hazardous constituents from nearby waste sites. Some 10,000 water samples were collected in 1993 at a laboratory analysis cost of $10,000,000. No widely accepted statistical method has been developed, to date, for estimating a technically defensible ground water sampling frequency consistent and compliant with federal regulations. Such a method is presented here based on the concept of statistical independence among successively measured contaminant concentrations in time.

INTRODUCTION

All custodians of go· ernment and commercial waste units permitted under the federal legislation called the Resource Conservation and Recovery Act (RCRA) are required to monitor the groundwater surrounding the waste unit. Groundwater samples are collected on a scheduled basis for the "life" of that waste unit and are sent to a laboratory for chemical analysis.

The RCRA legislation 40 CFR Part 264, Subpart F requires that monitoring wells be drilled both upgradient and downgradient to the waste unit groundwater flow. The upgradient wells represent background conditions.

During the initial or detection phase of waste unit operation, groundwater samples are collected regularly, usually quarterly but as often as "...four times semiannually...", from the upgradient wells. After at least four concentration measurements have been obtained for a given constituent, statistical comparisons are made between this set of "background" measurements and a similarly set of subsequent concentration measurements for the same constituent in each downgradient well. This two-sample statistical inference is designed to detect a significant exceedance of background conditions - i.e. a leaking waste site.

The basis for making credible two-sample inferences between upgradient and downgradient groundwater conditions, rests with the credence of the most sensitive assumption for such tests, viz., that sample measurements are statistically independent. Since all samples are collected as a time series at any given well, preserving statistical independence is a matter of choosing an appropriate sampling frequency.

The SRS annually collects about 10,000 groundwater monitoring samples and spends nearly $10 million on laboratory analyses. Presenting a sampling plan to federal and state regulators that meets regulatory intent while reducing the taxpayer expense is an aim of the Westinghouse, Savannah River Technology Center (SRTC).

This paper presents a method for selecting a technically defensible and cost effective sampling frequency based on the geostatistical concept of a temporal variogram.

STATISTICAL MODEL

Recent guidelines (EPA 1989, 1992) comprehensively review and present a broad range of statistical methods for assessing groundwater contamination impacts at RCRA waste units. Section 2.4.2 of the former of these two documents offers sampling procedure guidance as

> "Obtain a sequence of at least four samples taken at an interval that ensures, to the greatest extent technically feasible, an *independent* sample is obtained, by reference to the uppermost aquifer's effective porosity, hydraulic conductivity, and hydraulic gradient and the fate and transport characteristics of potential contaminants."

This document further states that

> "...a sequence of at least four samples taken at intervals far enough apart in time (daily, weekly, or monthly...) will help ensure the sampling of a discrete portion (i.e. an *independent* sample) of groundwater."

The former of these two guidance paragraphs suggests the use of the Darcy Equation (EPA 1989), a hydro-geological engineering model which when applied will ensure that samples are taken form completely separate units of groundwater flowing past the well head. However, applications of the model often yield unreasonable (i.e. inordinately expensive) sampling frequencies such as every 10 days.

A geostatistical method offers an alternative to the Darcy Equation. This method is called a temporal variogram. The variogram ($\hat{\gamma}(h)$) of a spatial process consists of the sum of successive squared differences among a specific linear transect or series of measurements taken h measurement units apart over a two-dimensional area (Cressie 1993, Issaks and Srivastava 1989). Analogously, the variogram of a temporal process or time series $\{X_t, X_{t+1}, \ldots, X_{t+N}\}$ is given as

$$\hat{\gamma}_t(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (X_i - X_{i+h})^2 \qquad (1)$$

for each sequence of paired differences h time units apart.

Like its spatial counterpart, commonly used in mining and geoscience, the temporal variogram seeks to estimate the variance of statistically independent process measurements.

An important relation exists between the variogram and the covariance ($C_t(h)$) of successive measurements , h time units apart. This relation is given as

$$\begin{aligned} \hat{\gamma}_t(h) &= C_t(0) - C_t(h) \\ &\cong \sigma_t^2 \quad \forall h \geq range \end{aligned} \qquad (2)$$

where

$$\begin{aligned} C_t(h) &= E\{(X_t - E(X_t))(X_{t+h} - E(X_{t+h}))\} \\ &= E\{(X_t)(X_{t+h})\} - \mu_t^2 \end{aligned} \qquad (3)$$

σ_t^2 is the variance of statistically independent measurements in the time series and the *range* is that value of h such that the covariance among measurements h or more units apart is asymptotically zero. E{.} is the expectation operator and denotes an average.

Typically, a plot is constructed depicting $\hat{\gamma}_t(h)$ as a function of h, $f(h)$, and whose points approach an apparent maximum value (σ_t^2) called the sill. The practical range (h_p) is sometimes (Issaks and Srivastava 1989) calculated as

$$f(h_p) = .95\hat{\sigma}_t^2 \qquad (4)$$

where

$$\hat{\sigma}_t^2 = \lim_{h \to \infty} \hat{\gamma}_t(h) \qquad (5)$$

i.e. the asymptote of the monotonically increasing function, $f(h)$.

Useful mathematical models of $f(h)$ are reviewed here that result in point or interval estimates of h_p, based on simulated time series data. Also presented is a preliminary examination of the effects of sample size on h_p estimates.

DATA SIMULATION

Two times series (X and Y) were generated using the ARMASIM function of the SAS/IML™ software. These data were simulated by first specifying an autocorrelation function of order 6 and using the Yule-Walker equations (Box and Jenkins 1976) to solve for the parameter estimates for the associated autoregressive time series model. This AR(6) model for the X-series is given as

$$X_t = \sum_{h=1}^{6} \varphi_h (X_{t-h}) + a_t \qquad (6)$$

and applies similarly in the generation of the Y-series but based on a different autocorrelation function.

The φ_h parameters are used as inputs to the ARMASIM function to generate 100 observations with a mean of 25 and a standard deviation of 5 for each series.

Table 1 presents the autocorrelation functions (ACF) used to produce each time series.

Table 1. X and Y autocorrelation functions (ACF), of order 6.

Lag (h)	X-series ACF	Y-series ACF
1	.80	.90
2	.64	.81
3	.41	.66
4	.17	.43
5	.03	.19
6	.00079	.03

Figure 1 is a plot of both simulated time series.

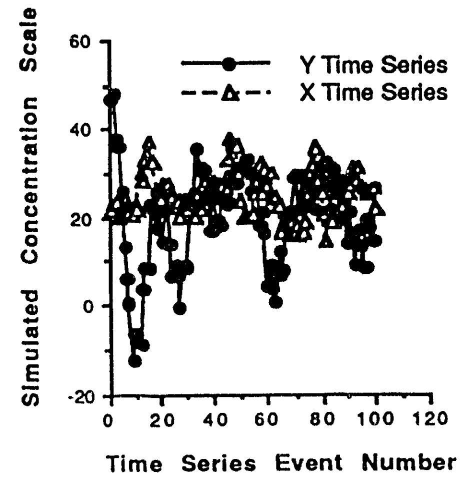

Figure 1. Plot of two simulated time series (X and Y) with a mean and variance of 25.

RESULTS

Temporal variograms were calculated for each series. Three typical models were fit to these variograms, viz., the spherical model, the exponential model and the gaussian model (see Issaks and Srivastava 1989). These models were fit using the iterative Gauss-Newton procedure of the nonlinear analysis platform in the JMP™ software.

In each case the gaussian model

$$\hat{\gamma}_t(h) = \beta_0 + \beta_1\left(1 - e^{-3\left(\frac{h}{\beta_2}\right)^2}\right) + \varepsilon_h \qquad (7)$$

produced the best fit. In the parlance of geostatistics, the $\{\beta_0, \beta_1, \beta_2\}$ parameters are the nugget, sill and range, respectively, of the variogram.

Figures 2 and 3 show the fit of the gaussian models for the X and Y time series using eq 7.

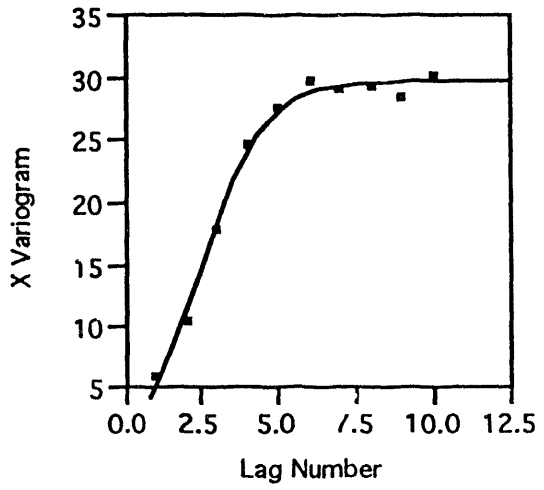

Figure 2. Gaussian model fit to the variogram based on the simulated X-time series

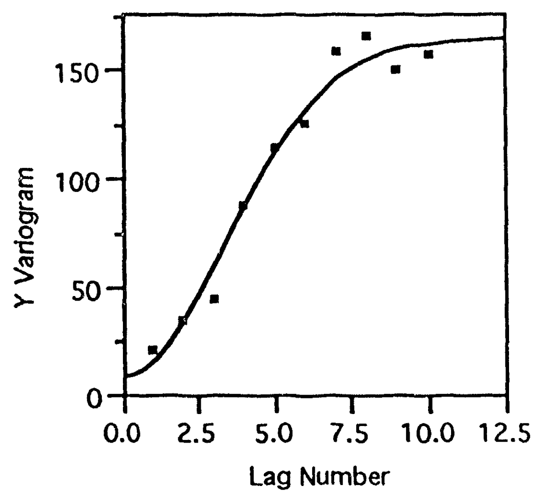

Figure 3. Gaussian model fit to the variogram based on the simulated Y-time series.

Tables 2 and 3 present the parameter estimates and associated 95% confidence litmus for the X- and Y-series variograms.

Table 2. Point and 95% confidence interval estimates for parameters of a gaussian model fit to the X-series variogram.

Param -eter	Estim- ate	Apprx StdErr	Lower CL	Upper CL
$\hat{\beta}_0$	2.6506	0.9336	0.4508	4.7600
$\hat{\beta}_1$	27.062	0.9389	24.888	29.288
$\hat{\beta}_2$	5.5331	0.2246	5.0567	6.0613

Table 3. Point and 95% confidence interval estimates for parameters of a gaussian model fit to the Y-series variogram.

Param -eter	Estim- ate	Apprx StdErr	Lower CL	Upper CL
$\hat{\beta}_0$	8.9263	8.7836	-11.791	28.484
$\hat{\beta}_1$	155.79	9.7711	132.72	178.94
$\hat{\beta}_2$	8.3499	0.7638	6.9133	10.389

Both the exponential and spherical models produced negative estimates of the nugget parameter. In addition, the exponential model did not produce a reasonable sill parameter within h = 10.

To examine the effects of sample size specifically on range ($\hat{\beta}_2$) estimation, the gaussian model was also used to fit variograms from the first 50 and the first 25 observations in both the X-series and Y-series. Table 4 presents the results from these analyses.

Table 4. Variogram (nugget, sill and range) parameters from gaussian model fits as a function of samples size (N) from the Y-series.

Series	N	Nugget	Sill	Range
X	100	2.65	27.1	5.53
X	50	6.30	15.9	8.01
X	25	4.82	29.3	9.06
Y	100	8.93	155.8	8.36
Y	50	0.78	245.2	7.76
Y	25	-7.36	472.4	8.36

DISCUSSION

It can be seen from the results that fitting nonlinear models to temporal variograms can produce reasonable estimates of a sampling frequency for hydro-geological processes such as groundwater monitoring. The simplest of nonlinear fitting algorithms (i.e. gauss-newton) also estimated parameter uncertainty and thus provided confidence interval estimates for temporal range or sampling frequency.

Recall that a sixth order autocorrelation process was used to simulate the two X- and Y-series. The estimated temporal range for the X-series was approximately six, as specified in the ACF. The estimated temporal range for the Y-series however, was approximately eight. The reason for this latter overestimate is not clear. Because of larger variation in the Y-series, the corresponding 95% confidence interval was also larger than the X-series - between 5 and 6 for the X-series and between 7 and 10 for the Y-series. Large first and second order autocorrelations in the Y-series ACF could be responsible for the larger variation and the larger than expected range parameter. Further examination of ACF patterns is warranted and may explain conditions under which the temporal range parameter will be over estimated.

The second finding of interest is that although small sample sizes dramatically affected both the nugget and sill parameters in the variogram fits, the range parameter was relatively stable. This is promising property of the temporal range in providing estimates of sampling frequency. It means that large sample sizes typically required for reliable time series analyses, may not be required if the primary intent is to estimate sampling frequency.

Groundwater samples are often collected quarterly and considered sufficient for regulatory purposes. But with only four observations per year, many years of sampling would be required to establish a sample size similar to the simulations used here. Although quarterly sampling may be adequate to assure statistical independence among sequential measurements, more frequent sampling may also be adequate. This fact, however, can never be determined from the quarterly data.

Patterns detected in the sampling history may also influence sampling frequency at a given well. Recent concentration trends that show relatively rapid increases or decreases will probably necessitate an increase in the sampling frequency. "Noisy" (i.e. highly variable) trends may suggest a higher sampling frequency than "quiet" ones.

Barcelona, et. al. (1989) presents a method for estimating a sampling interval (in weeks) based on an AR(1) model of groundwater time series data and the estimated standard error $Var(\overline{X}) = \hat{\sigma}^2 / n$ of that series. A ratio of the effective sample size (n_{ef}) to the actual sample size (n) is defined and is viewed as the relative loss in information due to autocorrelation. This ratio will be increasingly smaller with increasingly larger 1st order autocorrelations. Barcelona, et. al. (1989) restrict their method to AR(1) processes because in their view it is "...difficult to extend the analysis of water quality data beyond lag-one because the autocorrelation function becomes excessively noisy". The temporal variogram method presented here does not impose any restriction on the order of the ACF.

A complete method for estimating a sampling frequency should require the establishment of a "frequency estimation well" at every waste unit after its construction. Samples should be collected on a weekly or bi-weekly basis for either water quality metrics such as total organic carbon or specific conductance, or for some analyte known to be a constituent of concern. Within a year a sample size of at least 25 measurements will be available for analysis by the method of the temporal variogram.

Statistical models such as the gaussian model used to fit the hypothetical data used in this paper will provide technically defensible estimates for a groundwater sampling frequency. These frequency estimates can be proposed to environmental regulators as a part of the waste unit operating permit.

It is not intended that these statistical methods be the sole basis for a proposed sampling frequency. In fact, professional judgment from geologists and hydrologists in connection with the use of groundwater flow models is equally essential to a well supported proposal to regulators.

In order for the temporal variogram to become a mature technology, further examination of the initial assumptions is required, assumptions such as process stationarity. If there are trends in the time series data, parametric or nonparametric regression models could be used to remove the systematic variation due to the trend. A temporal range estimate could then be obtained from variogram analysis of the time-ordered residuals.

Other properties of real data include missing values and/or unequally spaced sampling intervals. The effects of data estimation (i.e.. replacement) or truncation methods on sampling frequency estimates will also require examination to provide greater method utility.

When the temporal variogram technology is mature and reliable, its application will promote more cost effective groundwater monitoring programs throughout industry and the government complex.

REFERENCES

Barcelona, M. J., Wehrman, H. A., Schock, M. R., Sievers, M. E. and Karny, J. R. 1989. Sampling Frequency for Ground-Water Quality Monitoring. EPA/600/4-89/032.

Box, G. E. P. and Jenkins, G. M. 1976. Time Series Analysis forecasting and control. Holden-Day, San Francisco.

Cressie, N. A. C. 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., New York.

EPA 1989. Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities. Interim Final Guidance. EPA/530 - SW-89-026.

EPA 1992. Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities. Addendum to Interim Final Guidance. EPA/530-R-93-003.

Issaks, E. H. and Srivastava, R. M. 1989. Applied Geostatistics. Oxford University Press, New York.