A JACKKNIFE APPROACH TO EXAMINE UNCERTAINTY AND TEMPORAL CHANGE IN THE SPATIAL CORRELATION OF A VOC PLUME

MARK D. VARLJEN[1], MICHAEL J. BARCELONA[2] and H. ALLEN WEHRMANN[3]
[1] *Applied Hydrogeologic Research, Inc., 11824 North Creek Parkway N, Suite 105, Bothell, WA 98011;* [2] *Dept. of Civil & Environmental Engineering, University of Michigan, 1221 IST Bldg., Ann Arbor, MI 48109–2099;* [3] *Office of Ground-Water Quality, Illinois State Water Survey, 2204 Griffith Drive, Champaign, IL 61820*
(author for correspondence, fax: +1 734 7636513)*

(Received 16 July 1997; accepted 2 June 1998)

Abstract. The application of geostatistics to spatial interpolation of time-invariant properties in ground-water studies (such as transmissivity or aquifer thickness) is well documented. The use of geostatistics on time-variant conditions such as ground-water quality is also becoming more commonplace.

Unfortunately, the detection of temporal changes in spatial correlation through direct comparison of experimental semivariograms is difficult due to the uncertainty in sample semivariograms constructed from field data. This paper discusses the use of the jackknife approach to estimate confidence limits of semivariograms of trichloroethane (TC) and other volatile organic compounds (VOC) in contaminated ground-water in northern Illinois. Examination of the 'spread' of the confidence limits about the semivariograms created from two types of sampling networks are discussed.

Keywords: confidence intervals, geostatics, ground-water, jackknife, spatial correlation, temporal variability, volatile organic compounds

1. Background

Practitioners have particularly appreciated the use of kriging techniques to interpolate contaminant concentrations where the number of sampling points is limited as well to provide information on the spatial correlation of those points. However, few practitioners are aware of the importance of temporal changes in the spatial correlation of ground-water quality when they apply these estimation techniques.

1.1. The jackknife method of semivariogram estimation

An important consideration regarding the design of geostatistically based sampling networks is that hydrogeochemical fields often vary in time as well as space (Fedorov, 1987). Because of this variation, a set of subregional or regional ground-water sampling locations that is considered optimum should not be assumed to be time-invariant without justification. The estimate of the semivariogram, $\hat{\gamma}$(h), is one means by which the spatial correlation of ground-water contamination can

Environmental Monitoring and Assessment **59:** 31–46, 1999.
© 1999 *Kluwer Academic Publishers. Printed in the Netherlands.*

be investigated (Cooper and Istok, 1988). Temporal change in spatial correlation can be identified by comparing semivariograms of geochemical concentrations in ground-water samples collected during different time periods. However, to do this effectively, the confidence limits of each estimate should be considered. Unfortunately, because of the presence of correlation in a single realization of a regionalized variable, classical statistical theory cannot be applied to make inferences regarding the confidence limits of the semivariogram (Shafer and Varljen, 1990). The jackknife method of parameters estimation offers a solution to this dilemma.

> "The jackknife estimator of the semivariogram for a particular time period gets its variability (the basis of the confidence limits) from the weighted differences between the semivariogram estimate computer from the entire sample data set for the time period of interest and the parameter estimate computed from all the data for that time period minus a partition of the sample data set for the time period. Computing this difference by repeatedly removing a partition results in a set of generated parameter estimates that are independent and nearly normally distributed, and therefore a variance can be readily calculated."

Let us suppose the entire sample data set of size *n* elements is partitioned into *g* subgroups of size *m* such that n = gm. Let $\hat{y}_{all}$ be the parameter estimated by using the entire data set n. Let $\hat{y}$ be the parameter estimated by using all the data remaining after removing the j^{th} partition, $n_j = (g - 1)m$. Partition-dependent estimates of the parameter, $J_j[\hat{y}]$, are then calculated according to:

$$J_j\ [\hat{y}] = g\,\hat{y}_{all} - (g - 1)\,\hat{y}_j \qquad j = 1,2, ..., g \tag{1}$$

The final jackknife estimate, $J\,[\hat{y}]$, of the parameter (Quenoille, 1956; Efron, 1982; Chung, 1984) is:

$$J[\hat{y}], = \frac{1}{g}\sum_{j=1}^{g} J - j[\hat{y}] \tag{2}$$

Substituting $\hat{\gamma}\ (h)$ for $\hat{y}$ results in the jackknife estimator of the semivariogram:

$$J_j[\hat{\gamma}(h)] = g[\hat{\gamma}(h)] - (g-1)[\hat{\gamma}_j(h)] \quad j = 1, 2, ..., g$$

$$J[\hat{\gamma}(h)] = \frac{1}{8}\sum_{j=1}^{g} J_j[\hat{\gamma}(h)] \tag{3}$$

Tukey (1958) extended Quenouille's work on reducing bias in estimation to the approximation of confidence limits. The partition jackknife estimates, $J_j[\hat{y}]$, may be used to construct confidence bands about the jackknife estimate of the parameter (Turkey, 1958; Efron and Gong, 1983). The variance, $\sigma_J^2(h)$, about the jackknife estimate $J[\hat{\gamma}(h)]$, or simply the jackknife variance, is estimated by (Chung, 1984):

$$\sigma_J^2(h) = \frac{1}{g(g-1)}\sum_{j=1}^{g}(J_j[\hat{\gamma}(h)] - J[\hat{\gamma}(h)])^2 \tag{4}$$

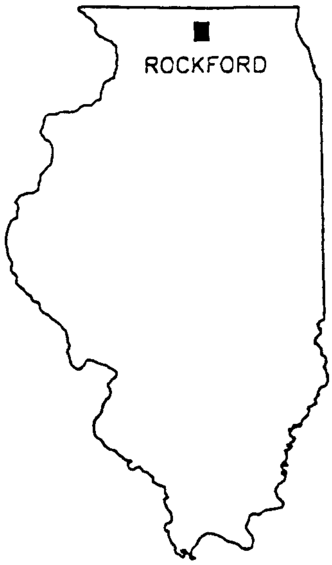

Figure 1. Location of study area, Rockford, Illinois.

where $\sigma_J\ (h)$, the root mean square error of estimation, can be used to approximate the confidence limits on $J[\hat{\gamma}(h)]$.

Shafer *et al.* (1989) discussed the jackknife technique for developing confidence limits on a sample semivariogram of shallow ground-water nitrate data. The computational aspects of the technique were described in more detail by Shafer and Varljen (1990). In the following case study, we applied the jackknife technique using the same number of pairs in each class to ground-water monitoring data from a large VOC plume to determine if the spatial correlation of VOC concentrations varied over time. The most significant aspect of this study is the illustration of the high level of uncertainty in sample experimental semivariograms. Also, the study sought to evaluate the sensitivity of that uncertainty to sample size and artificially-introduced variability, such as the variability from sampling and laboratory protocols.

2. Project Description

The data which form the basis of this study were collected after the discovery of ground-water contamination in several hundred domestic wells in an unincorporated area of southeast Rockford, Illinois (Figure 1). The field site covers an approximately 2 square mile area (5.2×10^6 m^2) within what is now known as the Southeast Rockford Superfund Site, an area of $\sim$4 square miles (10.4×10^6 m^2) of mixed residential, commercial, and industrial properties. The site is underlain by an

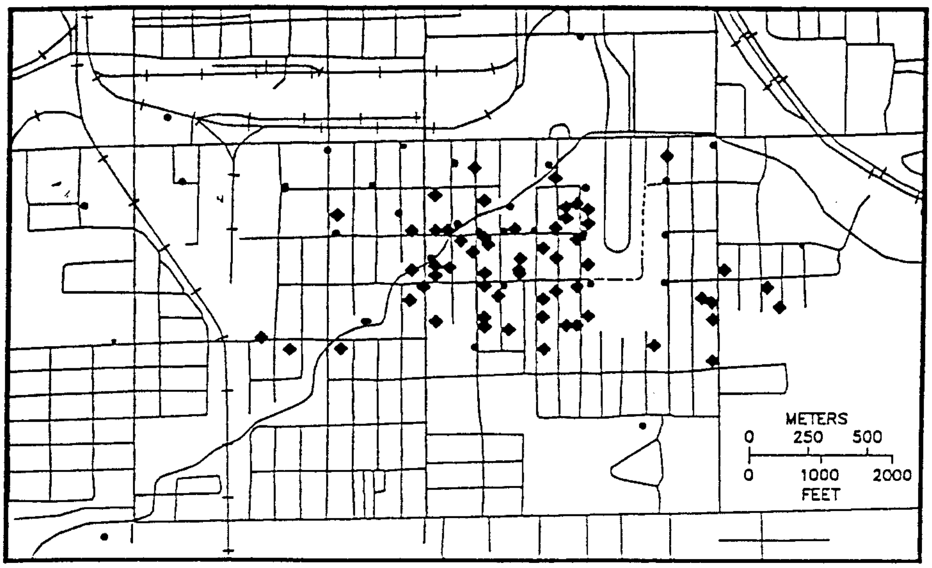

Figure 2. Locations of domestic wells (diamond) and monitoring wells (circles) sampled in this study.

extensive glacial outwash sand gravel aquifer and the contaminated domestic wells were generally less than 75 feet (23 m) deep. Among the principal contaminants is 1,1,1 trichloroethane (TCA).

Ground-water samples were initially collected from over 200 domestic wells in December 1989 by the Illinois Department of Public Health (IDPH). A follow-up sampling of these wells and others by the USEPA was conducted in July 1990. The study employed two networks of wells. The first network consisted of 59 contaminated domestic wells concentrated within an area of approximately one square mile. The second network was a dense but smaller network of 30 monitoring wells, constructed in locations based on the spatial structure derived from the domestic well sampling. We used results from the 59 wells which were common to both mass sampling periods. Shortly after this sampling was completed, the domestic wells were abandoned as municipal water was extended into the area. A network of over 40 monitoring wells was then constructed based on the geometry and spatial correlation of the volatile organic compound (VOC) plume as determined from the domestic well samplings. Thirty wells within the monitoring well network were sampled on nearly a quarterly basis for about 1.5 years (in May, August, December, 1991 and March and september 1992). The locations of the domestic and monitoring wells are shown on the map in Figure 2. A contour map of TCA concentrations in the monitoring wells in september 1992 is shown in Figure 3. TCA concentrations exceeded 250 μg/L along the central axis of the plume. The plume can be seen to extend from east to west which closely follows the direction of ground-water flow toward to Rock Rivers at the western edge of the site.

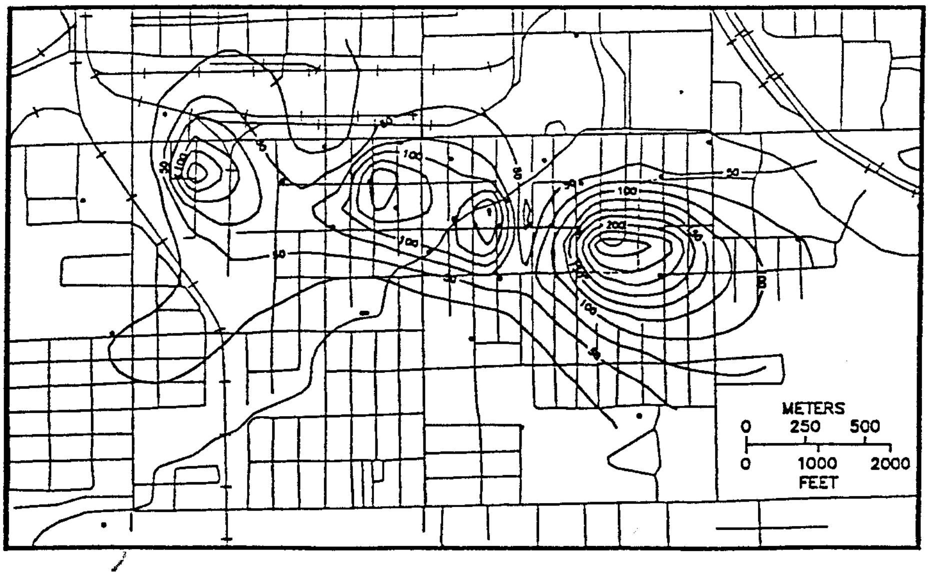

Figure 3. Concentrations of TCA within southeast Rockford study area, contour interval, 25 μg/L.

TABLE I

Precision and accuracy of July 1990 private well VOC sampling and analysis, and inter-laboratory comparison of analytical results for the major contaminant compounds

	Contaminant compound					Surrogate compound		
Statistics	DCE	DCA	C12DCE	TCA	TCE	BCM	B2C1PA	DCB14
ISWS Analysis								
Relative standard deviation[a]	20.0%	6.7%	6.7%	6.4%	7.2%	–	–	–
Relative bias[b]	–23.0%	+2.0%	–5.0%	–21.1%	–13.0	63.0%	6.3%	67.0%
IDPH Analysis								
Relative standard deviation[a]	154%	102%	76%	96%	88%	–	–	–
Relative bias[c]	–36.2%	–56.7%	–60.9%	–38.9%	–17.5%			

[a] Relative Standard Deviation (RSD) was estimated from the Range (R) and the Mean (M) of the reported concentration for field duplicate samples, assuming a formula of RSD=0.886*R/M. Average results on 6 pairs of samples;

[b] Average results on 7 field surrogate samples;

[c] Averaged results on 6 sets of samples split between IDPH and ISWS. (DCE-Dichloroethylene, DCA-Dichloroethane, c12DCE-cis 1,2 Dichloroethylene, TCA-Trichloroethane, TCE-Trchloroethylene, BCM-bromochloromethane, B2C1PA-Bromodichloropropane, DCB14-1,4 Dichlorobeneze). Both laboratories used purge and trap gas chromatography with Hall electrolytic conductivity detection.

TABLE II

Comparison of statistical parameters for 59 common sampling points sampled in Fall 1989 (IDPH) and Summer 1990 (ISWS) (concentrations in μg/L)

Compound	% Freq. Of detection[a]	Maximum[a]	Minimum[a]	ppb Mean[a]	Standard Deviation[a]
1,1-dichloroethene	80/98	63.4/87.9	0/0	10.4/23.6	16.1/21.8
1,1-dichloroethane	93/100	80.9/556	0/0.44	19.4/74.1	19.8/91.3
Cis-1,2-dichloroethene	1.7/100	108/233	0/056	1.84/57.7	14.1/52.6
1,1,1-trichlorethane	98/100	436/803	0/3.97	120/180	116/173
Trichloroethene	100/100	113/174	0.55/0.72	30.7/35.2	27.0/32.6

[a] IDPH Fall 1989 data statistics expressed first/ISWS Summer 1990 data statistics expressed second.

TABLE III

Overall mean concentration, relative standard deviation and percentage of total variance attributable to lab or field (sampling) error, and natural variability (November 1990–September 1992)

			Percent of total variability		
COMPOUND	Overall mean μg/L	Relative (std. dev.) %	Lab %	Field %	Natural %
TCA	119.5	(36%)	1.29	3.26	95.45
TCE	29.8	(43%)	1.95	12.75	85.30
C12DCE	45.2	(32%)	1.69	4.72	93.59
DCA	44.3	(28%)	1.02	5.22	93.76
DCE	16.3	(31%)	3.61	4.15	92.24

During our investigation we became interested in variations in VOC concentrations both spatially and temporally. Traditional structural analysis (semivariogram estimation) and kriging are tools to examine spatial variability. We implemented the jackknife technique to allow us to compare the structural analyses from different time periods.

> "Part of our research objectives were to understand contributions of different types of variability (i.e. sampling, laboratory, and natural) to the overall variability of VOC results and to assess the implications of this variability to various data analyses. After conducting these analyses, we found that idenetified sampling and/or laboratory variability contributed significantly to the uncertainty in estimating the semivariogram."

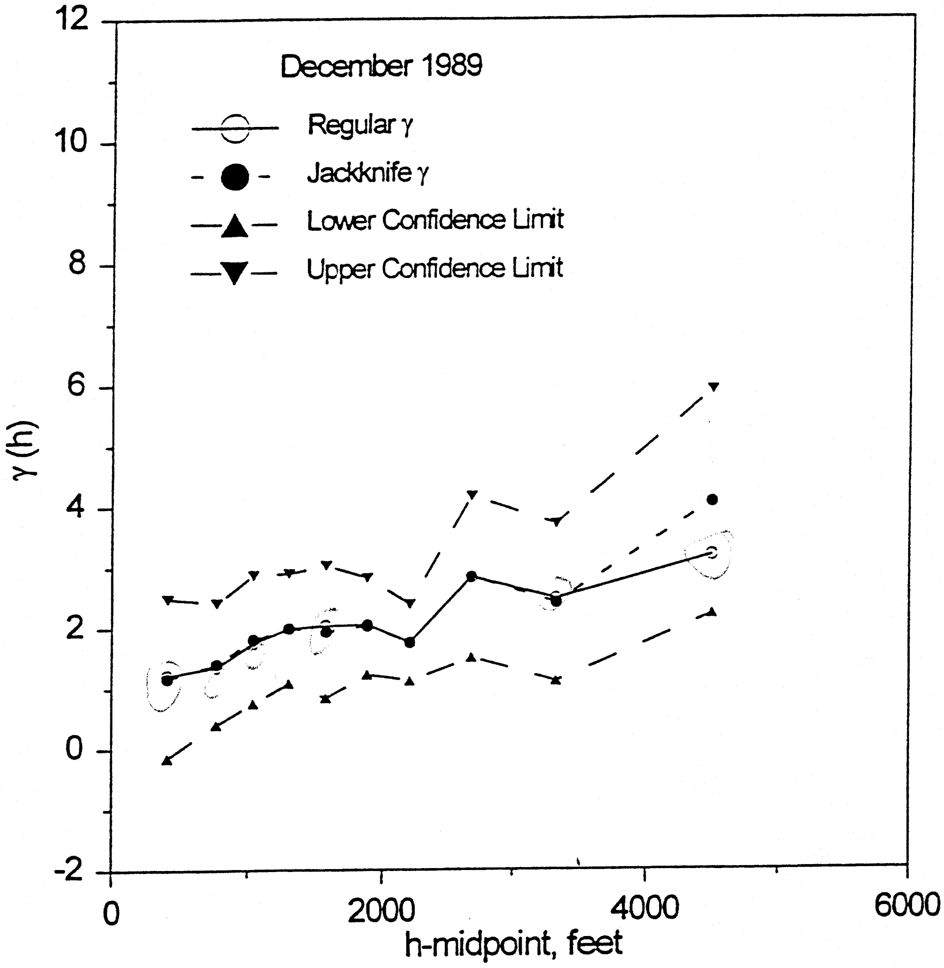

Figure 4. Jackknife confidence intervals on semivariogram of December 1989 kn[TCA] domestic well data.

3. Results and Discussion

In order to minimize controllable sources of error, the available data were used to analyze sampling and analytical variability. The methods for evaluating sampling and analytical contributions to overall variability are described in Barcelona *et al.* (1989).

The effect of less consistent sampling and analytical methods is shown in Tables I and II comparing the December, 1989 private well data collected by the IDPH to those collected in July, 1990. The precision (relative standard deviations) of the ISWS analysis for organic contaminants in samples split between the two labs demonstrate that the analytical method was in control (i.e. generally within $\sim \pm$

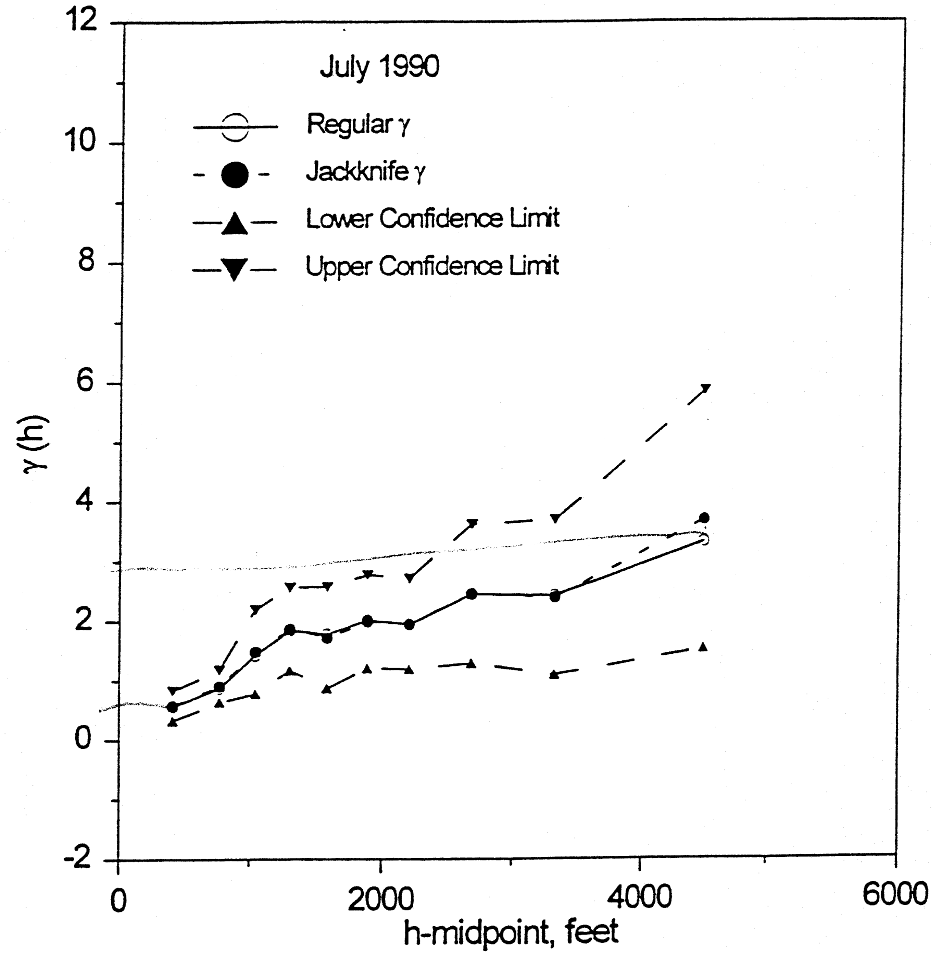

Figure 5. Jackknifed confidence intervals on semivariogram of July 1990 ln[TCA] domestic well data.

20% of the mean). The excessive negative bias for these compounds along with the generally lower detection frequency, mean and maximum concentrations (Table II) reported by the IDPH laboratory did not support confirmed analytical and sampling error control.

Given the sensitivity of geostatistical measures of trace organic contaminant variability to uncontrolled sources of error, a strict program of quality assurance/quality control was observed in the quarterly monitoring effort from November, 1990 through September, 1992 (see Table III). The monitoring wells were equipped with dedicated bladder pumps for controlled purging and sampling operations. The details of the sampling and analytical protocols are provided in Barcelona *et al.* (1994).

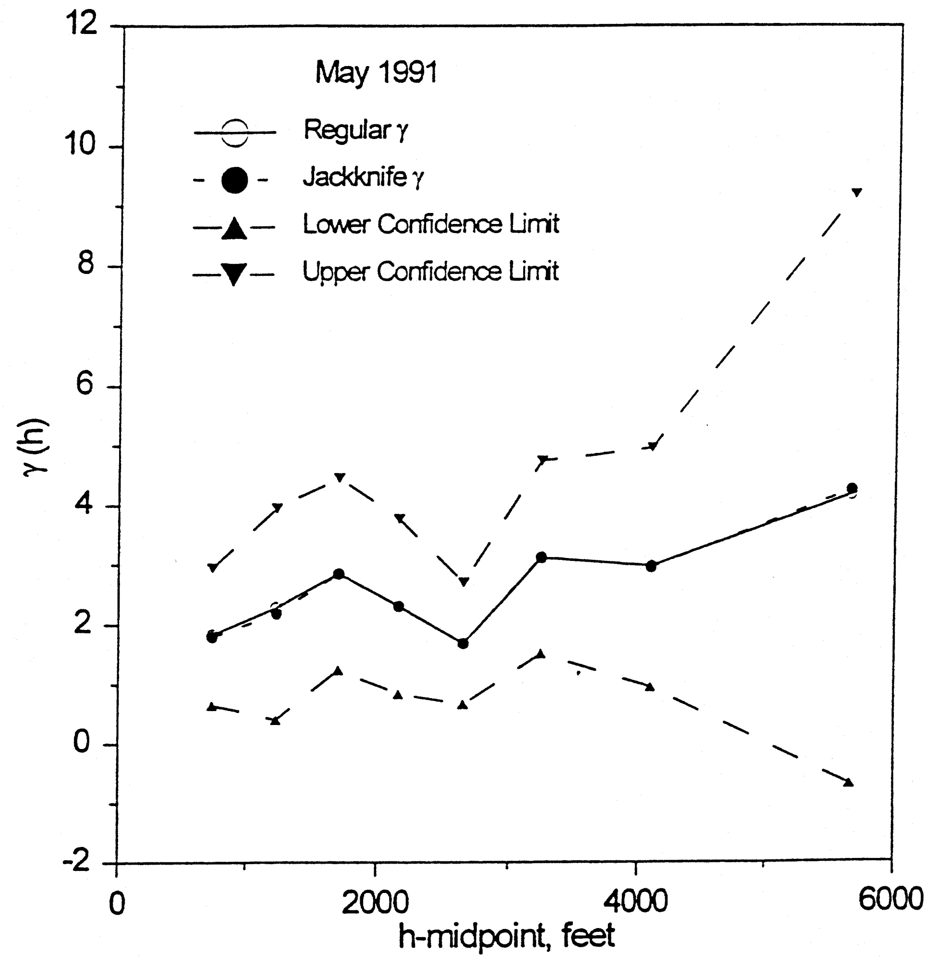

Figure 6. Jackknifed confidence intervals on semivariogram of May 1991 ln[TCA] monitoring well data.

The data in the table show that overall concentration variability across the intensive study areas for two years was $\leq$ 43% over the quarterly datasets. Of this total variability, lab and field error constituted less than ~15% of the error in the final concentration result. These results provide strong evidence that our QA/AC peocedures were in control over the study period. Also, they demonstrate that natural variability in VOC contaminant concentrations can be reproducibly observed with the simple sampling protocol used in this work.

Results of structural analyses for TCA concentrations for several sampling periods, with accomppanying jackknife confidence bands, are shown in Figures 4–10. The estimated semivariograms for each sampling interval are somewhat different (Figure 11), leading one to believe that temporal change in spatial correlation

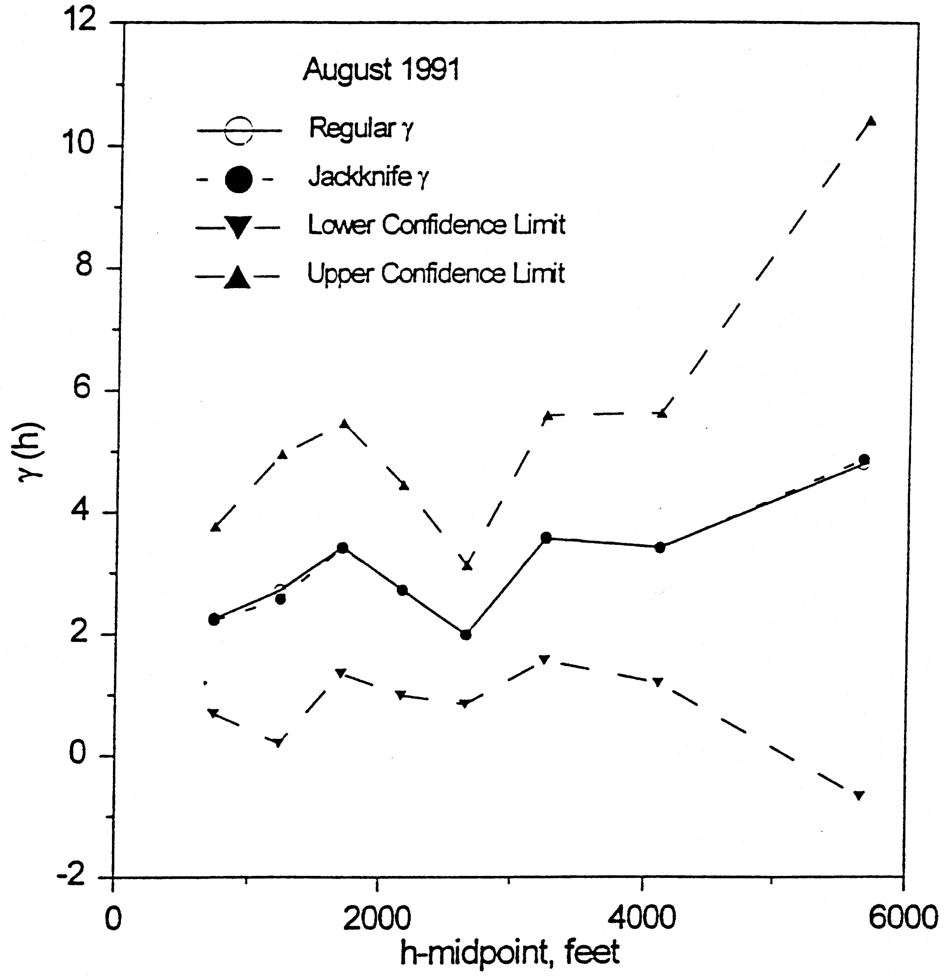

Figure 7. Jackknifed confidence intervals on semivariogram of August 1991 ln[TCA] monitoring well data.

structure had occurred. The confidence bands for each estimated semivariogram are large enough however to indicate that any differences are not significant. For example, Figure 12 shows the results for the December 1989 and July 1990 sampling events overlaid. Note that the overlap of confidence limits is complete.

The notion that there have been no temporal changes in spatial correlation is consistent with the hydrogeochemical processes that control the distribution of contaminants. The contaminant source is thought to be at least 1 mile (1,609 m) upgradient from the most upgradient monitoring well used for this study, and the plume is thought to be old (decades) and was apparently quite stable within the area we were monitoring.

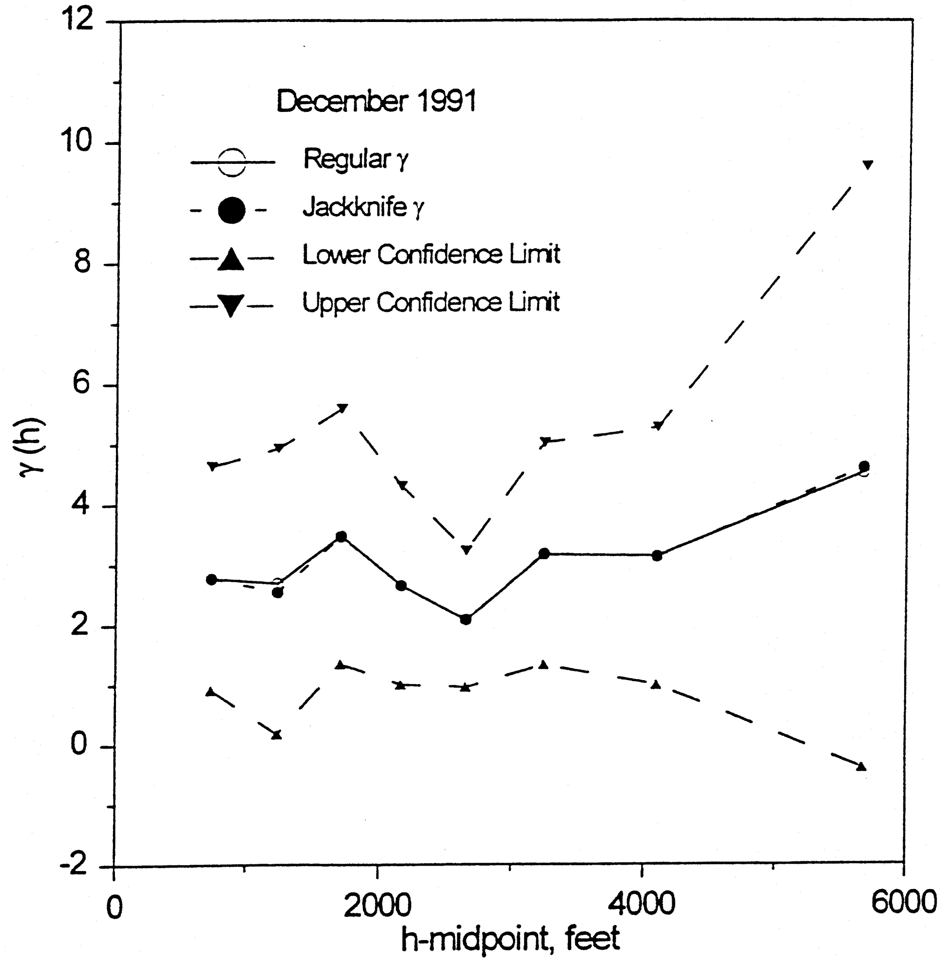

Figure 8. Jackknifed confidence intervals on semivrariogram of December 1991 ln[TCA] monitoring well data.

Perhaps the most interesting feature of the sample semivariograms and the associated jackknife confidence bands is the relative differences in the width of the confidence bands. Clearly, confidence in semivariogram estimates is extremely sensitive to sample size. Note the differences in the widths of the confidence bands for the sampling periods where 59 (December, 1989; July, 1990) wells were used vs. 30 wells for the remaining sample periods in 1991 and 1992. Doubling the number of sample points results in approximately a two-fold improvement in the 95% confidence interval.

The effect of sample size is somewhat intuitive and was expected. What was unexpected, however, was the difference in the confidence bands between the December 1989 and July 1990 sampling periods (Figures 4 and 5). We did not expect

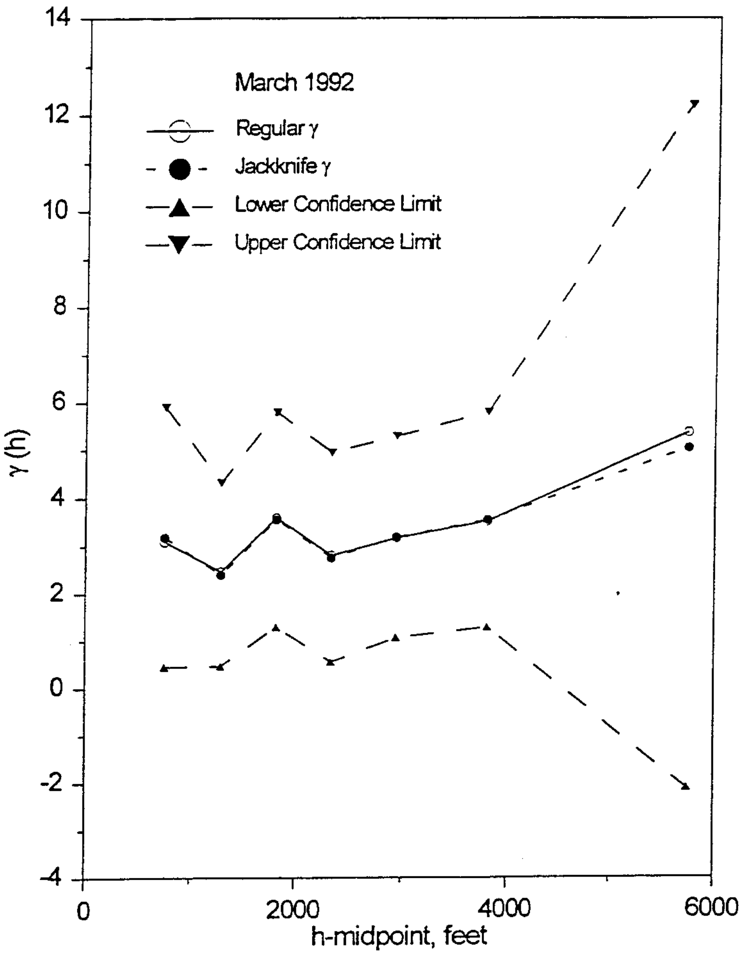

Figure 9. Jackknifed confidence intervals on semivrariogram on March 1992 ln[TCA] monitoring well data.

to see such a difference because the sampling points were used for each period. We interpret the difference in the confidence band to be due to the effect of sampling and laboratory variability. Different field personnel, sampling protocol, and laboratories were used for the two periods. We note that stricter protocol implemented by the July 1990 sampling team resulted in less noise in the data. The data set with less added variability allowed for more confidence in the estimation of the semivariogram.

It should be noted that virtually all of the semivariograms showed increasing uncertainty at longer lag distances, reflecting higher overall variance. Given the level of noise in the experimental spatial correlations across the entire range, the use of directional variograms (i.e. censoring the data) with far lower class sizes

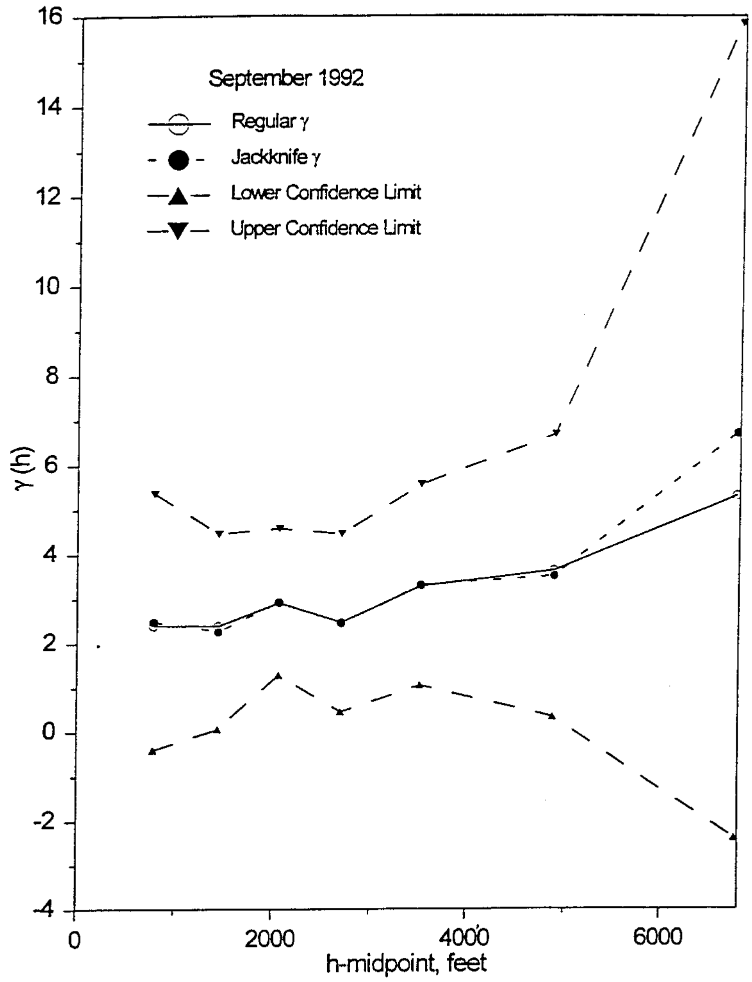

Figure 10. Jackknifed confidence intervals on semivrariogram of September 1992 ln[TCA] monitoring well data.

would be expected to increase the uncertainty associated with their use (e.g. constructing contour maps, estimating mass/concentrations distributions, etc.). Rules of thumb for minimum sample sizes or range selection should be evaluated very carefully in any specific application.

It may be argued that with known levels of sampling and analytical variabiity, one may simply remove (i.e. subtract) these errors and minimize the uncertainty in the spatial analyses. Sampling and analytical errors are essentially random and often concentration dependent. Therefore, it is not realistic to apply a simple correction without, in fact, introducing additional artificial variability into the spatial analyses.

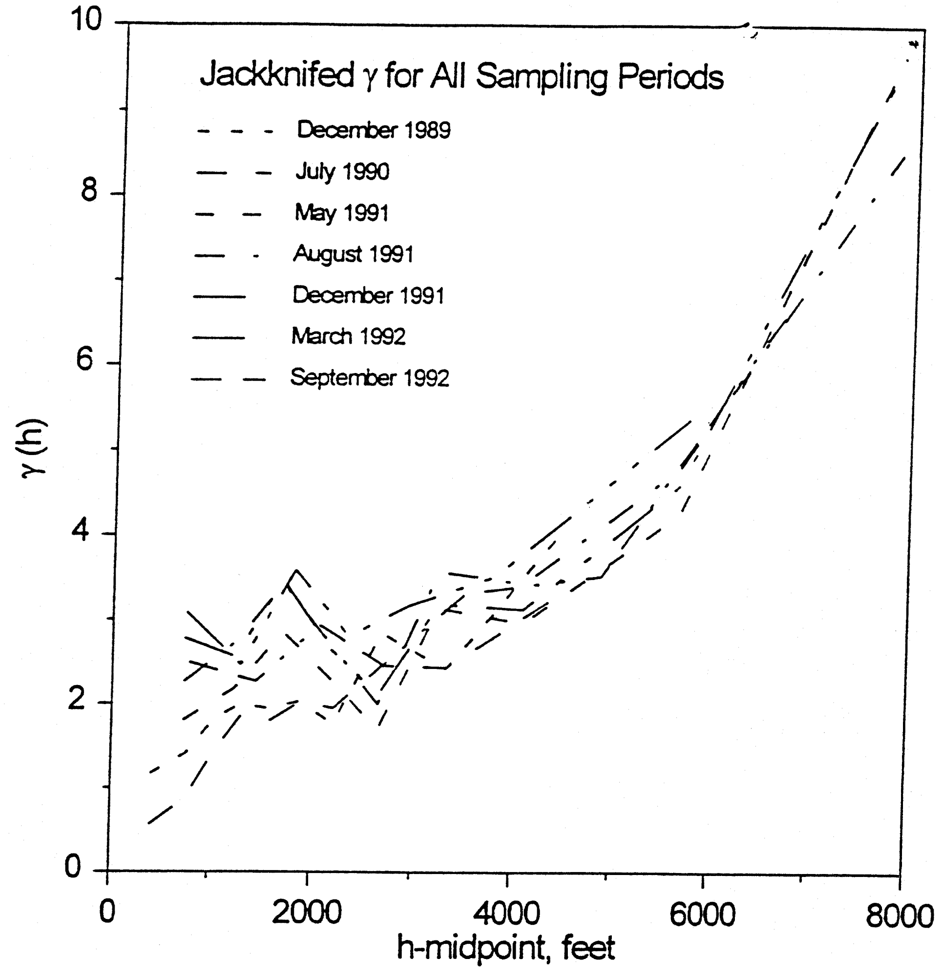

Figure 11. Jackknifed γ for all sampling periods.

4. Conclusions

It may be concluded from the analysis of jackknife stimates of sample semivariograms (and the associated confidence limits) that the data set could no identify any significant temporal changes in the spatial correlation of TCA over the three years that the study spanned. More importantly, however, the demonstrated high level of uncertainty that is associated with the estimation of semivariograms from field data sets was evident. Careful control of artificial variability can improve confidence in estimates of spatial correlation. The fact that sample size had such a profound effect on uncertainty in estimates of the semivariogram, coupled with the fact that we did not observe much significant temporal change in the plume, indicates that

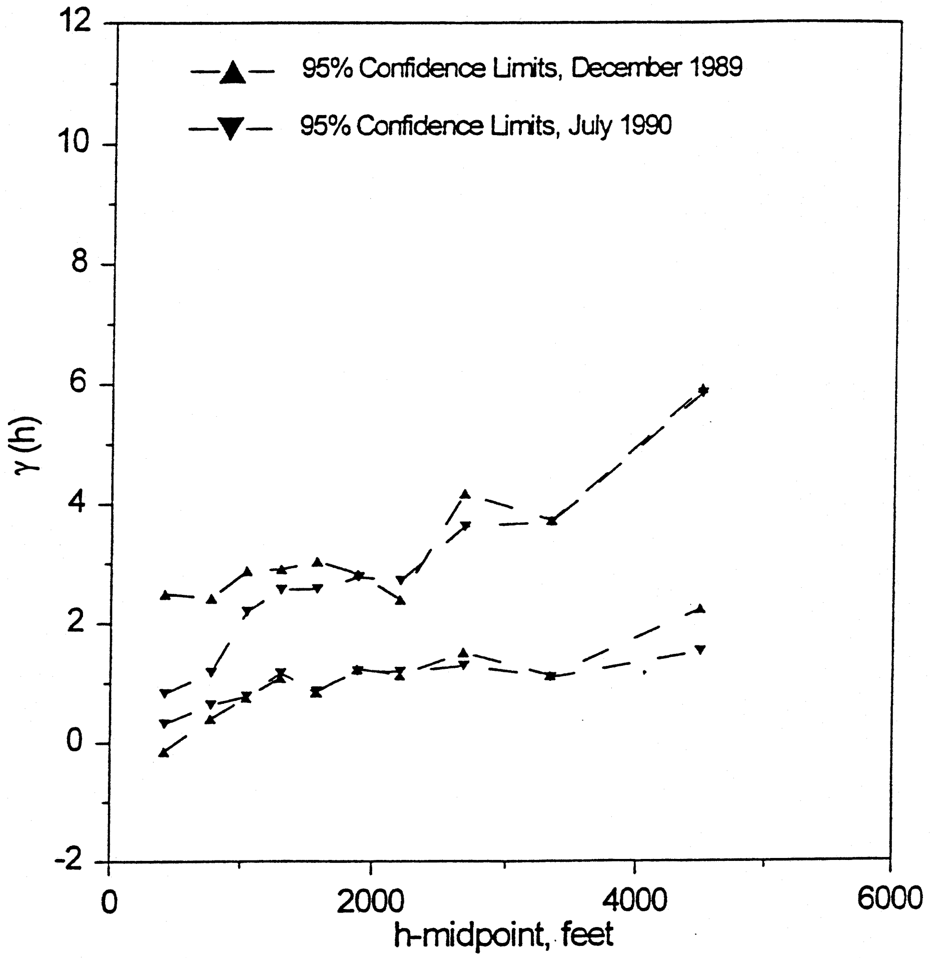

Figure 12. Jackknifed 95% confidence intervals on semivariograms of December 1989 and July 1990 ln[TCA] domestic well data,

the information return for our characterization would have been greater if we had used more sampling points and fewer sampling events.

It should be recognized that some of these conclusions are specific to this investigation, we conclude generally that variability and uncertainty in estimated spatial correlation is quite significant and should be considered in the application of geostatistics to ground-water contamination problems. Applications such as probability kriging to define clean-up boundaries or sampling network optimization based on kriging standard deviations are particularly sensitive to semivariogram uncertainty. The jackknife technique is an appropriate approach to quantifying this uncertainty.

Acknowledgments

The information for this study was collected through funding by the USEPA (EMSL) Las Vegas, NV under Cooperative Agreement CR-815681. This report has not been subjected to Agency review and, therefore, does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

Barcelona, M. J., Lettenmaier, D. P. and Schock, M. R.: 1989, 'Network Design Factors for Assessing Temporal Variability in Ground Water Quality', *Environ. Monit. And Assessment* **12**, 149.

Barcelona, M. J., Wehrmann, H. A. and Varljen, M. D.: 1994, 'Reproducible Well Purging and Stabilization Criteria', *Ground-Water* **32**(1), 12.

Chung, C. F.: 1984, 'Use of the jackknife method to estimate autocorrelation functions (or variograms)', in G. Verly, M. David, A. G. Journel and A. Marechal (eds.), *NATO ASI Series, Geostatistics for Natural Resources Characterization* **1**, D. Reidel, Hingham, MA, pp. 55–69.

Cooper, R. M. and Istok, J. D.: 1988, 'Geostatistics applied to ground-water contamination I: Methodology', *ASCE Journal of Environmental Engineering* **114**(2), 270.

Davis, B. M.: 1987, 'Uses and abuses of cross-validation in geostatistics', *Mathematical Geology* **19**(3), 214.

Efron, B.: 1982, 'The jackknife, the bootstrap and other plans', SIAM, Philadelphia.

Efron, B. and Gong, G.: 1983, 'A leisurely look at the bootstrap, the jackknife, and cross-validation', *American Statistician* **37**(1), 36.

Fedorov, V.: 1987, 'Kriging and other estimators of spatial field characteristics with special reference to environmental studies', WP–87–99, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Quenouille, M. H.: 1956, 'Notes on bias in estimation', *Biometrika* **43**, 353.

Shafer, J. M., Varljen, M. D. and Wehrmann, H. A.: 1989, 'Identifying temporal change in the spatial correlation of regional ground-water contamination', *Hydraulic Engineering '89 Proceedings*, National Conf. On Hydraulic Engineering/HY Div/ASCE, New Orleans, LA August 14–18, 1989, pp. 398–403.

Shafer, J. M. and Varljen, M. D.: 1990, 'Approximation of confidence limits in sample semivariograms from single realizations of spatially correlated random fields', *Water Resources Research* **26**(8), 1787.

Tukey, J. W.: 1958, 'Bias and confidence in not quite large samples', *Ann. Math. Stat.* **29**, 614.