ELSEVIER

Journal of Hazardous Materials 83 (2001) 1–10

Journal of Hazardous Materials

www.elsevier.nl/locate/jhazmat

Comparison of field and laboratory methods for monitoring metallic mercury vapor in indoor air

Raj Singhvi [a,*], Rod Turpin [a], Dennis J. Kalnicky [b], Jay Patel [b]

[a] *Environmental Response Team Center, U.S. Environmental Protection Agency, 2890 Woodbridge Avenue, Edison, NJ 08837, USA*

[b] *Lockheed Martin Technology Services Group, Environmental Services/REAC, 2890 Woodbridge Avenue, Edison, NJ 08837, USA*

Abstract

Real-time metallic mercury vapor levels of the indoor air were monitored at several mercury spill sites around the US in order to evaluate the effectiveness of site cleanup operations. Mercury readings taken in the field with a Jerome 431^{TM} Mercury Vapor Analyzer were compared with laboratory analysis using a modified National Institute for Occupational Safety and Health (NIOSH) 6009 method. Statistical evaluation showed that the data were highly comparable except at low concentrations, due to the large degree of uncertainty associated with measuring low levels of mercury with the Jerome analyzer. Regression analysis indicated that Jerome measurements of 10 $\mu g/m^3$ or greater are comparable for field analysis of mercury vapor in air. Published by Elsevier Science B.V.

Keywords: Mercury; Indoor air; Field monitoring; Laboratory analysis; Statistical evaluation

1. Introduction

The quality of indoor air and the resultant risk associated with accidental exposure to volatilized metallic mercury (Hg) are major concerns for building occupants. Indoor air monitoring programs that can produce high quality data with rapid turnaround of results are needed to effectively address these concerns. The field and laboratory analytical methods developed by the US Environmental Protection Agency's Environmental Response Team (U.S. EPA/ERT), through its Response Engineering Analytical Contract (REAC), provide timely, cost-effective elemental Hg analysis while maintaining rigorous quality assurance/quality control (QA/QC) procedures to ensure reliability of the analytical data. The

* Corresponding author. Tel.: +1-732-321-6761; fax: +1-732-321-6724.

0304-3894/01/$ – see front matter. Published by Elsevier Science B.V.
PII: S0304-3894(00)00334-4

Jerome 431™ Mercury Vapor Analyzer provides real-time screening of suspected contaminated areas to identify "hot spots" and to monitor progress of decontamination procedures. Once the Hg concentration falls below the Jerome detection level, clearance sampling is performed using a modified NIOSH 6009 method to ensure that the long-term exposure level is not exceeded [1].

2. Analysis methodology

2.1. Field monitoring (Jerome)

The Jerome 431™ Gold Film Mercury Vapor Analyzer (Arizona Instrument Company (AZI), Tempe, AZ) is designed for analysis of indoor air mercury vapor levels in the workplace environment and for the location of mercury spills [2]. Two real-time modes of operation are available: the SAMPLE mode for direct reading of Hg vapor concentration in milligrams per cubic meter (mg/m^3), and the SURVEY mode for quick screening to locate high concentration areas. A thin gold film, in the presence of mercury vapor, undergoes an increase in electrical resistance proportional to the mass of mercury in the sample. The gold film selectively adsorbs elemental mercury, which eliminates interferences common to UV mercury analyzers such as water vapor, particulates, cigarette smoke, and organic solvents. Activating either the SAMPLE or SURVEY mode starts an internal pump, which draws a precise volume of air over the Gold Film Sensor. The sensor adsorbs and integrates the Hg vapor, and the resulting signal is displayed on the digital readout meter. The Jerome 431™ is factory calibrated. A calibration kit supplied with the analyzer is used to verify proper operation.

2.2. Laboratory analysis (NIOSH 6009)

Field sampling: Indoor air samples of volatilized elemental Hg are collected on solid sorbent material (typically Hopcalite™ or Hydrar™) contained in glass tubes. Air is pumped through the sorbent with a personal sampling pump, which can be programmed for collection time and flow rate (typically 0.5–5 l/min). Pump flow rate is initially calibrated against a rotometer reference and is checked before/after sample collection. Sampling stations are typically set up in several locations within the structure.

Modified NIOSH 6009 method: the sorbent material from the collection tube (typically 200 mg in a single section) is quantitatively transferred to a 100 ml volumetric flask. The sample is digested by first adding 2.5 ml of concentrated nitric acid followed by 2.5 ml of concentrated hydrochloric acid. After digestion, the sample is diluted to volume with deionized water and analyzed by the cold-vapor atomic absorption spectroscopy technique with no additional dilutions. Results are reported as $\mu g/m^3$ based on the total air volume collected for the sample. Matrix effects are minimized by using sorbent material for preparation of blanks and calibration standards [3]. The modified NIOSH 6009 method incorporates more concentrated sample solutions than that for the standard method. This minimizes dilution effects while providing improved Hg detectability to meet the demanding action level requirements associated with emergency response actions.

3. Statistical methods

Several statistical analysis methods may be used for evaluating and comparing field and laboratory data [4,5]. A probability-value (*P*-value) is usually included in the output. Irrespective of the analysis being performed, the *P*-value is the lowest level at which the proposed hypothesis can be rejected. If the *P*-value is less than the given significance level (usually 0.05), the hypothesis can be rejected. If the *P*-value is greater than the significance level, there is no statistical significance and the hypothesis cannot be rejected. Prior to performing any statistical evaluations, a test of distribution is performed on the dataset to determine if parametric or non-parametric statistical analysis should be utilized.

3.1. Pairwise comparisons

Pairwise comparisons are useful for initial evaluation of field versus laboratory datasets. This is a hypothesis test, run at a significance level of 0.05, which determines if there are significant differences between two sets of paired data. During the test, one dataset is subtracted from the other to get a third set of differences. A statistical analysis is performed to test the null hypothesis that the mean of these differences equals zero. If the data is not normally distributed, a test about the median as opposed to the mean is performed. In both cases, the *P*-value determines the significance of the analysis. If the *P*-value is less than the significance level, the null hypothesis is rejected and there is significant difference between the datasets. If the *P*-value is greater than the significance level, there is no significant difference between the datasets. This does not mean that the datasets are equal, but, rather, that they are not significantly different from each other. Even if pairwise comparisons analysis indicates that field and laboratory datasets are significantly different, it does not mean that a strong relationship cannot exist between them.

3.2. Correlation analysis

Correlation analysis is related to regression analysis. It determines the degree of linearity between two sets of data and may be utilized prior to linear regression analysis. A correlation coefficient (R) is generated in the analysis which ranges in value from -1.0 (a perfect negative linear correlation) to 1.0 (a perfect positive linear relationship). A zero value indicates no linear relationship exists. If a strong linear relationship exists, linear regression analysis should be used to evaluate the datasets. If a non-linear relationship exists, a non-linear regression analysis may be considered.

3.3. Linear regression analysis

Regression analysis is used to fit a model between the independent variable (field data) and the dependent variable (laboratory data) to determine if a linear relationship exists and if that relationship is significant. Regression analysis yields the coefficient of determination (r^2), which defines the proportional amount of variability explained by the regression model. The r^2 value ranges from 0.0, which means no variability to 1.0, which indicates that 100% of the

variability is explained by the model. The regression also yields the *F*-statistic, which determines if the model explains a significant amount of the variation in the datasets. A *P*-value may also be generated for the *F*-statistic. If the *P*-value for the *F*-statistic is less than the significance level (0.05), and the r^2 value is high (>0.7), the regression model is significant.

The residuals of the regression model should be examined for outliers. The residuals are the differences between the predicted dependent values and the actual dependent values. A plot of residuals versus dependent values should be a random scattering of points. Anomalies or outliers are usually apparent. If any outliers are present, the regression analysis should be performed without these values to determine their impact upon the model. If the sample size for regression is small (less than eight observations) removal of data points should be avoided, irrespective of their impact, because their removal greatly increases the error associated with the regression analysis.

4. Mercury spill site case studies

Field and laboratory analytical data from several mercury spill sites were statistically evaluated to determine comparability of Jerome and NIOSH analysis of Hg vapor levels. Mercury datasets for the following sites were evaluated:

1. LCP chemical site; April–May 1995; 16 observations;
2. Grand Street site; January 1996; 124 observations;
3. Glenside mercury spill site; February–March 1996; 32 observations;
4. Dallas, PA mercury site; November 1996; 173 observations;
5. Grand Street site; August 1997; 63 observations.

4.1. Evaluation of mercury data

All pairwise comparisons and correlation analysis evaluations were performed using the SAS™ statistical analysis software package. The SAS™ correlation analysis output includes two coefficients: the Pearson coefficient for normal (bell shaped) data distributions and the Spearman coefficient for non-normal distributions. Regression analyses were done with Lotus 123 (Release 4.01), Quattro Pro (V 6.01) and SAS™ (V 6.10 and V 6.12). The SAS™ regression output includes a Student residual and Cook's *D* value for each observation [6]. The Student residual is the residual divided by the standard error. The Cook's *D* value is a relative measure of data quality. If the Student residual is between 2.0 and 3.0 in absolute value, the observation may be an outlier. If it is 3.0 or larger in absolute value, the observation is considered a probable outlier. When the Student residual is larger than 2.0 in absolute value and Cook's *D* is outside the range of the dataset, the observation may be considered a potential outlier and a new regression analysis should be performed without that observation.

4.2. LCP chemical site

Sixteen observations for Jerome versus NIOSH analysis were obtained from the LCP chemical site. In cases where multiple Jerome measurements were made, the average was

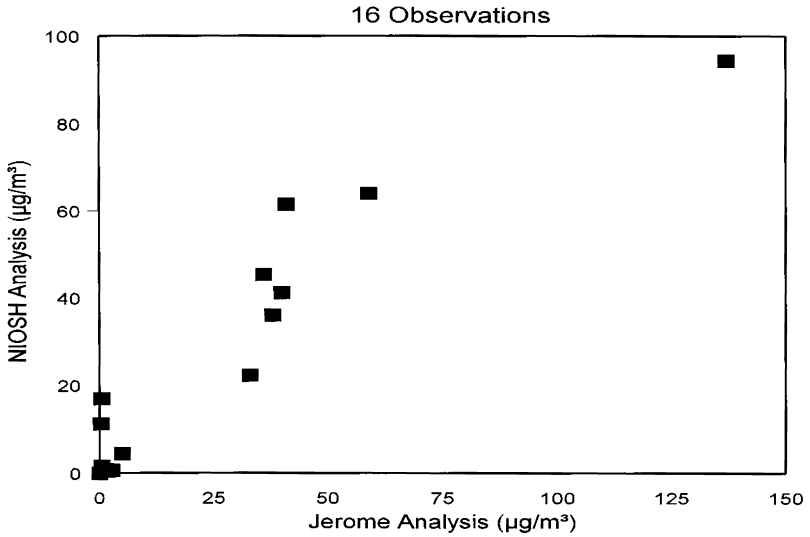

Fig. 1. Laboratory (NIOSH 6009) and Field (Jerome) Mercury results — LCP chemical site.

used for evaluation purposes. NIOSH and Jerome data are shown in Fig. 1. Pairwise comparisons analysis indicated that Jerome and NIOSH datasets were non-normal and they were not significantly different ($P = 0.910$). The Spearman coefficient ($R = 0.809$) indicated that the data were highly correlated and regression analysis was justified. Regression analysis results are in Table 1. One potential outlier was identified (Jerome $= 137$, NIOSH $= 94.4$) and the regression was repeated without the outlier. Fig. 2 shows residuals (predicted — NIOSH) versus the NIOSH (dependent) values for the regression without the outlier. Results without the outlier indicated that the Jerome and NIOSH data were highly comparable: $r^2 = 0.874$, slope $= 1.04$, $F = 98$. The RMS error (8.2) indicated that, in general, Jerome results of 9 $\mu g/m^3$ or greater were meaningful compared to laboratory analysis.

4.3. Grand Street site, January 1996

Fig. 3 shows Jerome versus NIOSH data for 124 observations at the Grand Street mercury site (January 1996). Pairwise comparisons analysis indicated that Jerome and NIOSH datasets were non-normal and they were significantly different ($P = 0.0001$). The Spearman coefficient ($R = 0.901$) indicated that the data were highly correlated and regression analysis was justified. Regression analysis (Table 1) identified four potential outliers and the regression was repeated without the outliers. Results without the outliers indicated that the Jerome and NIOSH datasets were highly comparable: $r^2 = 0.844$, slope $= 1.09$, $F = 645$. The RMS error (6.6) indicated that Jerome results of 7 $\mu g/m^3$ or greater were meaningful compared to laboratory analysis.

Table 1
Regression analysis results, NIOSH (dependent) vs. Jerome (independent)

Parameter	All data	Without outliers
LCP chemical site		
n[a]	16	15
(r^2)[b]	0.862	0.874
Slope	0.750	1.04
Intercept	6.60	2.59
RMS error[c]	10.8	8.24
F-value	94.3 ($P = 0.0001$)	98.0 ($P = 0.0001$)
Dallas, PA mercury site		
n	81	79
r^2	0.953	0.802
Slope	0.966	0.879
Intercept	−2.56	−2.18
RMS error	5.08	3.86
F-value	1614 ($P = 0.0001$)	317 ($P = 0.0001$)
Grand street site, January 1996		
n	124	120
r^2	0.788	0.844
Slope	1.12	1.09
Intercept	1.54	1.07
RMS error	8.10	6.56
F-value	458 ($P = 0.0001$)	645 ($P = 0.0001$)
Grand street site, August 1997		
n	63	62
r^2	0.901	0.955
Slope	0.960	1.00
Intercept	1.49	0.695
RMS error	19.3	13.1
F-value	536 ($P = 0.0001$)	1282 ($P = 0.0001$)

[a] n = number of observations.
[b] r^2 = coefficient of determination for the regression model.
[c] RMS error = the standard error of the Y estimate for the regression model.

4.4. Glenside mercury spill site

A total of 32 observations were obtained for comparison of Jerome versus NIOSH analysis at the Glenside mercury spill site. Two Jerome readings were taken at each location, one at initial placement of the sampling instrument and the other at sample pickup. All NIOSH results were less than 0.1 $\mu g/m^3$ and all Jerome readings were very low ($\leq$11) except for one observation, which had a Jerome reading of 323 at sample placement. Statistical analysis could not be performed on these datasets because there was no variation in the NIOSH results and very little variation in Jerome results. All but one of the Jerome measurements were less than 12 $\mu g/m^3$ and have relatively large uncertainty associated with them. The NIOSH data was qualitatively in agreement with Jerome readings except for one observation.

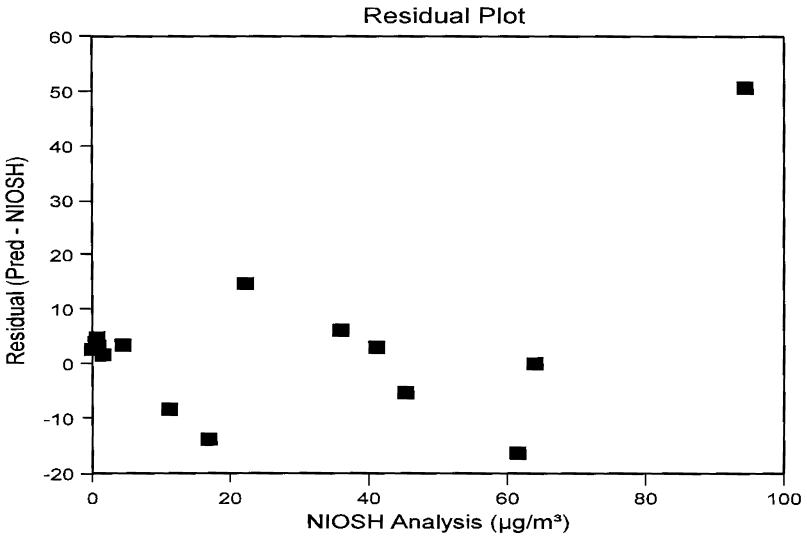

Fig. 2. Residual plot — LCP chemical site.

4.5. Dallas, PA mercury site

A total of 173 observations for Jerome versus NIOSH analysis data were obtained from the Dallas, PA mercury site. More than half of the data points (92 observations) had Jerome

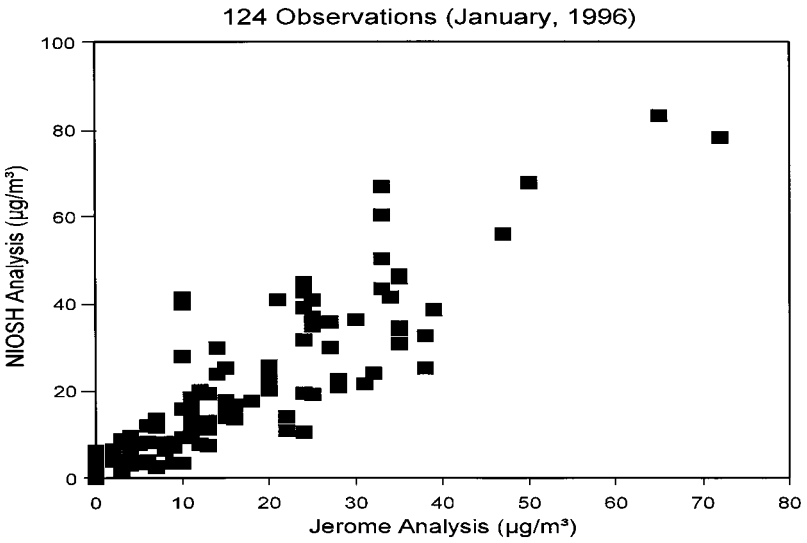

Fig. 3. NIOSH and Jerome mercury results — Grand Street site, January 1996.

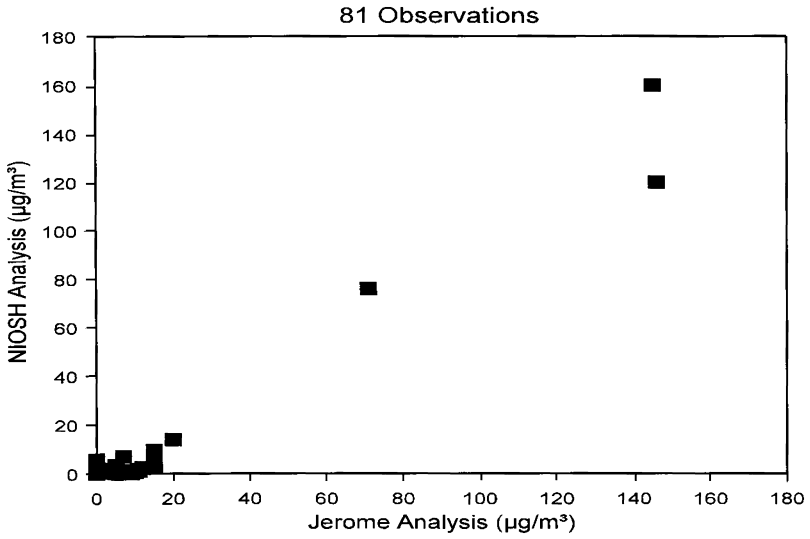

Fig. 4. NIOSH and Jerome mercury results — Dallas, PA mercury site.

readings reported as "less than" values, i.e. "<15", "<7", or "<3". Statistical analyses were performed on the remaining 81 observations. Fig. 4 shows NIOSH versus Jerome data for these observations. Pairwise comparisons analysis indicated that Jerome and NIOSH datasets were non-normal distributions, and were significantly different ($P = 0.0001$). Furthermore, the Spearman coefficient ($R = 0.511$) indicated that the data were not highly correlated. Regression analysis results (Table 1) indicated two potential outliers (Jerome $=$ 145, NIOSH $=$ 160 and Jerome $=$ 120, NIOSH $=$ 146) and the regression was repeated without the outliers. Removal of the potential outliers did not improve results; they were significantly worse without the outliers. Regression results for all data indicated that the Jerome and NIOSH data were highly comparable: $r^2 = 0.953$, slope $= 0.966$, $F = 1614$. The RMS error (5.1) indicated that Jerome results of 6 $\mu g/m^3$ or greater were meaningful compared to laboratory analysis.

4.6. Grand Street site, August 1997

Fig. 5 shows NIOSH versus Jerome results for 63 observations at the Grand Street mercury site (August 1997). In cases where multiple Jerome readings were taken, the average was used for evaluation purposes. Pairwise comparisons analysis indicated that Jerome and NIOSH datasets were non-normal and were not significantly different ($P = 0.541$). The Spearman coefficient ($R = 0.961$) indicated that the data were highly correlated. Regression analysis results (Table 1) identified one potential outlier (Jerome $= 143$, NIOSH $= 32$) and the regression was repeated without the outlier. Results without the outlier indicated that Jerome and NIOSH data were highly comparable: $r^2 = 0.955$, slope $= 1.00$, $F = 1282$.

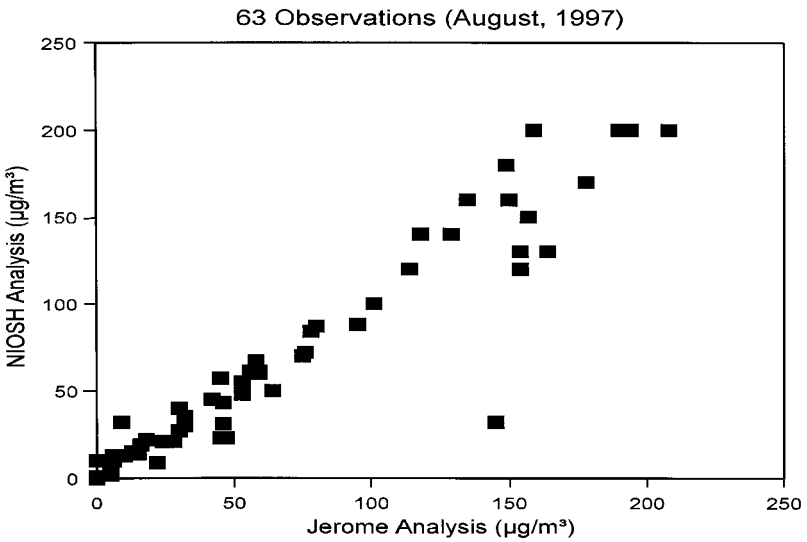

Fig. 5. NIOSH and Jerome mercury results — Grand Street site, August 1997.

The RMS error (13) indicated that Jerome results of 13 $\mu g/m^3$ or greater were meaningful compared to laboratory analysis.

5. Conclusions

Statistical analysis indicated that field (Jerome) and laboratory (NIOSH 6009) data for analysis of mercury in air samples were highly comparable for readings above the Jerome quantitation level (about 10 $\mu g/m^3$). The Glenside mercury spill site data could not be evaluated statistically, however, laboratory data was qualitatively in agreement with field readings except for one observation. Based on linear regression results, Jerome mercury readings of 6–13 $\mu g/m^3$ or greater are meaningful compared to laboratory analysis results. The typical instrument detection limit for the Jerome unit is approximately 3 $\mu g/m^3$, and a common field practice is to take three separate measurements and average the results. If the average is 3 $\mu g/m^3$ or greater, the reading is accepted. This corresponds to a quantitation limit of 9–10 $\mu g/m^3$, which is supported by statistical analysis. This suggests that Jerome mercury readings (average of three measurements) of 10 $\mu g/m^3$ or greater should be usable for field monitoring of mercury vapor in air.

Use of the Jerome 431^{TM} Mercury Vapor Analyzer provides real-time screening of suspected contaminated areas to assess initial extent of metallic mercury contamination, to identify “hot spots”, and to monitor progress of decontamination procedures at the spill site. For concentrations below 20 $\mu g/m^3$, the modified NIOSH 6009 method provides an effective way to measure low Hg vapor levels. The modified NIOSH method is simple enough to enable rapid sample turnaround, an important factor in making timely decisions, while conforming to accepted methodologies and QA/QC procedures.

Acknowledgements

The authors wish to thank Donna Getty of REAC for her statistical analysis support. The analytical methods described here were developed to meet US EPA/ERT/REAC field and laboratory requirements for monitoring indoor metallic mercury vapor and may not be applicable to the activities of other organizations. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

[1] R. Singhvi, D.A. Johnson, J. Patel, P. Solinski, Analytical method for indoor air monitoring for metallic mercury vapors, in: Proceedings of the 8th International Conference on Indoor Air Quality & Climate, Indoor Air 99, Edinburgh, Scotland, August 1999.
[2] Gold film mercury vapor analyzer, The Jerome 411, Operation Manual, Jerome Instrument Corporation, Division of Arizona Instruments Corporation (AZI), Rev. 9/87.
[3] US EPA/ERTC, Standard Operating Procedure #1827, Analysis of Mercury in Air with a Modified NIOSH 6009 Method, Rev. 2.0, 13 May 1999.
[4] R.O. Gilbert, Statistical Methods of Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, 1987.
[5] N. Draper, H. Smith, Applied Regression Analysis, 2nd Edition, Wiley, New York, 1981.
[6] S.D. Schlotzhauer, R.C. Little, SAS System for Elementary Statistical Analysis, SAS Institute Inc., Cary, NC, 1987.