REQUIRED NUMBER OF SAMPLES FOR ACCEPTABLE PRECISION OF COMPLIANCE DATA MEANS

WHITE PAPER

Prepared for:

National Rural Water Association
2915 South 13th Street
Duncan, Oklahoma 73533

Submitted By:

S. Cohen & Associates
(SC&A, Inc.)
1608 Spring Hill Road
Suite 400
Vienna, VA 22182

By:

Harry J. Chmelynski, PhD and Edwin L. Sensintaffar

October 13, 2011

TABLE OF CONTENTS

REQUIRED NUMBER OF SAMPLES FOR ACCEPTABLE PRECISION OF COMPLIANCE DATA MEANS

EXECUTIVE SUMMARY

A statistical analysis of data from the Environmental Protection Agency (EPA) Second Six-Year Review that was previously performed for 10 analytes identified a high degree of variability in the mean results when 1 sample in each of 4 calendar quarters were tested. The results of these tests often result in costly remediation for water systems when the mean of the four samples exceeds the Maximum Contaminant Levels (MCL) for identified contaminants. Improvement to the measurement process could provide a substantial savings to the Public Water System (PWS). A potential method of reducing the error associated with the measurement process is to increase the number of samples that are collected and analyzed each quarter from a specific PWS entry point.

Additional statistical analysis of these data has been performed to investigate the potential for improvement characterized by reduction of the error of the mean by increasing the number of samples beyond the required four quarterly samples. This analysis focused on identifying the number of potential samples that would be required to obtain an estimated error of the mean of ±10% or 25%. The results of the analysis, as summarized in Tables 2 and 3, identify a wide range in the number of samples required to meet these error limits for the 10 analytes. For example, fluoride tests meet the 25%, and nearly the 10% requirement, with only one sample per quarter. However, arsenic would require four samples per quarter to meet the 25% and 24 per quarter to meet the10% requirement. Some of the contaminants required numbers of samples that were beyond estimation in this analysis.

These results identify that increasing the number of samples obtained from a specific PWS entry point does not provide a simple answer to the issue of variability associated with the testing. However, it may provide a significant improvement in the error for some contaminants, such as radium, alpha particles, trichloroethylene (TCE), tetra-chloroethylene (PCE), arsenic and benzene. The increased cost of performing these additional analyses would need to be considered in light of the potential cost of remediation.

Even with the large amount of data included in the Second Six-Year Review and the earlier study, the amount of analyte-specific information for states and regions was limited. A geographical review by region of the United States shows that this problem is widespread, affecting PWSs in every part of the country. A regional analysis of the variability of test results found no evidence for regional differences.

1.0 INTRODUCTION

In accordance with current Safe Drinking Water Act and Standardized Monitoring Framework requirements, drinking water samples are collected from each entry point in a Public Water System (PWS) and analyzed for chemical and radiological contaminants on a quarterly basis. The average or mean concentration for these four quarterly measurements is used to determine compliance with the regulation. If the mean concentration of any of the contaminants exceeds the Maximum Contaminant Level (MCL), actions may be required that could be very costly, including additional sampling or remediation.

SC&A has previously conducted three evaluations of uncertainties in drinking water compliance data used for regulated analytes (SC&A 2005, SC&A 2007 and SC&A 2011). The first study contained an exploratory analysis of the problem. The second study evaluated the means of four quarterly measurements of as many contaminants (analytes) as data could be obtained by canvass of individual state regulatory agencies. In this study, SC&A gathered actual compliance data for several analytes from water systems or state enforcement agencies. SC&A made comparisons between the mean analyte levels and associated uncertainty. The results of the study revealed that the statistical uncertainty in the mean of quarterly measurements often exceeded the mean itself. These results were considered preliminary, since the available data for this study were limited in the number of systems for which we were able to obtain complete results for the evaluation. In addition, it was unclear if the results for samples taken from a given supply were from the same sampling point.

After the first two studies were completed, the Environmental Protection Agency (EPA) published a report, *The Analysis of Regulated Contaminant Occurrence Data from Public Water Systems in Support of the Second Six-Year Review of National Primary Drinking Water Regulations* (EPA 2009). This report included data for each of 69 chemical and radiological contaminants with over 15 million data records from water sampling over an 8-year period from January 1998 to December 2005. An EPA website for the Second Six-Year Review (http://water.epa.gov/lawsregs/rulesregs/regulatingcontaminants/sixyearreview/sixyearoccurencedata/index.cfm) provides downloadable datasets in Access® database format for these chemicals and data records. In a recent study (SC&A 2011), data were extracted from this database to estimate the sampling variability of the annual mean concentration for 10 selected analytes. Estimates of the variability of the mean were calculated under the current sampling protocol requiring four quarterly samples.

Appendix A to EPA 2009 provides a summary of the data submitted by PWSs with at least one measurement over the MCL. These PWSs were examined to identify states with sufficient data for analysis of each selected analyte. The 10 analytes and the states selected for the analysis and the MCL of each analyte are shown in Table 1.

Table 1: Analytes and States Selected for Analysis

SDWIS Code	Analyte	MCL		Selected States						
1005	Arsenic	0.01	mg/L	AK	AZ	ID	ME	MI	NH	NV
1025	Fluoride	4	mg/L	AZ	MO	NM	SC	VA		
1035	Mercury (inorganic)	0.002	mg/L	CA	NJ	SC	WI			
1085	Thallium	0.002	mg/L	CA	NC	UT				
2964	Dichloromethane (Methylenechloride)	5	µg/L	CA	MT	TX				
2984	Trichloroethylene (TCE)	5	µg/L	CA	CT	MA	NJ	NY	WI	
2987	Tetrachloroethylene (PCE)	5	µg/L	CT	MA	NJ				
2990	Benzene	5	µg/L	CA	IL	MD	NJ	NM		
4000	Alpha particles	15	pCi/L	ID	MN	NJ	NM	TX		
4010	Radium-226 and -228	5	pCi/L	IA	IL	MO	NJ	SC	WI	

2.0 REQUIRED SAMPLE SIZE TO REDUCE VARIABILITY IN THE ESTIMATED MEAN TO ACCEPTABLE LEVELS

A statistical analysis of data from the EPA Second Six-Year Review that was previously performed for 10 analytes identified a high degree of variability in the mean results when four quarterly samples were tested. Additional statistical analysis of these data has been performed in this study to investigate the potential reduction of the error in the mean by increasing the number of samples beyond the required four quarterly samples.

Data from the Second Six-Year Review were analyzed for each of 10 contaminants, including arsenic, fluoride, mercury (inorganic), thallium, dichloromethane, trichloroethylene, tetrachloroethylene, benzene, alpha particles, and radium-226 and -228. The analysis examined the mean (average) concentrations, as well as a standard deviation for each set of the four quarterly measurements. This information was then used to determine the percentage error in the estimate of the mean of each set of samples. The results of this analysis were reported in SCA 2011.

The error of estimation (E) for a mean computed using a sample of size n was found to be $E = t_{n-1}\,\sigma/\sqrt{n}$. Here, t_{n-1} is the tabulated value of the t-distribution for a sample size n, and σ is the standard deviation of the compliance measurements around their mean value. The results presented in SC&A 2011 were obtained using this formula with $n = 4$ samples per year. In this case, the t-value of t_3=3.182 is used to determine the estimation error for the mean. When the error of estimation is expressed as a percentage of the mean value, the percentage error in the estimated mean M is $\pm 100E/M$.

The variability in the estimated annual mean concentration for the 10 analytes is shown in Table 2. The percentage estimation errors for each analyte are shown in the right-hand column of the table. These estimates of variability were based on a sampling protocol using four quarterly compliance samples. When specific attention was directed at the estimation error when the mean concentration was near the MCL for each contaminant, the estimation error ranged from a high value of 255% of the mean for dichloromethane (methylene chloride) to approximately 45% for radium-226 and -228, alpha particles, and TCE, and 12% for fluoride.

Table 2: Estimation Error for the Annual Mean Using Four Quarterly Samples

SDWIS Code	Analyte	Estimation Error for Annual mean (% of Mean)
2964	Dichloromethane (Methylene chloride)	±255%
1035	Mercury (inorganic)	±165%
1085	Thallium	±157%
2990	Benzene	±104%
2987	Tetrachloroethylene (PCE)	±62%
4000	Alpha particles	±45%
4010	Radium-226 and -228	±45%
2984	Trichloroethylene (TCE)	±47%
1005	Arsenic	±72%
1025	Fluoride	±12%

The error of estimation for the annual mean may be reduced by increasing the number of samples collected. Increasing the value of n in the formula above has two effects: The t-value in the numerator is reduced (because the uncertainty in the estimate of σ is reduced) and the square root of n term in the denominator is increased. The reduction in the estimation error obtained by increasing the number of samples collected each quarter is shown in Figure 1 and Figure 2, which is an enlarged version of Figure 1. The curves in this figure show the decrease in the estimation error for the mean as the sample size is increased from one sample per quarter to higher values.

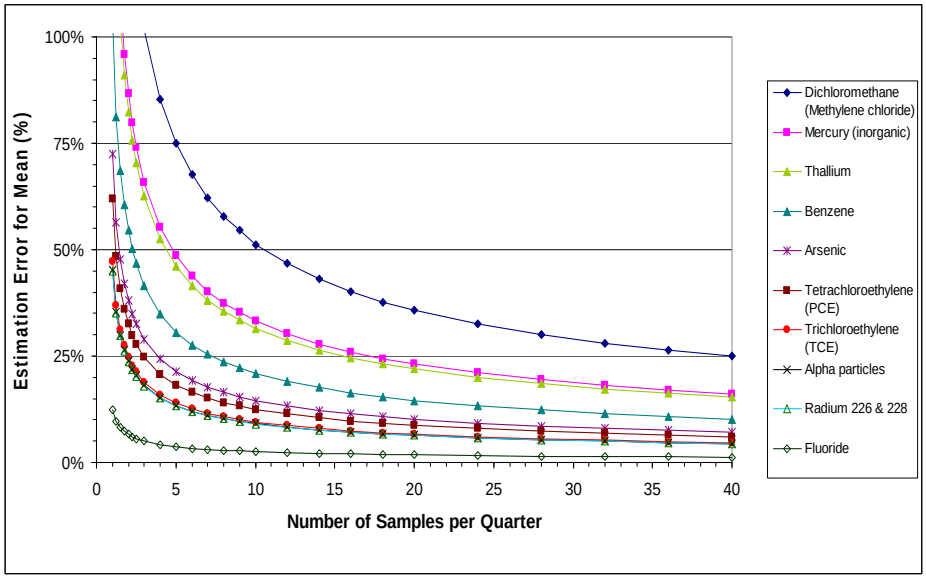


Figure 1: Estimation Error for Annual Mean vs. Number of Samples per Quarter

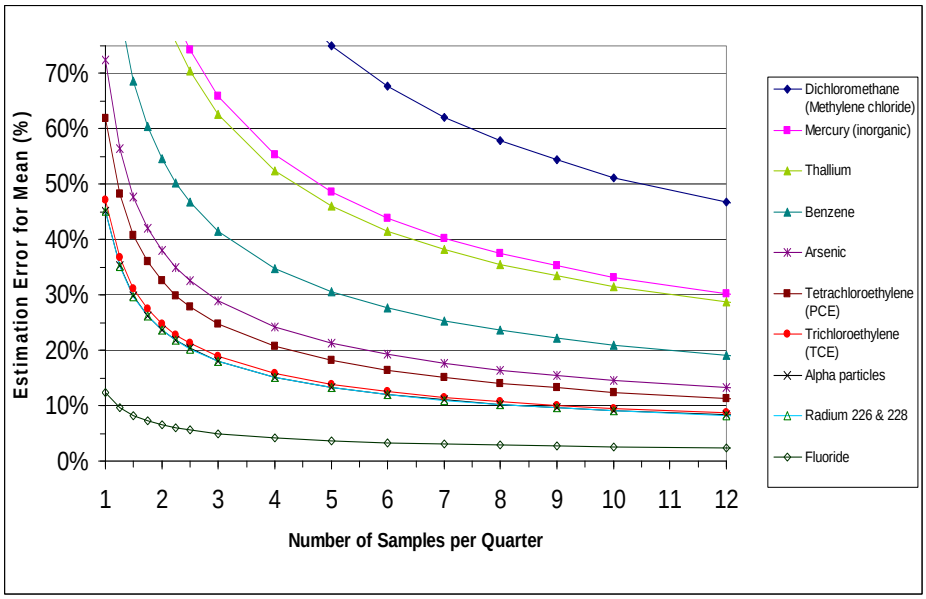


Figure 2: Estimation Error for Annual Mean vs. Number of Samples per Quarter (enlarged)

At the bottom of both figures, fluoride has the smallest estimation error of ±12% under the current sampling protocol using one sample per quarter. Next there is a cluster of three curves (radium, alpha particles, and trichloroethylene), which start at an estimation error of approximately 45%. Increasingly higher curves are shown for the remaining six analytes. Note that the highest four curves (benzene, thallium, mercury, and dichloromethane) all start at estimation errors larger than 100% in Table 2. These curves extend beyond the upper edge of Figure 1.

The size of the estimation error at one sample per quarter is reduced to lower levels by increasing the number of samples per quarter. The effectiveness in reducing the error by increasing by one more sample per quarter is largest at the beginning where the curves fall rapidly, and is reduced at higher sample sizes.

The number of samples per quarter required to achieve an estimation error of 10% or 25% is shown in Table 3. Fluoride already meets the goal of ±25% estimation error at one sample per quarter. The estimation errors for the five next lowest analytes can be reduced below 25% by using at most four samples per quarter: two samples per quarter for radium, alpha particles, and trichloroethylene; three for tetrachloroethylene; and four for arsenic. Benzene requires eight samples per quarter to achieve an estimation error ±25%. Increasingly larger sample sizes are required for the remaining three analytes: 16 for thallium, 18 for mercury, and 40 for dichloromethane.

Table 3: Number of Samples per Quarter Required for Estimation Errors of 10% or 25%

	Estimation Error for Annual Mean	
Analyte	**25%**	**10%**
Dichloromethane (Methylene chloride)	40	–
Mercury (inorganic)	18	–
Thallium	16	–
Benzene	8	40
Arsenic	4	24
Tetrachloroethylene (PCE)	3	16
Trichloroethylene (TCE)	2	10
Alpha particles	2	9
Radium-226 and -228	2	9
Fluoride	1	1.25

An estimation error of ±10% is not achievable at any realistic sample size for the highest three analytes (dichloromethane, mercury, and thallium). Approximately 10 samples per quarter would be required to achieve an estimation error of ±10% for radium, alpha particles, and trichloroethylene; 16 for tetrachloroethylene; 24 for arsenic; and 40 for benzene. An estimation error of ±10% is achievable for fluoride with only a small increase in the sample size.

3.0 GEOGRAPHICAL CHARACTERISTICS

The appendix to EPA 2009 contains extensive tables showing the percent of PWSs in each state with at least one detection over the MCL[1] during the 8-year period. Frequency plots for the four U.S. Census Regions (Northeast, South, Midwest, and West) are shown in Appendix A. The plots show the geographical distribution of the percent of PWSs with one or more detections above the MDA during the time period covered by EPA's Second Six-Year Review. Data for these plots were obtained from Appendix A of the EPA report.

The frequency counts show the distribution of the number of states in each region with percentages in the prescribed ranges. For example, the first graph in Appendix A for alpha particles shows that there was one state in the Midwest where approximately 60% of the PWSs reported one or more detections of alpha particles over the MCL.

For each of the 10 selected analytes, every region in the U.S. has at least one PWS with a detection over the MCL. Detections of arsenic, alpha particles, and radium are the most frequent, occurring in up to 20% to 60% of the PWSs in each region of the country. Less frequent detections (<5% of PWSs) are observed for the remaining analytes. Benzene and mercury have the least frequent detections, occurring in approximately 1% of the PWSs, but detections are found in every region.

Appendix B contains frequency plots by U.S. Census Region of the variability in the mean based on the average of four quarterly samples. The plots show the geographical distribution of the variability in the mean during the time period covered by EPA's Second Six-Year Review. In these plots, variability is expressed as a percentage of the mean (*100E/M*) for comparability. The frequency counts show the number of cases in each region with variability of the mean in the prescribed ranges.

For example, the first graph for alpha particles shows that the percentage variability of the mean in the Midwest appears to be lower than that in the other three regions where many of the PWSs show variability of over 100%. However, the Midwest also has relatively higher levels of alpha particles, as indicated in Appendix A. This would account for the lower variability when expressed in percentage form. Note the percentages shown in Table 2 were derived in SC&A 2011 using a regression technique designed to develop estimates for the percentage error when the PWS has levels near the MCL.

Data for the graphs in Appendix B include only the states with cases identified for the uncertainty analysis of each analyte shown in Table 1. Note the frequency of cases varied markedly across analytes, with only a few cases for some analytes. In the plots for benzene, dichloromethane, mercury, PCE, thallium, and TCE, many regions contain only a few cases. For these analytes, the data are inadequate to distinguish differences across regions.

These figures include data only for PWSs in the states listed in Table 1, which were selected for the error analysis. This set of PWSs does not cover all regions for each analyte. The uncertainty

[1] The occurrence of a detection over the MCL in a single sample does not indicate that an infraction has occurred. Such decisions are based on the mean of four quarterly samples.

in the estimated mean appears to affect all regions equally. For the three analytes with the most frequent detections (arsenic, alpha particles, and radium), the distributions are almost identical, with an upper tail extending to over 200% in almost all regions. The sparser data in the plots for the seven remaining analytes with less frequent detections cannot provide definitive results, but also suggest that the uncertainty in the mean is similar in most regions.

4.0 SUMMARY AND CONCLUSIONS

A statistical analysis of data from the EPA Second Six-Year Review that was previously performed for ten analytes identified a high degree of variability in the mean results when four quarterly samples were tested. Additional statistical analysis of these data has been performed to investigate the potential for improvement characterized by reduction of the error of the mean by increasing the number of samples beyond the required four quarterly samples. This analysis focused on identifying the number of potential samples that would be required to obtain an estimated error of the mean of ±10% or 25%. The results of the analysis as summarized in Tables 2 and 3 identify a wide range in the number of samples required to meet these error limits for the 10 analytes. For example, fluoride tests meet the 25% and nearly the 10% requirement with only one sample per quarter. However, arsenic requires four samples per quarter to meet the 25% and 24 to meet the ±10% requirements. Some of the contaminants required numbers of samples that were beyond estimation in this analysis.

These results identify that increasing the number of samples obtained from a specific PWS entry point does not provide a simple answer to the issue of variability associated with the testing. However, it may provide a significant improvement in the error for some contaminants, such as radium, alpha particles, trichloroethylene (TCE), tetra-chloroethylene (PCE), arsenic, and benzene. The increased cost of performing these additional analyses would need to be considered in light of the potential cost of remediation.

Even with the large amount of data included in the Second Six-Year Review and the earlier study, the amount of analyte-specific information for states and regions was limited. A geographical review by region of the United States shows that this problem is widespread, affecting PWSs in every part of the country. A regional analysis of the variability of test results found no evidence for regional differences.

5.0 REFERENCES

EPA 2009. *The Analysis of Regulated Contaminant Occurrence Data from Public Water Systems in Support of the Second Six-Year Review of National Primary Drinking Water Regulations,* EPA-815-B-09-006, U.S. Environmental Protection Agency, Washington, DC. October 2009. http://www.epa.gov/safewater/review/data.html#Six-Year%20Review%202%20occurr

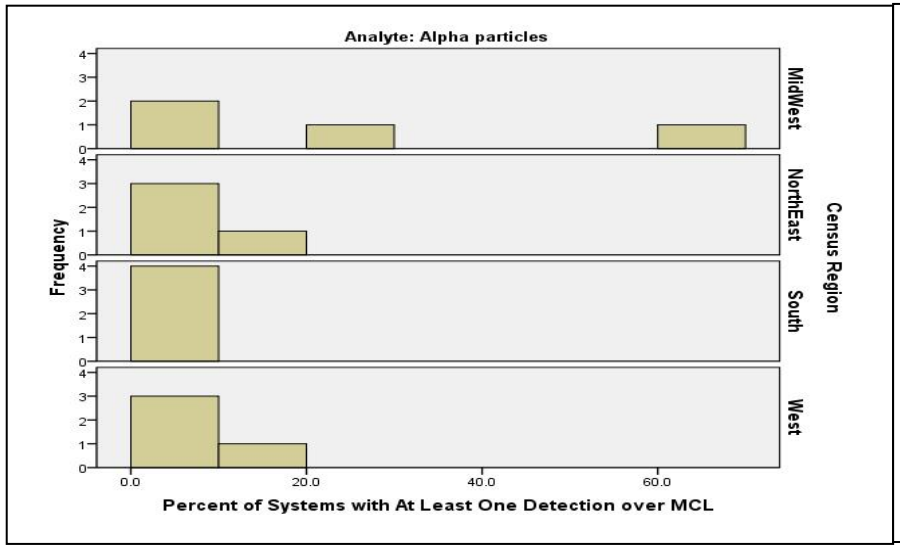
SC&A 2005. *Reliability of Laboratory Data Generated for Compliance with Drinking Water Regulations - A White Paper,* Prepared for the National Rural Water Association by SC&A, Inc., McLean, Virginia. Author: Charles Phillips. September 12, 2005.

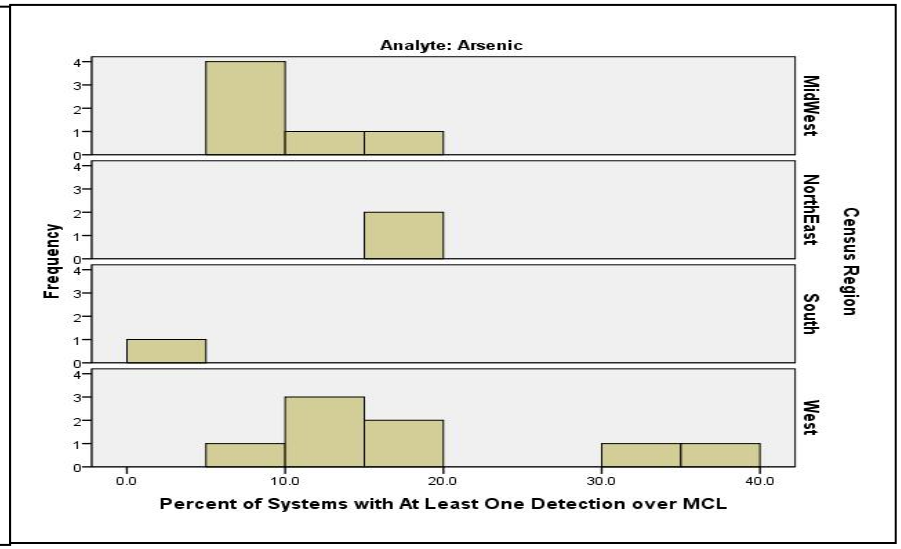
SC&A 2007. *Applicability of Laboratory Data Generated For Compliance with Safe Drinking Water Regulations - A White Paper,* Prepared for the National Rural Water Association by SC&A, Inc., Vienna, Virginia. Authors: Charles R. Phillips and Donald A. Chambless. May 8, 2007.

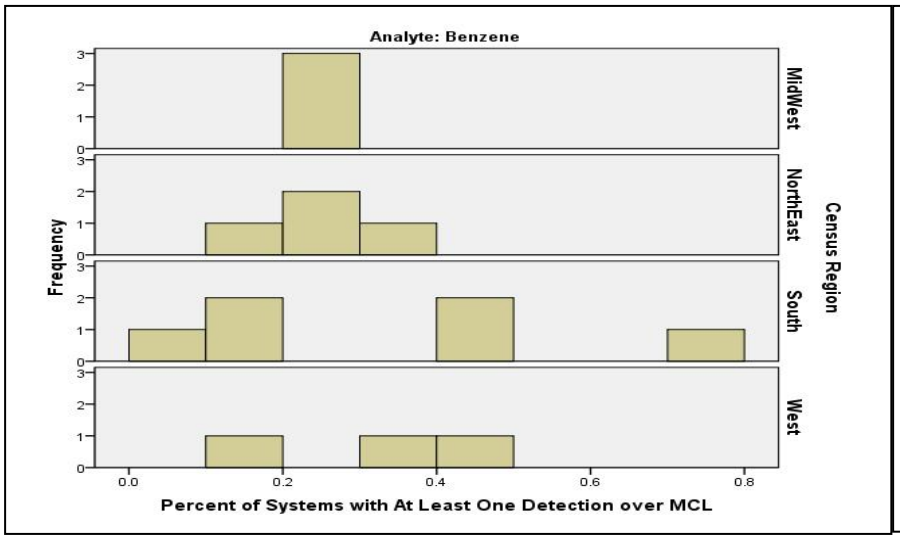
SC&A 2011. *Water Quality Compliance Decisions Based On Four Quarterly Measurements-An Analysis Based On The EPA Second Six-Year Review Of Primary Drinking Water Regulations,* Prepared for the National Rural Water Association by SC&A, Inc., Vienna, Virginia. Authors: Charles R. Phillips and Harry J. Chmelynski. February 16, 2011.

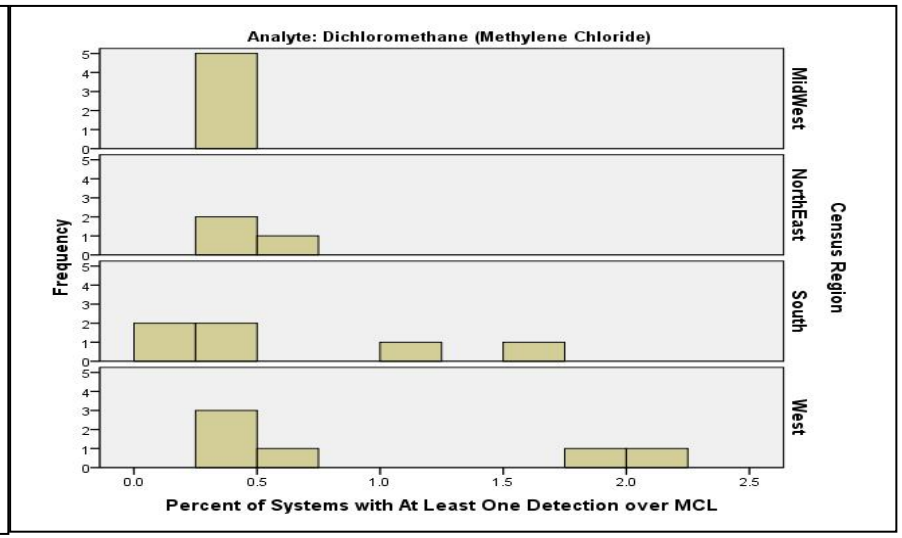
APPENDIX A: GEOGRAPHICAL DISTRIBUTION OF STATES BY PERCENT OF PWSs WITH AT LEAST ONE DETECTION OVER MCL

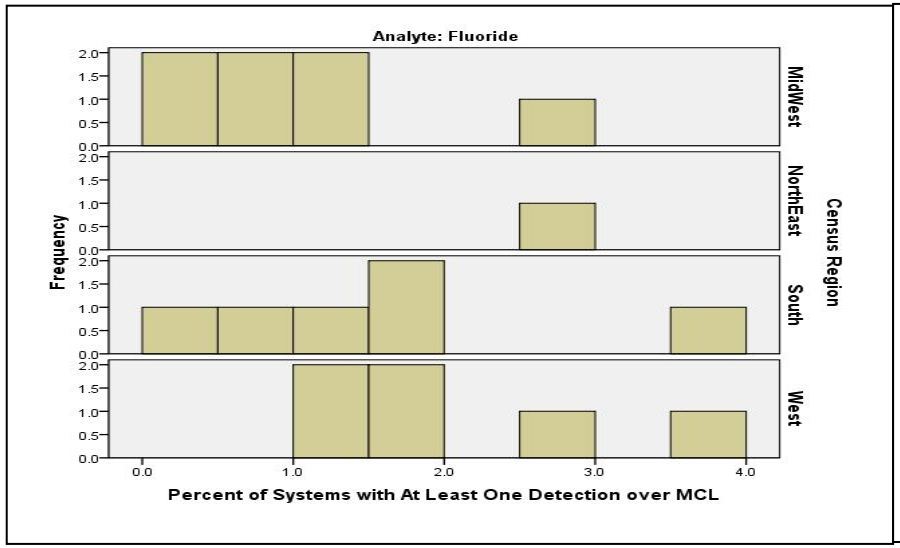
The plots in Appendix A show the geographical distribution of the percent of PWSs with one or more detections above the MDA during the time period covered by EPA's Second Six-Year Review. The frequency counts show the distribution of the number of states in each region with percentages in the prescribed ranges. For example, the first graph for alpha particles shows that there was one state in the Midwest where approximately 60% of the PWSs reported one or more detections of alpha particles over the MCL. Data for these plots were obtained from Appendix A of the EPA report.

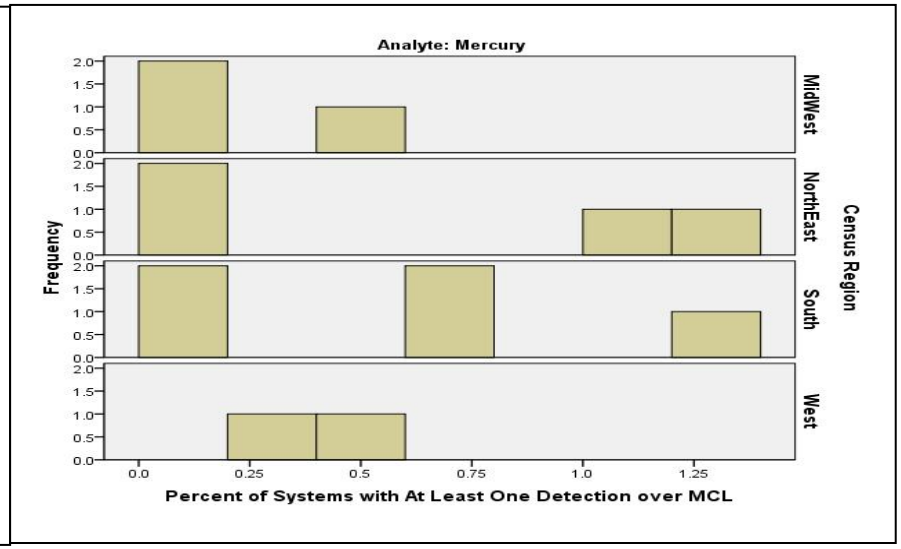

Analyte: Alpha particles
Frequency
Census Region
MidWest
NorthEast
South
West
0.0
20.0
40.0
60.0
Percent of Systems with At Least One Detection over MCL

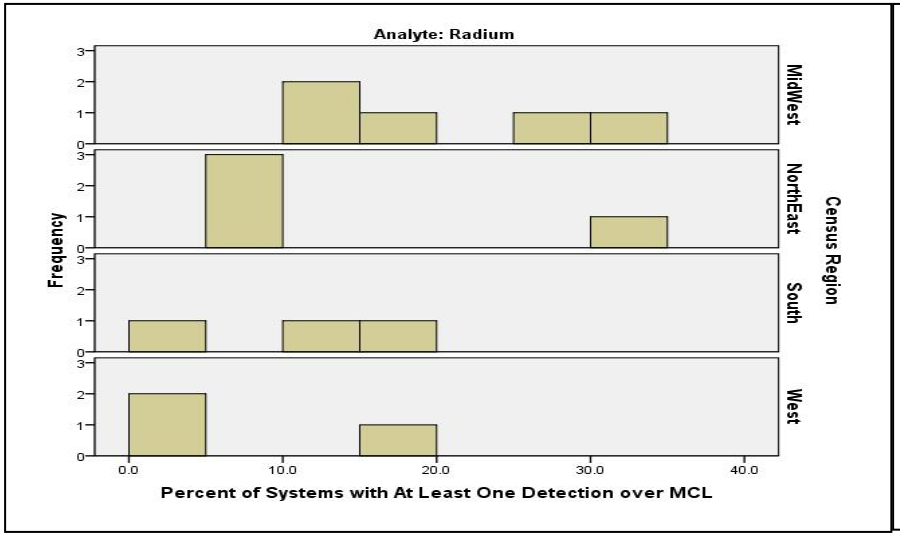

Analyte: Arsenic
Frequency
Census Region
MidWest
NorthEast
South
West
0.0
10.0
20.0
30.0
40.0
Percent of Systems with At Least One Detection over MCL

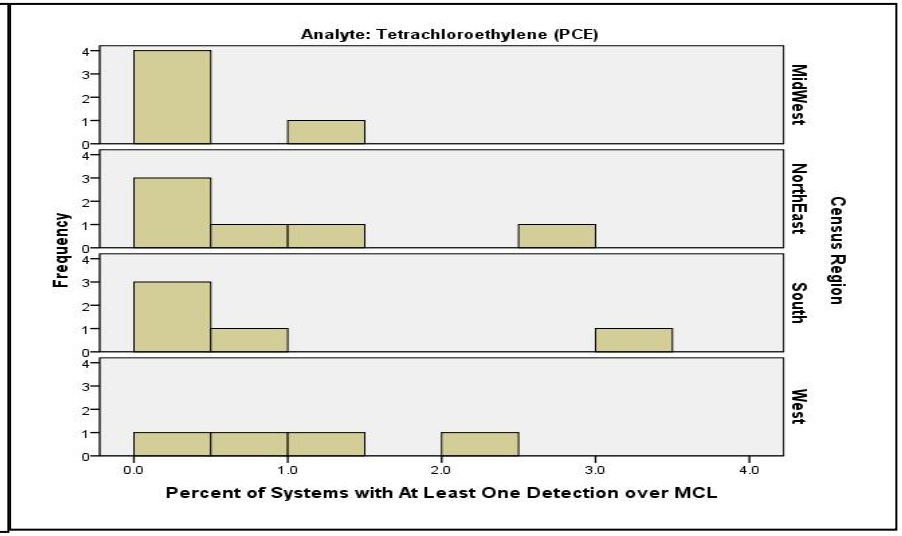

Analyte: Benzene
Frequency
Census Region
MidWest
NorthEast
South
West
0.0
0.2
0.4
0.6
0.8
Percent of Systems with At Least One Detection over MCL

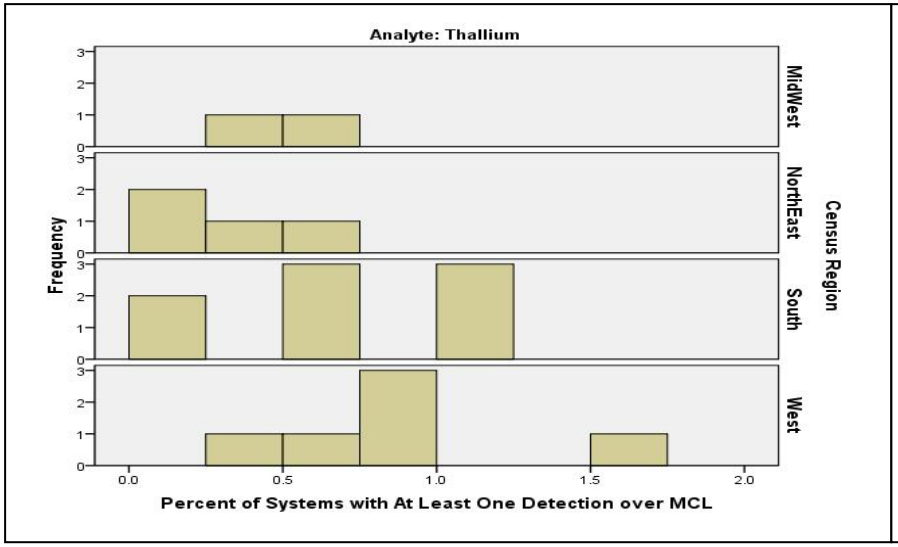

Analyte: Dichloromethane (Methylene Chloride)
Frequency
Census Region
MidWest
NorthEast
South
West
0.0
0.5
1.0
1.5
2.0
2.5
Percent of Systems with At Least One Detection over MCL

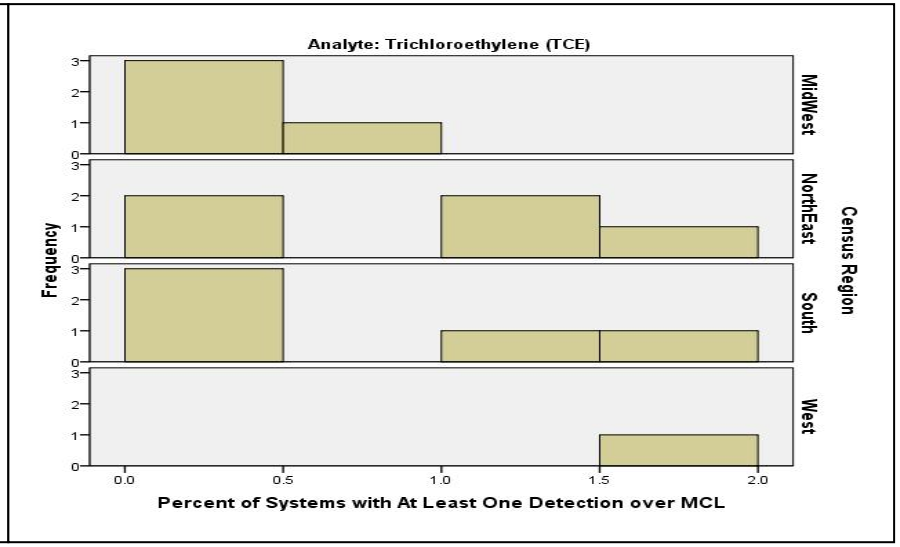

Analyte: Fluoride
Frequency
Census Region
MidWest
NorthEast
South
West
Percent of Systems with At Least One Detection over MCL


Analyte: Mercury
Frequency
Census Region
MidWest
NorthEast
South
West
Percent of Systems with At Least One Detection over MCL


Analyte: Radium
Frequency
Census Region
MidWest
NorthEast
South
West
Percent of Systems with At Least One Detection over MCL


Analyte: Tetrachloroethylene (PCE)
Frequency
Census Region
MidWest
NorthEast
South
West
Percent of Systems with At Least One Detection over MCL


Analyte: Thallium
Frequency
MidWest
NorthEast
South
West
Census Region
0.0
0.5
1.0
1.5
2.0
Percent of Systems with At Least One Detection over MCL

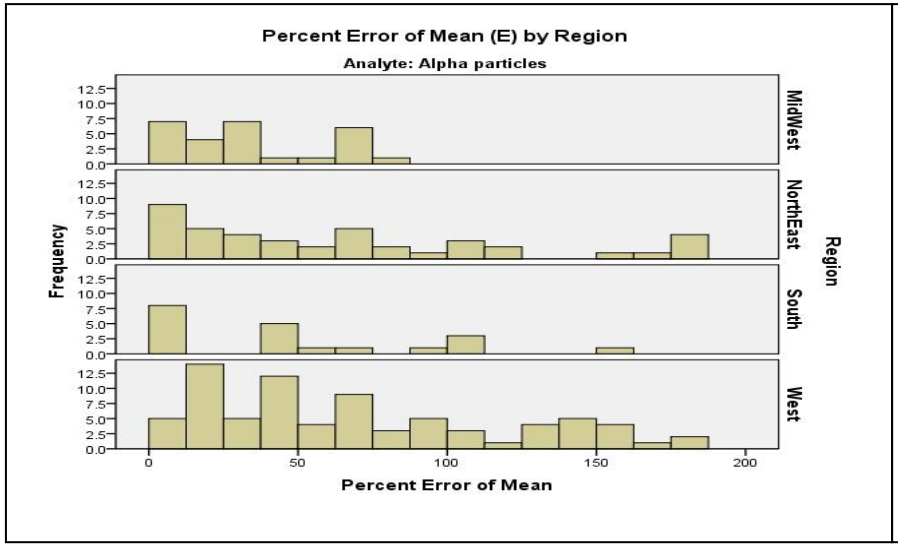

Analyte: Trichloroethylene (TCE)
Frequency
MidWest
NorthEast
South
West
Census Region
0.0
0.5
1.0
1.5
2.0
Percent of Systems with At Least One Detection over MCL

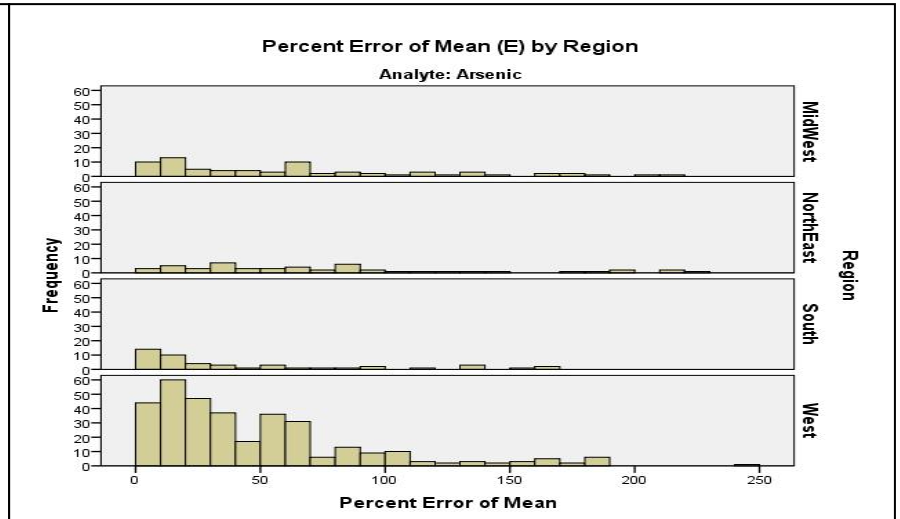
APPENDIX B: GEOGRAPHICAL DISTRIBUTION OF ERROR IN ESTIMATES OF MEAN BASED ON FOUR QUARTERLY SAMPLES

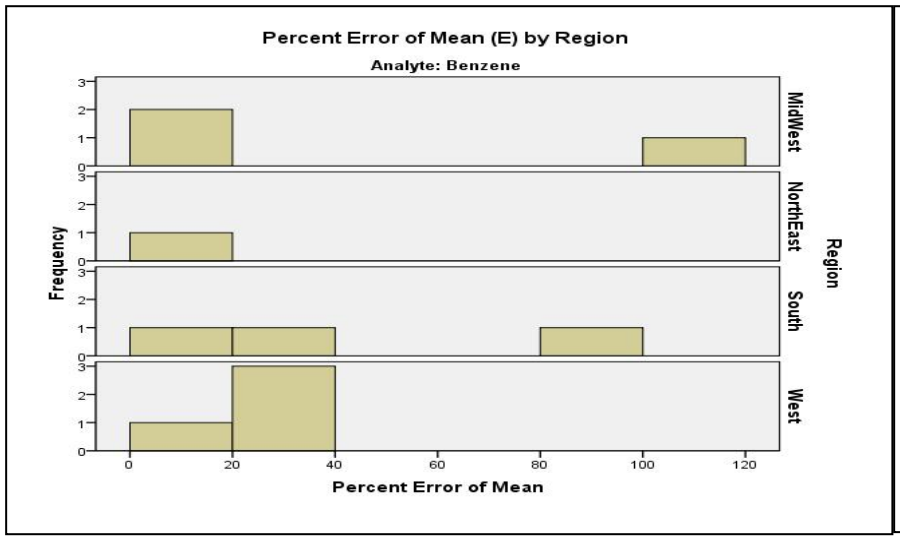
The plots in Appendix B show the geographical distribution of the variability in the mean (*E*) during the time period covered by EPA's Second Six-Year Review. The frequency counts show the distribution of the number of PWSs in each region with variability of the mean in the prescribed ranges. In these plots, variability is expressed as a percentage of the mean (*100E/M*) for comparability.

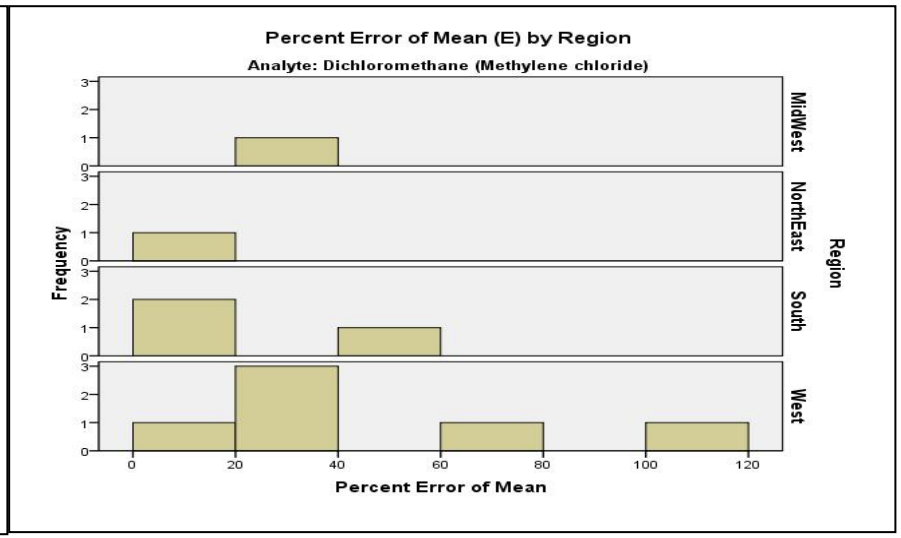
For example, the first graph for alpha particles shows that the percentage variability of the mean in the Midwest appears to be lower than that in the other three regions where many of the PWSs show variability of over 100%. However, it is likely that the Midwest also has relatively higher levels of alpha particles, as indicated in Appendix A. This would account for the lower variability when expressed in percentage form. The percentages shown in Table 2 were derived in SC&A 2011 using a regression technique to develop estimates for the percentage error when the PWS has levels near the MCL.

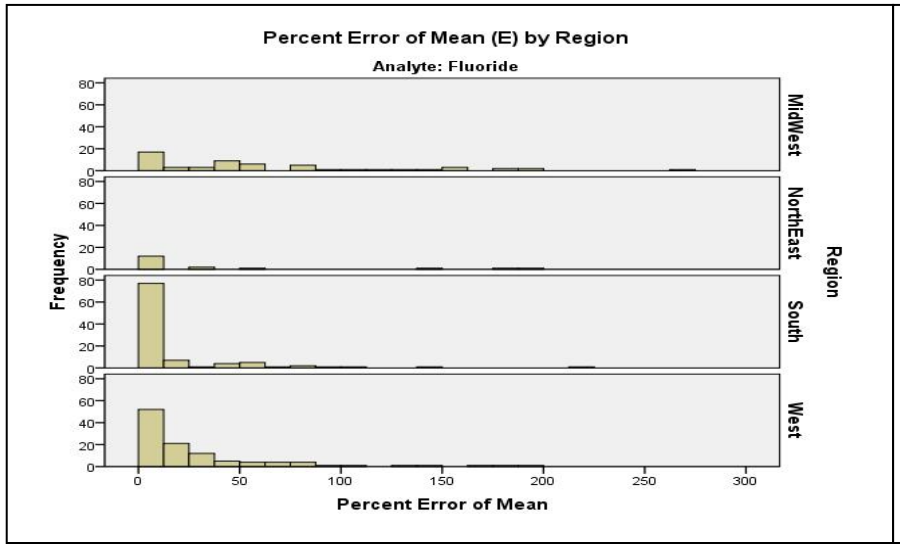
Data for these graphs include only the states with cases identified for the uncertainty analysis of each analyte shown in Table 1. Note the frequency of cases varied markedly across analytes, with only a few cases for some analytes. In the plots for benzene, dichloromethane, mercury, PCE, thallium, and TCE, many regions contain only a few cases. For these analytes, the data are inadequate to distinguish differences across regions.

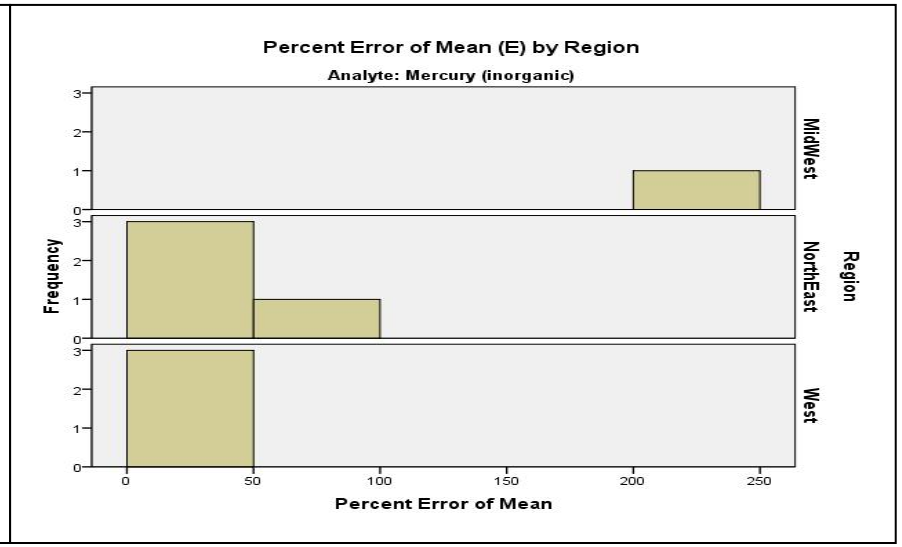

Percent Error of Mean (E) by Region
Analyte: Alpha particles
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean

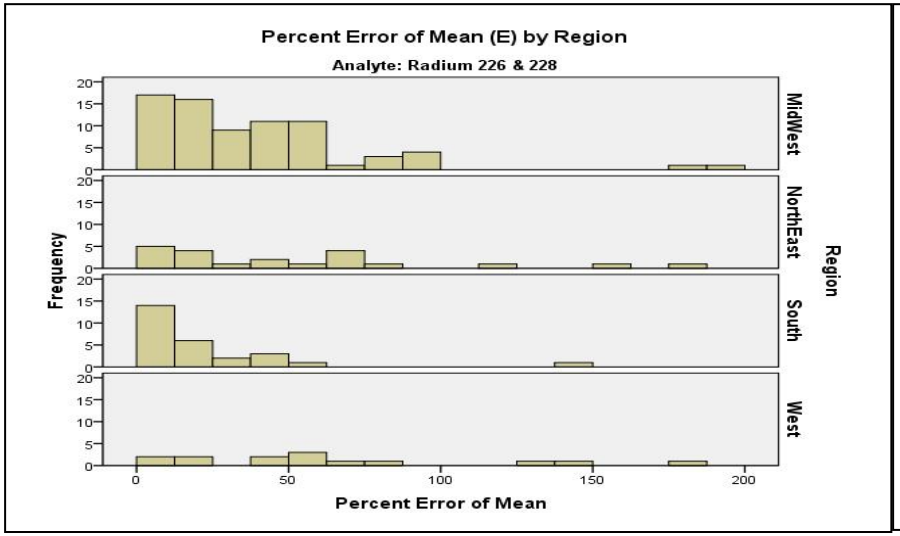

Percent Error of Mean (E) by Region
Analyte: Arsenic
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean

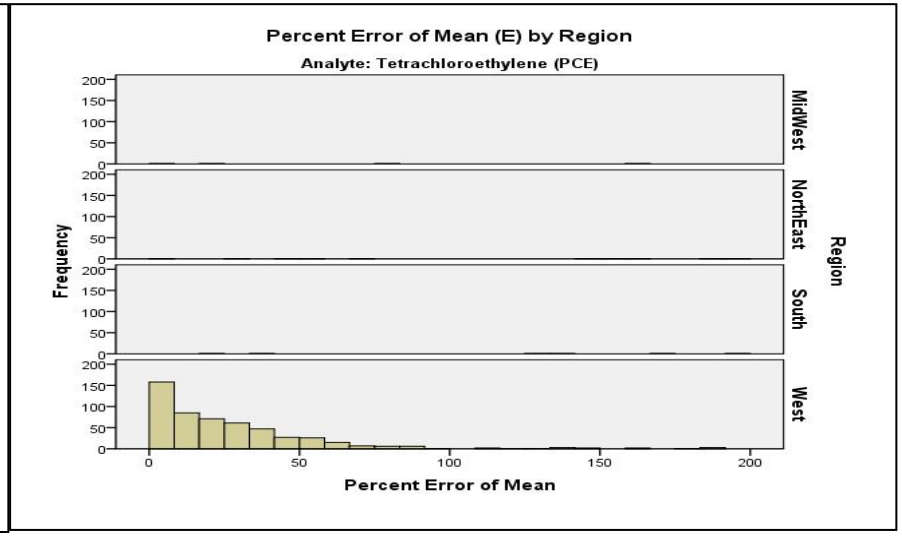

Percent Error of Mean (E) by Region
Analyte: Benzene
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean

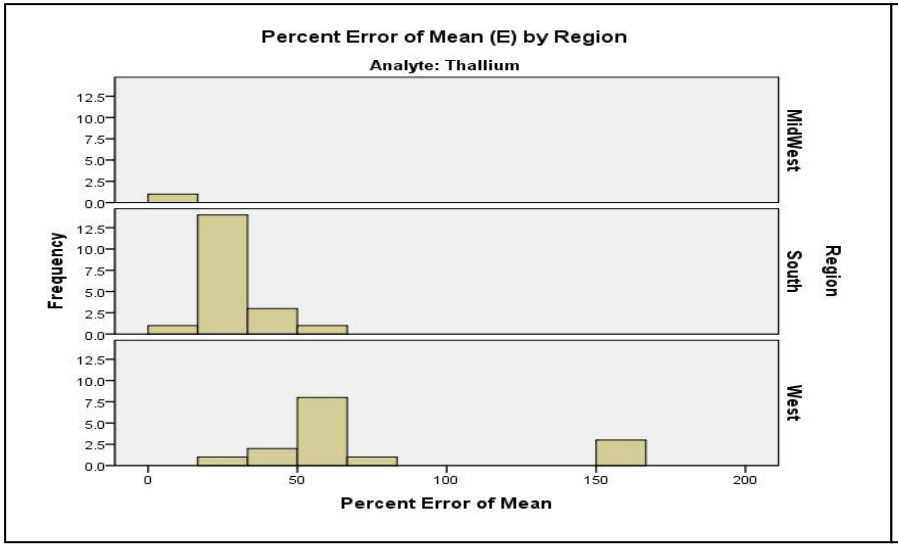

Percent Error of Mean (E) by Region
Analyte: Dichloromethane (Methylene chloride)
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean

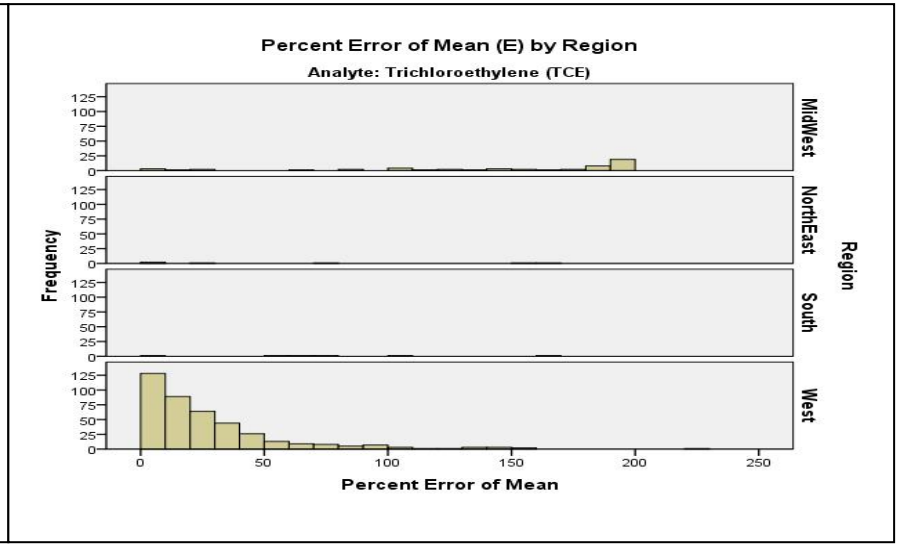

Percent Error of Mean (E) by Region
Analyte: Fluoride
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean


Percent Error of Mean (E) by Region
Analyte: Mercury (inorganic)
Frequency
Region
MidWest
NorthEast
West
Percent Error of Mean


Percent Error of Mean (E) by Region
Analyte: Radium 226 & 228
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean


Percent Error of Mean (E) by Region
Analyte: Tetrachloroethylene (PCE)
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean


Percent Error of Mean (E) by Region
Analyte: Thallium
Frequency
Region
MidWest
South
West
Percent Error of Mean


Percent Error of Mean (E) by Region
Analyte: Trichloroethylene (TCE)
Frequency
Region
MidWest
NorthEast
South
West
Percent Error of Mean